Merging bio-optical data from Biogeochemical-Argo floats and models in marine biogeochemistry

Elena Terzić[1,4], Paolo Lazzari[1], Emanuele Organelli[2], Cosimo Solidoro[1], Stefano Salon[1], Fabrizio D'Ortenzio[2], and Pascal Conan[3]

[1]Istituto Nazionale di Oceanografia e di Geofisica Sperimentale - OGS, Via Beirut 4, 34151 Trieste, Italy
[2]Sorbonne Universités, CNRS, Laboratoire d'Océanographie de Villefranche, LOV, F-06230, Villefranche-sur-Mer,France
[3]Sorbonne Université, Pierre et Marie Curie-Paris 06, CNRS - UMR7621 LOMIC, F66650 Banyuls-sur-Mer, France
[4]Università degli Studi di Trieste, Dipartimento di Matematica e Geoscienze, Via E. Weiss 2, 34128 Trieste, Italy

Correspondence to: Paolo Lazzari (plazzari@inogs.it)

Abstract. In numerical models for marine biogeochemistry, bio-optical data, such as measurements of the light field, may be important descriptors of the dynamics of primary producers and ultimately of oceanic carbon fluxes. However, the paucity of field observations has limited the integration of bio-optical data in such models so far. New autonomous robotic platforms for observing the ocean, i.e., Biogeochemical-Argo floats, have drastically increased the number of vertical profiles of irradiance, photosynthetically available radiation (PAR) and algal chlorophyll concentrations around the globe independently of the season. Such data may be therefore a fruitful resource to improve performances of numerical models for marine biogeochemistry. Here we present a work that integrates into a 1-dimensional model 1314 vertical profiles of PAR acquired by 31 BGC-Argo floats operated in the Mediterranean Sea between 2012 and 2016 to simulate the vertical and temporal variability of algal chlorophyll concentrations. In addition to PAR as input, alternative light and vertical mixing models were considered. We evaluated the models' skill to reproduce the spatial and temporal variability of deep chlorophyll maxima as observed by BGC-Argo floats. The assumptions used to set up the 1-D model are validated by the high number of co-located in-situ measurements. Our results illustrate the key role of PAR and vertical mixing in shaping the vertical dynamics of primary produces in the Mediterranean Sea. Moreover, we demonstrate the importance of modeling the diel cycle to simulate chlorophyll concentrations in stratified waters at the surface.

1 Introduction

In most biogeochemical models, which are successfully coupled with hydrodynamics, the description of optics is generally (over)simplified, therefore one of the necessary improvements still remains the integration of a more complex optical model, where inherent and apparent optical properties (IOPs and AOPs respectively) are already included as model state variables, (Fujii et al., 2007). The research community is emphasizing the importance of merging different methods in order to improve the skill of numerical models, such as the assimilation of remote sensing data or the use of in-situ data both for initialization and validation purposes. Until recently, the use of the latter was especially critical due the scarcity of observations, however

the emergence of autonomous robotic platforms such Biogeochemical Argo floats (hereafter BGC-Argo) helped filling the gap in bio-optical measurements acquired around the globe regardless of the season.

In particular, the Mediterranean Sea Monitoring and Forecasting Centre (Med-MFC) operatively produces analyses, forecasts and reanalyses of a series of biogeochemical state variables (e.g. chlorophyll, nutrients, pCO2) for the European Copernicus Marine Environment Monitoring Services (CMEMS) since 2015 using the MedBFM model (Lazzari et al., 2010, 2012, 2016), which assimilates surface chlorophyll from satellite observations (Teruzzi et al., 2014, 2018).

The introduction of BGC-Argo floats has led to a drastic increase in the number of radiometric measurements in the Mediterranean Sea, such as downward irradiance (Ed) and photosynthetically available radiation (PAR), for which specifically developed quality control procedures and refined sensor calibration (Organelli et al., 2016, 2017a) have widespread their use (Organelli et al., 2017b; Wojtasiewicz et al., 2018; Gerbi et al., 2016; Leymarie et al., 2018). BGC-Argo can therefore be an important source of high vertical spatial and temporal resolution data that can be integrated in the calibration and tuning of bio-optical numerical models for understanding marine biogeochemistry.

To this end the Mediterranean Sea proves to be an important region to study due to its bio-optically anomalous nature. It is characterized by complex trophic gradients (Crise et al., 1999; Lazzari et al., 2012; d'Ortenzio and Ribera d'Alcalà, 2009) and spatially heterogeneous inherent optical properties (Oubelkheir et al., 2005). Such gradients are mainly related to the inverse estuarine circulation of the area (Crispi et al., 2001) and to the varying distribution of optically significant substances (e.g. colored dissolved organic matter - CDOM; non-algal particles - NAP) that modulate the light penetration along the water column (Morel and Gentili, 2009b). Moreover, inherent optical properties (IOPs) could be affected also by important processes of Saharan dust deposition (Claustre et al., 2002).

At present, no studies have tried to assimilate radiometric quantities into numerical models to improve the simulation of chlorophyll dynamics in this basin and investigate the causes of the vertical, spatial and temporal variability eastward. Assimilating radiometry could prove more robust than chlorophyll assimilation as a result of a more accurate uncertainty characterization of optical measurements (Dowd et al., 2014; Organelli et al., 2016) compared to other biogeochemical variables, such as fluorescence-derived chlorophyll.

Specific studies are required to demonstrate to what extent the assimilation of radiometric data can improve the model skill in simulating key biogeochemical variables (e.g. chlorophyll, nutrients, primary productivity).

In this paper we develop a 1-dimensional model that assimilates PAR profiles acquired by BGC-Argo floats in order to replicate the vertical and temporal dynamics of phytoplankton chlorophyll concentrations. We analyse and validate model performances through a comparison of model outputs with the high number of co-located vertical profiles of chlorophyll concentrations (derived from fluorescence) measured by BGC-Argo floats. Subsequently, we test different mixing and bio-optical models that simulate downward irradiance and evaluate their skills in order to estimate how well they perform compared to in-situ measurements of PAR. The paper is organized as follows: in the Methods section, the Mediterranean Sea BGC-Argo floats network and the model configurations are presented. In the Results and Discussion section, we analyse the 1D biogeochemical simulations and their sensitivity according to the objectives of the work. General remarks are illustrated in the Conclusions section.

2 Methods

2.1 BGC-Argo floats data

The Mediterranean Sea BGC-Argo array operating in the period 2012-2016 was composed of 31 floats that acquired 1314 vertical profiles, Figure 1, of temperature (T) and salinity (S), chlorophyll *a* concentration (Chl, units of $\mathrm{mg\,m^{-3}}$), derived from fluorescence measurements between 0 and 1000 m (see Organelli et al., 2017b; Roesler et al., 2017), and radiometric quantities, such as downward planar irradiance (E_d) at three different wavelengths (λ = 380, 412 and 490 nm, units of $\mu\,\mathrm{W\,cm^{-2}\,nm^{-1}}$) and Photosynthetically Available Radiation (PAR, unit of $\mu\,\mathrm{mol\,quanta\,m^{-2}\,s^{-1}}$) integrated between 400 and 700 nm (Kirk, 1994). Radiometric measurements were obtained in the upper 250 m, with vertical resolution of 1 m between 10 and 250 m and 0.20 m between 0 and 10 m. All profiles were acquired around local noon.

The quality control (QC) procedure of radiometric profiles was specifically designed to identify and remove the dark signal, atmospheric clouds and wave focusing at the surface (Organelli et al., 2016). Note that the operational definition of PAR used by the BGC-Argo community takes into consideration the planar irradiance E_d rather than the scalar one E_o, therefore differing from its theoretical definition and leading to an underestimation of its values by 30% or more (Mobley et al., 2010). The scalar values of PAR were thus derived according to Baird et al. (2016), although the correction related to the irradiance scattering was neglected due to the lack of information on IOPs (see section 1 of supplementary materials).

Vertical profiles of chlorophyll concentration were quality-controlled according to the procedure of the international BGC-Argo program that removes spikes and corrects for non-zero deep values and non-photochemical quenching at the surface (Schmechtig et al., 2016; Organelli et al., 2017b). Due to a factory calibration bias for WETLABS ECO series Chl fluorometers, Chl a concentrations were corrected by a factor of 0.5 (Roesler et al., 2017; Organelli et al., 2017a, b; Barbieux et al., 2018).

All the data used in this study are freely available and compiled into the database published by Organelli et al. (2017b). To proceed with our study, 7 variables (T, S, Chl, Ed380, Ed412, Ed490, PAR) were vertically interpolated to a resolution of 1 m in the upper 400 m. Finally, we partitioned the profiles geographically into 13 (out of 16) subbasins (Fig1), with the majority of profiles located in the North Western Mediterranean (NWM, 332 profiles), followed by Northern Ionian (ION3, 172 profiles) and Southern Tyrrhenian (TYR2, 162 profiles). No data were available for the South-western Ionian (ION1) and the Eastern Levantine (LEV4) and only one profile was present in the Northern Adriatic (ADR1), as well as in the Western Levantine (LEV1). The WMO code specification for each BGC Argo float is provided in the section 2 of supplementary material.
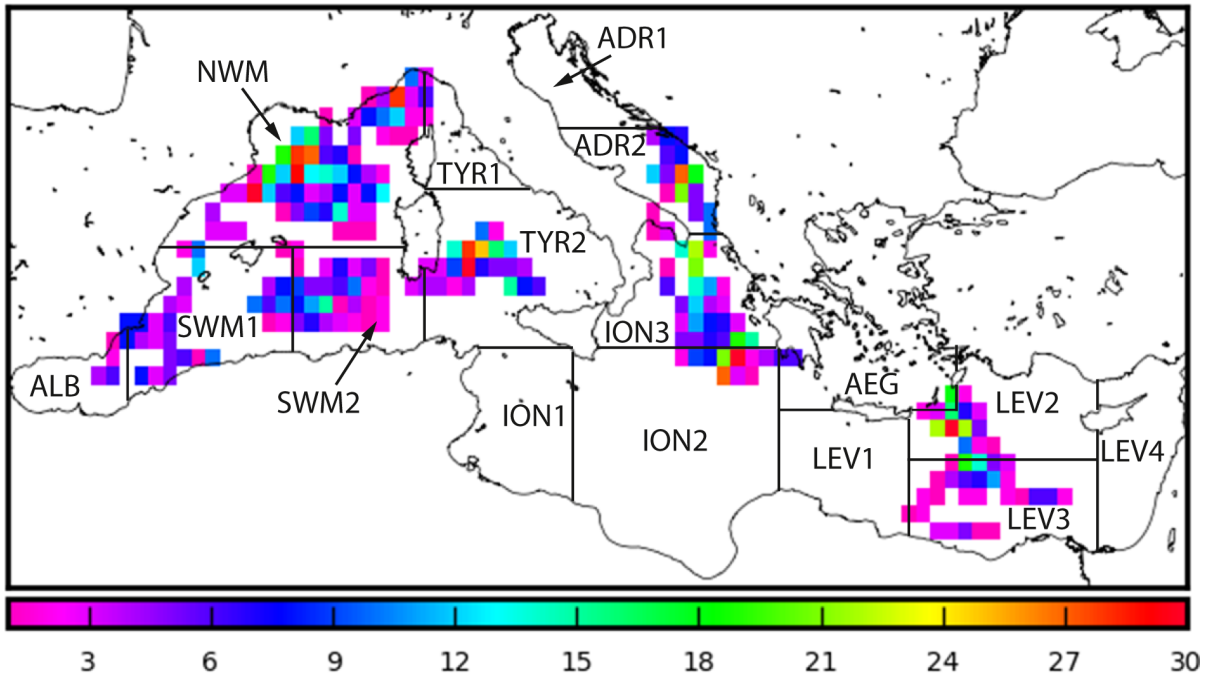
Figure 1. Spatial distribution of BGC-Argo float profiles superimposed to subbasin division used in the Mediterranean Copernicus Marine Environment Monitoring Service (CMEMS) system

2.2 1-D Biogeochemical Model

Biogeochemical processes have been simulated according to the voxel approach ("volume element with biological content and processes", Kohlmeier and Ebenhöh (2009)), discretized along the vertical in order to resolve vertical irradiance attenuation and nutrient gradients. Each voxel replicated light and mixing conditions according to the trajectory and measurements of the corresponding BGC-Argo float, thus simulating a pseudo-lagrangian experiment. No exchanges of mass between voxel and the surrounding field have been considered, which implies smaller mass exchanges due to horizontal diffusion and baroclininc components of the (upper ocean) advection field compared to vertical processes and biogeochemical dynamics. Conversely, voxel exchanges heat with the atmosphere and receives light in accordance with its moving position. Such an approach, similar to the one adopted by Kohlmeier and Ebenhöh (2009), has been already successfully applied by Mignot et al. (2018) in order to analyse BGC-Argo Floats in the North Atlantic.

Furthermore, it was assumed that major biogeochemical transformations can be described by the Biogeochemical Flux Model (BFM) parametrizations (see below), properly driven by a bio-optical model, which has been validated by contrasting model results and experimental data, as shown later. The model is formulated through a system of partial differential equations:

$$\partial_t C_i(z,t) = \partial_z[D_v(z,t)\partial_z C_i(z,t)] + v_{sink,i}\partial_z C_i(z,t) + BFM_i(T,S,PAR,\overline{C}(z,t)) \tag{1}$$

where C_i is the i-th biogeochemical tracer simulated (i=1,50), D_v is the vertical eddy diffusivity derived with the vertical mixing model described in subsection 2.2.1, v_{sink} is the sinking velocity and BFM_i is the reaction term corresponding to the tracer C_i. T, S, PAR are the data measured by BGC-Argo floats.

In the first set of simulations, the biogeochemical model was forced with PAR from BGC-Argo floats. Experimental values of temperature and density (computed from float profiles) were also taken into consideration.

A simulation for each of the BGC-Argo float trajectories was performed with this set-up, hereafter abbreviated as REF.

Four additional sets of simulations were performed on the REF configuration by applying different values of vertical eddy diffusivity coefficients (MLD1, MLD2, MLD3 and MLD4) in order to assess uncertainties due to different vertical diffusion parametrization.

Six additional sets of simulations were performed by forcing the biogeochemical model with PAR obtained by alternative bio-optical parametrizations (OPT1, OPT2a,b,c,d), one of which considering also the current modeling approach in the CMEMS Copernicus system (OPT3). In this way, the possibility of using biogeochemical models in the absence of PAR measurements was assessed.

Finally, a set of simulations was devoted to understand the impact of using a constant light approximation rather than following the diurnal light variation (CL1 and CL2 configurations) on chlorophyll distribution. Furthermore, we evaluated the impact on light propagation due to coloured phytoplankton degradation products, i.e., CDOM, (OPT4a,b,c,d and OPT5). We therefore tested a total of 17 classes of simulations that are summarized in Tab.1 and Tab.2.

The biogeochemical model BFM (Vichi et al., 2013) is a biomass-based numerical model that simulates the biogeochemical fluxes of carbon, phosphorus, nitrogen, silicon, and oxygen, characterizing the lower trophic level (producers, consumers, and recyclers) of the marine ecosystem. Its application is based on the coupled transport-biogeochemical model OGSTM-BFM (Lazzari et al., 2012, 2016). It includes four phytoplankton functional types (diatoms, nanoflagellates, picophytoplankton, and dinoflagellates), carnivorous and omnivorous mesozooplankton, bacteria, heterotrophic nanoflagellates, and microzooplankton. Each variable is described in terms of internal carbon, phosphorus and nitrogen concentrations. Phytoplankton functional types can be characterized regarding prognostic Chl and can additionally consider the silicate component for diatoms. Particulate and dissolved organic matter are also included, with the latter partitioned in labile, semi-labile and semi-refractory phases. The present study is focused mainly on Chl, reserving to future analysis (according to data availability and optical model complexity) a study of Plankton Functional Types (PFT) resource competition dynamics (Ryabov and Blasius, 2011, 2014). Initial conditions for all biogeochemical variables of BFM are provided by the CMEMS reanalysis of Mediterranean Sea biogeochemistry (period 1999-2015, Teruzzi et al., 2014) produced by the MedBFM model system. The initialization profiles are extracted from the MedBFM model output array, taking the nearest model point to the BGC-Argo position in time and space.

Simulations' time scale corresponds to a typical BGC-Argo time-series length during the period 2012-2016, i.e. 11 months on average, with a vertical resolution of 1m. After being initialized, the model evolves without further assimilation of biogeo-chemical data from the 3D configuration.

Table 1. Model configurations considered in the present work. All simulations include diurnal variability except the two cases with continuous light (CL1 and CL2), which use 24-hour averaged irradiance values.

SIM	MODEL DESCRIPTION
REF	PAR from BGC-Argo floats ; $D_v^{background} = 10^{-4} m^2 s^{-1}$
CL1	as REF with continuous daily light
CL2	as REF with continuous daily light and $D_v^{background} = 10^{-6} m^2 s^{-1}$
MLD1	as REF with $D_v^{background} = 5\,10^{-5} m^2 s^{-1}$
MLD2	as REF with $D_v^{background} = 10^{-5} m^2 s^{-1}$
MLD3	as REF with $D_v^{background} = 5\,10^{-6} m^2 s^{-1}$
MLD4	as REF with $D_v^{background} = 10^{-6} m^2 s^{-1}$
OPT1	Riley: $K_d(PAR) = 0.04 + 0.054\,Chl^{\frac{2}{3}} + 0.0088\,Chl$
OPT2a	
OPT2b	$K_d(PAR) = a\,Chl^b + c$
OPT2c	
OPT2d	
OPT3	$K_d(PAR)$ for the first optical depth $z_{od} = z_{eu}/4.6$
OPT4a	as OPT2a + Chl degradation to CDOM - time scale 1 day
OPT4b	as OPT2a + Chl degradation to CDOM - time scale 1 week
OPT4c	as OPT2a + Chl degradation to CDOM - time scale 1 month
OPT5	as OPT2a + CDOM following Dutkiewicz et al. (2015)

Table 2. Parameters derived for optical models using BGC Argo float data. For each version of OPT2 only data shallower than z_{max} were used to compute the regression.

Model	z_{max}	R^2	a	b	c
OPT2a	150	0.53	0.075 ± 0.0015	0.572 ± 0.018	0.027 ± 0.001
OPT2b	75	0.61	0.064 ± 0.0015	0.615 ± 0.021	0.040 ± 0.002
OPT2c	45	0.71	0.077 ± 0.002	0.469 ± 0.021	0.034 ± 0.002
OPT2d	30	0.75	0.088 ± 0.003	0.406 ± 0.023	0.029 ± 0.003

Vertical eddy diffusivity coefficient profiles $D_v(z)$ are here represented as Gaussian-shaped functions, using potential density values for the mixed layer depth (MLD) calculation with a density-based criterion (de Boyer Montégut et al., 2004; D'Ortenzio and Prieur, 2012). Such a shape is chosen due to its simplicity and in order to allow a gradual increase of vertical mixing through the pycnocline. Approaches and impacts of using different parametrizations to reconstruct mixing along the water column are shown and discussed in section 2.2.1. Since the surfacing of BGC-Argo floats is programmed at around local noon, the variability related to diurnal variation of solar irradiance is taken into consideration according to Kirk (1994).

2.2.1 Vertical Mixing Models

Unlike radiometric data, vertical mixing is an indirectly obtained quantity, described in terms of potential density (from temperature and salinity data) along the water column. Vertical eddy diffusivity coefficients (D_v) are defined as Gaussian-shaped functions in the form of:

$$D_v = D_v^{MLD} e^{-0.5\left(\frac{z}{(\sigma * MLD)}\right)^2} + D_v^{background} \tag{2}$$

σ was identified after an initial tuning procedure and equals 0.3 in all simulations. Values in REF model are equal to D_v^{MLD} = $1.0 m^{2s-1}$ and $D_v^{background}$ = $10^{-4} m^2 s^{-1}$.

The mixed layer depth (MLD) was defined with the density criterion at the threshold value (de Boyer Montégut et al., 2004; D'Ortenzio and Prieur, 2012):

$$\Delta\rho_\theta = |\rho_\theta(10m) - \rho_\theta(z)| = 0.03 kg m^{-3} \tag{3}$$

In simulations MLD1, MLD2, MLD3, and MLD4, $D_v^{background}$ values were perturbed for two orders of magnitude (from 10^{-6} to 10^{-4} m^2s^{-1}) in order to estimate the impact such variations have on modeled Chl profile shapes compared to measured ones (see Table 1).

2.2.2 Bio-Optical Models

Alternative parametrizations to measured PAR profiles were used in models OPT1, OPT2abcd, OPT3, OPT4abc and OPT5. They differ in methods used to evaluate the Beer-Lambert attenuation coefficient K_d(PAR), which is parametrized as a function of Chl concentration rather than being directly calculated from BCG-Argo irradiance data (see Tab.1 and Tab.2).

OPT1 uses the relationship obtained by a statistical analysis done by Riley (1956, 1975):

$$K_d(PAR) = 0.04 + 0.0088 [Chl] + 0.054 [Chl]^{\frac{2}{3}} \tag{4}$$

In OPT2 models, statistical regressions were carried out between $K_d(PAR)$ and Chl measured by BGC-Argo floats at four different depth ranges: 150 m, 75 m, 45 m and 30 m (OPT2a to OPT2d, see Table 2 for details):

$$K_d(PAR) = a [Chl]^b + c \tag{5}$$

a and c represent regression coefficients and b the exponent (values reported in Table 2. Confidence intervals were calculated with Student's two-sided t-test, where the significance level α was set equal to 0.05). Diffuse attenuation coefficients $K_d(PAR)$ were calculated for PAR measured by BGC-Argo floats as the local slope of the natural logarithm of downwelling irradiance for layers of 15 m thickness for the euphotic depth range, which corresponds to an attenuation of downward planar irradiance to 1% of the subsurface value (Kirk, 1994).

Albeit the regression based on the upper 30 m depth range measurements showed highest correlation, all four bio-optical models were considered and adopted in simulations OPT2a,b,c and d (Tab.2).

In model OPT3, based on the BGC-Argo data set, $K_d(PAR)$ is calculated for the first optical depth (Morel, 1988), the layer of interest for satellite remote sensing (Gordon and McCluney, 1975), and then adopted as a constant parameter for the entire water column. Such kind of light extinction definition has been used also in the 3-dimensional version of the OGSTM-BFM model, which integrates $K_d(490)$ data from satellite sensors as the external optical forcing in the exponential formulation of
5 downwelling irradiance (for more details see Lazzari et al., 2012, section 2.2.3).

OPT4 and OPT5 models include CDOM dynamics as in the Mediterranean Sea the latter can absorb more than 50% of blue light (Organelli et al., 2014; Morel and Gentili, 2009a), thus significantly impacting its attenuation along the water column. OPT4 assumes that CDOM is correlated to chlorophyll production (Organelli et al., 2014) and that light attenuation is therefore affected by a progressive accumulation of such a constituent ("dead" chlorophyll, initialized at zero concentration). In OPT4,
10 accumulation is compensated by decay (first order kinetic) set at different e-folding characteristic times: 1 day (OPT4a), 1 week (OPT4b) and 1 month (OPT4c).

OPT5 implemented a formulation of CDOM as described in Dutkiewicz et al. (2015): a 2% fraction of all dissolved organic matter (DOM) fluxes is directed to CDOM, including both temperature-related decay and a photodegradation term based on PAR (Bissett et al., 1999). Additional investigations are provided in section 3.3 to discuss CDOM dynamics along the water
15 column. Given the mono-spectral nature of the current description of light, the attenuation of CDOM on PAR is computed by averaging the exponential law of CDOM absorption (Bricaud et al., 1981) on the visible range.

2.3 Statistical Analysis

According to the work's objectives, four classes of simulations were considered, which correspond to the following subsections: the reference simulation, a subset with perturbed vertical mixing models, tests with different optical configurations, and a
20 last group of additional analyses involving CDOM description and diurnal variability. Outputs are validated qualitatively and quantitatively in terms of profile shapes and the deep chlorophyll maximum (DCM) depth. The DCM definition is based on the absolute maximum of Chl, excluding results of DCM shallower than 40 m or deeper than 200 m, as well as the ones with concentrations lower than $0.1~\mathrm{mgm^{-3}}$. All results, both for model and BGC-Argo floats, are averaged on a weekly basis. Model outputs are compared by means of match-up, Target and Taylor diagrams (Jolliff et al., 2009). In addition to the DCM
25 depth, the performance in reproducing the DCM thickness and Chl concentration in the DCM layer were also analysed. DCM thickness is operationally defined through a Gaussian fit as $\pm\sigma/2$ from the maximum, thus the Chl concentration at DCM is averaged over the DCM thickness. For a couple of simulations (REF and CL1), skills are also compared at the surface layer (0 - 25 m). In order to avoid corrections due to non-photochemical quenching, profiles acquired only during stratified periods were considered. The target diagram evaluates results with root mean square distance (RMSD) as the main statistical parameter,
30 which was calculated following equation 6:

$$RMSD = \sqrt{\frac{1}{n}\Sigma_{i=1}^{n}\Big(m_i - o_i\Big)^2} \tag{6}$$

where n is the number of data, m are the model data and o are the observables.

3 Results and Discussion

3.1 Reference Simulation

The assimilation of PAR profiles into the 1-D model helped to accurately estimate the deep chlorophyll maximum depth (Figure 2). The overall model skill in the REF configuration is shown in Figure 2, with the histogram within indicating a normal
5 distribution of residuals' deviation. Measured and modelled DCM depth showed high correlation (r=0.8, p-value <0.0005). Both model and measurements indicate that DCM depth varies typically between 50-70 m in western areas (ALB, SWM1, SWM2, NWM, TYR) and is generally deeper in eastern areas (ADR2, ION3, LEV2, LEV3), between 100-140 m. Model tends to slightly underestimate the DCM depth variability (Figure 2, regression slope = 0.81 < 1), in fact, deepest simulated DCM are around 125 m depth, whilst floats data reach 140 m (e.g. as shown in Figure 2 lovbio018c data).
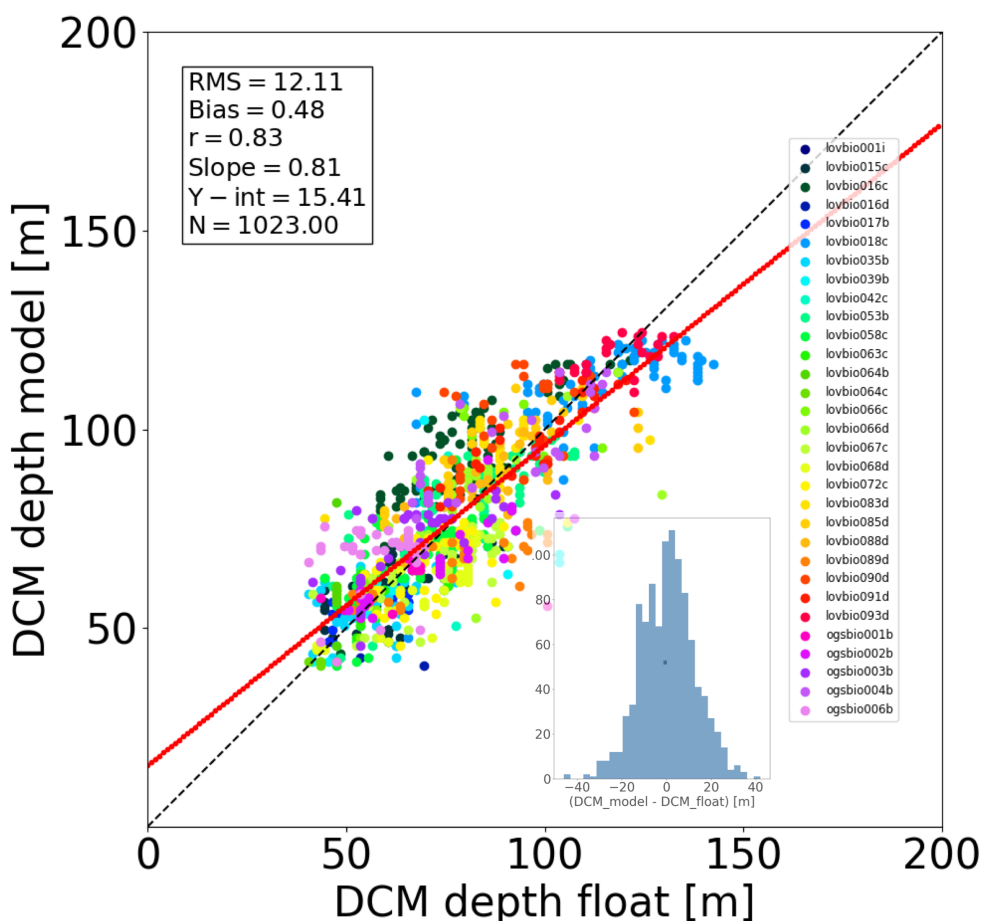
Figure 2. Match-up diagram comparing DCM depth obtained from BGC-Argo floats data versus REF model results. Each dot corresponds to a weekly profile. The red line depicts the linear regression between data and model values, defined by its slope and intercept (Y-int) shown in the box. Units of RMSD, Bias and Y-int are in meters. The correlation coefficient r is significant, with p-value < 0.005. The bottom sub-figure shows the residuals' histogram.

Chl patterns display high variability both at temporal and vertical scales, shown in Figures 3 to 6. The subsurface Chl pattern is formed by patchy structures and during stratification periods it is generally deeper moving eastward. BGC-Argo observations indicate that DCM is further eroded by vertical mixing occurring generally in autumn and early winter. Simulations provide an adequate reproduction of the Chl mixing timing and therefore the DCM erosion. By comparing point-to-point all Hovmöller

5 maps (considering both depth and time variability) for measured and simulated Chl (examples are reported in Figures 3, 4, 5 and 6), a significant average correlation of 0.75 is obtained: such a result quantitatively confirms that the alternation of mixing and stratification phases, as seen from BGC-Argo chlorophyll measurements, is well reproduced. At surface, the increase in Chl is triggered by rather shallow mixing (0-75 m layer). The initial condition statistics in reproducing DCM depth (R=0.62, slope of 0.53) is improved by the 1-D model for BGC-Argo data.
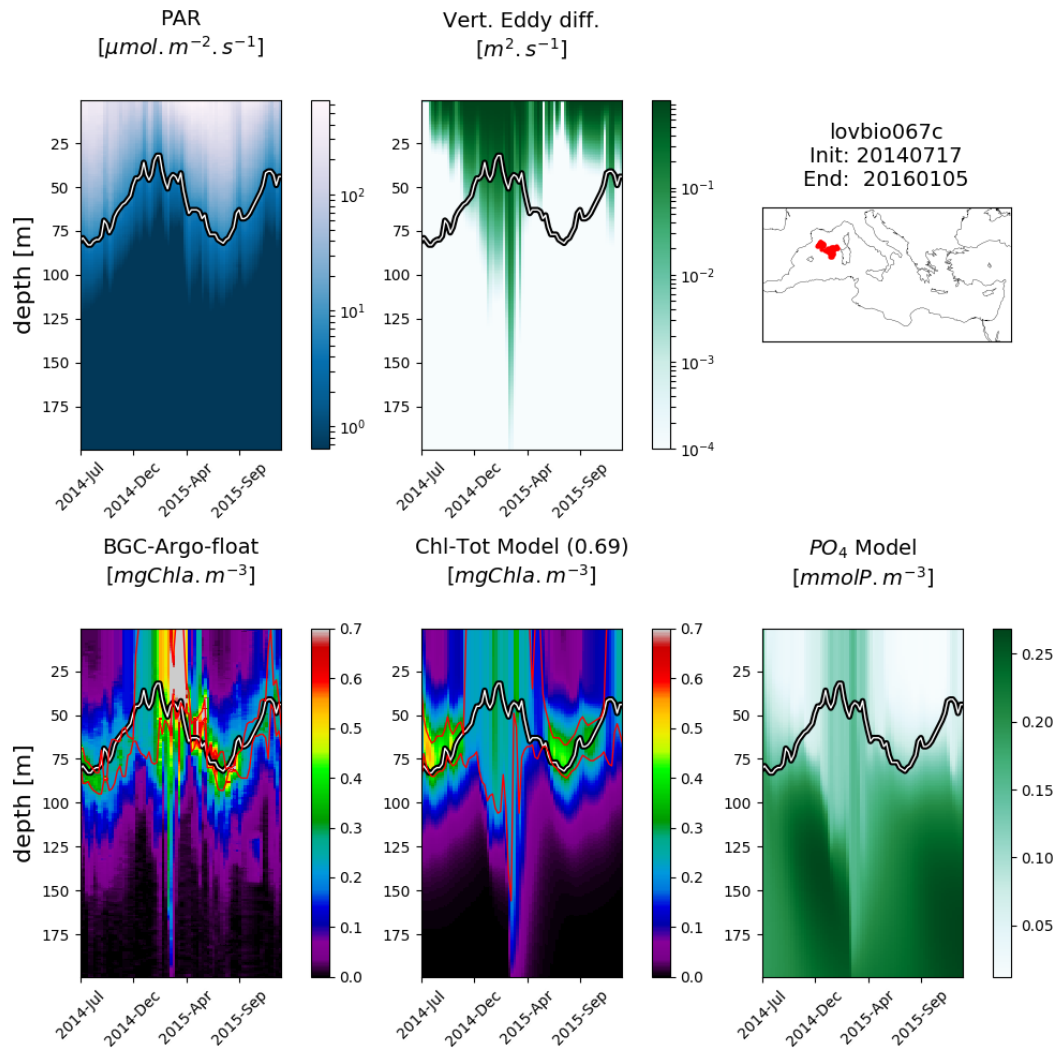
Figure 3. Hovmöller diagrams of BGC-Argo float lovbio067c (WMO code 6901649) comparing measured results and simulated ones (REF). The 6-imaged composite is organized as follows: top row shows PAR, vertical eddy diffusivity and the float trajectory; bottom row shows Chl derived from fluorescence measurements, simulated Chl and phosphate. The thick black-white line indicates the depth where PAR equals $0.5\ \mathrm{mol\,quanta\,m^{-2}\,day^{-1}}$ (Mignot et al., 2014). The number in parentheses in modelled Chl indicates point-by-point correlation with BGC-Argo float Chl.
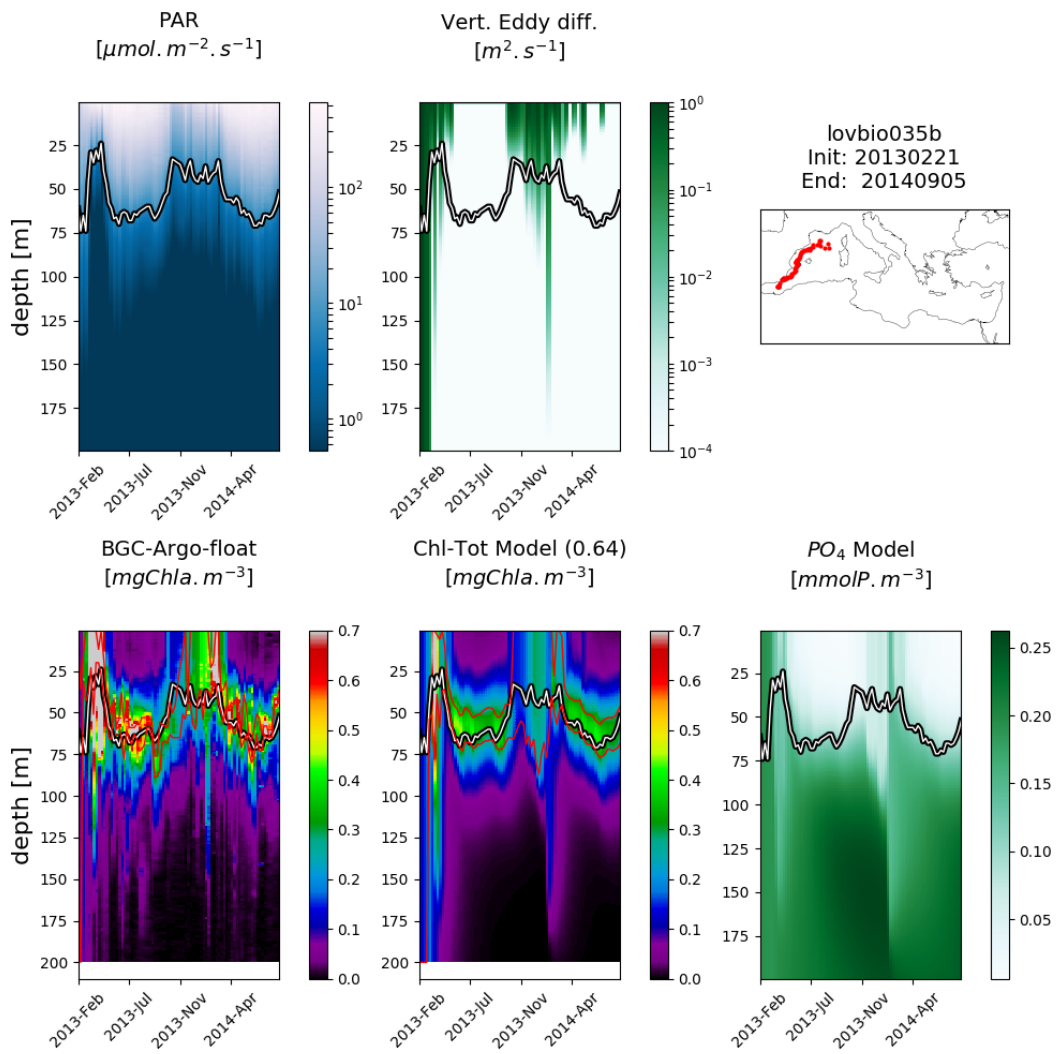
Figure 4. As Figure 3 but for the BGC-Argo float lovbio035b (WMO code 6901511).
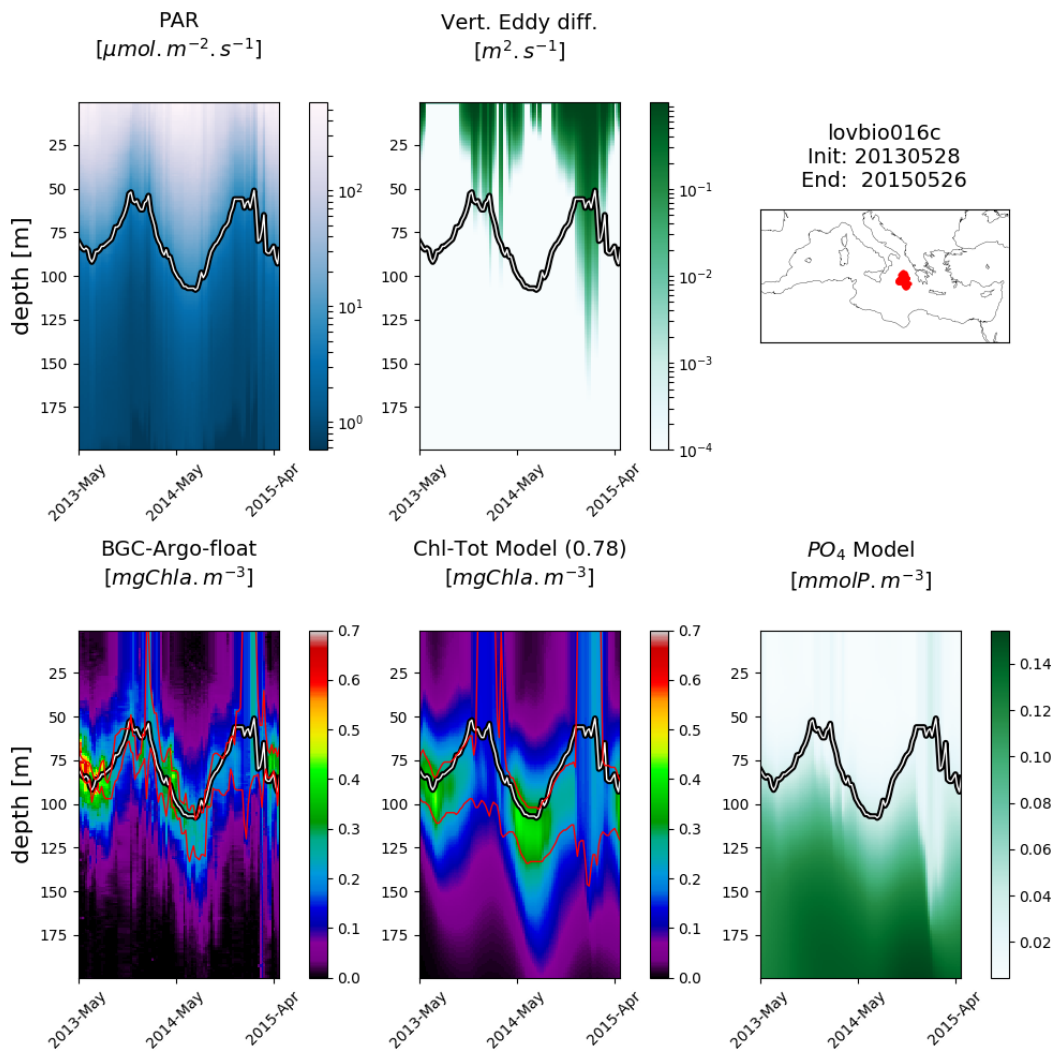
Figure 5. As Figure 3 but for BGC-Argo float lovbio016c (WMO code 6901510).
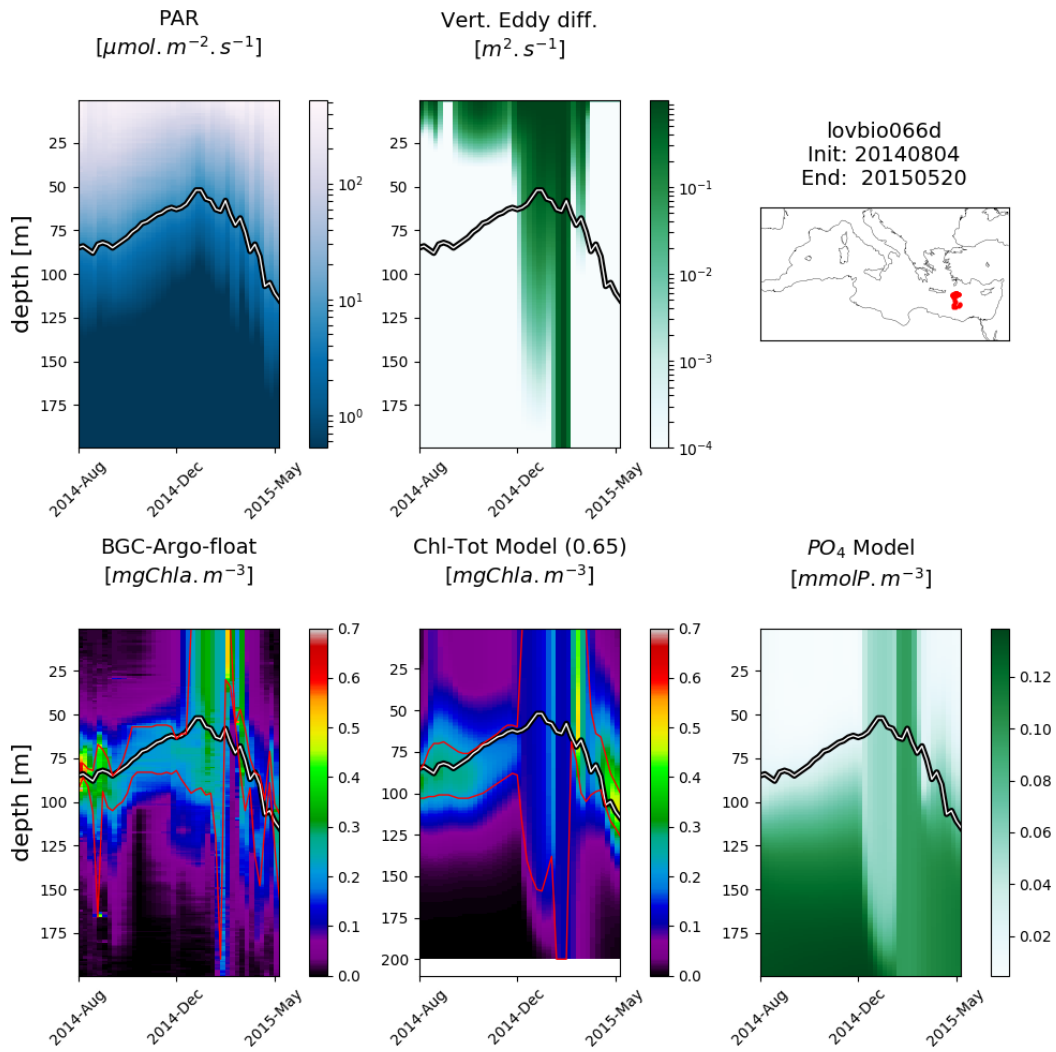
Figure 6. As Figure 3 but for BGC-Argo float lovbio066d (WMO code 6901655).

In addition to a correctly reproduced timing in the alternation of mixing and stratification phases, proved by high correlation, simulated Chl reproduces also episodic signals, such as Chl deepening due to specific mixing events. For example, a mixing event in the NWM subbasin, reaching approximately 200 m depth during winter in 2015, triggers an intrusion of Chl (0.2 $mg\,m^{-3}$) in deeper layers consistently to BGC-Argo float measurements (float lovbio067c, Figure 3). Similar dynamics is reproduced in winter 2014 (Figure 4) for the lovbio035b float drifting from NWM toward the ALB subbasin.

Considering float trajectories, two kinds of situations are possible: the BGC-Argo float trajectory is relatively stationary in the deployment area (as shown in Figures 3, 5 and 6), or the float passively migrates extensively, following a given water mass (as in Figure 4). It appears that also in the second case, when lateral dynamics effects could play an important role in

BGC-Argo float measurements, the approach applied allows an adequate representation of measured Chl patterns. However, it should be noted that in the present multi-float simulation there are no trajectories including both west and east Mediterranean basins. In such cases, strong gradients between deep water nutrient inventories could invalidate the approach, thus nudging or more sophisticated techniques would be required (Kohlmeier and Ebenhöh, 2009). Lateral advection processes could indeed

5 play an important role, although it appears that in the present case considering data-driven mixing and turbulence effects allow to simulate correctly the seasonal variability. The REF simulation can be therefore used as a reference for the following tests on mixing and bio-optical models analysed in the next sections. Furthermore, REF results demonstrate that irradiance along the water column, besides mixing, is the driving mechanism controlling DCM depth. Figure 7a shows a significant correlation between DCM and euphotic depths (i.e. where irradiance reaches 1% of surface PAR), both in cases of measured and simulated

10 Chl.
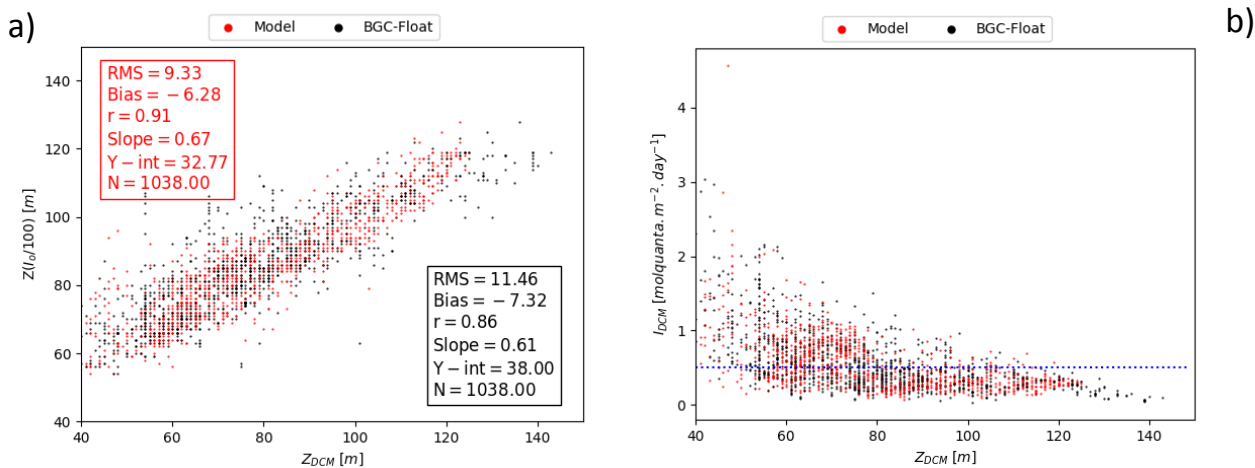
Figure 7. Panel a) DCM depth (z_{DCM}, x-axis) compared to the euphotic depth (z_{eu}, y-axis) both for modelled (red dot) and measured results (black dot). Red box (top left) reports statics for model z_{DCM} versus z_{eu}, whereas the black box (bottom right) shows statistics for z_{DCM} derived from chlorophyll data versus z_{eu}. Panel b) irradiance values (y-axis) at DCM depth (x-axis) both for modelled (red dot) and measured results (black dot). Horizontal blue line marks the 0.5 irradiance threshold (units $\mathrm{mol\,quanta\,m^{-2}\,day^{-1}}$) as identified in Mignot et al. (2014).

Similar results, valid on annual average conditions, were found by Mignot et al. (2014) in their Eq. 9, where euphotic depth results to be 0.73% rather than 1% of surface PAR, interpreting that the DCM is located at a fixed PAR value, oscillating near the $0.5\ \mathrm{mol\,quanta\,m^{-2}\,day^{-1}}$ isolume. Comparable conclusions can be derived for analyses presented hereby. Data and model outputs show similar results with a higher variability of critical PAR values in the case of shallower DCM, Figure 7.

15 However, the model-predicted DCM seems strongly constrained by light regime, Figure 12, whilst observed DCM fluctuates up and down over the euphotic depth (see transition from "T=29 weeks" to "T=31 weeks").

In order to further evaluate the dependence of model results on PAR forcing, two numerical experiments were carried out. In the first experiment, BGC-Argo floats were divided in couples composed by one trajectory located in the western basin and the other one in the eastern basin, by random selection. For each couple the initial conditions for nutrients were interchanged, which allows to estimate their impact on DCM depth. Results (see supplementary material, section 4) evidence that the inverted

5   initialization of nutrients does not significantly alter the outcome in terms of DCM depth, resulting in a slope reduction from 0.81 to 0.62, and maintaining similar correlation and bias. Thus it appears that the role of nutrients is secondary compared to the impact of light on DCM depth regulation on such time scales.

Performing the same operation by switching light data instead of nutrients proves to be technically more complex, thus an alternative approach was applied, which consists of a sensitivity analysis similar to the one described in Huisman et al.

10   (2004). For this purpose, two BGC-Argo floats (lovbio018c and lovbio067c for east and west respectively) and phosphate and PAR parameters were selected, constructing an array of 21x21=441 simulations (per float) for bivariate perturbations. Such technique allows to further understand the driving mechanisms of DCM depth variability.

A perturbation of 50% of the of initial PO_4 condition has only a minor effect on DCM depth position, Figure 8, while changes in light conditions show a large effect (approximately 10 meters difference).
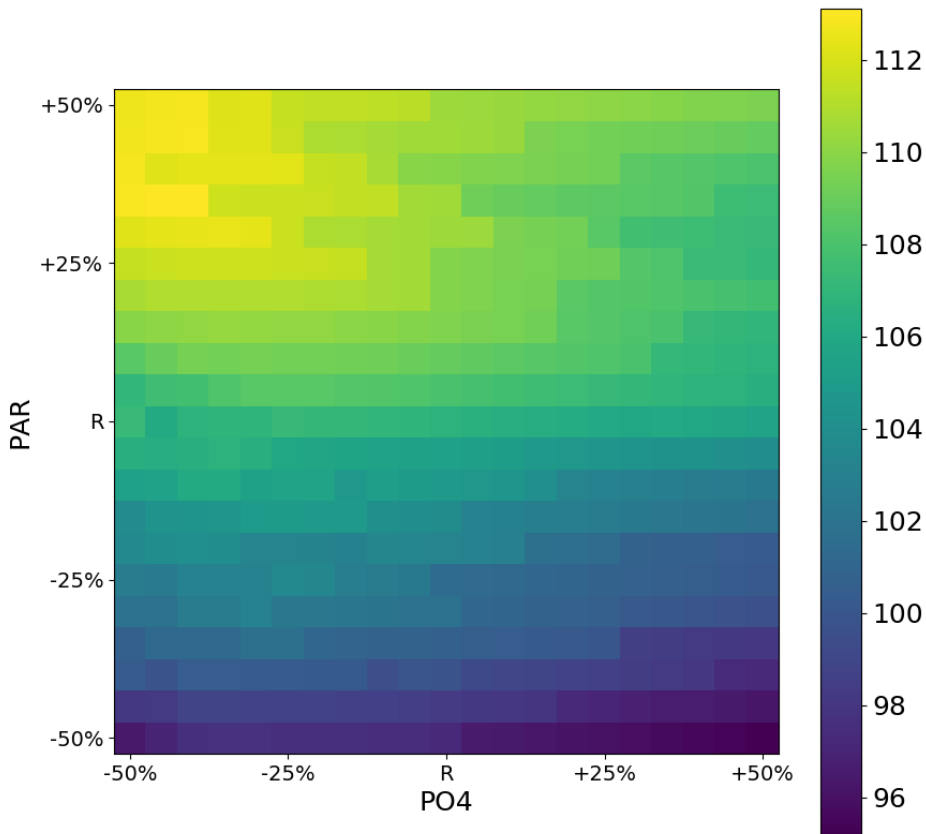
Figure 8. Sensitivity analysis of DCM depth perturbing light and initial conditions of PO4 (both by an uniform factor reported on axis in percentage) along the water column. 'R' marks the reference values. The BGC-Argo float here reported is the lovbio018c. Each pixel is a full simulation of a total of 21x21 simulations. The DCM depth is averaged over the simulation period.

The same sensitivity analysis is used to evaluate model performance in reproducing DCM width and chlorophyll at DCM (plots for both lovbio018c and lovbio067c are included in the supplementary material, section 5). Results indicate that the DCM width has a variability of 6 m in the perturbation range (±50%), as well as that the DCM magnitude is controlled by nutrient availability rather than light. Comparing measured chlorophyll concentrations and model results shows that the skill in reproducing the DCM thickness is lower compared to the DCM depth (Figure 9, r=0.55, slope=0.6).
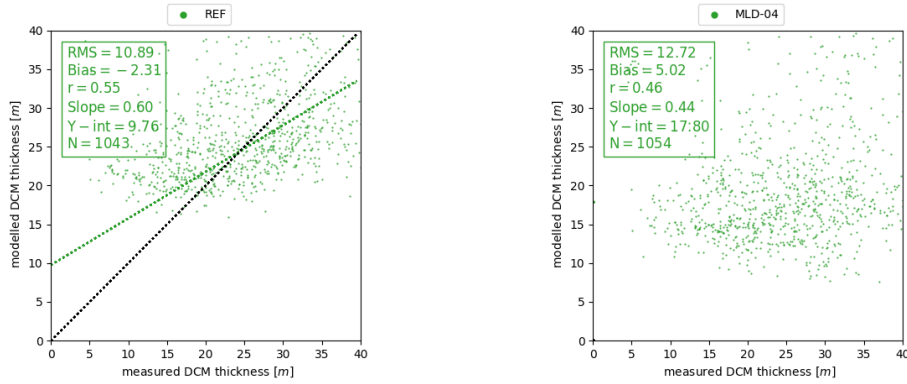
Figure 9. Scatter plot of DCM thickness. Left panel reports REF simulation ($D_v^{background} =10^{-4}\mathrm{m}^2\mathrm{s}^{-1}$), right panel shows MLD04 simulation ($D_v^{background} =10^{-6}\mathrm{m}^2\mathrm{s}^{-1}$). The thickness is defined as $\pm\sigma/2$ centered on the maximum computed on the vertical profiles by means of a Gaussian fit.
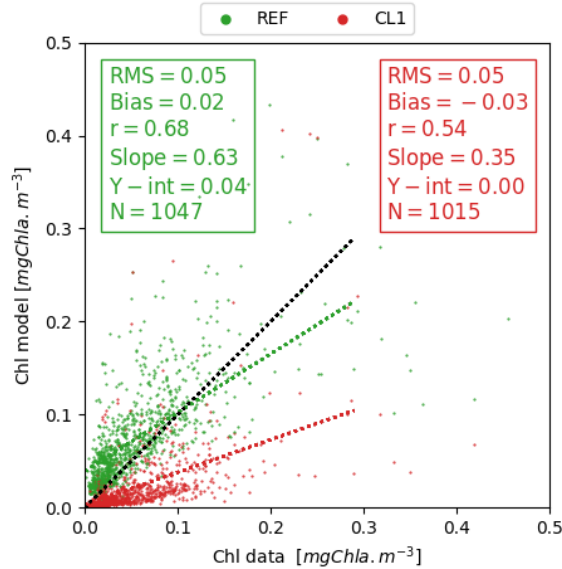
Figure 10. Scatter plot comparing 0-25 m average surface chlorophyll versus BGC-Argo float data for the stratified period condition (DCM > 40 m).

The DCM thickness varies between 20 and 40 m for the model, whereas a higher variability from 5 m to 40 m is measured by BGC-Argo (Figure 9). In section 3.2 we evidence how DCM thickness is controlled by the background vertical eddy diffusivity coefficient ($D_v^{background}$). 0-25 m average surface chlorophyll shows similar skill (r=0.68, slope=0.63) as in the case of DCM thickness, Figure 10.

5   The skill of the 1D model in reproducing DCM magnitude is lower than for the other indicators: measured chlorophyll con-

centration fluctuates in the DCM and the possible underlying mechanism (e.g. presence of Rossby waves or Kelvin waves) goes beyond the scope of the present paper. We show here the median chlorophyll in the DCM layer ($\pm\sigma/2$) for each BGC-Argo float (Figure 11). In general, simulations tend to underestimate chlorophyll concentration compared to BGC-Argo float in the Western Mediterranean Sea. Following the procedure of the sensitivity analysis shown before, we evaluated the effects

5 of perturbing nutrients for the BGC-Argo floats deployed in the West Mediterranean by increasing them by a factor 2 (orange dots, Figure 11). The skill in reproducing the DCM depth is almost the same between REF and REF with higher nutrients (see supplementary materials section 6), however clearly showing the relevant impact on chlorophyll concentration at DCM. Therefore, it could be possible to finely tune the initial conditions to maximize both skills in terms of DCM value and DCM depth. However, considering the fact that the measurements of concentration of chlorophyll as derived from fluorescence present some

10 uncertainties (Roesler et al., 2017; Barbieux et al., 2018; Organelli et al., 2017a), we prefer to keep the initialization based on reanalysis.
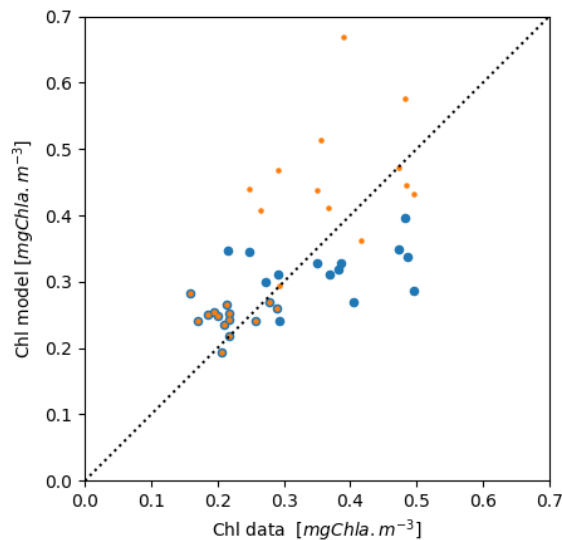
Figure 11. Scatter plot of DCM chlorophyll concentration as defined in the text: median concentration of the REF (blue dots) and from the simulation increasing PO4 (orange dots).

The Mediterranean Sea is a nutrient-limited basin (e.g. Crispi et al. (2001); Lazzari et al. (2016); Powley et al. (2017)), therefore an insight on the role played by nutrients requires further investigation. Phosphate dynamics shows an increase in surface Chl driven by nutrient uptake in upper layers due to convective mixing. During stratification periods, the phosphocline

15 follows the euphotic layer threshold. From results shown hereby, it can be ascertained that together with a strong correlation between light and DCM depth, nutrient concentration is an important driver in regulating phytoplankton biomass at DCM. Indeed, western subbasins exhibit significantly higher values, of both phosphate and biomass, compared to the eastern ones,

Figure 13. It should be noted that the REF simulation is forced by PAR measurements, hence we evaluate the direct impact of nutrients compared to light on DCM properties. The effect of self-shading by chlorophyll and CDOM can increase the role of nutrients in terms of DCM depth modulation, which can be evaluated only by using bio-optical models where attenuation is regulated by chlorophyll or CDOM as presented in section 3.3.
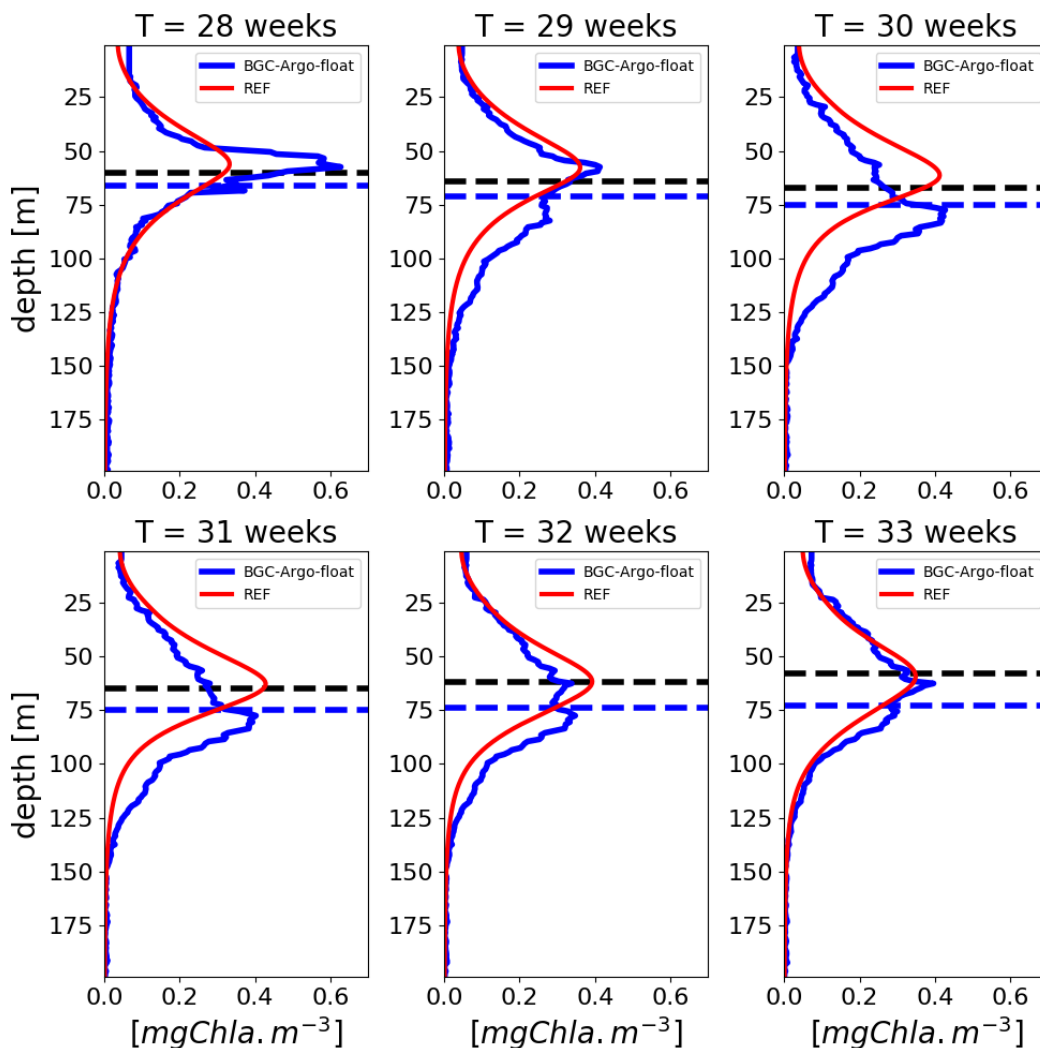
Figure 12. Example of a weekly time series of vertical profiles attributed to the REF simulation of lovbio035b BGC-Argo float (Figure 4) and compared to BGC-Argo float Chl values (thicker line). The horizontal dashed blue line represents the euphotic depth, whereas the horizontal dashed black line indicates the depth where measured PAR equals 0.5 $\mathrm{molquanta.m^{-2}.day^{-1}}$ as identified in Mignot et al. (2014).
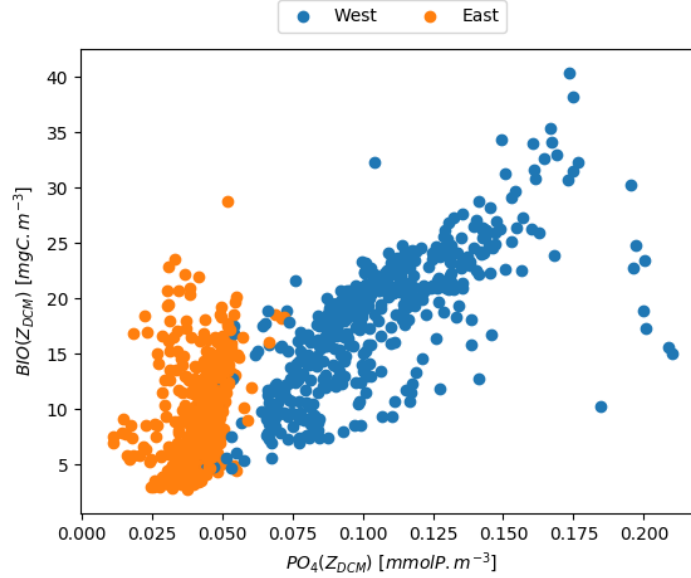
Figure 13. Phosphate (x-axis) and total biomass concentration (y-axis) of phytoplankton at DCM depth, including all modeled float trajectories.

3.2 Vertical Mixing Models

As shown in the previous section, the vertical distribution of chlorophyll displays a distinct variability, which can be at least partially ascribed to mixing. Typically, higher vertical eddy diffusivity values imply smoother structures. During the stratification phase, when DCM forms, the controlling mixing parameter is the background diffusivity $D_v^{background}$. Simplified theoretical models, such as the KiSS (after the names of Kierstead and Slobodkin (1953); Skellam (1951)), can provide rough quantitative scales in order to determine minimum vertical length scales (L_0) that allow formation of stable biomass patches (Ryabov and Blasius, 2008), including the DCM, in a steady state hypothesis:

$$L_0 \propto \sqrt{\frac{D_v}{\mu}} \tag{7}$$

where D_v is the vertical diffusivity coefficient and μ is the growth rate; in stratified conditions, $D_v = D_v^{background}$. Considering any compact vertical interval with favourable conditions for plankton growth (in terms of irradiance and nutrient availability), the increase of background diffusion over a critical value will produce a dispersal of patchy structures (i.e. a relative maximum of chlorophyll concentration), whereas an increase in growth rate μ can drive the formation of finer scale structures by a reduction of L_0.

The dynamics presented in this study is much more complex compared to KiSS, both in BGC-Argo floats data and in the 1-D medium-complexity biogeochemical model (BFM). Vertical eddy diffusivity can simultaneously affect nutrients, phytoplank-

ton, and mesozooplankton with intricate interactions, which in turn make difficult to derive analytical solutions. Moreover, unlike KiSS, both the model and environment are hardly ever in a steady state condition, as a result of daily and seasonal oscillations in physical forcings, which are essentially due to variability in diel irradiance and vertical mixing.

Several simulations, labelled as MLD1, MLD2, MLD3 and MLD4, were carried out by changing the background vertical eddy diffusivity coefficient $D_v^{background}$ values by two orders of magnitude (from $10^{-6}\,\mathrm{m}^2\mathrm{s}^{-1}$ to $10^{-4}\,\mathrm{m}^2\mathrm{s}^{-1}$, see Table 1). This subset of simulations (with float-derived PAR) clusters at a correlation of approximately 0.8 with a root mean square difference (RMSD) of DCM depth between 10-15 m. Modeled chlorophyll profiles appear much smoother than the observed ones, following a Gaussian shape for all tested values of eddy diffusivity. Small scale patterns are not detectable even when $D_v^{background}$ values are reduced to a minimum. Further analyses concerning these aspects are shown in section 3.4.

3.3 Bio-Optical Models
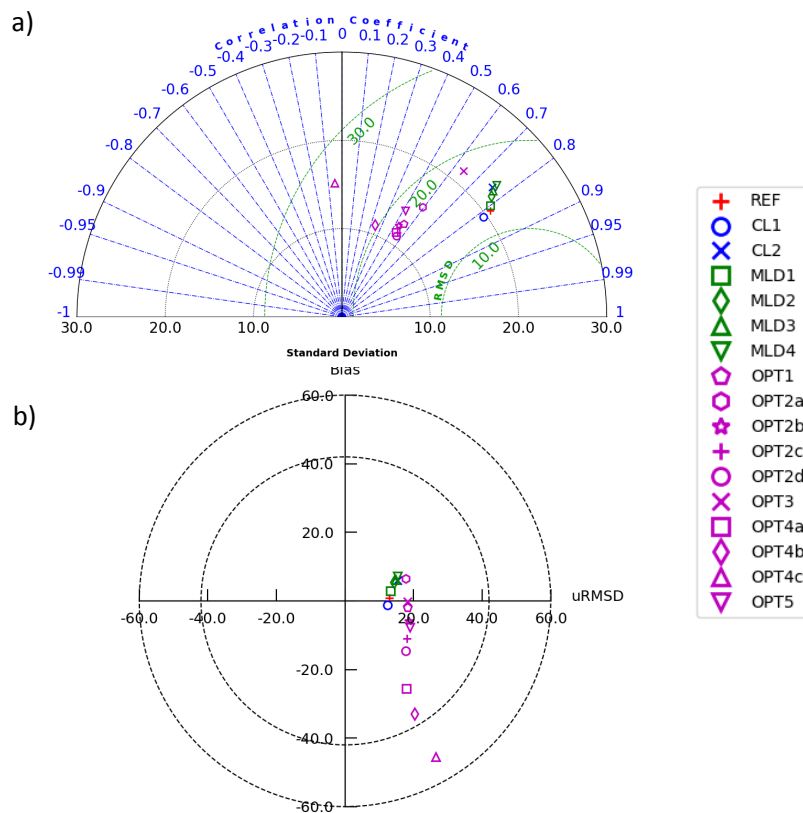
Figure 14. Panel a): Taylor diagram showing model skill in reproducing DCM depth compared to data. Correlation is represented by the angle with positive x axis, whereas distances from the origin depict standard deviations. Green circles illustrate iso-contours of RMSD levels. Panel b): Target diagram showing model skill in reproducing DCM depth compared to data. Distance to the origin defines the RMSD, all units are in meters. The position on x-axis is positive if the model standard deviation is higher than the one from data results and negative in the opposite situation. For the sake of completeness, all models considered are reported in these summarizing skill diagrams.

The alternative bio-optical models (OPT1, OPT2, OPT3) were slightly less accurate compared to REF: correlation decreases from 0.8 to 0.6-0.5, Figure 14. The OPT3 simulation showed a bias very close to zero, thus suggesting an intermediate skill compared to assimilated PAR simulations (e.g. REF) and the bio-optical models (OPT1 and OPT2). OPT1 and the OPT2 cluster of simulations show slightly lower correlations with a RMSD of approximately 20 m in all cases, with increase in bias (almost zero for OPT1 and from 6 m (OPT2a) to -14 m (OPT2d)). The latter may stem from the fact that statistics performed

for OPT2a to OPT2d models ranged from 150 m to 30 m respectively, therefore lowering the number of data considered due to a reduced depth interval. Despite an increasing correlation of the bio-optical model linear regression with decreasing depth range, it should be underlined that the equations for lower depth ranges (such as OPT2d for the first 30 m) most likely do not perform well at greater depths, hence a higher bias in spite of a higher correlation coefficient.

5   The ensemble of simulations with alternative optical models shows in all cases smoother curves compared to measured Chl profiles (see Figure 15). Chlorophyll self-shading effect increases from OPT2a to OPT2d, as explained above, due to different depth ranges of the dataset used to compute linear regressions. Some of the bio-optical models considered, in particular OPT1, OPT2a and OPT2b, reproduce the DCM depth gradient between western and eastern subbasins with a tolerance of ± 10 m (Figure 16). In previous studies (Crispi et al., 2002; Lazzari et al., 2012), the correct simulation of DCM depth longitudinal

10 gradient was obtained by forcing the system with a space-time dependent light attenuation parameter based on Secchi disk climatology or on satellite $K_d(490)$ data. Both empirical approaches prevent to understand whether the origin of such gradients is directly related to external forcings or, on the contrary, if it can be interpreted as a self-emerging property, i.e. related to the appearance of features which are not directly and explicitly imposed from the choice of boundary conditions or model parameters used in the numerical experiment (de Mora et al., 2016).

15 Results shown in Figure 16a suggest that a gradient in DCM depth could be partially reproduced and explained in terms of internal biogeochemical processes and partially due to external forcings (i.e. downward irradiance and nutrient initial conditions), even without considering lateral dynamics.

A direct analysis of the impact of alternative bio-optical models on light attenuation, Figure 16b, indicates that the simulated eastern basin waters present generally lower K_d values (and lower dispersion around the median) for REF and OPT3. In other

20 cases, where self shading is included, the variability is driven by chlorophyll (OPT1, OPT2a, OPT2b, OPT2c, OPT2d and OPT4a, OPT4b, OPT4c) or by chlorophyll and CDOM (OPT5), as bio-optical model parameters do not depend on space and time explicitly. West-east gradients are higher for maximum light attenuation along the water column (cross mark, Figure 16b) where the concentration of chlorophyll is higher. Note that for the OPT3 the average and maximum K_d overlap since K_d is for this simulation constant along the water column.

25 In fact, the average surface PAR of the dataset considered is higher in the eastern areas, especially during the months of January (40%), September (15%), October (22%), November (36%), December (16%), probably due to clearer atmospheric weather conditions. During summer, when DCM stabilizes, the west-east differences in measured surface PAR are lower and oscillate around 10%, however still contributing in increasing irradiance penetration at deeper layers. The western and eastern basins are also different in terms of nutrient regimes that in turn impact biogeochemical dynamics and the DCM depth gradient in

30 non-trivial ways. In particular, the role of nutrients can be evaluated by perturbing initial conditions for the trajectories starting in the western subbasin, as shown in section 3.1. Results indicate that increased nutrients in the western subbasin cause an amplification of the west-east light attenuation gradients (Figure 16c) related to the increase of chlorophyll.The OPT2a test (with increased nutrients) appears to be the most consistent one compared to REF and OPT3, in terms of Kd west-east gradients.

The emerging conceptual scheme is that the first-order controlling mechanism for DCM depth is related to light propagation

35 along the water column, as shown in REF and OPT3 simulations. Other tests indicate that nutrients modulate K_d consistently

with gradients simulated in REF. The temporal scale of subsurface nutrient variability controlling self-shading mechanisms is longer than the one of simulations, suggesting that such a mechanism is especially regulated through initial conditions.

Another key factor pertains to shorter wavelengths (400-450 nm) in the visible part of the spectrum: when light penetrates deeper along the water column, compounds like CDOM are more effective in absorbing light and might in turn enhance spatial gradients in irradiance regimes, which could synergistically contribute to a deeper DCM in eastern subbasins. However, with a current monospectral formulation, such aspects still cannot be addressed. Multi-spectral configurations linked with specific PFT and CDOM absorption terms are thus needed for future in-depth studies of the questions raised in the present work (Dutkiewicz et al., 2015).
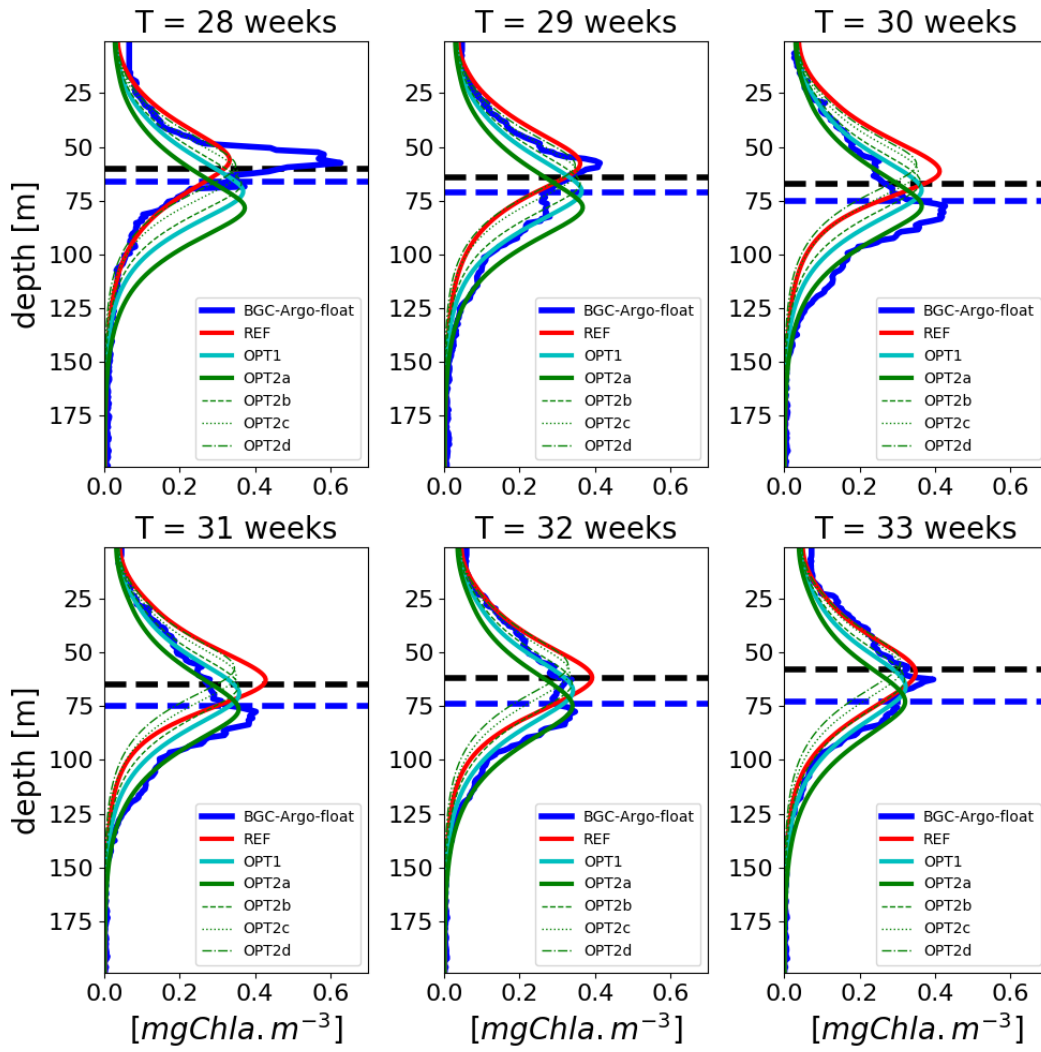
Figure 15. Example of a weekly time series of vertical profiles referred to lovbio035b BGC-Argo float (Figure 4) showing REF simulation and alternative bio-optical models OPT1 and OPT2 compared to BGC-Argo float Chl values (thicker line). The horizontal dashed blue line represents the euphotic depth z_{eu}, whereas the dashed black line indicates the depth where measured PAR equals $0.5 \, \mathrm{mol\,quanta\,m^{-2}\,day^{-1}}$ as identified in Mignot et al. (2014). The legend reports model configurations listed in Table 1.
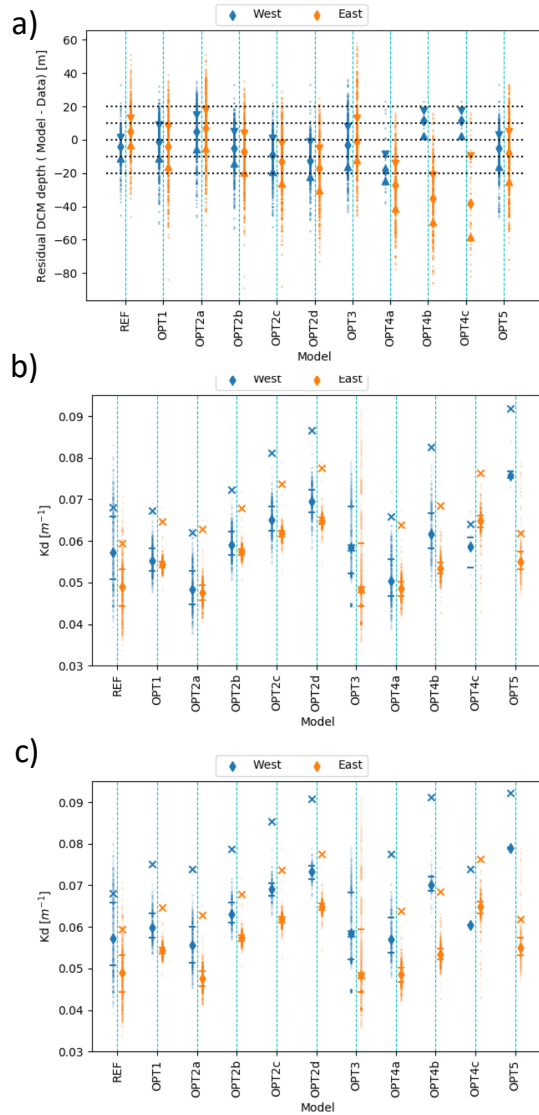
Figure 16. Panel a): Scatter plots of the residual difference between measured and modeled DCMs. The x-axis reports model configurations listed in Table 1. On the y-axis, residuals' median values for west (blue) and east (orange) profiles are shown. Triangles indicate the 25th and 75th percentiles. Panel b): K_d for west and east subbasin during stratified period, diamonds indicate the median over the vertical column, 25th and 75th percentiles are the horizontal lines. Crosses show the maximum over the vertical column. Panel c) is the same as b) but with double initial nutrient concentrations for the western basin simulations.

3.4 Daily variable versus constant PAR forcings

The use of daily averaged irradiance (i.e. with continuous light, CL1 and CL2) was compared against REF that includes the diurnal variability. A consistent reduction of surface Chl concentrations was observed in the former case (Figure 10), with a correlation lower than REF, affecting much less (in relative terms) the values around DCM (CL2 is shown in Figure 17). Near the surface, phytoplankton is limited by low nutrients (especially in eastern subbasins) whereas closer to DCM, the trophic limitation is weaker, sometimes null (Behrenfeld and Boss, 2003; Behrenfeld et al., 2004). One possible explanation could be that light limitation at DCM at a low-irradiance regime is almost linear, thus the averaging effects appear to be having a smaller impact than at surface, where light limitation is highly non-linear due to saturation. Furthermore, the Geider formulation for

Chl acclimation (Geider et al., 1998) in case of diurnal variability generates an increase in chlorophyll-to-carbon (Chl:C) ratio. This could in turn have important consequences in operational applications, where data assimilation is employed for model skill improvement. At the surface, the adoption of a diurnal cycle formulation could reduce corrections made by the assimilation scheme and therefore minimize possible spurious trends introduced by it (Gehlen et al., 2015).
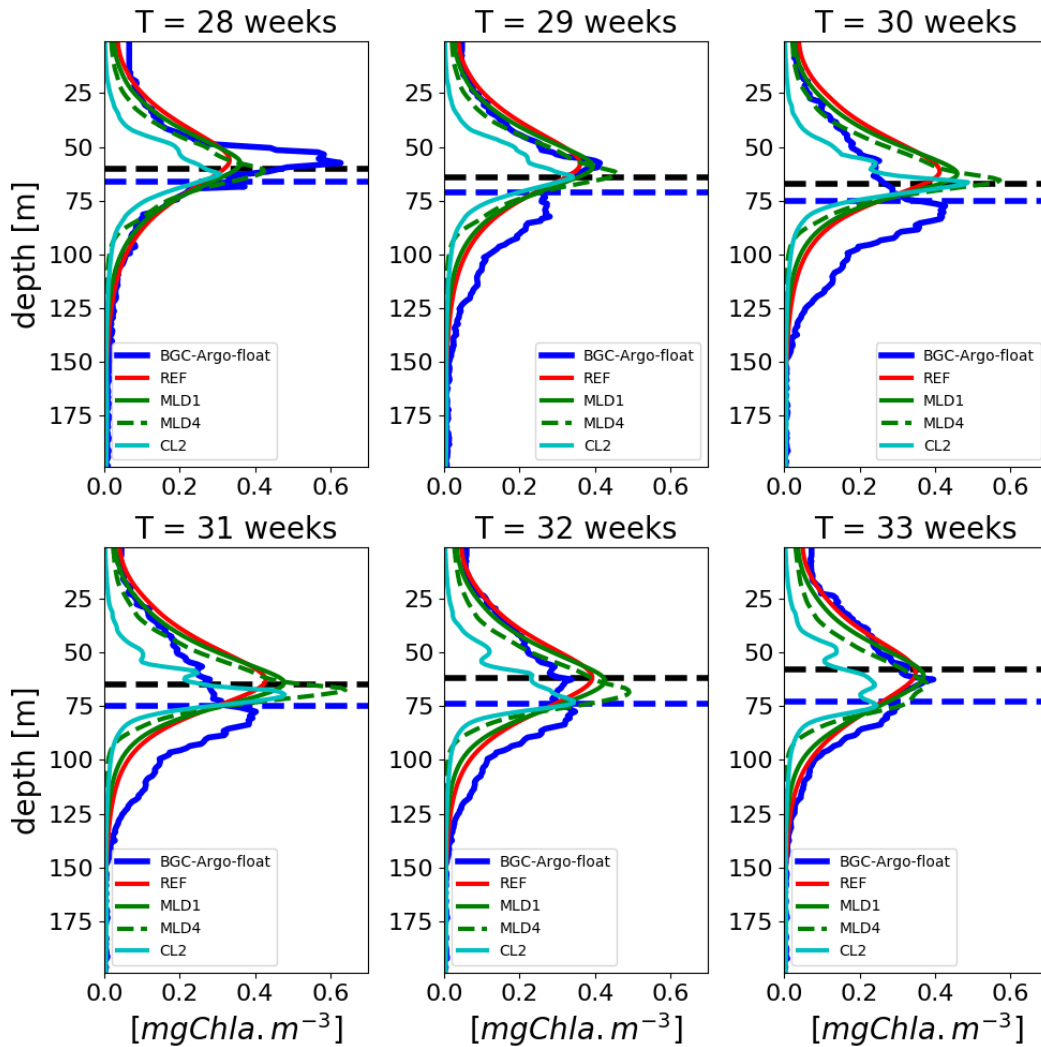
Figure 17. Example of a weekly time series of vertical profiles referred to lovbio035b BGC-Argo float (Figure 4) based on diel variability and constant daily light descriptions, compared to BGC-Argo float Chl values (thicker blue line). The horizontal dashed blue line represents the euphotic depth z_{eu}, whereas the dashed black line indicates the depth where measured PAR equals $0.5\,\mathrm{mol\,quanta\,m^{-2}\,day^{-1}}$ as identified in Mignot et al. (2014). The legend reports model configurations listed in Table 1.

Combining daily-averaged irradiances with lowest diffusivity rates ($D_v^{background}{=}10^{-6}\,\mathrm{m^2\,s^{-1}}$, simulation CL2) results in additional relative chlorophyll maxima at surface layers (see Figure 17, panel "T = 33 weeks"), as well as in increased patchiness of the entire vertical profile. Similar chlorophyll profiles with multiple subsurface maxima were identified in a comprehensive fluorescence data analysis in the Mediterranean Sea, Lavigne et al. (2015). Theoretical consideration predicts different maxima along the water column based on Tilman resource competition theory applied to a heterogeneous system (Ryabov and Blasius, 2011). At this stage, however, it is difficult to assess whether the patchy structures observed in data and

model are, for various reasons, realistic or artefactual. Nonetheless, it can be ascertained that the background diffusion needed to maintain such structures in model simulations is very low.

Within the framework of currently used mathematical formulations in the 1-D BFM model, the inclusion of diurnal variability tends to reduce the formation of fine-scaled structures that could be interpreted in terms of a reduction in diel growth (μ) or
5   seen as a possible perturbation that has an equivalent effect of an increased diffusion.

3.5 Bio optical models with CDOM formulation

OPT4 and OPT5 simulations take into consideration CDOM dynamics by including an additional term in OPT2a, where light attenuation by PAR was described only in terms of Chl. In OPT4a, b, and c, CDOM is parametrized as "dead" chlorophyll, by changing only the rate of Chl decay from 1 day to 1 month. Such simplified dynamics description, albeit arguably, derives
10   from high correlation observed between Chl and CDOM in Morel and Maritorena (2001). It should be noted, however, that no analysis, which could corroborate findings from Morel and Maritorena (2001), was carried out within the dataset examined hereby due to a lack of information on CDOM fluorescence. In all three model configurations, the "dead" chlorophyll accumulation results in higher turbidity levels that in turn reduce light penetration depths. This is quantified by significantly negative DCM biases (over 40 m in OPT04c), which result in shallower DCM compared to BGC-Argo derived profiles since the atten-
15   uation of Chl is overestimated even when considering fastest degradation rates. (Figure 14). The experiment OPT5 mimics the CDOM dynamics described in Dutkiewicz et al. (2015) where a lower bias is observed compared to the (over)simplified OPT4 tests (where correlation coefficients range from 0.6 to less than 0.1 for OPT4a to OPT4c respectively). OPT5 still results in a negative bias of around 10 m compared to the values from -25 m to -40 m for OPT4a to OPT4c.

In open ocean systems, at least three different mechanisms concerning CDOM entrainment in the euphotic layer are con-
20   sidered: lateral flux of CDOM from terrestrial waters (allochthonous origin), production of CDOM within the euphotic layer (autochthonous origin) and bottom-up flux of CDOM from the subsurface layer not affected by bleaching (Nelson and Siegel, 2013).

Figure 18 shows an example for a BGC-Argo float deployed in the North West Mediterranean subbasin (NWM). The model, regardless of initial conditions, correctly drives CDOM absorption coefficients in deeper layers to low values, while an en-
25   hanced surface production reinforces mineralization and bleaching and thus realizes a continuum of CDOM reactivity and lability. Results of CDOM variability from the BOUSSOLE site (north-west Mediterranean, Antoine et al. (2008)) show that CDOM absorption ranges to a maximum value of 0.07 m^{-1} and indicate that there is a temporal delay between phytoplankton bloom and a maximum in CDOM absorption (Figure 3 in Organelli et al. (2014)), whereas deeper layers (below 100 m) have generally lower CDOM absorption. The dataset shown in Organelli et al. (2014) evidences that cycles of CDOM accumula-
30   tion are followed by depletion in the upper 10 m due to photodegradation in summer. In modeling results presented hereby, bleaching has a deeper effect over the entire CDOM "productive" layer (see red and blue lines, Figure 18), while the subsurface CDOM maximum is not reproduced. Additional investigations of the OPT5 model configuration can address the autochthonous source dynamics, as well as the bottom-up flux of CDOM in this region. The lack of CDOM accumulation in deeper layers for

the OPT5 configuration hinders a proper analysis of the mechanism suggested in section 3.1 related to the emergence of CDOM from subsurface dark layers. Improving model dynamics calibrations could be possibly achieved by utilizing information on CDOM light absorption from BCG-Argo floats measurements (Xing et al., 2012; Organelli et al., 2017b).
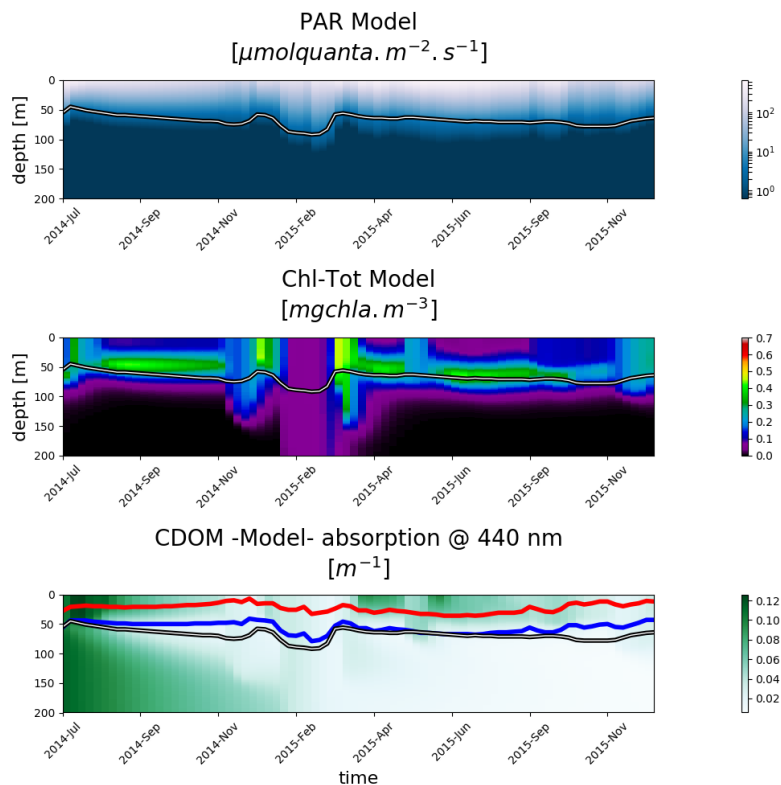
Figure 18. Hovmöller diagrams for BGC-Argo float lovbio068d (WMO code 6901648) showing: PAR (top), total chlorophyll (middle) and CDOM (bottom) simulated by model configuration OPT5. The white, red and blue lines depict the euphotic, 100% and 10% bleaching depths respectively.

4   Conclusions

The coupled modeling/experimental approach presented here provides a robust and accurate reproduction of DCM variability across the Mediterranean Sea. Such a combined configuration can integrate in a single framework multi-data measurements provided by BGC-Argo floats. DCM is a ubiquitous feature of the Chl vertical structure in the Mediterranean, and different forcing conditions generate geographical gradients in DCM characteristics (i.e. shallower DCM in western regions, deepening eastwards). Second-order features, such as impulsive vertical spikes or specific patterns observed in BGC-Argo profiles, are also qualitatively reproduced. Results for the reference simulation, where measured PAR is adopted, are summarized as follows:

– mixing and irradiance propagation control Chl dynamics;

- DCM position is mostly controlled by PAR.

- nutrients control the amount of biomass at DCM.

Moreover, it was demonstrated that vertical processes considered in the 1-D model, such as irradiance regimes and vertical mixing, allow to properly reconstruct a large part of Chl dynamics, which was quantified also by skill diagrams.

The role of nutrients in modulating self-shading (as inferred with bio-optical alternative experiments) appears relevant to shape west-east heterogeneity of vertical light attenuation.

The emerging conceptual scheme is that DCM gradients are directly controlled by irradiance modulation, than is in turn controlled through bio-optical processes which change attenuation according to optically active substances (e.g. chlorophyll, CDOM). Nutrients can impact attenuation by regulating chlorophyll content. The time scale of the nutrient pool variability is longer than the ones considered in the present simulation, thus enabling initial conditions to modulate west-east gradients.

Such kind of data-rich experiments, combined with a 1-D numerical model, could be considered as a useful tool also to a broader community, rather than only to biogeochemical modelers, in particular to address process studies.

The presented approach might be useful also to quantify the amount of measured signal related to vertical dynamics and the one derived from other processes, such as horizontal advection and subduction of water masses. The usage of PAR measured from BGC-Argo floats (used in REF, CL1, CL2, MLD1, MLD2, MLD3 and MLD4) provides higher correlations compared to configurations with alternative bio-optical models (used in OPT1, OPT2, OPT3, OPT4 and OPT5). CL1 (without diurnal cycle) shows overall highest correlation, comparable with REF (Figure 14a).

The comparison of different bio-optical models indicates that, when lacking direct measurements of PAR in subsurface layers, the most fitting alternatives would be OPT3, OPT2a and OPT1, that provide relatively lower bias and higher correlation coefficients (between 0.5 and 0.7), as well as lower RMSD values compared to REF.

Such an analysis can also suggest the rate of improvement when considering a value of light fully integrated in the visible range of the spectrum (400 to 700 nm, REF) versus simplified approaches (i.e. all the OPT simulations here considered).

These results further support the strategic relevance of BGC-Argo data. Temperature, salinity and radiometric parameters encapsulate fundamental information for the reconstruction of primary producers dynamics and are paramount to investigate hypotheses concerning DCM formation. CDOM fluorescence data measured by BGC-Argo floats could be integrated in simulations to further infer and reconstruct the observed biogeochemical processes.

Furthermore, considering a general 3-D biogeochemical model, it is not possible to have a full data coverage of the in-water PAR field without a fully coupled radiative transfer model. Such an approach could be thus exported to more complex 3-D biogeochemical models and generalized at a global scale.

Code and data availability. The BFM biogeochemical model and its documentation can be downloaded at the following address: http://bfm-community.eu/. The quality-controlled databases used in the present manuscript are publicly available from the SEANOE (SEA scieNtific Open data Edition) publisher at https://doi.org/10.17882/49388 and https://doi.org/10.17882/47142 for vertical profiles and products within the first optical depth, respectively.

Author contributions. E.T. and P.L. have designed the manuscript. E.T. performed the BGC-Argo data analysis, P.L. performed the simulations. All the authors contributed to the manuscript writing.

Competing interests. Authors declare that no competing interests are present.

Acknowledgements. This work is part of the PhD project of Elena Terzić, which was funded under the CMEMS contract for the Biogeochem-
5 istry Production Unit for the Mediterranean Sea. The simulations were performed in the framework of the ISCRA C project NOVBIOGE (HP10C8C9O6), granted by CINECA, Italy. This work was supported by the French "Equipement d'avenir" NAOS project (Novel Argo Ocean observing System - NAOS project funded by Agence Nationale de la Recherche (grant agreement ANR J11R107-F); the "Remotely-sensed biogeochemical cycles of the oceans- remOcean" project funded by the European Research Council (grant agreement 246777); the Argo-Italy project funded by the Italian Ministry of Education, University and Research; and the French Bio-Argo program - Bio-Argo
10 France funded by CNES-TOSCA, LEFE Cyber, and GMMC. We acknowledge sponsorship from the BIOPTIMOD CMEMS Service Evolution project and from the MISTRALS-MERMEX project.

References

Antoine, D., Guevel, P., Desté, J.-F., Bécu, G., Louis, F., Scott, A. J., and Bardey, P.: The "BOUSSOLE" buoy—A new transparent-to-swell taut mooring dedicated to marine optics: Design, tests, and performance at sea, Journal of Atmospheric and Oceanic Technology, 25, 968–989, 2008.

Baird, M. E., Cherukuru, N., Jones, E., Margvelashvili, N., Mongin, M., Oubelkheir, K., Ralph, P. J., Rizwi, F., Robson, B. J., Schroeder, T., et al.: Remote-sensing reflectance and true colour produced by a coupled hydrodynamic, optical, sediment, biogeochemical model of the Great Barrier Reef, Australia: comparison with satellite data, Environmental Modelling & Software, 78, 79–96, 2016.

Barbieux, M., Uitz, J., Bricaud, A., Organelli, E., Poteau, A., Schmechtig, C., Gentili, B., Obolensky, G., Leymarie, E., Penkerc'h, C., et al.: Assessing the Variability in the Relationship Between the Particulate Backscattering Coefficient and the Chlorophyll a Concentration From a Global Biogeochemical-Argo Database, Journal of Geophysical Research: Oceans, 123, 1229–1250, 2018.

Behrenfeld, M. J. and Boss, E.: The beam attenuation to chlorophyll ratio: an optical index of phytoplankton physiology in the surface ocean?, Deep Sea Research Part I: Oceanographic Research Papers, 50, 1537–1549, 2003.

Behrenfeld, M. J., Prasil, O., Babin, M., and Bruyant, F.: In search of a physiological basis for covariations in light-limited and light-saturated photosynthesis, Journal of Phycology, 40, 4–25, 2004.

Bissett, W., Walsh, J., Dieterle, D., and Carder, K.: Carbon cycling in the upper waters of the Sargasso Sea: I. Numerical simulation of differential carbon and nitrogen fluxes, Deep Sea Research Part I: Oceanographic Research Papers, 46, 205–269, 1999.

Bricaud, A., Morel, A., and Prieur, L.: Absorption by dissolved organic matter of the sea (yellow substance) in the UV and visible domains 1, Limnology and oceanography, 26, 43–53, 1981.

Claustre, H., Morel, A., Hooker, S., Babin, M., Antoine, D., Oubelkheir, K., Bricaud, A., Leblanc, K., Queguiner, B., and Maritorena, S.: Is desert dust making oligotrophic waters greener?, Geophysical Research Letters, 29, 2002.

Crise, A., Allen, J., Baretta, J., Crispi, G., Mosetti, R., and Solidoro, C.: The Mediterranean pelagic ecosystem response to physical forcing, Progress in Oceanography, 44, 219–243, 1999.

Crispi, G., Mosetti, R., Solidoro, C., and Crise, A.: Nutrients cycling in Mediterranean basins: the role of the biological pump in the trophic regime, Ecological Modelling, 138, 101–114, 2001.

Crispi, G., Crise, A., and Solidoro, C.: Coupled Mediterranean ecomodel of the phosphorus and nitrogen cycles, Journal of Marine Systems, 33, 497–521, 2002.

de Boyer Montégut, C., Madec, G., Fischer, A. S., Lazar, A., and Iudicone, D.: Mixed layer depth over the global ocean: An examination of profile data and a profile-based climatology, Journal of Geophysical Research: Oceans, 109, 2004.

de Mora, L., Butenschön, M., and Allen, J. I.: The assessment of a global marine ecosystem model on the basis of emergent properties and ecosystem function: a case study with ERSEM, Geoscientific Model Development, 9, 59–76, doi:10.5194/gmd-9-59-2016, https://www.geosci-model-dev.net/9/59/2016/, 2016.

d'Ortenzio, F. and Ribera d'Alcalà, M.: On the trophic regimes of the Mediterranean Sea: a satellite analysis, Biogeosciences, 6, 139–148, 2009.

Dowd, M., Jones, E., and Parslow, J.: A statistical overview and perspectives on data assimilation for marine biogeochemical models, Environmetrics, 25, 203–213, 2014.

Dutkiewicz, S., Hickman, A. E., Jahn, O., Gregg, W. W., Mouw, C. B., and Follows, M. J.: Capturing optically important constituents and properties in a marine biogeochemical and ecosystem model, Biogeosciences, 12, 4447–4481, doi:10.5194/bg-12-4447-2015, https://www.biogeosciences.net/12/4447/2015/, 2015.

D'Ortenzio, F. and Prieur, L.: The upper mixed layer, Life in the Mediterranean Sea: A look at habitat changes, edited by: Noga Stambler, Nova Science Publisher, pp. 127–156, 2012.

Fujii, M., Boss, E., and Chai, F.: The value of adding optics to ecosystem models: a case study, Biogeosciences Discussions, 4, 1585–1631, 2007.

Gehlen, M., Barciela, R., Bertino, L., Brasseur, P., Butenschön, M., Chai, F., Crise, A., Drillet, Y., Ford, D., Lavoie, D., Lehodey, P., Perruche, C., Samuelsen, A., and Simon, E.: Building the capacity for forecasting marine biogeochemistry and ecosystems: recent advances and future developments, Journal of Operational Oceanography, 8, s168–s187, doi:10.1080/1755876X.2015.1022350, https://doi.org/10.1080/1755876X.2015.1022350, 2015.

Geider, R. J., MacIntyre, H. L., and Kana, T. M.: A dynamic regulatory model of phytoplanktonic acclimation to light, nutrients, and temperature, Limnology and oceanography, 43, 679–694, 1998.

Gerbi, G. P., Boss, E., Werdell, P. J., Proctor, C. W., Haëntjens, N., Lewis, M. R., Brown, K., Sorrentino, D., Zaneveld, J. R. V., Barnard, A. H., et al.: Validation of ocean color remote sensing reflectance using autonomous floats, Journal of Atmospheric and Oceanic Technology, 33, 2331–2352, 2016.

Gordon, H. R. and McCluney, W.: Estimation of the depth of sunlight penetration in the sea for remote sensing, Applied optics, 14, 413–416, 1975.

Huisman, J., Sharples, J., Stroom, J. M., Visser, P. M., Kardinaal, W. E. A., Verspagen, J. M., and Sommeijer, B.: Changes in turbulent mixing shift competition for light between phytoplankton species, Ecology, 85, 2960–2970, 2004.

Jolliff, J. K., Kindle, J. C., Shulman, I., Penta, B., Friedrichs, M. A., Helber, R., and Arnone, R. A.: Summary diagrams for coupled hydrodynamic-ecosystem model skill assessment, Journal of Marine Systems, 76, 64–82, 2009.

Kierstead, H. and Slobodkin, L.: The size of water masses containing phytoplankton blooms, Journal of Marine Research, 38, 1953.

Kirk, J. T.: Light and photosynthesis in aquatic ecosystems, Cambridge university press, 1994.

Kohlmeier, C. and Ebenhöh, W.: Modelling the biogeochemistry of a tidal flat ecosystem with EcoTiM, Ocean dynamics, 59, 393–415, 2009.

Lavigne, H., D'Ortenzio, F., Ribera D'Alcalà, M., Claustre, H., Sauzède, R., and Gacic, M.: On the vertical distribution of the chlorophyll a concentration in the Mediterranean Sea: a basin-scale and seasonal approach, Biogeosciences, 12, 5021–5039, doi:10.5194/bg-12-5021-2015, https://www.biogeosciences.net/12/5021/2015/, 2015.

Lazzari, P., Teruzzi, A., Salon, S., Campagna, S., Calonaci, C., Colella, S., Tonani, M., and Crise, A.: Pre-operational short-term forecasts for Mediterranean Sea biogeochemistry, Ocean Science, 6, 25–39, 2010.

Lazzari, P., Solidoro, C., Ibello, V., Salon, S., Teruzzi, A., Béranger, K., Colella, S., and Crise, A.: Seasonal and inter-annual variability of plankton chlorophyll and primary production in the Mediterranean Sea: a modelling approach, Biogeosciences, 9, 217, 2012.

Lazzari, P., Solidoro, C., Salon, S., and Bolzon, G.: Spatial variability of phosphate and nitrate in the Mediterranean Sea: A modeling approach, Deep Sea Research Part I: Oceanographic Research Papers, 108, 39–52, 2016.

Leymarie, E., Penkerc'h, C., Vellucci, V., Lerebourg, C., Antoine, D., Boss, E., Lewis, M. R., D'Ortenzio, F., and Claustre, H.: ProVal: A new autonomous profiling float for high quality radiometric measurements, Frontiers in Marine Science, 5, 437, 2018.

Mignot, A., Claustre, H., Uitz, J., Poteau, A., D'Ortenzio, F., and Xing, X.: Understanding the seasonal dynamics of phytoplankton biomass and the deep chlorophyll maximum in oligotrophic environments: A Bio-Argo float investigation, Global Biogeochemical Cycles, 28, 856–876, 2014.

Mignot, A., Ferrari, R., and Claustre, H.: Floats with bio-optical sensors reveal what processes trigger the North Atlantic bloom, Nature communications, 9, 190, 2018.

Mobley, C., Boss, E., and Roesler, C.: Ocean optics web book, 2010.

Morel, A.: Optical modeling of the upper ocean in relation to its biogenous matter content(Case I waters), Journal of Geophysical Research, 93, 749–10, 1988.

Morel, A. and Gentili, B.: The dissolved yellow substance and the shades of blue in the Mediterranean Sea., Biogeosciences, 6, 2009a.

Morel, A. and Gentili, B.: A simple band ratio technique to quantify the colored dissolved and detrital organic material from ocean color remotely sensed data, Remote Sensing of Environment, 113, 998–1011, 2009b.

Morel, A. and Maritorena, S.: Bio-optical properties of oceanic waters- A reappraisal, Journal of Geophysical research, 106, 7163–7180, 2001.

Nelson, N. B. and Siegel, D. A.: The global distribution and dynamics of chromophoric dissolved organic matter, Annual Review of Marine Science, 5, 447–476, 2013.

Organelli, E., Bricaud, A., Antoine, D., and Matsuoka, A.: Seasonal dynamics of light absorption by chromophoric dissolved organic matter (CDOM) in the NW Mediterranean Sea (BOUSSOLE site), Deep Sea Research Part I: Oceanographic Research Papers, 91, 72–85, 2014.

Organelli, E., Claustre, H., Bricaud, A., Schmechtig, C., Poteau, A., Xing, X., Prieur, L., D'Ortenzio, F., Dall'Olmo, G., and Vellucci, V.: A Novel Near-Real-Time Quality-Control Procedure for Radiometric Profiles Measured by Bio-Argo Floats: Protocols and Performances, Journal of Atmospheric and Oceanic Technology, 33, 937–951, 2016.

Organelli, E., Barbieux, M., Claustre, H., Schmechtig, C., Poteau, A., Bricaud, A., Boss, E., Briggs, N., Dall'Olmo, G., D'Ortenzio, F., et al.: Two databases derived from BGC-Argo float measurements for marine biogeochemical and bio-optical applications, Earth System Science Data, 9, 861, 2017a.

Organelli, E., Claustre, H., Bricaud, A., Barbieux, M., Uitz, J., D'Ortenzio, F., and Dall'Olmo, G.: Bio-optical anomalies in the world's oceans: An investigation on the diffuse attenuation coefficients for downward irradiance derived from Biogeochemical Argo float measurements, Journal of Geophysical Research: Oceans, 2017b.

Oubelkheir, K., Claustre, H., Sciandra, A., and Babin, M.: Bio-optical and biogeochemical properties of different trophic regimes in oceanic waters, Limnology and oceanography, 50, 1795–1809, 2005.

Powley, H. R., Krom, M. D., and Van Cappellen, P.: Understanding the unique biogeochemistry of the Mediterranean Sea: Insights from a coupled phosphorus and nitrogen model, Global Biogeochemical Cycles, 2017.

Riley, G. A.: Oceanography of Long Island Sound, 1952-54, vol. 15, Bingham Oceanographic Laboratory, 1956.

Riley, G. A.: Transparency-chlorophyll relations, Limnology and Oceanography, 20, 150–152, 1975.

Roesler, C., Uitz, J., Claustre, H., Boss, E., Xing, X., Organelli, E., Briggs, N., Bricaud, A., Schmechtig, C., Poteau, A., et al.: Recommendations for obtaining unbiased chlorophyll estimates from in situ chlorophyll fluorometers: A global analysis of WET Labs ECO sensors, Limnology and Oceanography: Methods, 2017.

Ryabov, A. and Blasius, B.: Population growth and persistence in a heterogeneous environment: the role of diffusion and advection, Mathematical Modelling of Natural Phenomena, 3, 42–86, 2008.

Ryabov, A. B. and Blasius, B.: A graphical theory of competition on spatial resource gradients, Ecology Letters, 14, 220–228, 2011.

Ryabov, A. B. and Blasius, B.: Depth of the biomass maximum affects the rules of resource competition in a water column, The American Naturalist, 184, E132–E146, 2014.

Schmechtig, C., Poteau, A., Claustre, H., D'Ortenzio, F., Dall'Olmo, G., and Boss, E.: Processing Bio-Argo particle backscattering at the DAC level, 2016.

5   Skellam, J. G.: Random dispersal in theoretical populations, Biometrika, 38, 196–218, 1951.

Teruzzi, A., Dobricic, S., Solidoro, C., and Cossarini, G.: A 3-D variational assimilation scheme in coupled transport-biogeochemical models: Forecast of Mediterranean biogeochemical properties, Journal of Geophysical Research: Oceans, 119, 200–217, 2014.

Teruzzi, A., Bolzon, G., Salon, S., Lazzari, P., Solidoro, C., and Cossarini, G.: Assimilation of coastal and open sea biogeochemical data to improve phytoplankton simulation in the Mediterranean Sea, Ocean Modelling, 132, 46 – 60,

10   doi:https://doi.org/10.1016/j.ocemod.2018.09.007, http://www.sciencedirect.com/science/article/pii/S1463500318303184, 2018.

Vichi, M., Cossarini, G., Gutierrez, M., Lazzari, P., Lovato, T., Mattia, G., Masina, S., McKiver, W., Pinardi, N., Solidoro, C., and Zavatarelli, M.: The Biogeochemical Flux Model (BFM): Equation Description and User Manual, BFM Report series 1, BFM version 5 (BFM-V5), Bologna, Italy, http://bfm-community.eu, 2013.

Wojtasiewicz, B., Hardman-Mountford, N. J., Antoine, D., Dufois, F., Slawinski, D., and Trull, T. W.: Use of bio-optical profiling float data

15   in validation of ocean colour satellite products in a remote ocean region, Remote Sensing of Environment, 209, 275–290, 2018.

Xing, X., Morel, A., Claustre, H., D'Ortenzio, F., and Poteau, A.: Combined processing and mutual interpretation of radiometry and fluorometry from autonomous profiling Bio-Argo floats: 2. Colored dissolved organic matter absorption retrieval, Journal of Geophysical Research: Oceans, 117, 2012.