We thank Referee #3 for their helpful comments and review of our paper; we have responded to all of their comments below.

The reviewer comments are in black and our responses are in blue.

Referee #3: Goddijn-Murphy and colleagues present a gridded monthly climatology of sea surface CO2 fugacity based on SOCAT measurements. Prior to interpolation, the data are adjusted from their in-situ temperature to climatological temperature derived from satellite observations. While the interpolation method seems less advanced than pre-existing approaches like Takahashi et al (2009), the consideration of the appropriate temperature is novel and interesting, and in my opinion worth publishing.

Reply: We thank the reviewer for highlighting the novelty and success of our approach.

Response in paper:

None specific

Referee #3: Nevertheless, I have several concerns about this temperature adjustment, both regarding the way it is presented in the manuscript and the steps actually taken, as follows. The mentioning of a bias in Takahashi et al (2009) on page 1898 line 10-11 seemed to suggest that it is the fCO2 change along this vertical temperature gradient you aim to correct for (beginning of Sect 1.2 seems to be in this line as well).

Reply: As the focus of this paper is the complexities of SST measurements and their implications for sea surface f_{CO2} climatologies it is important that our temperature adjustment is well understood. The reviewer supposes that the bias mentioned on page 1898 lines 10-11 is the vertical temperature gradient we aim to correct for. However, in the preceding sentence (lines 8-10) we say that this is the bias from under-sampling and their interpolation method according to Takahashi et al (2009). In other words, there is a difference between the 'true' monthly grid box mean of f_{CO2} (at whichever depth) and the mean of a number of f_{CO2} samples during that month in that grid box. This is not the same as differences due to the vertical temperature gradient. In the following sentence (lines 12-14) we say that Takahashi et al. (2009) acknowledge that by using SST at depth in their calculations, surface-layer effects could introduce systematic errors in the sea-air p_{CO2} differences but leave that for future research. Later, in the beginning of Section 1.2, all the complexities of using in situ SST including the vertical SST gradient are described (see next author response).

Response in paper:

See following responses

Referee #3: Later however (page 1902 line 11-13 and methods) you actually seem to adjust from subskin in-situ temperature to subskin monthly mean temperature, ie. along temporal variations within the months and between years. These are very different targets. It should be clarified what your target actually is.

Reply: Those are different targets but they both need to be addressed to derive 'truly mean' monthly grid-box values of f_{CO2} at which gas transfer occurs. The full procedure applied to the *in situ*, instantaneous f_{CO2} values at depth involves:

1) minimize the bias due to under-sampling;

2) adjust for errors due to the vertical SST gradient;

3) minimize errors due to the different methods and instrumentation used for measuring f_{CO2}.

The text in the beginning of Section 1.2 can be replaced by these bullet points to clarify this. All three targets can be addressed by calculating f_{CO2} for a monthly grid box mean of subskin SST from satellite, $T_{ym,i}$, instead of for in situ SST, in this paper indicated by $f_{CO2,ym,i} = f_{CO2}(T_{ym,i})$.

Response in paper:

See following responses

Referee #3: If I understand correctly that your target is adjusting for the temporal variations (one of the other reviewers understood it differently), you should clarify that this temperature adjustment is only relevant because you use the data in the chosen climatological year rather than in their actual year.

Reply: As explained above the temperature adjustment is not only relevant because we derived f_{CO2} data for the chosen climatological year (in our case 2010) from f_{CO2} data in their measurement year but for other reasons as well. First of all we had to derive a monthly grid box mean $f_{CO2,ym,i}$ using the respective monthly mean subskin temperature from satellite, $T_{ym,i}$. Then the $f_{CO2,ym,i}$ values were extrapolated to the year 2010 ($f_{CO2,cl,i}$) and finally the $f_{CO2,cl,\,i}$ data were grouped by month and averaged over 1° ×1° squares before the data were kriged.

The explanation of our re-conversion and the reasons for it are spread out through the paper and we can see that a short summary of our reanalysis in Section 1.5 would help clarify this to the reader.

Response in paper:

The several referee comments above and from other referees make it clear that we must work on our explanations. As outlined above a short summary of our reanalysis in Section 1.5 will be added to aid clarity.

Referee #3: In contrast to what some formulations (e.g. beginning of Sect 1.2) may suggest, the instantaneous sea-air CO2 flux (mentioned as primary target of the work) is actually determined by the CO2 fugacity at its in-situ temperature (not at some "true climatological mean" [page 1898 line 9] which is actually an oxymoron as a mean is a conceptual quantity only). So far it remains open how the actual sea-air CO2 flux would be re-calculated from the presented mean monthly fugacity climatology. The authors should address how they envisage to do that, and spend a few more words on their motivation to create a mean climatology only.

Reply: The flux of slightly soluble nonreactive gases across the air-sea interface, *F* (mol $m^{-2}s^{-1}$), can be defined as the product of a quantity called the gas transfer velocity, *k* (m s^{-1}), and the concentration difference between the top and bottom of the liquid boundary layer. This can be expressed as

$$F = k\alpha_W\left(f_{CO2W} - f_{CO2A}\right) \quad (1)$$

with a_W the aqueous-phase of CO_2 in water and f_{CO2w} and f_{CO2a} the fugacity in water and in air respectively. It is common practice to use the time- and area-averaged gas exchange equation

$$\overline{F} = \overline{k}\,\overline{\alpha_W}\,\overline{\Delta f_{CO2}} \quad (2)$$

(Schuster et al. 2009), though this does necessarily neglect covariances. We conform to that practice and the necessity of unbiased estimates of the monthly composite values is then manifest. The actual production of “climatological fluxes” is beyond the scope of this paper, but the requirement for monthly composites should be clear. Reducing fluxes to “a climatology in a reference year” again is standard, but we agree somewhat unsatisfactory. In the absence of sufficient sampling of dissolved carbon dioxide globally in the upper ocean, a “climatology” will remain a necessary if clumsy product.

Response in paper:

The importance of properly defined monthly composite values for fluxes integrated over a month is clarified. The definition of the gridded fields in the reference year will also be clarified

Referee #3: Concerning the vertical temperature gradient, a constant cool bias of 0.14K is assumed at all times and locations. This is of similar order of magnitude than the dT values reported in Fig 1, such that any deviations of the actual vertical temperature gradient from this assumed number will be a substantial error of fCO2 introduced through the temperature adjustment. I did not find this error being discussed in Sect 5, but it should certainly be considered and quantified.

Reply: the error of $T_{ym,i}$ (grid box mean of subskin SST from satellite) and its propagation in $f_{CO2,cl}$ is discussed and estimated in Section 5.8., page 1916 line 13-21. The total uncertainty in $T_{ym,i}$ (standard deviation of the mean + uncertainty in the cool skin bias) is estimated to be ±0.2 K propagating into an error of ±3 µatm in $f_{CO2,cl}$ (Table 3).

Response in paper:

None specific but we review the clarity of our explanations.

Referee #3: There are various other fugacity (partial pressure) gridded products in the literature (global products, e.g., by Takahashi et al. (2009), Park et al (2010), Rödenbeck et al (2013), as well as regional studies, e.g. Telszewski et al. (2009), Landschützer et al (2013), Ishii et al (2014) and various others). How do the presented results compare to these previous results? I feel it would be appropriate to at least mention this body of work, and to explicitly compare to at least one of them, to allow the reader to see how this study relates to the work of the community.

Reply: As explained to Reviewer #2, we will briefly discuss other products and compare the applied methodologies but we will not compare their results as this is not within the scope of our paper.

Response in paper:

As outlined in the reply above

Minor comments:

Referee #3: *p 1896 l 10: The term "climate quality" may need a short explanation (is it synonymous to climatological"?)*

Reply: We will replace 'climate quality' with 'single year monthly composite'. We apologise for any confusion (we used climate quality because it is not synonymous with climatological)

Response in paper:

As outlined in the reply above

Referee #3: Sect. 1.4: Please give the definition of "mass boundary layer" and "thermal boundary layer". Are they used synonymously?

Reply: The terms are not synonymous and we agree their meaning should be made clearer to the reader. Thank you for pointing out this oversight.

Response in paper:

We will add a figure showing how the different layers relate to SST definitions. Please see the figure below.

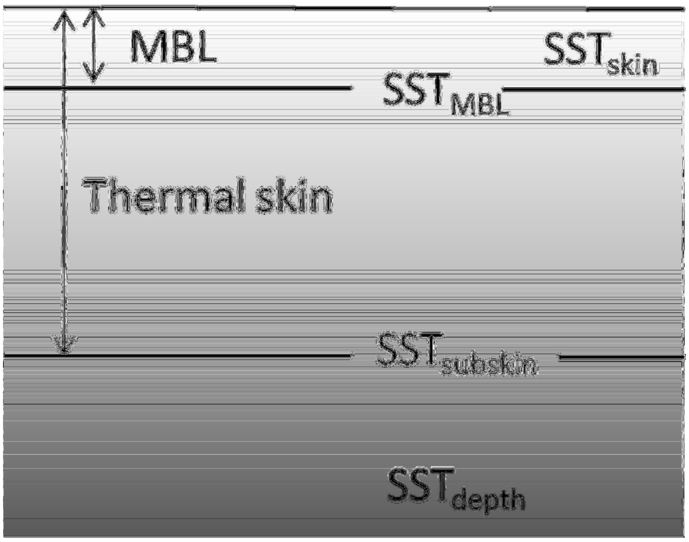

Figure R1. A schematic of the surface ocean, depicting the definition of the molecular boundary layer (MBL), thermal skin and various temperatures at depths.

Referee #3: p 1902 l 11: Does index i refer to the SOCAT point? Please clarify in the text.

Reply: Again, we should have been clearer here.

Response in paper:

We will change "These SST_{skin} grid points were linearly interpolated to the SOCAT measurement locations ($SST_{skin,\,i}$)" to "These SST_{skin} grid points were linearly interpolated to the i-th SOCAT measurement location ($SST_{skin,\,i}$)"

Referee #3: p 1902 l 12-17: Give the units of the numbers (here and throughout).

Reply: We apologize for the missing units and we will add them into the paper.

Response in paper:

Units are added

Referee #3: p 1902 l 18: The judgement of overestimation is only true if the 0.14K adjustement is correct.

Reply: The 0.14 degree is taken from (Donlon et al. 1999). The subskin temperature is widely accepted by the international SST community to be 0.14-0.17 K lower than the skin temperature. Using 0.17 (Donlon et al. 2002) would result in an over estimation of 0.06 K instead of 0.09 K.

Response in paper:

We will add a sentence noting the potential bias related to the assume "skin effect".

Referee #3: Sect 1.5: The text is a bit out of context. If this is meant as an announcement of later sections, explicitly refer to there.

Reply: OK, thank you for pointing this out. We will give a short summary (preview) of our re-analysis here instead.

Response in paper:

As outlined in the reply above

Referee #3: Sect 2: The description of SOCAT (Sect. 2 and elsewhere) is formulated like a method's description and, as I feel, can be misunderstood as if the authors were part of that work. Please use formulations that clarify that SOCAT is not part of the present study.

Reply: OK, thanks for pointing this out .This is consistent with a comment by Reviewer #1. We will move the section that describes the SOCAT methods to the Appendix and make clear that it is not part of our work.

Response in paper:

The paper will be restructured as outlined in the reply above.

Referee #3: p 1906 l 14: The difference between Eqns (1) and (8) can be explicitly calculated, so why do you just "expect"?

Reply: Apologies, we agree.

Response in paper:

We will remove the word ‘expect’.

Referee #3: p 1909 l 3-5: I don’t think this expection is justified. Please show representative data records to convince the reader.

Reply: We think that the reviewer is entirely justified to raise this as a query, but we are guided by precedent (e.g. Takahashi et al., 2009, Olsen et al., 2004) and it is beyond the scope of the paper to do more than take this as a reasonable assumption. We use the term “reasonable assumption” advisedly since we think it “reasonable” but not “certain”. Faced with inconsistencies in temperatures, it is clear that a correction of some sort is required and it is certain that an isochemical assumption is better founded than assuming a constant fugacity. The argument is theoretical rather than experimental and it is questionable whether any data can meaningfully test the assumption. Some processes conserve DIC and Alkalinity, while others do not (see for example, Chapter 6 of Williams and Follows, “Ocean Dynamics and the Carbon Cycle”, Cambridge University Press, 2011; or Chapter 4 of Emerson and Hedges, “Chemical Oceanography and the Marine Carbon Cycle”, Cambridge University Press, 2008). We can apply an isochemical assumption with confidence to some of the more important processes behind temperature inconsistencies, such as the formation of warm layers, e.g. Olsen et al., 2004. The assumption is less satisfactory given processes of photosynthesis, respiration and air-sea exchange, but we are guided by the gradual seasonal change in DIC and Alk associated with these processes and the implicit assumption is that there is no systematic bias resulting for corrections within each calendar month. It would be clearly inappropriate to extrapolate values from one month to the next on an isochemical assumption and that has been avoided.

Response in paper:

We will review our explanations

Referee #3: p 1909 l 18: As atmospheric growth has accelerated, I don’t think the 1.5uatm/yr slope is still justified for 2010.

Reply: Yes, this is probably correct, but the issue is being actively debated by the international community. However, the focus of this work was not to study trends in atmospheric pCO2, so we based our work on published results (ie Takahashi et al., 2009). However, more recently Takahashi et al. (2014) presented an updated oceanic p_{CO2} trend of 1.9 µatm/year. This would result in a difference of 0.8, 0.4 and 0 µatm for 2008, 2009 and 2010 respectively and this information is explained and included in Section 5.5 of the paper.

Response in paper:

We will review our explanations but no specific change is proposed.

Referee #3: Sect 3.3, item 3: Aren't you concerned by that? How many values are affected?

Reply: This issue is discussed and the error calculated in Section 5.7. The proportion of affected values is illustrated in Fig. 11, but we will replace this figure with giving the fractions of missing values over all months and all years (salinity 14%, T_{eq} 17%, P 37%, and P_{eq} 41%). This is the unfortunate issue of working with sparse datasets and it is simply due to that feature of the in situ dataset.

Response in paper:

As outlined in the reply above

Referee #3: Sect 4: I feel reference to a piece of software is not a valid method description, because this software may not be available in the future any more. Please give a more mathematical description of what you did.

Reply: Our apologies, we forgot to give the reference "Pebesma, E. J., 2004: Multivariable geostatistics in S: the gstat package. *Comput. Geosci.,* **30**, 683-691." in which the mathematics of the kriging is described. We will add this reference into the text.

Response in paper:

We will add the reference into the text.

Referee #3: p 1911 l 6: Please give units of these parameters.

Reply: Apologies, we will add the units.

Response in paper:

Units added

Referee #3: p 1911 l 20-22: I feel it is neither fair nor meaningful to compare the actual numerical bias of Takahashi et al (2009) to the theoretically expected zero bias of Kriging. Please give the actual numerical bias of your results. If you want to rate the two methods, the bias would need to be calculated in an identical way.

Reply: Our response to this comment is two-fold. Firstly, we acknowledge that this might be read as a rather unfair comparison. Thus, we will take out the reference to Takahashi here. Secondly, there is an important point that was intended and does need to be included. There are clearly issues with temperature inconsistencies and from the example of Takahashi et al. (2009) that may be exacerbated by the interpolation method. Thus, it is important that we have used a kriging method where the expectation value of the bias is zero. A more complete estimation of the bias in our results is calculated in Section 5.9. It was not possible to calculate the biases in identical ways because the two methods are different.

Response in paper:

As outlined in the reply above, an appropriate mention of the importance of consistent and unbiased temperatures will replace the slightly clumsy reference to Takahashi et al.

Referee #3: p 1913 l 1: Does min and max refer to space or between months?

Reply: The min and max refer to both space and between months. For each month the global grid box mean, -min and -max were calculated and then the mean of the mean, min and max over the twelve months were calculated. We will clarify this in the paper.

Response in paper:

As outlined in the reply above

Referee #3: Sect 5.7: Clarify (also in the headline) that you are talking about the auxiliary values here, not about fCO2.

Reply: We agree that this should be clarified.

Response in paper:

We will replace 'missing values' with 'missing values in the full fCO2 re-analysis due to missing auxillary values in the original SOCAT data'

Referee #3: p 1916 l 22: It is not transparent to me where the numbers in the square root come from. It would be easier if you used symbols (rather than numbers). (here and at several other places)

Reply: We agree that this should be clarified.

Response in paper: We will add after '… uncertainty.' on page 1916, line 10 "The total of *n* independent errors Δx_1, Δx_2, … Δx_n, is estimated by $\sqrt{(x_1)^2 + (x_2)^2 + \ldots + (x_n)^2}$."

***Referee #3:** Fig 11: Is this the fraction per cruise?*

Reply: We agree that this should be clarified. The fraction is the mean over all years per 1°x1° grid box.

Response in paper: The figure caption will be revised as outlined above.

REFERENCES

Donlon, C. J., P. Minnett, C. L. Gentemann, I. J. Barton, B. Ward, J. Murray, and P. D. Nightingale, 2002: Towards operational validation of satellite sea surface skin temperature measurements for climate research. *J. Climate,* **15**, 353-369.

Donlon, C. J., P. D. Nightingale, T. Sheasby, J. Turner, I. S. Robinson, and J. Emery, 1999: Implications of the oceanic thermal skin temperature deviation at high wind speeds. *Geophys. Res. Lett.,* **26**, 2505-2508.

Schuster, U., A. J. Watson, N. R. Bates, A. Corbiere, M. Gonzalez-Davila, N. Metzl, D. Pierrot, and M. Santana-Casiano, 2009: Trends in Nort Atlantic sea-surface fCO2 from 1990 to 2006. *Deep-Sea Res. II,* **56**, 620-629.