

## Response to Reviewer 1's Comments

Alain de Verneil      Zouhair Lachkar      Marina Lévy      Shafer Smith

June 16, 2021

We all would like to thank Reviewer 1 for providing their comments in reviewing this manuscript (MS). It is very clear from the detail and thoughtfulness of the comments and questions that a lot of time and effort were spent in the review, which is both fair and will help produce a better paper. Below, we reproduce the reviewer comments (lines preceded with >), with our responses interspersed within along with updated text in the MS shown in italics where necessary. Changes made to the original text show red and blue for deletions and additions, respectively.

>Review of the "Evaluating the Arabian Sea as a regional source of atmospheric CO<sub>2</sub>: seasonal variability and drivers" by Alain de Verneil<sup>1</sup>, Zouhair Lachkar<sup>1</sup>, Shafer Smith<sup>2</sup>, and Marina Lévy<sup>3</sup>

>This study analyzes the spatial and seasonal distribution of the air-sea CO<sub>2</sub> exchange in the Arabian sea. This study is relevant and original both because, on the one hand, it is carried out in the Arabian Sea; region poor in data and poorly known. On the other hand, this study covers a large panel of aspects both in space and time. They first validated their model to observational pCO<sub>2</sub> data and climatologies in addition to other variables such as SST, SSS DIC. The spatio-temporal variability of the CO<sub>2</sub> flux is then analyzed for the entire domain of the AS as well as for different subregions with a quantification of the wind speed and pCO<sub>2</sub> contribution. A further details analysis of the pCO<sub>2</sub> spatio-temporal variability is also performed with a finer decomposition into processes such as temperature, circulation (upwelling effect), biology, ... However, this study presents a lot of information that goes a bit in all directions with a lot of sentences without justifications behind their strategies (e.g., line 106 value of 2 matm) and a lack of quantitative values to support the text in the results and discussions sections. I recommend this publication in "Biogeosciences" after major revision.

>1

>Lines 25-31: clearly need quantitative values to support the text.  
>Lines 25-26: "and up to the present has on average acted to uptake excess anthropogenic CO<sub>2</sub> (Ciais et al., 2013; Khatiwala et al., 2009)" Should add a value  
>+ references need to be updated (e.g., GCP)  
>Line 27: "The Arabian Sea (AS) is a region of the ocean that has been found to naturally release CO<sub>2</sub> to the atmosphere (Sarma et al., 1998), mitigating the ocean's role in moderating atmospheric CO<sub>2</sub> accumulation." Reference 1998 is very old.  
>What about e.g., Cao et al 2020? What is the value of the AS source of Sarma?  
>Without a value, it is difficult to evaluate the contribution of the AS.

We agree that including quantitative values provides better context to motivate the paper. We've added values to these lines as follows:

The global ocean represents a major reservoir of inorganic carbon on the planet's surface (40x atmosphere), and up to the present has on average acted to uptake ~23% of the 11Gt excess anthropogenic carbon  $\text{CO}_2$  each year (Friedlingstein et al., 2020; Ciais et al., 2013; Khatiwala et al., 2009). The Arabian Sea (AS) is a region of the ocean that has been found to naturally release  $\text{CO}_2$  to the atmosphere ( $\sim 90 \text{ MtCyr}^{-1}$ , Sarma et al., 1998), mitigating the ocean's role in moderating atmospheric  $\text{CO}_2$  accumulation. While the AS as a regional basin is considered too small to greatly impact global budgets of air-sea  $\text{CO}_2$  exchange (Naqvi et al., 2005), it attracts attention because ~~some of the highest~~ rates of air-sea  $\text{CO}_2$  flux  $7\text{-}33 \text{ molCm}^{-2}\text{yr}^{-1}$  and values  $>700 \mu\text{atm}$  of partial pressure of  $\text{CO}_2$ , or  $p\text{CO}_2$ , have been observed there, in addition to unique features such as the world's thickest oxygen minimum zone (OMZ) (Lachkar et al., 2016) and corresponding Carbon Maximum Zone (CMZ) (Paulmier et al., 2011).

The reference from 1998 is old, indeed. Thank you for bringing Cao et al., 2020, to our attention. It's true that the Arabian Sea is a case study in this paper, but if you look into the methods they use the same 1995 JGOFS data as essentially all other data-driven studies. In addition, since that paper focuses on the nearshore regions  $\leq 100\text{km}$  of the coast (although admittedly a decent amount of the upwelling occurs there), it doesn't address the Arabian Sea as a whole.

---

>Line 65: "The need is further emphasized when quantities such as  $p\text{CO}_2$  can be utilized as a proxy for other things, such as community compensation depth (Sreeush et al., 2019b)." Not clear, need to be developed if you mention that in the text.

Yes, we included this to simply show the importance of better determining  $p\text{CO}_2$  and air-sea  $\text{CO}_2$  flux because these very same quantities are being used to estimate more complicated quantities related to an ecological state, compounding the uncertainty. We propose to expand upon it as follows:

*The need to reduce uncertainty is further emphasized when modeled carbon chemistry quantities such as  $p\text{CO}_2$  can be are utilized as a proxy for other things. For example, a recent modeling study in the AS found that  $p\text{CO}_2$  could be used to indicate such as community compensation depth, which reflects the complicated balance between primary production and respiration in the water column (Sreeush et al., 2019b). As a result, the possibility exists to propagate uncertainties beyond carbon chemistry.*

---

>Line 66: "However, most recent studies..." Should add references. And most recent >"modelled" studies or "observational" studies? Not clear

Yes, by "recent studies" we are referring to modeling studies, since observational carbon data is lacking. The citations in mind are the three just cited in the lines above. We've added the following to clarify:

*However, ~~most recent~~ these AS modeling studies compare...*

---

>Line 71-72: "Sarma et al. (2013) found that coupled ocean biogeochemical models underestimated the air-sea  $\text{CO}_2$  flux." should add "in the AS."

This change will be in the revised MS.

---

>line 75: "peak of flux". Which flux? Air-sea  $\text{CO}_2$  exchange? Advecting carbon flux? >Not clear

Yes, "peak of flux" refers to Air-sea CO<sub>2</sub> exchange, and will be changed to read:

*Additionally, Sarma et al. (2013) found that ~~the~~ peak air-sea CO<sub>2</sub> ~~of~~ flux observed in...*

---

>Line 83: should say "air-sea CO<sub>2</sub> flux."

This change will be in the revised MS.

---

>Line 84: "the quantification of the air-sea flux will focus on the contributing factors  
>of pCO<sub>2</sub> ...". Spatial contribution or seasonal contribution? Not clear

While the focus in this MS is on the seasonal cycle, geographic differences are also highlighted, at least for pCO<sub>2</sub> and air-sea CO<sub>2</sub> flux. However, to highlight the seasonal focus, we will amend the MS to read:

*Quantification of **seasonal** air-sea CO<sub>2</sub> flux will focus on the contributing factors of  $\Delta pCO_2$ , the difference in seawater and atmospheric pCO<sub>2</sub>, and wind.*

---

>Line 84: should say "deltapco2 and wind" instead of pco2 and wind. What about  
>the contribution of atmospheric pco2?

The change will be implemented in the MS (see directly above). Atmospheric pCO<sub>2</sub> is clearly a factor in determining  $\Delta pCO_2$  and has a seasonal signal of its own. However, in this region  $\Delta pCO_2$  is almost always positive (see Figure 2a inset), meaning that the magnitude of annual oscillations in atmospheric pCO<sub>2</sub> is generally not enough to reverse air-sea CO<sub>2</sub> flux direction. Thus, atmospheric pCO<sub>2</sub> variability will only impact the magnitude of flux. Considering that during the year sea surface pCO<sub>2</sub> can vary by  $\sim 200 \mu atm$  vs.  $\sim 4-6 \mu atm$  in the well-mixed atmosphere (range for Keeling curve), most of the variability will be driven by the sea surface.

---

>Line 84: should say "sea surface temperature". Same for salinity

We understand the confusion and the need for clarification, and note that the need for consistent phrasing is repeated in the rest of the MS. So, we propose the following changes:

*In particular, the role of **sea surface** temperature (SST), **sea surface** salinity (SSS), DIC, and TA in determining the seasonal cycle...*

For the rest of the MS, will use SST and SSS.

---

>Line 85: "which of course also varies from region to region within the AS". Should  
>say "for the entire domain of the AS as well as its spatial heterogeneity within the  
>AS".

The suggested changes will be implemented in the revised MS.

---

>Line 87: "biological production/respiration". What about calcification  
>remineralization or Nitrification denitrification?

CaCO<sub>3</sub> calcification, remineralization, and nitrification/denitrification are all processes included in the biogeochemical model used here. However, by use of the phrase "such as," the intention was to simply provide an example, not all of the physical vs. biological process comparisons that could be made. We will specify this in the revised sentence:

*...such as advection and mixing versus biological production and respiration, among others.*

---

>Line 91: "resolved by the data" Not clear which data? from which study?

We will change this to specify:

*resolved by the *in situ* data*

Generally speaking, in our revised text we will try to make sure data refers to the limited, in situ data, whereas model "data" will be referred to as "output"

---

>line 91 In the beginning of the introduction, you refer that data in the AS are very old (20 years ago) and here you refer to study from 2015. Not clear

Yes, we understand the confusion. In some instances, the "old" data from the 1990s are used in contemporary studies, such as the Cao et al., 2020 reference you have provided. Additionally, many studies, such as the references in this sentence, are all model results, and as mentioned in the previous response we will rectify this to specify these are model results, not "data":

*resolved by the *in situ* data, although models suggest interannual...*

---

>Line 92: "The rest of the paper" should use another sentence/to be reformulate.

We understand that this may be too informal, the sentence will be rephrased as:

*The ~~rest of the paper will start~~ study begins with a description...*

---

>2

>2.1

>Question: How many data in SOCAT and LDEO database? Monthly? Daily? annual values? what period of study? What is the spatial resolution? Do you include the coastal datasets?

The SOCAT database has 186,970 observations in the AS domain, whereas LDEO has 90,870, meaning SOCAT has roughly double the observations. The reason for this difference is that the LDEO database collects from a specific methodology for measuring pCO<sub>2</sub>, whereas SOCAT merges multiple methods. Since these are both collections of data from multiple cruises, there are no regularly recurring monthly/daily/annual observations. SOCAT includes observations from the years: 1962-3, 1991-2, 1995, 1997, 1999, and 2014. LDEO has observations from the years: 1962-63, 1980-81, 1991-2, 1995, and 1999. Being ungridded data, there is also no regular spatial resolution.

The hope was that the monthly and annual distribution of sampling in the AS would be visible from Figure 2b. All the data are included, and no difference is made between open ocean vs coastal sampling locations. In order to demonstrate how both the SOCAT and LDEO datasets cover similar areas and include coastal areas, please see the plot below showing the sampling locations.

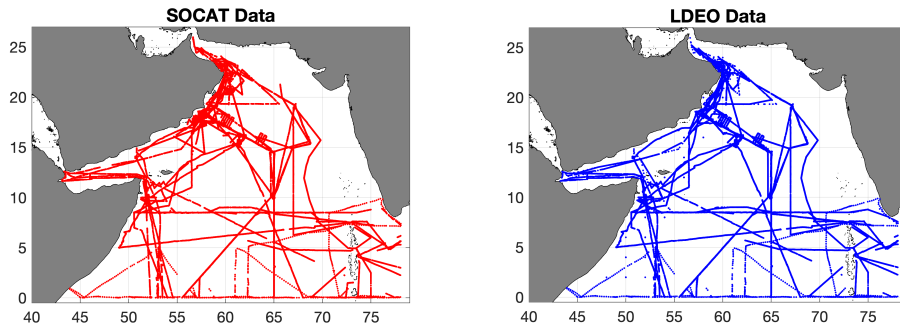


Fig. R1-1. Sampling locations for (a) SOCAT and (b) LDEO.

In the revised MS, we will include these numbers in Sect. 2.1 so that our choices regarding datasets (such as preferring SOCAT) are more transparently motivated.

>Question: paragraph from line 110: Why in addition to the Socat and LDEO  
>observation data, do you also add the climatology of tk09 and L15? what more do  
>they bring? And why you used the climatology of TK09 since others and more recent  
>are available (see e.g., Rodenbeck et al., 2015 which compare different pco2  
>climatologies). I do not understand why you choose TK09?

As mentioned in the introduction, since there are few in situ data, we made the effort to validate our model output with these data directly. The AS modeling studies cited in the introduction compare their models to TK09 as the validation step. Therefore, the comparison with SOCAT/LDEO pCO<sub>2</sub> is our attempt to be more thorough in our model validation, and the comparison with TK09 for air-sea CO<sub>2</sub> flux is an attempt to provide context that allows for comparison with previous modeling work in the AS, despite the fact that more recent climatologies are available such as Rödenbeck et al., 2015. Additionally, in Rödenbeck et al., 2015 it appears to be that only the southern portion of the AS is represented.

Regarding L15, even though it is also calibrated with available LDEO/SOCAT data, as a forward-feed neural network, we include it because it is an independent methodology to the model and TK09 with relatively high spatial resolution (1°x1°), and also because it is a de facto standard for global pCO<sub>2</sub> estimates, such as the GCP and IPCC reports.

Please see our response below to the question "Line 109: "The gridded products" monthly?daily?annual grid? For which period?" for our proposed text to clarify these choices of TK09 and L15, as well as to highlight why they (and the other pCO<sub>2</sub> datasets) are needed.

>Question: Is the coastal domain included in your definition of the AS domain? if yes  
>(e.g., line 157 you use the word coast), the pco2 climatology of Laruelle et al 2017  
>and Roobaert et al (2019) for the coastal flux should be considered here? You refer  
>to TK09 and L15 but what about the coastal regions?

The AS domain does include both the open ocean and coastal regions. The spatial resolution of TK09 is indeed problematic for the coasts, whereas L15 does include values within 100km of the coast. We appreciate that some climatologies are better suited for the open oceans whereas others are specific to coastal regions. Considering how the AS is ocean-dominated as a coastal region (the conclusion of the Cao et al., 2020 reference), and that the climatology of L15 and the in situ pCO<sub>2</sub> data include the coast, adding a sub-regional air-sea CO<sub>2</sub> flux analysis specifically for the coasts, while possibly useful, would also risk further confusing the reader with the addition

of another dataset. This concern arises especially in light of the amount of reorganization and clarification necessary for the major revisions in this MS.

---

>Comment: In the text, a clarification is need on the period in which you are going to  
>work + spatial resolution + temporal and how you have treated the different  
>datasets to switch, for example, to seasonal. Why are some databases used only for  
>model validation while others for the flux estimation (table 1)?

We are happy to elucidate further the period/spatial-temporal resolution in which we focus in the paper. This study is on the baseline seasonal cycle of pCO<sub>2</sub> and air-sea CO<sub>2</sub> flux. The model is initialized and forced with a mix of climatologically averaged data which do not belong to a specific time period, such as 1990-2011, apart from the time period that any data at all exist in the AS. As discussed in a later response below, the implied assumption is that the "baseline" seasonal pCO<sub>2</sub> and air-sea CO<sub>2</sub> flux has not changed considerably over the past decades. As a result, model output is designed to be representative of the AS during any of the past few decades. To this end, we will add to the MS:

*... pCO<sub>2</sub> values were calibrated to the year 2005, the representative year used for the model's atmospheric xCO<sub>2</sub>. The year 2005 was chosen for the model's xCO<sub>2</sub> concentration because it is the end of the historical period for the Intergovernmental Panel of Climate Change (IPCC) models. The earliest SOCAT data comes from 1962, and different databases used in this study stem from similarly different timespans. As a result, in using them this present analysis both assumes the existence of and attempts to quantify a baseline seasonal pCO<sub>2</sub> and air-sea CO<sub>2</sub> cycle that has remained stable over the past decades.*

Regarding data treatment, all calculations and manipulations on the model output take place at the monthly level, and are then aggregated into seasons to reflect the monsoon and inter-monsoon periods. To compare in situ data to the model output, we take each observation within a given month at a given lon/lat and interpolate the model's monthly pCO<sub>2</sub> value for the position. When calculating air-sea CO<sub>2</sub> flux from other datasets, again monthly values are used and then aggregated into seasons when desirable. This treatment to the in situ data will be added to the first paragraph of Sect. 2.1, and the treatment of comparison datasets is in a response below.

---

>Comment: This manuscript needs a paragraph about the spatial and temporal pco2  
>and FC02 uncertainties.

We agree that more needs to be discussed on the uncertainties in these quantities. The SOCAT protocol is designed so that observations with quality control flags A and B are within 2μatm. Within our dataset, this is the case for <20,000 observations, about 10%. Since wind is prescribed in our study, we assume there is no uncertainty there. Using the average SST and SSS from the A/B quality SOCAT observations, we get 2.68e<sup>-2</sup>mmolC/m<sup>3</sup>/μatm, resulting in changes of 0.0018, 0.0443, and 0.177molCm<sup>-2</sup>yr<sup>-1</sup> of deviation with 1, 5, and 10 ms<sup>-1</sup> winds, respectively. We will include these numbers in the revised MS to provide the reader some insight into uncertainty.

---

>Comment: I suggest moving section 2.4.3 before 2.4.1

We understand your suggestion, and those made later in this review, to re-organize the paper so that air-sea flux is presented first before pCO<sub>2</sub> variability. Granted, air-sea CO<sub>2</sub> flux is the putative focus of the paper, and so this order makes sense, and other studies such as Sarma et al. (2013) do the same. We ordered the paper this way primarily because a lot of text is dedicated to first validating the model through pCO<sub>2</sub>. Moving on to air-sea CO<sub>2</sub> flux after pCO<sub>2</sub> validation, followed by revisiting pCO<sub>2</sub> seasonality, felt less continuous in the narrative sense. However, we are open to the consideration and will make an attempt in the revision to follow this order.

>Line 98: "pco2" should specify that it is sea surface pco2

This will be changed in the revised MS.

---

>Line 100: "model pCO2 is calculated from DIC and TA" how?

The methods used in the model are discussed in the following section, so for the curious we will add the following:

*...model pCO<sub>2</sub> is calculated from DIC and TA (see Sect. 2.2), pCO<sub>2</sub>...*

---

>Line 101: "in situ un-gridded data merged from SOCAT v. 2019" why you do not  
>discuss about this dataset in the Introduction. In the introduction, you said that  
>there is in the AS only observational data from 20 years ago (JGOFS) and here you  
>discuss about LDEO and SOCAT. Not clear

Both LDEO and SOCAT contain (and in the AS mostly consist of) the JGOFS data. We would hope that this is self-evident, with both products being attempts at being collections of all available in-situ data of a given methodology. This is reflected in the fact that most of the in situ data we use come from the 1990s, as is visible in Figure 2b. We briefly mention SOCAT in the introduction to make the point that the AS is under-represented in part because of the lack of follow-up in data collection following JGOFS. To clarify the point, we will add that SOCAT and LDEO both contain the JGOFS data and whatever few data there exists outside of JGOFS:

*Here, pCO<sub>2</sub> validation stems from in situ un-gridded data merged from SOCAT v. 2019 (downloaded from <https://www.socat.info/index.php/version-2019/> September 2019) and the Lamont-Doherty Earth Observatory (LDEO) surface pCO<sub>2</sub> database (Takahashi et al., 2019). Both databases aggregate all available in situ surface pCO<sub>2</sub> data, including JGOFS.*

---

>Line 103: "SOCAT data was preferred to LDEO observations" why?

In answering your first question for Sect. 2, we addressed this concern by showing that SOCAT is twice as dense, which we will add in the revised MS:

*Due to the large overlap in the two databases, SOCAT and LDEO contain >180,000 and ~90,000 data in the AS, respectively. SOCAT has more data because it includes multiple methodologies. As a result, SOCAT data ~~were~~ preferred to ~~LDEO observations~~, and LDEO observations were included for years where SOCAT data are unreported.*

---

>Line 103: "large overlap" how many data? you should be more quantitative and more  
>precise in the words used.

As seen in the response above, the "large overlap" argument will be removed from the MS, and we appreciate the need to be more precise in our wording.

---

>Line 105: "... using reported sea surface temperature (SST) and S data included in  
>the products" S= sea surface salinity? which equation do you use to convert fco2 to  
>pCO2? Not clear how fugacity and partial pressure are related to SST and S

As addressed previously, S will become SSS for sea surface salinity. The equations to convert fugacity to pCO<sub>2</sub> are included below for your reference. Only SST is needed for the conversion from fCO<sub>2</sub> to pCO<sub>2</sub>, whereas salinity enters the picture for the vapor correction in converting



pCO<sub>2</sub> to xCO<sub>2</sub>. Since these equations are both somewhat long due to their thermodynamic nature and are not particularly illuminating for the reader, we will include a reference in the revised MS:

*...reported ~~sea surface temperature (SST)~~ and SSS data included in the products using routines from the CO2SYS software package (Van Heuven et al., 2011).*

*Here are the equations:*

$$pCO_2 = fCO_2 \cdot e^{-(b+2\delta) \cdot P / (R \cdot (SST+273.15))} \quad (1)$$

where *b* is:

$$b = -1636.75 + 12.0408 \cdot (SST+273.15) - 0.0327957 \cdot (SST+273.15)^2 + 3.16528 \cdot 10^{-5} \cdot (SST+273.15)^3 \quad (2)$$

and  $\delta$  is  $(57.7 - 0.18 \cdot (SST+273.15))$ . *P* is 1.01325 bar (1 atm), and *R*=83.1451 ml bar K<sup>-1</sup> mol<sup>-1</sup>.

>Line 106: why a value of 2 muatm yr-1. Should be justified.

We apologize for the wording, the anthropogenic effect on pCO<sub>2</sub> was calculated to be 2 by linear fit to the merged data; this value was not prescribed a priori. The revised version will be:

*The anthropogenic effect of increasing surface pCO<sub>2</sub> was removed by calculated ~~by~~ removing a fit linear trend of 2 muatm yr<sup>-1</sup>, slightly...*

>Line 109: "The gridded products" monthly?daily?annual grid? For which period?

We understand that each dataset used in this paper should contain information regarding its spatiotemporal resolution and scope. For example, both TK09 and L15 are global, monthly climatological products, but have different calibration years of pCO<sub>2</sub>, as well as slightly different decadal spans. This information, all written out, becomes both hard to follow and tedious, which is why Table 1 is included. We will re-write this paragraph so that at the beginning the reader will be directed to Table 1 where the pertinent information (reference xCO<sub>2</sub> year, gridded or not, spatiotemporal resolution, use case, etc.) for each dataset is visible:

*Alternative pCO<sub>2</sub> products are used for comparison purposes. A complete list of these datasets and their characteristics is provided in Table 1. For all the comparison datasets, air-sea CO<sub>2</sub> flux is calculated from monthly values. ΔpCO<sub>2</sub> values are calculated using Keeling curve values (downloaded from <https://www.esrl.noaa.gov/gmd/ccgg/trends/gI-data.html>, downloaded September 2019) of atmospheric xCO<sub>2</sub> for the respective calibrated year of each data set. The same climatological winds as used in the model (Sect. 2.2) are applied to the pCO<sub>2</sub> products. The gridded products ~~from~~ TK09 is chosen because previous modeling studies in the AS use it as validation (see Introduction). ~~and~~ The L15 climatology, while based upon the same in situ data mentioned above, represents different processing methodologies, and as a high-resolution, global pCO<sub>2</sub> dataset, also serves to provide independent context to the model validation. pCO<sub>2</sub> is also calculated from DIC and TA provided by the statistical fits to JGOFs data by Sarma (2003) and to the gridded GLODAP climatological product. The statistical fits of Sarma (2003) were used twice, first using model monthly SST, SSS, and Chl-a, and second with World Ocean Atlas (WOA) 2009 monthly SST, SSS with SeaWiFS Chl-a. GLODAP-derived pCO<sub>2</sub> also uses WOA2009 monthly SST, SSS applied to the annual DIC, TA values. Calculations of pCO<sub>2</sub> are performed using the*



CO2SYS software package (Van Heuven et al., 2011). Since all calculations are conducted at the near-surface, differences between this software suite and Orr and Epitalon (2015) are minimal. ~~For air-sea flux calculations, all  $\Delta pCO_2$  values were calculated using Keeling curve values (downloaded from <https://www.esrl.noaa.gov/gmd/ccgg/trends/gI-data.html>, downloaded September 2019) of atmospheric  $xCO_2$  for the respective calibrated year of each data set (1995 for Sarma (2003), 2001 for L15, 2002 for GLODAP, 2005 for TK09). A summary of these datasets and their characteristics is provided in Table 1.~~

---

>Line 115: Which products/datasets do you used to calculate the air-sea flux?

Please see the response above.

---

>2.2

>Line 121: "System-AGRIF (ROMS-AGRIF) 3.1.1." version 3.1.1? what means AGRIF?

AGRIF is an acronym for "Adaptive Grid Refinement In Fortran," which is a tool to create one and two-way nesting of model grids. Though this particular capability is not used in this study, we include it as part of the name. The full name, along with the explicit "version" will be added to the MS.

---

>Line122: "Previously used in the region (Lachkar et al., 2016), the..." should be by >Lachkar et al.,

This change will be implemented in the revised MS.

---

>Line 125: "advection of tracers" in repeat twice in the same sentence

The repeated phrase will be deleted in the revision.

---

>Line 125: what is the K-profile?

>Line 126: what is KPP?

K-profile, and K-profile parameterization (KPP), is a vertical mixing scheme, described in Large et al., 1994. A real description of it is well beyond the scope of this forum, but the K stands for diffusivity, and the theory revolves around K's vertical profile and its relation to the gradient Richardson number. Since KPP is relatively well-established in the field, we leave it to the reader to look at the reference for further information.

---

>Line 127: 33° W or E?

We apologize, it is 33°E, and this will be added to the MS.

---

>Line 127: which previous studies? Need references.

This choice is primarily to match with Sarma et al. (2013), which at one point in the analysis explicitly considers the AS north of the equator. This will explanation will be included in the revised version:

*For the sake of comparison with ~~previous studies~~ Sarma et al., (2013), we will...*

---

>Line 127: sometimes you speak at present sometime in future→same comment for  
>all text.

We understand that sometimes in the course of writing that tenses can be wrongly intermixed, which further confuses meaning. The revised MS will be re-read carefully to make sure there is consistency in grammatical structure.

---

>Line 130: what is the name of the OBGCM model?

The biological model has no colloquial name based off an acronym, such as PISCES or NE-MURO. Internally, we refer to it as N<sub>2</sub>PZD<sub>2</sub> since there are two nutrient and detrital components, along with a single component each of phyto- and zooplankton.

---

>Line 131: "Biological parameters for the model are the same as those used in Gruber  
>et al. (2011)." That is to say? Which parameters? Should be enumerated.

We sympathize with the reviewer that modeling papers often simply include a reference and expect the reader to wade through the details. To this end, we will add a Table that includes these values for reference.

---

>Line 136: what about nitrification and denitrification?

Both nitrification and denitrification are included in the model; this is implied by "removal and creation" or nitrate, but we will include them explicitly in the revised version:

*TA changes with the removal and creation of nitrate (NO<sub>3</sub>), including nitrification and denitrification, as well as dissolution/precipitation of CaCO<sub>3</sub>*

---

>Line 144: "climatological mode". Not clear, what is the period?

Apologies, "climatological mode" is short-hand for running the model with climatological, averaged forcing. Since it is confusing, this will be removed in the revised MS:

*The model was run ~~in climatological mode~~, with 360-day years and interpolated, climatologically averaged monthly forcing.*

---

>Line 145: "surface salinity". In the text before it was S. Not clear

Similar to previous responses, this will now be SSS in the revised MS.

---

>Line 146: SST is from 1985-1997 but then you make a FC02 comparison for  
>another period. Still confused about the period of this study. What is the spatial  
>resolution of SST?

>Line 147: Why do you need wind stress? The period of the wind stress is  
>1999-2009, different than SST. Not clear again.

The questions posed are quite logical and pertinent. Indeed, the SST forcing is based from a climatology dating from 1985-1997, and the climatology for wind stress comes from a separate period, from 1999-2009. In the ideal, this makes no sense, and every source of forcing in the model should correspond to the same time period of study. Unfortunately, this is not the case as surely the reviewer is aware. However, in using these climatological forcing datasets, we understand the

need to explicitly state the underlying assumption that the dynamics in the forcing variables have not changed. This assumption is stronger for some variables and less so for others. For example, with ongoing climate change mean SSTs have probably changed (increased) in the time since 1997, whereas the strength and direction of summer monsoon winds have not changed much. In order to more directly state our assumption, in the revised MS we write (starting Line 144):

*... monthly forcing. The different climatological products derive from datasets spanning slightly different periods, and so here we assume in this study that the dynamics represented within them have not changed in the time since. Heat flux, ...*

The SST spatial resolution is  $\sim 0.1^\circ \times 0.1^\circ$ , less than 10km at the equator. Wind stress is the form in which wind forcing is implemented in ROMS (of course, u/v wind speeds can be converted to wind stress), and so it is necessary to run the model.

---

>Line 148: "T, S" same comment than before. SST or T .. not clear

Yes, thank you for keeping track, this will be remedied in the revision.

---

>2.3

>Line 157-158: need reference from the literature to support summer monsoon and enhanced biology.

Please find an added reference below:

*...but also enhanced biological productivity (Schott and McCreary Jr, 2001).*

---

>Line 162: same comment need reference for winter monsoon NPP

Please find an added reference below:

*...where the winter monsoon's primary productivity is most intense Kumar et al., 2001).*

---

>Figure 1: what is the period of the study?

The SeaWiFS mission spanned from 1997 to 2010. As mentioned in a previous response regarding climatologies, the assumption inherent in this comparison is that average Chl-a does not change much outside this time period.

---

>2.4

>2.4.1

>Line 168: "The proximate variables impacting pCO<sub>2</sub> in the" should be "..that affect pco<sub>2</sub> change in the"

The wording will be changed to reflect this in the revised MS.

---

>Line 170: What is the depth of the water column where the pco<sub>2</sub> decomposition into processes is performed in the model?

The decomposition is calculated for the topmost cells in the grid. Since ROMS has an sigma-coordinate system in the vertical, these cell depths vary but the average is 2.6m deep.

>Line 171: In equation 1, terms on the RHS are not described in the text and not clear  
>when you said "deviation from an average."

Yes, you are quite right, the partial derivatives are not described. Additionally, we will clarify what we mean by "deviation from an average." The revision will be:

*...likewise express deviations from ~~an average~~ prescribed value depending on whether the deviations are spatial or temporal in nature (see below). The coefficients of the  $\Delta$  terms are partial derivatives of  $pCO_2$  with respect to these variables, namely DIC, TA, SST, and SSS, and are calculated via centered differences described below. However, in order...*

---

>Line 181: how partial derivative are calculated? Which coefficient are used?

The description of how partial derivatives are calculated starts at line 190. To better illustrate the calculation, we will include an example calculation:

*...determined by Orr et al., (2018). For example, to calculate the monthly  $pCO_2$  anomaly due to SST for a gridpoint with annual mean  $pCO_2$  of  $430 \mu atm$ , annual mean SST of  $24^\circ C$  and monthly SST of  $26^\circ C$ :*

$$\Delta pCO_2 \approx \frac{\partial pCO_2}{\partial SST} \Delta SST + \dots \approx \frac{pCO_2(24 + 1 \cdot 10^{-4}, \dots) - pCO_2(24 - 1e - 4)}{2 \cdot 1e - 4} \cdot (26 - 24) + \dots \quad (3)$$

*where  $1 \cdot 10^{-4}$  is the recommended SST deviation.*

---

>2.4.2

>Line 197-198: what about the river flux? Did you include the contributions from  
>rivers?

Yes, indeed, riverine flux impacts salinity as well as carbon variables such as DIC and TA, and certainly there are freshwater inputs coming from India during the summer monsoon. However, these fluxes are not directly implemented as land/ocean boundary conditions. Instead, DIC and TA are altered by virtual fluxes that are proportional to SSS changes as prescribed by the COADS climatology. We will note this at the appropriate place in line 139:

*Surface fluxes of DIC and TA due to evaporation, ~~and~~ precipitation, and river input were included as virtual fluxes proportional to SSS forcing.*

---

>Line 202: Eq 6 what are "New" and "Reg" and why  $CaCO_3 precip.remin$  has the same  
>sign than  $PPnew+reg$  (both with a -)?

"New" and "Reg" are the two forms of primary production, "new" production from nitrate assimilation and "regenerated" production from ammonium. Primary production and  $CaCO_3$  production all remove inorganic carbon from the water column, so their contribution to the DIC budget is negative.

---

>2.4.3

>Line 226:"five-year average" why five? Not clear

The five-year average is necessary because of the definition of the Reynolds decomposition. In order for the equations in the decomposition to be exact, the mean needs to be defined against the entire timeseries. We use five years of model output to account for internal variability in the model, and as a result the mean needs to be calculated around the entire five years of output. We will make reference to this in the revised version:

*...where ' indicates an anomaly and  $\bar{x}$  is a five-year average of variable  $x$ , which are calculated at each grid point. The five-year average is necessary for exact closure in the Reynolds decomposition.  $F'_{CO_2}$  is the...*

---

>3

>3.1

> Question: Why did you analyze the variables in depth in Fig. S1 and Fig. S2 and not just the spatial distribution at the sea surface as it is performed for Fig. S3? What is the reason for the choice to do in depth?

There was no conscious decision to show depth profiles in temperature and salinity but not with currents, but it is rather a consequence of data availability. Firstly, sea surface spatial distributions of temperature and salinity would reflect the surface forcing, whereas distributions of temperature and salinity at depth better demonstrate model performance and there is at least data available. The only measurements for currents stem from drifting floats that reflect the near-surface, and so this is the comparison we are forced to use.

---

>Question: I do not understand between figure 2 and figure 3 for the data spatial distribution. Are they the same data between both figures? It seems there is more surface covered by the data on figure 2 than when they are divided into seasons (figure 3)?

It is understandable that figures 2 and 3 might visually give the impression that these are not the same data. In figure 2, all available data are binned into a  $1^\circ \times 1^\circ$  grid, which in itself inflates the sense of data coverage and may be the cause for the dissonance.

---

>Comment: Lots of imprecision on the form and problems with the numbering of the panels. e.g., Figure 3. "Seasonal surface pCO<sub>2</sub> (muatm) from data (a-d)" panels are not ok between the figures and the caption.

Yes, many apologies, this caption is reflective of the initial distribution of subpanels, which was changed last-minute to maximize figure size in the submitted MS since .png files are not vectorized. The revised caption will reflect the updated placement of subpanels.

---

>Comment: the pco2 Seasonality has already been analyzed by the previous figures and compared to the SOCAT database spatially and temporally with a Taylor diagram. Why do a comparison to LA15 again? What more does this bring? not clear in the text

Figure 5 is included as part of the validation process because while Figures 3 and 4 show comparisons with the data (visual in Figure 3, quantitative in the Taylor diagram of Figure 4), the model's performance still needs further context. For example, Figure 3 visually demonstrates that the model reproduces summertime pCO<sub>2</sub> near Oman, while over-estimating pCO<sub>2</sub> most of the time elsewhere. Figure 4 shows that correlations are quite low and that model output's variability is less than the data's. Considering these weaknesses, does the model represent an improvement?

This is where Figure 5 shows that the model better captures the extremely high summer values seen in SOCAT than an alternative, popular pCO<sub>2</sub> product. We will revise the text to highlight the motivation behind this figure in the Results section:

*Direct comparisons between the in situ data and model output demonstrate the positive bias and middling correlations of the model with respect to the data, as well as the model's tendency to under-represent variability. As a result, it is necessary to investigate how these shortcomings compare with alternative pCO<sub>2</sub> estimates in the AS. Figure 5 shows monthly comparisons of the pCO<sub>2</sub>. Comparisons between the monthly probability distribution functions of pCO<sub>2</sub> from in situ data, model output, and L15 demonstrate temporal variability between pCO<sub>2</sub> products (Fig. 5). For most of the year, the data (Fig. 5a) stays within a relatively narrow range (375-425 μatm), except for the summer monsoon where values can exceed 500 μatm and the median value has its peak. In the model ...*

---

>Comment: I do not understand the added value of the Taylor diagram. What more >does this figure add? moreover, you do not explain the various terms of the taylor >diagram such as the Pearson coefficient. Why is it not done on the other variables, >especially when you go directly to line 255: Since the model successfully replicates >other tracers, physics, and biological processes". A little fast, no? In addition, the >seasonality of these processes is not analyzed.

The Taylor diagram is added because it is a compact, visual way of representing multiple statistical quantities of interest to score a model's fidelity. As you have mentioned, it contains the Pearson correlation coefficient, as well as normalized standard deviation with model bias added as a color. In the revised MS we will include an explanation of the terms in the Taylor diagram:

*A Taylor diagram (Taylor, 2001) comparing in situ pCO<sub>2</sub> data and model output shows the relative performance of the model (Fig. 4). The distance from the origin,  $r$ , is model variability normalized by standard deviation of in situ data. The angle created from the y-axis,  $\theta$ , is the Pearson correlation coefficient between the model and in situ data. If the model were to perfectly reproduce the data, it would appear at the position (1,0), equivalent to a normalized standard deviation of 1, and correlation coefficient of 1. For the entire dataset...*

We think that doing a Taylor diagram on other variables, such as SOCAT SST and SSS, would be a wonderful addition. First, it would show the relatively high performance of the model in representing the physics, and provide comparison demonstrating the difficulty in reproducing pCO<sub>2</sub>. As a result, the model's apparently lackluster performance in pCO<sub>2</sub> would better motivate further comparison with L15 in Figure 5. This way, we can provide better model validation of other tracers and physics without seeming to gloss it over, and also better motivate the three figures of validation done for pCO<sub>2</sub>.

---

>Comment: You start with fig S1 to S3 discussing variables then you discuss about >pco2 and then you discuss again with variables (line 255). Not clear. Should be first >pco2 then variable or vice versa

Yes, we understand that the text should not move from one subject to another and back again. Jumping back to SST, SSS, TA, and DIC is part of an attempt to attribute the source of model bias. Understandably, this should all be part of the validation process, and the text will be moved before discussion of pCO<sub>2</sub> spatial and seasonal distributions.

---

>Line 230 "The implementation of ROMS-AGRIF presented here has been used in >previous studies (Lachkar et al., 2016)." And? What is the conclusion of this >sentence? Is it in the AS?

Yes, we should be more explicit. In saying "the implementation presented here", we meant this specific implementation in the AS that has been published in previous studies. We will be more precise in our wording in the revised MS:

*The implementation of ROMS-AGRIF presented here has been used in previous studies of the AS (Lachkar et al., 2016).*

---

>Line 233: Add panels with the mismatch (model minus data) in FIG S1 to S3 would be useful to highlight difference between model-data. Same comment for figure 3.

This is a good idea, especially since the data in Figure 3 can be crowded together and colors are difficult to directly compare between subpanels. As a result, we will add subpanels showing model minus data for Figs. S1-3 and Figure 3.

---

>Fig. S 2. "Transects, similar to Fig. S1, but with salinity instead of temperature" should be temperature.

Thank you, the typographic error will be corrected in the revised MS.

---

>Lines 230-233: e.g. "Reproduce well the ..." need to add quantitative values.

Understandably, quantitative values are needed so that the reader can judge the validation for themselves apart from visual inspection of Figs. S1-S3. We will add values to the text in reference to the subpanels we will add showing model - data, as well as correlations and standard deviations from the Taylor diagrams to be added.

---

>Line 235: "Regarding pCO<sub>2</sub>, in situ data from the merged SOCAT/LDEO database shows that average binned pCO<sub>2</sub> values in the region are positive for most of the AS (Fig. 2a). The ensemble of observations show that 90% of pCO<sub>2</sub> observations are positive, indicating positive flux to the atmosphere (Fig. 2a, inset)." These two sentences should be merged in one.

Understood, these two sentences will be combined as follows:

*Regarding pCO<sub>2</sub>, in situ data from the merged SOCAT/LDEO database shows that ~90% of average binned  $\Delta pCO_2$  values in the region AS are positive (Fig.2a, inset), indicating a positive flux to the atmosphere that is applicable geographically (Fig.2a). for most of the AS (Fig. 2a). The ensemble of observations show that ~90% of  $\Delta pCO_2$  observations are positive, indicating positive flux to the atmosphere (Fig. 2a, inset).*

---

>Line 239: "pCO<sub>2</sub> value..." we do not know if it is modelled pco2 value or observational pco2 value. Maybe you could use "data-based pco2" and "modeled based pCO<sub>2</sub>" everywhere in the text.

Understood, as mentioned in an earlier response, efforts will be made in the revision to specify in situ pCO<sub>2</sub> data versus model output, including in this section.

---

>Line 242: Again, this paragraph needs quantitative values. In the pco2 model description, there is no value in the text. Use values instead of "Is lowest".



Yes, we understand the desire for more quantitative values. To this end, in addition to the subpanels suggested for Fig. 3, we will add a table that contains the means and standard deviations for both data and model so that there is more transparency in the results and so that the results section of the paper is not cluttered with repetition of values.

---

>Fig. 4 "Taylor diagram of modeled vs. observed surface pCO<sub>2</sub>, both in total and seasonal sub-sampling. Data are from merged SOCAT and LDEO databases, corrected to year 2005. Distance from origin .." should be from

This typographical error will be corrected in the revised MS.

---

>Line 260: "close to reported measurement error" give a value.

Yes, there needs to be more context. As stated in an earlier response, the best accuracy for SOCAT data is  $\sim 2\mu\text{atm}$ , and this quality applies to only 10% of the current database. We will revise this sentence to better reflect reality:

*...results in a pCO<sub>2</sub> shift of -6.8 and -3.5  $\mu\text{atm}$  for SST and SSS, respectively, close in magnitude to the best-case ~~reported~~ measurement error of  $\sim 2\mu\text{atm}$ .*

---

>Line 255: "tracers, physics, and biological processes" which are?

We will revise this to state:

*Since the model successfully replicates ~~other tracers, physics, and biological processes~~ temperature, salinity, surface currents, and seasonality of primary production (Fig. 1, Fig. S1-S3), we look for the source of bias...*

---

>Line 257: Why the SST from SOCAT is used here from the socat database while in Fig. S1 it derives from the WOA? Not clear

SST and SSS are used from SOCAT because 1) this is as close as possible to the conditions of in situ data collection, and 2) while it is possible some of these data are used in generating the WOA, it is somewhat independent because WOA is used to initialize the model (even though satellite data forcing at the surface is probably more important). Additionally, in Fig. S4 we are hunting for the bias in these parameters at the SURFACE, whereas in Fig. S1-S2 a better measure of model performance is at depth, necessitating the use of the WOA dataset.

---

>Line 258: "sst" sometime sst sometime T. Not clear. Same comment for every variable e.g. Fig S4, SAL instead of S or SSS. Be consistent all text and figures.

Yes, thank you for pointing another instance of the terminology not being consistent, in the revised version it will be SSS.

---

>Fig. S4. same comment SST or temperature... not clear

Thank you, this will become SST in the revised MS.

---

>Line 267: Figure 5, I suggest merging the three plots in one and use monthly boxplot for instance for median, 5th and 95th. Easier for the reader.

Thank you for the idea, we are willing to alter Figure 5 so that it is all in one plot and consisting of monthly boxplots for the three datasets.

---

>Line 277: "competing product" Not clear. what that mean? Observational? Model?  
>Not clear

The revised text will be more precise:

*...the model captures the high summer monsoon pCO<sub>2</sub> values better than the competing alternative L15 climatology products.*

---

>3.2

>3.2.1

>Comment: need to be more quantitative.

In order to be more quantitative, we propose to add a table for Sect. 3.1. For this section, we will add more values:

*Spatial pCO<sub>2</sub> anomalies calculated from the regional annual mean highlight the geographic hotspots of pCO<sub>2</sub> inside the domain (Fig 6a). pCO<sub>2</sub> anomalies range from -89 to +415µatm, indicative of a positive skew in the distribution. Within the regions of analysis prescribed in this study, it is clear that Oman, the Indian coast, and the North AS including the Gulf of Oman host enhanced pCO<sub>2</sub>, with average positive anomalies of 8.6, 21.5, and 49µatm, respectively. In contrast, both the oligotrophic central AS and Somalia regions have negative pCO<sub>2</sub> anomalies (-13.7 and -2.9µatm). The contributing factors to these pCO<sub>2</sub> anomalies, temperature, DIC, TA, and freshwater components, display differing distributions. Temperature (Fig. 6c) contributes toward negative pCO<sub>2</sub> anomalies in a southwest-to-northeast band along the coasts of east Africa and the Arabian peninsula, up to the coasts of Pakistan and northern coast of India near Gujarat. The cold SST structure contributes to a -20µatm effect on pCO<sub>2</sub>, and largely overlaps the stronger summer monsoon winds (Fig. S5). The opposite trend is found in the central oligotrophic and Indian regions, where the average temperature contributiones to pCO<sub>2</sub> is 20µatm positively to pCO<sub>2</sub> anomaly, despite upwelling in the southern Indian coast. The distribution of DIC-induced anomalies (Fig. 6d) shows a positive influence near coastal regions and the western AS off the coast of Somalia (+25µatm), whereas a strong minimum is found in an oval region encompassing the central, open-ocean AS(average -36.6µatm effect). TA effects (Fig. 6e) show a north-south gradient with positive contributions to pCO<sub>2</sub> occurring in the north and negative towards the south, in a similar distribution to surface salinity gradients in the AS and with a +20 to -20µatm difference, similar to temperature. Freshwater contributions (Fig. 6f) show a similar distribution as TA, but weaker in magnitude ±10µatm for all regions.*

---

>Comment: You should be careful when you use the term "dpcO2" for change in pco2  
>spatial or temporally. You also use dpcO2 for the air-sea CO2 gradient. Not clear  
>e.g., Fig. 6 and Fig.2.

True, in the scientific literature  $\Delta p\text{CO}_2$  refers to the air-sea CO<sub>2</sub> gradient used in the calculation of air-sea CO<sub>2</sub> flux, and the casual reader may completely misinterpret the meaning of Fig. 6 and 7. As a result, the figure labels will be changed to read pCO<sub>2</sub> anomaly instead of  $\Delta p\text{CO}_2$ .

>Comment: maybe specify in the text what corresponds to a positive anomaly and a  
>negative anomaly on pco2

In asking what a positive vs negative anomaly corresponds to, perhaps this means to better define what the averaging process is. The map of spatial anomalies is derived by subtracting the spatial average of pCO<sub>2</sub> from the map of temporal average (annual mean) pCO<sub>2</sub>. Thus, a positive anomaly indicates a region has higher pCO<sub>2</sub> than the domain as a whole. Please see the proposed revision above for the specification of positive vs negative anomalies.

---

>Line 286: T, S alk are variable but FW is a process. Maybe salinity instead of FW?

We understand that the phrasing for freshwater is more evocative of a process as opposed to a variable. We have no objections to using salinity, but it is necessary to stress that a change in pCO<sub>2</sub> due to "salinity" includes effects where DIC and TA increase linearly with salinity.

---

>Line 284: In fig 6, maybe add numbers to regions for each box, then in panel b also add the number instead of name. Easier for the reader to say in panel b where is the regions in panel a

If the reviewer finds that numbers instead of the names to be useful for a quick reference, we as the authors have no strong objection to this and the change will be made in the revised figure.

---

>Line 286: "The cold SST structure largely overlaps the stronger summer monsoon winds." Not clear

Yes, in the text we simply state this, but do not show it. We will add a supplementary figure that shows the summer wind stress forcing used in this study to show that indeed, the colder SST matches with the strongest winds in the AS.

---

>Line 301: "The previous section outlines why certain geographic regions within the AS have overall high or low pCO<sub>2</sub> values" This does not explain why. Just how the variables change within different regions but not in terms of process the section does not explain why. It DESCRIBES the pCO<sub>2</sub> changes associated with the variables but not processes; pay attention to the word used. Question: what are the processes that explain the spatial distribution of pCO<sub>2</sub>? Not clear

Yes, thank you, we take your point that these variables are descriptive, perhaps indicative, but not explanatory in the way of processes or mechanisms. We will re-word this so as not to mislead readers:

*The previous section outlines ~~why certain~~ geographic regions within the AS that have overall high or low pCO<sub>2</sub> values, but in order to investigate the strong seasonal monsoon ~~impact cycle~~ in the AS, the decomposition of ~~variables~~ factors affecting monthly pCO<sub>2</sub> values is calculated at each model grid point and averaged into each analysis region (Fig. 7).*

Reviewer 2 has also stressed the importance of highlighting processes. They have suggested that in addition to a detailed analysis of DIC fluxes, to look at TA and SST so that the overall impact of mechanisms such as advection/mixing, primary production, CaCO<sub>3</sub> calcification/dissolution, etc. can be linked first to SST, SSS, TA, and DIC, and then ultimately to pCO<sub>2</sub>. Until this analysis is implemented in the revision, the answer to your question is not immediately clear.

---

>3.2.2

>Comment: Again, this section needs numbers to support the text. E.g., what is the typical amplitude of the respective seasonal signal?

Similar to section 3.2.1, numbers will be added to this section, such as the extrema of domain-wide pCO<sub>2</sub> anomalies (-9.3 to 20.3 μatm), DIC contributions to pCO<sub>2</sub> in Oman (-40.7 to 81.1 μatm), etc.

---

>Question: How do you explain that the important contribution of DIC in the upwelling regions is not seen on the seasonality of the entire domain (Fig. 7a)

Yes, considering how strong the upwelling signal in DIC is near the coast of Oman, it is curious that there is not more of an impact on the domain as a whole. The simplest answer is that the upwelling region both consists of a smaller area (Oman region is 5% of the domain total), and that excess DIC is not exported as much from the region (see Horz. Circ. in Fig. 8b).

---

>Line 302: "the decomposition of factors affecting monthly" should be "variable" instead of "factors"?

The word "factors" will become "variables" in the revised ms.

---

>Line 303: Figure 7. "Timeseries of pCO<sub>2</sub> anomalies (μatm)" units should be μatm month<sup>-1</sup>

We understand that these are monthly anomalies, but they do not represent fluxes, and we disagree that the units should be changed to μatm month<sup>-1</sup>, and suggest that it remains μatm.

---

>Line 304: "pCO<sub>2</sub>" should be called "pco2 anomaly"

This change will be implemented in the revised MS.

---

>Line 305: "DIC acts" DIC is not a process. Should be "change in pCO<sub>2</sub> associated to DIC change"

Yes, we understand, the language will be changed to reflect this and not mislead the reader.

---

>Line 307: you first discuss about Fig.7a in the text for the entire domain. Then you discuss about the oligotrophic central region. Hence, Fig.7d should be Fig. 7b. The different panels in figure 7 should be in order of what they are called in the text.

Fair enough, in the revised MS we will change the order of subpanels to reflect when they are mentioned in the text, so that the order will be: Domain, Oligotrophic, Oman, India, Somalia, and the North AS.

---

>Lines 320-326 Maybe this section should be in the discussion section instead of the result section.

These lines were added to briefly summarize the section's main points, but we are amenable to move this to the Discussion section, especially considering that the Reviewer suggests reducing the Results section.

---

>Line 328: "is controlled by the physical processes of surface forcing, mixing and advection" not clear what means the term "surface forcing"?

Surface forcing refers to heating/cooling. We will use this wording in the revised MS:

*Whereas SST and its effect on  $p\text{CO}_2$  is controlled by the physical processes of surface ~~forcing~~ heating and cooling, mixing ...*

---

>Line 330: Figure 8. "Timeseries of DIC fluxes ( $\text{PgCyr}^{-1}$ )" why the DIC flux is in  $\text{Pg C yr}^{-1}$ ? Should be in  $\text{Pg C month}^{-1}$ ? Why do you call this variable "DIC flux" instead of ">delta DIC" such as "deltapco2" before?

The unit of DIC flux in  $\text{PgCyr}^{-1}$  was chosen to match the presentation of monthly air-sea  $\text{CO}_2$  flux in Sarma et al., (2013). However, since in Fig. 12 the implied rate is in  $\text{month}^{-1}$ , we will change this to be consistent throughout the MS.

We call this variable DIC flux because, well, it represents the amount of DIC, in units of mass, that are added or removed per unit time through the volume's 2D boundaries or in situ due to biogeochemical processes. Considering that  $\Delta p\text{CO}_2$  is to be removed in favor of  $p\text{CO}_2$  "effect" or "deviation," there is no need to use  $\Delta\text{DIC}$ .

---

>3.3

>Comment: Again, this section needs numbers and more precise.

Similar to Sects. 3.1 and 3.2, we understand this need and will add appropriate numbers to the presentation.

---

>Line 330: "Budgets of DIC fluxes in the upper 20 m (Fig. 8) show that two major >processes dominate, vertical circulation (light blue lines) and net biological >processes (green lines)." I'm not sure that the term "budget" is appropriate here. >This figure shows the monthly decomposition of DIC change into processes which is >not a budget. Also, it is not clear when you said, "the two major processes >dominate". They do not dominate; they show the largest seasonal variations but >they not dominate the final seasonal signal.

We understand that in some contexts a budget consists of accounting for all the reservoirs of a given variable to estimate its totality (e.g. finding total organic carbon in a pond by adding up all phytoplankton, macrophyte, zooplankton, fish biomass, etc.). Since the model already has total DIC, we instead tally up all the monthly fluxes that add and remove DIC in a given volume over the course of a year, such that they balance out to create a closed cycle and DIC "budget."

However, if the precise use of "budget" in this instance is deemed inappropriate, then in the revised MS it can be easily foregone:

~~*Budgets of DIC fluxes in the upper 20m...*~~

---

>Line 330: Fig. 8 complicated to compare one region over another. All panels should >be the same yaxis scale.

We understand that for the sake of comparison the y-axis should be the same across all sub-panels. However, due to the differently sized areas that the regions represent, the y-axis scale for India is about two order of magnitude smaller than for the entire domain. As a compromise, we will express the DIC flux as a volume-specific rate, and include on the top-right of each panel a number to indicate the relative size of each volume to that of the domain.

---

>Line 332: "acts as a source of DIC" when? All time? Which month? Should be more  
>precise.

The following revision will be in the final MS:

*In the entire domain and all sub-regions, and for all months, vertical circulation (advection and mixing) acts as a source of DIC, with the sum of all biological processes acting as a sink...*

---

>Line 341: "escape is mostly smaller than biological flux," not clear. Smaller of what?  
>In term of seasonal amplitude? For a specific month? Again, need to be more precise  
>in the text.

Once more numerical values are introduced into the revised MS, this will be more clear:

*DIC flux due to atmospheric escape, while reaching its maximum magnitude of  $\sim 27\text{TgCmonth}^{-1}$  in June and July for the whole domain (Fig. 8a), only surpasses ~~is mostly smaller than~~ biological flux in May, when  $19\text{TgC}$  is released to the atmosphere compared to  $12.3\text{TgC}$ .*

---

>Line 351-357 Maybe this section should be in the discussion section instead of the  
>result section.

Similar to the last paragraph of Sect. 3.2, these sentences are designed to briefly summarize the main results. We can also move this to the Discussion section in the revised MS.

---

>3.4

>Comment: Again, this section needs more value to support results than only a  
>qualitative description.

As with all the other sections in the Results, more numbers will be added in the revised text.

---

>Comment: I propose to change the order of the different sections. First, what  
>governs the variability of  $f\text{CO}_2$ , then  $p\text{CO}_2$ . So, I suggest to first start with 3.1, then  
>3.4 then 3.2 and 3.3.

As per the suggestion earlier on to re-organize the methods section, we are willing to change the order in the revised MS to begin with air-sea  $\text{CO}_2$  flux before presenting  $p\text{CO}_2$  variability and DIC fluxes.

---

>Line 370: what about the contribution of atmospheric  $p\text{CO}_2$ ?

As stated previously, the variability due to atmospheric  $p\text{CO}_2$  is much less than that seen in the surface ocean. For example, if atmospheric  $p\text{CO}_2$  were to remain constant in this model, the change to  $\Delta p\text{CO}_2$  would be at most  $<5\%$ . However, it is true that the term in Eqn. 8 is  $\Delta p\text{CO}_2$ , not  $p\text{CO}_2$ , so this will be amended in the MS:

*...can be attributed to the contributions of winds,  $\Delta p\text{CO}_2$ , and interacting cross-terms...*

---

>Line 372: Figure 10 "(a) Anomaly of air-sea  $\text{CO}_2$  flux during summer monsoon JJAS  
>(molCm<sup>-2</sup>yr<sup>-1</sup>. Summer flux anomaly contributions due to (b) wind, (c)  $p\text{CO}_2$ , and  
>(d) cross-terms in Eqn. (8)". Could say what means a positive/negative anomaly.  
>Also (molCm<sup>-2</sup>yr<sup>-1</sup>).

Thank you, we will include this information in the figure caption:

*Anomaly of air-sea CO<sub>2</sub> flux during summer monsoon JJAS (molCm<sup>-2</sup>yr<sup>-1</sup>). Summer flux anomaly contributions due to (b) wind, (c) pCO<sub>2</sub>, and (d) cross-terms in Eqn.(8). Anomalies are calculated relative to the five-year average.*

---

>Line 374: "The \_pCO2 contribution to" dpco2 or pco2? Not clear. What about the >atmospheric CO2 contribution?

Yes, this will be corrected to be  $\Delta p\text{CO}_2$  in the revised MS.

---

>Line 378: Figure 11. "Seasonal CO2 flux anomaly (purple) for winter DJFM monsoon >(top-left), springAM(top- right), summer monsoon JJAS (bottomleft), and fall ON >(bottom-right). Contributors to the flux are solubility/winds (k\_,blue), pCO2 (red), >and cross-terms (yellow)." Results are shown for entire domain and sub- regions." >Problem with panel "top left" "top right" in the caption and the figure.

Yes, many apologies again, similar to Figure 3, the organization of this figure was changed to maximize its size in the submitted MS, but the caption missed being updated. It will become:

*Seasonal CO<sub>2</sub> flux anomaly (purple) for winter DJFM monsoon (top-left), spring AM (second from top-right), summer monsoon JJAS (bottom-left third from top), and fall ON (bottom-right). Contributors to the flux are solubility/winds (k<sub>a</sub>,blue), pCO<sub>2</sub> (red), and cross-terms (yellow). Results are shown for entire domain and sub-regions.*

---

>Line 378: Figure 11. It is complicated based on this figure to see which process >controls the seasonal flux variability. This figure represents the budget for each >season. It would be easier to represent the seasonal flux variability into processes >using the same strategy than for the seasonal pCO2 anyalsis with seasonal profiles >as realized for pco2 (Fig 7 and 8).

We understand your concern, we have no objection to recasting Figure 11 to follow Figures 7 and 8 for better presentation of the model output.

---

>Line 378: "The seasonal flux anomalies for all regions throughout the year are >displayed in Fig. 11" should be "aggregated by season" instead of "throughout the >year".

Yes, thank you, this smooths out the prose:

*The seasonal flux anomalies for all regions ~~throughout the year~~ are aggregated by season and displayed in Fig. 11*

---

>Line 393-395: Maybe this section should be in the discussion section instead of the >result section.

Similar to previous sections, we are amenable to shifting these brief summaries to the discussion instead of results.

---

>4

>Comment: Lots of repetition of results in the discussion section. Should be reduced.  
>I also propose to reorganize the discussion section in 1) model vs data, 2) spatial >analysis (FCO2 and pCO2) and processes, 3) seasonal analysis (FCO2 and pCO2) and >processes and include in 3) the point 4.3



We appreciate that, especially with the inclusion of more values in the Results section, that the Discussion can be pared down. Additionally, as with the Methods and Results, the revised MS will be organized as proposed.

---

>Question: according to your results, does the seasonal FC02 and pC02 variabilities and processes diverge between the open ocean and the coastal domain in the AS?

For air-sea CO<sub>2</sub> flux, the open ocean and coastal regions in the AS have the same seasonal signal, with a summertime peak. the pCO<sub>2</sub>, however, diverges between the coastal seas and open ocean. In the open ocean, pCO<sub>2</sub> (and ΔpCO<sub>2</sub>) has a peak in the spring whereas coastal pCO<sub>2</sub> has its seasonal peak in the summer. Hopefully, this result is clear from Figures 7 and 9.

---

>Line 406: what is the order of magnitude of the bias of Valsala? same order as you? >Need quantitative values.

In Valsala and Maksyutov (2010), in their Figure 4 (see their "North Indian" subpanel) the free-run model is (from visual estimation) ~5-15μatm above TK09. By comparison, our model has an overall bias of 2.1μatm with respect to *in situ* data, with spring having the largest seasonal bias of 48.4μatm. We shall add these numbers to the revised MS:

*For instance, in the global data assimilation study of Valsala and Maksyutov (2010), they found an overall positive bias in the North Indian ocean, ~+5 to +15μatm above TK09 compared to -3.1 to +48.4μatm with respect to in situ data. Additionally, that study found ~~with~~ a similar...*

---

>Line 428: "Since upwelling regions are limited in geographic extent near the coast, capturing their high pC02 values can be difficult for other approaches, such as TK09 with its coarse grid". Same comment than before. Why did you use the TK09 climatology instead of others finer spatial resolution climatologies. Also, since the upwelling is close to the coast, what about the coastal climatology of Laruelle et al. (2017)?

As we have mentioned previously, TK09 was included because many previous modeling studies use it, and while it is not part of our model validation, it serves as a useful baseline to compare to those studies. However, since the suggestion to use Laruelle et al., (2017) has repeatedly entered discussion, find below a figure comparing the monthly pdf distributions of the in situ data, the L15 climatology, and the combined L15 and Laruelle climatology published in Landschützer et al. (2020) doi:10.5194/essd-12-2537-2020, and downloaded Jun 2021 from NOAA NCEI.

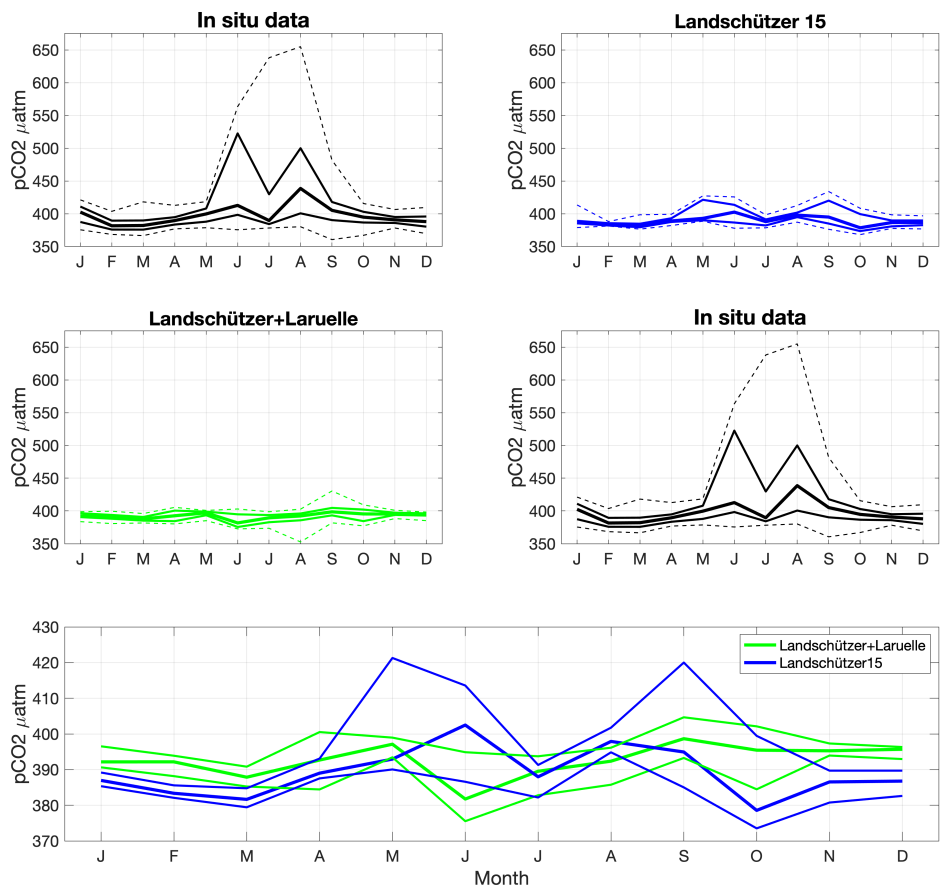


Fig. R1-2. Monthly PDFs showing 5-25-50-75-95 percentile  $p\text{CO}_2$  values of (top-left, mid-right) in situ SOCAT/LDEO data, (top-right) L15 climatology, (mid-left) L15+Laruelle climatology, and (bottom) the 25-50-75 percentile  $p\text{CO}_2$  values of L15 and L15+Laruelle.

With the coastal climatology merged into L15, we can see that outside of summer the  $p\text{CO}_2$  is slightly higher and probably more accurate. During the summer, however,  $p\text{CO}_2$  is underestimated even more so than in L15. As the figure shows, including the new, higher-resolution climatology does not a priori increase skill in the AS. This exercise should further underscore the need for more *in situ* data in the region.

>4.2

>Line 434-445: should be in section 4.1

Fair enough, this section of text does compare the model to previous studies with regard to the bias and higher  $p\text{CO}_2$  values, and would make sense in Sect. 4.1.

>Line 434: "from previous studies" observational or modelled studies? Not clear

The references in the following text contain a mix of both modeling and data-based studies. As a result, we will clarify as follows:

*The distribution of model pCO<sub>2</sub> is both similar to and different from previous data-based and modeling studies.*

---

>Line 446: what are the consequences on this seasonal mismatch on your decomposition into processes?

The consequences of seasonal mismatch vis à vis our decomposition into processes in the AS depends on the magnitude. For instance, if maximum pCO<sub>2</sub> is in June instead of August near Oman, this does not rule out upwelling as a process, but rather the wind forcing used in the model should be investigated to determine its timing. The greatest mismatch in our study is the springtime pCO<sub>2</sub> maximum caused by heating; there is no plausible reason to doubt that heating occurs (and as far as we can tell from the SOCAT data the model is not over-estimating SST), and multiple models presented show this feature. However, the fact that the data do not reflect a springtime maximum requires an explanation that is not readily available. This specific discrepancy, which does reflect an unresolved issue in the region, is explored in the following paragraphs.

---

>Line 462: "influence of outflows from marginal seas" how?

In this sentence, we merely mention that part of the reason for high salinity in the north AS is due to outflow from the Persian/Arabian Gulf. We will add references in the revised MS:

*...with strong tropical heating and the influence of outflows from marginal seas Prasad et al., 2001; L'Hégaret et al., 2015).*

---

>Line 470: "the dominance of temperature is clear" spatially or seasonally?

In this section we are discussion the seasonal pCO<sub>2</sub> anomalies, and so in the revised MS we will stress the seasonal aspect:

*Zooming out from the upwelling regions and looking at the whole AS, the dominance of temperature on the seasonal pCO<sub>2</sub> cycle is clear.*

---

>Line 471: Be careful when you compared processes on different time scale (seasonal vs interannual)

True, just because temperature is important for the annual pCO<sub>2</sub> cycle does not translate into its inter-annual variability being the determining variable in the AS. To avoid further confusion, this sentence will be removed:

*...temperature is clear. ~~This observation, temperature overriding DIC in the AS to determine temporal pCO<sub>2</sub> changes, can also be seen in the interannual variability analysis of Doney et al. (2009). In the domain average, ...~~*

---

>4.3

>Line 497: "and so no biologically-induced decrease of pCO<sub>2</sub> occurs". Maybe add "in the final pCO<sub>2</sub> signal".

In the revised MS, this changed will be made:  
*and so no biologically-induced decrease of pCO<sub>2</sub> occurs textcolorbluein the final pCO<sub>2</sub> signal.*

---

>4.4

>4.4.1

>Line 512: "where uptake of CO<sub>2</sub> takes" should be "atmospheric CO<sub>2</sub> uptake takes"

Thanks, this will be re-written in the revised version:

*...where uptake of CO<sub>2</sub> takes...*

---

>Line 518: Fig. 9d instead of Fig. 10d

Thank you, this will be corrected in the revised MS.

---

>Line 521: "The fact that model CO<sub>2</sub> flux peaks in summer despite a wide-ranging spring peak in pCO<sub>2</sub>" For the entire region? Which region? Not clear

Yes, in this sentence we were referring to the entire region; we will be more precise in the revised MS:

*The fact that model CO<sub>2</sub> flux for the entire domain peaks in summer despite a ~~wide-ranging~~ spring peak in pCO<sub>2</sub> for the domain as a whole, along with the Somalia and oligotrophic regions, is the first sign that...*

---

>Line 535: need more reference for the use of different k-parametrization and wind speed on the flux calculation (e.g., wanninkhof, Ho, Roobaert, ...)

True, there has been much more recent work on improving the air-sea flux calculation, and these studies should be highlighted:

*...the possibility that incomplete representation of winds and the various parameterizations of piston velocity must be considered in addition to pCO<sub>2</sub>, especially in light of recent work in the field (Roobaert et al., 2018; Wanninkhof 2014; Ho et al., 2006).*

---

>Line 521-535: Need more comparison to other studies. e.g, you could compare your results with the study of Roobaert et al. (2019) that decompose the CO<sub>2</sub> flux seasonality into processes in the region.

Yes, thank you for pointing out the recent global studies that decompose CO<sub>2</sub> flux and its constituents in a manner comparable (and in ways superior) to our own study. We will add Roobaert et al. (2019) to our discussion, since judging from their Figure 7  $\Delta p\text{CO}_2$  and cross-terms are more important than wind near the Horn of Africa, and in Figure 8 winds do not factor largely into FCO<sub>2</sub> for 0-10°N or 10-40°N (although, granted, this is a global result so perhaps the AS signal is not clear):

*to the point that wind contribution to the domain's summer anomaly was 90.8% in magnitude relative to the total. We should keep in mind, however, that these results conflict with the analysis of Roobaert et al. (2019). In their global study of coastal waters, while seasonal CO<sub>2</sub> flux variability in the AS is relatively high compared to other regions (their Figure 6), the largest contributions come from  $\Delta p\text{CO}_2$  and cross-terms (their Figure 7), especially near the Horn of Africa. As a result, further work should be conducted to reduce uncertainty in sea surface pCO<sub>2</sub> values to determine whether winds,  $\Delta p\text{CO}_2$ , or cross-terms are significant drivers of air-sea flux. ~~Therefore~~ Additionally, when considering the inconsistencies of models in estimating air-sea CO<sub>2</sub>*

flux (Sarma et al., 2013), *the uncertainties from possibility that incomplete representation...*

As an aside, we would be curious to see where the data come from that allow for estimation of pCO<sub>2</sub> and flux coming from the Persian/Arabian Gulf...

---

>4.4.2

>Question: What is the real purpose of this section 4.4.2? is the objective to calculate  
>an uncertainty on the flux? if so, why only focused on the choice of pco2? why you  
>chose TK09 instead of that of Rodenbeck (2014) which is finer in terms of spatial  
>resolution and more recent. What about the uncertainty associated with the other  
>terms involved in the flux calculation? the wind? its spatial and temporal resolution?  
>the k-parametreization? Etc...

The point of Section 4.4.2 is not to calculate the total uncertainty in flux. Rather, it is to highlight that despite the fact that in our analysis winds appear to control the timing and magnitude of CO<sub>2</sub> flux, there is non-negligible variability in our estimate due to the uncertainty in pCO<sub>2</sub>, and so we show that variability by keeping constant the wind product, wind parameterization, etc. At some level, this is an important point to make considering that as discussed above Roobaert et al. (2019) show that ΔpCO<sub>2</sub> is more important than winds in determining flux!

We understand that different wind products and parametrizations would lead to different flux magnitudes, but that is not the point of our study (plus it would be numerically quite onerous to run a complete sensitivity analysis with the full model). Again, TK09 was chosen for comparison with previous studies in the region, and the L15 climatology is both recent and has a finer resolution.

---

>Question: This study highlights that the effect of the wind governs the seasonal  
>variability of the flux. Why is it important to do, as it is performed here, a finer  
>pCO2 decomposition into processes?

True, at face value if the CO<sub>2</sub> flux was the sole variable of concern, then the importance of winds should obviate the need to enter further detail into the other components. However, since many other studies (including the Roobaert et al.2019 analysis above) stress ΔpCO<sub>2</sub> and surface pCO<sub>2</sub> dynamics in general, it is included here. This is both for the sake of completeness but also so that if future studies contradict ours, then perhaps the weakness in our study (whether it is the wind product or the pCO<sub>2</sub> model output) will be more clear.

---

>Line 540: "Considering the important role of winds" on the seasonality? Spatiality?  
>Not clear

In this sentence the stress is on seasonality:  
*Considering the important seasonal role of winds...*

---

>Line 542: Fig 12: should be "Tg C month<sup>-1</sup>" instead of "TgC"

Yes, formally the unit is a flux, and so it will be changed to TgCmonth<sup>-1</sup>.

---

>Line 546: "Despite differing pCO2 seasonality". In this figure only LA15 is referred.  
>Not the other products. So you can not refer as something common between all  
>pco2 products

Fair enough, the reference to Figure 5 was to show an example of how pCO<sub>2</sub> seasonality is different, without having to provide a figure showing the pCO<sub>2</sub> seasonality of ALL the pCO<sub>2</sub> products. In order to be consistent, the reference to Fig. 5 will be removed in the revised MS.

---

>Line 560: "1) gridded data-based pCO<sub>2</sub> products will under-estimate the upwelling >zone maxima of pCO<sub>2</sub> and CO<sub>2</sub> flux during the summer" how can you conclude that >since you said before that your model produces higher pCO<sub>2</sub> values compared to >other studies. What is the value/study for this conclusion? Not clear

The conclusion that pCO<sub>2</sub> products under-estimate summertime pCO<sub>2</sub> should be clear from Fig. 5; if necessary, a similar plot can be made for all the other pCO<sub>2</sub> products considered in this study in the Supplementary material. The model produces an overall positive bias, true, but the model still has a slight *negative* bias during the *summer*.

---

>5

>Line 573: 160 while in line 550 it is between 57-120. Not clear.

The true model estimate is 160TgCyr<sup>-1</sup>. This discrepancy arises because the model, with its high resolution, has a greater total surface area, whereas the more coarse products do not. To this end, model output was interpolated to the coarser grid for a more fair comparison. This needs to be clarified in the methods section (end of Sect. 2.1):

*differences between this software suite and Orr and Epitalon (2015) are minimal. Furthermore, for air-sea CO<sub>2</sub> flux intercomparison purposes, all pCO<sub>2</sub> values except for TK09 are interpolated to the same 1°x1° grid already shared by GLODAP, WOA, and L15. This also means model CO<sub>2</sub> flux in the comparison context is reduced due to the model's high spatial resolution.*

---

>Line 574: need references.

In the revised MS we will add these references:

*with the range of other findings (Sarma, 2003; Naqvi et al., 2006; Sarma et al. 2013, ).*

---