Dear Referee 1,

Thank you for providing your review, which we considered very helpful to strengthen the presentation of our study. Most of your comments request an extension of the manuscript with additional information, which we are happy to implement. In particular, we aim to add a dedicated discussion section addressing the biogeochemical interpretation of our NCP estimates. Several points you raised were also in agreement with RC2, and we cross reference our answers where this applies.

Please find our detailed answers (bold font) and proposed text edits (bold italic font) next to your comments (normal font) below. Line numbers refer to the initially submitted version of the manuscript.

We hope to have addressed all of your comments appropriately, but welcome additional feedback if required.

Best wishes
Jens Daniel Müller, on behalf of all co-authors

***

General major comments (no particular order):

Traditionally NCP is constrained to the depth of the euphotic zone, compensation depth or MLD. Differences between these and their implications for the export of organic matter should be explained and discussed, even though you chose to use TPD in your approach.

We will introduce traditional NCP concepts by inserting following text in L30 of the introduction:

“... we define NCP as the net amount of carbon fixed in organic matter (gross production minus respiration) that is produced in a defined water volume over a defined period. This definition implies that the choice of an integration depth is a
critical component of any NCP estimate. Traditionally, NCP is constrained either to the depth of the euphotic zone, the compensation depth at which gross production equals respiration, or the mixed layer depth (Sarmiento and Gruber, 2006). Of those approaches, only the integration to the compensation depth is directly linked to the vertical distribution of carbon fixation and remineralisation and therefore quantifies the amount of formed organic matter that can potentially be exported.”

To align the terminology used to describe our methods with the description of traditional concepts, we will clarify in L180 that our best-guess is constrained to the compensation depth:

“... we derive the column inventory of incremental changes of $\Delta C_T^*$ (i$\Delta C_T^*$) between two cruise events through vertical integration of $\Delta C_T^*$ from the sea surface to the compensation depth (cd), i.e. the depth (z) at which no net drawdown of CO$_2$ was observed”

This clarification in the Methods will further be supported by an equation summarizing our NCP calculation, as proposed by R2.

We will further clarify in L214 that only for our NCP reconstructions, we replace the traditional compensation depth with a MLD/TPD constraint:

“In the lack of vertical $C_T$* observations that would allow us to determine the compensation depth, we tested two alternative approximations of the integration depth, which are: ...”

Finally, we will add a dedicated section in L399 of the discussion to describe the biogeochemical relevance and interpretation of the chosen NCP constraint with respect to organic matter export:

“Our best–guess of cumulative NCP on July 24 (~1.2 mol m$^{-2}$) represents the net amount of organic matter that was produced throughout the bloom event in the surface waters above the compensation depth at 12 m. After subtracting ~20 % dissolved organic carbon (DOC) production, our NCP estimate equals the produced particulate organic carbon (POC) that is potentially available for export. In contrast, NCP estimates derived from other traditional methods for the integration across depth (such as the lower bound of the euphotic zone or the mixed layer depth) would not directly relate to the POC export potential.”

This additional information was also requested by RC2.

Discuss how NCP is related to export production and deoxygenation. Which processes and phenomena do alter/modulate organic matter supply to the deep?

In the new section that will be included in L399 to discuss the biogeochemical relevance and interpretation of our findings, we will address this question as follows:

“However, the potential POC export constraint by our NCP estimate is not equivalent to the supply of organic matter to the deep waters of the Gotland Basin, because POC might be (partly) remineralised before sinking beneath the permanent halocline. Remineralisation of POC that occurs during the bloom event above the compensation depth is – according to our definition of NCP –
already included in our estimate. In contrast, any additional remineralisation of POC that occurs between the compensation depth and the halocline, or above the compensation depth after the end of the bloom event, reduces the organic matter supply to the deep waters and thereby mitigates deoxygenation. Indeed, our profiling measurements indicate a steady accumulation of $C_T^*$ beneath the compensation depth (Fig. 4), likely fueled by the remineralisation of organic matter. However, our measurements do neither allow to constrain the budget of this $C_T^*$ accumulation, nor could we attribute the source of organic matter.”

This additional information was also requested by RC2.

It is not clear how getting NCP for the previous years will aid in better understanding the bloom controls since there have been no simultaneous (to SOOP operation) studies of bloom phenology, POC profiles or nutrient concentrations.

We will clarify this by adding the following information to our conclusions (L437):

“The application of this approach will allow for the detection and attribution of trends in cyanobacteria NCP across decades. In particular the comparison of NCP estimates of bloom events that occurred under different environmental conditions will provide a better understanding of the controlling factors. Factors to be tested include the environmental parameters used to constrain NCP (pCO$_2$, SST, and TPD), but also additional observations of nutrients and phytoplankton composition routinely determined on SOOP Finnmaid and in the framework of the Baltic Sea monitoring program. The recently started initiative to deploy biogeochemical ARGO floats in the Baltic Sea will further aid to link surface NCP estimates and deep water deoxygenation, and thereby constrain biogeochemical budgets in the Baltic Sea.”

This additional information was also requested by RC2.

Which measurements, on which platforms and on which scales will help you ‘disentangling the drivers of NCP’ in the future.

Please refer to the reply above.

Please discuss the early-spring bloom and its contribution to the annual NCP.

Enabling reliable NCP estimates for the mid-summer cyanobacteria bloom is the core aim of this study. A reliable hindcast of the mid-summer NCP will only be possible when the findings of this study are applied to almost two decades of available SOOP pCO$_2$ data. In the absence of this information, the assessment of the importance of the spring relative to the mid-summer bloom is highly uncertain. Nevertheless, we agree that some more information on the spring-bloom and a rough approximation of its contribution to the annual NCP should be given. Accordingly, we will extend section 1.3 of the introduction with following information:

“The first production event is the spring bloom, which is controlled by the availability of nitrate and shifted from being dominated by diatoms to dinoflagellates in the late 1980s (Wasmund et al., 2017; Spilling et al., 2018).
After a so-called bluewater period with close–to–zero NCP rates, the second type of production events are mid–summer blooms of nitrogen-fixing cyanobacteria that develop in most years depending on meteorological conditions. Although cyanobacteria NCP is yet poorly constrained, its relative contribution to the annual NCP in the Eastern Gotland Sea in 2009 was estimated in the order of 40% (Schneider and Müller, 2018; Schneider et al., 2014), though the uncertainty is high. This preliminary estimate further needs to be interpreted with care, as cyanobacterial NCP varies significantly between years and regions."

This additional information was also requested by RC2.

In the paper you mention that your approach to NCP (using surface pCO2 and the modelled MLDs/TPDs) is limited to the time-periods of a stable or shoaling thermocline. Would you go as far as to estimate the % of time when the bloom falls outside this pattern based on your modelled MLD/TPD? This could give you some information/confidence about the applicability of the NCP approach using surface pCO2 only in the lack of full profiles.

As pointed out in the discussion (L420ff) and in line with several previous studies, we did not observe signs for a continued net production of organic matter during or directly after the deepening of the mixed layer. We therefore conclude that bloom events are entirely confined to periods of a stable or shoaling thermocline. We will explicitly state this by adding in L421:

"We thus conclude that reconstructed NCP estimates are not affected by a systematic underestimation due to this temporal restriction."

The authors may want to include the calculations of heat flux and use it as a proxy for CT* drawdown. These will likely correlate.

If our understanding of this comment is correct, you are asking for the correlation between heat fluxes across the air-sea interface and the CT* drawdown near the sea surface. In first approximation, the heat flux should be strongly correlated with the SST increase over the bloom period and under the condition of a stable thermocline, i.e. from June 6 to 24 in this study. Therefore, we argue that the heat flux can be replaced by SST increase to address this comment.

Indeed, previous studies revealed a strong positive correlation between the CT* drawdown and the increase in SST (rather than absolute SST) for individual bloom events (Schneider and Müller, 2018; and references therein). Although a strong correlation exists, the slope of this relationship revealed significant variability in space and time. As a consequence, the CT* consumption cannot be accurately predicted from a known SST increase or heat flux. The observational constraint of the CT* consumption therefore remains an essential component to determine NCP. In order to avoid the potential misconception that the CT* consumption can be constrained when only the heat flux or SST is known, and because we see no added value in re-addressing the previously discussed correlation, we do not intend to include the correlation between heat fluxes and CT* drawdown in this study. However, the previously found correlation between CT* drawdown and SST increase will be addressed as follows in Sect. 1.5 of the introduction:
“... it was found that the $C_{\text{fr}}$ drawdown correlates well with the co-occurring increase in sea surface temperature (SST), rather than with absolute SST. This relationship was attributed to a common driver, which is the light dose received by the water mass under consideration (Schneider and Müller, 2018).”

***

Specific comments:

L15: specify units of NCP (mol m$^{-2}$ (?)

“in moles of carbon per surface area” will be inserted next to the first mention of NCP in L5.

L20: within 10% (if taking into account the period of a stable thermocline?) please specify

For clarification, “... across the bloom event” will be added to the sentence. We consider the bloom event duration to be identical to the period of a stable thermocline. This coincidence was further clarified in L421. (Please refer also to our reply to the respective major comment.)

L30: see the general comments re conventional definition of NCP. Please add the missing info here.

This will be done. Please refer to our answer to your respective general comment.

L33: 1.2. Baltic Sea: please add more references citing studies of drivers and causes of hypoxia/anoxia in the Baltic Sea and sealed estuaries in general.

The following additional information and references will be added to support the readers access to the most relevant literature:

“The export of organic matter into the deep waters is considered the ultimate cause for the expansion of anoxic areas in the Baltic Sea, which are nowadays among the largest anthropogenically induced anoxic areas in the world (Carstensen et al., 2014). Although the actual oxygenation state of the deep basins of the Baltic Sea is modulated by the frequency and strength of inflow events (Mohrholz et al., 2015; Neumann et al., 2017) and the biogeochemical properties of the inflowing waters (Meier et al., 2018), the long-term expansion of the anoxic water body was primarily attributed to increased nutrient inputs from land (Jokinen et al., 2018; Meier et al., 2019; Carstensen et al., 2014; Mohrholz, 2018) that fueled the organic matter production in surface waters.”

This additional information and the paraphrasing of the original quote taken from Carstensen et al. (2014) was also requested by RC2.
L47: how important is the spring bloom in the Baltic Sea relative to the summer bloom? Which species dominate during the spring bloom?

Both questions will be addressed. Please refer to our answer to your respective general comment.

L58: ‘The limited understanding of the factors...’ This study does not explicitly address the factors controlling the bloom, instead examines the ways to quantify NCP. This paragraph needs rewording to include background information relevant to this study.

This issue was also raised by RC2 and we agree that this study itself does not explicitly address the controlling factors of the blooms. However, we expect a major contribution to this question when applying the new NCP reconstruction approach to almost two decades of SOOP observations. In order to clarify this, we will add the following sentence in L62:

“A long–term hindcast of cyanobacteria NCP and the attribution of its strength to prevailing environmental conditions in particular years could improve our understanding of controlling factors and facilitate more reliable predictions of the blooms. However, such a hindcast of cyanobacteria NCP was so far impossible due to missing vertically-resolved observations that would allow to constrain their organic matter production.”

Please refer also to our answer to your respective general comment.

L73: cite the literature examining equilibration time of gases.

The following citation will be included: Wanninkhof (2014)

L75-L80: do these statements refer to the Baltic Sea? Need to make this clear to the reader. Global N-based NCP estimates are very popular)

The statement refers to the general limitation that NCP can not be determined from the consumption (i.e. decreasing concentration) of nitrate for blooms of N-fixing organisms, which also applies to Baltic Sea cyanobacteria blooms. To clarify this, the respective sentences will be rephrased to:

“However, time series of nutrient consumption do not allow for determining NCP of algae blooms dominated by nitrogen-fixing organisms and those with highly variable C:P ratios. As both characteristics are typical for Baltic Sea cyanobacteria blooms (Nausch et al., 2009), the well established CT approach is the favorable method to determine mid-summer NCP in this region.”

L136: what are the accuracies and precision of the sensors after their calibrations?

The sentence in L136 will be rephrased as follows to address the measurement quality of the CTD instrument:
“Pre- and post-deployment calibrations of the instrument were carried out in the accredited calibration lab of IOW in the time span of a few month around the deployments and confirmed that the temperature and conductivity sensors achieved the typical accuracy of better than ±0.01 °C and ±0.01 S m⁻¹, respectively.”

L148: ‘of 1% of reading’. Is this true for this particular study? Did you compare pCO² (sensor) to pCO² (DIC-TA)? If so, this comparison should be included instead.

To address the sensor performance more clearly in the methods section, we will move the following information from the appendix to L148:

“Given the statistics of the pre- and post-deployment calibration, the small drift encountered throughout the deployment and the otherwise smooth performance and regular cleaning of the sensor during the deployment, the accuracy of the measurements is considered to be within 1% of reading as also found by Fietzek et al. (2014).”

Further details illustrating the excellent sensor performance are given in Appendices A1 – A3.

Unfortunately the suggested comparison of measured pCO² and pCO² calculated from DIC and TA is not meaningful to address the uncertainty of the measured values in the Baltic Sea. It was previously shown that calculated pCO² can deviate from measured pCO² by more than 200 µatm (Kuliński et al., 2014), due to contributions of dissolved organic acids to TA, which are not accounted for in routine CO² system calculations. We avoid a direct comparison, because the accuracy of calculated pCO² values is far lower than the expected accuracy of measured pCO².

L149: Please justify using A4.1 for tau determination and A4.2 for the correction? Why not to stick to one equation? Both work great, although A4.2 add more noise to the data.

We agree that our equation A4.2 can be replaced by A4.1 solved for the in situ pCO² (i.e. the true value), as originally proposed by (Miloshevich et al., 2004) and previously applied to measurements performed with the HydroC pCO² sensor by Fiedler et al. (2013) and Atamanchuk et al. (2015). Comparing our original response time corrected pCO² values based on A4.2 with those calculated based on the rearranged version of A4.1, we found that the vertical profiles gridded to 1m depth intervals are identical, i.e. the choice of the equation used has no impact on the biogeochemical interpretation of our measurements. However, we agree that there is no reason to apply two different types of equations and therefore replace the previous A4.2 with the rearranged version of A4.1 in our calculations and the appendix.

L159: based on (the length) of the release line??

Yes. The sentence was rephrased accordingly.
L160: HgCl2

Corrected.

2.4.1 should go before L 176

The order of the sections was changed.

L177: in my opinion, the term ‘estimate’ already refers to an indirect, non-empirical way to assign value to a certain phenomenon. ‘Best-guess’ should, therefore, be dropped.

We agree that our combined use of the terms “best-guess” and “estimate” represents a pleonasm. Still, we deem it very important to clearly distinguish our two types of estimates, i.e. the “best–guess” based on vertically resolved observations and the “reconstruction” based on surface observations. As dropping the term “best–guess” would reduce clarity, we will drop “estimate” when used in conjunction with “best–guess”. For example in L6 “... providing a best–guess NCP estimate ...” will be replaced with “... providing a NCP best–guess ...”.

L276: ‘Since July 6th ...’ what portion or % of the production you expect prior to July 6th?

Due to the availability of SOOP observation before July 6, it is indeed possible to approximate the production of this bloom that occurred before the first sampling event of the field campaign. The following information will be added to the description of the SOOP observations in L348ff:

“Based on SOOP observations before July 6, first signs of the onset of the investigated bloom event were detected on July 3. Between July 3 and 6, an SST increase of ∼1 °C was accompanied by a C\textsubscript{T}* drawdown of ∼10 μmol kg\textsuperscript{-1} (data not shown). Still, in the absence of any vertically resolved observation for this time period, the following comparison of the reconstructions to the best-guess needs to be restricted to the period July 6 – 24 during which the bulk of NCP occurred.”

L293: ‘that NCP has ceased’ ...do you mean there is no production (NPP), nutrient limitation or that NPP is balanced by respiration?

As the C\textsubscript{T}* approach resolves only the net effect of production minus respiration of organic matter, this statement refers to the fact that production is roughly balanced by respiration. This will be clarified by rephrasing the sentence to:

“... indicating that the production and respiration of organic matter were balanced during this period”
f)-g) Fair-sea or Fatm for clarify on Y-axis

h) flux = air-sea flux. ‘Flux’ is too ambiguous in this case. flux mixing corrected ( -NCP ) – production has a positive sign, so negative to integrated CT*

The labels will be changed as you recommend, both in figures and the corresponding manuscript text.

L400: Please also explain how to overcome the caveats. This info can be useful for the reader. What other measurements are needed?

In the new section (L399) we will include a discussion of the biogeochemical relevance and interpretation of our findings, where we will address this question as follows:

“We conclude that NCP estimates determined with the methods developed in this study are of direct relevance to quantify the drivers for deep water deoxygenation. However, a better understanding of the organic matter remineralisation processes would be required to close the budget of biogeochemical transformations. New observational platforms, such as recently deployed biogeochemical ARGO floats (Haavisto et al., 2018), will complement the existing SOOP infrastructure and help to provide the required observational constraints throughout the water column.”

L420: Indeed, it is unlikely that production will continue during the periods of deepening thermocline. However, deepening-restratification events may have an impact on both entrainment of CT and detrainment of POC. The latter, especially, can moderate more efficient transport from the photic zone to the depth bypassing the respiration stage. Therefore, it appears crucial to monitor the frequency of occurrence of such events as they impact both the production and efficiency of carbon export.

In the new section (L399) we will include a discussion of the biogeochemical relevance and interpretation of our findings, where we will address this question as follows:

“In contrast to shallow remineralisation processes, the deepening of the mixed layer that marked the end of the studied bloom event may facilitate the efficient transport of POC from the surface layer to depth. Focusing on the accumulation of remineralisation products beneath 150 m in the Gotland basin, a previous study revealed that – in accordance with the main input of POC during the productive period – remineralisation rates exhibit a pronounced seasonality.
This seasonality was found to be most pronounced in the water layers closest to the sediment surface, suggesting that beneath 150 m the remineralisation takes place mainly at the sediment surface and is of minor importance during particle sinking through the deep water column. The pronounced seasonality further confirms that surface organic matter production and deep water oxygen consumption are indeed tightly coupled, despite a potential degradation of POC before export across the permanent halocline.

This additional information was also requested by RC2.

***

Additional references used in this reply


