A modelling and data assimilation approach to estimate the carbon/chlorophyll ratio in a coupled hydrodynamical-biological model.

Blaise Faugeras, Olivier Bernard, Antoine Sciandra, Marina Lévy

June 29, 2004

Abstract

The principal objective of hydrodynamical-biological models is to provide estimates of the main carbon fluxes such as total and export oceanic production. These models are nitrogen based, that is to say that the variables are expressed in terms of their nitrogen content. Moreover models are calibrated using chlorophyll data sets. Therefore carbon to chlorophyll (C:Chl) and carbon to nitrogen (C:N) ratios have to be assumed. This paper addresses the problem of the representation of these ratios. In a 1D framework at the DYFAMED station (NW Mediteeanean Sea) we propose a model which enables the estimation of the basic biogeochemical fluxes and in which the spatio-temporal variability of the C:Chl and C:N ratios is fully represented in a mechanical way. This is achieved through the introduction of new state variables coming from the embedding of a phytoplankton growth model in a more classical NNPZD-DOM model. Following this modelling step, the parameters of the model are estimated using the adjoint data assimilation method which enables the assimilation of chlorophyll and nitrate data sets collected at DYFAMED in 1997.
Comparing the predictions of the new model with those of the classical NNPZD-DOM model which was calibrated with the same data sets, we find that both models reproduce the reference data in a comparable manner. Both fluxes and stocks can be equally well predicted by either model. However if the models are coinciding on an average basis, they are diverging from a variability prediction point of view. In the new model biology adapts much faster to its environment giving rise to higher short term variations and finally to a higher variability.

1 Introduction

The estimation of the amount of carbon fixed by oceanic phytoplankton during primary production is a key point to quantify the future evolution of atmospheric carbon. To address this issue biogeochemical models are increasingly being used and applied to a variety of regions in the ocean. All these models are more or less complex variants of the Fasham *et al.* [6] model and are nitrogen based. That is to say that the compartments or state variables of the model, and in particular phytoplankton, are expressed in terms of their nitrogen content ($mmolN/m^3$). The main reason for this is that nitrogen plays a critical role in ocean biology as an important limiting nutrient in a wide range of areas. Therefore it is a natural currency for modelling biological fluxes.

However, if the aim of such a model is to estimate and predict carbon fluxes, it means carbon to nitrogen ratios have to be assumed. It is therefore quite common to use the Redfield ratio to convert the nitrogen fluxes computed from the phytoplankton

compartment to carbon fluxes. However the nitrogen to carbon ratio is known to be highly variable [5] both in time and space.

A second weakness of the classical approach is due to the lack of available measurements to calibrate and validate the model. Since the nitrogen of the phytoplanktonic compartment cannot be measured, the chlorophyll a measurements are broadly used to assess the behaviour of this compartment. Therefore a nitrogen to chlorophyll ratio has to be assumed, or using the C:N already hypothesised, a C:Chl must be supposed. In most simple models this ratio is assumed to be constant [6], and in more detailed approaches it can be represented as a static function light [4]. However from laboratory studies it was demonstrated that the C:Chl is also highly variable, especially in conditions of nitrogen limitation [3, 2].

Then the main question that we want to address in this paper rises. What are the effect of the variability of the C:N and C:Chl ratio on the predictions of the biogeochemical models and on their calibration ? What is the benefit of complexifying the model by predicting these ratio with a dynamical model integrating the phenomenon of coupling between light and nitrogen limitation ?

The goal of this paper is to address this aspect using a coupled hydrodynamical-biological modelling. This work follows another study where only the spatial variability of the C:Chl was considered [7], and was estimated through an assimilation step. Here we consider the same 1D framework at the DYFAMED station, we propose a model which enables the estimation of the basic biogeochemical fluxes and in which the spatio-temporal variability of the C:Chl and C:N ratios is fully represented in a mechanical way. This is achieved through the introduction of new state variables. The phytoplankton compartment is replaced by 3 compartments : carbon C, nitrogen associated to chlorophyll L and remaining non-chlorophyllian nitrogen U. This new model results from the embedding of this phytoplankton growth model (the BIOLOV model [20]), developed and validated in chemostat experiments, in a 1D NNPZD-DOM model (Nitrate, Ammonium, Phytoplankton, Zooplankton Detritus, Dissolved Organic Matter) already validated at DYFAMED.

The objective of this paper is not to fully compare this "classical" 1D NNPZD-DOM model with our new model in terms of the quality of the simulated most important biogeochemical fluxes such as production and export. Actually this would have been an interesting question to address but as shown in [7] these fluxes estimates can not be recovered at DYFAMED from stock (chlorophyll and nitrate) data. Indeed we have shown that some measured fluxes need also to constrain the assimilation process for parameter estimation in order to guaranty consistent predictions. As a consequence both models produce comparable average fluxes values since these are constrained by data assimilation. Therefore we clearly concentrate on the spatial temporal variability of the fluxes and ratios as predicted by the models. Then we estimate the benefit of using more complicated models if we want to account both the average value and the variability of these quantities.

They are several reasons why the DYFAMED station, located in the Northwestern Mediterranean Sea, is an interesting test case. First, several biogeochemical production regimes that take place in the world ocean are found here. Deep convection occurs during winter, leading to a spring bloom. Oligotrophy prevails during summer while perturbations in the meteorological forcing generate a secondary bloom in fall. Secondly, the station is far enough away from the Ligurian Current to be sufficiently protected from lateral transport, thereby permitting a 1D study. Moreover, DYFAMED is a JGOFS time-series station which means that a data base of biogeochemical and physical parameters is available to carry out and validate simulations. A final reason for using the DYFAMED station is that it is relatively well known (Deep-Sea Res. II, special issue, 49(11), 2002), and has been the subject of previous model studies ([17], [19]).

The paper is organized as follows: Section 2 is devoted to the construction of the new model by embedding the BIOLOV model into the 1D NNPZD-DOM model. In Section 3 the calibration of the model using data assimilation is discussed. Numerical results are presented in Section 4 and a Discussion section ends the paper.

2 Model presentation

2.1 A one-dimensional NNPZD-DOM model

The main goal of this section is to briefly present the one-dimensional NNPZD-DOM biological model proposed in [19], originally qualitatively calibrated for DYFAMED and then used to assimilate data from this station in [7]. This model is one of the two basis elements of the new model proposed in this paper and for sake of clarity we will recall a few elements about it. However, we refer to [19] and [17] for details of the different modelled processes and for the parameterisations. Compared to other more complete models ([6]; [17]), the relative simplicity of this model is a trade-off to obtain a first approximation of the basic biogeochemical fluxes with a minimum number of prognostic variables. Nitrate and ammonium allow the estimation of new and regenerated production, zooplankton mortality and detrital sedimentation feed the particle export flux.

The biological model is embedded in a 1D physical model, which simulates the evolution over time of velocity, temperature, salinity and turbulent kinetic energy (TKE). As for dynamical processes, the only one taken into account is vertical diffusion. The mixing coefficient, K, is obtained diagnostically from TKE, with a 1.5 closure scheme in the Mellor Yamada nomenclature [9]. The model covers the first 400 m of the water column, with a vertical discretisation of 5 m.

Biological tracers are vertically mixed with the same diffusion coefficient as temperature and salinity. A specific biological reaction term, F_B, is added to the diffusion equation. Tracers are expressed in terms of their nitrogen contents ($mmolNm^{-3}$). For each of the state variables, NO_3, NH_4, P, Z, D and DOM, the prognostic equation reads as follows:

$$\frac{\partial B}{\partial t} - \frac{\partial}{\partial z}(K\frac{\partial B}{\partial z}) = F_B, \tag{1}$$

where B is the biological tracer concentration. A schematic representation of the model is shown on Fig. 1 and the parameters are presented in Table 1.

The formulation of phytoplankton growth takes into account limitation by both nutrients and light. Nutrient limitation follows the Hurtt and Armstrong [15] kinetics:

$$L_{NH_4} = \frac{NH_4}{k_n + NH_4}. \tag{2}$$

Based on the hypothesis that the total limitation follows the same law, i.e.

$$L_{NO_3} + L_{NH_4} = \frac{NO_3 + NH_4}{(k_n + NO_3 + NH_4)}, \tag{3}$$

we obtain,

$$L_{NO_3} = \frac{k_n NO_3}{(k_n + NH_4)(k_n + NO_3 + NH_4)}. \tag{4}$$

Light limitation is expressed as follows:

$$L_I = 1 - \exp(-I/k_{par}). \tag{5}$$

The photosynthetic available radiation, I, is predicted from surface irradiance and phytoplankton pigment content according to a light absorption model. Two different wavelengths are considered and the absorption coefficients depend on the local

phytoplankton concentrations:

$$I(z,P) = I_r(z,P) + I_g(z,P), \tag{6}$$

$$I_g(z,P) = I_g(z-\Delta z,P)(1-exp(-k_g\Delta z)), \tag{7}$$

$$I_r(z,P) = I_r(z-\Delta z,P)(1-exp(-k_r\Delta z)), \tag{8}$$

$$k_g = k_{go} + k_{gp}(\frac{12Pr_d}{r_{pig}r_c})^{l_g}, \tag{9}$$

$$k_r = k_{ro} + k_{rp}(\frac{12Pr_d}{r_{pig}r_c})^{l_r}. \tag{10}$$

Grazing of phytoplankton and detritus is formulated following Fasham *et al.* [6]:

$$G_p = G_p(P,Z,D) = \frac{g_z r P Z}{k_z(rP+(1-r)D)+rP^2+(1-r)D^2}, \tag{11}$$

$$G_d = G_d(P,Z,D) = \frac{g_z(1-r)DZ}{k_z(rP+(1-r)D)+rP^2+(1-r)D^2}. \tag{12}$$

Other modelled biogeochemical interactions include phytoplankton mortality, phytoplankton exudation, zooplankton mortality (considered as large particles which are supposed to be instantaneously exported below the productive layer and remineralised in the water column), zooplankton exudation, fecal pellet production, detritus sedimentation, detritus breakdown, nitrification, and dissolved organic matter remineralisation (Fig. 1).

Below a depth of 150 m, remineralisation processes are preponderant and the surface model does not apply. Instead, decay of phytoplankton, zooplankton and detritus in nutrients with a rate τ, and a vertical redistribution of zooplankton mortality according to Martin's profile [18] parameterise remineralisation below the surface layer. This parameterisation conserves total nitrogen.

An important point in the formulation of this NNPZD-DOM model concerns the C:Chl ratio, r_c. It is a parameter which enters the model through the light limitation term L_I (Eqs 9 to 10). It is used to convert phytoplankton P into Chl which is one of the measured quantity to which the model's output can be compared. The simplest way to choose r_c is to take a constant (in space and time) value. However as with earlier studies [10] it is quantitatively shown in [7] assimilating data from the DYFAMED station that it is impossible to simulate a correct summer surface chlorophyll bloom intensity together with a correct summer subsurface chlorophyll maximum in an oligotrophic regime using this constant C:Chl ratio. This is also in agreement with experimental observations showing a large increase of the C:Chl with respect to light [3, 2]. To overcome this difficulty the authors gave a vertical variability to this ratio (Fig 2). This lead to a significant improvement of the data assimilation results. However this is not cheap from a numerical point of view since a function of depth has to be estimated instead of a single parameter. Some other studies ([4], [17]) propose to parameterise the C:Chl ratio as a static function of light in the water column. It is therefore varying in space and time. Here we propose a mechanical and more realistic representation of the dynamical variations of the $C : Chl$ ratio with respect to environmental conditions. The main tool we use to achieve this objective is the BIOLOV model presented in the following section.

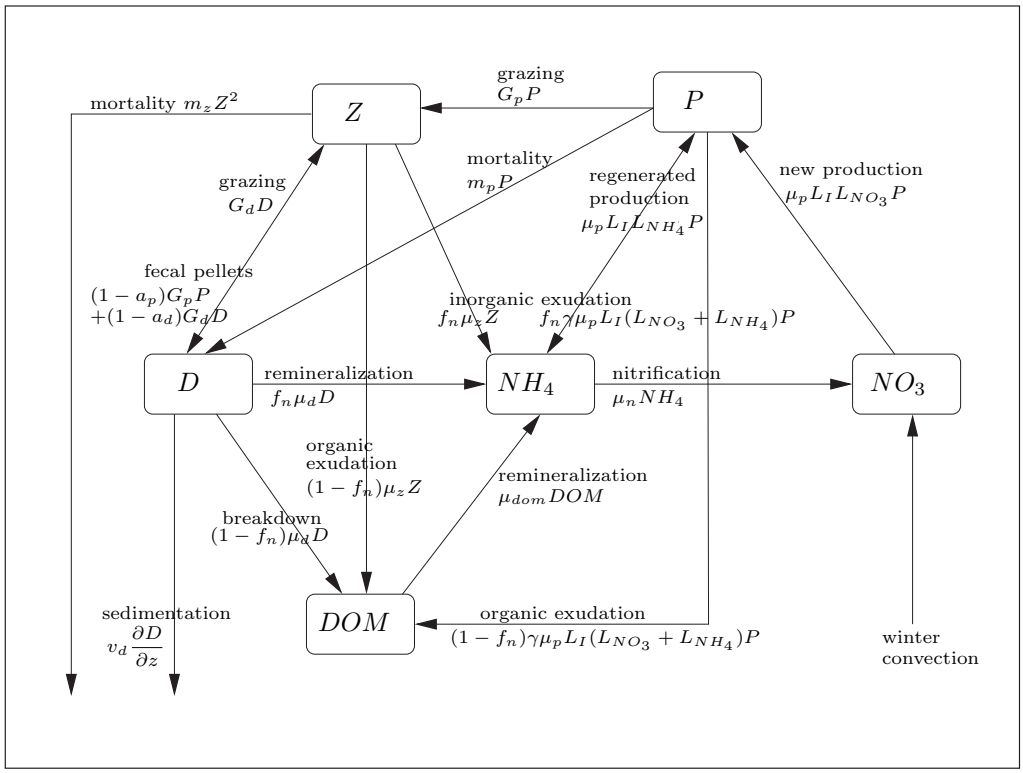

Figure 1: Schematic representation of the compartments and processes of the NNPZD-DOM surface layer model.

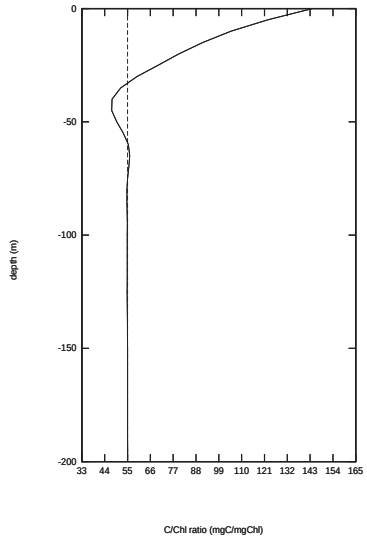

Figure 2: Dashed line : first guess constant C:Chl ratio. Continuous line : C:Chl ratio allowing vertical variability estimated through the assimilation of the 1997 DYFAMED data set (from [7]). Also see Fig. 10 for a plot of this ratio as a function of light.

parameter	name	value	unit
half-saturation const. for nutrients	k_n	0.01	$mmolm^{-3}$
carbon/chlorophyll ratio	r_c	55	$mgC.mgChla^{-1}$
phyto. exudation fraction	γ	0.05	
zoo. nominal preference for phyto food	r	0.7	
max. specific zoo. grazing rate	g_z	8.68E-6	s^{-1}
half-saturation const. for grazing	k_z	1	$mmolN.m^{-3}$
non-assimilated phyto. by zoo.	a_p	0.3	
non-assimilated detritus by zoo.	a_d	0.5	
zoo. specific exudation rate	μ_z	1.16E-6	s^{-1}
phyto. mortality rate	m_p	9.027E-7	s^{-1}
zoo. mortality rate	m_z	1.0E-7	$mmolN.m^{-3}.s^{-1}$
detritus breakdown rate	μ_d	1.04E-6	s^{-1}
detritus sedimentation speed	v_d	5.8E-5	$m.s^{-1}$
max. phyto. growth rate	μ_p	2.31E-5	s^{-1}
light half-saturation const.	k_{pr}	33.33	$W.m^{-2}$
decay rate below the euphotic layer	τ	5.80E-5	s^{-1}
ratio of inorganic exudation	f_n	0.8	
nitrification rate	μ_n	3.81E-7	s^{-1}
slow remineral. rate of DOM to NH_4	μ_{dm}	6.43E-8	s^{-1}
coeff. for Martin's remineralization profile	h_r	-0.858	
pigment absorption in red	k_{rp}	0.037	$m^{-1}.(mgChl.m^{-3})^{-l_r}$
pigment absorption in green	k_{gp}	0.074	$m^{-1}.(mgChl.m^{-3})^{-l_g}$
power law for absorption in red	l_r	0.629	
power law for absorption in green	l_g	0.674	
contribution of Chl to absorbing pigments	r_{pig}	0.7	
water absorptions in green	k_{go}	0.0232	m^{-1}
water absorptions in red	k_{ro}	0.225	m^{-1}
carbon/nitrogen ratio	r_d	6.625	$mmolC.mmolN^{-1}$

Table 1: The 28 parameters of the NNPZD-DOM model

2.2 The BIOLOV model

The BIOLOV model [20] is a phytoplankton growth model which was developed and validated in chemostat experiments. In the BIOLOV model phytoplankton is represented by 3 variables: the carbon C, the non-chlorophyllian nitrogen U and the chlorophyllian nitrogen L. It was originally formulated for a chemostat as a system of ordinary differential equations describing the evolution of 4 state variables C (expressed in $mmolC.m^{-3}$), U, L and S, the nitrogen source (expressed in $mmolN.m^{-3}$). Let f ($mmolN.mmolChla^{-1}$) denote the conversion factor from chlorophyllian nitrogen to chlorophyll a (see Table 3): Chl=$f\,L$. The model integrates both limitations by light and nitrogen and has been designed to have a simple formulation with few parameters and state variables representing measurable quantities in a photobioreactor or in the open ocean. Figure 3 shows a schematic representation of the model and its parameters are presented in Table 2.

Phytoplankton growth is assumed to be triggered by two distinct metabolic pathways: uptake and assimilation of nitrogen on one hand and carbon fixation through photosynthesis on the other hand.

The full nitrogen pathway consists in three steps. First nitrogen (S) is uptaken by the cell into the cellular nitrogen pool (U). The uptake rate is classically represented by a Monod kinetics,

$$\rho_m \frac{S}{K_S + S} C. \tag{13}$$

In a second step, the nitrogen pool is used to produce chlorophyllian proteins (L). The chlorophyll synthesis is assumed to be dependent on light intensity, I, and on the ratio $\frac{L}{C}$. It is also proportional to the nitrogen pool U. The rate of the reaction is therefore assumed to be given by $\gamma(I)U\frac{L}{C}$ where

$$\gamma(I) = \frac{\alpha K_L I}{K_I + I} \frac{K_C}{K_C + I}. \tag{14}$$

Finally, a natural degradation of chlorophyll is also taken into account through a rate βL.

The carbon pathway results from two reactions. First, inorganic carbon (CO_2) is incorporated into the cell through photosynthesis to form particulate carbon, C. The rate of this reaction is $a(I)L$ with

$$a(I) = \frac{\alpha I}{K_I + I}. \tag{15}$$

This rate is highly dependent on light level and chlorophyll pigments which act as catalyzers for this reaction. Inorganic carbon is considered as non-limiting for phytoplankton growth in marine environment and therefore it does not intervene. Second, a proportion of the carbon is lost by respiration. The respiration rate is proportional to the carbon biomass and is written λC.

parameter	name	value	unit
max. uptake rate	ρ_m	5.787 E-6	$mmolN.mmolC^{-1}.s^{-1}$
half-saturation const. for absorption	K_S	0.43	$mmolNm^{-3}$
max. synthesis rate	K_L	6.59	
threshold coefficient	K_C	33.0	$\mu mol\ quanta.m^{-2}.s^{-1}$
max. rate for carbon	α	2.7894 E-4	s^{-1}
half-saturation const. for fixation	K_I	208.5	$\mu mol\ quanta.m^{-2}.s^{-1}$
respiration rate	λ	6.25 E-7	s^{-1}
degradation rate	β	3.9931 E-6	s^{-1}

Table 2: Parameters of the BIOLOV model

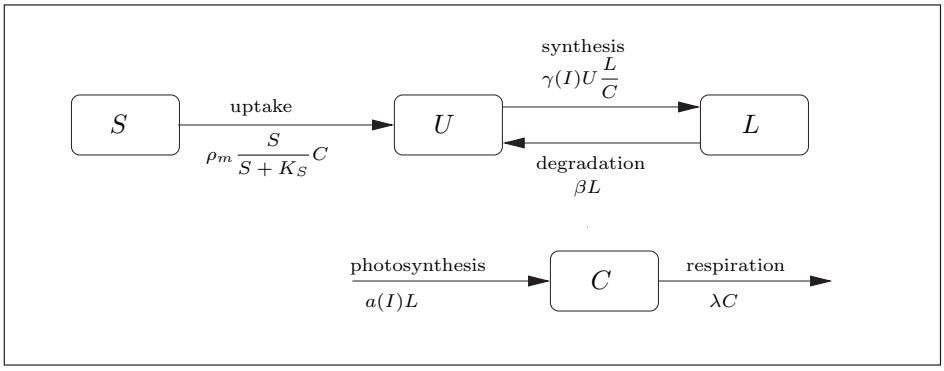

Figure 3: Schematic representation of the compartments and processes of the BIOLOV model.

2.3 Combining the 1D NNPZD-DOM model and the BIOLOV model

Now that we have briefly described the 1D NNPZD-DOM model and the BIOLOV model we are able to construct our new model. Our goal is to embed the BIOLOV model into the NNPZD-DOM model, in order to obtain a model which can be compared to the DYFAMED data set and in which the carbon/chlorophyll ratio evolves dynamically. In line with our objective to compare the effect of different phytoplankton parametrisation on prediction variability we kept unchanged the hydrodynamical model such as the other compartments of the biological model.

Phytoplankton is not anymore represented by a single variable P but by the 3 variables U, L and C. Figure 4 shows a complete representation of the compartments and processes of the new model, and all parameters are presented in Table 3. The different reaction terms can now be written as follows.

$$F_{NO_3} = -\rho_p(I)L_{NO_3}C + \mu_n NH_4, \tag{16}$$

$$\begin{aligned} F_{NH_4} &= -\rho_p(I)L_{NH_4}C + f_n\gamma\rho_p(I)(L_{NO_3}+L_{NH_4})C \\ &\quad +f_n\mu_z Z + f_n\mu_d D - \mu_n NH_4 + \mu_{dom}DOM, \end{aligned} \tag{17}$$

$$\begin{aligned} F_N &= \rho_p(I)L_{NO_3}C + \rho_p(I)L_{NH_4}C - f_n\gamma\rho_p(I)(L_{NO_3}+L_{NH_4})C \\ &\quad -(1-f_n)\gamma\rho_p(I)(L_{NO_3}+L_{NH_4})C - \gamma(I)U\frac{L}{C} + \beta L \\ &\quad -G_p(U+L,D,Z)U - m_pU, \end{aligned} \tag{18}$$

$$F_L = \gamma(I)U\frac{L}{C} - \beta L - m_pL - G_p(U+L,D,Z)L, \tag{19}$$

$$F_C = a(I)L - \lambda C - m_pC - G_p(U+L,D,Z)C, \tag{20}$$

$$\begin{aligned} F_Z &= G_p(U+L,D,Z)(U+L) + G_d(U+L,D,Z)D \\ &\quad -(1-a_d)G_d(U+L,D,Z)D - (1-a_p)G_p(U+L,D,Z)(U+L) \\ &\quad -m_zZ^2 - \mu_zZ, \end{aligned} \tag{21}$$

$$\begin{aligned} F_D &= (1-a_p)G_p(U+L,D,Z)(U+L) + (1-a_d)G_d(U+L,D,Z)D \\ &\quad -G_d(U+L,D,Z)D + m_p(U+L) \\ &\quad -f_n\mu_dD - (1-f_n)\mu_dD - v_d\frac{\partial D}{\partial z}, \end{aligned} \tag{22}$$

$$\begin{aligned} F_{DOM} &= (1-f_n)\gamma\rho_p(I)(L_{NO_3}+L_{NH_4})C + (1-f_n)\mu_zZ \\ &\quad +(1-f_n)\mu_dD - \mu_{dom}DOM, \end{aligned} \tag{23}$$

where

$$G_p = G_p(U+L,D,Z), \tag{24}$$

and

$$G_d = G_d(U+L,D,Z). \tag{25}$$

Let us make a few comments on some important modelled processes:

- **nutrient uptake**.
  Contrary to the BIOLOV model not only nitrate but also ammonium is represented in the model. However the uptake rate still uses a Monod kinetics as in the BIOLOV model,

  $$L_{NO_3} + L_{NH_4} = \frac{NO_3 + NH_4}{k_n + NO_3 + NH_4}.$$

  Concerning the BIOLOV model the new variable C appears in the formulation of the uptake rate,

  $$\rho_p(I)(L_{NO_3} + L_{NH_4})C.$$

  However in the BIOLOV model the maximum uptake rate is a constant, ρ_m, and the uptake rate does not depend explicitly on light. Numerical sensitivity studies reinforced by light dark chemostat experiments suggested that this is a too raw approximation in our 1D DYFAMED context. We therefore choose to allow light dependence of the maximum uptake rate through the following formulation:

  $$\rho_p(I) = \frac{\rho_m}{K_{I2} + I}$$

- Concerning the BIOLOV model, light dependence appears in the chlorophyll synthesis and in the photosynthesis rates (Eqs 19 and 20).
- Concerning the NNPZDDOM model phytoplankton organic and inorganic exudation are proportional to the uptake of nitrate and ammonium (Eq. 18).
- **Phytoplankton mortality and zooplankton grazing**.
  Phytoplankton mortality appears in the 3 equations for U, L and C. The mortality rate, m_p, is the same for each of these variables.
  In the same way, zooplankton grazing on phytoplankton now appears in each of the 3 equations for U, L and C. The grazing rate $G_p(U + L, D, Z)$ is the same for each of these variables ($U + L$ represents total phytoplanktonic nitrogen).

The carbon/chlorophyll ratio in $mgC.mgChla^{-1}$ (computed as $12f\frac{C}{L}$) is now a diagnostic variable of the model and evolves dynamically in space and time. Chlorophyll data can be directly compared to the variable $f\,L$ of the model, whereas in former studies phytoplankton and chlorophyll were related through a linear relation, $\frac{12Pr_d}{r_c}$ (where r_d is the C:N ratio and r_c the C:Chl ratio), in which the poorly known C:Chl ratio played a crucial role.

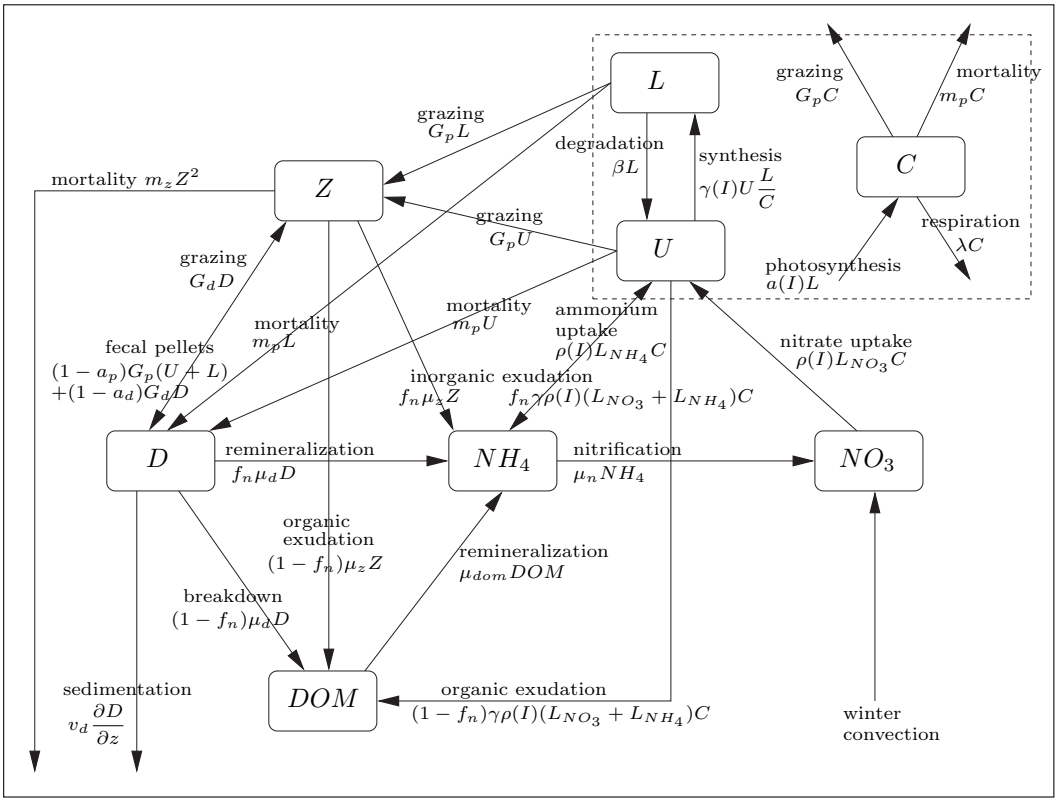

Figure 4: Schematic representation of the compartments and processes of the new model.

parameter	name	value	unit
half-saturation const. for nutrients	k_n	0.3208	$mmolNm^{-3}$
phyto. exudation fraction	γ	0.0649	
zoo. nominal preference for phyto food	r	0.6279	
max. specific zoo. grazing rate	g_z	1.2152E-5	s^{-1}
half-saturation const. for grazing	k_z	1.291	$mmolN.m^{-3}$
non-assimilated phyto. by zoo.	a_p	0.333	
non-assimilated detritus by zoo.	a_d	0.555	
zoo. specific exudation rate	μ_z	8.120E-7	s^{-1}
phyto. mortality rate	m_p	1.354E-6	s^{-1}
zoo. mortality rate	m_z	1.058E-7	$mmolN.m^{-3}.s^{-1}$
detritus breakdown rate	μ_d	1.352E-6	s^{-1}
detritus sedimentation speed	v_d	4.1122E-5	$m.s^{-1}$
decay rate below the euphotic layer	τ	5.5738E-5	s^{-1}
ratio of inorganic exudation	f_n	0.7488	
nitrification rate	μ_n	3.0213E-7	s^{-1}
slow remineral. rate of DOM to NH_4	μ_{dm}	4.8675E-8	s^{-1}
coeff. for Martin's remineralization profile	h_r	-0.858	
pigment absorption in red	k_{rp}	0.037	$m^{-1}.(mgChl.m^{-3})^{-l_r}$
pigment absorption in green	k_{gp}	0.074	$m^{-1}.(mgChl.m^{-3})^{-l_g}$
power law for absorption in red	l_r	0.629	
power law for absorption in green	l_g	0.674	
contribution of Chl to absorbing pigments	r_{pig}	0.7	
water absorptions in green	k_{go}	0.0232	m^{-1}
water absorptions in red	k_{ro}	0.225	m^{-1}
max. uptake rate	ρ_m	8.6805E-6	$mmolN.mmolC^{-1}.s^{-1}$
half-saturation const. for uptake	K_{I2}	104.25	$\mu mol\ quanta.m^{-2}.s^{-1}$
threshold coefficient	K_C	33.0	$\mu mol\ quanta.m^{-2}.s^{-1}$
max. rate for carbon	α	2.7894 E-4	s^{-1}
max. synthesis rate	α_K	1.800E-3	s^{-1}
half-saturation const. for fixation	K_I	208.5	$\mu mol\ quanta.m^{-2}.s^{-1}$
respiration rate	λ	6.25 E-7	s^{-1}
degradation rate	β	7.9862 E-7	s^{-1}
L/Chl ratio	f	0.43	mmolN.m-3/mgChl.m-3

Table 3: 33 parameters of the new model

3 Calibration of the new model using data assimilation

3.1 The DYFAMED data set

The model general set up is the same as in [19] and [7]. The standard run consists of the simulation of year 1997. The simulation is forced with ECMWF atmospheric data, which give the wind stresses and heat fluxes every 6 hours.

The stock data used in this study are monthly chlorophyll and nitrate profiles collected during year 1997 at the DYFAMED station, and surface fluorescence data measured by the Carioca buoy [14] moored at DYFAMED in 1997 (Figs. 5, 6, 7). Following the conclusions of [7] concerning the necessity to use fluxes data as well as zooplankton data in the parameter estimation process, we also make use of annual estimates of total production (TP), nitrate uptake (NU) and export fluxes (EF, estimated from disparate measurements undertaken in the 90s), and of the zooplankton content estimated during a special cruise in May 1995 [1].

3.2 Parameter estimation

The purpose of parameter estimation is to find a set of optimal parameters that minimises a cost function, J, which measures the distance, in a weighted least-squares sense, between the model's solution and the observations. The optimisation is carried using the quasi-Newton algorithm implemented in the $n1qn3$ Fortran subroutine of Gilbert and Lemaréchal [12]. The computation of the gradient of J with respect to control parameters is required at each step of the minimisation. This gradient results in one integration of the adjoint model. The adjoint code was partially obtained using the automatic differentiation program Odyssée ([8]; [13]), which is an efficient tool for obtaining adjoint codes since it enables the automatic production of adjoint instructions. This so called variational adjoint method has already been applied in marine biogeochemistry by several authors (see [16] for example).

As shown in Table 3, the biological parameters have very different orders of magnitude. To avoid any numerical difficulties which might arise from this during the minimisation, we adimensionalise the parameter vector K, dividing each parameter K_i by its first guess value K_i^0, $k_i = K_i/K_i^0$. Such a non-dimensionalisation procedure can be regarded as a preconditioning for minimisation. The control variable is k of size p, where p is the number of parameters, and k is dimensionless.

The model-data misfit part, J_0, of the cost function can be written as the sum of five terms:

$$J_0(k) = J_{sc}(k) + J_{cp}(k) + J_{np}(k) + J_{zoo}(k) + J_{flux}(k).$$

If we denote by ϕ the direct operator which, to a set of parameter k, associates the model equivalents to the data and the data by d, then J_{sc} the cost related to the nsc surface chlorophyll observations, dsc reads:

$$J_{sc}(k) = \frac{1}{2}\sum_{i=1}^{ncs} w_{sc}(\phi_{sc}(k)_i - dsc_i)^2.$$

J_{cp} is related to the ncp chlorophyll profile observations, dcp (mcp measurement points on the vertical),

$$J_{cp}(k) = \frac{1}{2}\sum_{i=1}^{ncp}\sum_{l=1}^{mcp} w_{cp}(\phi_{cp}(k)_{i,l} - dcp_{i,l})^2.$$

J_{np} is related to the nnp NO_3 profile observations, dnp (mnp measurement points on the vertical),

$$J_{pn}(k) = \frac{1}{2}\sum_{i=1}^{nnp}\sum_{l=1}^{mnp} w_{np}(\phi_{np}(k)_{i,l} - dnp_{i,l})^2.$$

J_{zoo} is related to the single zooplankton data,

$$J_{zoo}(k) = \frac{w_{zoo}}{2}(\phi_{zoo}(k) - dzoo)^2.$$

J_{flux} is related to the TP, NU and EF data estimates.

$$J_{flux}(k) = \frac{w_{tp}}{2}(\phi_{tp}(k) - dtp)^2 + \frac{w_{nu}}{2}(\phi_{nu}(k) - dnu)^2 + \frac{w_{ef}}{2}(\phi_{ef}(k) - def)^2.$$

The different weights w are composed of the square of the assumed a priori observation errors and of a scaling factor accounting for the number of each type of observation. Therefore, we have $w_{cs} = \dfrac{1}{ncs.\sigma_{cs}^2}$, $w_{cp} = \dfrac{1}{(ncp.mcp)\sigma_{cp}^2}$, and $w_{np} = \dfrac{1}{(nnp.mnp)\sigma_{np}^2}$.

We also add two penalty terms to J_0. The first term,

$$J_p(k) = \frac{1}{2}\sum_{i=1}^{p} w_i(k_i - k_i^0)^2,$$

accounts for the a priori parameter values and their standard deviations, $\sigma_i = \frac{1}{\sqrt{w_i}}$. This term forces the minimisation to avoid biologically absurd optimal parameter values.

The second term,

$$J_{rc}(k) = \frac{w_{rc}}{2}\sum_j\sum_n((12f(C/L)_{j,n} - 250)^+)^2,$$

where n refers to time and j to space, is added to penalise carbon/chlorophyll ratio values greater than 250 $mgC.mgChla^{-1}$ which is considered as an upper bound by biologists. The choice of the weight w_{rc} is not straightforward, since a compromise has to be found between the quality of the minimisation of J_0 and realistic bound for the carbon/chlorophyll ratios. In practice, w_{rc} was chosen so that the data misfit part of the cost and its regularisation part are balanced.

4 Numerical results

4.1 Model fit to the DYFAMED data

An a priori parameter set is constructed in the following way: parameters which were already included in the NNPZD-DOM model are given the values estimated using data from the DYFAMED station in [7]. Parameters coming from the BIOLOV model are given the values proposed in [20]. A few parameters were then hand-tuned empirically in order to obtain a qualitatively correct first guess simulation. Note that since parameter K_L (Tab. 2) only enters the equations of the BIOLOV model in the αK_L product form (Eq. 14), for the sake of parameter estimation those two parameters are concatenated in a single parameter α_K.

A first run is carried out using the a priori parameter set and without any data assimilation (Figs. 5, 6, 7). The seasonal variability is well reproduced, but the spring bloom appears to be too strong and the transition towards oligotrophy poorly represented. This first guess run predicts reasonably good TP and NU fluxes estimates (Tab. 4). On the other hand, the export flux is underestimated and is associated with an underestimation of the zooplankton content (Tab. 5).

The evolution over time of the chlorophyll and nitrate data profiles (Figs. 6, 7) reflects the seasonal variability at DYFAMED. Winter mixing brings nutrients to the surface, but the short residence time of algae in the euphotic layer, swept along by strong vertical motions, prevents the development of biomass. As the year progresses, the surface layer becomes more stable, thus allowing the winter nutrient enrichment to be utilised continuously and finally allowing the algae bloom. From mid-May to November, the situation remains fairly stable with the upper layer nutrient content very low, and the system mainly oligotrophic and characterised by a deep chlorophyll maximum. It may be noticed that the nitrate profiles show strong variability below 100m. This variability cannot be attributed to biological processes since these occur closer to the surface, and it is therefore most likely due to horizontal advection. It cannot therefore be captured by the model. During the oligotrophic period, the location of the nitracline and the deep chlorophyll maximum is fairly constant. This may be an indication of the absence of strong Ekman pumping and a crude justification for the neglect of vertical advection.

In order to calibrate the simulation we estimated 25 parameters among the 33 parameters of the model. Following the sensitivity analysis conducted in [7] which proved that the optical model parameters as well as parameter h_r can not be recovered from the DYFAMED data set, we did not include them in the control vector. The optimal parameter set is shown on Table 6.

The parameter estimation procedure resulted in several improvements in the simulation. Bloom intensity is divided by a factor of two, in agreement with the observations (Fig. 5). Chlorophyll profiles during the bloom (dates 03/02 and 03/21 on Fig. 6) also show a better fit with the data. The subsurface chlorophyll maximum (dates 06/19, 07/12 and 09/01) is well simulated. Clearly this is an improvement compared to the classical NNPZD-DOM model which gave poor results in the simulation of this oligotrophic regime. Concerning NO_3 profiles the improvement is less significant. This problem already occurred in former studies and was attributed to the fact that these profiles are affected by 3D physics not taken into account in the model. Data assimilation also forces the model to predict correct zooplankton quantities (Tab. 5) and correct fluxes estimates (Tab. 4). Concerning the total production flux (TP), it is worth to notice that if in the NNPZD-DOM model it is computed in nitrogen unit and then converted to carbon unit using a constant C:N ratio, in the new model this flux is directly computed in carbon unit through the formulation

$$TP = a(I)L - \lambda C \ (mmolC/m^3/s)$$

This formulation can also lead to negative production values when respiration is stronger than photosynthesis. This situation occurs at least every night in the model when $a(I) = 0$.

Figure 8 shows the C:Chl ratio profiles at the same dates than the Chl profiles of Figure 6. The model simulates values which are in accordance with what one should have expected. Globally the ratio decreases with depth that is to say with light. This is certainly a consequence of the BIOLOV model's structure which was designed to simulate these features observed experimentally at steady state ([20], [11], [2]). The bloom and post-bloom period correspond to relatively low C:Chl ratio values (lower than 100 $mgC/mgChl$) and relatively homogeneous profiles. On the contrary during summer when the oligotrophic regime occurs surface values are high (between 200 and 250 $mgC/mgChl$) and the ratio rapidly decreases with depth from 0 to about 50 m which corresponds to the depth of the subsurface chlorophyll maximum.
With this new model the C:N ratios, computed as $C/(U + L)$, evolves dynamically. Figure 9 shows the C:N ratio profiles at the same dates than the Chl profiles of Figure 6. As for the C:Chl ratio, it strongly varies with depth and time which is not

at all the usual hypothesis made in such a biogeochemical model where this ratio is taken equal to 6.625 $mmolC/mmolN$. This variability is consistent with experimental observations [5, 21], even if the values reach sometimes too low thresholds probably due to the higher carbon losses due to respiration than nitrogen losses in strong light limitation conditions.

4.2 Variability of the predicted ratios

In order to better assess the results of our model, and to first discuss the variability of the C:Chl ratio we compare the obtained values with those computed as a function of light, as proposed in ([4], [17]). With obvious notations this parametrisation is written as follows:

$$R_{C:Chl}(I) = 12.R_{C:N}/R_{Chl:N}(I), \tag{26}$$

where $R_{C:N}$ (in $molC/molN$) is fixed and

$$\begin{aligned} R_{Chl:N}(I) &= [R^{min}_{Chl:N} - (R^{max}_{Chl:N} - R^{min}_{Chl:N})I/I_{max}] \quad \text{if } I < I_{max}, \\ &= R^{min}_{Chl:N} \qquad\qquad\qquad\qquad\qquad\qquad\quad \text{if } I \geq I_{max}. \end{aligned} \tag{27}$$

where $R^{min}_{Chl:N}$, $R^{max}_{Chl:N}$ and I_{max} are fixed parameters.

Figure 10 shows the C:Chl ratio computed with our model. The light intensity was discretised in intervals of 10 W/m^2. In each interval the value of the considered ratio was computed so as its standard deviation. Two other C:Chl are also plotted on the same figure: the one that would have been obtained with parameterisation (26), and the one obtained in [7]. This latter which was originally computed as function of depth was transposed to a function of light on the basis of mean light intensity in the water column. This is the reason why no value is given for large values of light intensity which do not occur often in the simulation. The first remark is that the general form of the curve obtained with (26) fits the data cloud issued from the model. It is very probable that adjusting the parameters provided in ([4], [17]) we would be able to have a good adjustment. The second remark is that the curve obtained in [7] fits the data cloud even better. However, the main outcome of this comparison is the huge variability of this ratio as predicted by our model. The variation can easily reach 100% of the predictions meaning thus that the coupling between carbon and chlorophyll can be very loose.

Now Figures 11 shows the C:N ratio, and highlight the high variability of this ratio which is probably the main cause of the variability of the C:Chl ratio (N/Chl being less variable).

4.3 Variability of the predicted fluxes

The inter-monthly variability of the fluxes has been estimated for both models as presented on Figures 12 and 13. First it appears that this variability is much higher close to the surface, with a maximum for the nitracline in summer. The fluxes vary generally much more for our model than for the previous model, especially at the surface. It is worth noting that the total production is approximately the same for both model since this data has been used during the assimilation step to constrain the data.

This higher variability probably reflects the ability of the phytoplankton compartment in BIOLOV to promptly react to optimal conditions for carbon incorporation. This faculty results from the partial uncoupling between nitrogen and carbon incorporation as described by the BIOLOV model.

5 Discussion : benefits of a more complex model

We have seen in [7] that the assimilation results could not be satisfying if the C:Chl *a* ratio was kept constant for the Monod type phytoplankton models. The introduction of an empirical C:Chl *a* ratio parameterized by depth led to good results, especially when the fluxes data were assimilated.

In our approach the idea consisted in providing a mechanical parameterization to this ratio by taking into account the nonlinear coupling between limitation by light and by nitrogen. For this we used the BIOLOV model that was validated in chemostat experiments.

It is worth noting that even if the models are comparable, the involved mechanisms are really different since in the new model the algal physiology is taken into account especially by integrating the mechanisms of photoadaptation. As a consequence, the new model needs 8 state variables (3 of them representing phytoplankton) and 33 parameters whereas the previous model needs 6 state variables (one of them being phytoplankton) and 28 parameters plus the 20 parameters used to discretized the C:Chl *a* ratio along depth.

The first conclusion of our study was that the obtained model reproduces the reference data in a comparable manner than the previous model. Both fluxes and stocks can be equally well predicted by both models.

However if the models are coinciding on an average basis, they are diverging from a variability prediction point of view. In the new model biology adapts much faster to its environment giving rise to higher short term variations and finally to a higher variability.

As we have shown with *e.g.* C:Chl *a* ratio versus light intensity the average model predictions remain in line with other empirical formulations.

At this point it can appear rather surprising that both simple and complicated model can reproduce accurate average fluxes. Of course the benefit of the mechanical more complex model is to provide a thinner time resolution of the phenomena.

Now the question of a possible higher variability on the variables and fluxes should be discussed in future studies. Especially the possible implication on a sampling strategy and data interpretation.

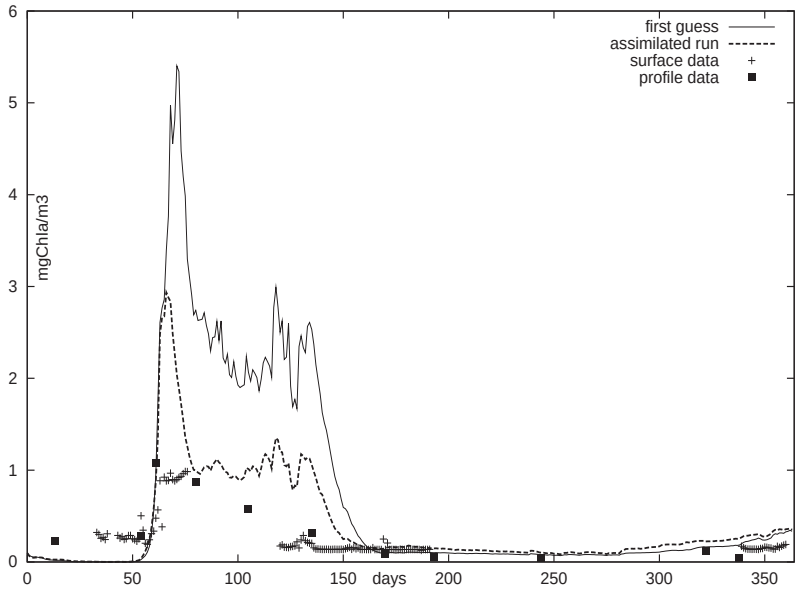

Figure 5: Surface chlorophyll ($mgChla/m^3$). Crosses are the surface observations. Squares are the surface measurements taken from the chlorophyll profile set (Fig. 6).

flux	data	first guess	opt
NU	42 ± 15	48.7	47.2
TP	156 ± 30	92.8	106.3
EF	2 ± 0.8	0.2	2.5

Table 4: Flux data (from Marty and Chiaverini, 2002) and calculated fluxes at 200m. Values are given in $gC.m^{-2}.y^{-1}$. NU stands for NO_3 uptake, TP for total production, EF for export flux.

data	first guess	opt
20 ±10	1.3	19.1

Table 5: Zooplankton content data and calculated values at 200m, for May 1997, in $mmolNm^{-2}$.

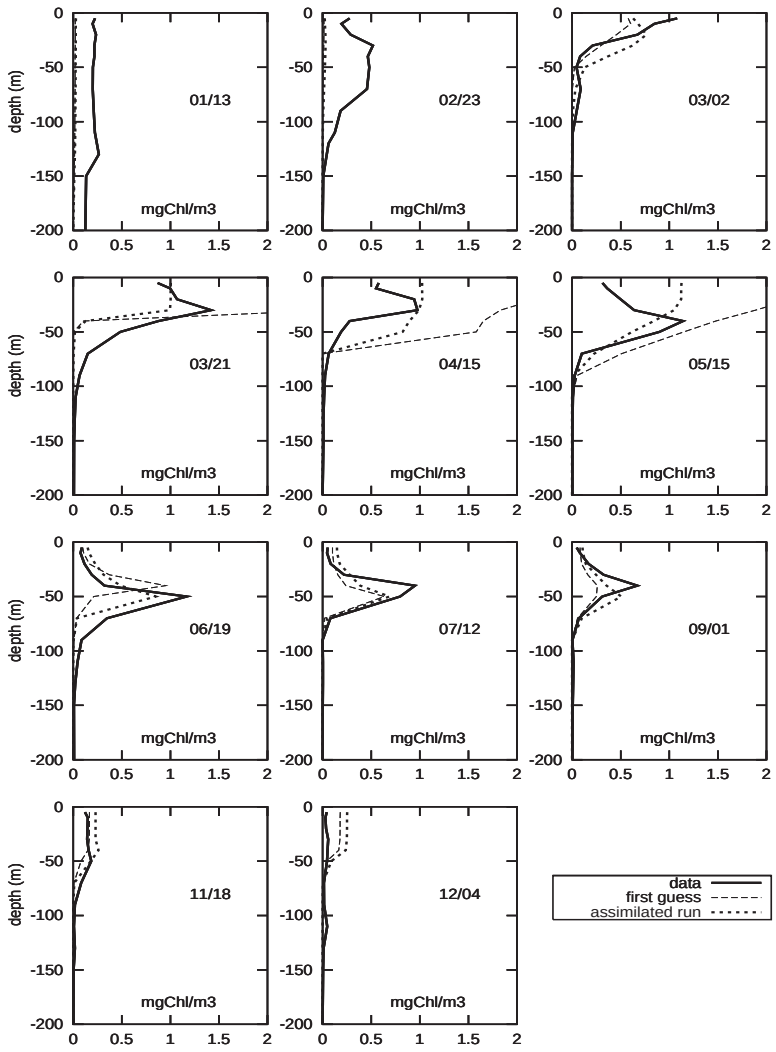

Figure 6: Chlorophyll profiles ($mgChla/m^3$). Depth in meters.

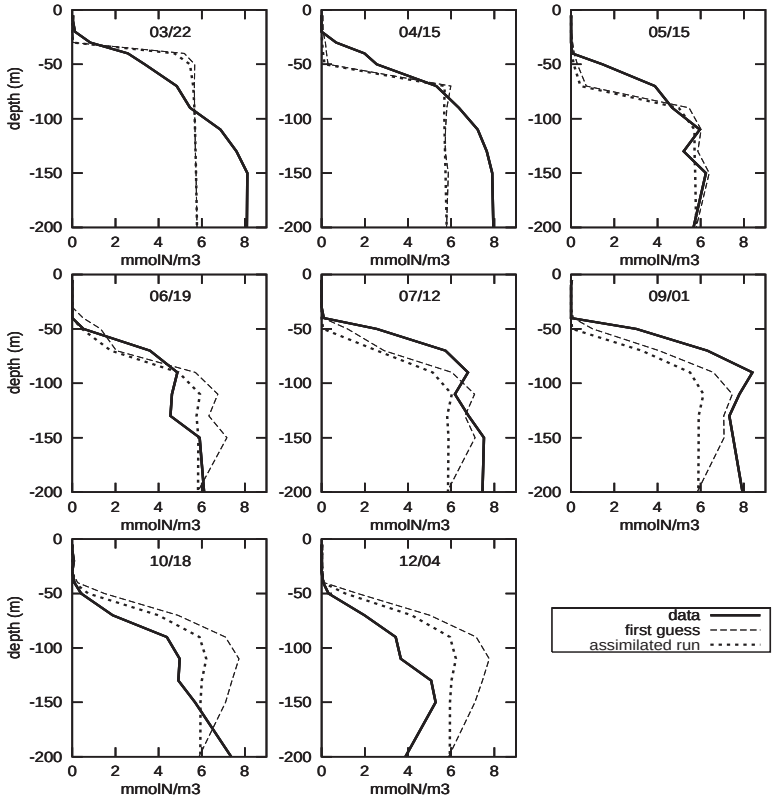

Figure 7: NO3 profiles ($mmolN/m^3$). Depth in meters.

parameter	optimal set k^{opt}
k_n	0.944
γ	0.971
r	0.404
g_z	1.314
k_z	1.135
a_p	0.855
a_d	0.802
μ_z	1.220
m_p	1.394
m_z	0.850
μ_d	0.867
v_d	0.347
τ	1.003
f_n	0.947
μ_n	0.993
μ_{dm}	1.299
ρ_m	1.335
K_{I2}	0.590
K_C	0.740
α	1.218
α_K	0.464
K_I	1.041
λ	1.121
β	1.059
f	0.948

Table 6: Optimal adimensionalized (K/K^0) parameters

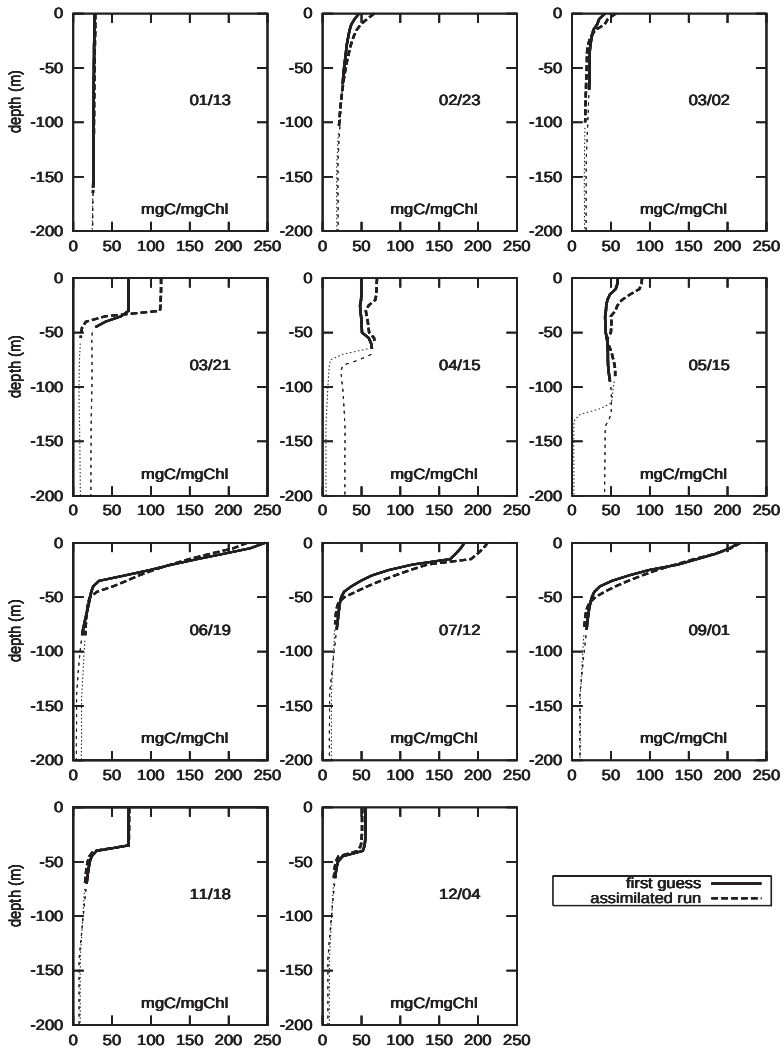

Figure 8: Mean carbon to chlorophyll ratio profiles for year 1997 at DYFAMED. Thin dashed lines indicate that the algal biomass is very weak, and thus that the ratio is calculated with very small values of C or Chl.

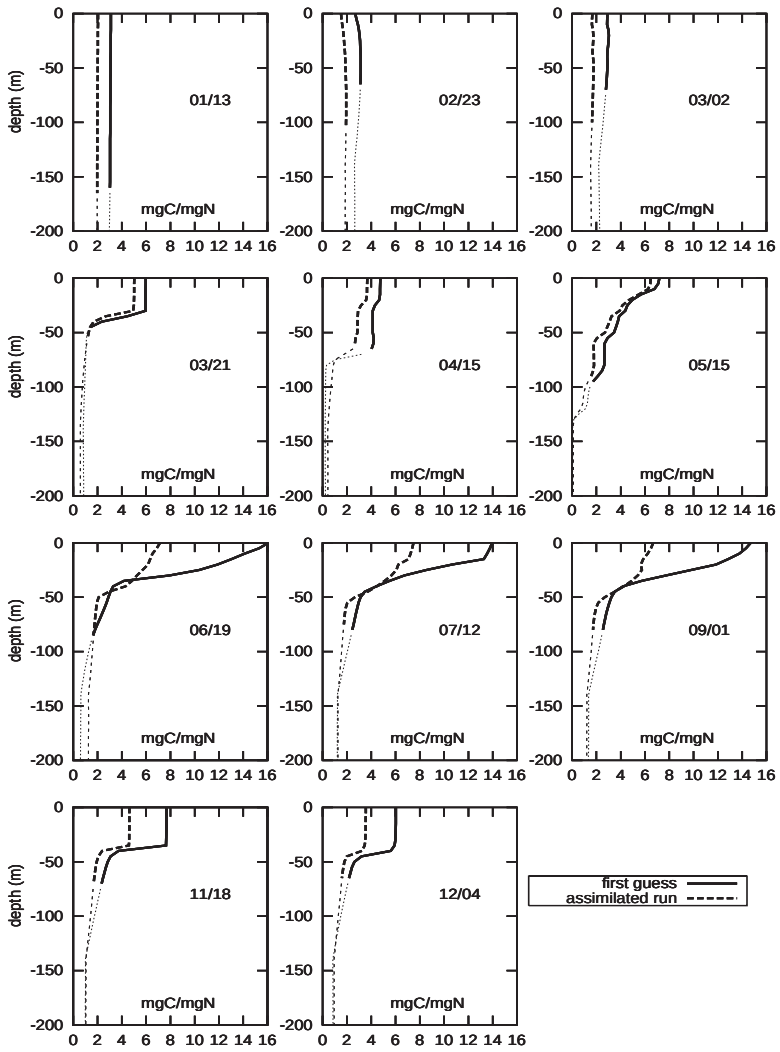

Figure 9: Mean carbon to nitrogen ratio profiles for year 1997 at DYFAMED. Thin dashed lines indicate that the algal biomass is very weak, and thus that the ratio is calculated with very small values of C, Chl or N.

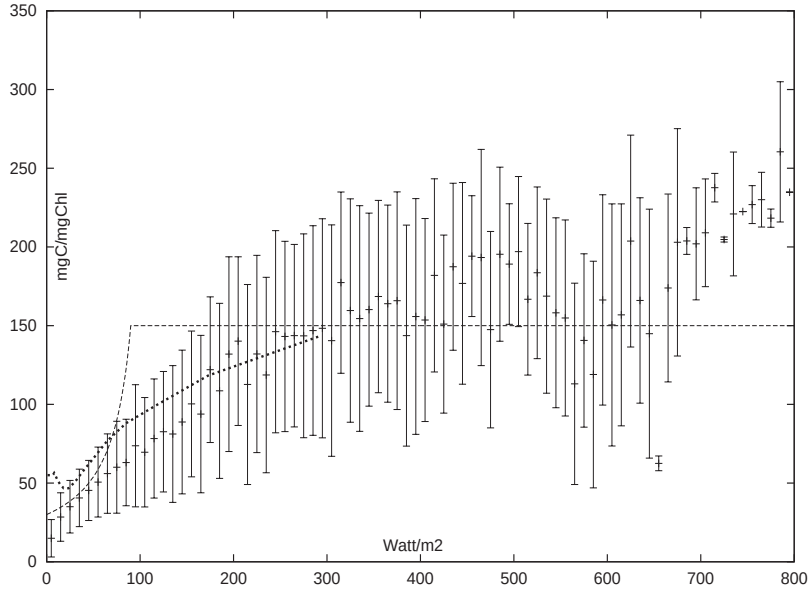

Figure 10: Carbon to chlorophyll ratio versus light intensity. Mean values and standard deviation by interval of light intensity of length 10 $Watt/m^2$ calculated for the whole water column during year 1997. The thin dashed line corresponds to the parameterisation of this ratio as function of light given by Eqs. 26 and 27. The bold dashed line corresponds to the ratio obtained in [7] plotted against the mean light intensity in the water column.

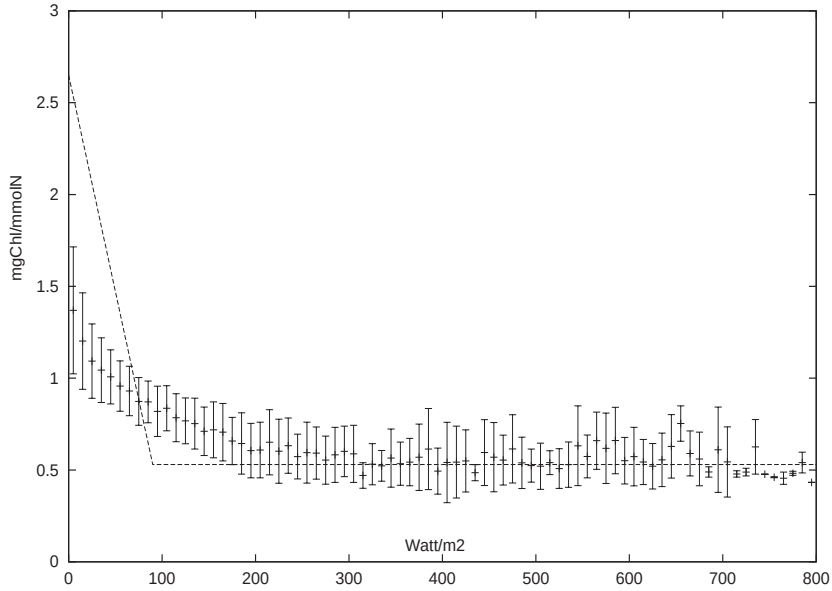

Figure 11: Carbon to nitrogen ratio versus light intensity. Mean values and standard deviation by interval of light intensity of length 10 $Watt/m^2$ calculated for the whole water column during year 1997. The dashed line corresponds to the parameterisation of this ratio as function of light given by Eqs. 26 and 27.

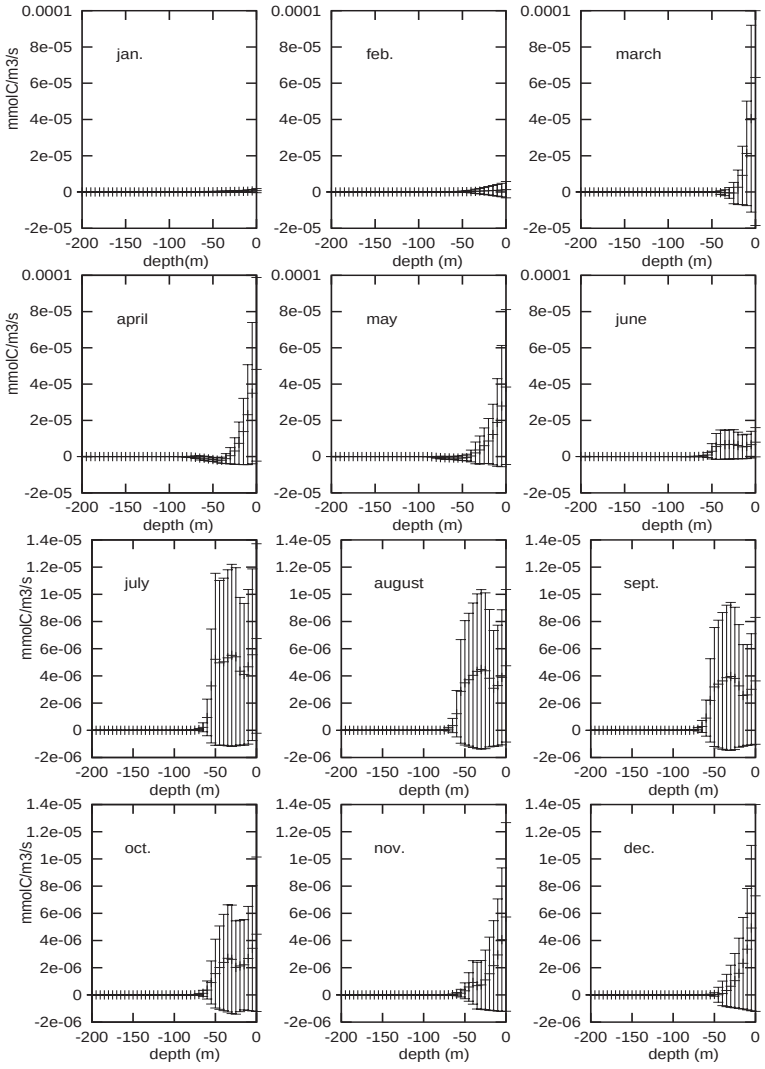

Figure 12: Monthly total production profiles for the new model.

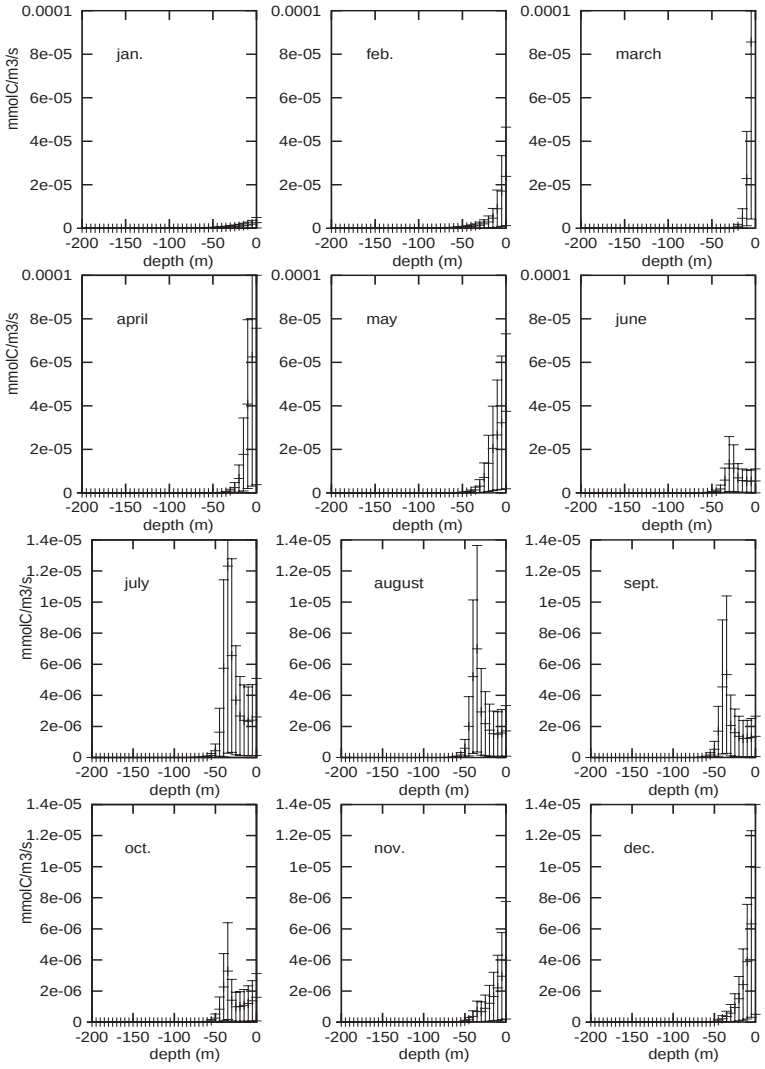

Figure 13: Monthly total production profiles for the NNPZD-DOM model.

References

[1] V. Andersen and L. Prieur. One-month study in the open NW Mediterranean Sea (DYNAPROC experiment, May 1995): overview of the hydrobiologeochemical structures and effects of wind events. *Deep-Sea Research I*, 47:397–422, 2000.

[2] M.S. Chalup and E.A. Laws. A test of the assumptions and predictions of recent microalgal growth models with the marine phytoplankter. *Pavlova Luthri. Limnol. Oceanogr.*, 35(3):583–596, 1990.

[3] J.J. Cullen. On models of growth and photosynthesis in phytoplankton. *Deep-Sea Research*, 37:667–683, 1990.

[4] S.C. Doney, D.M. Glover, and R.G. Najjar. A new coupled, one-dimensional biological-physical model for the upper ocean:Applications to the JGOFS Bermuda Atlantic Time-series Study (BATS) site. *Deep-Sea Research II*, 43(2-3):591–624, 1996.

[5] M.R. Droop. Vitamin B12 and marine ecology. IV. the kinetics of uptake growth and inhibition in *monochrysis lutheri*. *J. Mar. Bio. Ass. U.K.*, 48:689–733, 1968.

[6] M.J.R. Fasham, H.W. Ducklow, and S.M. McKelvie. A nitrogen-based model of plankton dynamics in the oceanic mixed-layer. *J. Mar. Res.*, 48:591–639, 1990.

[7] B. Faugeras, M. Lévy, L. Mémery, J. Verron, J. Blum, and I. Charpentier. Can biogeochemical fluxes be recovered from nitrate and chlorophyll data? A case study assimilating data in the Northwestern Mediterranean Sea at the JGOFS-DYFAMED station. *J. Mar. Sys.*, 40-41:99–125, 2003.

[8] C. Faure and Y. Papegay. Odyssée Version 1.6, the language reference manual. Rapport Technique 211, INRIA, 1997.

[9] P. Gaspar, Y. Gregories, and J.M. Lefevre. A simple eddy kinetic energy model for simulations of the oceanic vertical mixing: tests at station Papa and long term upper ocean study site. *Journal of Geophysical Research*, 95:16179–16193, 1990.

[10] R.J. Geider, H.L. MacIntyre, and T.M. Kana. Dynamic model of phytoplankton growth and acclimation: responses of the balanced growth rate and the chlorophyll a:carbon ratio to light, nutrient-limitation and temperature. *MEPS*, 148:187–200, 1997.

[11] R.J. Geider, H.L. MacIntyre, and T.M. Kana. A dynamic regulatory model of phytoplanktonic acclimatation to light, nutrients and temperature. *Limnol. Oceanogr.*, 43(4):679–694, 1998.

[12] J.C. Gilbert and C. Lemaréchal. Some numerical experiments with variable storage quasi-newton algorithms. *Mathematical Programming*, 45:407–435, 1989.

[13] A. Griewank. *Evaluating Derivatives. Principles and Techniques of Algorithmic Differentiation*. Frontiers in applied mathematics. SIAM, 2000.

[14] E.M. Hood and L. Merlivat. Annual to interannual variations of f in the northwestern Mediterranean Sea : results from hourly measurements made by CARIOCA buoys, 1995-1997. *J. Mar. Res.*, accepted, 2002.

[15] G.C. Hurtt and R.A Armstrong. A pelagic ecosystem model calibrated with BATS data. *Deep-Sea Research II*, 43:653–683, 1996.

[16] L.M. Lawson, Y.H. Spitz, E.E. Hofmann, and R.B. Long. A data assimilation technique applied to a predator-prey model. *Bulletin of Mathematical Biology*, 57:593–617, 1995.

[17] M. Lévy, L. Mémery, and J.-M. André. Simulation of primary production and export fluxes in the Northwestern Mediterranean Sea. *J. Mar. Res.*, 56:197–238, 1998a.

[18] J.H. Martin and S.E. Fitzwater. Dissolved Organic Carbon in the Atlantic, Southern and Pacific Oceans. *Nat.*, 356:699–700, 1992.

[19] L. Mémery, M. Lévy, S. Vérant, and L. Mervilat. The relevant time scales in estimating the air - sea CO_2 exchange in a mid latitude region. *Deep-Sea Research II*, 49(11):2067–2092, 2002.

[20] L. Pawlowski, O. Bernard, E. Le Floc'h, and A. Sciandra. Qualitative behaviour of a phytoplankton growth model in a photobioreactor. In *International Federation of Automatic Control. 15th World Congress. Barcelone, Espagne.*, 2002.

[21] A. Sciandra, L. Lazzara, H. Claustre, and M. Babin. Responses of the growth rate, pigment composition and optical properties of *cryptomonas* sp. to light and nitrogen stresses. *MEPS*, 201:107–120, 2000.