Deep chlorophyll maximum and nutricline in the Mediterranean Sea: emerging properties from a multi-platform assimilated biogeochemical model experiment

By Anna Teruzzi, Giorgio Bolzon, Laura Feudale, Gianpiero Cossarini

Point by point replies to review

Replies are in italic green, while we report in italic blue parts of the new manuscript version. Line numbers into brackets refer to the new version of the manuscript.

Review by Sarah Schlunegger

Referee comments on "Deep chlorophyll maximum and nutricline in the Mediterranean Sea: emerging properties from a multi-platform assimilated biogeochemical model experiment" by Teruzzi et al., currently in discussion at Biogeosciences.

Teruzzi et al presents the results from a data-assimilating biogeochemical model of the Mediterranean Sea. This work provides a novel case-study of the dual use of remotely-sensed (ocean color) and in situ (BGC-Argo) biogeochemical constraints in reconstruction of one year of the biogeochemical state of the Mediterranean Sea. Firstly, the data-informed model solution demonstrates fidelity with observations for a number of surface and depth-resolved biogeochemical metrics, such as the vertical position and longitudinal gradients of the deep chlorophyll maximum and its co-variance with the nutricline. Secondly, the study presents compelling evidence for the synergistic benefits of assimilating estimates of remotely-sensed chlorophyll concentrations in tandem with in situ, depth-resolved estimates of chlorophyll. Thirdly, the study discuses some of the nuanced differences between the impacts of assimilating satellite versus BGC-Argo chlorophyll upon the model solution. The most striking difference discussed is the strong seasonal signatures the different observational streams have upon the solution, with remotely-sensed chlorophyll providing stronger constraint during winter months, and in situ chlorophyll providing stronger constraint during summer months. Finally, implications for optimized sampling strategies, such as the recommendation that BGC-Argo increase sampling frequencies during the 'influential' summer season, are discussed.

This work represents a timely contribution to the underway community efforts towards optimizing the design of a global, autonomous biogeochemical observing network for use in constraining reconstructions of the evolving ocean state. I recommend this manuscript for publication after a few minor clarifications and elaborations are incorporated, as I outline below.

We thank Sarah Schlunegger for the positive and constructive comments on the manuscript. All the points raised in her review are addressed below.

Points of clarification:

The model setup describes the biogeochemical model as being coupled offline to the dynamical model. Coupled implies two-directional influence, i.e. that the biogeochemistry is both impacted by and impacts the dynamical fields. Is this the case?

If it is the case that the model set up is "coupled" in the true sense of the word, then the impact on the dynamical fields should also be presented.

If it is not the case, and in fact the biogeochemical model or assimilation of biogeochemical fields does not feed-back or modulate the underlying dynamical fields, then a more appropriate term to use is "forced" or "driven" – i.e. Line 135 should read "the MedBFM was forced or driven offline with output from MENO3.4-OceanVar…"

In MedBFM, the biogeochemical variables are tracers advected and diffused by ocean dynamics and transformed by biogeochemical processes, adopting a source splitting operator (Butenschön et al., 2012) without feed-back on the dynamical fields. In particular, transport and advection are driven by precomputed daily 3D fields of currents and diffusivity, while T and S are used in the biogeochemical module. We modified the manuscript following the Reviewer's suggestion to better explain this aspect as follows:

(L. 142-153) "In the current application, the MedBFM was forced offline with outputs from the NEMO3.2 model of the Mediterranean CMEMS model system (Simoncelli et al., 2016)"

To improve accessibility, an additional paragraph of description or context of Eq. 1 should be included. This would involve, for instance, expalaination of the "innovation" term (Line 159), and discussing the significance or intuitive purpose of the different covariance vectors.

In the new version of the manuscript we revised the description of terms used in eq. (1):

(L. 161-170) "In 3DVarBio, assimilation is performed through the minimization of a cost function that is defined on the basis of Bayes' theorem (Lorenc, 1986) as the weighted sum of the square mismatches between the model background state $\boldsymbol{x_b}$ (the model state before the assimilation) and the analysis $\boldsymbol{x_a}$ (the assimilation result) and the observations $\boldsymbol{y}$. Each square mismatch is weighted according to its accuracy estimations, meaning that $\boldsymbol{x_a} - \boldsymbol{x_b}$ is weighted by the background error covariance matrix $\boldsymbol{B}$ while ($\boldsymbol{y}$– $H(\boldsymbol{xb})$) by the observation error covariance matrix $\boldsymbol{R}$:

$$J(\boldsymbol{x_a}) = (\boldsymbol{x_a} - \boldsymbol{x_b})^T \boldsymbol{B^{-1}} (\boldsymbol{x_a} - \boldsymbol{x_b}) + (\boldsymbol{y} - H(\boldsymbol{x_b}))^T \boldsymbol{R^{-1}} (\boldsymbol{y} - H(\boldsymbol{x_b})). \quad (1)$$

In eq. (1) $\boldsymbol{y} - H(\boldsymbol{x_b})$ is usually named innovation and H the observational operator that maps the values of the model background state $\boldsymbol{x_b}$ in the observation space. In our application $H(\boldsymbol{x_b})$ are model values of the variables observed by satellite or floats at observation locations. Through the minimization of the cost function (1), the assimilation provides the analysis $\boldsymbol{x_a}$, i.e., the optimal weighted distance from both $\boldsymbol{y}$ and $\boldsymbol{x_b}$."

Regarding the "Impact indicator" metric:

(a) It needs to be explained that the impact indicator given by Eq 2, although it does not contain an explicit 'directionality' of the impact (i.e. towards better or worse agreement with the data-constraints), that the assimilation methodology works to push the model solution towards the data, therefore any non-zero value of I(t) represents a nominal improvement in model-data fit. From the equation alone, and without sufficient prior understanding of the methodology, this is not obvious, making it difficult to interpret if the additional streams of assimilated data are merely influencing or in fact improving the solution.

Thank you for the very constructive comments about the Impact Indicator. Concerning the first comment (a), it is true that the impact is in the direction of the assimilated observations for grid-points corresponding to observation locations. On the other hand, grid-points without observations are corrected consistently with the solution of the DA scheme (e.g., the effects of surface chlorophyll assimilation on the water column profiles) with impacts vanishing for grid-points far from any observation locations. Thus, the impact indicator is higher where the assimilation brings the simulation toward the observations and this can be considered an *improvement of the simulation with respect to nearby assimilated observation. As suggested by the Reviewer, modified the description of the impact indicator in order to highlighting this aspect:*

(L. 286-290) "The subscript 200 represents the integral over the 0-200 m layer, while the overbar represents the average over the whole Mediterranean and over seasonal periods. The impact indicator $I_{xy}(t)$ was calculated for each assimilation date and each grid point, and then, statistically analysed and summarized on a seasonal base. The indicator $I_{xy}(t)$ quantifies how much an assimilated run deviates from the REF simulation, thus it is higher where and when the simulation deviates from REF, and is closer than REF to the assimilated observations."

Moreover, to clarify that impact indicator is calculated for each model grid-point, we added the xy subscript in its definition and in all the occurrences:

(L. 284) "$I_{xy}(t) = \frac{|ScFcn(t) - REF(t)|_{200}}{\overline{REF(t)_{200}}}$ *(2)"*

(b) The maps of Fig. 6 and Fig. 7 present a very derived / abstract metric. These figures present the 50th and 90th percentile 'impact' of assimilated the given field during the given season. Interpretability of this abstracted metric would improve if an additional subpanel was included in each figure which showed a representative distribution of the impact-indicator for a single grid-cell (or averaged over a region), with markings at the 50th and 90th percentile. This could also provide opportunity to contrast the summer vs. winter distributions of the indicator. For instance, for chlorophyll, the winter distribution of the indicator would be shifted toward "1" while the summer distribution would be closer to "zero". See a mock-up below. This will help orient the reader as to what the maps are presenting.

We thank the Reviewer for this constructive comment about the Impact Indicator. Following the Reviewer's suggestion, we investigated the distribution of the Impact Indicator over winter and summer. In particular, we analysed frequency distributions of the impact indicator at some locations, and we provide in Fig. R1 two examples. The histograms confirm results shown in Fig. 6 and 7, with higher Impact factor percentiles in winter with respect to summer for chlorophyll and the opposite for nitrate (since the impact of sparse float observations is negligible at the simulation beginning). Further, the histograms are quite scattered since they depend on a number of factors (e.g., the number of observations in the area and assimilation changes during subsequent dates). Therefore, we think that the non-parametric statistics (percentiles) shown in the manuscript are a sound choice to quantify the impact of the assimilation on model solution.

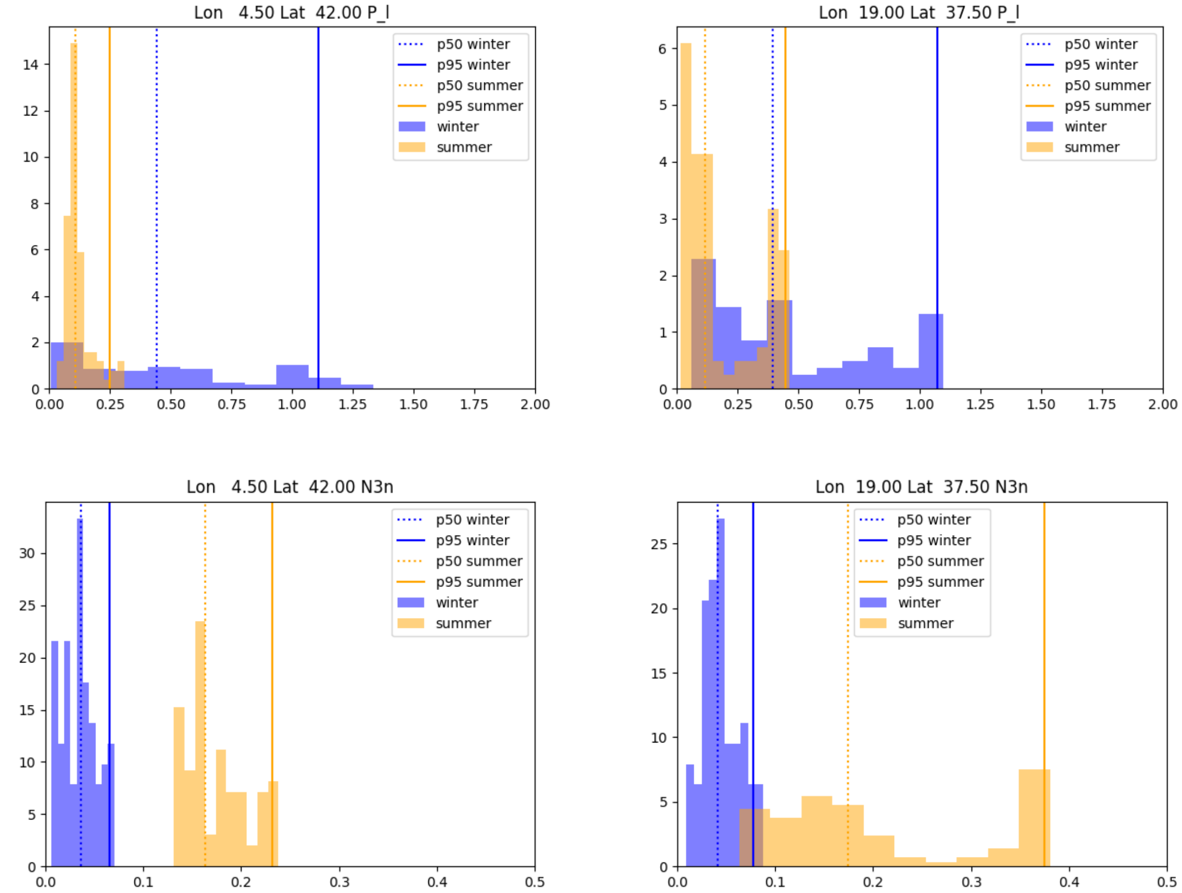

Fig. R1. Frequency histograms of the Impact Indicator at two locations for chlorophyll (top) and nitrate (bottom) in winter (blue) and summer (orange). Vertical lines indicates the values of the 50th (p50, dashed line) and 95th (p95, continuous line) percentiles.

The conclusions section would benefit from a final sentence that poses the significance of the study within the context of future advances in biogeochemical data assimilation within basin and global domains, i.e. something like " The multi-platform assimilation yielded improvements in model representation of large-scale (hundreds to thousands of kilometers) bio-dynamical features and is suggestive of the applicability of this advancement to reconstructions of other ocean regions and the global domain."

We thank the reviewer for this suggestion that helps to enlarge the perspective of our manuscript. We added a comment about the applicability of multi-platform DA to other domains: :

(L. 517-521) "The impacts of multi-variate profile assimilation are directly linked to the sampling frequency and dimension of the BGC-Argo network, which should increase to match the consolidated importance and relevance of satellite observation assimilation. Thus, in a perspective view, the multi-platform assimilation can improve model representation of both large-scale (hundreds to thousands of kilometres) to mesoscale features and be beneficial for robust reconstruction in global and regional reanalysis."

Minor editorial and stylistic suggestions:

L7-10: Rewrite to something like: "Data assimilation has lead to advancements in biogeochemical modelling and scientific understanding of the ocean. The recent operational availability of data from BGC-Argo floats, which provide valuable insights into key vertical biogeochemical processes, stands to further improve biogeochemical modelling through assimilation schemes that include observations from floats in addition to traditionally assimilated satellite data." (bold is new)

Thank you for the suggestion. We included it in the revised manuscript (L. 7-9).

L16: "maximum depth, intensity and nutricline depth" (added a comma and removed the first 'and')

We modified the manuscript accordingly (L. 17).

L39: Add the following reference, which also provides motivaiton for the direct use of optical properties in data assimilation:

Dutkiewicz, Stephanie, Anna E. Hickman, Oliver Jahn, Stephanie Henson, Claudie Beaulieu, and Erwan Monier. 2019. "Ocean Colour Signature of Climate Change." Nature Communications. 10 (1). https://doi.org/10.1038/s41467-019-08457-x.

We inserted the reference (L. 38)

L40: Ocean-colour observation assimilation takes advantage of the frequent, large-scale satellite observations of ocean properties related to the microbial biology of the upper ocean.

L42: "deeper ocean layers requires approximations and assumptions."

We changed the manuscript according to the two previous Reviewer's suggestions (L. 40-41 and L. 42).

L45: "localization" – what does this mean?

In Ensemble Kalman filter (EnKF)-like data assimilation, "localization" is a method applied to limit the impact of the assimilation (increments) to areas relatively close to the observations. In particular, localization is applied when knowledge of spatial covariance is relevantly approximated due to the limited ensemble size. In this sense, localization avoid to insert spurious and non-realistic effects and can be applied not only to limit the spatial covariance but also on covariance among variables or on the time dimension. Concerning the use of "localization" in the sentence at L. 45, we are referring to vertical localization applied in some DA application of satellite ocean colour, meaning that the assimilation impacts are limited to a portion of the water column through localization. We modified the sentence as follows to make it more clear about "localization" in EnKF-like data assimilation:

(L. 42-46) "Vertical covariance must be parameterized by synthetic precalculated vertical profiles in variational schemes (Teruzzi et al., 2018), while EnKF-like (ensemble Kalman filter) schemes may have limitations in effectively impacting deeper ocean layers (Fontana et al., 2013; Hu et al., 2012). Indeed, some EnKF-like applications introduce limitation to the increments in subsurface layers through localization in the vertical direction to address spurious correlations (Goodliff et al., 2019; Pradhan et al., 2019)."

L66: “ideal” – replace with “suitable” or “appropriate” – as it could be easily argued that a region with less coastal margins and more open-ocean conditons, where ocean color is more reliable, would be a more ‘ideal’ location to do a multi-platform assimilation.

We replaced “ideal” with “suitable” (L. 67).

L348: “In our results, this hypothesis was supported by higher nutrient uptakes in the western…”

We replaced “scenario” with “hypothesis” (L. 364).

L409: Remind readers of what Vv means, i.e. “ Vv (vertical co-varience error)”

L415: Remind readers of what Vb means, i.e. “ Vb (biogeochemical co-varience error)”

In the reviewed Discussion the meaning of the operators is reminded at the first occurrence of each operator (L. 429 and L. 459).

Review by Anonymus Referee #2

Review of "Deep chlorophyll maximum and nutricline in the Mediterranean Sea: emerging properties from a multi-platform assimilated biogeochemical model experiment" by Teruzzi et al.

The manuscript addresses the performance of a 3D-Var biogeochemical data assimilation system, constrained with both chlorophyll data from satellite and chlorophyll and nitrate data from BGC-Argo floats, that is applied to a realistic simulation of the Mediterranean Sea for the year 2015. After demonstrating the validity of the method, the authors use their product to investigate the spatial and seasonal variability in the vertical structure of chlorophyll and nitrate fields. The problem is introduced clearly, methodology appears sound, and results are compelling. I recommend the manuscript for publication in Biogeosciences after revisions. Below are a couple of important results that need to be clarified, a list of issues to be addressed, and some minor comments.

We thank the Anonymous Reviewer for the positive comment on our manuscript and the careful reading. All the points raised by the Reviewer are addressed in the following.

1) Fig 2: With the exception of the "LEV" region, it seems that assimilating only float data yields a better fit with chl obs than assimilating floats + satellite. It seems like it would be best to ignore satellite data, even at the surface. On the other hand (Fig 3), assimilating satellite Chl yields a better fit with nitrate obs, which is counter intuitive, and assimilating float Chl barely has an impact. Please comment!

This Reviewer's comment helped us to clarify some aspects that in the current manuscript version are not probably appropriately presented and discussed:

In Fig. 2, RMSDs are calculated with respect to float chlorophyll. It is not obvious that satellite chlorophyll assimilation would positively affect metrics with respect to float chlorophyll because of erroneous propagation of surface information along the vertical and because of potential inconsistency between satellite and float data. Indeed, we discussed the inconsistency between satellite and float chlorophyll in the Discussion section, where a comparison between values of chlorophyll concentration from satellite and float highlights their differences. Thus, since satellite and BGC-Argo float inconsistency at surface and because vertical covariance is a prescribed propriety, the change in chlorophyll profiles due to satellite assimilation is either reducing and increasing the distance between modelled chlorophyll and BGC-Argo chlorophyll profiles (Fig. 2). On the other hand, considering RMSDs with respect to satellite chlorophyll (Fig.5), the effect of inconsistency between satellite and float chlorophyll is highlighted by the slight increase in RMSDs in simulations with float chlorophyll assimilation. Thus, we think that the model-assimilation can act as a filter solving inconsistency between sensors, meaning that even if the performance of the multi-platform assimilation is lower than anyone of the single assimilation, it produces a balanced solution with respect all the available information.

Concerning Fig. 3:

- *With the exception of TYR, the assimilation of satellite chlorophyll reduces the RMSD of nitrate (computed on BGC-Argo data) with respect to REF simulation. This is because satellite assimilation correct surface modelled phytoplankton dynamics continuously during late winter/spring (i.e., reducing bloom maxima), and, as a results the entire profile of the phytoplankton is nearly uniformly modified (i.e., reduced) in the euphotic layer. In turn, by reacting to the new phytoplankton conditions, less nitrogen is eventually re-mineralized to nitrate and nitrate concentration is modified in the direction of reducing the REF overestimation in the upper layer. This mechanism has been investigated in previous applications of satellite chlorophyll assimilation (Teruzzi et al., 2018, 2014).*

- *In the case of BGC-Argo chlorophyll assimilation, the changes in the phytoplankton profiles are less uniform (Cossarini et al., 2019), often alternating positive and negative increments along the same profile. The effects on nitrate profiles (through new uptake or release after mortality and exudation) are non-linear and not uniform. It follows that also RMSDs with respect to float nitrate can both decrease or increase with respect to REF. Effects on non-assimilated biogeochemical variables are discussed in a number of works (e.g., Ciavatta et al., 2014; Ford, 2020; Mattern et al., 2017; Santana-Falcón et al., 2020; Simon et al., 2015; Teruzzi et al., 2018; Tsiaras et al., 2017; Yu et al., 2018), where the non-degradation of non-assimilated variables is considered a good result of the assimilation process. In the present manuscript, we briefly discussed the effects of satellite chlorophyll assimilation on nitrate RMSDs (L. 225-228) but we did not summarized these effects in the Discussion.*

Considering the points highlighted above, we modified the manuscript adding paragraphs dedicated to the effects on non-assimilated variables. In particular, we focused on the mixed effects of satellite chlorophyll assimilation, which slightly degrade metrics with respect to float chlorophyll but improves those with respect to float nitrate. In perspective, when the inconsistency between satellite and float will be solved, the multi-platform assimilation will provide improvements over large areas thanks to the relevant spatial coverage of satellite observations. In the meantime, as discussed in the manuscript (and reinforced by new comments) the model, acting as a dynamical filter, effectively integrates both sources of information. Hereafter the new paragraphs added in the manuscript:

(L. 422-431) "As a consequence of these discrepancies and of the propagation of information through the prescribed DA vertical covariance , the change in chlorophyll profiles due to satellite assimilation either reduced and increased the distance between modelled chlorophyll and BGC-Argo chlorophyll profiles (Fig. 2). On the other hand, considering RMSDs with respect to satellite chlorophyll (Fig. 5), the effect of inconsistency between satellite and float chlorophyll was highlighted by the slight increase in RMSDs in simulations with float chlorophyll assimilation. In perspective, when the inconsistency between satellite and float will be solved, the multi-platform assimilation will provide improvements over large areas thanks to the relevant spatial coverage of satellite observations. In the meantime, the model, acting as a dynamical filter, integrates both sources of information together with the vertical covariance V_V operator implemented in the assimilation scheme. Indeed, the vertical covariance is a key element that allows to integrate surface with sub-surface information."

(L. 439-449) "Considering effects on nitrate, the assimilation of satellite chlorophyll reduced the nitrate RMSDs (computed on BGC-Argo data) with respect to REF simulation in all the sub-basins with exception of TYR (Fig. 3). The rather persistent and broad DA increments during late winter and early spring were acting to reduce overestimation of the bloom maxima, resulting in nearly uniform reductions of the phytoplankton biomass in the euphotic layer. In turn, by reacting to the new phytoplankton concentration, less nitrogen was eventually re-mineralized to nitrate. Thus, the nitrate concentration was modified in the direction of reducing the REF nitrate overestimation in the upper layer. An analogous mechanism has been investigated in previous applications of satellite chlorophyll assimilation in the MedBFM system (Teruzzi et al., 2018, 2014). In the case of BGC-Argo chlorophyll assimilation, the changes in the phytoplankton profiles were less uniform (Cossarini et al., 2019), often alternating positive and negative increments along the same profile. It follows that the effects on nitrate profiles (through new uptake or release after mortality and exudation) were non-linear and not uniform, with relatively small impacts on the overall nitrate RMSDs."

2) L. 242-244, 260-261: In Fig 4 I don't see a reduction in RMSD, but instead an increase in levels 4-6. So data assimilation does reduce the model skill in fitting oxygen? It looks like the pink and red curves are on top of each other, suggesting that satellite Chl is responsible for degrading the O2 solution.

Thanks to this Reviewer's comment on the slight degradation of oxygen solution (Fig. 4), we went carefully through the oxygen validation results. Firstly, we considered a new recently updated dataset of BGC-Argo *oxygen measurements available at the Coriolis/Ifremer data assembly centre (float trajectories are shown in Fig. R2). Using the updated oxygen dataset, the recalculated RMSD values (Fig. R3) are lower than in Fig. 4 of the original manuscript, indicating that the simulation better compares with respect to the more recent and possibly more reliable observations. At the same time, RMSDs very slightly differ among simulations, especially in the eastern sub-basins. Moreover, the effect of satellite chlorophyll assimilation is not univocal, since RMSDs are both slightly reduced (TYR) and increased (NWM) . On the other hand, the float assimilation has a very little effect on oxygen RMSDs. The limited and non-univocal effects of the assimilation on oxygen metrics are related to the interaction of a number of trophic processes (e.g., phytoplankton production and respiration, zooplankton and bacteria respiration) after the assimilation changes on phytoplankton biomass. We think that the non-degradation of the oxygen RMSDs is a good result of the assimilated simulations.*

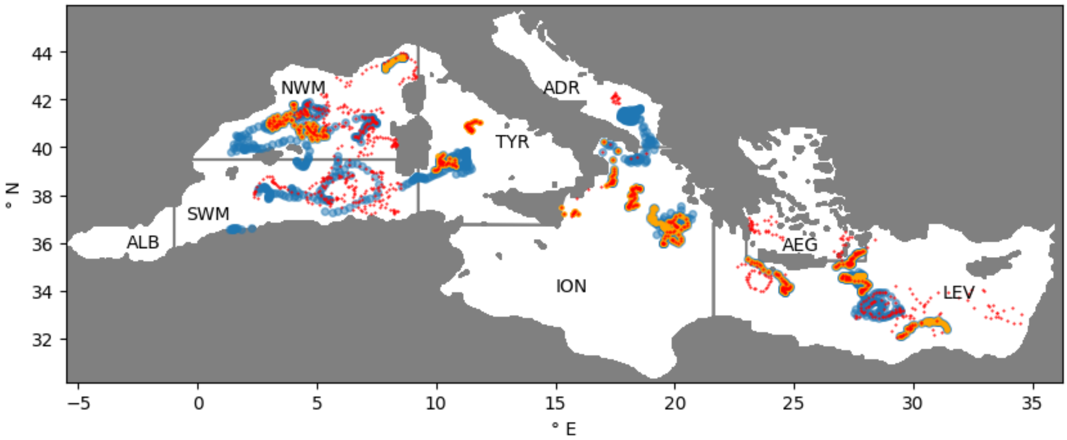

Fig. R2. Positions of BGC-Argo floats equipped with sensors to provide chlorophyll (blue), nitrate (orange) and oxygen (red) and limits of the subbasins.

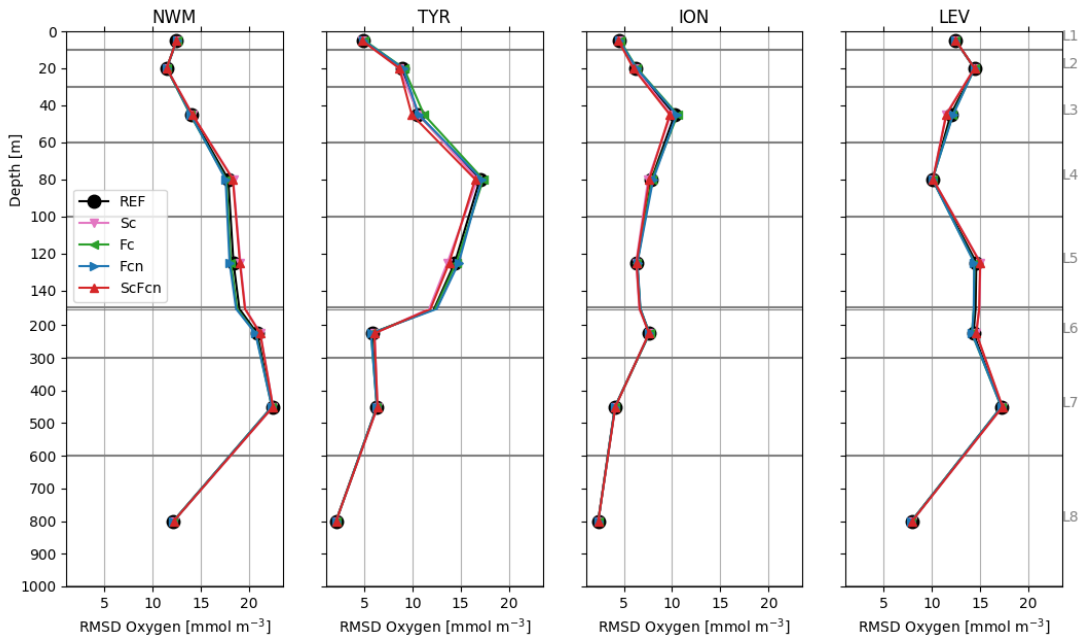

Fig. R3. RMSD between model simulations and BGC-Argo oxygen data in four sub-basins. Grey lines indicate the limits of layers L1-L8 used to calculate the RMSD. The depth scale is different above and below 150 m (double grey line).

We also calculated RMSDs using only oxygen profiles at locations where floats were assimilated (Fig. R4 and R5), thus excluding profiles far from float assimilations. While, oxygen RMSDs are further reduced in this case in the surface layer in NWM and in almost all layers of LEV, differences of RMSDs between simulations are very small also for this dataset and similar to those of Fig. R3.

According to the above considerations, we updated the manuscript introducing metrics based on the updated BGC-Argo oxygen dataset. In particular, Fig. 1 and Fig. 4 were replaced by Fig. R2 and Fig. R3, respectively. Moreover, we commented on the slight effect on oxygen in the assimilated simulations, which do not insert degradation on the non-assimilated variable. The relevant parts in the new version of the manuscript read as:

(L. 251-257) "The RMSD between the float oxygen data and REF simulation (Fig. 4) increased from the surface (approximately 5-15 mmol m-3) to the sub-surface and deeper layers (approximately 15-20 mmol m-3) in the NWM and TYR sub-basins , while it was almost uniform along the vertical in the eastern sub-basins with ranges 2-10 mmol m-3 and 7-17 mmol m-3 in ION and LEV, respectively. In particular, RMSDs very slightly differed among simulations, especially in the eastern sub-basins. In particular, the float assimilation had a very little effect on oxygen RMSDs and the effect of satellite chlorophyll assimilation is not univocal, since RMSDs are both slightly reduced (TYR) and increased (NWM) in Sc simulation. On the other hand, the float assimilation had a very little effect on oxygen RMSDs."

(L. 449-458) "Concerning the assimilation effects on the non-assimilated oxygen variable, the non-degradation of the oxygen RMSDs was a positive aspect of the assimilated simulations, while the limited and non-univocal effects of the assimilation on oxygen RMSDs were related to the interaction of a number of trophic processes (e.g., phytoplankton production and respiration, zooplankton and bacteria respiration) after the assimilation increments on phytoplankton biomass. Effects on non-assimilated biogeochemical variables are discussed in a number of works (e.g., Ciavatta et al., 2014; Ford, 2020; Mattern et al., 2017; Santana-Falcón et al., 2020; Simon et al., 2015; Teruzzi et al., 2018; Tsiaras et al., 2017; Yu et al., 2018), where the non-degradation of non-assimilated variables is considered already a good result of the assimilation process. Moreover, the model-assimilation system acts as a filter so that, even if the performance of the multi-platform assimilation is lower than anyone of the single assimilation, it produces a balanced solution with respect to all the available information."

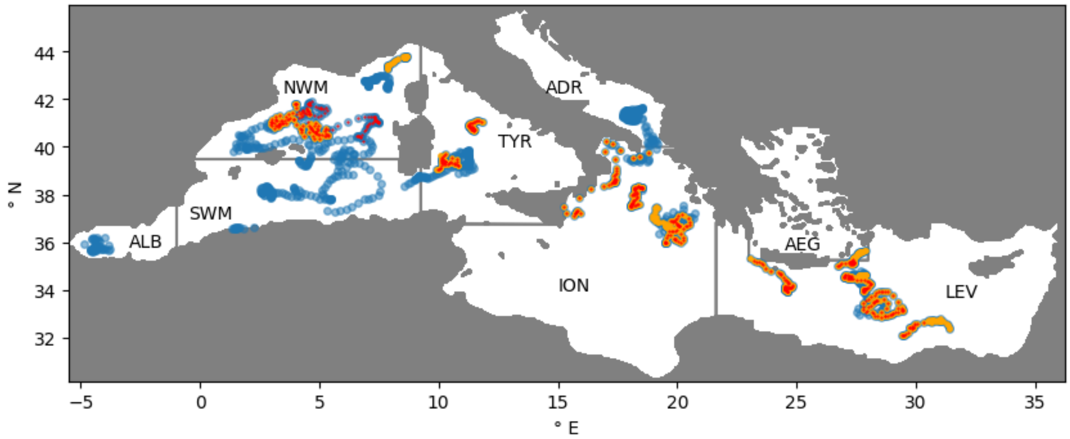

Fig. R4. Positions of BGC-Argo floats equipped with sensors to provide chlorophyll (blue), nitrate (orange) and oxygen (red) and limits of the subbasins. Only oxygen profiles where floats are assimilated are indicated.

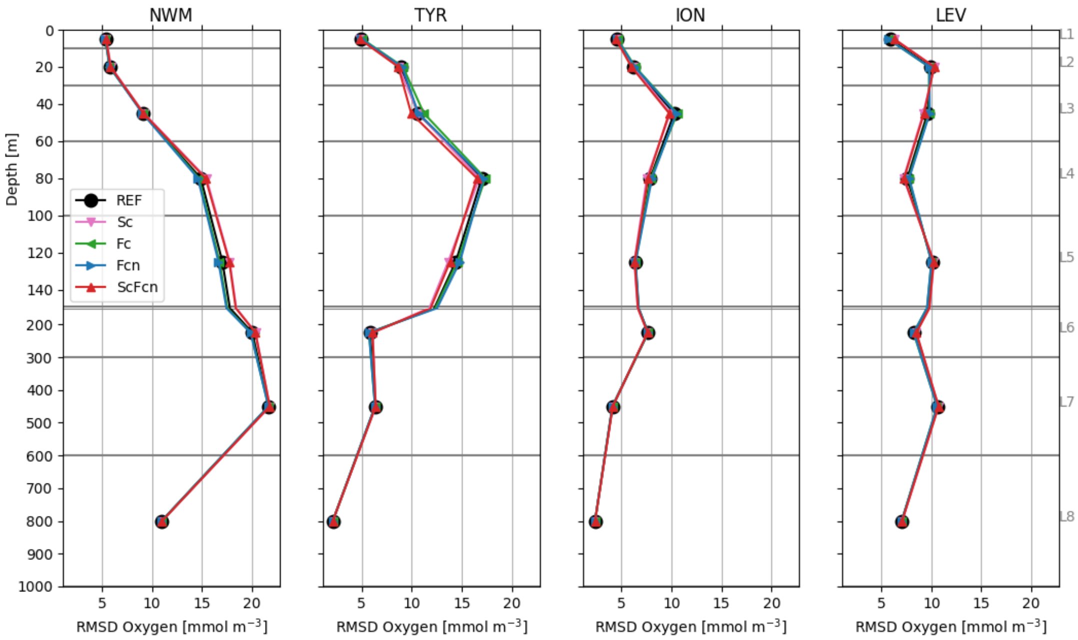

Fig. R5. RMSD between model simulations and BGC-Argo oxygen data in four sub-basins only at locations with float assimilation. Grey lines indicate the limits of layers L1-L8 used to calculate the RMSD. The depth scale is different above and below 150 m (double grey line).

Other comments:

3) Abstract and L. 68: is "semi-independent data" a common way to refer to observations used in model-data comparisons before and after assimilation? I suggest changing "semi-independent data (before assimilation)" to "assimilated data (before and after assimilation)".

"Semi-independent data" refers to the use for validation of observations before the assimilation. We changed the term in "assimilated data (before the assimilation)" (L. 13-14 and L. 69).

4) When data is excluded from assimilation based on a DA criterion, as stated in L. 98-100, is that a "quality check" on the data? Or is it removing observations that can't be fit because of model inadequacies? Is there a reference for the threshold values (5 mg/m3 for Chl and 1 or 2 mmol/m3 for NO3)? Similarly on L. 87-88 "A further quality check on satellite values before the assimilation resulted in the exclusion satellite chlorophyll observations whose mismatch value with respect to the model was higher than 10 mg m-3." - Is this a check of the quality of the data or a DA-based exclusion criterion?

5) L. 96-97: "chlorophyll profiles were checked for negative values (rejection)" - I imagine that a bias correction is applied to the Chl data, resulting in sometimes negative values at some depths. Does that necessarily mean the profile can't be used? Maybe negative values could be viewed as "zero"?

We thank the Reviewer for these comments . In the revised manuscript we clarified that two levels of quality check are performed on both satellite and float observations. The first quality check is made on observations independently from data assimilation and is intended to the exclusion of spikes and possible unrealistic values. For satellites, this step consists in "removing observations whose anomalies were higher than 3 times the daily climatology standard deviation" (L. 85-86). Float observations were instead checked with a more *complex (and somehow empirical) procedure, indeed, the relative novelty of BGC-Argo data sets release with respect to consolidated procedures adopted in ocean colour data processing may lead to a higher occurrence of poorly reliable observations. In particular: "nitrate profiles were rejected if the surface value was higher than 3 mmol m^{-3}, chlorophyll profiles were checked for negative values (rejection), and quenching correction was performed by imposing a constant chlorophyll value in the mixed layer" (L. 96-98). In this phase, the exclusion of single negative values (not of the entire profile) is thus based on the fact that negative concentrations cannot be accepted in the float dataset and on the choice to not insert any artificial changes on concentrations. The exclusion of nitrate values higher than 3 mmol m^{-3} in the surface layers is based on the analysis of climatological values provided in the EMODnet dataset (Buga et al., 2018).*

The second phase of observation checks is a pre-data-assimilation procedure, intended to exclude values not suitable for assimilation due to the range of variability model-observation differences to be consisted with the assumed uncertainties levels. For satellite, this check consists in the exclusion of "observations whose mismatch value with respect to the model was higher than 10 mg m^{-3}" (L. 89), while for floats the threshold values were set to 5 mg m^{-3} for and 1 or 2 mmol m^{-3} for chlorophyll and nitrate, respectively. These values have been calibrated after a statistical analysis of the distribution of the model-observations mismatches of the REF run. The exclusion based on DA criteria were of the order of 2%, l<1% and 3% for satellite chlorophyll and float chlorophyll and nitrate, respectively. In the new version of the manuscript we explained more clearly that the quality checks are sub-divided in two different phases and we added information on exclusion occurrences and comment about them:

(L. 88-94) "Two levels of quality check were performed on observations. The first quality check was made independently from data assimilation, and consisted in removing observations whose anomalies were higher than 3 times the daily climatology standard deviation in order to remove spikes. A second pre-data-assimilation check rejected satellite observations not suitable for assimilation, excluding observations whose mismatch value with respect to the model was higher than 10 mg m-3 to keep the range of model-observation differences consistent with the assumed uncertainties levels. The threshold was calibrated by a statistical analysis of the model-observations mismatches of the REF run. Through the pre-assimilation criterion, about 2% of satellite chlorophyll values were considered not suitable for the assimilation and rejected."

(L. 101-111) "As for satellite observations, two levels of quality check were applied on float observations. In the first check, which excluded unrealistic values, an expert judgment procedure was applied, because the relative novelty of BGC-Argo data sets release may lead to a higher occurrence of poorly reliable observations. In particular, nitrate profiles were rejected if the surface value was higher than 3 mmol m-3, chlorophyll profiles were checked for negative values (rejection of single negative observations), and quenching correction was performed by imposing a constant chlorophyll value in the mixed layer. The exclusion of nitrate values higher than 3 mmol m-3 in the surface layers was based on the analysis of climatological values provided in the EMODnet dataset (Buga et al., 2018). In the pre-data-assimilation check (second quality check phase) observations were rejected when the mismatch with respect to the model was higher than 5 mg m-3 and 2 mmol m-3 for chlorophyll and nitrate, respectively. As for satellite, these values were based on a statistical analysis of the model-observations mismatches in REF. The pre-assimilation check excluded nearly 3% and less than 1% of the float nitrate and chlorophyll observations were excluded, respectively."

6) L. 115-116: "3DVarBio is the data assimilation scheme for the correction of phytoplankton functional type and nutrient variables (i.e., nitrate and phosphate)" - It sounds like phytoplankton type is a control variable, which to me suggests that the type can be changed by assimilation. However, I believe you mean that the biomass in each phytoplankton class can be optimized. It would be useful to know how many functional types there are (it's mentioned later, but at this point I was wondering). Also, this is one of a few places where

phosphate is mentioned. Why control phosphate and not other bgc variables, like oxygen? Does the model include other nutrients?

The term "phytoplankton functional types" is usually used in multi-phytoplankton multi-nutrient complex model and satellite works (Butenschön et al., 2016; Nair et al., 2008) but we agree that in this contest can be misleading. For sake of clarity, we mention that BFM has 4 PFTs (diatom, picoplankton, nanoflagellates and dinoflagellates; Lazzari et al., 2012). However, since this information is not relevant at that part of the manuscript, we re-formulated the sentence:

(L. 137-139) "3DVarBio is the data assimilation scheme for the correction of phytoplankton biomass and nutrient concentrations (i.e., nitrate and phosphate) using surface chlorophyll from satellite observations and vertical profiles of chlorophyll and nitrate from the BGC-Argo floats."

Even if other nutrients are simulated in BFM model, the update of phosphate is a nutrient a key element in the Mediterranean basin, where both nutrients can act as limiting factors of phytoplankton growth (Lazzari et al., 2016) (L. 190). On the other hand, actual silicate concentrations in OGST-BFM applications in the Mediterranean Sea do no lead to limitation conditions, thus we did not applied assimilation update on this variable. Oxygen as well is not a limiting factor for plankton dynamics in our simulations of the Mediterranean Sea biogeochemistry, moreover, since the number of BGC-Argo floats equipped with oxygen sensors, it would be preferable to implement an assimilation scheme directly based on oxygen observations instead of on pre-computed covariance between oxygen and float chlorophyll or nitrate. Indeed, BGC-Argo oxygen assimilation is one of the foreseen update of the OGSTM-BFM system. Finally, it should be considered that the addition of variables in the assimilation scheme would increase its computational costs, compared to relatively small improvements in case of non-limiting nutrients or relatively high uncertainty in the definition of covariance between the assimilated and non-assimilated nutrients.

In the new version of the manuscript, we added details on the choice to apply assimilation updates based on nitrate observations to nitrate and phosphate only.

(L. 178-180) "In the present study, the 3DVarBio assimilation scheme was adapted to assimilate float nitrate data and both satellite and float chlorophyll data and to provide corrections on all the phytoplankton variables and on nitrate and phosphate concentrations."

(L 200-201) "Adopting a conservative approach, other nutrients were not updated by DA since their less relevant role as limiting factor in the Mediterranean Sea"

(L. 466-470) ". In the present application the covariance operator V_B provided impacts on all the phytoplankton variables together with those on two nutrients (phosphate and nitrate), which can act as limiting factors of phytoplankton growth in the Mediterranean Sea (Lazzari et al., 2016). In perspective V_B could be developed to include other variables, however, considering silicate, it should be noted that in OGST-BFM applications in the Mediterranean Sea silicate limitation is less relevant compared to nitrate or phosphate."

7) Section 2.2 is hard to follow, with so many models and acronyms. I would start by describing MedBFM, then OGSTM and BFM, then 3DVarBio. What does it mean that the transport model is "fully consistent with

the off-line coupling of the NEMO3.6 vvl"? On L. 135 the "MENO3.4-OceanVar model" is mentioned: should it be NEMO? If so, why mention NEMO3.6 earlier when you're using version 3.4?

We modified Section 2.2 according to the Reviewer. Concerning NEMO version, we corrected the versioning: we used NEMO3.2 (L. 120-139, L. 142).

8) L. 141-145: What does it mean for climatological profiles to "integrate" data? Is it different from initializing a model from data and spinning it up for a couple years? What is a "Newtonian dumping term"?

We were meaning that the initial conditions were calculated using data from EMODnet and those from Lazzari et al. (2016) and Cossarini et al. (2016). After initializing the model with these initial conditions, we let the model spin up for a couple of year, as the Reviewer pointed out. We rephrased the sentence to make this point clearer:

(L. 148-151) "Initial conditions were provided by a two-year spin-up simulation forced by 2015 physical fields in perpetual mode. The two-year spin-up was initialized by profiles of biogeochemical variables as provided by the EMODnet_int climatology, which merges the in situ EMODnet data collections (Buga et al., 2018) and the datasets listed in Lazzari et al. (2016) and Cossarini et al. (2015)."

The Newtonian dumping term is explained in Lazzari et al., 2010, and refers to the treatment of biogeochemical variable concentrations at boundaries. Since it is not relevant to detail this aspect in the manuscript, we modified the sentence avoiding to refer to the Newtonian dumping:

(L.151-153) "In the Atlantic area (i.e., west of the Strait of Gibraltar to the 9° longitude) tracer concentrations were relaxed to climatological seasonally varying profiles. Seasonal profiles of phosphate, nitrate, silicate, and dissolved oxygen were derived from an analysis of the climatological World Ocean Atlas 2018 data (Garcia et al., 2019) and the EMODnet_int dataset."

9) L. 145: Tracer concentrations need to be relaxed to climatology even though the simulation is only 1-year long? Is that typical for the Med Sea?

Any relaxation of tracer concentration is done in the model domain. This refer only on the Atlantic boundary, that has a buffer zone of 2 degrees in longitude (between 9W to 7W) to have stable boundary effects. Thus, tracer concentrations are relaxed in the Atlantic buffer zone only, as the change proposed in the answer 8 clarifies.

10) Comment on the fact that RMSD increases at the surface in TYR when only satellite chlorophyll is assimilated?

Manuscript changes according to your comment 1) added comments also on this aspect (L. 422-431, L. 439-449).

11) L. 281: What is the significance of the 0.3 value?

We decided on this threshold after testing on other values, since it highlights the effect on the vertical profile in Fig. 7. Moreover, using other threshold values would not relevantly modify the shape of profiles of Fig. 7 but only intensify or weaken the differences between REF and ScFcn simulation profiles.

12) Fig 7: Why does the y-axis stop at 300m? There is a large change at 300m in summer, TYR. Caption: "5 left panels" should be "5 right panels". It would be helpful to label figures panels "a", "b", "c", etc. Why are some panels empty?

We decided to limit the profiles to 300 m to better focus on the assimilation impact in the euphotic zone were biogeochemical processes are more rapid and intense. Moreover, differences between REF and ScFcn of Fig. 7 propagate nearly constantly until 500 m and then tend to vanish between 500 and 600 m (Fig. R6). We added a comment on the impact below 300 m in the new version of the manuscript (L. 322-323) but we prefer to keep the limit to 300 m in the manuscript Fig. 7 to highlight the assimilation effects in the euphotic zone.

We corrected "5 left panels" to "5 right panels" and add label to figures.

Some panels are empty, since not all the sub-basins have areas with I(t) higher than 0.3 during winter.

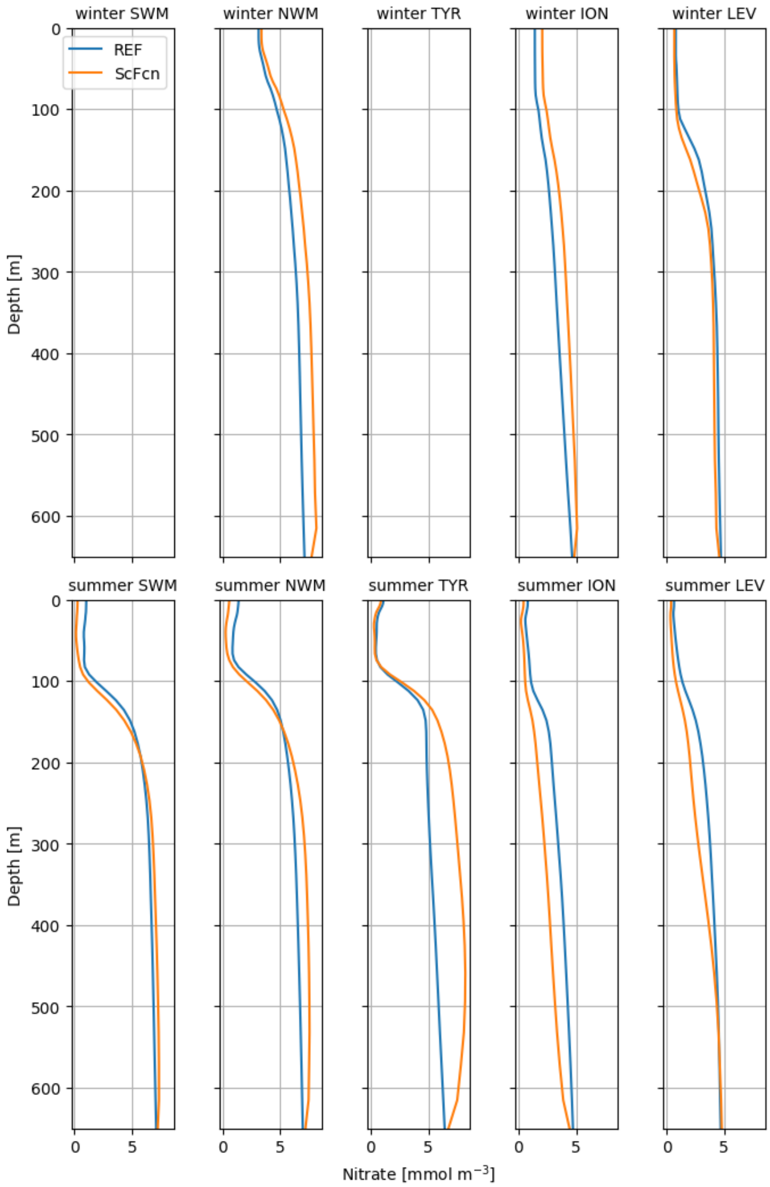

Fig. R6. Mean profiles at time and location with $I_{xy}(t)$ higher than 0.3 in five sub-basins for the ScFcn and REF simulations in winter (top panels) and summer (bottom panels).

13) L. 429-430: The results are consistent with previous studies. Can you elaborate on whether the results are quantitatively different? Or do they simply agree with previously known features of the Med Sea?

Our results on DCM features are quantitatively consistent with previous studies (Lavigne et al., 2013; Lazzari et al., 2012; Mignot et al., 2014) (L. 337-339). Moreover, we discussed differences of nutricline depth with respect to the findings of Barbieux et al. (2019) (L. 493-497. The added values of our results consists mainly in the capability to provide a full 3D validated descriptions of a number of features of the Mediterranean Sea DCM. We clarified these issues in the new version of the manuscript:

(L. 359-360) "The west to east decreasing values of PAR at DCM is documented in a study based on BGC-Argo observation (Mignot et al., 2014), even if the reported values were slightly lower than in Fig. 9."

(L. 477-481) "The results were quantitatively consistent with previous estimations of DCM-depth over the Mediterranean Sea (Lavigne et al., 2013; Lazzari et al., 2012; Mignot et al., 2014) and qualitatively with the results of studies that investigated the variability in the main DCM and nutricline features according to different nutrient and light availability regimes (e.g., Aksnes et al., 2007; Barbieux et al., 2019; Beckmann and Hense, 2007; Cossarini et al., 2019; Cullen, 2015; Gong et al., 2017; Terzić et al., 2019)."

Minor comments:

- L. 84-86: "Original products […] were reviewed for spikes excluding observations whose anomalies were higher than 3 times the daily climatology standard deviation" - The sentence is not clear. Maybe rephrase to "In the original products […] observations whose anomalies were higher than 3 times the daily climatology standard deviation were excluded in order to remove spikes".

We rephrased the sentence according to the Reviewer's suggestion in the new version of quality checks description (L. 88-90).

- L. 122-123: Is it only the optical model this is not included in the present application (but the other 2 are)? Why isn't it included?

Among the improvements listed at L. 122-123, the optical model only is not included in the present application since the development of BGC-Argo assimilation was carried out in a separated branch of the model. However, it is foreseen that the two model branches will converge in the near future. We removed the reference to the optical model from the sentence since it is not applied in this specific work .

- Fig 2: The gray lines hard to see; I suggest removing other grid lines.

We removed the other horizontal grid lines in the figure (as in Fig. R3).

- L. 286 p5th = 95th?

- L. 384-385: "Maintaining the diagonal the observation error covariance matrix" - remove second "the"?

- L. 393: "Thus, while it is desirable an increase of nitrate sensors number" - should be "while it is desirable to increase the number of nitrate sensors"?

- The Verdy and Mazloff reference is missing from the bibliography.

We corrected all the four previous typos as suggested by the Reviewer.

References

Barbieux, M., Uitz, J., Gentili, B., Pasqueron de Fommervault, O., Mignot, A., Poteau, A., Schmechtig, C., Taillandier, V., Leymarie, E., Penkerc'h, C., D'Ortenzio, F., Claustre, H., Bricaud, A., 2019. Bio-optical characterization of subsurface chlorophyll maxima in the Mediterranean Sea from a Biogeochemical-Argo float database. Biogeosciences 16, 1321–1342. https://doi.org/10.5194/bg-16-1321-2019

Buga, L., Sarbu, G., Fryberg, L., Magnus, W., Wesslander, K., Gatti, J., Leroy, D., Iona, S., Larsen, M., Koefoed Rømer, J., Østrem, A.K., Lipizer, M., Giorgietti, A., 2018. EMODnet Chemistry Eutrophication and Acidity aggregated datasets v2018. https://doi.org/10.6092/EC8207EF-ED81-4EE5-BF48-E26FF16BF02E

Butenschön, M., Clark, J., Aldridge, J.N., Allen, J.I., Artioli, Y., Blackford, J., Bruggeman, J., Cazenave, P., Ciavatta, S., Kay, S., Lessin, G., van Leeuwen, S., van der Molen, J., de Mora, L., Polimene, L., Sailley, S., Stephens, N., Torres, R., 2016. ERSEM 15.06: a generic model for marine biogeochemistry and the ecosystem dynamics of the lower trophic levels. Geosci. Model Dev. 9, 1293–1339. https://doi.org/10.5194/gmd-9-1293-2016

Butenschön, M., Zavatarelli, M., Vichi, M., 2012. Sensitivity of a marine coupled physical biogeochemical model to time resolution, integration scheme and time splitting method. Ocean Model. 52–53, 36–53. https://doi.org/10.1016/j.ocemod.2012.04.008

Ciavatta, S., Torres, R., Martinez-Vicente, V., Smyth, T., Dall'Olmo, G., Polimene, L., Allen, J.I., 2014. Assimilation of remotely-sensed optical properties to improve marine biogeochemistry modelling. Prog. Oceanogr. 127, 74–95. https://doi.org/10.1016/j.pocean.2014.06.002

Cossarini, G., Mariotti, L., Feudale, L., Mignot, A., Salon, S., Taillandier, V., Teruzzi, A., D'Ortenzio, F., 2019. Towards operational 3D-Var assimilation of chlorophyll Biogeochemical-Argo float data into a biogeochemical model of the Mediterranean Sea. Ocean Model. 133, 112–128. https://doi.org/10.1016/j.ocemod.2018.11.005

Ford, D.A., 2020. Assessing the role and consistency of satellite observation products in global physical–biogeochemical ocean reanalysis. Ocean Sci. 16, 875–893. https://doi.org/10.5194/os-16-875-2020

Lazzari, P., Solidoro, C., Salon, S., Bolzon, G., 2016. Spatial variability of phosphate and nitrate in the Mediterranean Sea: A modeling approach. Deep Sea Res. Part Oceanogr. Res. 108, 39–52. https://doi.org/10.1016/j.dsr.2015.12.006

Mattern, J.P., Song, H., Edwards, C.A., Moore, A.M., Fiechter, J., 2017. Data assimilation of physical and chlorophyll a observations in the California Current System using two biogeochemical models. Ocean Model. 109, 55–71. https://doi.org/10.1016/j.ocemod.2016.12.002

Nair, A., Sathyendranath, S., Platt, T., Morales, J., Stuart, V., Forget, M.-H., Devred, E., Bouman, H., 2008. Remote sensing of phytoplankton functional types. Remote Sens. Environ., Earth Observations for Marine and Coastal Biodiversity and Ecosystems Special Issue 112, 3366–3375. https://doi.org/10.1016/j.rse.2008.01.021

Santana-Falcón, Y., Brasseur, P., Brankart, J.M., Garnier, F., 2020. Assimilation of chlorophyll data into a stochastic ensemble simulation for the North Atlantic Ocean. Ocean Sci. 16, 1297–1315. https://doi.org/10.5194/os-16-1297-2020

Simon, E., Samuelsen, A., Bertino, L., Mouysset, S., 2015. Experiences in multiyear combined state–parameter estimation with an ecosystem model of the North Atlantic and Arctic Oceans using the Ensemble Kalman Filter. J. Mar. Syst. 152, 1–17. https://doi.org/10.1016/j.jmarsys.2015.07.004

Teruzzi, A., Bolzon, G., Salon, S., Lazzari, P., Solidoro, C., Cossarini, G., 2018. Assimilation of coastal and open sea biogeochemical data to improve phytoplankton simulation in the Mediterranean Sea. Ocean Model. 132, 46–60. https://doi.org/10.1016/j.ocemod.2018.09.007

Teruzzi, A., Dobricic, S., Solidoro, C., Cossarini, G., 2014. A 3-D variational assimilation scheme in coupled transport-biogeochemical models: Forecast of Mediterranean biogeochemical properties. J. Geophys. Res. Oceans 119, 200–217. https://doi.org/10.1002/2013JC009277

Tsiaras, K.P., Hoteit, I., Kalaroni, S., Petihakis, G., Triantafyllou, G., 2017. A hybrid ensemble-OI Kalman filter for efficient data assimilation into a 3-D biogeochemical model of the Mediterranean. Ocean Dyn. 67, 673–690. https://doi.org/10.1007/s10236-017-1050-7

Yu, L., Fennel, K., Bertino, L., Gharamti, M.E., Thompson, K.R., 2018. Insights on multivariate updates of physical and biogeochemical ocean variables using an Ensemble Kalman Filter and an idealized model of upwelling. Ocean Model. 126, 13–28. https://doi.org/10.1016/j.ocemod.2018.04.005