

Assimilation of ocean colour data into a Biogeochemical Flux Model of the Eastern Mediterranean Sea

G. Triantafyllou¹, G. Korres¹, I. Hoteit², G. Petihakis¹, and A. C. Banks¹

¹Hellenic Centre for Marine Research, Anavissos, Greece

²Scripps Institution of Oceanography, La Jolla, CA, USA

Received: 2 August 2006 – Published in Ocean Sci. Discuss.: 21 September 2006

Revised: 18 January 2007 – Accepted: 14 August 2007 – Published: 21 August 2007

Abstract. An advanced multivariate sequential data assimilation system has been implemented within the framework of the European MFSTEP project to fit a three-dimensional biogeochemical model of the Eastern Mediterranean to satellite chlorophyll data from the Sea-viewing Wide Field-of-view Sensor (SeaWiFS). The physics are described by the Princeton Ocean Model (POM) while the biochemistry of the ecosystem is tackled with the Biogeochemical Flux Model (BFM). The assimilation scheme is based on the Singular Evolutive Extended Kalman (SEEK) filter, in which the error statistics were parameterized by means of a suitable set of Empirical Orthogonal Functions (EOFs). To avoid spurious long-range correlations associated with the limited number of EOFs, the filter covariance matrix was given compact support through a radius of influence around every data point location. Hindcast experiments were performed for one year over 1999 and forced with ECMWF 6 h atmospheric fields. The solution of the assimilation system was evaluated against the assimilated data and the MedAtlas climatology, and by assessing the impact of the assimilation on non-observed biogeochemical processes. It is found that the assimilation of SeaWiFS data improves the overall behavior of the BFM model and efficiently removes long term biases from the model despite some difficulties during the spring bloom period. Results, however, suggest the need of subsurface data to enhance the estimation of the ecosystem variables in the deep layers.

ration, low rainfall and river runoff, in conjunction with the outflow of the bottom layer at Gibraltar, result in an oligotrophic gradient toward the east. The observed oligotrophy is thought to be due to low phosphorus concentrations (decreasing from west to east) limiting phytoplankton and bacterial growth (Krom et al., 1992). The state of the art in modelling the circulation of the Mediterranean Sea is an operational system based on physical components (Pinardi et al., 2003). While these systems exist for ocean physics, the scientific knowledge and technological capacity to construct such a system for the ecosystem is currently lacking. The three dimensional modelling of marine ecosystems is lagging behind the modelling of marine physics, as it requires robust hydrodynamic models and adequate computing resources. Furthermore, ecosystem models need to be adjusted with biological, physical and chemical data at relevant space-time scales to achieve accurate simulation of the variability of the area under study. The framework of data assimilation provides the appropriate tools for improving models' agreement with data and for enhancing their predictive capabilities. An assimilation system is composed of two basic components, an observing system and a numerical model complemented with a data assimilation scheme that can efficiently extract the reliable information from the observations to optimally initialize the forecast. Although assimilation systems for meteorological and oceanic models are well established, the use of data assimilation techniques with marine ecosystem models is far less developed.

One of the few information sources on the state of coastal and pelagic ecosystems comes from "ocean colour" satellite remote sensing (Platt et al., 1995). Since the launch of the Coastal Zone Colour Scanner on the Nimbus 7 satellite in 1978, satellite borne ocean colour sensors have become the standard tool for determining distributions of phytoplankton and other biogeochemical parameters in the ocean (IOCCG, 1999). Satellite data benefits from a high temporal resolution with repeated coverage of the same area of the sea

1 Introduction

The Mediterranean Sea is characterized by prominent morphological discontinuities defining sub-basins with different hydrodynamic and ecological conditions. The high evapo-

Correspondence to: G. Triantafyllou
(gt@ath.hcmr.gr)

surface on a daily basis. The quality of the data is, however, limited by the ability of remote sensing of ocean colour, through the analysis of ocean leaving radiance, to yield information on water-quality parameters such as phytoplankton pigments (more precisely chlorophyll-a and phaeophytin-a), suspended sediment, and yellow substance (gelbstoff) in the euphotic layer (Tassan, 1994). Chlorophyll-a is a key parameter in the model as it is calculated from each phytoplankton group separately and describes the total primary biomass. Additionally with the use of remote sensing techniques, adequate data of Chlorophyll-a will become more and more available offering significant opportunities for data assimilation systems. Therefore the appropriate use of satellite data may guide the modelling and forecasting processes closer to realistic conditions. However, the key to the use of these data in ecosystem modelling/forecasting research relies on the availability of two premises: a well calibrated algorithm to retrieve accurate chlorophyll-a estimates from the original spectral data, and an assimilation technique capable of efficiently assimilating these surface data into the ecosystem model. In the domain of marine ecology, early studies used four-dimensional variational data assimilation techniques for estimating poorly known parameters in the model. Such techniques basically seek for the unknown parameters that minimize the misfit between model simulations and data, e.g. the adjoint method (Fennel et al., 2001; Friedrichs, 2001; Gunson et al., 1999; Lawson et al., 1996; Spiitz et al., 1998; Vallino, 2000), direct minimization methods (Fasham and Evans, 1995; Prunet et al., 1996), and Monte Carlo methods (Harmon and Challenor, 1996; Matear, 1995). Recently, focus shifted toward the use of Kalman filter based sequential assimilation techniques to directly compute estimates of the system state, as these methods allow for efficient handling of the model uncertainties while intermittently adjusting the model trajectory each time new observations are available (Ghil and Malanotte-Rizzoli, 1991). Taking into account (even partially) the model error is a key step for building a successful ecosystem assimilation system because of significant uncertainties in the current ecological models. For instance, Anderson et al. (2000) used optimal interpolation to assimilate both physical and biological data into a mesoscale-resolution 3-D ocean model. Allen et al. (2002) and Natvik and Evensen (2002) demonstrated the effectiveness of the well-known ensemble Kalman filter for data assimilation with a 1-D, three-component model. The Singular Evolutive Extended Kalman (SEEK) filter, which is a suboptimal extended Kalman (EK) filter developed by Pham et al. (1997), was particularly popular and successfully implemented in several marine ecosystem data assimilation studies. The SEEK filter operates with low-rank matrices to avoid the prohibitive computational cost of the EK filter while only adjusting the model forecast in the directions of error growth. It further supports different degrees of simplifications in the evolution of its “correction directions” at a minimal loss of performance (Hoteit et al., 2002, 2004).

Carmillet et al. (2001) used the SEEK filter with an invariant set of Empirical Orthogonal Functions (EOFs) correction directions to assimilate pseudo-ocean colour data into a 3-D physical–biogeochemical model of the North Atlantic Ocean. Hoteit et al. (2003a) used the same filter with a 1-D complex ecosystem model of the Cretan Sea assimilating real observations of oxygen and nitrate and validating the filter with behavior independent chlorophyll data. Triantafyllou et al. (2003) implemented the ensemble variant of the SEEK filter in a three dimensional ecosystem model of the Cretan Sea. In the same area, (Hoteit et al., 2005) successfully tested the SEEK filter with semi-evolutive correction directions composed of global and local EOFs. The reader is referred to Triantafyllou et al. (2005) for a review on the implementation of the SEEK filter and its variants in different shelf and regional areas of the Mediterranean Sea.

One of the major goals of the Mediterranean ocean Forecasting System (MFS) project during its second phase (2003–2006), named Toward Environmental Predictions (MFSTEP), was the development of numerical forecasting systems at basin and regional scale. Within the framework of this project, one of the tasks was to implement the SEEK filter to assimilate Sea viewing Wide Field of view Sensor (SeaWiFS) data into a coupled physical (Princeton Ocean Model) – biological (Biogeochemical Flux Model) model of the Eastern Mediterranean, developed during the first phase of the MFSTEP project. This study presents the first attempt to use an advanced Kalman filtering technique for the assimilation of ocean colour data into a complex state-of-the-art three-dimensional marine ecosystem model. The controllability of the ecosystem variability using satellite measurements is a major question in marine ecology and will be addressed here. After presenting the physical and ecological components of the Eastern Mediterranean ecosystem model in Sect. 2, a general overview of the SeaWiFS data is provided in Sect. 3. The assimilation scheme is described in Sect. 4. Assimilation results of SeaWiFS data into the coupled model are reported and the behavior of the assimilation system is discussed in Sect. 5. A general conclusion, including a discussion on the progress made thus far and the problems that still need to be addressed, is offered in Sect. 6.

2 The Eastern Mediterranean ecosystem model

The Eastern Mediterranean ecosystem model consists of two, on-line coupled sub-models: the Princeton Ocean Model (POM), which describes the hydrodynamics of the area, and provides the physical forcing to the second sub-model, the Biogeochemical Flux Model (BFM).

2.1 The physical model

The hydrodynamic model is based on the Princeton Ocean Model (POM), a primitive equation, 3-D circulation model.

POM has been extensively described in the literature (Blumberg and Mellor, 1987; Horton et al., 1997; Lascaratos and Nittis, 1998) and is accompanied by a comprehensive User's guide (Mellor, 1998). It has been previously used in the Mediterranean area by Drakopoulos and Lascaratos (1997) and Zavatarelli and Mellor (1995) and in the eastern Levantine basin by Lascaratos and Nittis (1998) and Korres and Lascaratos (2003). The model has a bottom – following vertical sigma coordinate system, a free surface and a split mode time step. Potential temperature, salinity, velocity and surface elevation, are prognostic variables. Horizontal diffusion terms are evaluated using the Smagorinsky (1963) diffusion formulation while the vertical mixing coefficients are computed according to the Mellor-Yamada 2.5 turbulence closure scheme (Mellor and Yamada, 1982). The model has one open boundary located at 20° E as shown in Fig. 1 where open boundary conditions apply. The computational grid has a horizontal resolution of $1/10^\circ \times 1/10^\circ$ and 25 sigma levels in the vertical with a logarithmic distribution near the sea surface, which results in a better representation of the surface mixed layer. As for the min/max depths of the first sigma level these range from 0.08 to 11.88 m. Within the first 30 m of the water column there are at minimum 1 sigma levels (at grid points with depths greater than 3250 m) and at maximum 24 levels (at grid points with depths of 25–29 m). Considering the size (10–14 km) of the internal Rossby radius of deformation for the Eastern Mediterranean basin, such a model resolution (~ 10 km) can marginally resolve the mesoscale eddy activity. The U.S. Navy Digital Bathymetric Data Base 5 ($1/12^\circ \times 1/12^\circ$) was used for building up the model's bathymetry using bilinear interpolation to map the data onto the model's grid. The model is very similar to the ALERMO model used in Korres and Lascaratos (2003) with the only exception being the horizontal resolution (coarser) as the computational burden in this particular study was very high due to coupling with the BFM, the incorporation of data assimilation and the execution of several sensitivity tests. The model includes parameterization of the Dardanelles outflow into the Aegean Sea, the runoff of the major rivers of the Thermaikos Gulf (Aliakmonas, Axios and Loudias) and the runoff of Nestos and Evros rivers to north-central and north-eastern Greece. The Dardanelles outflow into the Aegean Sea is a dominant factor for the freshwater budget of the basin, providing approximately 300 km^3 of brackish water on an annual basis. The main Greek rivers (Axios, Aliakmonas, Gallikos, Pinios, Sperchios, Evros, Strimonas and Nestos) on the other hand, with a total runoff of $\sim 19 \text{ km}^3/\text{yr}$, have a much lower contribution. Even lower is the contribution of the Turkish rivers with a total runoff of $\sim 5 \text{ km}^3/\text{yr}$.

2.2 Biogeochemical Flux Model (BFM)

BFM is a generic highly complex model based on the European Regional Seas Ecosystem Model (ERSEM) (Baretta et al., 1995; Vichi et al., 2004). A detailed description of

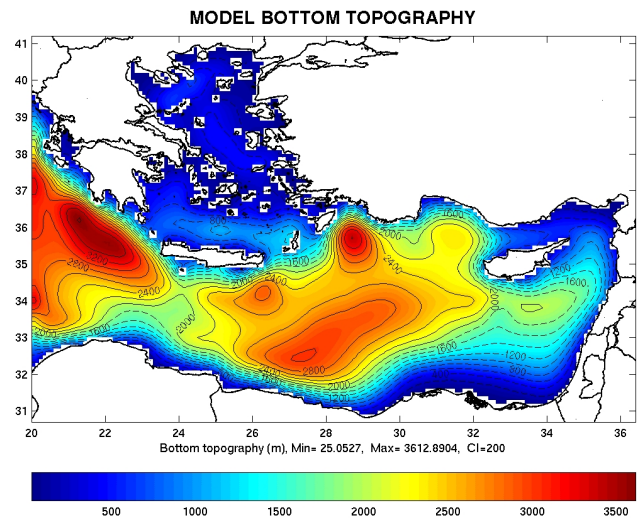


Fig. 1. Bathymetric map of the Eastern Mediterranean model domain.

the BFM is provided by Vichi et al. (2007a, b), and its implementation in the Eastern Mediterranean can be found in a companion paper (Petihakis et al., 2007). Hereafter a brief description is given. As in ERSEM the model uses a functional group approach separating the organisms according to their trophic level (producers, consumers and decomposers) and further subdivided on the basis of their trophic links and/or size. Although within each trophic level the groups have the same processes, differentiation is achieved through the different parameter values. All the important physiological (ingestion, respiration, excretion and egestion), and population (growth, migration and mortality) processes are included, and are described by fluxes of carbon and nutrients. Carbon is the basic unit cycled in the system, followed by macronutrients, chlorophyll and oxygen, with variable carbon/nutrients and carbon/chl-*a* ratios. Following the model's food web, diatoms are preyed on by microzooplankton and omnivorous mesozooplankton, nano-phytoplankton to an extent by heterotrophic nanoflagellates but mostly by microzooplankton, pico-phytoplankton mostly by heterotrophic nanoflagellates and to a lesser extent by microzooplankton and finally flagellates by microzooplankton and omnivorous mesozooplankton. Bacteria consume Dissolved Organic Carbon (DOC) both labile and semi-labile, act as decomposers on Particulate Organic Carbon (POC) and compete with phytoplankton for inorganic nutrients. Their main predators are the heterotrophic nanoflagellates. Heterotrophic nanoflagellates are preyed on by microzooplankton which in turn is eaten by omnivorous mesozooplankton. Omnivorous mesozooplankton is preyed on by carnivorous mesozooplankton which is the top predator of the food chain. For all consumers there is feeding within the same functional group (cannibalism), which acts as a stabilizing mechanism.

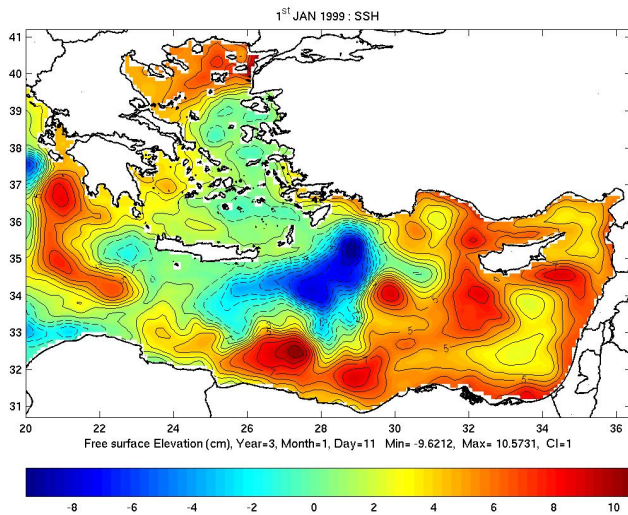


Fig. 2. The free surface elevation distribution for the 1 January 1999.

2.3 Climatological run

The physical model climatological run was initialized with the Mediterranean Ocean Data-Base (MODB-MED4) (Brasseur et al., 1996) which contains seasonal profiles of temperature and salinity mapped on a $1/4^\circ \times 1/4^\circ$ horizontal grid. These data were mapped onto the model's horizontal grid using bilinear interpolation. Temperature and salinity profiles at the eastern boundary were also derived from the same database. The biogeochemical model initial conditions for the nutrients were taken from Levitus (1982) while the other biogeochemical state variables were initialized using the same method as the one described in Petihakis et al. (2002).

The ecosystem model initialized from spring conditions was run for four years forced with a perpetual year atmospheric data set in order to reach a quasi steady state and to obtain inner fields fully coherent with the boundary conditions. More specifically the momentum budget at the sea surface was specified according to the ECMWF wind stress monthly climatology (Korres and Lascaratos, 2003). The heat flux and water flux boundary conditions at the surface were set as follows:

$$\rho_o c_p K_H \left. \frac{\partial T}{\partial z} \right|_{z=0} = Q_T - Q_{\text{SOL}} + c_1 (T^* - T_1) \quad (1)$$

$$w_{\sigma=0} = E - P - R + c_2 \frac{S^* - S_1}{S_1} \quad (2)$$

where Q_T is the monthly average total heat flux field, E is the evaporation rate taken from the Kondo-Bignami monthly climatology (Kondo, 1975), precipitation rate P is taken from Jaeger's monthly climatology (Jaeger, 1976) and T^* , S^* fields are taken from MODB-MED4 SST and SSS seasonal

climatology. Solar radiation flux Q_{SOL} is calculated with the Reed formula (Reed, 1977) using the ECMWF monthly cloud cover data.

The ecosystem pelagic state variables along the open boundary are described by solving water column 1-D ecosystem models at each grid point on the open boundary. For the physical model, open boundary conditions were set as follows: Zero gradient condition for the free surface elevation. Flather boundary condition (Flather, 1976) for the barotropic velocity normal to the open boundary. Sommerfeld radiation for the internal (baroclinic) velocities (Sommerfeld, 1949). Temperature and salinity at the open boundary are advected upstream. When there is inflow through the open boundary, temperature and salinity profiles are prescribed from the MODB-MED4 seasonal climatology.

2.4 Hindcast experiment

In this experiment the model was integrated for year 1999 (January–December 1999) initialized from the climatological run. At the same time the model was asynchronously (off-line) coupled with the coarse resolution ($0.5^\circ \times 0.5^\circ$) ECMWF 6 h atmospheric data (wind velocity, air temperature, relative humidity and cloud cover) for the same period of time covering the whole Mediterranean basin. This set of atmospheric data was used by the air-sea interaction scheme of the physical model for the estimation of heat, freshwater and momentum fluxes at the sea surface. In order to adjust the basin climatological dynamics to the interannual forcing, the model was integrated for two years using the same atmospheric data set. The free surface elevation distribution for the 1st Jan 1999 (during the second year of model integration) is shown in Fig. 2. The physical model setup is able (despite its rather coarse resolution) to represent several key features of the Eastern Mediterranean general circulation. As shown in Fig. 2, the model produces an intense and elongated Rhodes gyre with a strong two-lobe Mersha Matruh anticyclone to the south. At the eastern end of the basin a stretched Shikmona anticyclone is reproduced by the model while to the southwest of Peloponissos the signature of the Pelops gyre is evident.

3 Ocean color data

Daily data from the SeaWiFS sensor onboard the SeaStar satellite at local area coverage spatial resolution (~ 1.1 km by 1.1 km pixels) for one year (1999) were used for this study. This represented more than 20 GB of data which were acquired from NASA Goddard through the NASA Ocean Colour Web and ftp service (Feldman and McClain, 2004). The data were processed by the NASA SeaDAS software v4.0 (Baith et al., 2001) from the original spectral data using default values from the SeaWiFS level 2 product processing chain which includes atmospheric correction, georeferencing

and chlorophyll-a retrieval OC4v4 (O'Reilly et al., 2000; Patt et al., 2003). The OC4v4 algorithm, which was designed for biological studies of the global oceanic environment, uses the following fourth order polynomial equation to calculate chlorophyll-a estimates

$$\log_{10}(\text{chl-a}) = 0.366 - 3.067R_{4S} + 1.930R_{4S}^2 + 0.649R_{4S}^3 - 1.532R_{4S}^4 \quad (3)$$

where R is reflectance, R_{4S} is the maximum value of R_{443}/R_{555} , R_{490}/R_{555} or R_{510}/R_{555} and 443, 490, 510 and 555 represent the wavelengths of the four SeaWiFS bands used. This empirical method of retrieving chlorophyll-a provided reasonably accurate estimates for most case-I waters (Bricaud et al., 2002), except for concentrations under certain oligotrophic conditions, and was found inaccurate for case-II waters of the Eastern Mediterranean (Bricaud et al., 2002; Sancak et al., 2005). This means that the data used in this study may occasionally be spoiled by important observational errors. At the time of processing and data assimilation, these SeaWiFS data were, however, the best available satellite derived chlorophyll-a concentration estimates for the area of interest. In principle, observational errors can be accounted for in the assimilation system if accurate information about the statistics of the observational errors were available (see Sect. 4). In this feasibility study, errors statistics were estimated from the data sets, while simply identifying large observational errors over areas of strong ecosystem variability. More accurate estimates of these errors will be however beneficial for any future operational analysis/forecasting system. The daily chlorophyll-a estimates were remapped to a flat grid using a cylindrical equidistant projection, again using the NASA SeaDAS software. Finally, 8-day averages were produced that took into account a land mask and data masked by clouds, and provided gridded chlorophyll estimates suitable for assimilation into the coupled hydrodynamic-ecosystem models. These pre-processed data were used every 8 days to correct the state of BFM using a sequential data assimilation procedure that we briefly describe in the following section.

4 The assimilation scheme

The assimilation scheme is sequential and is based on the Singular Evolutive Extended Kalman (SEEK) filter developed by Pham et al. (1997). The SEEK filter is a simplified Extended Kalman (EK) filter suitable for applications with high dimensional systems (the system dimension is denoted by N), as in meteorology and oceanography. The filter avoids the prohibitive computational burden associated to the significant size of the EK filter $N \times N$ -error covariance matrices (denoted by P) by operating with low-rank error covariance matrices. More precisely, the SEEK filter uses the standard decomposition $P = LUL^T$ of a low-rank matrix, where L and U are respectively $N \times r$ and $r \times r$ matrices, so that

numerical calculations involving P can be likewise achieved by means of L and U . This allows drastic computational savings in time and storage without requiring major changes in the EK filter's algorithm. Starting from an initial low-rank r error covariance matrix obtained via an Empirical Orthogonal Functions (EOF) analysis (see below), Pham et al. (1997) showed that when the model dynamics are perfect (no model error), the EK filter error covariance matrices always remain of the same rank r . The EK filter analysis step is then only applied along the directions of L ; hence its columns are referred to as the correction directions of the filter. When the model is imperfect, the model error can be projected onto the subspace spanned by the correction direction to avoid continuous increase in the rank of the error covariance matrices. In its most general form, the SEEK filter evolves its correction directions in time with the tangent linear model to follow changes in the model dynamics. In this study, however, these directions were kept invariant for reasons explained below. The filter's algorithm is summarized below. A more detailed description can be found in Pham et al. (1997).

4.1 The filter algorithm

After every observation cycle (8 days in the present study), a vector of observations Y_k^o , represented as a function of the system state X_k^f and a measurement of uncertainty ε_k

$$Y_k^o = H_k(X_k^f) + \varepsilon_k, \quad (4)$$

is available at a given time t_k . Assuming that a model forecast X_k^f is available at time t_k together with the associated (low-rank) error covariance matrix P_k^f decomposed as $L_k U_k L_k^T$, the SEEK filter serially entrains these observations with the forecast of the model X_k^f to produce the analysis state X_k^a with the formula

$$X_k^a = X_k^f + \mathbf{G}_k [Y_k^o - H_k(X_k^f)]. \quad (5)$$

The gain matrix \mathbf{G} , which linearly interpolates between the observations and the forecast, is given by

$$G_k = P_k^f \mathbf{H}_k^T [\mathbf{H}_k P_k^f \mathbf{H}_k^T + R_k]^{-1} = L_k U_k (\mathbf{H}_k L_k)^T R_k^{-1}, \quad (6)$$

where \mathbf{H}_k denotes the linearization of the observation operator H_k about X_k^f , and the matrix U_k is recursively updated according to

$$U_k^{-1} = [U_{k-1} + P_{L_k}^T Q_k P_{L_k}]^{-1} + (\mathbf{H}_k L_k)^T R_k^{-1} \mathbf{H}_k L_k \quad (7)$$

with Q_k being the covariance matrix of model uncertainties and $P_{L_k} = (L_k^T L_k)^{-1} L_k$ the projection operator onto the subspace spanned by the columns of L_k . Once the analysis is made, a new forecast of the state is obtained by advancing the previous analyzed state X_k^a with the model

$$X_{k+1}^f = M_{k,k+1}(X_k^a), \quad (8)$$

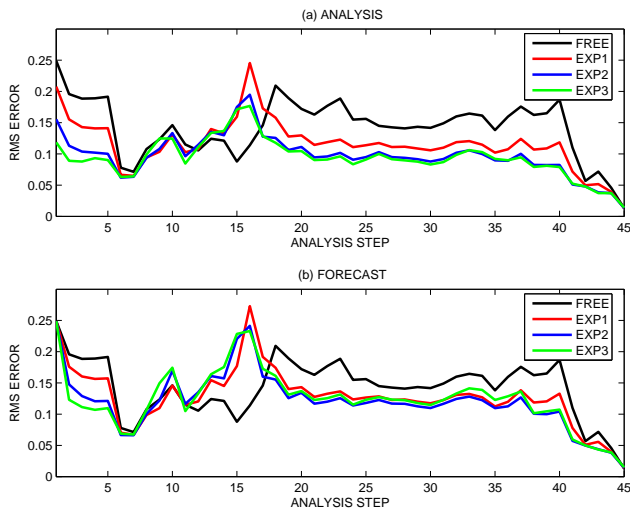


Fig. 3. Chlorophyll-a RMS error for the analysis and the forecast for three different choices of the radius of influence (EXP1: 70 km, EXP2: 150 km, and EXP3: 250 km) along with the free run.

where M is the transition operator representing the model dynamics. The covariance matrix of the forecast error is then

$$P_{k+1}^f = L_{k+1}U_kL_{k+1} + Q_k, \quad (9)$$

where the new correction directions L_{k+1} evolve in time with the tangent linear model $\mathbf{M}_{k,k+1}$ (evaluated about the analyzed state X_k^a) according to

$$L_{k+1} = \mathbf{M}_{k,k+1}L_k. \quad (10)$$

The evolution of the correction directions is generally beneficial to keep track of changes in the model dynamics (Hoteit et al., 2002; Hoteit and Pham, 2003b). The numerical integration of Eq. (9) requires, however, $r+1$ times forecast model runs, which can be rather significant with a heavily loaded coupled physical-biogeochemical model such as the one used in this study. Following Brasseur et al. (1999) who found that a reliable set of invariant EOFs describing the dominant modes of the system may provide a good correction subspace for quasi-linear dynamical models, the correction directions of the SEEK filter were kept invariant in the present study. Theoretically, this can be supported by assuming that the ecosystem state generally undergoes little change between two consecutive observations, which allows approximating $\mathbf{M}_{k,k+1}$ by the identity matrix. In practice, several studies (e.g. Carmillet et al., 2001; Hoteit et al., 2003) suggested that performance losses associated with this approximation were not significant given the achieved drastic reduction in the computational burden of the SEEK filter as only one model integration is required in this case for the computation of the forecast state.

4.2 Localization of the filter analysis

The low-rank approximation generally results in very few degrees of freedom for the filter analysis to fit available observations. Another difficulty in the assimilation system is that the initial EOFs correction directions are not updated with model dynamics. These functions, especially those associated with the least energetic modes, can be spoiled with spurious auto/cross correlations, which inevitably introduces noise into the filter analysis. As suggested by Houtekamer and Mitchel (2001), a simple strategy to deal with this problem is to exclude observations greatly distant from the grid point being analyzed. This allows the retention of the structures of the short-range correlations in the filter's error covariance matrices, which are assumed to be more reliable, while filtering out long-range correlations. This "localization" of the filter analysis can be efficiently implemented through a Schur product (an element by element multiplication) of the error covariance matrix and a correlation function with local support (Gaspari and Cohn, 1999). In this approach, the filter's gain in Eq. (7) is reformulated as

$$G_k = (\gamma \circ P_k^f) \mathbf{H}_k^T [\mathbf{H}_k (\gamma \circ P_k^f) \mathbf{H}_k^T + R_k]^{-1}, \quad (11)$$

where $\gamma \circ P_k^f$ denotes the Schur product of the forecast covariance matrix P_k^f with the localization function γ . Although this formulation entails an approximation in the filter's algorithm, it is naturally supported by the fact that only data points located in the "neighbourhood" of an analyzed grid point should contribute to the analysis at this point.

In the current system, the localization function is defined by means of a radius of influence d (in km) around the analyzed grid point. All data located outside this horizontal ($\times 24$ vertical levels) area of influence are not retained in the analysis. Assimilation experiments were performed in order to find an appropriate value for d . Figure 3 shows the time evolution of the Chl-a Root Mean Square (RMS) estimation error (data/model misfit) for the forecast (i.e. just before the assimilation of the new observations) and the analysis as they result from three assimilation runs with different choices of the radius of influence: 250 km (EXP1), 150 km (EXP2) and 70 km (EXP3), respectively. RMS of the filter runs are compared to those obtained from a model free-run (i.e. model run without assimilation). The assimilation system behaves poorly with $d=70$ km. Although the choice of 250 km seems to be the best choice in terms of the analysis RMS error, the forecast system shows some weaknesses for such large value of d , as can be noticed from the forecast RMS misfit plotted in the lower panel of Fig. 3. This increase in the forecast RMS might be explained by some spurious large-range correlations in the EOFs correction directions that affected the quality of the filter's analysis, occasionally providing defective initial conditions to compute the forecast. These results suggest that a radius of influence should be chosen neither too large to filter out spurious large-range correlations, nor

too small in order to obtain smoother analysis states. Following these findings, a radius of influence $d=150$ km was selected in the sequel.

4.3 Model and observational errors

The observational and model error covariance matrices R and Q need to be specified in the filter's algorithm. These matrices are generally very poorly known. It is common to consider a time-invariant diagonal observation error covariance matrix \mathbf{R} . This means that the observational errors are assumed to be spatially uncorrelated, but this can be partly accounted for by overestimating the diagonal coefficients of R . Following the discussion for the observational errors in Sect. 3, the diagonal coefficients of R were set as a fraction of the variance of the SeaWiFS data at each observed location. The specification of Q is considerably more complex because very little information are available about the model error, and because of the significant number ($N \times N$) of parameters that need to be estimated. Following Pham et al. (1997), a simple compensation technique is used to replace the term $P_{L_k}^T Q_k P_{L_k}$ in Eq. (10) by a forgetting factor ρ which artificially amplifies the background error covariance matrix. This leads to a new update formula for the matrix U :

$$U_k^{-1} = \rho U_{k-1}^{-1} + (\mathbf{H}_k L_k)^T R_k^{-1} \mathbf{H}_k L_k \quad (12)$$

The value of ρ depends on the system under study, since it can be further used to account for other sources of errors in the filter, such as the underestimation of the error covariance matrices by low-rank matrices or linearization errors, by assigning more weights to recent observations. ρ was empirically set to 0.3 after several sensitivity assimilation experiments with different values of ρ ranging between 0 and 1. As an example, Fig. 4 plots the RMS analysis and forecast errors as obtained from the filter with three different choices of ρ (EXP4: 0.15, EXP2: 0.3, EXP5: 0.6) to show that, in our experiments, the lowest analysis and forecast RMS misfits were obtained using $\rho=0.3$.

4.4 The correction directions

A common strategy to determine the filter's correction directions is to use model statistics as an approximation of the true system statistics (Pham et al., 1997). Then by appropriate sampling of model state vectors one can obtain an approximation of the filter's covariance matrix through the dominant empirical orthogonal functions (EOFs). In the present study, the EOFs were generated as follows. After the 4-year climatological run (described in Sect. 2.3), another 2-year integration of the coupled system (physical & biogeochemical model) was carried out (described in Sect. 2.4) to generate a historical sequence of model states, sampled every 2 days during the last year of the integration. Since the state variables are of different nature, a multivariate EOF analysis was applied to the sampled set of 180 state vectors. In this analysis, model state variables were normalized by the inverse of

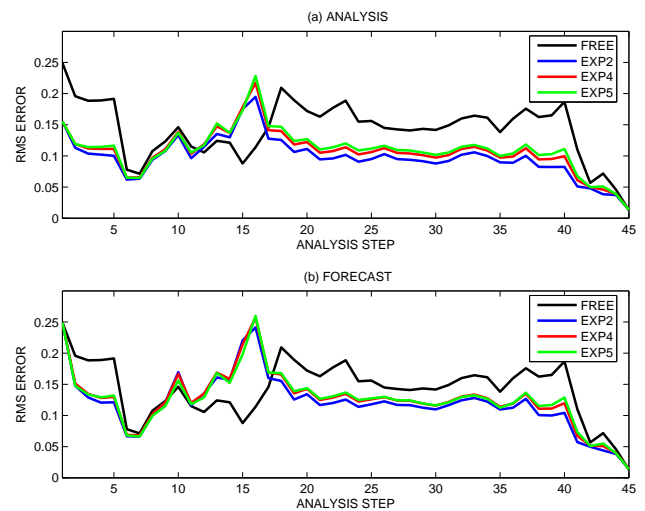


Fig. 4. RMS error for two experiments with different choices of the forgetting factor (EXP2: 0.3; EXP4: 0.15, EXP5: 0.6).

the square-root of their domain-averaged variances. For the assimilation experiment, 25 EOFs were retained explaining 95% of the sampled ecosystem variability.

5 Assimilation of satellite ocean color data for the period 1999

This section presents and discusses the results of the main assimilation experiment in which the 1999 SeaWiFS data were assimilated into the Eastern Mediterranean ecosystem model using the SEEK filter with a radius of influence $d=150$ km and a forgetting factor $\rho=0.3$. Morel (1998) suggested that the depth sensed by SeaWiFS depends upon the concentration of chlorophyll and the wave band, and for waters with chlorophyll ≤ 0.1 mg/m³, as in the case of the Levantine, this depth is about 30 m. In this study, the observation operator H_k integrates vertically the model chlorophyll profiles for the four phytoplankton groups over the first 30 m in the filter's equations. Hereafter, the overall behavior of the assimilation system as compared to the assimilated data and the MedAtlas climatology (MEDAR group, 2002) is first presented and analyzed. The impact of the assimilation of ocean colour data on the other ecological components of the model is then discussed. This allows us (i) to assess the relevance of the assimilation system to efficiently propagate surface observations into the deep ocean, and (ii) to study the ability of the BFM model to properly assimilate the information from the SeaWiFS data.

Figure 5 plots the time evolution of the RMS error between the assimilated chlorophyll data and the estimated chlorophyll concentrations as it results from the model free-run (without assimilation), and the filter run before (forecast) and after (analysis) the filter's correction. Overall the

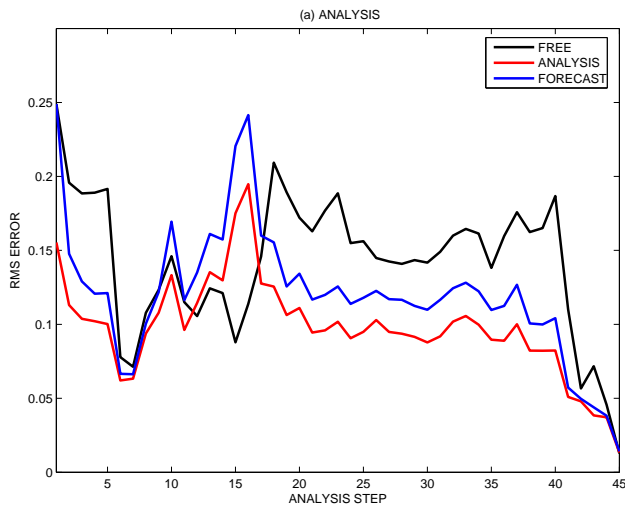


Fig. 5. Chlorophyll Root Mean Square difference (RMS) of the free run model data, the forecast and the analysis over the entire computation field.

filter's behavior is quite satisfactory, obviously improving the model/data consistency. The RMS error for both forecast and analysis is always smaller than the RMS error of the free-run, with only exception the period of spring bloom between the end of March and the end of April, when the filter, particularly the forecast state, was not able to follow the rapid changes in the ecosystem state. The filter correction step efficiently brings the model closer to the assimilated data after every filtering cycle as the analysis RMS error is always smaller than the forecast RMS error. The filter analysis state is also shown to respect the dynamics of the model, as the forecast RMS error remains stable after each analysis step, which allows the BFM model to smoothly assimilate the SeaWiFS data, suggesting that the assimilation system is properly working.

The poor behavior of the filter during the spring bloom period is a common difficulty for a Kalman filter technique, generally inefficient with highly intermittent and fastly varying processes, having probability distributions not well characterized by means and variances (Bennett, 1992; Hoteit et al., 2005). This particular period is also known to be characterized by increased variability in both nutrients and biological variables with significant gradients being developed over very short time periods (a few days only), shorter than the 8 days assimilation window. The estimation error might therefore quickly increase in the lack of adequate data to guide the system toward the true trajectory during this unstable period. The limited impact of the filter's correction step during this period might be also due to the misrepresentation of the bloom event in the EOFs based correction directions. Indeed, the dynamics of the assimilated variable are mainly controlled by the supply of phosphorus and nitrogen, while the grazing pressure becomes significant only

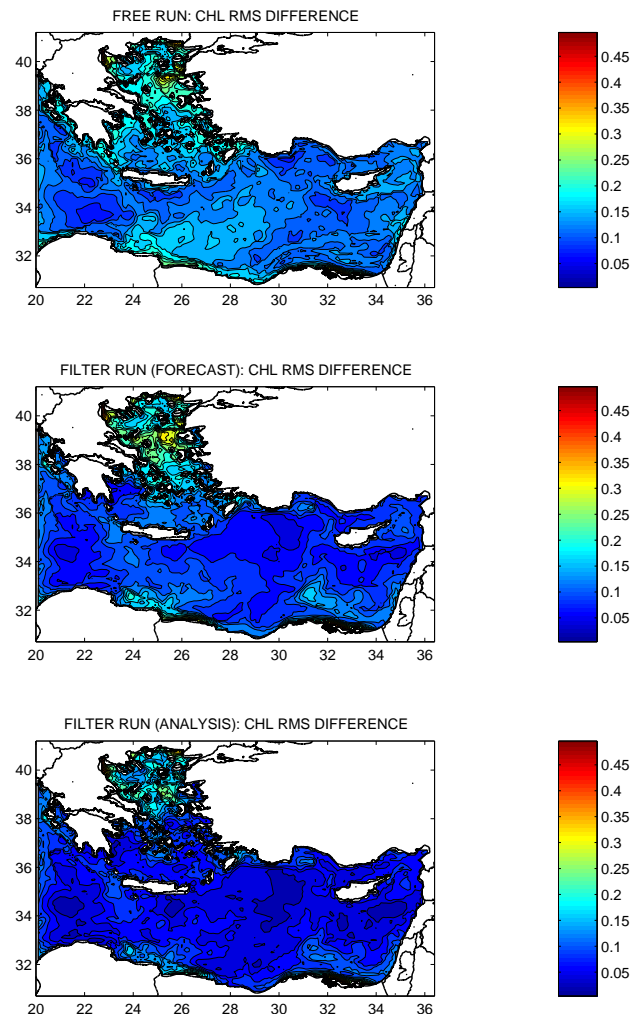


Fig. 6. Spatial distribution of the chlorophyll RMS differences over the entire assimilation period (Free run, Forecast, Analysis).

during the bloom period. Thus, the calculation of the dominant EOFs from a phosphate limited system with relatively small top down control on primary production for most of the year might be insufficient to efficiently approximate the covariance matrix of the filter during the bloom period. The use of seasonal sets of EOFs that can be associated with the major ecological events, and the evolution of the correction directions with the model dynamics are expected to improve the behavior of the filter during this period.

The spatial distribution for the chlorophyll RMS model-filter/data misfits averaged over the entire assimilation window is shown in Fig. 6. The top panel clearly shows that the model free-run/data misfits are the largest, with the worse model performance observed in the northern Aegean Sea and the area close to the northern coast of Africa. The North Aegean Sea is characterised by an important ecosystem variability during most of the year. This characteristic stems from the combination of shallow depths at the north and deep

basins at the central part of the Aegean in conjunction with the significant inputs from Dardanelles, Evros, Nestos, Aliakmomas and Axios rivers, which are not well represented in the model (lack of data). The lower panel of Fig. 6 suggests that the filter correction step is always able to efficiently improve the forecast, even over the North Aegean area; the analysis error was, over the whole domain, less than that of the free and the forecast runs. The fact that the filter forecast error significantly increases after every analysis step in the area of the North Aegean (middle panel) means that the model is not able to properly assimilate the information contained in the data. Additionally, the filter did not provide a good analysis to restart the model, although the analysis was close enough to the data in the area. A plausible explanation for the poor forecast is therefore a dynamically imbalanced analysis caused by the important model/data differences or, in other words, the model error.

As a demonstration of the good performance of the assimilation system, Fig. 7 plots surface chlorophyll concentrations from satellite observations (top panel), model free-run (middle panel) and analysis (lower panel) for the period 21–28 June. In order to compare satellite observations with model results, surface chlorophyll concentrations produced by the model were integrated over the first 30 m of depth and averaged over the same period. The model free run significantly overestimates the surface chlorophyll concentrations in the Levantine Sea, and particularly in the Aegean Sea. The filter successfully corrects the forecast driving the model closer to the satellite data over the whole model domain, although the improvement over the Aegean Sea is less significant. It is interesting to note that the filter efficiently reduces the strong variability in the Levantine Sea.

To further validate the behavior of the assimilation system, annual mean model chlorophyll concentrations from the free and the assimilated runs were compared with MedAtlas data (MEDAR group, 2002) across an East-West transect at latitude 34° N (Fig. 8). In the upper part of the euphotic zone the improvement in the assimilated run is evident with the misfit being significantly smaller compared to the values from the free run. An interesting outcome is the poorer behavior of the filter at the deeper parts around the DCM where the free run seems to be closer to the MedAtlas data. Although one has to take into consideration that the attempted comparison is between climatological field data and model results for 1999, a possible explanation is that in cases the ecological model is more productive compared to the field measurements, the assimilation of sea color data is constraining it by reducing primary production. As a result, functional groups feeding on the removed phytoplankton cannot sustain their acquired biomass and sink as detritus to deeper layers forming a deposition area just under the euphotic zone. However, to fully understand this process, a set of modelling experiments in a data rich area, are one of our immediate tasks.

A well known and reasonably studied hydrological feature over the modelling area is the Rhodes Gyre, offering a good

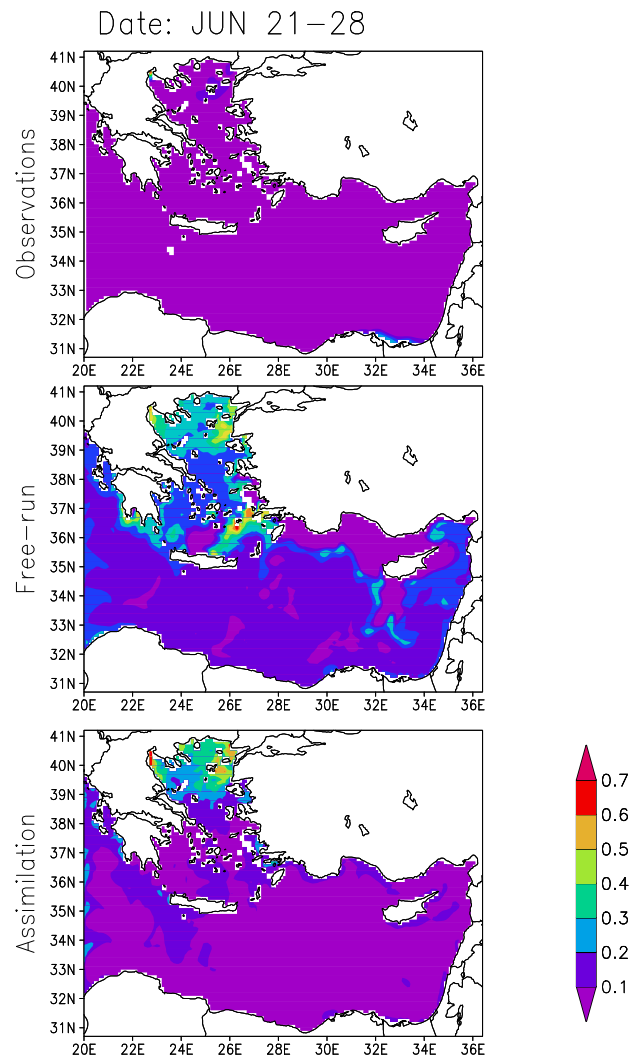


Fig. 7. Satellite surface chlorophyll observations (top panel), model free run (central panel) and analysis (lower panel) for the period 21–28 June.

opportunity to further assess the ecosystem functioning of the assimilation system in the deep layers. Figure 9 plots a North–South cross section of different ecological variables at longitude 28.5° E during the summer period as they result from the assimilation system. Knowledge about the biogeochemical processes in this area (Krom et al., 2003) can be used to validate the behavior of the assimilation system. The Rhodes Gyre is a cold core eddy that tends to change intensity, size and location with time. The main characteristic of the Gyre is the upwelling of nutrients in the center that results in increased phytoplankton biomass and primary productivity. The top panel shows the vertical distribution of the chlorophyll concentrations where the DCM is depicted at 60 m with a magnitude of $\sim 0.3 \text{ mg/m}^3$ in close agreement with Salihoglu et al. (1990). The middle panel shows the integrated chlorophyll, exhibiting a peak with a magnitude

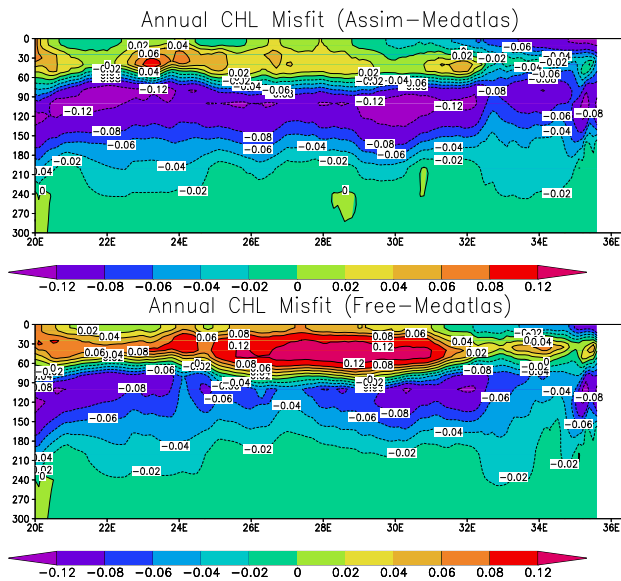


Fig. 8. Cross section of the upper 300 m of the water column along a section at latitude 34° N: Annual chlorophyll-a misfit of the 1999 assimilation run (top panel) and of the free run (bottom panel) from MedAtlas climatology.

of $\sim 28 \text{ mg/m}^2$ which is a slight underestimation compared to 39 mg/m^2 as estimated by Ediger and Yilmaz (1996). The phosphate concentrations plotted in the lower panel shows a formation of nutricline at 75 m. Phosphorous has been depleted in the upper layers exhibiting concentration values less than $0.02 \mu\text{g-at/l}$, while in cross section phosphate increases with depth to a maximum of $0.22 \mu\text{g-at/l}$. Model phosphate concentrations (free-run and assimilated) for this particular area when compared with insitu data from MedAtlas (Fig. 10), show a clear improvement in the assimilation system reducing the misfit at the top 300 m of the water column.

The performance of the assimilation system is also assessed in Fig. 11 which plots the time evolution of the depth integrated chl-a concentrations for the Rhodes gyre as it results from the model free-run and as estimated by the assimilation system. The model free-run without any assimilation provides quite good estimates of the integrated chl-a concentrations (35 mg/m^2) compared to the ones provided by the field measurements (39 mg/m^2). The assimilation of ocean colour data clearly reduces the surface chl-a concentrations pushing the model top layers towards a more oligotrophic condition according to the information extracted from the observations. However this drives the system dynamics away from the “truth” in the deep layers, not-visible by the satellite, subsequently reducing the depth integrated chl-a concentrations, being in disagreement with the observed values. For reasons already explained there are certain limitations in the assimilation of only surface ocean colour data which call for the need of constraining the model dynamics with subsurface ecological data as well.

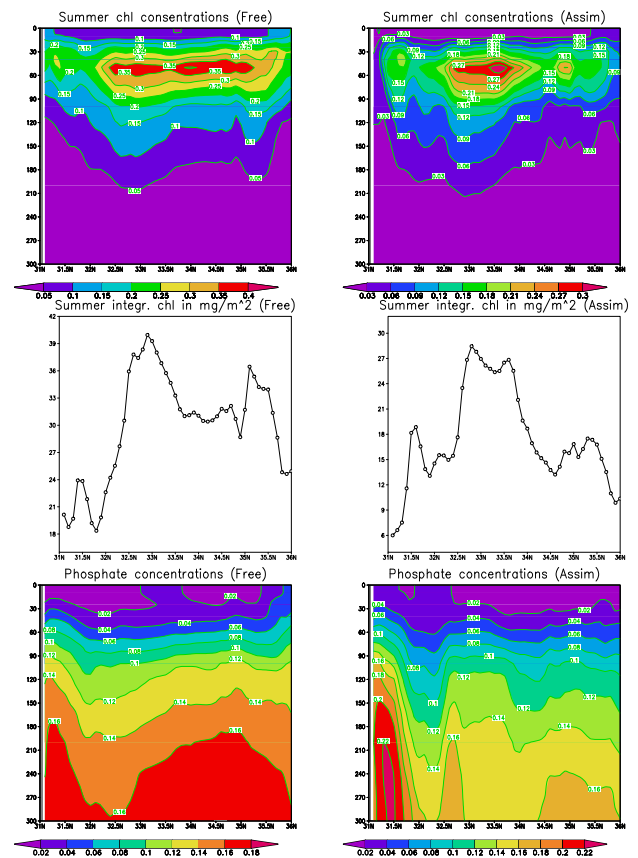


Fig. 9. Cross section at longitude 28.5° E for Free run (Left panels) and Assimilation run (Right panels). (Top) Summer distribution of chlorophyll concentrations (mg/m^3). (Middle) Integrated chlorophyll concentrations over 0–120 m averaged over the summer period (mg/m^2) (Lower) Summer vertical distribution of phosphorous concentrations (mmols/m^3).

Having stated the above limitation of the assimilation system in oligotrophic areas, which is mainly related to complex and highly variable ecological processes associated with cyclonic gyres and upwelling areas and the fact that the hydrodynamical part (which is an important driving mechanism for the ecology) needs to be constrained with assimilation of some key physical observations (satellite measured SLA for example), the overall performance of the assimilation system is in general quite positive as can be seen for example in the spatially integrated chl-a concentration during August (Fig. 12). The values resulting from the assimilation system are quite closer to the measured concentrations of Ignatiades et al. (2002); Siokou-Frangou et al. (2002) than those simulated by the model free-run. Additionally the decomposers in both free-run and assimilation run are within the range of measured values in the North, North-East and South Aegean as shown in Table 1. Moreover, it was found that the free run exhibits a more uniform spatial pattern, in contrast to the one estimated by the assimilation system which exhibits

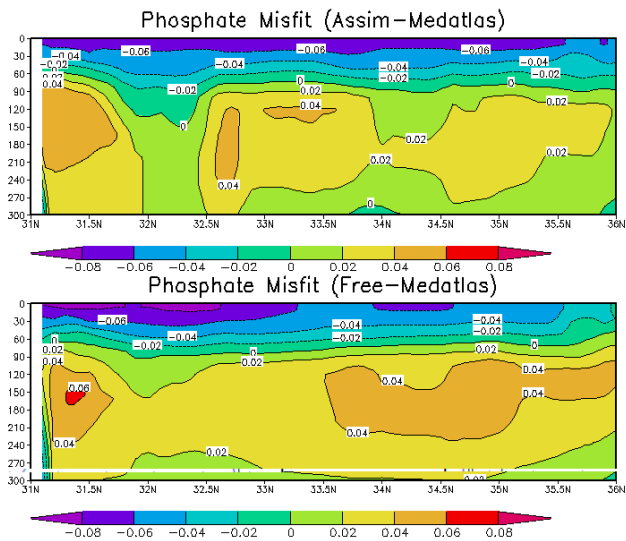


Fig. 10. Cross section of the upper 300 m of the water column at longitude 28.5° E. Top panel shows the summer phosphate misfit of the 1999 assimilation run from MedAtlas summer climatology while bottom panel shows the summer phosphate misfit of the free run from MEDATLAS summer climatology.

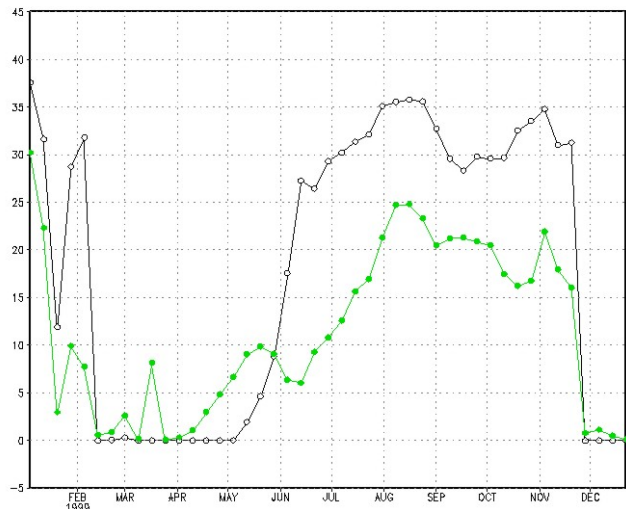


Fig. 11. Integrated chlorophyll (mg/m^2) from the model free-run (black line) and the assimilation run (green line).

significantly more spatial variability. Bacteria in oligotrophic systems play a very important role as they remineralise and consume nutrients (which are in short supply), sustaining an active microbial loop. As their growth depends on the available nutrient concentrations and the supply of DOC, which in turn are strongly coupled to the hydrodynamic fields, one would expect such strong spatial variability to be in agreement with the characteristics of the complex system of gyres and jets of the Eastern Mediterranean.

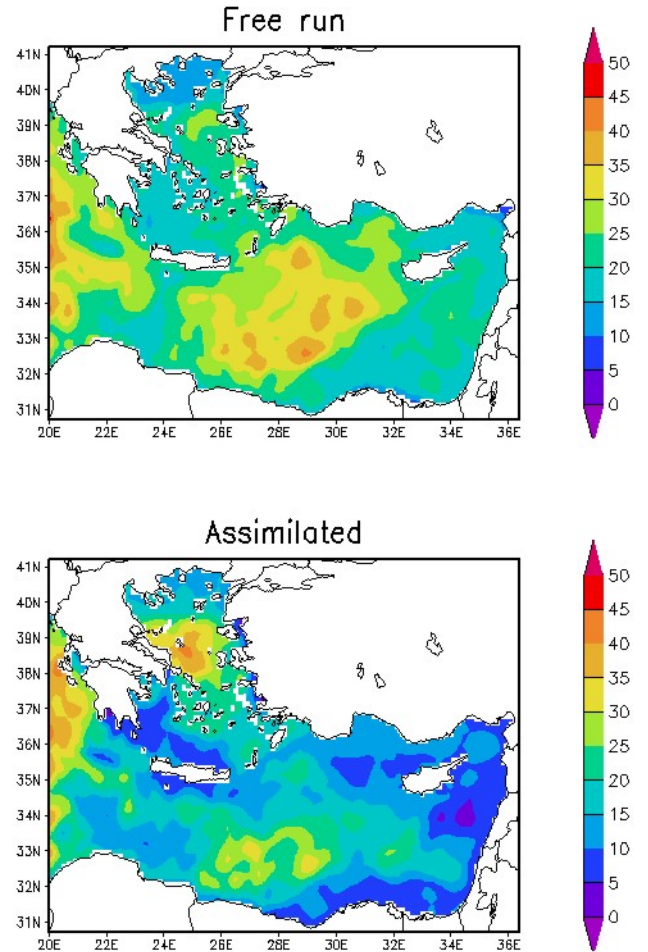


Fig. 12. Integrated chlorophyll (mg/m^2) over (0–150 m) from the model free-run and the assimilation run during August.

6 Conclusions

This study describes the implementation of an advanced assimilation system for the marine ecosystem of the Eastern Mediterranean based on a complex three-dimensional ecological model and a simplified Extended Kalman filter to assimilate SeaWiFS ocean colour data. The ecosystem model is composed of two coupled sub-models: the physical Princeton Ocean model (POM) and the Biogeochemical Flux Model (BFM). The filter is based on the Singular Evolutive Extended Kalman (SEEK) filter, in which the error statistics were parameterized by means of a suitable set of Empirical Orthogonal Functions (EOFs). A localization of the filter analysis step was implemented to filter out any spurious long-range correlations in the EOFs. After several sensitivity experiments which were performed in order to find appropriate values for some of the filter parameters, a hind-cast experiment was conducted for the year 1999 with the aim of demonstrating the effectiveness of this system and to validate its outputs. The results of this experiment clearly

Table 1. In situ and integrated bacterial biomass from the free-run and the assimilation run of 1999 during March and September at three areas in the Aegean Sea.

BACTERIA (mgC/m ²)	N. AEGEAN		S. AEGEAN		NE. AEGEAN	
	March	Sep	March	Sep	March	Sep
Siokou et al. (2002)	1406±327	1470±273	1423±43	1505±432	1403±142	1251±107
Free-run	1022±172	799±42	1197±59	1072±42	1298±126	808±27
Assimilation	1022±182	1245±123	1357±112	1072±134	1395±123	1181±79

demonstrate that the assimilation system operates in a satisfactory way: the assimilation system was capable of successfully fitting the observational data, and to propagate to some extent the surface observations to the deep layers. Furthermore, the assimilation improved the model behaviour and the impact of the satellite ocean colour data on all ecological components of the model was mainly positive. This is in agreement with the multivariate property of the assimilation scheme used in the present study.

However, some weaknesses were also observed, especially in the complex oligotrophic areas where adjustments were necessary in order to improve the model/data consistency. The filter nevertheless did not adjust several ecological components in the deep layers well enough to cover significant model errors not represented in the assimilation system. Even though the availability of a limited subsurface data set is not enough to solve this problem, the assimilation of these data would constrain the model variability in the deep layers and help preventing any deviation from reality. An important issue related to this problem is the assumption of “perfect physics” retained in this study. The improvement of the physical solution through the assimilation of physical data is expected to improve the behavior of the coupled model, which means less model errors in the ecological solution. Another poor performance of the assimilation system was observed during the spring bloom period during which the filter failed to follow the rapid change in the dynamics of the ecosystem. However, it is expected that the assimilation of more data, and the evolution of the filter correction directions would improve the behavior of the system during such periods. The quality of the data was also an issue in the present study. One solution could be to assimilate the colour data directly, before converting them to chlorophyll, by including a bio-optical algorithm to predict the colour from the model phytoplankton values, which hopefully will result in a reduction of the uncertainties in the data. The use of better estimate of the covariance matrix of the observational errors can be also beneficial for the assimilation system and should be considered in future studies.

Despite the use of a state-of-the-art coupled physical-biogeochemical marine ecosystem model constrained with the most synoptic ecological data sets using an advanced assimilation scheme, the overall results of this study are still at a preliminary stage, though giving all the improvements that

can be reported to the system. This study, however, clearly indicates that the development of an assimilation system capable of providing reliable estimates of the ecosystem state is achievable. This is an important finding since an alike system can be also implemented for any area of the global ocean. This conclusion is of particular interest for the marine ecosystem community and provides us with encouraging and promising results for future developments.

Acknowledgements. This work was supported by the Mediterranean Forecasting System – Towards Environmental Predictions Project (MFSTEP). Contract no. EVK3-CT-2002-00075 NASA Goddard Distributed Active Archive Centre is gratefully acknowledged for the supply of the SeaWiFS data.

Edited by: N. Pinardi

References

- Allen, J. I., Ekenes, M., and Evensen, G.: An Ensemble Kalman Filter with a complex marine ecosystem model: Hindcasting phytoplankton in the Cretan Sea, *Ann. Geophys.*, 21, 399–411, 2002, <http://www.ann-geophys.net/21/399/2002/>.
- Anderson, L. A., Robinson, A. R., and Lozano, C. J.: Physical and biological modelling in the Gulf Stream region: Data assimilation methodology, *Deep Sea Res.*, 47, 1787–1827, 2000.
- Baith, K., Lindsay, R., Fu, G., and McClain, C. R.: SeaDAS, A Data Analysis System for Ocean-Color Satellite Sensors, *EOS Trans. AGU*, 202 pp., 2001.
- Baretta, J. W., Ebenhoh, W., and Ruardij, P.: The European Regional Seas Ecosystem Model, a complex marine ecosystem model, *Neth. J. Sea Res.*, 33, 233–246, 1995.
- Bennett, A.: *Inverse methods in physical oceanography*, Cambridge Univ. Press, 346 pp., 1992.
- Blumberg, A. F. and Mellor, G. L.: A description of a three dimensional coastal ocean circulation model, in: *Three Dimensional Coastal Ocean Model*, edited by: Heaps, N. S., AGU, Washington, D.C., 1–16, 1987.
- Brasseur, P., Ballabrera-Poy, J., and Verron, J.: Assimilation of altimetric observations in a primitive equation model of the Gulf Stream using a singular evolutive extended Kalman filter, *J. Mar. Syst.*, 22(4), 269–294, 1999.
- Brasseur, P., Brankart, J.-M., Schoenauen, R., and Beckers, J.-M.: Seasonal Temperature and Salinity Fields in the Mediterranean Sea, *Climatological Analyses of an Historical Data Set*, *Deep Sea Res.*, 43, 159–192, 1996.

- Bricaud, A., Bosc, E., and Antoine, D.: Algal Biomass and Sea Surface Temperature in the Mediterranean Basin. Intercomparison of Data from Various Satellite Sensors, and Implications for Primary Production Estimates, *Rem. Sens. Environ.*, 81, 163–178, 2002.
- Carmillet, V., Brankart, J. M., Brasseur, P., et al.: A singular evolutive extended Kalman filter to assimilate ocean color data in a coupled physical-biochemical model of the North Atlantic ocean, *Ocean Model.*, 3(3–4), 167–192, 2001.
- Drakopoulos, P. G. and Lascaratos, A.: Modelling the Mediterranean Sea, climatological forcing, *J. Mar. Syst.*, 20, 157–173, 1997.
- Ediger, D. and Yilmaz, A.: Characteristics of deep chlorophyll maximum in the Northeastern Mediterranean with respect to environmental conditions, *J. Mar. Syst.*, 9, 291–303, 1996.
- Fasham, M. J. R. and Evans, G. T.: The use of optimization techniques to model marine ecosystem dynamics at the JGOFS station at 47°N 20°W., *Trans. R. Soc. Lond. Ser. B*, 348, 203–209, 1995.
- Feldman, G. C. and McClain, C. R.: Ocean Color Web, SeaWiFS Reprocessing, 4, 2004.
- Fennel, K., Losch, M., Schroter, J., and Wenzel, M.: Testing a marine ecosystem model: sensitivity analysis and parameter optimization, *J. Mar. Syst.*, 28, 45–63, 2001.
- Flather, R. A.: A tidal model of the northwest European continental shelf, *Mem. Soc. R. Sci. Liege, Ser. 6*, 10, 141–164, 1976.
- Friedrichs, M. A. M.: Assimilation of JGOFS EqPac and SeaWiFS data into a marine ecosystem model of the Central Equatorial Pacific Ocean, *Deep-Sea Res. II*, 49(1–3), 289–319, 2001.
- Gaspari, G. and Cohn, S. E.: Construction of correlation functions in two and three dimensions, *Q. J. Roy. Meteor. Soc.*, 125, 723–757, 1999.
- Ghil, M. and Malanotte-Rizzoli, P.: Data assimilation in meteorology and oceanography, *Adv. Geophys.*, 33, 141–266, 1991.
- Gunson, J., Oschlies, A., and Garçon, V.: 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, 1999.
- Harmon, R. and Challenor, P.: A Markov chain Monte Carlo method for estimation assimilation into models, *Ecol. Model.*, 101, 41–59, 1996.
- Horton, C., Clifford, M., Schmitz, J., and Kantha, L. H.: A real-time oceanographic nowcast/forecast system for the Mediterranean sea, *J. Geophys. Res.*, 102(C11), 25 123–25 156, 1997.
- Hoteit, I., Pham, D.-T., and Blum, J.: A simplified reduced order kalman filtering and application to altimetric data assimilation in tropical pacific, *J. Mar. Syst.*, 36, 101–127, 2002.
- Hoteit, I., Triantafyllou, G., Petihakis, G., and Allen, J. I.: A singular evolutive extended Kalman filter to assimilate real in situ data in a 1-D marine ecosystem model, *Ann. Geophys.*, 21, 389–397, 2003a.
- Hoteit, I. and Pham, D. T.: Evolution of the reduced state space and data assimilation schemes based on the Kalman filter, *J. Meteorol. Soc. Japan*, 81, 21–39, 2003b.
- Hoteit, I., Triantafyllou, G., and Petihakis, G.: Towards a data assimilation system for the Cretan Sea ecosystem using a simplified Kalman filter, *J. Mar. Syst.*, 45, 159–171, 2004.
- Hoteit, I., Triantafyllou, G., and Petihakis, G.: Efficient Data Assimilation into a Complex 3-D Physical-Biogeochemical Model Using a Semi-Evolutive Partially Local Kalman Filter, *Ann. Geophys.*, 23, 1–15, 2005, <http://www.ann-geophys.net/23/1/2005/>.
- Houtekamer, P. and Mitchell, H.: A sequential ensemble Kalman filter for atmospheric data assimilation, *Mon. Wea. Rev.*, 23, 1–15, 2001.
- Ignatiades, L., Psarra, S., Zervakis, V., et al.: Phytoplankton size-based dynamics in the Aegean Sea (Eastern Mediterranean), *J. Mar. Syst.*, 36, 11–28, 2002.
- IOCCG: Status and Plans for Satellite Ocean-Colour Missions: Considerations for Complementary Missions, Report No. 2, Dartmouth, Canada, 43 pp., 1999.
- Jaeger, L.: Monatskarten des Niederschlags für die Ganze Erde (Monthly precipitation maps for the whole earth), *Berichte des Deutscher Wetterdienstes*, 33 pp., 1976.
- Jazwinski, A. H.: Stochastic processes and filtering theory, Academic Press, New York, 1970.
- Kondo, J.: Air-sea bulk transfer coefficients in diabatic conditions, *Bound.-Lay. Meteorol.*, 9, 91–112, 1975.
- Korres, G. and Lascaratos, A.: An eddy resolving model for the Aegean and Levantine basins for the Mediterranean Forecasting System Pilot Project (MFSP): Implementation and climatological runs, *Ann. Geophys.*, 21, 205–220, 2003, <http://www.ann-geophys.net/21/205/2003/>.
- Korres, G. and Lascaratos, A.: An eddy resolving model of the Aegean and Levantine basins for the Mediterranean Forecasting System Pilot Project (MFSP): Implementation and climatological runs, *Ann. Geophys.*, 21, 205–220, 2003, <http://www.ann-geophys.net/21/205/2003/>.
- Krom, M., Groom, S., and Zohary, T.: The Eastern Mediterranean, in: *Biogeochemistry of Marine Systems*, edited by: Black, K. D. and Shimmield, G. B., Blackwell Publishing, Oxford, 91–126, 2003.
- Krom, M. D., Brenner, S., Kress, N., Neori, A., and Gordon, L. I.: Nutrient dynamics and new production in a warm-core eddy from the Eastern Mediterranean Sea, *Deep-Sea Res.*, 39(3/4), 467–480, 1992.
- Lascaratos, A. and Nittis, K.: A high-resolution three-dimensional study of intermediate water formation in the Levantine Sea, *J. Geophys. Res.*, 103(C9), 18 497–18 511, 1998.
- Lawson, L. M., Hofmann, E. E., and Spitz, Y. H.: Time series sampling and data assimilation in a very simple ecosystem model, *Deep-Sea Res. II*, 43(2–3), 625–651, 1996.
- Levitus, S.: Climatological Atlas of the World Ocean, NOAA/ERL GFDL, Professional Paper 13, Princeton, N. J., 173 pp., 1982.
- Matear, R. J.: Parameter optimization and analysis of ecosystem models using simulated annealing: A case study at station P, *J. Mar. Res.*, 53, 571–607, 1995.
- THE MEDAR GROUP – 2002 MEDAR/MEDATLAS: 1998–2001 Mediterranean and Black Sea database of temperature, salinity and bio-chemical parameters and climatological atlas (4 CDROMs), European Commission Marine Science and Technology Programme, and internet server www.ifremer.fr/sismer/program/medarIFREMER/TMSI/IDM/SISMER, Centre de Brest, 2002.
- Mellor, G. L.: User's guide for a three-dimensional primitive equation, numerical ocean model (July 1998 version), Princeton University, 1998.
- Mellor, G. L. and Yamada, T.: Development of a Turbulence

- Closure Model for Geophysical Fluid Problems, *Rev. Geophys. Space Phys.*, 20, 851–875, 1982.
- Morel, A.: Optical modelling of the upper ocean in relation to its biogenous matter content (case I waters), *J. Geophys. Res.*, 93, 10 749–10 768, 1998.
- Natvik, L. J. and Evensen, G.: Assimilation of ocean color data into a biochemical model of the North Atlantic. Part 1., Data assimilation experiments, *J. Mar. Syst.*, 40–41, 155–169, 2002.
- O'Reilly, J. E., Maritorena, S., O'Brien, M. C., et al. (Eds.): *SeaWiFS Postlaunch Calibration and Validation Analyses, Part 3, Vol. 11*, NASA Goddard Space Flight Centre, Greenbelt, Maryland, 2000.
- Patt, F. S., Barnes, R. A., Eplee Jr., R. E., et al.: Algorithm Updates for the Fourth SeaWiFS Data Reprocessing, in: *SeaWiFS Postlaunch Technical Report Series*, edited by: Hooker, S. B. and Firestone, E. R., NASA Technical Memorandum 32003–206892, NASA Goddard Space Flight Centre, Greenbelt, Maryland, 2003.
- Petihakis, G., Triantafyllou, G., Allen, J. I., Hoteit, I., and Dounas, C., Modelling the Spatial and Temporal Variability of the Cretan Sea Ecosystem, *J. Mar. Syst.*, 36, 173–196, 2002.
- Petihakis, G., Triantafyllou, G., Korres, G., Pollani, A., and Hoteit, I.: Eastern Mediterranean biogeochemical flux model: simulations of the pelagic ecosystem, *Ocean Sci. Discuss.*, 3, 1349–1398, 2006, <http://www.ocean-sci-discuss.net/3/1349/2006/>.
- Pham, D. T., Verron, J., and Roubaud, M. C.: Singular evolutive Kalman filter with EOF initialization for data assimilation in oceanography, *J. Mar. Syst.*, 16, 323–340, 1997.
- Platt, T., Sathyendranath, S., and Longhurst, A.: Remote Sensing of Primary Production in the Ocean, Promise and Fulfilment, *Phil. Trans. R. Soc., B(348)*, 191–202, 1995.
- Pinardi, N., Allen, I., and Demirov, E.: The Mediterranean ocean Forecasting System: first phase of implementation (1998–2001), *Ann. Geophys.*, 21, 3–20, 2003, <http://www.ann-geophys.net/21/3/2003/>.
- Prunet, P., Minster, J.-F., Ruiz-Pino, D., and Dadou, L.: Assimilation of surface data in a one-dimensional physical physical-biogeochemical model of the surface ocean (1). Method and preliminary results, *Global Biogeochem. Cy.*, 10(1), 111–138, 1996.
- Reed, R. K.: On estimating insolation over the ocean, *J. Phys. Oceanogr.*, 17, 854–871, 1977.
- Salihoğlu, I., Saydam, C., Baştürk, O., et al.: Transport and distribution of nutrients and chlorophyll-a by mesoscale eddies in the Northeastern Mediterranean, *Mar. Chem.*, 29, 375–390, 1990.
- Sancak, S., Beşiktepe, S. T., Yilmaz, A., Lee, M., and Frouin, R.: Evaluation of SeaWiFS chlorophyll-a in the Black and Mediterranean Seas, *Int. J. Remo. Sens.*, 26(10), 2045–2060, 2005.
- Siokou-Frangou, I., Bianchi, M., Christaki, U., et al.: Carbon flow in the planktonic food web along a gradient of oligotrophy in the Aegean Sea (Mediterranean Sea), *J. Mar. Syst.*, 33–34, 335–353, 2002.
- Smagorinsky, J.: General circulation experiments with the primitive equations, I, The basic experiment, *Mon. Wea. Rev.*, 91, 99–164, 1963.
- Spiitz, Y. H., Moisan, J. R., Abbot, M. R., and Richman, J. G.: Data assimilation and a pelagic ecosystem model: parameterization using time series observations, *J. Mar. Syst.*, 16, 51–68, 1998.
- Sommerfeld, A.: Partial differential equations, *Lecture Notes on Theoretical Physics, Vol. 6.*, Academic Press, San Diego, CA, 1949.
- Tassan, S.: Local Algorithms using SeaWiFS Data for the Retrieval of Phytoplankton Pigments, Suspended Sediment, and Yellow Substance in Coastal Waters, *Appl. Opt.*, 33, 2369–2378, 1994.
- Triantafyllou, G., Hoteit, I., Korres, G., and Petihakis, G.: Ecosystem modelling and data assimilation of physical-biogeochemical processes in shelf and regional areas of the Mediterranean Sea, *App. Num. Anal. Comp. Math*, 2, 262–280, 2005.
- Triantafyllou, G., Hoteit, I., and Petihakis, G. A.: Singular evolutive interpolated Kalman filter for efficient data assimilation in a 3-D complex physical-biogeochemical model of the Cretan Sea, *J. Mar. Syst.*, 40–41, 213–231, 2003.
- Vallino, J. J.: Improving marine ecosystem models, use of data assimilation and mesocosm experiments, *J. Mar. Res.*, 58, 117–164, 2000.
- Vichi, M., Baretta, J., Baretta-Bekker, J., et al.: European Regional Seas Ecosystem Model III., Review of the Biogeochemical Equations, 2004.
- Vichi, M., Pinardi, N., and Masina, S.: A generalized model of pelagic biogeochemistry for the global ocean ecosystem. Part I: theory, 64, 89–109, 2007a.
- Vichi, M., Ruardij, P., Baretta, J. W., Zavatarelli, M., Pinardi, N., Polimene, L., Butenschoten, M., Masina, S., and Patara, L.: The Biogeochemical Flux Model (bfm). Technical report, CMCC-INGV, <http://www.bo.ingv.it/bfm> 2007b.
- Zavatarelli, M. and Mellor, G. L.: A numerical study of the Mediterranean sea circulation, *J. Phys. Oceanogr.*, 25(6), 1384–1414, 1995.