Authors' Responses to Reviewers' comments (RC1, RC2 and RC3)

Dear Associate Editor Prof. Dr. Carol Robinson, dear three anonymous Reviewers,

First of all, on behalf of all co-authors, I would like to gratefully acknowledge your time and efforts spent reviewing and considering our manuscript. In general, the Reviewers identified our contribution as valuable, although important concerns were raised. We have taken their criticism seriously and following their recommendations prepared a substantially revised manuscript. In summary, the following main changes were implemented in the revised manuscript version; a) the manuscript language and structure were carefully revised; b) further observations from literature were included to complement our data, discuss how representative of the seasonal changes our profiles may or may not be, and provide a better overall context; c) the highlighted method aspects, calculations and uncertainties were clarified; and d) the supplement was revised to include all data from this study. Thanks to the Reviewers insightful and constructive feedback we were able to prepare an improved manuscript version, which we look forward to submitting for your further consideration for publication in Biogeosciences.

Below, please find below our point-to-point comments to all Reviewers' comments (with our responses are in blue).

Yours sincerely,

Hana Jurikova

Responses to Reviewer 1 (bg-2020-448-RC1-supplement)

This manuscript present results of O2/Ar ratios and 17D (17-O excess) in water samples collected in profiles on the SEATS (SouthEast Asian Time-series Study) station located in the South China Sea. This method has been widely used in the ocean to determine aquatic primary productivity. The net oxygen production (NOP) is determined by discriminating the physical contribution to the dissolved oxygen concentration through measurements of the O2/Ar ratio, and determine solely the biological input. The gross oxygen production (GOP) is determined by the quantification of the stable isotopic abundance in dissolved oxygen (16O, 17O and 18O) and on the basis of the mass fractionation identify the presence of biotic (from photosynthesis) or abiotic sources (from the exchange with the atmosphere) of O2 in the sample. The authors aimed to study the influence of seasonal monsoon forcing to the local aquatic primary productivity and to better identify oligotrophic and eutrophic stages.
For this, vertical profiles were collected to represent three seasonal phases: 1) during the transition of the monsoon seasons (16th October, 2013), 2) during the southwest monsoon (SWM) in summer (June to September; samples collected in 5th and 6th August, 2014), and 3) during the northeast monsoon (NEM) in winter (November to April; samples collected in 24th and 25th April, 2015). Complementary measurements of dissolved oxygen and fluorescence were collected to support the observations from the isotopic analysis. The authors found that winter conditions allow for a shift from net heterotrophy to net autotrophy in the course of 24 h due to the influences of colder temperature and stronger winds. Although the authors did not quantify specifically vertical exchanges between the mixed layer and deeper waters, their vertical profiles hint to the exchange of productive deeper

waters with less productive mixed layer water from summer. However, the highest productivity estimates were found during the inter-monsoon sampling in October 2013. The NEM might play an important role to control the metabolic balance at the SEATS in the South China Sea by shifting to net autotrophy mode, in contrast to the predominant net heterotrophic state in summer.

Although the O2/Ar and TOI methods have been widely applied in ocean research, few studies in the past have focused on vertical sampling as the authors pointed out well (e.g. Juranek and Quay, 2005; Wurgaft et al., 2013) in the water column, despite several discussions have focused on the relevance of vertical influences to the mixed layer productivity determination. In this regard, the contribution of the vertical profiles from this study, and the focus on analyzing the changes in the different seasons, made a nice short manuscript that certainly falls within the scope of BG. My major concern is that the authors draw conclusions to characterize seasons (or annual trend) from scarce time data points (one profile per day in max. two days in summer or winter), and other previous observations should be used to better understand the SEATS station in the context of aquatic productivity. In addition, the way that the manuscript is currently written does not meet yet the quality necessary to accept it for publication, and major changes are needed. Besides it requires a thorough language editorial review, it is oddly arranged with paragraphs that are included in the discussion but that actually belong to results, and vice versa. To be able to consider this manuscript for publication in BG, I encourage the authors to improve the current version substantially. I list below general and specific comments with the hope that these can support the authors to improve their current version.

Reply: We thank the Reviewer for recognizing our contribution as valuable, the constructive feedback and the numerous suggestions for improvement. In brief, following the Reviewer's suggestions, in the revised manuscript version we have included further observations and data from SEATS station to provide a better context for our results and the understanding of aquatic productivity in this region. Likewise, we have carefully revised the written quality, language and structure of the manuscript to meet the standard for publication in BG and make it more accessible to the community.

General comments.

- The text requires a thorough English editorial check. In many places, the articles are missing to refer to e.g. "THE" SCS, or "THE" limit of the photic zone. I encourage the authors to let the revised version be read by a professional English editor or a native speaker to correct this recurrent grammar errors. Also, some sentences are hard to understand or the fluency of the text is hard to follow. This can also be alleviated by the contribution of a professional editor.

Reply: The revised manuscript text has been thoroughly revised and edited.

- The authors sampled in October 2013 (intermediate season), August 2014 (summer) and April 2015 (winter), and discuss these as a sequential annual set up, but do the authors think they would have found the same pattern if e.g. August 2013 and April 2014, or simply October 2015, were instead sampled instead? How the results could have changed if a true sequential year was sampled? Are there any anomalous events in between the presented sampling dates that could indicate that the results are not representative of true sequential seasons in a "normal" year? For example, the 2014-2016 ENSO influence?

Reply: This is an important point that has also been raised by Reviewer 2 and Reviewer 3. In the revised manuscript we have included further information and observations from the literature that provide a better context for our data. It is known that ENSO plays an important role in the regional climate, however, its exact role and impact on marine biogeochemistry is yet to be fully understood. Indeed, one of the main aims of the SEATS program is to understand how ENSO modulation of the monsoon strength influences the biogeochemical cycles in the SCS. Such assessment, however, requires more years of data, in particular during the ~2016 ENSO year. We have made routine (roughly once a year) cruises to the SCS for water analysis since the project started in 2013, but we are yet to analyse these and finalise the samples collected. We hope the analysis can be done soon so that we can report the results to the community. Current available data (not included in this paper) seem to show a difference between peak El Niño and off-El Niño years, but whether the difference is significant requires further analysis, which is yet to be done. Besides, as mentioned, any ENSO's influences, at a first order, are likely going to be transmitted through the monsoon forcing (El Niño years -> warm Pacific -> late arrival and weaker monsoon in the SCS; e.g. Zhou & Chan, 2006, Int. J. Climatol. 27: 157-167, https://doi.org/10.1002/joc.1380). Hence monsoon can be considered the central climatic phenomenon and is also the reason for the focus of our study. In terms of short-term variabilities, apart from the timing of the monsoon arrival (hence Fig. 2), typhoons also play an important role in the regional biogeochemistry.

- Aren't there any historical records in the SEATS station that could support the observations presented in the current study to better characterize the seasonal changes? For example, nutrients distribution, record of dissolved O2 profiles? The authors made references to other studies, so they could make better use of the past data to interpret their current results.

Reply: We have expanded on the mentioned point in the discussion section and added further relevant data from the literature to support out observations.

- The authors discuss their results in the context of steady state mixed layer, but clearly one of the major contributors to their observations seems to be the diapycnal mixing and contribution from deeper water to the mixed layer primary productivity. The authors could estimate this mixing as it has been done previously (e.g. Castro- Morales et al., 2013; Seguro et al., 2019) on the basis of their vertical profiling data and quantify the vertical contributions.

Reply: Please see also our replies to Reviewer 2 below. We have added a statement that acknowledges that this has not been considered in our calculations. However, at least during the observation period, we did not find any significant upwelling of water (internal wave) at SEATS (although unfortunately our dataset is limited to properly explore this over the O_2 residence time), hence we do not think this component of mixing from the deep played an important role.

Introduction.
The authors need to sustain their arguments with numerical evidence from the listed references. I make some specific cases in the list of comments below.

Reply: We thank the Reviewer for the suggestions and include our replies to the listed different specific cases below.

Methods.

There is no indication as to how Chlorophyll a was measured in the vertical profile. Also, a lot of relevant information such as: how many samples were collected from each depth (i.e. duplicate analysis)? They should also at least mention briefly the sample processing (preservation, extraction of gases).

Reply: We have added information on Chlorophyll a measurement and further details on sample collection. Our sample processing and extraction protocols are detailed in Jurikova et al. (2016) which we have stressed in the revised manuscript.

How was the PAR obtained to define their PML depth? The authors mention repeatedly "daily variation" in their results, but the methods description hint to a one-time sampling of each profile in one sampling date, so how are they able to discuss a time change variation during a day?

Reply: As written in the manuscript, the PAR was obtained from shipboard measurements, and with no to negligible daily variations (see Table 1), but with indeed seasonal variations (see also Fig. 3). We are not quite sure to which specific part the Reviewer hints at as to our knowledge any daily variations whenever discussed point to the variations in the composition of the dissolved oxygen (i.e. $^{17}\Delta$ and the $\Delta(O_2/Ar)$) and not the physical water column properties. The sampling dates are also provided in Section 2.1.

Discussion.
All the paragraphs in lines 211-220, 221-228, and parts of the "October 2013" should belong to results and not to discussion section. Also, it is oddly arranged, since I would start chronologically from October 2013, August 2014 and April 2015.

Reply: We have reorganized the text and moved the relevant paragraphs to the results section as well as changed the order to follow chronologically, as suggested. Just for information, the reason we used the previous presentation order is that we principally focused on the two different monsoon season (august 2014 and April 2015) and due to less information (lack of O_2-Ar data) from October 2013 treated this data somewhat separately.

How can the authors interpret an increasing value of 17D with depth while D(O2/Ar) is decreasing? So apparent net heterotrophy by the consumption of O2 is assumed due to negative D(O2/Ar) (also vertically exchanged O2 depleted deep water), but at the same time, increasing photosynthetic production of O2 is assumed due to increasing 17D with depth? I think this is the representation of the vertical contributions to the ML both in summer and winter.

Reply: In all cases, the PLD, which is also where the Chl a maximum was situated was well below the MLD depth. Therefore, we interpret these values as indicative of biological production, however also high consumption (hence the high $^{17}\Delta$, but the low $\Delta(O_2/Ar)$, respectively). We have elaborated on this in the revised text.

Use of notation.
There seems to be a mix up in the definition of the parameters. The variable 17D is the notation to represents the 17O-excess or the triple oxygen isotopic anomaly of dissolved oxygen, defined with the Eq. 1. However, in some places the authors refer to 17D as the "triple oxygen isotope composition technique" (L62), the "triple isotopic analyses of dissolved O2" (L70), or simply the "triple oxygen isotope composition" (L101). The authors

should be careful on how they make use of the notation and how they refer to this parameter in the text.

Reply: We have modified this in the relevant text parts for consistency, the $^{17}\Delta$, as defined, refers to the triple oxygen isotope anomaly or ^{17}O-excess.

List of detailed comments.
L22 – I wouldn't call 17D a "conservative" tracer since it is influenced by photosynthetic or atmospheric sources.

Reply: Modified. We originally used this term because in the deep $^{17}\Delta$ can no longer be affected by photosynthesis or atmospheric sources. However, and as also pointed out by the Reviewer 2, although not directly, $^{17}\Delta$ in the deep can indeed be affected if by respiration and mixing (if two parcels of water with very different δ-values mix), hence indeed the use of the term conservative was an unfortunate choice from our side and has to be taken and utilized more carefully.

L25 – How is the SCS influencing the global biogeochemical cycling of C and O2? Can the authors sustain this statement with some references? Especially in L29 the authors mention that the properties in marginal seas cannot be extrapolated to global scales.

Reply: We have added further references to the text to support the importance of SCS in global cycling. The two statements, however, by context do not oppose each other, since L29 refers to extrapolation in terms of acting as a sink or source, and regardless whether SCS acts as a sink or source its role would be important (L24).

L29 – which range of latitudes represent "mid-latitude shelves"?

Reply: Mid-latitudes are a commonly used term and defined as from the tropic of Cancer to the Arctic circle and from the Tropic of Capricorn to the Antarctic circle (22 °27' to 66 °33' N and S, respectively).

L33 – this statement lacks of numerical evidence, how much is "high" annual surface temperature?

Reply: The reason for this is because in the context of the text the statement is comparative (rather than aimed at providing absolute values); this sentence introduces tropical and subtropical shelves, following on from the previous sentence where mid-latitude shelves are discussed. Therefore, in comparison to mid-latitude shelves, subtropical and tropical shelves have high annual surface temperature.

L37 – can the authors elaborate what do they mean with "unusual seasonal pattern in phytoplankton biomass"?

Reply: Clarified.

L36 – providing the values of the CO2 fluxes will be helpful.

Reply: We have elaborated on this in the discussion section 4.2 (as also suggested by the Reviewer below in specific comments)

L44 – “Owing to its geographical position ....”

Reply: Added.

L45 – “the weaker southwest monsoon (SWM) in summer”

Reply: We would prefer to keep the original wording.

L46 – “the strong northeast monsoon (NEM) in winter”

Reply: We would prefer to keep the original wording.

L49 – How much are “medium chlorophyll a” concentrations? L66 - which other sources?

Reply: Clarified.

L67 – “enables an estimation of the integrated gross productivity...”

Reply: Corrected.

L69 – “In order to evaluate the photosynthetic O2 production, and its contribution to the local carbon balance...”

Reply: Corrected.

L72 – how are the authors defining here “deep water”? as of below the seasonal mixed layer depth?

Reply: Clarified.

L76 – “on board” instead of “aboard”

Reply: We prefer to keep aboard which is also correct.

L99 – The authors provide here a very specific detail in the pre-processing of samples, but lack of more general and necessary information on the method, so either elaborate and complete this sentence or simply remove the detail.

Reply: We are sorry but not sure what the Reviewer exactly refers to. The sample pre-processing is a routine procedure, and in our lab exactly that of the references cited therefore we consider it appropriate to refer to the published sources. The $^{17}\Delta$ is however something that is not fixed and depends on definition (e.g. can be linear or log) and hence we consider it important to clearly describe the definition used. Moreover, while this paper can be understood without the sample pre-processing routine, it would be hard to follow without providing the $^{17}\Delta$ definition, hence we consider the present description of methods pertinent.

L104 - Equation 1 should have 17D on the left to the equation.

Reply: Thanks for pointing out this oversight from our side.

L124 – what are the mean ± 1 standard deviation values of the 17D in the equilibrated water samples?

Reply: Added.

L129 – Throughout the manuscript the authors refer to gross oxygen production as GP, so please also change it in Equation 2.

Reply: We used GOP when referring to production rates in O_2 units, and GP to the scaled units, we have now clarified this.

L133 - Why mentioning Wanninkhof et al., 2009 if in the end they used Ho et al., 2006?

Reply: We believe it is useful to mention both, as the parametrization from Wanninkhof et al., 2009 is commonly used, but in our particular case with high windspeeds we used the one from Ho et al., 2006.

L135 – how much is "high" wind speeds? Also, these lines are hard to follow. They mention "O2 production time", wouldn't they mean O2 residence time or O2 production rate? By adding the units of the O2 concentration and GP, as well of K it will help the reader to understand this calculation.

Reply: Clarified.

L139 – here needs to refer for the first time to Table 1.

Reply: Added.

L141 – throughout the manuscript the authors refer to net oxygen production as NP, please change here NOP.

Reply: See also the previous reply about GOP/GP - We used NOP when referring to production rates in O_2 units, and NP to the scaled units, we have now clarified this.

L154 – "the standard approach"? I am not sure there is a standard approach in this, this line might be rephrased.

Reply: Rephrased.

L160 – I don't think is useful to start the description of results by contrasting SEATS to HOT. Also, further comparisons to HOT should be done with caution since it is a tropical station with very different seasonal characteristics.

Reply: We have revised this section to make the comparison between SEATS and HOT fit better. Nonetheless we would prefer to keep the comparison as it helps to put our data into context and also allows to contrast results from different time series stations.

L166 – here it should be "August 2014"

Reply: Yes, thanks, sorry about the typo.

L169 – What depths refer to a "shallow" mixed layer?

Reply: We have clarified this; in this case a "shallow mixed layer" is to be understood in the context of the global ocean.

L176-181 – what is the reference value to say that Chl-a was generally low? With respect to what? Can the authors include here some reference values for this statement? Also, there are no uncertainty values added to these measurements? If the authors compare a range of 0.2-0.3 mg m-3 to 0.5 mg m-3, are these values really different based on the uncertainty given in the measurements? Maybe also some statistical analysis will help here to understand if these values are significantly different or not.

Reply: Again, this should be considered in context of the expected ranges for the ocean, which we believe are well defined and understood. We have clarified this and added further details on the uncertainty. Statistical analysis in this case is not appropriate as we only had few data points.

L184 - how can the authors define "a daily component" if only a one-time point sampling was done?

Reply: We are not sure what the reviewer means, the vertical sampling is explained in Section 2.1. For both August 2014 and April 2015 samples were collected on two consecutive days, with some variations in between observed, hence "daily component".

L185-186 – how much is "low" 17D, which depth range is defined for "upper 17D profiles", which "values below"? the "mixed layer values" of what?

Reply: Clarified.

L190-197 – this paragraph is very hard to follow.

Reply: Revised to improve readability.

L199-206 – this paragraph should belong to discussion and still I am not sure that a comparison to HOT brings something substantial to the results of this manuscript.

Reply: Revised, see previous comment regarding comparison to HOT.

L215 – So far, the authors did not provide anywhere the typical 17D values for atmospheric O2 (or at equilibrium) and for photosynthetic O2, this will help the readers to understand the shift in values. Also, referring to Table 1 here will help to the reader and many of these lines can be skipped in results and only focus on the discussion.

Reply: Added.

L254-261 – This paragraph should rather belong to introduction.

Reply: We believe that the current position of the paragraph in the text, at the start of the discussion of the production rates, might be more helpful to the reader and would prefer to keep as it is.

L263 – This paragraph should come earlier in the discussion and contains data that should be moved to results or referred in results to table 1.

Reply: Revised.

L265 – Directing the reader to Table 1 should become a lot earlier in the manuscript in the results section. In this table the authors present production rates in terms of mg C m-2 d-1, whereas in the text the present the results in terms of g C m-2 d-1. Please decide one or the other for both presentations of the data. Table 1 will also benefit if uncertainty values are added in the form of e.g. ± standard deviation if possible.

Reply: Added.

L262-279 – How these results can be compared to the CO2 fluxes reported by Tseng et al., 2005 (and briefly mentioned in the introduction?).

Reply: We have expanded this section of discussion and also elaborated on comparison to Tseng et al.

L268 – I think that to say that two days of sampling with little variation between measurements might be representative of summer is an overstatement.

Reply: Revised.

L275-278 – the authors could estimate the vertical contributions of oxygen on the basis of their vertical profiling data, in this way they can rule out deep local production or vertical transport. Which processes can sustain new production at 30-50 m depth? What role can play processes such as photoinhibition and phytoplankton vertical migration due to enhanced mixing to the observed subsurface Chl-a maximum and deep higher productivity?

Reply: The interpretation of these data from October 2013 is unfortunately more complicated as we lack $\Delta(O_2/Ar)$ data. Considering the depth of the PLD, below the MLD it is likely that these values are a result of biological O_2 contribution to the mixed layer from the thermocline; we have expanded on this in the revised manuscript.

L281 – If the mixed layer depth is defined solely on the basis of temperature gradient, this can lead to shallower layers than in contrast are very different to mixed layers defined on other parameters related to metabolic processes. A mixed layer defined on the basis of density gradients or dissolved oxygen (e.g. Castro-Morales and Kaiser, 2012), might support better the mixed layer productivity observations, and will be perhaps be more similar to their photic layer depth. The authors might consider changing their ML definition as test.

Reply: As described in Section 2.1 "Mixed layer was defined by 1 °C (ΔT) threshold from the temperature at 10 m depth, and further verified by visual inspection of vertical temperature, density and dissolved oxygen profiles.", our MLD was not solely defined on temperature but verified also on the basis of density gradients and dissolved oxygen.

L299-300 – Besides that the comparison between 14C and the O2 isotope is limited due to their methodological nature, the authors hint to the relevance of the seasonal shifts in the region to the local productivity, hence they should look at the sampling periods when the published 14C incubations were made before considering a general decline in productivity rates.

Reply: We have expanded on this as suggested.

L316 - something seems to be missing in the finishing sentence.

Reply: Revised.

L320 – I think this section in the discussion must be the center of the discussion since is the most important contribution in the manuscript (vertical 17D quantification), hence this should be the beginning of their discussion section.

Reply: We are glad the Reviewer acknowledged the values of our deep data and thank for the suggestion. We would however prefer to keep this section in the current place; it is our view that the current organization is more appropriate as the description of $^{17}\Delta$ is directly relevant to primary production, and the deeper part is more complimentary. We will however consider making the $^{17}\Delta$ section more visible.

L330-338 - This paragraph is oddly written with some missing or wrongly placed words. It needs to be improved, but also the placing within the manuscript should be considered. It provides a lot of details on the physics in the SCS, and it should be moved to a site description section at best. Throughout the manuscript, the authors did not present any data that intended to identify the water masses at SEATS in relation to the productivity data, so trying to interpret their results in this context is a stretch towards the end of the manuscript.

Reply: We have improved the readability of this paragraph. As explained in the previous reply we would prefer to keep the different discussion sections in the current order. Also we are not exactly sure what the Reviewer refers to in the last sentences; this section discusses the $^{17}\Delta$ in relation to deep water mixing, not productivity, as at this depth any changes in $^{17}\Delta$ can only results from mixing of waters with different $^{17}\Delta$.

L340-354 – Following the above comment, this paragraph should be moved to an earlier part of the manuscript if the authors think it contributes substantially to their productivity observations. Further, the authors suggest that 17D might be a useful tracer of water mixing processes, hence the interpretation of their results in this regard should be made prominent in the manuscript. However, I think the data set is still too limited for this purpose.

Reply: As explained the aim of this dataset is to trace potential mixing processes rather than productivity, hence we would prefer to keep it separate. Also, because the dataset is limited, we prefer the current placing towards the end of the manuscript rather than making it more prominent.

L359 – Was not the highest GP recorded during the inter-monsoon period?

Reply: Indeed, however, in this sentence we refer to the net production, which was highest during the winter monsoon season.

L360-362 – I think the authors cannot yet conclude this statement with the few data points they present, especially if no other data from traditional conservative tracers for this purpose is discussed in this manuscript.

Reply: We have toned down this statement, also please note that the following sentences also highlight that further research in this direction is necessary.

Responses to Reviewer 2 (bg-2020-448-RC2-supplement)

Summary.
Jurikova et al provide measurements from the SEATS station in the South China Sea of O2 triple isotopes (Oct 2013, Aug 2014 and Apr 2015) and O2/Ar (Aug and Apr only). They quantify rates of NP and GP and discuss seasonal changes. The topic of the manuscript is relevant to the Biogeosciences audience, as indicated by the companion paper Jurikova et al. (2016). I am always interested to read papers with triple oxygen isotope data from a new region and to see new research groups conducting these very challenging and valuable measurements.

Summary of major comments.
Below I summarize the major issues that I believe should be addressed before acceptance. More details on each point follows on subsequent pages.

1) Although the article focuses on the calculated GP and NP rates, there uncertainties in these rates are not quantified. Comprehensive error analysis of the calculated rates is required. The manuscript should acknowledge that they cannot correct for the effects of complex physical processes (vertical mixing, lateral advection) and non-steady state on their mass balance. The authors state that the system was net heterotrophic in Aug 2014 (L16); in fact, the NP rate was effectively 0 considering the uncertainty. I think the article would be more useful to the broader community and future investigators if they also described how future studies could constrain the largest uncertainties in the NP and GP estimates.

Reply: We thank for the insightful comments, and as suggested now included an error analysis and propagation, and explicitly acknowledge that the method does not correct for the effects of complex physical processes. We have also added text that describes how the large uncertainties could be better constrained in the future.

2) The authors describe the use of "17Δ of deep O2 as a valuable novel conservative tracer for probing mixing processes" (L22) but simple calculations and previously published papers have shown that 17Δ is nonconservative in the subsurface when the effects of mixing and respiration considered together (Nicholson et al., 2014).

Reply: This is correct. As already stated in our reply to Reviewer 1, we originally used this term because in the deep $^{17}\Delta$ can no longer be affected by photosynthesis or atmospheric sources. However, as pointed out by the Reviewer and shown in Nicholson et al. 2014, although not directly, $^{17}\Delta$ in the deep can indeed be affected if by respiration and mixing (if two parcels of water with very different δ-values mix), hence the use of the term conservative was an unfortunate choice from our side, and we have now modified this.

3) The authors report 17Δ values for the thermocline that are "much higher than any previously reported values" and very high values in the deep ocean at low O2 levels. I am concerned these results may be an analytical artifact related to size/pressure effects (nonlinearities in the mass spectrometer response when the sample and standard contain different amounts of gas). The authors should describe the calibration procedures and directly address the possible uncertainties.

Reply: In the current mass spectrometer setting, the below pressure adjustment is made every acquisition (i.e. about ~20 min analysis time). With this setting, no observable bleeding effect was seen (i.e. less than the nominal precision of small delta and cap delta reported in the text). Furthermore, in the deepest samples (3500 m) the O_2 concentration was still relatively high (~100 umol/kg for the deepest samples vs. ~190 umol/kg at the surface) and hence sufficient for analyses without sacrificed precision or introduction of analytical artifacts. We thank the Reviewer for pointing this out and have now included the details in the revised manuscript text.

4) The authors are not in compliance with the journal's data policy. More details are required on the methods for calculating GP and NP so that they can be reproduced by others. The mixed layer depth calculations may need to be modified.

Reply: We have revised the manuscript as suggested to comply with the journal data policy.

Detailed comments.
First, a comprehensive error analysis of the dataset is required. In my opinion, the claim in the abstract of the system being net heterotrophic in Aug is not supported by the analysis. Looking at Table 1 and the supplementary data, the O2/Ar disequilibrium at the surface in Aug 2014 was ≤0.5% (the stated analytical uncertainty, L120-124), so they cannot conclude whether the system is net autotrophic or heterotrophic.

Reply: We have now included the error analysis and modified the text accordingly.

Beyond the analytical uncertainty, they have not acknowledged many other larger sources of uncertainty, such as the gas transfer velocity (e.g., the parameterization choice, and the method of time integration), the conversion from O2 to C units, mixed layer depth, and potential impact of physical processes (e.g., vertical mixing and lateral advection), and non-steady state dynamics on their calculated rates.

Reply: Acknowledged.

Indeed, corrections for physical processes and non-steady state would be very difficult to estimate given these given they only have profiles at one location, and only on one or two consecutive dates in each season. However, the 20-30 per meg difference in 17Δ on consecutive days in April and August at 5-10 m depth, as well as the broad range of literature on circulation in this region demonstrates the system is very dynamic and that a mixed layer NP and GP estimates based on the current dataset will have significant uncertainties which should be acknowledged.

Reply: Indeed, unfortunately physical processes and non-steady state is very hard to estimate at the moment (though we hope to be able to elaborate on this in future work), and we now acknowledge this in the text.

On L150 they state "Assuming the mixed layer is at steady state" – I think it would be appropriate to add "and there is not significant entrainment/upwelling of low-O2 subsurface water into the mixed layer, nor lateral advection from adjacent waters."

Reply: Added.

On L280 they state "Furthermore, we note that the calculated production rates should be considered as minimum production only. At SEATS the euphotic zone was persistently deeper than the mixed layer during our sampling (Fig. 3), which may lead to an underestimation of the rates." A similar statement is in the caption of Table 1. I think this would be better framed by stating that you are calculating mixed layer production rates. Mixed layer GP will indeed be less than total water column GP if production occurs below the mixed layer. However, their calculated mixed layer productivity may underestimate the true mixed layer NP values due to mixing/entrainment of low-O_2 waters into the mixed layer, and overestimate the true mixed layer GP values due to mixing/entrainment of high-$^{17}\Delta$ waters into the mixed layer.

Reply: We have revised the text as suggested.

I think the article will be of more use to future investigators if they describe how future studies could be designed to better constrain the largest uncertainties in the NP and GP estimates.

Reply: We thank for pointing this out and have added this information to the revised manuscript text.

Second, the authors incorrectly claim several times that $^{17}\Delta$ in the subsurface is conservative and only affected by mixing (e.g. L22, 72, 322, 363). For example, L322:
"Due to the conservative behaviour of O_2 in a parcel of deep water, where it may no longer be influenced by air–sea gas exchange or photosynthesis, the $^{17}\Delta$ could also present a valuable tracer for deep water mixing processes, since any variations in $^{17}\Delta$ can only result from mixing of waters with different $^{17}\Delta$."

This statement is unfortunately not true. Please see Nicholson et al. (2014), Figure 6 and section 4.2 which states "while respiration alone does not alter $^{17}\Delta_{dis}$, the tracer is nonconservative when the effects of respiration and mixing are combined".

Reply: See also our previous reply, we have modified this now.

The authors' discussion of $^{17}\Delta$ as a subsurface tracer should be substantially revised. The authors could perform simple calculations to see whether their observed $^{17}\Delta$ values can be explained by mixing and respiration. If the observations can be explained by simple modeling, it would give the reader more confidence that their unprecedented subsurface values are accurate (see below).

Reply: We have revised this section.

Third, the authors report $^{17}\Delta$ values in the subsurface that are "much higher than
17 any previously documented upper ocean values" (e.g. Δ of 218 per meg and
$\Delta O_2/Ar$ of -48% at 100 m depth on 5 Aug 2014). They also report very high values
17 in the deep ocean (e.g. Δ of 215 per meg and $\Delta O_2/Ar$ of -71% on 6 Aug 2014).
Modeling results (Nicholson et al., 2014) and the observations of Hendricks (2005) predict $^{17}\Delta$ in low-O_2 waters can actually reach or approach negative values due to the combined effects of mixing and respiration.

Based on my own experience measuring 17Δ, I am concerned these highly unusual 17Δ values at O2 undersaturation are an analytical artifact related to variability in the sample size (number of moles of O2 in the flask). The methods are not described in this paper but the authors reference Jurikova et al. (2016) where it appears that they applied a correction for the O2/Ar ratio in the sample (Ar correction), but not the total amount of gas in the sample (size correction).

Reply: Please see also our above reply on this issue. First of all, we apologise for any misunderstandings that may have arisen due to missing details on this in the manuscript text, which have now been added. Indeed, as in Jurikova et al., 2016 we performed a correction for the O_2/Ar ratio in the sample. We have however not performed a correction for the total amount of gas in the sample, as this was not necessary with the current mass spectrometer setting, and the amount of O_2 available in the sample.

In their response to the reviewers of the 2016 paper, the authors stated:
"We did not perform any corrections due to differential gas depletion between the bellows, as we also did not observe any fractionation (within the current precision)."

I am not sure I understand this statement. The way that researchers typically perform a sample size calibration is by putting the reference standard in both bellows but varying the total amount (volume) of gas in each bellow by adjusting the bellows compression and plotting a calibration curve that spans the full range of sizes in their sample set. The size correction is discussed in the appendix of Stanley et al. (2010). Note that this size correction also simultaneously corrects for the "pressure baseline" effects described by Yeung et al. (2018). This correction should be performed periodically as it varies with time.

For the system I have used, the size calibration was the largest correction (much bigger than the Ar calibration), whenever the size difference between sample and standard was greater than roughly 10% (i.e., most samples from below the mixed layer). However, this correction will vary with time and between instruments, and it is possible for it to be negligible, even at low O2 saturation.

Reply: Indeed, we are aware of this correction, however, this is something we did not observe with our current mass spectrometer settings. By varying the sample size from >30 umol to ~10 umol, we did not observe any significant difference in $\delta^{18}O$, nor $^{17}\Delta$ values, within our reported precision (the maximum $^{17}\Delta$ difference found was 10 per meg). Considering that the O_2 concentration at depth was around ~100 umol/kg based on the CTD data, and at least 150 ml of water was collected for the extraction of dissolved O_2 then our sample sizes fall well within this range.

The bellow pressure adjustment was made every acquisition (i.e. about ~20 min analysis time) and remained well balanced throughout. All analyses were performed at 6000 mV, and by the end of the analyses the imbalance between the sample and the reference was maximum some 200 mV, and hence our precision remained unaffected.

Our statement from the Response letter on the 2016 paper therefore intends to convey this message. We have clarified this in the manuscript text.

The authors need to state in the manuscript whether these size correction tests were done across the full range of sample sizes measured in this dataset, both for the Ar-free and with-Ar methods. If the size corrections were performed with both methods and the offset was negligible, that will give the reader confidence in the data. If they were not performed, then this potential uncertainty should be clearly communicated to the readership base who may not have experience with these very challenging measurements.

Reply: See also our above reply. We have added this information to the manuscript text and thank the Reviewer for pointing this out.

It is also worth noting that because the calculated NP and GP values are based on mixed layer data only (they are not corrected for mixing/entrainment of the subsurface waters), any errors due to a lack of size correction may not significantly affect their NP and GP calculations, if the size of sample and reference was similar for all of the mixed layer samples.

Reply: We thank the Reviewer for highlighting this point, indeed any potential uncertainties arising from sample size correction would only affect the samples from below the mixed layer. However, as previously described this is something we have tested, and we are confident that our deep samples were not affected by measurement artifacts.

On L95 they mention that for the October data (Ar-free method) some samples were run at Nagoya University – did the Nagoya laboratory perform a size calibration? What were the result of the inter-lab comparison? (e.g., what was the mean absolute difference between duplicates run in the two labs)? In general, how many samples were collected at each depth (error bars are not reported on the figures or supplemental data)? What is the precision of replicate water field samples (not lab-equilibrated waters)?

Reply: We have added this information to the revised manuscript text. The $^{17}\Delta$ determination at Nagoya University was indeed conducted as O_2 gas, not O_2-Ar mixture, in this case also the size calibration was not required (note also that the Stanley et al., 2010 appendix refers to a calibration for an O_2-Ar mixture).

On L120, they report the reproducibility of the equilibrated water samples. What was the mean value of 17Δ in equilibrated water?

Reply: This information was reported in Jurikova et al. (2016), but we have now also included it to the revised manuscript text.

Finally, the authors are not in compliance with the journal's data policy (https://www.biogeosciences.net/policies/data_policy.html) and have not provided sufficient details on the GP and NP calculation methods. Stating that the CTD data is available by contacting the authors is not acceptable. All data needed to reproduce the GP and NP calculations should be published to a repository capable of issuing a DOI such as PANGAEA or Zenodo.

Authors should archive the following:
1) Date and time of each sample (since they discuss in the paper that time of day is important; note the date of the final profile is incorrectly listed as April

25th, 2016; the year should be 2015)
2) latitude and longitude
3) all CTD data used in the paper: exact depth/pressure, temperature, salinity, O2, chlorophyll fluorescence for each sample discrete 17Δ sample, as well as the high-resolution profiles plotted in Fig 4.
4) optional: mixed layer depth, wind speed, and k values used to calculate GP/NP and the final GP and NP results, and a description of how each was determined.
5) Define in the metadata how 17Δ is calculated (what lambda is used)

Reply: We have now included the request data to the supplement.

The authors have not stated how the mixed layer depth (MLD) is defined. I believe that the calculated NP on 24 Apr is too low due to an incorrect definition of the MLD. I suggest that they should use an O2-based mixed layer depth as described in Castro-Morales and Kaiser (2012). This would prevent the authors from including undersaturated O2 samples in the NP calculation. In April 2015, it appears the authors used the same MLD of 23 m on both sampling dates (Table 1), even though the O2 profiles on the 24 and 25 Apr are significantly different. On 24 April, ΔO2/Ar is 1.6% at 5 m, 1.3% at 10 m, and -4.5% at 20 m. Including the 20 m depth sample in the mixed layer leads to an underestimate of NP. They report the mixed layer ΔO2/Ar = -2% and NP -160 mg C m-2 d-1 in Table 1, when mixed layer NP should be positive based on the samples at 5 and 10 m depth. On 25 April, ΔO2/Ar is ~2% (range 1.4 to 2.2%) from 0 to 30 m depth so a MLD of 23 m is reasonable. If they used a shallower MLD on 24 Apr, based on the O2 profile, then the calculated NP rates on 24 and 25 Apr would be very similar.

On L137, please clarify how "mixed-layer O2 production time" is calculated and how it differs from mixed-layer O2 residence time. I have never heard this term before.

Reply: Clarified. The mixed layer production time (= O_2 in the water column/ O_2 gross production rate) is included to reflect the rate at which O_2 is produced biologically against the residence time which included physical processes (see also lines 244-246 in the original manuscript version for interpretation).

Additionally, I did not understand their method of calculating k (gas transfer velocity) and request they provide more details on this. On L133 they state "K was derived from wind speeds measured on the ship using an anemometer and verified against NCEP data" but then on L135 say that K was "averaged over the O2 residence time in the mixed-layer preceding sampling (16, 7, and 4 days for October 2013, August 2014 and April 2015, respectively). How did they get 16 days worth of data if using the sonic anemometer?

Reply: This is indeed a misunderstanding from our side for which we apologise and have now clarified in the text. We used satellite data and where possible verified it against the data available from the sonic anemometer which were too limited to be used alone (as noted by the Reviewer).

Please note that the widely-accepted method of calculating k for O2 mass balances involves a weighted approach accounting for the fraction of the mixed layer ventilated each day rather than a simple average over the residence time of the mixed layer (Reuer et al., 2007; Teeter et al., 2018).

Reply: Indeed, we have however not originally considered this approach as the O_2 residence time in the MLD was very short and we also lack the information on the MLD history. We also thank for the suggestion of the Teeter et al., 2018 approach of which we were not aware and which introduces a modification for shorter time periods that could be helpful in our case. We will consider making assumptions of the MLD history and also include the weighted-approach for calculating k and the impact of the different k on the final production rates.

In the data supplement, one of the cruise dates is incorrectly listed as April 25th, 2016; the year should be 2015.

Reply: Thanks, corrected.

In Figure 4 and elsewhere in the text please specify the time of day each O2 and triple oxygen isotope profile was collected (which of the fluorescence profiles does it correspond to)? Could differences in the time of day (diel variability) likely to cause some of the differences between consecutive days that they observe?

Reply: Revised.

L 29: replace "unrepresentative" with "highly uncertain" or similar.

Reply: Replaced.

Final points.
I hope that the authors find my feedback helpful and constructive. I realize that I have provided many recommendations to improve the manuscript, and that making these changes will not be a simple case of minor revisions that authors often hope to receive. I am suggesting these revisions to ensure that the published manuscript is scientifically accurate and to maximize future reuse of the manuscript and dataset by other researchers, which should hopefully lead to more citations for the authors.

Reply: We are grateful for all the thoughtful and constructive feedback received and the Reviewer's time and energy spent on this review. The recommendations were extremely helpful and have enabled us to prepare a much-improved manuscript version which we hope will now be more accessible and useful to the community.

References.
Castro-Morales, K., & Kaiser, J. (2012). Using dissolved oxygen concentrations to determine mixed layer depths in the Bellingshausen Sea. Ocean Science, 8(1), 1–10. https://doi.org/10.5194/os-8-1-2012

Hendricks, M. B., Bender, M. L., Barnett, B. A., Strutton, P., & Chavez, F. P. (2005). Triple oxygen isotope composition of dissolved O2 in the equatorial Pacific: A tracer of mixing, production, and respiration. Journal of Geophysical Research: Oceans, 110(12), 1–17. https://doi.org/10.1029/2004JC002735

Jurikova, H., Guha, T., Abe, O., Shiah, F. K., Wang, C. H., & Liang, M. C. (2016). Variations in triple isotope composition of dissolved oxygen and primary production in a subtropical reservoir. Biogeosciences, 13(24), 6683–6698. https://doi.org/10.5194/bg-13-6683-2016

Nicholson, D., Stanley, R. H. R., & Doney, S. C. (2014). The triple oxygen isotope tracer of primary productivity in a dynamic ocean model: Triple oxygen isotopes in a global model. Global Biogeochemical Cycles, 28(5), 538–552. https://doi.org/10.1002/2013GB004704

Reuer, M., Barnett, B., Bender, M., Falkowski, P., & Hendricks, M. (2007). New estimates of Southern Ocean biological production rates from O2/Ar ratios and the triple isotope composition of O2. Deep Sea Research Part I: Oceanographic Research, 54, 951–974. https://doi.org/10.1016/j.dsr.2007.02.007

Stanley, R., Kirkpatrick, J., & Cassar, N. (2010). Net community production and gross primary production rates in the western equatorial Pacific. Global Biogeochemical Cycles, https://doi.org/10.1029/2009GB003651

Teeter, L., Hamme, R. C., Ianson, D., & Bianucci, L. (2018). Accurate Estimation of Net Community Production From O2/Ar Measurements. Global Biogeochemical Cycles, 32(8), 1163–1181. https://doi.org/10.1029/2017GB005874

Yeung, L. Y., Hayles, J. A., Hu, H., Ash, J. L., & Sun, T. (2018). Scale distortion from pressure baselines as a source of inaccuracy in triple-isotope measurements. Rapid Communications in Mass Spectrometry, 32(20), 1811– 1821. https://doi.org/10.1002/rcm.8247

Responses to Reviewer 3 (bg-2020-448-RC3)

Summary.
In this manuscript, the authors present triple oxygen isotope data from the SouthEast Asian Time-series Study (SEATS, 18Ë□N, 116Ë□E) in the South China Sea. The samples were taken during three different years and three different states of the monsoon in the area. Based on this data and O2/Ar measurements, they determine gross primary production and net community production rates.

General Comments.
There are a number of severe points that need to be addressed before a possible publication. Most of these are covered by the two already submitted, detailed referee comments. Generally, this study lacks detailed description about how the sampling and analysis was conducted (e.g. the missing Chl a sampling in the methods section) and a proper error analysis.

Reply: We thank the Reviewer for reviewing our manuscript and the feedback. Please see also our responses to Reviewer 1 and 2 on the specific issues. As also highlighted by the other Reviewers and in our replies to their comments, this missing information on the different aspects of the methods section as well as a proper error analysis have now been added.

The other main critique point, that I want to stress, is that this study is lacking a discussion about whether or not the sampled profiles are representative for the region and the season. This study is based on triple oxygen isotope measurements from three different seasons and three different years: 1 cast in October 2013 (11 samples from 5 to 500m), 2 casts on consecutive days from August 2014 (24 samples between 5 and 3500m) and 2 casts on consecutive days in April 2015 (21 samples between 5 and 3500m). O2/Ar data is only available for the stations in 2014 and 2015.

This is a very small amount of data to base such a study on. The authors need to give a detailed discussion about why they think that the presented profiles are representative for the area and the season. This could for example be done by comparing the measured profiles of temperature, chlorophyll and dissolved oxygen with previous observations in the area.

Reply: This has also been highlighted by previous Reviewers and we have carefully revised it to include the requested discussion and further data and observations from literature.

Minor comment.
Figure 4 needs to be changed. Especially for the shown oxygen data, the chosen diagram type and colorbar is not suited. It would also be interesting to see oxygen saturation data.

Reply: Changed as suggested.