



HAL
open science

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

Blaise Faugeras, Marina Lévy, Laurent Mémery, Jacques Verron, Jacques Blum, Isabelle Charpentier

► To cite this version:

Blaise Faugeras, Marina Lévy, Laurent Mémery, Jacques Verron, Jacques Blum, et al.. 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. *Journal of Marine Systems*, 2003, 40-41, pp.99-125. 10.1016/S0924-7963(03)00015-0 . hal-00230148

HAL Id: hal-00230148

<https://hal.science/hal-00230148>

Submitted on 10 Feb 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

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

Blaise Faugeras^{a,*}, Marina Lévy^b, Laurent Mémery^b, Jacques Verron^c, Jacques Blum^d, Isabelle Charpentier^a

^a*Laboratoire de Modélisation et Calcul, IMAG, Tour IRMA, 51 rue des Mathématiques, Domaine Universitaire, BP 53, 38041 Grenoble Cedex 09, France*

^b*LODYC, Université Pierre et Marie Curie, 5 Place Jussieu, 75252 Paris Cedex 05, France*

^c*Laboratoire des Ecoulements Géophysiques et Industriels, BP53X, 38041 Grenoble, France*

^d*Laboratoire J.A. Dieudonné, Université de Nice-Sophia-Antipolis, 06108 Nice Cedex 02, France*

One of the principal objectives of studying biogeochemical cycles is to obtain precise estimates of the main fluxes, such as total, new and export oceanic productions. Since models can incorporate the a priori knowledge of the most important processes, they are increasingly used for this purpose. However, biogeochemical models are characterized by a large number of poorly known parameters. Moreover, the available data are rather sparse in both time and space, and represent concentrations, not fluxes. Therefore, the major challenge is to constrain the relevant fluxes using information from a limited number of observations and from models incorporating poorly known internal parameters.

The present study attempts to meet this challenge. In a 1D framework at the DYFAMED station (NW Mediterranean Sea), near-monthly nitrate and chlorophyll profiles and daily surface chlorophyll concentrations are assimilated in a coupled dynamical–biological model using the tangent linear and adjoint models. Following sensitivity analyses that show that some parameters cannot be recovered from the data set used, assimilation of observed 1997 data is performed. The first inversion considered clearly shows that, in agreement with previous studies, (1) the data impose a C/Chl ratio that varies with depth (i.e. light) and (2) the “initial” conditions (e.g. winter nitrate profile) strongly constrain the annual biogeochemical fluxes. After assimilation of the 1997 data, the agreement between the data and the model is quantitatively improved in 1995 and 1996, which can be considered a good validation of the methodology. However, the order of magnitude of the biogeochemical fluxes, and especially of the particulate export and regenerated production, are not correctly recovered. An analysis of the simulations shows that this result is associated with a strong decrease in zooplankton concentrations. An additional constraint of maintaining acceptable levels of zooplankton is therefore added. The results are improved, but remain unsatisfactory. A final inversion, which takes into account the a priori estimates of the major annual fluxes, is then performed. This shows that there is no

* Corresponding author.

E-mail address: faugeras@imag.fr (B. Faugeras).

inconsistency between the NO_3 and chlorophyll data, the order of magnitude of the fluxes and the model. The work therefore demonstrates that recovering biogeochemical fluxes from available data of concentrations and stocks is not a straightforward exercise: the coverage and type of observations, and the nonlinearities of the biogeochemical model all contribute to this difficulty.

Keywords: Biogeochemical model; Parameter optimization; DYFAMED

1. Introduction

Together with ocean dynamics, the main biogeochemical fluxes such as total, new and export productions, control the total inorganic content of the mixed layer, which in turn affects carbon dioxide exchanges between the ocean and the atmosphere. The study of biogeochemical cycles endeavours to estimate these fluxes accurately. Modeling approaches have the advantage of being able to make use of a priori knowledge, and are increasingly used for such purposes. But the solutions of biogeochemical models depend to a large extent on the values of the biological parameters. These parameters are not only numerous, but most of them are also poorly known since they often represent unmeasurable variables used in the parameterization of complex exchanges between biological compartments. Moreover, these compartments represent the averaged behavior of different biological species. Traditionally, the parameters are tuned until the best agreement between model results and data is reached. Data assimilation techniques provide a powerful means to achieve this tuning, and are becoming a mandatory step in model development. However although theoretically appealing, the practical use of data assimilation is a difficult task, and exploratory studies are needed before data assimilation can be routinely used in 3D biogeochemical models.

The available biogeochemical data are rather sparse in time and space, and often represent concentrations and not fluxes. Indeed, biogeochemical fluxes are difficult to measure on a routine basis. Primary production measurements require in situ incubations, while export measurements are generally evaluated from the deployment of sediment traps. Moreover, the error associated with these measurements is quite high (Richardson, 1993;

Buesseler et al., 1992, 2000). Nitrate and chlorophyll data, on the other hand, are more accessible, more widespread, and more precise.

The major challenge is therefore to try to constrain the biogeochemical fluxes using information from a limited number of stock observations and nonlinear models associated with poorly known internal parameters. Such an attempt is presented in this study. In a 1D framework at the DYFAMED station, near-monthly nitrate and chlorophyll profiles and daily surface chlorophyll concentrations are assimilated in a coupled dynamical–biological model using the tangent linear and adjoint models.

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 set 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., Part 2, special issue, 29 (12), 2002, in press), and has been the subject of previous model studies (Lévy et al., 1998a; Mémerly et al., 2002).

Several inverse methods have been applied to the estimation of biogeochemical parameters. These methods are all based on the minimization of a cost function that depends on the parameters to be estimated and measures the distance between the observations and the corresponding values calculated by

the model. But as pointed out by Athias et al. (2000), the result of the assimilation may depend on the choice of the inverse method used to find the minima of the cost function. It is thus of primary importance to choose an adequate method and test its performance. Most previous studies have used a variational data assimilation approach, based on a linearization of the cost function around an a priori solution (Prunet et al., 1996a,b), and a gradient descent method (Fennel et al., 2001; Gunson et al., 1999; Lawson et al., 1995, 1996; Schartau et al., 2001; Spitz et al., 1998, 2001). Other studies have used nonlinear optimization techniques (Evans, 1999; Fasham and Evans, 1995) or global optimization algorithms such as simulated annealing (Athias et al., 2000; Hurtt and Armstrong, 1996, 1999; Matar, 1995).

Mathematical arguments favor the choice of nonlinear techniques. Indeed, the inverse parameter estimation problem is always a nonlinear problem, even in the case of linear dynamics (Evensen et al., 1998). Therefore, the cost function may possess different local minima. There are, however, three reasons why, in our opinion, the optimal control method, often referred to as the adjoint method, is promising. First of all, global optimization algorithms require a large number of cost function evaluations and this does not seem possible using a full three-dimensional model, which is the final goal of ocean color data assimilation. Secondly, the adjoint method is already being used with some success in ocean global circulation models, which is a first step towards its implementation in 3D biogeochemical models. Thirdly, the a priori parameter values or “first guess” values required in the adjoint method are not completely unknown, even if they are attributed large error bars. It therefore seems natural to treat the inverse problem locally around these a priori values. The ideas and mathematical concepts of optimal control theory were formalized about 30 years ago (Lions, 1971) and have since received a lot of attention with respect to their applications in meteorology (Le Dimet and Talagrand, 1986; Courtier and Talagrand, 1987; Talagrand and Courtier, 1987) and oceanography (Long and Thacker, 1989a,b; Luong et al., 1998).

For simplicity, most previous biogeochemical data assimilation studies have considered 0D model set-ups, in which biogeochemical variables corre-

spond to mixed-layer values. This approach is suitable when most biogeochemical activity occurs within the mixed layer. This is the case in the spring phytoplankton bloom, provided that the upload of nutrients is correct. It is, however, not the case for oligotrophic regimes, where the chlorophyll maximum is below the mixed layer. The novelty of our approach is in the use of the adjoint technique in a 1D context to assimilate observed data, as well as the first use of the DYFAMED data set for data assimilation purposes.

The biogeochemical model used in this study comprises six nitrogen compartments. In a study where data from the Bermuda Atlantica Time Series (BATS) station are assimilated, Spitz et al. (2001) warn that data cannot be satisfactorily assimilated if the structure of the biogeochemical model is not adequate. We follow the procedure adopted by Spitz et al. (2001) and use the results of a first assimilation attempt to modify the model. We show that a significant improvement in model performance is obtained when the C/Chl ratio is allowed to vary in the vertical, and also when the winter nitrate upload is estimated.

The paper is organized as follows: after a presentation of the model and the method, sensitivity analyses are carried out to determine which parameters can be recovered from the data set used. Then an assimilation of observed data for the year 1997 is performed. After the assimilation, some of the fluxes (e.g. particulate export production) are shown to be very poorly estimated.

An analysis of the simulation shows that zooplankton values have fallen to a very low level, which would explain this result. An additional constraint is then added: the zooplankton stock during the post bloom regime (Andersen and Prieur, 2000). Although the results are improved, they are still not satisfactory. The method is then validated using independent data from years 1995 to 1996. A final inversion is then carried out that also takes into account the a priori estimates of the major annual fluxes (total and new productions, and particulate and dissolved exports). This inversion shows that there is no inconsistency between the nitrate and chlorophyll data, the estimates of the fluxes and the model used. The difficulty of recovering the biogeochemical fluxes from real data assimilation of

concentrations and stocks (nitrate, chlorophyll, and eventually zooplankton) is then discussed.

2. Description of the model and data

2.1. Model description

The biogeochemical model used in this study is a NNPZD-DOM model (Nitrate, Ammonium, Phytoplankton, Zooplankton Detritus, Dissolved Organic Matter), qualitatively calibrated for DYFAMED by Mémery et al. (2002). Compared to other more complete models (Fasham et al., 1990; Lévy et al., 1998a), its relative simplicity 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, and winter mixing of accumulated semi-refractory DOM is associated with the dissolved export flux, which can be the major export flux in the Mediterranean Sea (Copin-Montégut and Avril, 1993). A schematic representation of the model is shown in Fig. 1, and the basic equations

are presented below. More details about the modeled biogeochemical processes can be found in Mémery et al. (2002) and in Lévy et al. (1998a).

The biogeochemical 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_T , is obtained diagnostically from TKE, with a 1.5 closure scheme in the Mellor Yamada nomenclature (Gaspar et al., 1990). The model covers the first 400 m of the water column, with a vertical discretization of 5 m.

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

$$\frac{\partial C}{\partial t} - \frac{\partial}{\partial z} \left(K_T \frac{\partial C}{\partial z} \right) = F_C \quad (1)$$

where C is the tracer concentration.

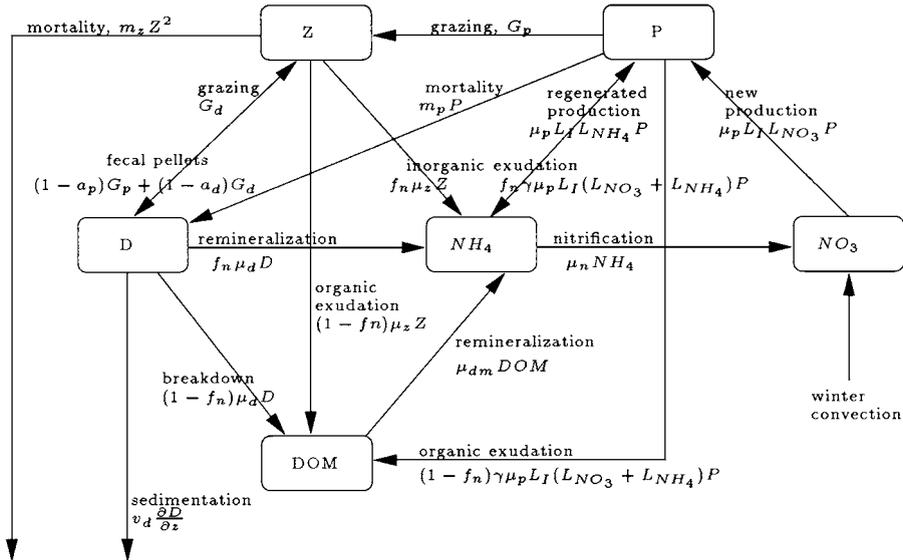


Fig. 1. Schematic representation of the compartments and processes of the NNPZD-DOM surface layer model.

For the surface layers (0–150 m), the reaction terms read as follows:

$$F_{\text{NO}_3} = -\mu_p L_I L_{\text{NO}_3} \text{P} + \mu_n \text{NH}_4 \quad (2)$$

$$F_{\text{NH}_4} = -\mu_p L_I L_{\text{NH}_4} \text{P} + f_n (\gamma \mu_p L_I (L_{\text{NO}_3} + L_{\text{NH}_4}) \text{P} + \mu_z Z + \mu_d \text{D}) - \mu_n \text{NH}_4 + \mu_{\text{dm}} \text{DOM} \quad (3)$$

$$F_{\text{P}} = (1 - \gamma) \mu_p L_I (L_{\text{NO}_3} + L_{\text{NH}_4}) \text{P} - G_p - m_p \text{P} \quad (4)$$

$$F_Z = a_p G_p + a_d G_d - m_z Z^2 - \mu_z Z \quad (5)$$

$$F_{\text{D}} = (1 - a_p) G_p - a_d G_d + m_p \text{P} - \mu_d \text{D} - \nu_d \frac{\partial}{\partial z} \quad (6)$$

$$F_{\text{DOM}} = (1 - f_n) (\gamma \mu_p L_I (L_{\text{NO}_3} + L_{\text{NH}_4}) \text{P} + \mu_z Z + \mu_d \text{D}) - \mu_{\text{dm}} \text{DOM} \quad (7)$$

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 (1996) kinetics:

$$L_{\text{NH}_4} = \frac{\text{NH}_4}{k_n + \text{NH}_4} \quad (8)$$

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

$$L_{\text{NO}_3} + L_{\text{NH}_4} = \frac{\text{NO}_3 + \text{NH}_4}{(k_n + \text{NO}_3 + \text{NH}_4)} \quad (9)$$

we obtain,

$$L_{\text{NO}_3} = \frac{k_n \text{NO}_3}{(k_n + \text{NH}_4)(k_n + \text{NO}_3 + \text{NH}_4)} \quad (10)$$

Light limitation is expressed as follows:

$$L_I = 1 - \exp(-\text{PAR}/k_{\text{par}}) \quad (11)$$

The photosynthetic available radiation (PAR) is predicted from surface irradiance and phytoplankton

Table 1
Parameter values

Parameter	Name	Value	Unit
Half-saturation const. for nutrients	k_n	0.01	mmol m^{-3}
Carbon/chlorophyll ratio	r_c	55	mg C mg Chla^{-1}
Phyto. exudation fraction	γ	0.05	
Zoo. nominal preference for phyto. food	r	0.7	
Max. specific zoo. grazing rate	g_z	$8.68\text{e}-6$	s^{-1}
Half-saturation const. for grazing	k_z	1	mmol N 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	$1.16\text{e}-6$	s^{-1}
Phyto. mortality rate	m_p	$9.027\text{e}-7$	s^{-1}
Zoo. mortality rate	m_z	$1.0\text{e}-7$	$\text{mmol N m}^{-3} \text{s}^{-1}$
Detritus breakdown rate	μ_d	$1.04\text{e}-6$	s^{-1}
Detritus sedimentation speed	ν_d	$5.8\text{e}-5$	m s^{-1}
Max. phyto. growth rate	μ_p	$2.31\text{e}-5$	s^{-1}
Light half-saturation const.	k_{pr}	33.33	W m^{-2}
Decay rate below the euphotic layer	τ	$5.80\text{e}-5$	s^{-1}
Ratio of inorganic exudation	f_n	0.8	
Nitrification rate	μ_n	$3.81\text{e}-7$	s^{-1}
Slow remineral. rate of DOM to NH_4	μ_{dm}	$6.43\text{e}-8$	s^{-1}
Coeff. for Martin's remineralization profile	h_r	-0.858	
Pigment absorption in red	k_{rp}	0.037	m^{-1} $(\text{mg Chl m}^{-3})^{-I_r}$
Pigment absorption in green	k_{gp}	0.074	m^{-1} $(\text{mg Chl m}^{-3})^{-I_g}$
Power law for absorption in red	I_r	0.629	
Power law for absorption in green	I_g	0.674	
Contribution of Chl to absorbing pigments	r_{pg}	0.7	

pigment content according to a light absorption model. Two different wavelengths are considered and the absorption coefficients depend on the local phytoplankton concentrations:

$$\text{PAR}(z, \text{P}) = \text{PAR}_r(z, \text{P}) + \text{PAR}_g(z, \text{P}) \quad (12)$$

$$\text{PAR}_g(z, P) = \text{PAR}_g(z - \Delta z, P)(1 - \exp(-k_g \Delta z)) \quad (13)$$

$$\text{PAR}_r(z, P) = \text{PAR}_r(z - \Delta z, P)(1 - \exp(-k_r \Delta z)) \quad (14)$$

$$k_g = k_{go} + k_{gp} \left(\frac{12Pr_d}{r_{pg}r_c} \right)^{l_g} \quad (15)$$

$$k_r = k_{ro} + k_{rp} \left(\frac{12Pr_d}{r_{pg}r_c} \right)^{l_r} \quad (16)$$

where $k_{go} = 0.0232 \text{ m}^{-1}$ and $k_{ro} = 0.225 \text{ m}^{-1}$ are the water absorptions in green and red.

Grazing of phytoplankton and detritus is formulated following Fasham et al. (1990):

$$G_p = \frac{g_z r P^2 Z}{k_z (rP + (1-r)D) + rP^2 + (1-r)D^2}, \quad (17)$$

$$G_d = \frac{g_z (1-r)D^2 Z}{k_z (rP + (1-r)D) + rP^2 + (1-r)D^2}. \quad (18)$$

Other modeled 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 remineralized in the water column), zooplankton exudation, fecal pellet production, detritus sedimentation, detritus breakdown, nitrification, and dissolved organic matter remineralization (Fig. 1).

Below a depth of 150 m (referred to as z_{bio} hereafter), remineralization processes are preponderant and the surface model does not apply. Instead, decay of phytoplankton, zooplankton and detritus in nutrients, and a vertical redistribution of zooplankton mortality according to Martin and Fitzwater's (1992) profile parameterize remineralization

below the surface layer. This parameterization conserves total nitrogen.

$$F_{\text{NH}_4} = -\tau \text{NH}_4 \quad (19)$$

$$F_{\text{NO}_3} = \tau(\text{NH}_4 + P + Z + D + \text{DOM}) + \frac{\partial f}{\partial z} \quad (20)$$

$$F_P = -\tau P \quad (21)$$

$$F_Z = -\tau Z \quad (22)$$

$$F_D = -\tau D - v_d \frac{\partial D}{\partial z} \quad (23)$$

$$F_{\text{DOM}} = -\tau \text{DOM} \quad (24)$$

The remineralization of the export flux due to zooplankton mortality (large particles) is expressed as:

$$f(z_{\text{bio}}) = \int_0^{z_{\text{bio}}} m_z Z^2 dz, \quad (25)$$

$$f(z) = f(z_{\text{bio}})(z/z_{\text{bio}})^{h_r}. \quad (26)$$

Stating that no deposition occurs on the ocean floor, a further condition is added:

$$f(z_{\text{bottom}}) = 0. \quad (27)$$

2.2. The DYFAMED simulation set-up

The model set up is the same as in Mémery et al. (2002). The standard run consists of a spin-up period of two repeated years (1995), followed by two more years, 1995 and 1996. The end of year 1996 is the restart point for the working year 1997. A complete description of seasonal and interannual variability is documented in Mémery et al. (2002). The simulation is forced with ECMWF atmospheric data, which give the wind stresses and heat fluxes every 6 h. Although

the DYFAMED station can be considered, at a first approximation to behave like a 1D vertical system (Lévy et al., 1998a; Andersen and Prieur, 2000), deep convection in winter involves barotropic–baroclinic instabilities, which are three-dimensional processes (Madec et al., 1991; Lévy et al., 1998b). In winter, therefore these 3D processes must be parameterized. We use a crude parameterization consisting of restoring temperature, salinity and dissolved tracers (nitrate and DOM) towards homogeneous profiles in February when convection is at its peak and the mixed-layer is deepest. For example, the NO_3 reaction term, F_{NO_3} , becomes $F_{\text{NO}_3} + \alpha(N^r - \text{NO}_3)$. N^r is the restoring profile, and α equals $1/7 \text{ day}^{-1}$ in February and 0 otherwise. For temperature and salinity, we use the February data as restoring profiles, since they are fairly homogeneous. For nitrate, the only available winter profile is not homogeneous. Therefore, we construct a homogeneous profile with the a priori averaged nitrate value below 200 m (6 mmol N m^{-3}). This restoring profile can be seen as the nitrate initial condition or preconditioning, since it determines to a large extent the amount of nitrate made available to the system. For DOM, restoring is carried out towards zero, which implicitly assumes that all the semi-labile DOM is exported in winter through convection.

In this work, as well as in the previous studies carried out at DYFAMED (Lévy et al., 1998a; Mémery et al., 2002), advection is neglected. This might result in a crude approximation in summer during strong wind events. Indeed, Andersen and Prieur (2000) and Chifflet et al. (2001) suggest that Ekman pumping may be responsible for an upward advective shift of the nitracline and the deep chlorophyll maximum (DCM) of the order of 10 to 15 m.

2.3. The DYFAMED data set

The data used in this study are monthly chlorophyll and nitrate profiles collected for the years 1995–1997 at the DYFAMED station, and surface fluorescence data measured by the Carioca buoy (Hood and Merlivat, 2001) moored at DYFAMED in 1997 (Figs. 2–4). We also make use of various annual estimates of new production, regenerated production, dissolved and particulate export (estimated from disparate measurements undertaken in the 1990s), and of the zooplankton content estimated during a special cruise in May 1995 (Andersen and Prieur, 2000).

The evolution over time of the chlorophyll and nitrate profiles (Figs. 3 and 4) reflects the seasonal variability at DYFAMED. Winter mixing brings nutrients to the surface, but the short residence time

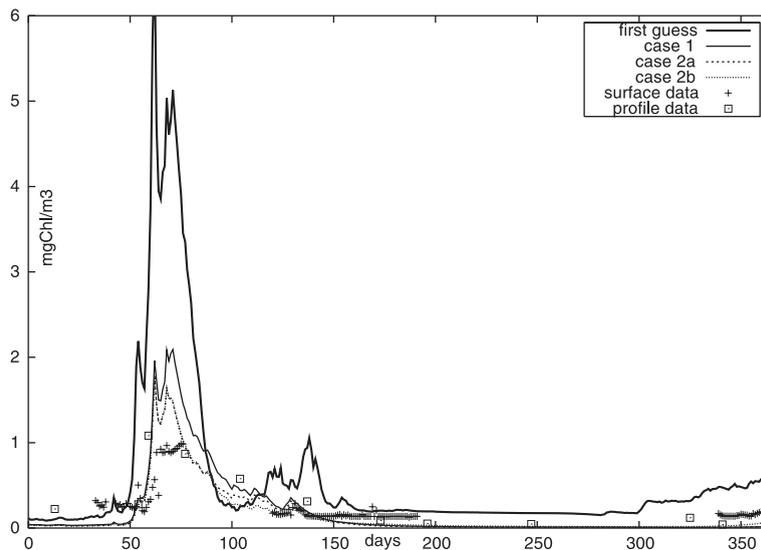


Fig. 2. Surface chlorophyll (mg Chl a m^{-3}) for year 1997. Crosses are the surface observations. Squares are the surface measurements taken from the chlorophyll profile set (Fig. 3). The first guess, Case 1, Case 2a and Case 2b runs are represented according to the legend above.

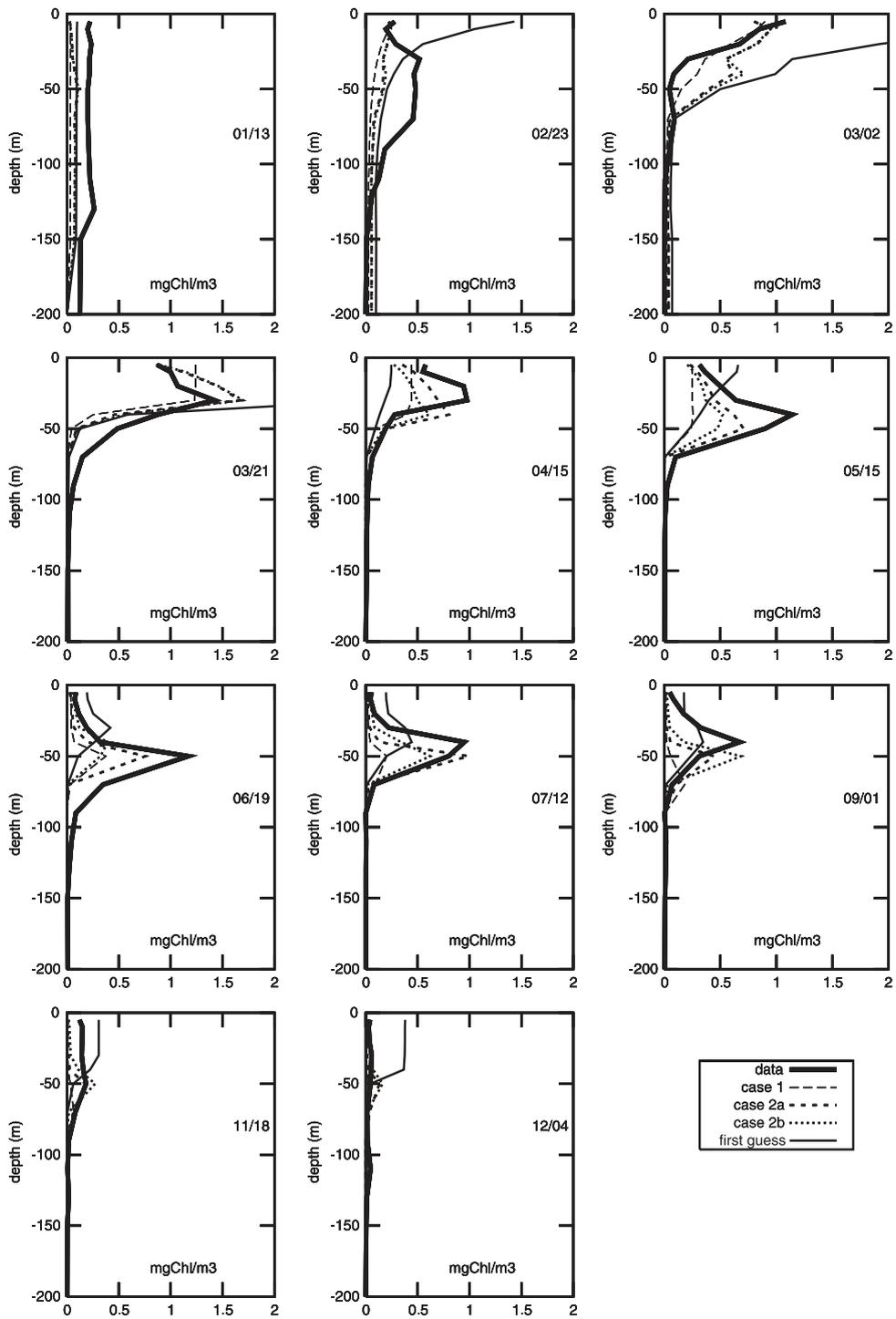


Fig. 3. Chlorophyll profiles ($\text{mg Chl } a \text{ m}^{-3}$) for year 1997. Depth in meters. The observed profiles as well as the first guess, Case 1, Case 2a and Case 2b profiles are represented according to the legend above.

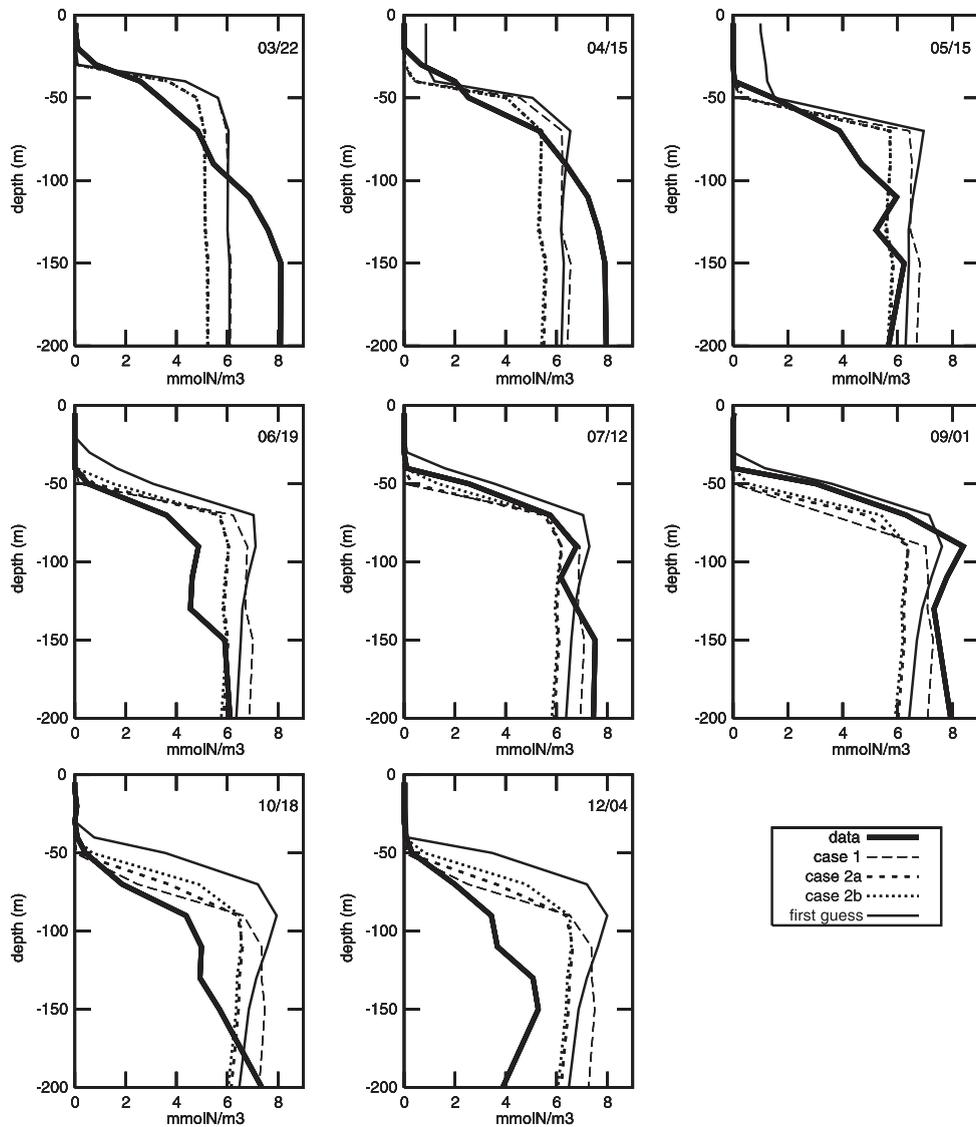


Fig. 4. NO_3 profiles (mmol N m^{-3}) for year 1997. Depth in meters. The observed profiles as well as the first guess, Case 1, Case 2a and Case 2b profiles are represented according to the legend above.

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 utilized continuously. As a consequence, 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 characterized by a

deep chlorophyll maximum. It may be noticed that the nitrate profiles show strong variability below 100 m. This variability cannot be attributed to biological processes since these occur closer to the surface, and it is therefore most likely due to advection. It cannot therefore be captured by the model. During the oligotrophic period, the location of the nitracline and the DCM 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.

Estimates of total production (TP) vary considerably. Minas (1970), using in situ measurements, estimated the Northwestern Mediterranean TP to be about $78 \text{ g C m}^{-2} \text{ year}^{-1}$, whereas Antoine et al. (1995) obtained a value of $157.5 \text{ g C m}^{-2} \text{ year}^{-1}$ from satellite data and Marty and Chiavérini (2002) an average value of $156 \text{ g C m}^{-2} \text{ year}^{-1}$ over the 1990s from C14 incubations (although they warn that this value should be considered as in the upper range). Estimates of new production (NP) range from 20 (Bethoux, 1989) to $42 \text{ g C m}^{-2} \text{ year}^{-1}$ (Marty and Chiavérini, 2002). This latter value is based on nitrate uptake (NU) and does not therefore correct for nitrification, which explains the difference with the Bethoux (1989) estimate. With regard to export fluxes, the DOC export flux has been estimated by Copin-Montegut and Avril (1993), based on 1991 DOC data, to be around $14.8 \text{ g C m}^{-2} \text{ year}^{-1}$, while sediment trap data from 1987 (Miquel et al., 1993) predict a particulate export at 200 m of $4 \text{ g C m}^{-2} \text{ year}^{-1}$.

The data assimilation period is restricted to 1997, when all three types of synoptic data are available.

For coherence with the model set-up strategy, the non-homogeneous February nitrate profile is not accounted for in the assimilation. Data from years 1995 and 1996 are kept for validation.

In order to compare model results and data, calculated phytoplankton concentrations are converted to mg Chl a m^{-3} using a constant (in time, space and during parameter optimization) Red field ratio C/N of $r_d=6.625$, and a C/Chl ratio (r_c in Table 1) that will be optimized. A linear relationship is assumed between fluorescence and chlorophyll concentrations. The proportionality coefficient was chosen so that surface fluorescence data and monthly chlorophyll profiles were coherent with one another.

2.4. An initial simulation

A first run is carried out using the a priori parameter set and without any data assimilation (Figs. 3–5). The seasonal variability is well reproduced, but the spring bloom appears to be too strong and the transition towards oligotrophy poorly represented. On the other hand, and because the parameters have been tuned for this purpose (Lévy et al., 1998a), this “first guess” run predicts reasonably good fluxes: 110.6 g C

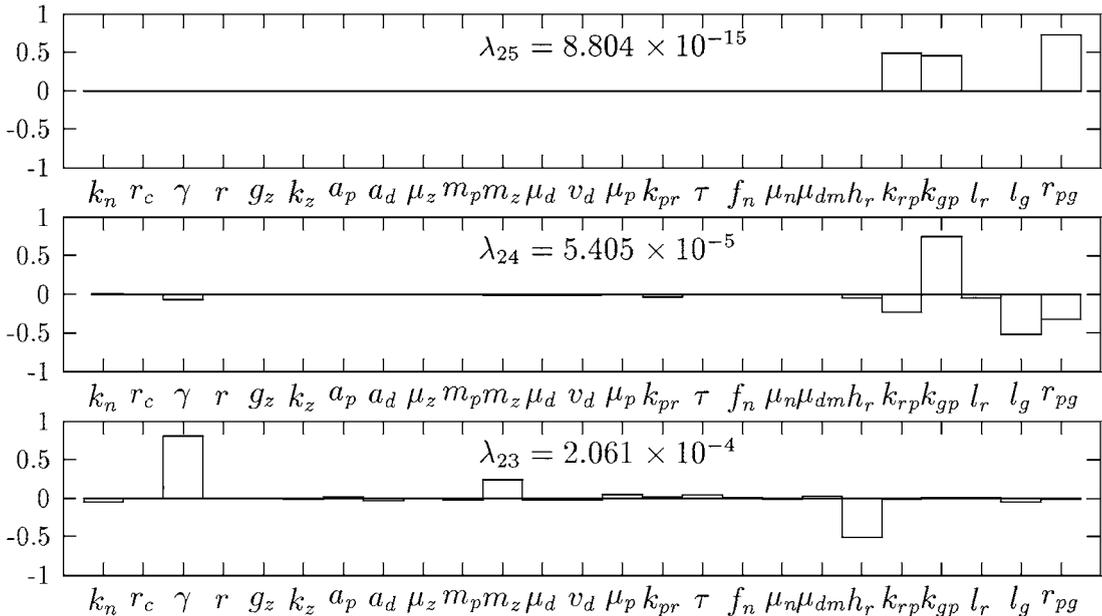


Fig. 5. Three smallest eigenvalues and corresponding eigenvectors of the Hessian for the 25-parameter optimization case.

$\text{m}^{-2} \text{ year}^{-1}$ for TP, $22.5 \text{ g C m}^{-2} \text{ year}^{-1}$ for NP, $13.2 \text{ g C m}^{-2} \text{ year}^{-1}$ for DOC export and $5.2 \text{ g C m}^{-2} \text{ year}^{-1}$ for particulate export.

3. Method

3.1. Control variables and cost function

The variational data assimilation method mainly consists in finding an optimal control minimizing a cost function which measures the distance, in a weighted least-squares sense, between the model's solution and the observations. In the current literature, authors usually consider the biological parameters as being the control variables. It is therefore assumed that the discrepancy between the model and the observations can be attributed mainly to these poorly known quantities. However, implicitly assuming that the model equations are perfect except for the unknown parameters may be a problem (Evensen et al., 1998). Firstly, minimization may lead to an optimal parameter set attempting to correct for errors other than those related to the particular parameterizations. Secondly, the tuning of the parameters may not be sufficient to fit real observations and it may be necessary to introduce more degrees of freedom in the model equations.

In our case, the winter restoring terms present in the equations are clearly crude parameterizations made to take into account 3D processes which are not explicitly resolved. The restoring value of 6 mmol N m^{-3} for NO_3 was chosen empirically on the basis of the DYFAMED data. The NO_3 restoring term is particularly important in the biological dynamics (Mémery et al., 2002). It can be regarded as an initial condition from which we want to start at the end of winter. It can also be regarded as a particular form of a model error term, and we may need to control it.

The 1D formulation of the model as well as a first real data assimilation experiment (Case 1, Section 4.4.1) also lead us to consider a C/Chl ratio r_c that varies with depth. We have therefore considered different cases in this study. Case 1 is classical: the cost function depends on the biological parameters, which are undistributed, or constant in time and space. In Case 2, the cost function depends on all the

previous parameters but r_c can vary with depth. It also depends on the NO_3 restoring profile, N^r , which is also space-distributed.

As shown in Table 1, the biological parameters have very different orders of magnitude. To avoid any numerical difficulties which might arise from this during the minimization, we adimensionalize the parameter vector \mathbf{K} , dividing each parameter \mathbf{K}_i by its first guess value \mathbf{K}_i^0 , $k_i = \mathbf{K}_i / \mathbf{K}_i^0$. In the same way, we consider the dimensionless NO_3 restoring profile, $n^r = (n_i^r)_{i=1, \dots, nz} = N^r / 6$, where nz is the number of space discretization points. Such a non-dimensionalization procedure can be regarded as a preconditioning for minimization.

To sum up, in Case 1, the control variable is $x = k$ of size p , where p is the number of parameters, and k is dimensionless. In Case 2, the control variable is $x = (\tilde{\mathbf{k}}, r_c, n^r)$ of size $(p-1) + nz + nz$, where $\tilde{\mathbf{k}}$ represents the parameter vector without r_c and n^r . In both cases, the first guess value for x is $x^0 = (1, \dots, 1)$.

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

$$J_0(x) = J_{\text{sc}}(x) + J_{\text{pc}}(x) + J_{\text{pn}}(x).$$

J_{sc} is related to the ncs surface chlorophyll observations, dcs.

$$J_{\text{sc}}(x) = \frac{1}{2} \sum_{i=1}^{\text{ncs}} w_{\text{cs}} (\phi_{\text{cs}}(x)_i - \text{dcs}_i)^2.$$

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

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

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

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

More simply we have,

$$J_0(x) = \frac{1}{2} \langle W(\phi(x) - d), (\phi(x) - d) \rangle_{\text{N obs}}.$$

$\langle \cdot, \cdot \rangle_{\text{N obs}}$ denotes the euclidian scalar product in $\mathbb{R}^{\text{N obs}}$, with $\text{N obs} = \text{ncs} + \text{ncp} \times \text{mcp} + \text{nnp} \times \text{mnp}$. ϕ is the

direct operator which, to an admissible control variable x , associates the model equivalents to the data. d is the data. W is a weighting diagonal matrix. The weights w_{cs} , w_{cp} and w_{np} 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 $w_{cs} = \frac{1}{ncs \times \sigma_{cs}^2}$, $w_{cp} = \frac{1}{(ncp \times mcp) \sigma_{cp}^2}$, and $w_{np} = \frac{1}{(nnp \times mnp) \sigma_{np}^2}$.

We also add 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, $\sigma_i = \frac{1}{\sqrt{w_i}}$, given in Table 2. This term enables the minimization to avoid biologically absurd optimal parameter values.

The second term, which only applies in Case 2, is a Tikhonov regularization term (Tikhonov and Arsenin, 1977).

$$J_r(n^r) = \frac{1}{2} w \sum_{i=1}^{nz} \left(\frac{\partial^2 n^r}{\partial z^2} \right)_i^2$$

This is added to avoid instability phenomena inherent in such inverse problems. Its role is to smooth the

Table 2

A priori parameter errors σ , reference parameter set k^0 (adimensionalized with Table 1 values) and perturbed parameter set $k^0 + \delta k$

Parameter	A priori error σ	k^0	First guess error δk
k_n	1	1	0.4
r_c	2	1	0.4
γ	1	1	0.4
r	0.5	1	0.2
g_z	0.3	1	0.3
k_z	0.9	1	-0.4
a_p	1	1	0.3
a_d	1	1	0.4
μ_z	0.25	1	-0.1
m_p	0.4	1	-0.6
m_z	1	1	0.5
μ_d	0.5	1	0.2
v_d	1	1	0.4
μ_p	0.5	1	-0.5
k_{pr}	0.5	1	0.2
τ	0.5	1	0.2
f_n	0.2	1	0.2
μ_n	1	1	0.4
μ_{dm}	1	1	-0.2

space-distributed control n^r . The choice of the weight w is not straightforward, as a compromise has to be found between numerical stability and the quality of the minimization of J_0 . In practice, w was chosen so that the data misfit part of the cost and its regularization part are balanced. The same regularization term is added for the distributed parameter r_c .

3.2. Automatic differentiation

To minimize the cost function J , we applied the quasi-Newton algorithm implemented in the *n1qn3* Fortran subroutine of Gilbert and Lemaréchal (1989). The computation of the gradient of J with respect to control variables is required at each step of the minimization. This gradient results in one integration of the adjoint model. The tangent and adjoint codes were obtained using the automatic differentiation program *Odyssée* (Faure and Papegay, 1997; Griewank, 2000), which is an efficient tool for obtaining adjoint codes since it enables the automatic production of adjoint instructions. However, codes produced by automatic differentiation do not usually use computer memory in a very efficient way. Saving the direct model trajectory is the major problem. A differentiation program has to follow systematic methods to provide the evaluation trajectory. Thus, *Odyssée* systematically uses a local calculation and storage technique for the trajectory. Automatically differentiating a 3D model and using the adjoint code directly seems impossible. This problem partially disappears using a 1D model, so that the work to be done on the automatically produced adjoint routines is more reasonable. In this study, we used an integration period of 1 year, the choice being made to cover a whole annual cycle of biological activity when all three types of data were available. Moreover, it involves affordable computer memory resources and acceptable calculation times. An adjoint run is about 10 times more expensive than a direct run in terms of memory requirements and CPU time.

3.3. Sensitivity analysis method

As was highlighted in Fennel et al. (2001), a sensitivity analysis should be an integral part of any attempt to optimize the parameters of a biological model. Fennel et al. identified different inconsisten-

cies in the formulation of their inverse problem such as parameters which could not be determined independently.

The a priori approach consists in using a sensitivity analysis based on twin experiments, that is in the case where the global minimum of the cost function is known, to help determine the most sensitive and consequently the most important parameters to optimize in order to solve a numerically well-posed inverse problem.

Let us consider that the cost function only depends on parameter k and that no penalty term is added:

$$J(k) = \frac{1}{2} \langle W(\phi(k) - d), (\phi(k) - d) \rangle_{N_{\text{obs}}} . \quad (28)$$

A small perturbation h on parameter set k induces a perturbation on ϕ ,

$$\phi(k + h) - \phi(k) \approx \phi'(k)h$$

$\phi'(k)$ is the sensitivity matrix, of size $N_{\text{obs}} \times p$. It can easily be calculated by p integrations of the tangent linear model, perturbing the parameters one after another. The i th integration leads to the i th column of matrix $\phi'(k)$ that represents the influence of parameter k_i on the model's output used to calculate the cost function. Then, if we expand J in its Taylor series about the optimal solution \mathbf{k}^0 and ignore terms of a higher order than two, we obtain the following, as the first derivative vanishes:

$$J(\mathbf{k}^0 + h) \approx J(\mathbf{k}^0) + \frac{1}{2} \langle \phi'(\mathbf{k}^0)^T W \phi'(\mathbf{k}^0) h, h \rangle_p + \frac{1}{2} \langle W \phi''(\mathbf{k}^0)(h, h), (\phi(\mathbf{k}^0) - d) \rangle_{N_{\text{obs}}} .$$

In the context of twin experiments, where data are generated by a model run using \mathbf{k}^0 as the parameter set, the formula reduces to

$$J(\mathbf{k}^0 + h) \approx \frac{1}{2} \langle \phi'(\mathbf{k}^0)^T W \phi'(\mathbf{k}^0) h, h \rangle_p \approx \frac{1}{2} h^T H h,$$

where $H = \phi'(\mathbf{k}^0)^T W \phi'(\mathbf{k}^0)$ is the Hessian of the cost function in this particular case. H is easily calculated once $\phi'(\mathbf{k}^0)$ has been obtained. It may be noted that H

does not depend on the quality of the data but only on the intrinsic properties of the model structure, or in other words on the way in which the parameters are mutually related in the conceptual representation of the biological phenomena. The Hessian matrix provides a good approximation of the covariance matrix for the model parameters (Thacker, 1989).

Its diagonal element h_{ii} gives the sensitivity of the cost function to change in parameter k_i , whereas its off-diagonal elements indicate the degree to which pairs of model parameters are correlated. In fact, all the needed information is contained in the eigenvalues and vectors of H . The condition number, $\lambda_{\text{max}}/\lambda_{\text{min}}$ gives an indication of the degree of singularity in the problem. Small eigenvalues correspond to large uncertainties in the identification of the parameters that make a significant contribution in the related eigenvectors. Such an analysis provides a systematic method for determining parameters which may be difficult to identify.

4. Results

4.1. Description of the experiments

This short subsection briefly presents all the different numerical experiments which follow in the text. Let us first remember that in Case 1 the cost function is assumed to depend on constant parameters, whereas in Case 2 it is assumed to depend on constant parameters and on two parameters (n^r and r_c) allowing vertical variability. We first conduct a sensitivity analysis based on a twin experiment identical to Case 1 (Section 4.2). A synthetic data set is created with the model, by running the reference 1997 simulation with the parameter set \mathbf{k}^0 . These synthetic data have the same spatio-temporal distribution as the real 1997 data. We then conduct twin experiments for both Case 1 (Section 4.3.1) and Case 2 (Section 4.3.2). Once these validation steps have been successfully completed, we come to the assimilation of real (i.e. observed) data. Section 4.4.1 deals with Case 1 optimization. Case 2 is divided into Case 2a in Section 4.4.2 and Case 2b in Section 4.4.3. As explained in Introduction, in Case 2b zooplankton data is added to the cost function, and in Case 2c annual flux estimates are also assimilated.

4.2. Twin experiment sensitivity analysis

We apply the sensitivity analysis methodology described in Section 3.3 to the twin experiments to ensure that we are not trying to solve an ill-posed inverse problem. Within this framework, the cost function is defined without any penalty term (Eq. (28)). The Hessian, H , is calculated at the optimal point \mathbf{k}^0 , and then its eigenvalues and eigenvectors.

If the cost function is assumed to depend on the 25 parameters in Table 1, the condition number of its Hessian is 1.149×10^{17} , indicating that the inverse problem is extremely ill-conditioned. The three smallest eigenvalues, $\lambda_{25} = 8.804 \times 10^{-15}$, $\lambda_{24} = 5.405 \times 10^{-5}$ and $\lambda_{23} = 2.061 \times 10^{-4}$ and their corresponding eigenvectors, v_{25} , v_{24} and v_{23} are plotted in Fig. 5. v_{25} has contributions mainly for k_{rp} , k_{gp} and r_{pg} . As the associated eigenvalue is almost 0, these parameters cannot be determined with the data under consideration. All of them belong to the optical model formulation (Eqs. (12)–(16)). Eigenvalue $\lambda_{24} = 5.405 \times 10^{-5}$ is still quite small, and the important contributions in v_{24} again correspond to the optical parameters. These parameters enter the model only in the combinations $k_{\text{gp}}(12r_{\text{d}}/$

$r_{\text{pg}}r_{\text{c}})^{\text{g}}$ and $k_{\text{rp}}(12r_{\text{d}}/r_{\text{pg}}r_{\text{c}})^{\text{r}}$ (Eqs. (15) and (16)). It is therefore not surprising that these parameters appear to be dependent on one another. As a considerable improvement is needed in the condition number of H , we will not try to estimate the five optical parameters in all of the following experiments. Their values are kept fixed at the initial guesses. Only the C/Chl ratio r_{c} will therefore play a role in the calculation of light limitation L_{I} during the optimization procedure.

Under this new assumption, the condition number of the Hessian matrix falls to 7.489×10^6 which is much more acceptable and comparable to the condition number values presented in Fennel et al. (2001). The worse resolutions are observed for h_{r} , γ and m_{z} . The coefficient for Martin’s remineralization profile, h_{r} , only concerns the remineralization model (below 150 m), and is kept constant in the following sections.

Finally, the chosen formulation includes 19 parameters. The condition number of the Hessian is 9.108×10^5 . Eliminating two or three additional parameters, such as γ , m_{z} or k_{n} , which are the least sensitive (Fig. 6), does not considerably improve the conditioning of the problem and may, on the con-

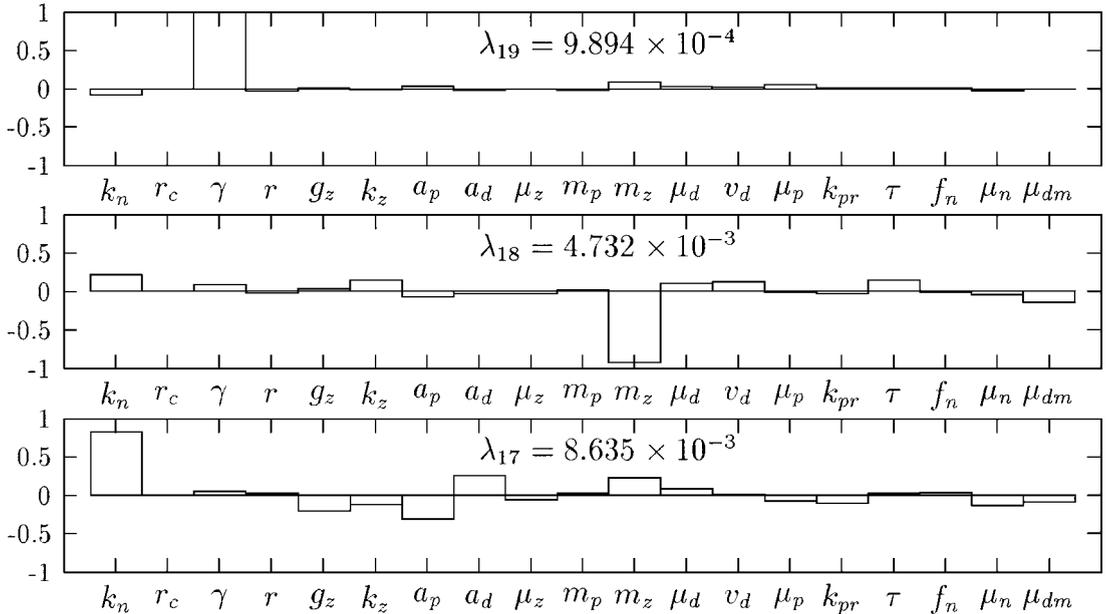


Fig. 6. Three smallest eigenvalues and corresponding eigenvectors of the Hessian for the 19-parameter optimization case.

trary, deprive the model of useful degrees of freedom.

4.3. Twin experiment data assimilation

4.3.1. Case 1

The second and final validation step before assimilation of observed data is to conduct twin experiments. Synthetic data are produced by the model using the first guess parameter vector \mathbf{k}^0 containing the 19 selected parameters. In order to fully test the possibility of recovering these parameters from the synthetic data, no penalty term is added. The variances assumed are as follows: $\sigma_{cs}=0.3 \text{ mg Chla m}^{-3}$, $\sigma_{cp}=0.1 \text{ mg Chla m}^{-3}$, and $\sigma_{np}=0.4 \text{ mmol N m}^{-3}$. This provides a good balance between the three terms of J_0 . Different first guesses for the parameter vector were obtained by perturbing \mathbf{k}^0 within the a priori error deviation ranges given in Table 2. All the corresponding optimizations converged to the minimum of the cost function, $J(\mathbf{k}^0)=0$.

The results of this experiment are shown in Fig. 7. Table 2 presents the first guess parameter set $\mathbf{k}^0 + \delta k$. The convergence criterion ($\|\nabla J\| \leq \varepsilon$, where ε is a small value) is satisfied after 129 iterations. A cost function value of $7.35e-11$ is obtained, indicating

that it has reached its global minimum. All the 19 parameters have been recovered. In the first 20 iterations, the cost function decreases rapidly then enters a long period of more gradual decline. This corresponds to the rapid recovery of the most sensitive parameters, followed by a difficult search for the optimal values of the less sensitive parameters (k_n , γ , and m_z).

4.3.2. Case 2

We are now concerned with the optimization of 18 space- and time-constant parameters and 2 space-dependent parameters, r_c (the C/Chl ratio) and n^r (the NO_3 winter preconditioning). This greatly increases the number of control variables from 19 to 178. However, the problem still remains overdetermined as the control dimension does not exceed the number of data. Nevertheless, this does not mean that it is possible to recover the model controls given the set of data we have for year 1997 at the DYFAMED station. Some controls may not have a real influence everywhere on the simulation results. Moreover, the method has to be able to distinguish between wrong biological parameter values, wrong C/Chl ratios and wrong NO_3 preconditioning to explain the discrepancy between

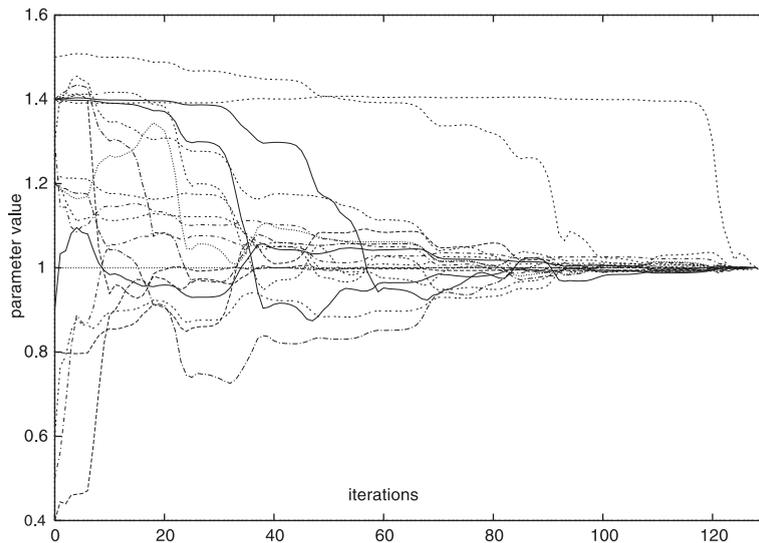


Fig. 7. Convergence of the 19 parameter set from $\mathbf{k}^0 + \delta k$ to \mathbf{k}^0 .

model and data. To check this, a twin experiment is conducted in which no penalty term is added to the cost function relating to the constant parameters. A Tikhonov regularization term is used for the two distributed controls r_c and n^r .

$$J(x) = J_0(x) + J_r(n^r) + J_r(r_c)$$

Exact equivalents of year 1997 DYFAMED data were created using the reference a priori k^0 vector for the 18 parameters, the distributed r_c ratio and the distributed n^r profile plotted in Fig. 8. First guess control variables consist of a profile n^r homogeneous on the vertical with value $n^r=1$ and the parameter set $k^0 + \delta k$ in Table 2, except for r_c which is also chosen to be homogeneous in the vertical. The first guess for r_c is 1. Below 200 m, this value is the same as that used to generate the data. This choice was made because below 200 m

r_c can take any value without influencing the simulation or, therefore, the cost function value. The twin experiment proved to be successful, with all biological parameters, as well as r_c , completely recovered and the restoring profile, n^r , also very well reproduced. The r_c and n^r used to generate the data, first guess and recovered values, are plotted in Fig. 8. It takes 543 iterations for the optimization to converge. This increase in the number of iterations required compared with Case 1 is due to the increase in the size of the control vector.

4.4. Assimilation experiments using real (observed) data

4.4.1. Estimation of constant parameters—Case 1

A real data assimilation experiment is performed using nitrate and chlorophyll data from the DYFAMED station. The first guess is chosen as k^0 , which corresponds to the parameter values

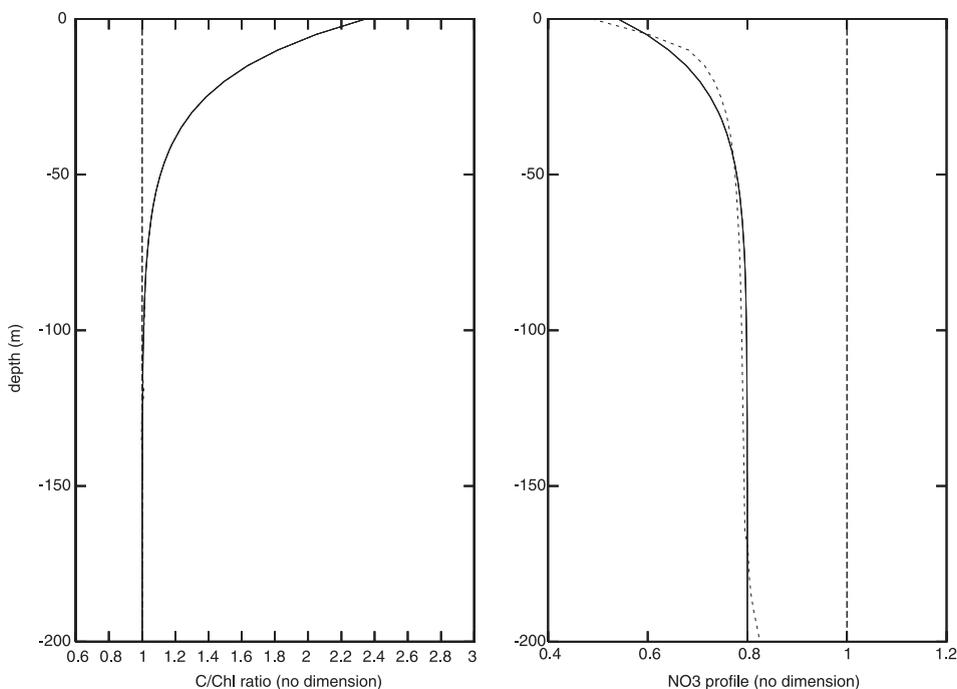


Fig. 8. Left: r_c , first guess (long dashes), reference (continuous line) and optimal (short dashes superimposed on the reference). Right: n^r , first guess (long dashes), reference (continuous line) and optimal (short dashes). The reference values, $R_c = 55 \text{ mg C mg Chl}^{-1}$ and $N^r = 6 \text{ mmol N m}^{-3}$, are used for adimensionalization.

Table 3

First guess and optimal adimensionalized parameters in Case 1, 2a, 2b and 2c optimizations

Parameter	First guess k^0	Optimal set k^{opt} (Case 1)	Optimal set k^{opt} (Case 2a)	Optimal set k^{opt} (Case 2b)	Optimal set k^{opt} (Case 2c)
k_n	1	0.964	0.992	0.992	0.746
r_c	1	2.452			
γ	1	1.086	1.073	1.065	1.299
r	1	0.950	0.962	0.896	0.897
g_z	1	0.920	0.944	0.965	0.937
k_z	1	1.285	1.245	1.172	1.291
a_p	1	1.209	1.235	0.915	1.110
a_d	1	1.029	1.022	1.015	1.117
μ_z	1	1.039	1.028	1.001	1.049
m_p	1	1.010	1.113	1.080	1.092
m_z	1	1.001	0.993	0.992	1.058
μ_d	1	0.435	0.597	0.557	1.149
v_d	1	1.732	1.556	1.160	0.709
μ_p	1	0.727	0.806	0.795	0.880
k_{pr}	1	1.081	1.050	1.040	1.009
τ	1	0.900	1.035	1.036	0.961
f_n	1	0.951	0.971	0.970	0.936
μ_n	1	0.880	0.967	0.960	0.793
μ_{dm}	1	0.895	0.893	0.860	0.757

commonly used in this area (Lévy et al., 1998a; Mémery et al., 2002), and the data used are those described in Section 2.2. A penalty term is added to the cost function of the twin experiments:

$$J(k) = J_0(k) + J_p(k) \quad (29)$$

This regularizes the inverse problem and the optimization converges in 33 iterations. The optimal parameter set, k^{opt} , is given in Table 3, and the overall quality of the optimization is quantified by the normalized cost related to each type of data in Table 4.

Table 4
Optimization results

Data type	First guess cost	Optimal cost (Case 1)	Optimal cost (Case 2a)	Optimal cost (Case 2b)	Optimal cost (Case 2c)
Surface chlorophyll	1	0.059	0.029	0.027	0.027
Chlorophyll profile	1	0.170	0.100	0.134	0.126
NO ₃ profile	1	0.626	0.525	0.554	0.717

Costs are normalized.

Correlation coefficients have also been computed to measure the shift between the model and the data. For surface chlorophyll, a time correlation coefficient is defined by:

$$r_t = \frac{\sum_{i=1}^{\text{ncs}} (\phi_{cs}(k^{\text{opt}})_i - \overline{\phi_{cs}^{\text{opt}}})(dcs_i - \overline{dcs})}{\left(\sum_{i=1}^{\text{ncs}} (\phi_{cs}(k^{\text{opt}})_i - \overline{\phi_{cs}^{\text{opt}}})^2 \sum_{i=1}^{\text{ncs}} (dcs_i - \overline{dcs})^2 \right)^{1/2}} \quad (30)$$

where $\overline{\phi_{cs}^{\text{opt}}}$ and \overline{dcs} denote the mean values. In the same way, space correlation coefficients are computed for each profile.

The largest change in the parameter set is obtained for r_c , the C/Chl ratio, which is used to convert phytoplankton into chlorophyll units. While the first guess for r_c is 55 mg C mg Chl⁻¹, the optimal value is 135 mg C mg Chl⁻¹. Other important changes concern parameters related to detritus, and thus with the export. Detritus breakdown rate, μ_d , decreases from 1.04e-6 s⁻¹ to 4.524e-7 s⁻¹, while detritus sedimentation speed, v_d , increases from 5 to 8.7 m s⁻¹. These changes in parameters have the effect of

decreasing the chlorophyll concentrations and of increasing the export.

Globally, the assimilation resulted in a considerable improvement in the model/data match, compared to the first guess experiment (Table 4). Bloom intensity is divided by a factor of three, in agreement with the observations (Fig. 2), and the time correlation coefficient increases from 0.87 to 0.92, indicating that the starting date and bloom duration are simulated more accurately. Chlorophyll profiles during the bloom (dates 02/23, 03/02 and 03/21 in Fig. 3) also show a better fit with the data. These improvements are mainly associated with the change in the C/Chl ratio. As with chlorophyll, nitrate is also better simulated and this is particularly apparent above the nitracline (Fig. 4). Nitrate concentrations at the surface were overestimated, and are now reduced. The reduction in nitrate concentrations is essentially achieved through the change in parameters μ_d and ν_d , and the associated increase in export. It may be noticed that the cost with respect to nitrate is globally higher than that relative to chlorophyll (Table 4). This is due to the poor fit between simulated and observed nitrate concentrations below the nitracline, where advective events seem to be important.

Despite the improvements noted above, however, certain problems remain. First, there is no improvement in summer, during the oligotrophic regime. The model has difficulties in reproducing the subsurface chlorophyll maximum. On 09/01, the space correlation coefficient falls from 0.865 to 0.061, indicating a substantial shift between observations and model results in terms of subsurface maximum depth. Second, the exported flux is too large by almost one order of magnitude (Table 5).

To address these shortcomings, the model is modified as follows. In order to better represent the chlorophyll subsurface maxima, the C/Chl ratio is allowed to vary on the vertical. Indeed, this ratio is known to be highly dependent on the photosynthetic available radiation (PAR) (Geider et al., 1997). In the Mediterranean Sea, it can vary from 30 to 200 mg C mg Chl⁻¹. Moreover, earlier optimization studies have already emphasized the need for a variable C/Chl ratio (Hurtt and Armstrong, 1996, 1999; Schartau et al., 2001; Spitz et al., 2001). The amount of NO₃ present in the model is very dependent on the winter preconditioning profile (Mémery et al., 2002). In

Table 5

Flux data (from Marty and Chiavérini, 2002) and calculated fluxes at 200 m

Flux	Data	First guess	Case 1	Case 2a	Case 2b	Case 2c
NU	42 ± 15	43.8	42.1	39.5	36.7	43.9
NP		22.5	31.6	26.8	23.1	24.2
TP	156 ± 30	110.6	59.5	59.6	68.4	115.4
SF		3.6	17.8	12.6	8.3	1.6
ZF		1.6	10 ⁻³	10 ⁻²	0.3	0.3
PF	2 ± 0.8	5.2	17.8	12.6	8.6	1.9
DOMF	15 ± 5	13.2	13.1	13.1	13.2	13.7

Values are given in g C m⁻² year⁻¹. NU stands for NO₃ uptake, NP for new production (NU corrected from nitrification), TP for total production, PF for total particulate export flux (SF, the sedimentation flux + ZF, the zooplankton mortality flux) and DOMF for DOM export flux.

order to improve the match between observed and simulated NO₃ profiles without excessively increasing the export, a space-distributed NO₃ winter preconditioning profile is introduced in the control vector.

4.4.2. Estimation of parameters allowing vertical variability—Case 2a

Using the space-dependent r_c and n^f described above as control variables improves all three parts of the cost J_0 (Table 4). Compared to Case 1, the summer chlorophyll profiles are better represented and improvements are also noted during other periods of the year (Fig. 3). Nitrate profiles are also in better agreement with the observations (Fig. 4).

Fig. 9 shows the retrieved r_c and n^f profiles. The r_c ratio is high (up to 143 mg C mg Chl⁻¹) near the surface, and rapidly decreases with depth (45 mg C mg Chl⁻¹ at about 40 m). Due to the regularization term, $J_r(r_c)$ in the cost function, r_c remains close to its first guess value (55 mg C mg Chl⁻¹) below 50 m. This decrease in r_c with depth is realistic and physically related to the interdependence of r_c and PAR. Obviously, allowing r_c to vary only with depth is an oversimplification, since it does not account for the seasonal variations of PAR. However, this experiment clearly shows that the approximation already greatly improves the results in relation to the constant r_c case.

As for the optimal nitrate preconditioning profile, this is approximately homogeneous in the first 200 m of the water column (Fig. 9), consistent with the fact that the winter mixed layer is almost homogeneous down to this depth. It contains less nitrate than the

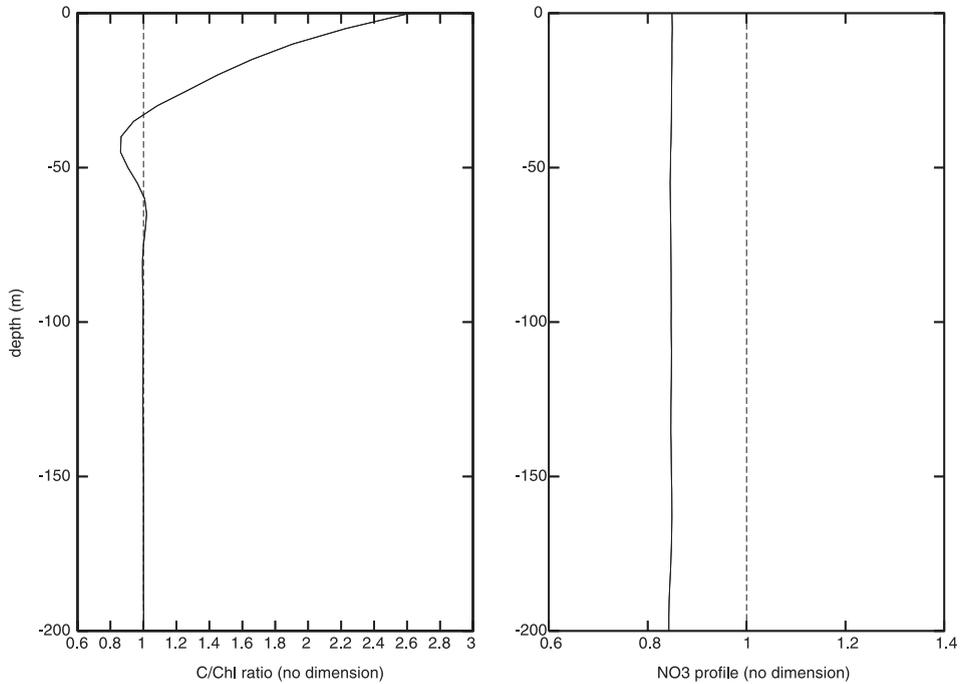


Fig. 9. Left: first guess (dashed line) and optimal (continuous line) r_c . Right: first guess (dashed line) and optimal (continuous line) n^f . The reference values, $R_c = 55 \text{ mg C mg Chl a}^{-1}$ and $N^f = 6 \text{ mmol N m}^{-3}$, are used for adimensionalization.

first guess (4.9 compared to 6 mmol N m^{-3}), which again indicates that there was too much nitrate in the first guess experiment.

The optimal parameters are given in Table 3. They do not differ greatly from those of Case 1, but it should be noted that the parameters which were thought to vary because of the presence of too much NO_3 at the surface (μ_d, v_d) have optimal values closer to their first guess value than was the case in Case 1. In other words, the optimal values of these parameters do not suffer any more from the excess of nitrate, and we may expect them to be more representative of reality.

Regarding the fluxes, DOC export and NP are barely modified by the assimilation, and keep their acceptable first guess value. However, the detritus flux is still too high, and TP too low (Table 5). TP has fallen to $59.6 \text{ g C m}^{-2} \text{ year}^{-1}$, which is well below the data range. The detritus flux is more than three times too high ($12.6 \text{ g C m}^{-2} \text{ year}^{-1}$) and consists only of the sedimentation of slow sinking particles. Export through sedimentation of fast sinking particles (ZF), which is parameterized through the zooplankton mortality term, has almost vanished,

whereas it represented one third of total particulate export in the first guess experiment.

The question arises as to why the model chooses to favor sedimentation of slow sinking particles over zooplankton mortality fluxes. Globally, in Cases 1 and 2a, zooplankton concentrations, which are not constrained by any data, are much weaker than in our first guess run (Table 6). This explains the fall in ZF, which behaves like Z^2 . Sensitivity studies show that the measurements, which cause the zooplankton concentrations to decrease, are the surface values of the chlorophyll profiles corresponding to the dates 03/21 and 04/15 (the two squares corresponding to days 80 and 105 in Fig. 2). These values impose a slow decrease in surface chlorophyll after the bloom maximum. This decrease is much more rapid in the first

Table 6
Zooplankton content data and calculated values at 200 m, for May 1997, in mmol N m^{-2}

Data	First guess	Case 1	Case 2a	Case 2b	Case 2c
20 ± 10	33.7	1.9	3.1	17.9	22.7

guess run. The data assimilation forces the model to simulate relatively high surface chlorophyll values until the month of April. This is achieved through a strong decrease in grazing, mainly resulting from the decrease in zooplankton concentrations. In Case 1, the model runs right between the two abovementioned post-bloom data (Fig. 2) and $ZF \approx 10^{-3} \text{ g C m}^{-2} \text{ year}^{-1}$, whereas in Case 2a the model runs below these measurements and $ZF \approx 10^{-2} \text{ g C m}^{-2} \text{ year}^{-1}$. Given the strong influence of these two post-bloom measurements on the assimilation results, it may intuitively be thought that the sampling frequency at this period must play an important role. However, this cannot be easily tested and generalized in a rigorous manner. Twin experiments could be performed but this was decided against as one of the objectives of the study was to use the available DYFAMED data set.

4.4.3. Adding zooplankton information—Case 2b

The above observations all point to the need to constrain zooplankton concentrations in some way. Although no synoptic zooplankton data is available, the zooplankton content was estimated to be 20 mmol N m^{-2} in May 1995 (Andersen and Prieur, 2000). Simulated zooplankton concentrations in May in the first guess experiment are consistent with this value, but they are too low in Case 1 and Case 2a (Table 6). We therefore added this zooplankton data to the cost function and conducted a Case 2 type optimization, referred to below as Case 2b.

As expected, zooplankton concentrations in May ($17.9 \text{ mmol N m}^{-2}$) are now more acceptable. The fit between observed and simulated chlorophyll and nitrate data is not altered much, except during the post-bloom period (day 04/15, Fig. 3). Indeed, as discussed above, the increase in zooplankton results in an increase in the grazing pressure on phytoplankton, particularly during the post-bloom. However, the results are still much better than in Case 1 (Table 4).

The estimated C/Chl ratio, r_c , and NO_3 restoring profile, n^r , are very similar to those of Case 2a (Fig. 9). Observation of the other parameters shown in Table 3 also reveals that modifications are only slight compared with Case 2a. The most significant changes concern the detritus sedimentation speed, which decreases from 7.8 to 5.8 m day^{-1} , the assimilated fraction of phytoplankton, which increases from 63% to 72%, and the nominal preference for phytoplank-

ton, which decreases from 0.67 to 0.63. Zooplankton grazing rate increases by only 2%, while the zooplankton loss rates decrease by less than 1%. This behavior emphasizes the strong nonlinearities in the zooplankton equation.

Improvement in the zooplankton content benefits the main fluxes (Table 5). The sedimentation of slow sinking particles (SF) decreases in response to the decrease in v_d compared to Cases 1 and 2a. This leads to a more realistic value for total particulate export (EF), which now includes a weak contribution from fast sinking particles (ZF). TP has increased to an intermediate value between that of Case 2a and the first guess experiment.

5. Discussion

5.1. Validation against independent data

To validate the methodology, the model is run for years 1995 and 1996 with the optimal parameter set and the C/Chl profile obtained after assimilation of year 1997 data (Case 2b).

The inter-annual variability in winter conditions is taken into account by not using the NO_3 preconditioning profile estimated for 1997, to simulate years 1995 and 1996. Instead, for each of the years 1995 and 1996 we set the parameters to their values observed in Case 2b and only estimate the NO_3 preconditioning profile. This results in a significant improvement of the model fit to chlorophyll profiles for years 1995 and 1996 (Table 7, Figs. 10 and 11). Surface and subsurface maxima in spring and summer are globally better simulated, although several profiles in summer–fall 1995 (09/14/95, 10/05/95, 11/15/95) are not really

Table 7
Comparison of the first guess run and the validation run with independent chlorophyll and NO_3 profile data for year 1995 and 1996

Data type	Year	First guess cost	Validation run cost
Chlorophyll profile	1995	1	0.278
Chlorophyll profile	1996	1	0.092
NO_3 profile	1995	1	0.646
NO_3 profile	1996	1	0.320

Costs are normalized.

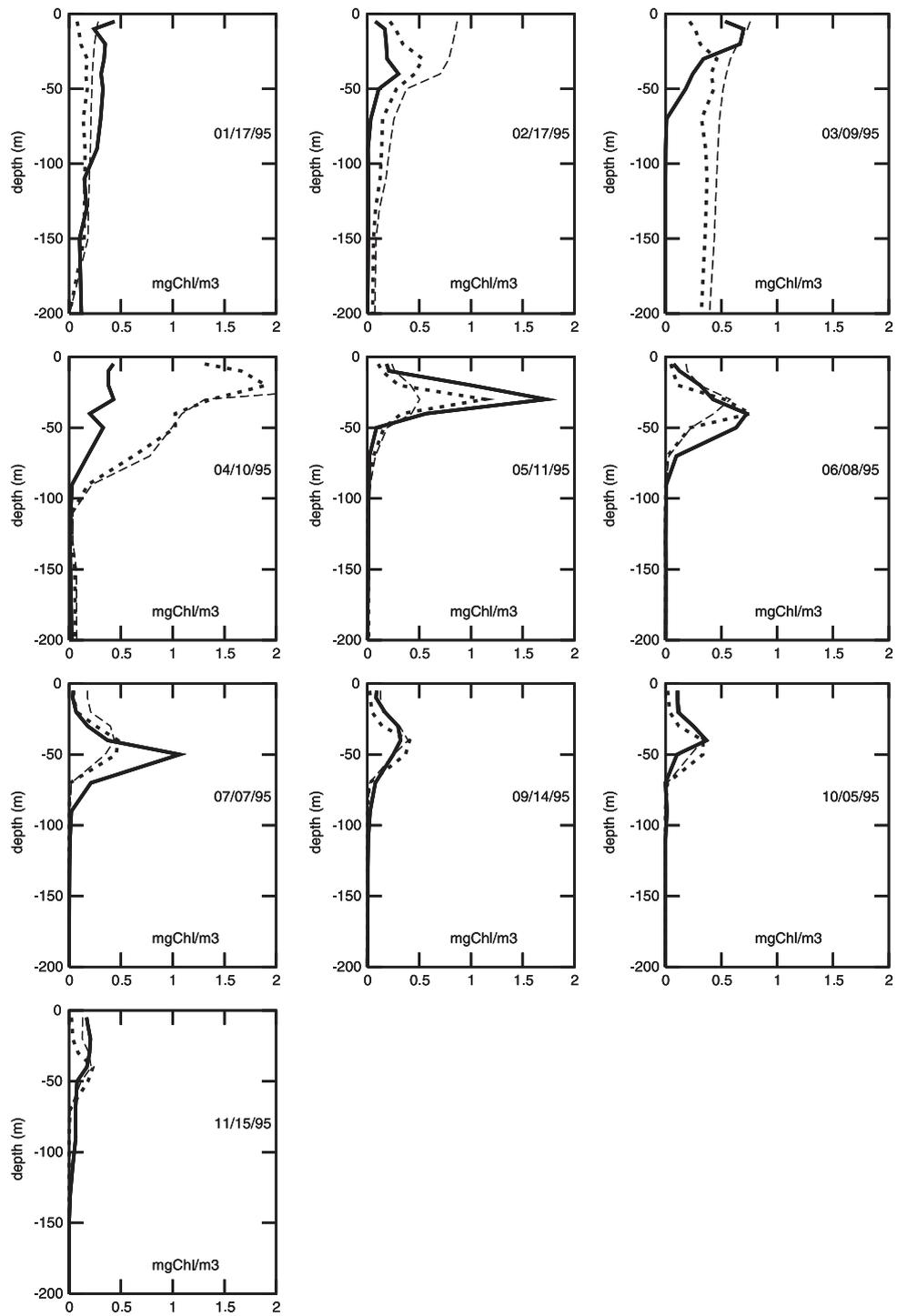


Fig. 10. Chlorophyll profiles ($\text{mg Chl } a \text{ m}^{-3}$) for year 1995. Depth in meters. The continuous line is the observation, the thin dashed line is the first guess run and the bold dashed line is the Case 2b optimal parameter set run.

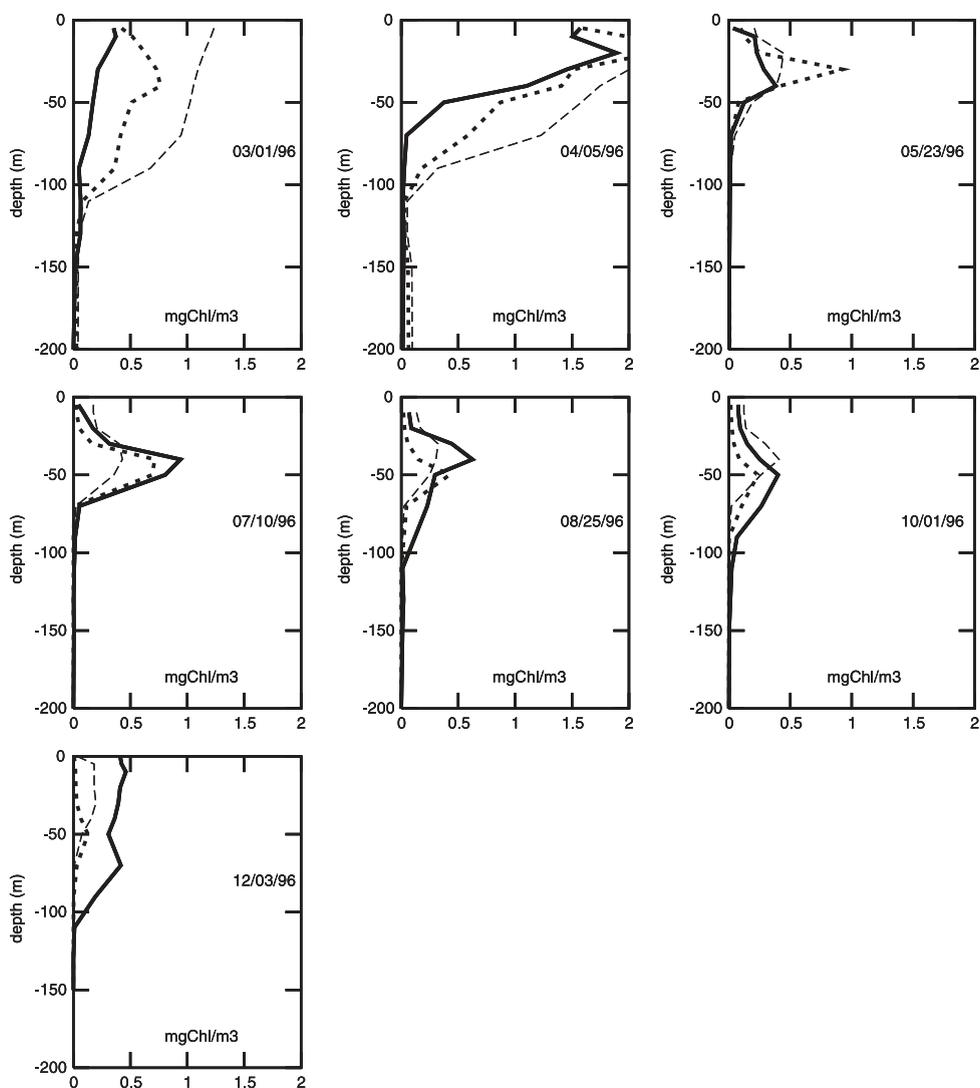


Fig. 11. Chlorophyll profiles (mg Chl a m^{-3}) for year 1996. Depth in meters. The continuous line is the observation, the thin dashed line is the first guess run and the bold dashed line is the Case 2b optimal parameter set run.

improved. NO_3 profiles are also closer to the data (Table 7, Figs. 12 and 13). Compared to year 1997 (Table 4), the agreement between data and model is globally better in 1996, but not as good in 1995. This might be linked to the ocean dynamics. Atmospheric forcing is characterized by much lower high frequency variability in 1996 than in the other years, which means that the evolution of the tracers over time is smoother (Mémery et al., 2002): vertical movements due to Ekman pumping or other dynamical processes, which are not considered in these

simulations, could be weaker in 1996. The quantitative improvement of the simulation results in 1995 and 1996 after assimilation of the year 1997 can be considered a validation of the methodology.

5.2. Stock data and flux values—Case 2c

The series of numerical experiments conducted in this work shows that the principal biogeochemical fluxes can be significantly modified by the assimilation of stock data. In addition, the experiments illustrate the

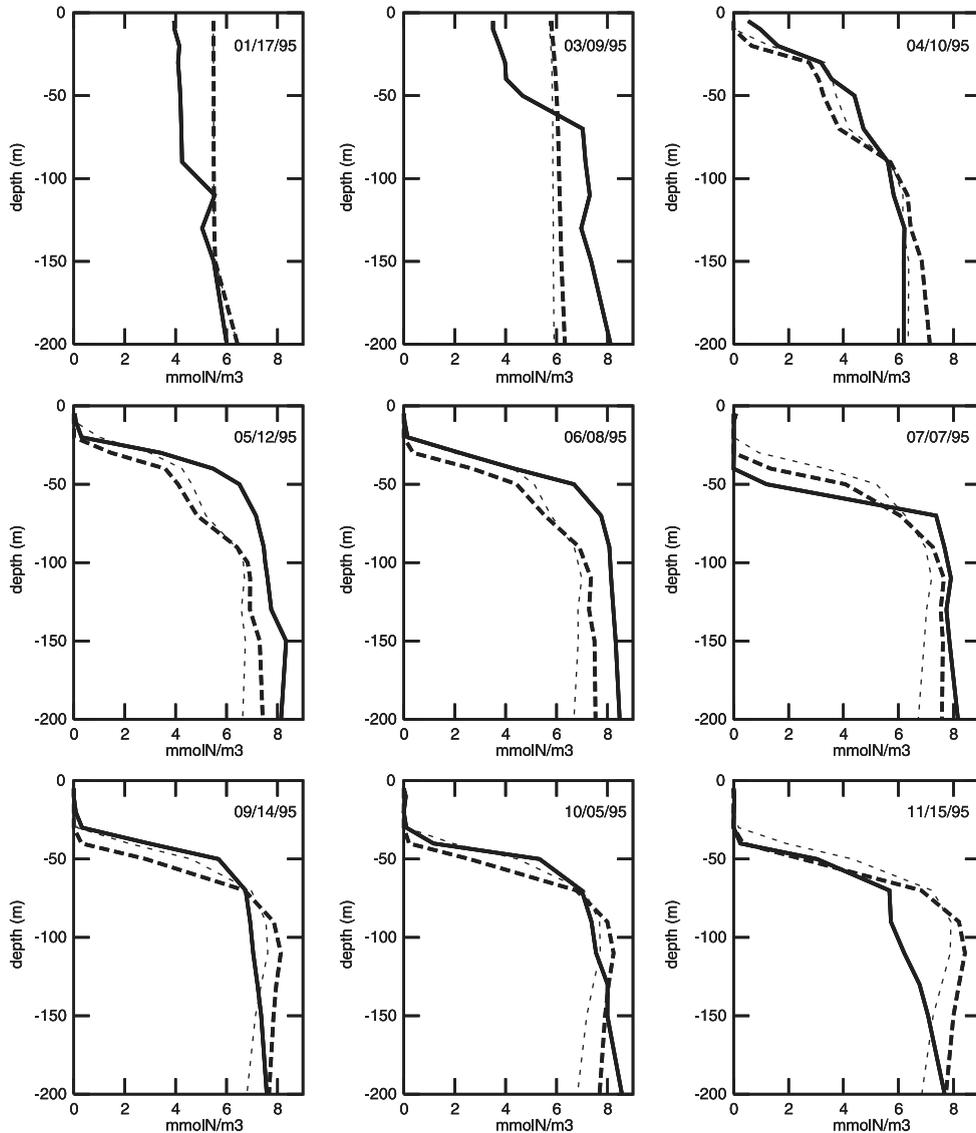


Fig. 12. NO_3 profiles (mmol N m^{-3}) for year 1995. Depth in meters. The continuous line is the observation, the thin dashed line is the first guess run and the bold dashed line is the Case 2b optimal parameter set run.

impossibility of recovering correct fluxes from the assimilation of NO_3 and chlorophyll data alone (Case 2a). When a single zooplankton measurement is added (Case 2b), flux prediction is improved. However, this improvement is not sufficient, and leads to a slight deterioration of the fit to chlorophyll data.

One possible reason for the lack of success in retrieving the fluxes could be that the model structure prohibits solutions consistent with both the stock data

and the order of magnitude of the fluxes. To test this hypothesis, a final experiment was conducted in which all types of data were assimilated: stock measurements as well as the data on the four fluxes presented in Table 5, taken from Marty and Chiavérini (2002). This experiment, referred to as Case 2c, proved to be successful. Surface chlorophyll, chlorophyll and NO_3 profile optimization results are comparable to those obtained earlier (Table 4), while the zooplankton con-

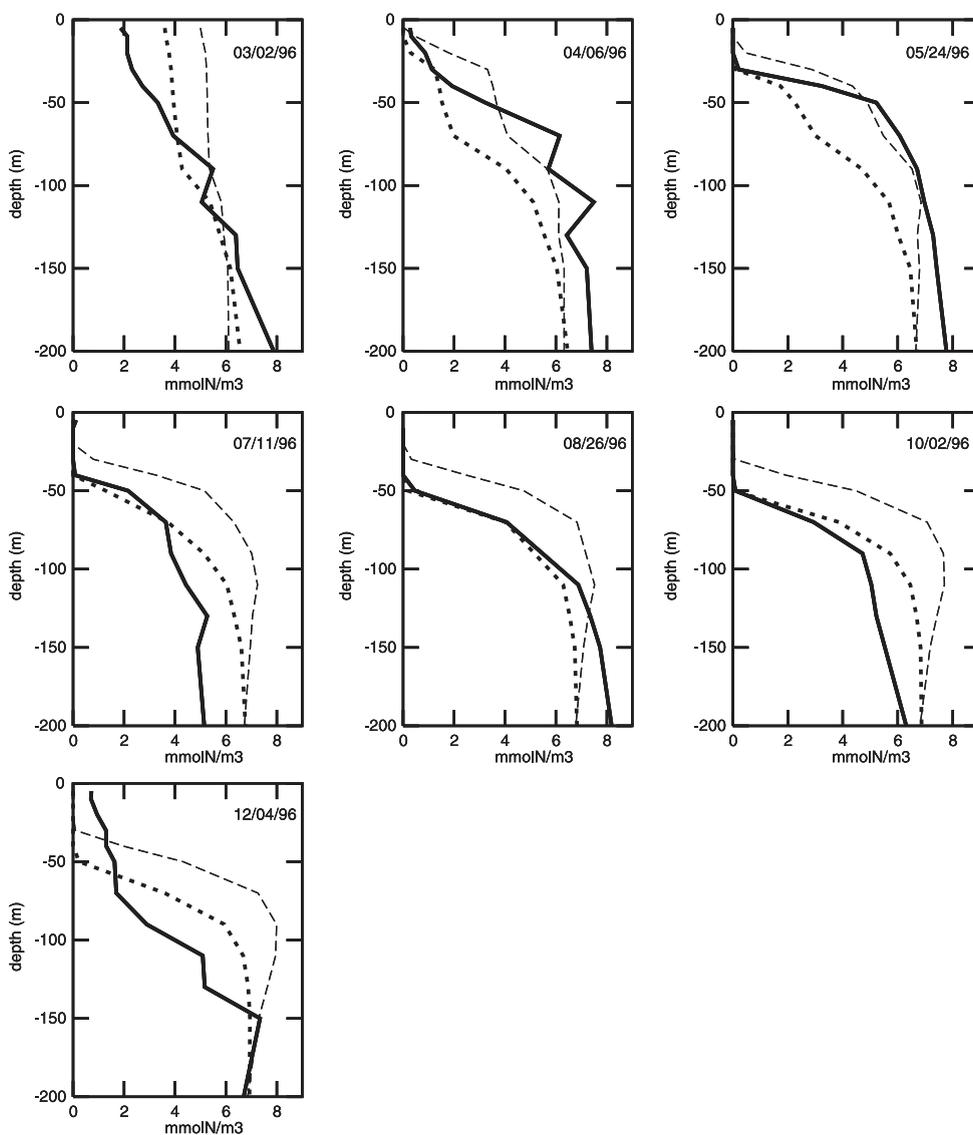


Fig. 13. NO_3 profiles (mmol N m^{-3}) for year 1996. Depth in meters. The continuous line is the observation, the thin dashed line is the first guess run and the bold dashed line is the Case 2b optimal.

tent in May is $22.7 \text{ mmol N m}^{-2}$, which is the best we obtain with the data (Table 6). More importantly, the predicted fluxes are now correct (Table 5). The hypothesis concerning the role of the model structure in the lack of success in retrieving the fluxes can therefore be rejected.

The estimated C/Chl ratio and NO_3 preconditioning profile are very similar to those found in Case 2a and Case 2b. With regard to the other parameters, presented

in Table 3, the most striking changes compared with the previous cases appear for k_n , γ , μ_d and ν_d . The half-saturation constant, k_n , and the exudation fraction, γ , were found to be among the less sensitive parameters in Section 4.2 and, not surprisingly, were not greatly modified by the data assimilation in Case 1, Case 2a and Case 2b. This was no longer true in Case 2c where production flux data seem to constrain these parameters. Sedimentation speed, ν_d , for which the optimal

value was always found to be higher than the first guess value, is now lower, an observation that can be related to the decrease in the export flux imposed by the data. Furthermore, detritus feeds the NH_4 and DOM compartments at a higher rate μ_d .

Unlike for chlorophyll data, the relative cost of the NO_3 profiles is higher in Case 2c than in Case 2b (Table 4). Unlike in the first guess, the flux that had to be modified in Case 2c was particle export production (Table 5). To achieve this new balance, the inversion degraded the fit of the NO_3 profiles. Although a bias due to dynamical processes cannot be entirely discounted, this finding could result from an underestimation of particle export production, obtained from shallow sediment traps (Buesseler, 1991).

6. Conclusions

The present study has demonstrated the feasibility of using the adjoint data assimilation method to constrain the biological part of a one-dimensional coupled physical–biogeochemical model with observed data from the DYFAMED station in the Northwestern Mediterranean Sea.

A preliminary sensitivity analysis showed that not all parameters can be estimated and enabled the inverse problem to be correctly formulated. The correctness of the formulation was demonstrated using identical twin experiments.

Real surface chlorophyll data as well as chlorophyll and NO_3 profile measurements were assimilated for the year 1997. The assimilation method required choosing a judicious first guess. Different assimilation exercises then led to a number of conclusions.

As with earlier modeling studies (Geider et al., 1997), it was quantitatively shown that it is impossible to simulate correct surface chlorophyll bloom intensity together with a correct summer subsurface chlorophyll maximum in an oligotrophic regime using a constant C/Chl ratio. We used a C/Chl ratio allowing vertical variability to improve data assimilation results.

It was also quantitatively demonstrated that in our IDDYFAMED context, the optimization of winter NO_3 conditions is of prime importance to data assimilation efficiency. The winter restoring profile governs the quantity of nitrate present in the model and needs to be controlled. Some estimated parameter values

appeared to be dependent on this quantity. Moreover, NO_3 concentrations below the surface layer are largely imposed by this restoring profile.

Finally, a good improvement in the chlorophyll model-data fit was obtained. As for NO_3 data, the improvement was less significant below a depth of 100 m where data are affected by 3D physics not taken into account in the model. The method was then validated using an independent data set for years 1995 and 1996, and a new simulation conducted. This calculated acceptable fluxes and also predicted chlorophyll and NO_3 profiles more correctly than the first guess simulation.

The last but perhaps most important point concerns the simulation of production and export fluxes. It was shown that, using a simple NNPZD-DOM model, it is not possible to recover correct fluxes simply through the assimilation of NO_3 and chlorophyll stock data alone, although the model was able to simulate correct stocks together with correct fluxes. This failure was due to the collapse of zooplankton in the model. Non-linearities, particularly in the zooplankton conservation equation, are certainly responsible for this behavior: minor changes in the parameters can considerably modify system evolution. Nevertheless, the addition of zooplankton data to the cost function did not prove to be entirely successful or enable the correct fluxes to be recovered. As stocks result from a balance between input and output fluxes, it may be that information associated with stocks alone cannot constrain the different fluxes independently. This conclusion implies that it may be impossible to estimate the biogeochemical fluxes accurately if a first guess of these fluxes is not taken into account and assimilated. Further studies, using other models and different data coverage, are needed to determine whether this statement can be generalized.

References

- Andersen, V., Prieur, L., 2000. One-month study in the open NW Mediterranean Sea (DYNAPROC experiment, May 1995): overview of the hydrobiogeochemical structures and effects of wind events. *Deep-Sea Res.*, Part 1 47, 397–422.
- Antoine, D., Morel, A., André, J.-M., 1995. Algal pigment distribution and primary production in the Eastern Mediterranean Sea as derived from Coastal Zone Color Scanner observations. *J. Geophys. Res.* 100, 16193–16209.

- Athias, V., Mazzega, P., Jeandel, C., 2000. Selecting a global optimization method to estimate the oceanic particle cycling rate constants. *J. Mar. Res.* 58, 675–707.
- Bethoux, J.-P., 1989. Oxygen consumption, new production, vertical advection and environmental evolution in the Mediterranean Sea. *Deep-Sea Res., Part 2* 36, 769–781.
- Buesseler, K.O., 1991. Do upper-ocean sediment traps provide an accurate record of particle flux? *Nature* 353, 420–423.
- Buesseler, K.O., Bacon, M.P., Cochran, J.K., Livingston, H.R., 1992. Carbon and nitrogen export during the JGOFS North Atlantic Bloom Experiment estimated from 234TH/238U disequilibria. *Deep-Sea Res.* 39, 1115–1137.
- Buesseler, K.O., Steinberg, D.K., Michaels, A.F., Johnson, R.J., Andrews, J.E., Valdes, J.R., Price, J.F., 2000. A comparison of the quantity and composition of material caught in a neutrally buoyant versus surface-tethered sediment trap. *Deep-Sea Res., Part 1* 47, 277–294.
- Chifflet, M., Andersen, V., Prieur, L., Dekeyser, I., 2001. One-dimensional model of short-term dynamics of the pelagic ecosystem in the NW Mediterranean Sea: effects of wind events. *J. Mar. Syst.* 30, 89–114.
- Copin-Montegut, G., Avril, B., 1993. Vertical distribution and temporal variation on dissolved organic carbon in the Northwestern Mediterranean Sea. *Deep-Sea Res.* 40, 1963–1972.
- Courtier, P., Talagrand, O., 1987. Variational assimilation of meteorological observations with the adjoint vorticity equation: II. Numerical results. *Q. J. R. Meteorol. Soc.* 113, 1329–1347.
- Evans, G.T., 1999. The role of local models and data sets in the Joint Global Ocean Flux Study. *Deep-Sea Res.* 46, 1369–1389.
- Evensen, G., Dee, D., Schröter, J., 1998. Parameter estimation in dynamical models. In: Chassignet, E.P., Verron, J. (Eds.), *Ocean Modeling and Parameterizations*. NATO ASI. Kluwer Academic Publishers, Dordrecht.
- Fasham, M.J.R., Evans, G.T., 1995. The use of optimization techniques to model marine ecosystem dynamics at the JGOFS station at 47 deg N 20 deg W. *Philos. Trans. R. Soc. Lond., B* 348, 203–209.
- Fasham, M.J.R., Ducklow, H.W., McKelvie, S.M., 1990. A nitrogen-based model of plankton dynamics in the oceanic mixed-layer. *J. Mar. Syst.* 48, 591–639.
- Faure, C., Papegay, Y., 1997. *Odyssée Version 1.6, the language reference manual*. Rapport Technique 211, INRIA.
- Fennel, K., Losch, M., Schröter, J., Manfred, W., 2001. Testing a marine ecosystem model: sensitivity analysis and parameter optimization. *J. Mar. Syst.* 28, 45–63.
- Gaspar, P., Gregories, Y., Lefevre, J.M., 1990. A simple eddy kinetic energy model for simulations of the oceanic vertical mixing: tests at station Papa and long term upper ocean study site. *J. Geophys. Res.* 95, 16179–16193.
- Geider, R.J., MacIntyre, H.L., Kana, T.M., 1997. 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.
- Gilbert, J.C., Lemaréchal, C., 1989. Some numerical experiments with variable storage quasi-Newton algorithms. *Math. Program.* 45, 407–435.
- Griewank, A., 2000. Evaluating derivatives. Principles and techniques of algorithmic differentiation. *Frontiers in Applied Mathematics*. SIAM.
- Gunson, J., Oschlies, A., Garçon, V., 1999. Sensitivity of ecosystem parameters to simulated satellite ocean color data using a coupled physical–biological model of the North Atlantic. *J. Mar. Res.* 57, 613–639.
- Hood, E.M., Merlivat, L., 2001. Annual to interannual variations of fCO₂ in the northwestern Mediterranean Sea: high frequency time series data from CARIOCA buoys, 1995–1997. *J. Mar. Res.* 59, 113–131.
- Hurtt, G.C., Armstrong, R.A., 1996. A pelagic ecosystem model calibrated with BATS data. *Deep-Sea Res., Part 2* 43, 653–683.
- Hurtt, G.C., Armstrong, R.A., 1999. A pelagic ecosystem model calibrated with BATS data and OWSI data. *Deep-Sea Res., Part 1* 46, 27–61.
- Lawson, L.M., Spitz, Y.H., Hofmann, E.E., Long, R.B., 1995. A data assimilation technique applied to a predator–prey model. *Bull. Math. Biol.* 57, 593–617.
- Lawson, L.M., Hofmann, E.E., Spitz, Y.H., 1996. Time series sampling and data assimilation in a simple marine ecosystem model. *Deep-Sea Res., Part 2* 43, 625–651.
- Le Dimet, F.X., Talagrand, O., 1986. Variational algorithms for analysis and assimilation of meteorological observations: theoretical aspects. *Tellus* 38A, 97–110.
- Lévy, M., Mémery, L., André, J.-M., 1998a. Simulation of primary production and export fluxes in the Northwestern Mediterranean Sea. *J. Mar. Res.* 56, 197–238.
- Lévy, M., Mémery, L., Madec, G., 1998b. The onset of a bloom after deep winter convection in the Northwestern Mediterranean Sea: mesoscale process study with a primitive equation model. *J. Mar. Syst.* 16, 7–21.
- Lions, J.L., 1971. *Optimal Control of System Governed by Partial Differential Equations*. Springer, Berlin.
- Long, R.B., Thacker, W.C., 1989a. Data assimilation into a numerical equatorial ocean model, Part I: the model and the assimilation algorithm. *Dyn. Atmos. Ocean.* 13, 379–412.
- Long, R.B., Thacker, W.C., 1989b. Data assimilation into a numerical equatorial ocean model: Part II. Assimilation experiments. *Dyn. Atmos. Ocean.* 13, 313–440.
- Luong, B., Blum, J., Verron, J., 1998. A variational method for the resolution of data assimilation problem in oceanography. *Inverse Probl.* 14, 979–997.
- Madec, G., Chartier, M., Delecluse, P., Crepon, M., 1991. A three-dimensional numerical study of deep water formation in the Northwestern Mediterranean Sea. *J. Phys. Oceanogr.* 21, 1349–1371.
- Martin, J.H., Fitzwater, S.E., 1992. Dissolved organic carbon in the Atlantic Southern and Pacific Oceans. *Nature* 356, 699–700.
- Marty, J.-C., Chiavérini, J., 2002. Seasonal and interannual variations in phytoplankton production at DYFAMED time-series station, Northwestern Mediterranean Sea. *Deep-Sea Res., Part 2*, 49/11, 2017–2030.
- Matear, R.J., 1995. Parameter optimization and analysis of ecosystem models using simulated annealing: a case study at Station P. *J. Mar. Res.* 53, 571–607.
- Mémery, L., Lévy, M., Vérant, S., Mervilat, L., 2002. The relevant

- time scales in estimating the air–sea CO₂ exchange in a mid latitude region. *Deep-Sea Res., Part 2*, 49/11, 2067–2092.
- Míñas, H.J., 1970. La distribution de l'oxygène en relation avec la production primaire en Méditerranée nord-occidentale. *Mar. Biol.* 7, 181–204.
- Miquel, J.-C., Flower, S.W., La Rosa, J., 1993. Vertical particle fluxes in the Ligurian Sea. *Ann. Inst. Oceanogr.* 69, 107–110.
- Prunet, P., Minster, J.F., Ruiz-Pino, D., Dadou, I., 1996a. Assimilation of surface data in a one-dimensional physical–biogeochemical model of the surface ocean: 1. Method and preliminary results. *Glob. Biogeochem. Cycles* 10, 111–138.
- Prunet, P., Minster, J.F., Echevin, V., Dadou, I., 1996b. Assimilation of surface data in a one-dimensional physical–biogeochemical model of the surface ocean: 2. Adjusting a simple trophic model to chlorophyll, temperature, nitrate, and pCO₂ data. *Glob. Biogeochem. Cycles* 10, 139–158.
- Richardson, K., 1993. The ICES 14C primary production measurement intercomparison exercise. *MEPS* 72, 189–201.
- Schartau, M., Oschlies, A., Willebrand, J., 2001. Parameter estimates of a zero-dimensional ecosystem model applying the adjoint method. *Deep-Sea Res., Part 2* 48, 1769–1800.
- Spitz, Y.H., Moisan, J.R., Abbot, M.R., Richman, J.G., 1998. Data assimilation and a pelagic ecosystem model: parameterization using time series observations. *Deep-Sea Res., Part 2* 16, 51–68.
- Spitz, Y.H., Moisan, J.R., Abbot, M.R., 2001. Configuring an ecosystem model using data from the Bermuda Atlantic Series (BATS). *Deep-Sea Res., Part 2* 48, 1733–1768.
- Talagrand, O., Courtier, P., 1987. Variational assimilation of meteorological observations with the adjoint vorticity equation: I. *Theory. Q. J. R. Meteorol. Soc.* 113, 1311–1328.
- Thacker, W.C., 1989. The role of the Hessian in fitting models to measurements. *J. Phys. Oceanogr.* C94, 6177–6196.
- Tikhonov, A.N., Arsenin, V.Y., 1977. *Solutions of Ill-posed problems.* Winston, Washington, DC.