This manuscript presents a detailed account of glider and buoy measurements of dissolved O_2 and inorganic carbon (DIC) concentrations, which, together with physical fluxes due to horizontal advection, air-sea exchange, and mixing, are used to calculate net community production (N) in the euphotic zone of the NW Mediterranean Sea in spring 2016. Both the data, which are or will be openly accessible through BOUSSOLE, MISTRALS and BODC databases, and the derived N rates are important contributions to our knowledge on the biotic contribution to the ocean carbon sink, which is an important and timely scientific topic. The presentation of methods is detailed and comprehensive, including calibration of sensor, and results and calculations are also clearly presented. I am however less enthusiastic about the conclusions and discussion.

We are pleased that the reviewer found the manuscript interesting to read and we would like to thank them for the helpful comments. Please see below our comments on the discussion.

Conclusions seem a little vague to me, mainly focus on what has been done, rather than what has been observed. E.g., "this was the first time that high-resolution vertical profiles covering the wider DyFAMed area provided insights into the biogeochemical and physical processes during a spring bloom." However which are those insights provided by this study is not mentioned.

We will reorganise the conclusion section and add some further details.

I particularly don't agree with the conclusion that "this study demonstrates the capability of estimating N using measurements obtained by an autonomous glider". I do fully support the approach and agree that gliders have a unique potential to provide estimates of N at spatial and temporal scales that are unsuitable for other methods/platforms.

We will modify the text on line 305 as follows:

> "This study demonstrates the **potential** of using autonomous glider measurements for estimating N."

In this regard, I think that the paper is a very substantial scientific contribution. I also acknowledge the conscientious calculations of both rates and uncertainty, and their detailed description in the paper.

Thank you.

[…]I think that the paper would improve with a discussion on the causes for the large differences between the different N estimates, which are the rates that better explain the backscatter and chlorophyll-a build-up in the region, and a more complete comparison of bloom N estimates with the literature (including blooms elsewhere). Results here are only compared with Coppola et al. (2018) and Copin-Montégut (2000), yet it is acknowledged that such a comparison is not possible "because each study is focused on different timescales (from years to days) or different seasons."

The above comments provided by the reviewer highlight that calculating *N* and advection using glider pH and O_2 measurements is possible, but difficult. Bearing in mind their uncertainties the results are mutually compatible. Differences primarily arise from challenges relating to glider spatial coverage and sensor calibration. We have not shied away from this, considering that we have discussed such issues and potential contradictions in the manuscript. For example, we discuss the uncertainty of *N* in detail on lines 289 – 298. Previous studies may also have been too optimistic about uncertainties associated with their estimates.

We'd also like to refer to our discussion on l. 283, which refers to the good agreement between the March and April data of Coppola et al. (2018) and our results. The notion that there are large differences between *N* estimates is perhaps overstated; we would rather conclude that there are large uncertainties in *N* estimates, and that deriving *N* estimates from in situ measurements with a comprehensive uncertainty budget is comparatively hard.

Relating to the length of the discussion, it is also worth noting that some discussion has appeared outside of Section 7. For example, on lines 169 – 173 we discuss the opposing north-south gradients for glider O2 and satellite-derived ocean color, and on lines 264 – 271 we discuss the non-Redfield ratios in the context of previous studies and Redfield-derived carbon fluxes.

It is challenging to determine which *N* estimate is "better", however we will add a short paragraph further discussing the potential sources of errors for the *N* estimates. We will also add a conclusion that it is important to consider all systematic and random uncertainties when deriving *N* estimates.

Other specific comments:

- Figure 1 presents surface chlorophyll a concentrations on 24 March 2016. I think that the climatology for the period of study or the bloom period would be more useful as context for N estimates.

We show chlorophyll a concentrations on 24 March 2016 to highlight the patchiness of blooms in this region. This was useful for the discussion on lines 169 – 173 where we reference the figure.

- Which are the consecuentes of using a single mean euphotic depth of 46 m for calculating N throughout the study. Large temporal differences in backscatter (Fig.5) and spatial differences in chlorophyll-a (Fig.1) suggest that the actual euphotic depth should have changed substantially during the period of study, particularly associated to the phytoplankton bloom. Do this have an impact on N estimation under different scenarios? More specifically, is the calculated N an unbiased estimation of euphotic zone net community production both during periods when ZeuZlim? On the other hand, photosynthetic gross production (GP) is limited to the euphotic layer, however the respiration (R) of the organic matter produced is not; beyond entrainment, do the large

changes observed in the ratio between the euphotic and mixed depths (Fig.5) have an effect in the interpretation of N?

We have investigated the sensitivity of *N* estimates to z_{lim} (See figure 1 below). The overall pattern is similar in each case, however the choice of Z_{lim} can at times affect *N*, particularly for DIC-based *N* estimates (e.g. after 25/03). The largest absolute difference between single-day *N* estimates using a z_{lim} of 36 m and 56 m was 135 mmol m^{-2} d^{-1} on 26/03 for N_b(DIC).

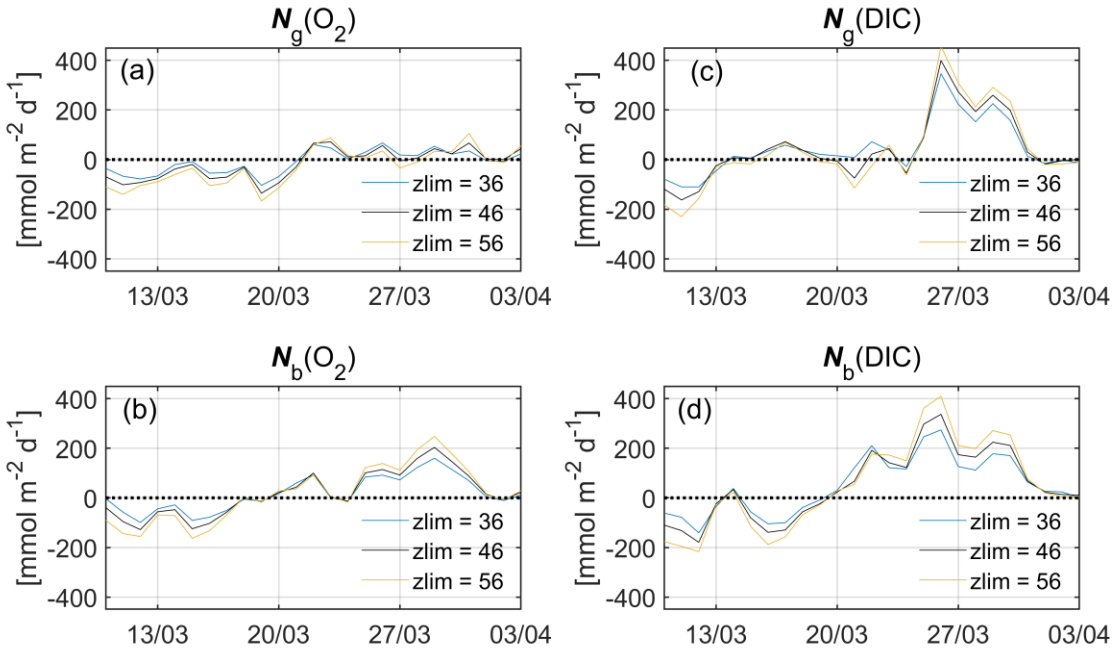

Figure 1 Comparing N_g and N_b estimates using different Z_{lim}.

We will add a brief discussion on this to the revised manuscript.

The 4-day smoothed MLD is > Z_{eu} most of the time before 25/03, and < Z_{eu} afterwards. We have found no relationship between the MLD/ Z_{eu} ratio and the *N* sensitivity to z_{lim}.