Available online at www.sciencedirect.com
ScienceDirect
Journal of Marine Systems 69 (2008) 205–225
JOURNAL OF MARINE SYSTEMS
www.elsevier.com/locate/jmarsys

# Assimilation of SeaWiFS ocean chlorophyll data into a three-dimensional global ocean model

Watson W. Gregg

*NASA/Goddard Space Flight Center/Global Modeling and Assimilation Office, Code 610.1, Greenbelt, MD 20771, USA*

Received 25 May 2005; received in revised form 7 December 2005; accepted 27 February 2006
Available online 1 March 2007

**Abstract**

Sea-viewing Wide Field-of-view Sensor (SeaWiFS) chlorophyll data were assimilated with an established three-dimensional global ocean model. The assimilation improved estimates of chlorophyll relative to a free-run (no assimilation) model. Compared to SeaWiFS, annual bias of the assimilation model was 5.5%, with an uncertainty of 10.1%. The free-run model had a bias of 21.0% and an uncertainty of 65.3%. In situ data were compared to the assimilation model over a 6-year time period from 1998 through 2003, indicating a bias of 0.1%, and an uncertainty of 33.4% for daily coincident, co-located data. SeaWiFS bias was slightly higher at −1.3% and nearly identical uncertainty at 32.7%. The free-run bias and uncertainty at −1.4% and 61.8%, respectively, indicated how much the assimilation improved model results. Annual primary production estimates for the 1998–2003 period produced a nearly 50% improvement by the assimilation model over the free-run model as compared to a widely used algorithm using SeaWiFS chlorophyll data. These results suggest the potential of assimilation of satellite ocean chlorophyll data for improving model results.


## 1. Introduction

There is no substitute for observations in the effort to understand the oceans' biogeochemical cycles. However, observations alone cannot allow a full understanding. They are necessarily limited in time and space. This is true even for remote-sensing observations, which are generally limited to the surface and to once-daily temporal frequencies. Often the most important variables in the oceanic system, such as carbon and nutrients, are not directly observed. For example, ocean color sensors detect chlorophyll concentrations, which are related to organic carbon. But chlorophyll is not necessarily carbon, surface observations are not necessarily representative of the water column, and biomasses are not necessarily indicative of fluxes.

Numerical models can potentially connect the satellite observations, and provide meaningful information that observations cannot. Numerical models have no time or space limitations. They can translate the observations into the key geophysical variables that directly affect the oceans' biogeochemical cycles. And they can convert the remote sensing snapshots into fluxes. Finally, since numerical models are constructed with fundamental principles of ocean physics, biology, geochemistry, and radiative transfer, they can potentially provide understanding of the causes for distributions and changes seen in the remote sensing observations.

In practice, models are deficient in their representation of processes and interactions, and consequently their outputs stray from the observations. If they can be

*E-mail address:* watson.gregg@nasa.gov.

0924-7963/$ - see front matter © 2007 Elsevier B.V. All rights reserved.
doi:10.1016/j.jmarsys.2006.02.015

inextricably linked to the observations, then models can provide greatly enhanced understanding of biogeochemical cycling, by identifying the nature of the deficiencies and providing clues to improvement, as well as by nudging model variables toward realism.

Coupling models and data together through data assimilation is among the tightest and most intimate interrelationships as exists in computational Earth sciences research. In the data assimilation method, the model results are constrained by the observations. It is a field of research that has many aspects and challenges, but it can potentially be rewarding, since it maximizes the value of the data, but at the same time allows the fundamental processes in the model to act naturally.

Here an existing coupled general circulation, biogeochemical, and radiative model of the global oceans is used as a platform to assimilate Sea-viewing Wide Field-of-view Sensor (SeaWiFS) chlorophyll data products. The model can potentially enable us to investigate biogeochemical pathways, fluxes, and reservoirs, as well as the underlying processes, to produce a better understanding of the ocean system and how it operates. However, given the lack of maturity of data assimilation for ocean biology and biogeochemistry, the more limited objective in this effort is to improve estimates of surface chlorophyll and depth-integrated primary production, respectively. This effort represents an initial attempt to assimilate remote sensing ocean color data in a global model.

## 2. Background

Data assimilation can be classified into two groups, variational (inverse) methods and forward (sequential) methods (Anderson et al., 2000). Although a complete description of the specific methods is beyond the scope of this paper, a brief overview can help distinguish the classes. The variational, or inverse, methods construct a cost function, which is a measure of the difference between model output and observations over a specified time and space interval. Minimization of this cost function is the goal of the inverse methodology subject to constraint by the model (McGillicuddy et al., 1998), in which model parameters, boundary and initial conditions, and forcing functions can be involved. Minimization methods are quite varied in ocean ecosystem studies, including gradient steepest descent (Natvik et al., 2001), conjugate gradient method (Fasham et al., 1995, 1999), simulated annealing (Hurtt and Armstrong, 1996, 1999), and a micro-genetic algorithm (Schartau and Oschlies, 2003), among others (see Table 1). The most popular method in ocean biology has been the adjoint method. This is an efficient optimization method that reduces the large number of required iterations by first finding the gradient of the cost function, although iteration is still required. The inverse methods optimally modify model parameters using information from the data at the time of the model output, and reverse the flow of information back to the model at the time of its initial condition, hence the term inverse.

The focus of most work on variational marine biological applications has been on *parameter estimation*, i.e., to analyze and improve model performance by adjusting parameters to conform to observations. Variational methods are quite well-suited to parameter optimization and also have the advantage of mass conservation (Anderson et al., 2000). Improvement in model parameterizations has largely been achieved as measured by reducing model-data misfits as the direct result of application of variational methods. However, the parameter adjustments tend to be specific to the particular model formulation, and may not be applicable for other models or applications. The methods can also be computationally expensive for complex models and large areas.

The sequential class of data assimilation methods operates on the outputs of the model, rather than its internal parameters. A model field is created after an integration (time) step, which is then combined with the observations to produce an analysis or best state estimate. In the simplest case, the analyzed field is then used to re-initialize the model for the next execution step. The model parameters remain fixed in the method but the model outputs are driven toward the observations through constant confrontation with data. Examples of this type of assimilation are direct data insertion (Ishizaka, 1990), nudging (Armstrong et al., 1995), optimal interpolation (Popova et al., 2002), and various implementations of the Kalman filter. This class of methods is most appropriate for improved *state and flux estimation*, i.e., to produce more accurate derived variables, such as chlorophyll and primary production. Sequential methods have limited ability for parameter estimation beyond that which can be achieved though a reasonably comprehensive validation effort and they do not conserve mass or flux, but they do have a reasonable computational cost.

Assimilation of satellite ocean color data is a relatively new phenomenon in ocean sciences. There now exists a long uninterrupted time series of high quality data beginning with SeaWiFS (data set beginning in September 1997), and the Moderate Resolution Imaging Spectroradiometer (MODIS)-Aqua (data set beginning in July 2002). Despite the recent proliferation of remote sensing data, relatively few data assimilation

Table 1
Previous efforts in assimilation of in situ and simulated data for ocean biological/ecological studies

| Authors | Assimilation method | Model dimension | Location | Assimilation data source |
|---|---|---|---|---|
| *Variational methods* | | | | |
| Fasham et al. (1995) | Conjugate gradient method | 0D | Northwest Atlantic (BATS) | In situ |
| Fasham et al. (1999) | Conjugate gradient method | 0D | Northeast Atlantic | In situ |
| Natvik et al. (2001) | Gradient steepest descent/conjugate gradient method | 0D | Arbitrary | Simulated |
| Harmon and Challenor (1997) | Markov chain Monte Carlo | 0D | Arbitrary | Simulated |
| Hurtt and Armstrong (1996) | Simulated annealing | 0D | Northwest Atlantic (BATS) | In situ |
| Hurtt and Armstrong (1999) | Simulated annealing | 0D | North Atlantic (BATS and OWSI) | In situ |
| Matear (1995) | Simulated annealing | 0D | Northeast Pacific (Station P) | In situ |
| Weber et al. (2005) | Micro genetic algorithm | 0D | Northwest Atlantic (BATS) | In situ |
| Lawson et al. (1996) | Adjoint | 0D | Arbitrary | Simulated |
| Fennel et al. (2001) | Adjoint | 0D | Northwest Atlantic (BATS) | In situ |
| Schartau et al. (2001) | Adjoint | 0D | Northwest Atlantic (BATS) | In situ |
| Spitz et al. (1998) | Adjoint | 0D | Northwest Atlantic (BATS) | In situ |
| Spitz et al. (2001) | Adjoint | 0D | Northwest Atlantic (BATS) | In situ |
| Vallino (2000) | Adjoint | 0D | Arbitrary | In situ |
| Kuroda and Kishi (2004) | Adjoint | 0D | Northwest Pacific | In situ |
| Leredde et al. (2005) | Adjoint | 0D | Arbitrary | Simulated |
| Oschlies and Schartau (2005) | Micro genetic algorithm | 1D | North/Equatorial Atlantic | In situ |
| Schartau and Oschlies (2003) | Micro genetic algorithm | 1D | North Atlantic (3 stations) | In situ |
| Freidrichs (2001) | Adjoint | 1D | Equatorial Pacific | Simulated |
| Faugeras et al. (2003) | Adjoint | 1D | Mediterranean Sea | In situ |
| Faugeras et al. (2004) | Adjoint | 1D | Mediterranean Sea | In situ |
| Prunet et al. (1996) | Adjoint | 1D | Northeast Pacific (Station P) | In situ |
| Miller et al. (2000) | Green's function | 3D | California coast | In situ |
| Gunson et al. (1999) | Adjoint | 3D | North Atlantic | Simulated |
| Schlitzer (2002) | Adjoint | 3D | Southern Ocean | In situ |
| *Sequential methods* | | | | |
| Losa et al. (2003) | SIR Sequential Importance Resampling filter | 0D | Northwest Atlantic (BATS) | In situ |
| Eknes and Evensen (2002) | Ensemble Kalman filter | 1D | Arbitrary | Simulated |
| Allen et al. (2002) | Ensemble Kalman filter | 1D | Cretan Sea | In situ |
| Hoteit et al. (2003) | Singular evolutive extended Kalman filter | 1D | Cretan Sea | In situ |
| Ibrahim et al. (2004) | Singular evolutive extended Kalman filter | 1D | Cretan Sea | In situ |
| Magri et al. (2005) | Singular evolutive extended Kalman filter | 1D | Ligurian Sea | In situ |
| Carmillet et al. (2001) | Singular evolutive extended Kalman filter | 3D | North Atlantic | Simulated |
| Triantafyllou et al. (2003) | Singular evolutive interpolated Kalman filter | 3D | Cretan Sea | Simulated |
| Anderson et al. (2000) | Optimal interpolation | 3D | Gulf Stream | In situ |
| Besiktepe et al. (2003) | Optimal interpolation | 3D | Massachusetts Bay | In situ |
| Popova et al. (2002) | Optimal interpolation | 3D | Northeast Atlantic | In situ |

The table is divided into those utilizing variational methods and those using sequential methods.

studies have utilized it (Table 2). Hemmings et al. (2003, 2004) used variational methods in a 0-dimensional representation of the North Atlantic to refine model parameters. Results were mixed, as the assimilation improved chlorophyll concentrations in some areas, but seasonal variability was poorly represented. Losa et al. (2004) used variational methods with Coastal Zone Color Scanner (CZCS) data in a 0-dimensional simulation of the North Atlantic. Spatial patterns of chlorophyll were much improved using the optimized model as compared to a CZCS composite for 1979–1985. Some of the problems, especially in coastal areas and high latitudes, were attributed to data error. Freidrichs (2002) used SeaWiFS data in a 1-dimensional adjoint assimilation in the equatorial Pacific Ocean. Her emphasis was on model diagnosis and parameter

Table 2
Data assimilation efforts in ocean biological/ecological studies using satellite data

| Authors | Assimilation method class | Specific assimilation method | Model dimension | Location | Satellite assimilation data source |
|---|---|---|---|---|---|
| Hemmings et al. (2003) | Variational | Conjugate direction set method | 0D | North Atlantic (30 stations) | SeaWiFS |
| Hemmings et al. (2004) | Variational | Conjugate direction set method | 0D | North Atlantic (30 stations) | SeaWiFS |
| Losa et al. (2004) | Variational | Maximum data cost criterion | 0D | North Atlantic | CZCS |
| Freidrichs (2002) | Variational | Adjoint | 1D | Equatorial Pacific | SeaWiFS |
| Garcia-Gorriz et al. (2003) | Variational | Adjoint | 3D | Adriatic Sea | SeaWiFS |
| Ishizaka (1990) | Sequential | Insertion | 3D | Southeast US coast | CZCS |
| Armstrong et al. (1995) | Sequential | Nudging | 3D | Atlantic | CZCS |
| Natvik and Evensen (2003a,b) | Sequential | Ensemble Kalman filter | 3D | North Atlantic | SeaWiFS |
| Present effort | Sequential | Conditional relaxation analysis method | 3D | Global | SeaWiFS |

CZCS indicates the historical sensor Coastal Zone Color Scanner that flew from late 1978 to mid-1986.

estimation, and used the time series from September 1997 to April 1998. She found that assimilation of SeaWiFS 8-day composites could only provide realistic parameter sets when a portion of the time series was excluded. This was attributed to a change in ecosystem dynamics during this period and inadequacy of the 5-component model to represent it. She emphasized the importance and utility of inverse methods for model formulation. Garcia-Gorriz et al. (2003) also assimilated SeaWiFS data using the adjoint method. They studied the Adriatic Sea in January and June 1998. By varying recovered parameter sets, months, and portions of Adriatic Sea, they found misfits (biases in this case) between the assimilation and SeaWiFS ranging from 79% (more than the unassimilated model) to 22%. The best case improvement occurred in the Southern portion of the sea in June. Data error was found to be an important consideration for application of the adjoint method.

Although Schlitzer (2002) and Oschlies and Schartau (2005) used in situ data in their inverse assimilations, they used SeaWiFS data to evaluate their results. Schlitzer (2002) estimated export production (derived from primary production from SeaWiFS) in the Southern Ocean to be 2–5 times higher in the assimilation model than in SeaWiFS. This was attributed to the inability of ocean color sensors to detect subsurface chlorophyll and/or errors in the conversion from primary to export production. Oschlies and Schartau (2005) used an inverse method at three stations in the North Atlantic, and applied it to the entire basin. Primary production results from the assimilation model compared favorably with estimates using CZCS data, but agreement of spatial patterns and temporal variability of chlorophyll between the model and SeaWiFS 5-year mean chlorophyll (1997–2002) was lacking.

Utilizing a sequential data assimilation methodology, Ishizaka (1990) directly inserted CZCS chlorophyll into a 3-dimensional model of the southeast US coast. Immediate improvements in chlorophyll were observed this multi-variate assimilation but did not hold for more than 2 days. Cross-shelf fluxes of chlorophyll were reduced, because the unassimilated model appeared to overestimate chlorophyll, but the temporal variability was unaffected. Armstrong et al. (1995) used a nudging method with CZCS data in the North Atlantic. Initial results motivated a major change to the model configuration, which then showed good comparisons with CZCS chlorophyll. The ratio of assimilated chlorophyll to CZCS chlorophyll (a measure of the bias) ranged from 0.9 to 1.1 over most of the North Atlantic, when expressed as zonal means, with values $>1.5$ in the equatorial region.

A sequential assimilation approach for state estimation using modern ocean color data was employed by Natvik and Evensen (2003a,b). They used SeaWiFS data with an Ensemble Kalman Filter (EnKF) assimilation in a three-dimensional model of the North Atlantic. SeaWiFS errors were specified at 35% based on pre-launch goals (Hooker et al., 1992). Model errors were assumed to be related to atmospheric forcing data, specifically wind stress and surface temperature. Their results showed the EnKF was capable of providing an updated state consistent with SeaWiFS, including both phytoplankton and nitrate in a multi-variate analysis, over the period April and May 1998. However, quantitative expressions for the agreement were difficult to interpret.

In the present effort, a type of sequential assimilation model is used in a global context with multi-year SeaWiFS data with explicit observation error accounting. The emphasis is on state and flux estimation; specifically, improved estimates of surface chlorophyll and depth-integrated primary production, respectively. Quantitative measures and statistical analyses are utilized to evaluate the effects of data assimilation in a global context.

## 3. Methods

### *3.1. Coupled three-dimensional circulation/biogeochemical/radiative model of the global ocean*

A diagrammatic representation of a fully coupled general circulation/biogeochemical/radiative model, called the NASA Ocean Biogeochemical Model (NOBM), illustrates the complex interactions among the three major components, ocean general circulation, radiative, and biogeochemical processes models (Fig. 1). The Ocean General Circulation Model (OGCM) is a reduced gravity representation of circulation fields (Schopf and Loughe, 1995). It is global in scale, extending from near the South Pole to 72° N, in increments of 2/3° latitude and 1.25° longitude, comprising all regions where bottom depth >200 m. The model contains 14 vertical layers, in quasi-isopycnal coordinates, and is driven by wind stress, sea surface temperature (SST), and shortwave radiation (Table 3).

The biogeochemical processes model contains 4 phytoplankton groups, 4 nutrient groups, a single herbivore group, and 3 detrital pools (Fig. 2). The phytoplankton groups differ in maximum growth rates, sinking rates, nutrient requirements, and optical properties. The 4 nutrients are nitrate, regenerated ammonium, silica to regulate diatom growth, and iron. Three detrital pools provide for storage of organic material, sinking, and eventual remineralization back to usable nutrients. This results in 12 state variables in the fully coupled model. Atmospheric deposition of iron and sea ice are required as an external forcing fields (Table 3). The biogeochemical processes model is fully described in Appendix A.

Radiative transfer calculations provide the underwater irradiance fields necessary to drive growth of the phytoplankton groups, and interact with the heat budget. The Ocean-Atmosphere Radiative Model (OARM; Gregg, 2002a) contains a treatment of the spectral and directional properties of radiative transfer in the oceans, and explicitly accounts for clouds. The atmospheric radiative model is based on the Gregg and Carder (1990) spectral model, extended to the spectral regions 200 nm to 4 μm. It requires external monthly climatologies of cloud properties (cloud cover and liquid water path),

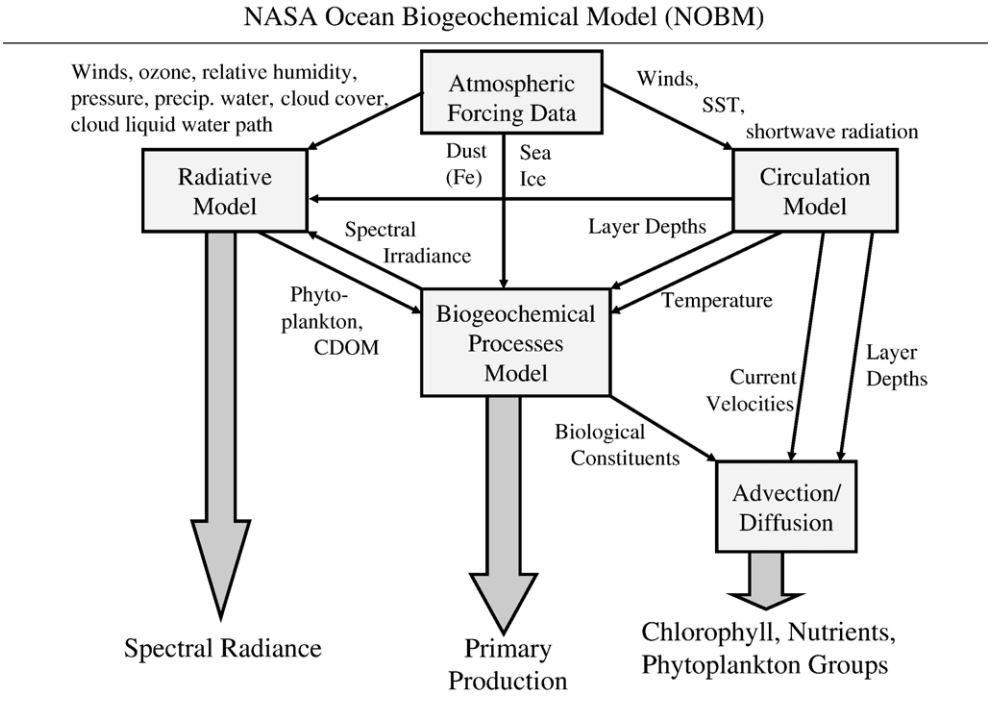


Fig. 1. Pathways and interactions among the components of the NASA Ocean Biogeochemical Model (NOBM).

Table 3
Forcing data sets required to force NOBM, their purpose, and sources of data

| Variable | Purpose | Source |
|---|---|---|
| *General circulation model* | | |
| Wind stress | Surface forcing | NCEP reanalysis |
| Sea surface temperature | Surface forcing | OISST |
| Shortwave radiation | Surface forcing | NCEP reanalysis |
| *Biogeochemical process model* | | |
| Aerosol composition | Surface input (iron) | GOCART |
| Sea ice | Surface forcing | OISST |
| *Radiative transfer model* | | |
| Wind speed | Surface reflectance/aerosols | NCEP reanalysis |
| Precipitable water | Water vapor absorption | NCEP reanalysis |
| Surface pressure | $O_2$ absorption/Rayleigh scattering | NCEP reanalysis |
| Relative humidity | Marine aerosols | NCEP reanalysis |
| Ozone | Gaseous absorption | TOMS |
| Cloud cover | Cloud distribution | ISCCP |
| Cloud liquid water path | Cloud attenuation properties | ISCCP |

NCEP is the National Center for Environmental Prediction, TOMS is the Total Ozone Mapping Spectrometer, ISCCP is the International Satellite Cloud Climatology Project, OISST is the Optimum Interpolated Sea Surface Temperature product, and GOCART is the Global Ozone Chemistry Aerosol Radiation and Transport model (Ginoux et al., 2001). Daily data are used for the 2001 assimilation analyses, and monthly data are used for the 1997–2003 analyses.

surface pressure, wind speeds, relative humidity, precipitable water, and ozone (Table 3). Aerosols are considered to be strictly of marine origin and are computed as in Gregg and Carder (1990).

Oceanic radiative properties are driven by water absorption and scattering, the optical properties of the phytoplankton groups, and chromophoric dissolved organic matter (CDOM). Three irradiance paths are enabled: a downwelling direct path, a downwelling diffuse (scattered) path, and an upwelling diffuse path. All oceanic radiative calculations include the spectral nature of the irradiance.

### 3.2. Data assimilation

The data assimilation methodology used here is the Conditional Relaxation Analysis Method (CRAM; Oort, 1983). The method is used for bias correction in Optimal Interpolation Sea Surface Temperature (OISST) data (Reynolds, 1988; Reynolds and Smith, 1994), and has been used successfully for ocean color in situ-satellite applications (Gregg and Conkright, 2001, 2002; Conkright and Gregg, 2003). CRAM uses data to provide an internal boundary condition, which here is the satellite ocean chlorophyll and solves for an analyzed chlorophyll field of model and data

$$\nabla^2 C_{\mathrm{T}}(\mathrm{ana}) = \nabla^2 C_{\mathrm{T}}(\mathrm{model}) \quad (1)$$

where $C_{\mathrm{T}}$(ana) is the final analyzed field of total chlorophyll and $C_{\mathrm{T}}$(model) is the model field (sum of the 4 phytoplankton groups). Insertion of satellite chlorophyll into the model field serves a bias-correction function. The matching of Laplacian's of the model chlorophyll and model/satellite chlorophyll extends the bias correction away from the satellite data points, while maintaining the higher order model variability. Because of the wide range of chlorophyll over the global oceans (>3 orders of magnitude), model and satellite data are logarithmically-transformed (base 10) before application of Eq. (1). The analyzed chlorophyll is transformed back to natural units for re-initialization of the next model integration. The analyzed chlorophyll is the sum of the 4 phytoplankton components, and is distributed among the functional groups to retain the previous model-derived relative abundances:

$$\Delta C_{\mathrm{T}} = C_{\mathrm{T}}(\mathrm{ana}) - C_{\mathrm{T}}(\mathrm{model}) \quad (2)$$

$$C_{\mathrm{T}}(\mathrm{model}) = \Sigma_i C_i \quad (3)$$

$$f_i = C_i(\mathrm{model}) / C_{\mathrm{T}}(\mathrm{model}) \quad (4)$$

$$C_i(\mathrm{ana}) = C_i(\mathrm{model}) + f_i \Delta C_{\mathrm{T}} \quad (5)$$

where $\Delta C_{\mathrm{T}}$ (Eq. (2)) is the difference between the analyzed total chlorophyll, $C_{\mathrm{T}}$(ana), using CRAM and the model, $C_{\mathrm{T}}$(model). This is called the analysis increment. $C_{\mathrm{T}}$(model) is the total chlorophyll (sum of all 4 phytoplankton components, Eq. (3)), $C_i$ is the $i$th phytoplankton chlorophyll component, and $f_i$ is the fraction of the $i$th phytoplankton component of the total chlorophyll.

Data assimilation is performed daily, to remove aliases associated with sampling by SeaWiFS (i.e., cloud cover, sun glint, inter-orbit gaps), that are incorporated in 8-day and monthly data products. Assimilation occurs at model midnight.

CRAM is computationally very fast, so much that there is nearly negligible additional processing time in its use. However, it is very strongly weighted toward the data. Thus data errors are an important problem in its

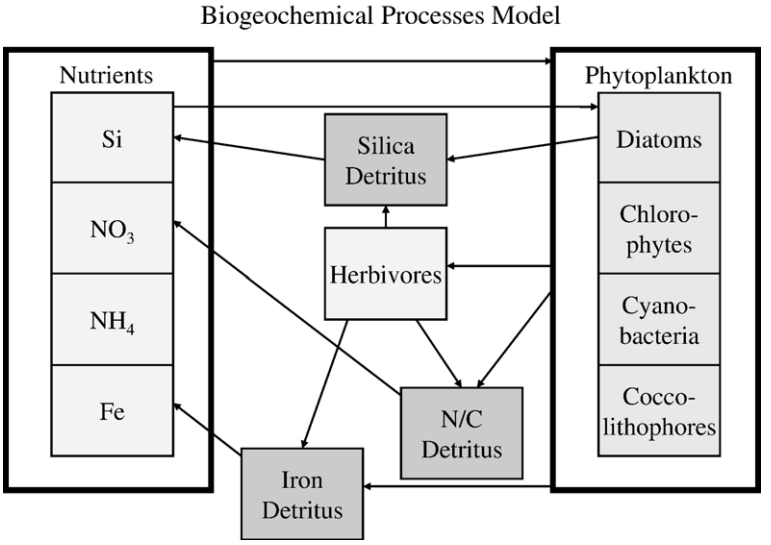


Fig. 2. Pathways and interactions among the components of the Biogeochemical processes model, comprising 4 phytoplankton groups, 4 nutrient groups, a single herbivore group and 3 detrital components.

application. For this reason, data errors must be minimized to the extent possible. In the present application, data error minimization efforts involve:

1) All daily SeaWiFS chlorophyll >2 times the monthly mean are excluded
2) SeaWiFS data are weighted 25% monthly mean to 75% daily data
3) SeaWiFS data occurring within a model grid point containing ice are excluded
4) Regional weighting of model and SeaWiFS chlorophyll is enforced (Fig. 3)

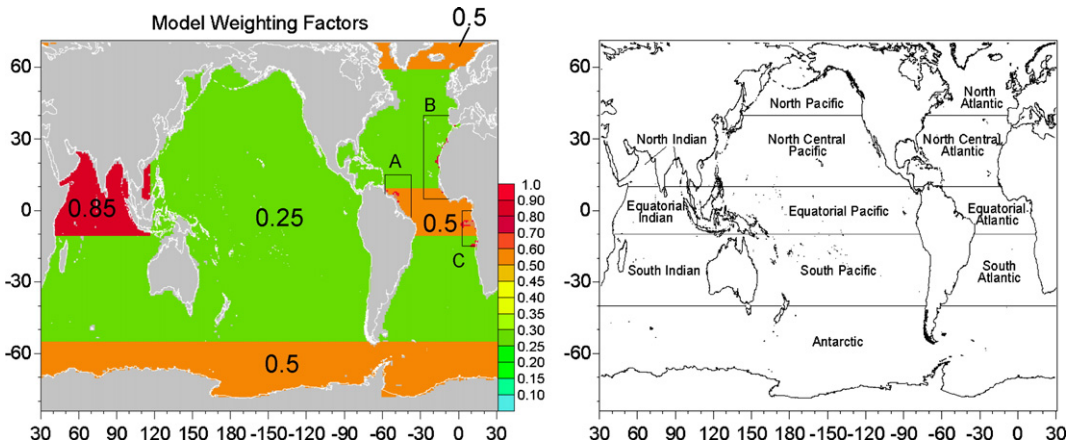


Fig. 3. Regional model weighting factors along with delineation of the major oceanographic basins. There are 3 special cases, with outlines delineated on the figure: (A) Amazon River outflow region, (B) Mauritanian offshore region, and (C) Congo River outflow region. In these sub-regions, higher weighting is used depending upon the satellite chlorophyll concentration, $C$(sat). In the Amazon and Congo outflow region (A and C)

if $C$(sat)>2 mg $m^{-3}$, then weight=1.0
else if $C$(sat)>1 mg $m^{-3}$, then weight=0.9
In the Mauritanian offshore region (B)

if $C$(sat)>1 mg $m^{-3}$, then weight=0.9
else if $C$(sat)>0.5 mg $m^{-3}$, then weight=0.75.

The fourth data error minimization is based on analyses by Gregg and Casey (2004) indicating regions where SeaWiFS tends to perform poorly compared to in situ data. It is also partly based on assimilation trial-and-error: where the assimilation produces negative values of any of the model variables or where unrealistic values occur, heavier weighting toward the model is enforced. Typically the two circumstances overlap. For example, excessively high chlorophyll concentrations are produced by the assimilation in the Congo and Amazon/Orinoco River discharges. These are regions dominated by CDOM, which produce erroneous chlorophyll values in satellite retrievals. These regions were shown to have a poor comparison with in situ data (Gregg and Casey, 2004). Similar problems occur with respect to regions where light-absorbing dust is prevalent, such as the tropical Atlantic and North and Equatorial Indian Oceans (Gregg and Casey, 2004). The regional model weighting factors used in the assimilation are shown in Fig. 3.

### 3.3. Data sets

Global chlorophyll data from SeaWiFS were obtained from the NASA Ocean Color Web site at daily and monthly, 9-km resolution. The data set version number was 5.1. The data were re-mapped to the model grid before assimilation and comparison.

Forcing data sets are shown in Table 3. Soil dust data sets were available only for the period January 2000 through July 2002. Climatologies were created to provide data when needed outside this period.

In situ chlorophyll data were obtained from the SeaWiFS Bio-Optical Archive and Storage System (SeaBASS; Werdell and Bailey, 2002) and the NOAA/National Oceanographic Data Center (NODC)/Ocean Climate Laboratory (OCL) archives (Conkright et al., 2002a). This was an updated version of the same combined data set used by Gregg and Casey (2004). The in situ data were re-mapped to the model grid on a daily basis.

### 3.4. Performance evaluation

SeaWiFS data assimilation is evaluated in the context of chlorophyll (state estimation) and primary production (flux estimation). For chlorophyll, monthly mean values of the assimilation model are compared with monthly mean SeaWiFS chlorophyll. Analyses involve the percent error

$$\text{Percent Error (PE)} = \frac{C(\text{assim}) - C(\text{sat})}{C(\text{sat})} \times 100 \qquad (6)$$

where $C_T$(assim) is the assimilation model total chlorophyll (which differs from $C_T$(ana) because it is the result of the assimilation process where $C_T$(ana) is used to re-initialize the model), and $C$(sat) is the satellite (SeaWiFS) chlorophyll.

Monthly percent errors are computed over the entire model domain where SeaWiFS and assimilation model chlorophyll values are co-located, and the bias is estimated using the median of the percent errors:

$$\text{Monthly Median Percent Error (MMPE)} = \text{median(PE)} \qquad (7)$$

The median was chosen for error analysis because of the lognormal distribution of chlorophyll data (Campbell, 1995). Logarithm transforms are common in such circumstances but percent errors are difficult to obtain and interpret. The median is nearly independent of the distribution of the data and is thus a useful, simple, and easy-to-interpret representation of the bias regardless of the distribution, and naturally incorporates the percent error.

The annual error is computed as the mean of the monthly median percent error over the 12 months of the year

$$\text{Annual Mean Percent Error (AMPE)} = \frac{\text{MMPE}}{12} \qquad (8)$$

Use of the mean of the monthly medians to obtain an annual value is reasonable because the MMPEs are normally distributed. This is confirmed by observation that the median of the MMPEs produces nearly the same value.

The uncertainty, or dispersion of the data, is represented by the Annual Semi-Interquartile Range (ASIQR)

$$\text{Monthly Semi} - \text{Interquartile Range (MSIQR)} = \text{IQR(PE)}^{*}0.5 \qquad (9)$$

$$\text{Annual Semi} - \text{Interquartile Range (ASIQR)} = \frac{\text{MSIQR}}{12} \qquad (10)$$

The interquartile range (IQR) encompasses all data between the 25th percentile and the 75th percentile of the data. One-half this value, the Semi-Interquartile Range (SIQR), is analogous to the standard deviation for normally-distributed data, in that the median ± SIQR contains 50% of the data. The difference is that the mean ± standard deviation encompasses 66% of the data.

Assimilation model annual errors are also evaluated by comparison with those from the free-run (control, i.e., no assimilation) model. The assimilation frequency is also adjusted, by assimilating every 2 days, 3 days, etc. instead of daily assimilation, to observe error growth.

Additionally, SeaWiFS, free-run model, and assimilation model chlorophyll are compared against the large data base of in situ chlorophyll data from NASA/SeaBASS and NOAA/NODC. Analyses involve use of the bias and uncertainty defined similarly to the monthly analysis defined in Eqs. (6)–(10) except compiling all daily data over the 6-year time span into a single representation of error, where in situ data and satellite/model data are coincident and co-located.

$$\text{Percent Error}(\text{PE}_{\text{is}}) = \frac{C(\text{assim}) - C(\text{is})}{C(\text{is})} \times 100 \qquad (11)$$

$$\text{Median Percent Error (MPE)} = \text{median}(\text{PE}_{\text{is}}) \qquad (12)$$

$$\text{Semi} - \text{Interquartile Range (SIQR)} = \text{IQR}(\text{PE}_{\text{is}})^{*}0.5 \qquad (13)$$

where $C$(is) indicates chlorophyll concentration of in situ data and $\text{PE}_{\text{is}}$ is the percent error of the assimilated chlorophyll relative to the in situ chlorophyll.

Primary production provides a means to evaluate the ability of the assimilation model to improve flux estimates. Primary production is computed in the model as a function of growth rate multiplied by the carbon:chlorophyll ratio:

$$\text{PP} = \int \Sigma \mu_i C_i \; \Phi \; dz \qquad (14)$$

where $\mu_i$ is the realized new growth rate of phytoplankton component $i$, $C_i$ is the chlorophyll concentration of component $i$, $\Phi$ is the carbon:chlorophyll ratio, and the product is integrated over depth. Assimilation of chlorophyll affects the total chlorophyll but not the relative abundances of the phytoplankton groups, $\mu$, or $\Phi$ directly. All three can be affected by the assimilation of chlorophyll indirectly, however, by changing the irradiance in the water column and the horizontal and vertical gradients of phytoplankton and nutrients. Free-run model-computed primary production is compared with model-computed primary production derived from assimilated chlorophyll and finally against primary production derived directly from satellite chlorophyll data using the Vertically Generalized Production Model (VGPM; Behrenfeld and Falkowski, 1997). The VGPM requires chlorophyll, SST, and photosynthetically available radiation (PAR) as inputs. Chlorophyll is taken from SeaWiFS, SST is the same source as used for model forcing (Table 3)), and PAR is derived from the atmospheric component of OARM, with wavelength region 350–700 nm selected and converted to quanta.

## 4. Results and discussion

### 4.1. NOBM

Minor changes in NOBM (see Appendix A) necessitated re-evaluation to ensure its performance did not degrade from Gregg et al. (2003). In the 50th year of model execution using climatological monthly forcing, basin-scale seasonal chlorophyll variability from the model was statistically positively correlated ($P<0.05$) with those determined from SeaWiFS monthly climatological chlorophyll in each of the 12 major oceanographic basins of the world (see Fig. 3), except the Equatorial Pacific, which exhibited very little seasonal variability. Global annual chlorophyll was 18.2% lower than SeaWiFS. Annual mean log-transformed dissolved iron concentrations in the model surface layer were positively correlated with observations ($P<0.05$) over the 10 (out of 12) major oceanographic basins where data were available (1951 in situ data records derived from the general literature, see Gregg et al., 2003 for details). The South Indian and South Atlantic were the basins where dissolved iron data were lacking.

Overall patterns of phytoplankton functional group distributions exhibited broad qualitative agreement with in situ data (359 surface layer observations, see Gregg et al., 2003 for details). Diatoms, cyanobacteria, and coccolithophores each exhibited statistically significant correlation with the in situ data across basins. Chlorophytes did not. Chlorophytes are a transitional group in the model, and they represent a wide range of phytoplankton, such as flagellates, *Phaeocystis* spp., etc. This expectation is probably unrealistic, and probably accounts for the lack of statistical significance in their relative abundances.

### 4.2. SeaWiFS assimilation 2001

The year 2001 was used to evaluate the effectiveness of chlorophyll assimilation. Here the free-run and assimilation models were forced with daily data. These were from the same sources as in Table 3, except at daily time-varying frequencies. Daily wind stresses were weighted 80:20 percent monthly:daily to minimize transient high wind events.

Basin-scale seasonal variability for both the free-run and assimilation models were statistically positively correlated with SeaWiFS in all 12 major basins, but the correlation coefficients for the assimilation model were much higher (Table 4). These results suggest the lack of significance for the Equatorial Pacific in the climatological model was due to use of climatological forcing.

Table 4
Mean annual basin difference from SeaWiFS for the free-run and the assimilation model, and the correlation coefficients ($r$) for the correlation with SeaWiFS seasonal variability for 2001

| Basin | Free-run model | | Assimilation model | |
|---|---|---|---|---|
| | Difference | $r$ | Difference | $r$ |
| North Atlantic | −25.7% | 0.891* | −16.8% | 0.998* |
| North Pacific | −32.5% | 0.724* | −20.1% | 0.991* |
| North Central Atlantic | −32.5% | 0.816* | −16.8% | 0.819* |
| North Central Pacific | 36.1% | 0.959* | 2.1% | 0.992* |
| North Indian | −64.4% | 0.779* | −50.1% | 0.871* |
| Equatorial Atlantic | −44.0% | 0.704* | −26.2% | 0.903* |
| Equatorial Pacific | −4.3% | 0.645* | −4.5% | 0.963* |
| Equatorial Indian | −22.2% | 0.941* | −13.3% | 0.974* |
| South Atlantic | 20.3% | 0.865* | −3.6% | 0.991* |
| South Pacific | 70.6% | 0.794* | 3.4% | 0.990* |
| South Indian | 49.8% | 0.725* | 6.0% | 0.999* |
| Antarctic | 24.2% | 0.903* | −10.7% | 0.971* |

An asterisk indicates the correlation is significant at $P<0.05$.

Daily assimilated satellite chlorophyll from SeaWiFS for April 1 compared favorably with monthly mean SeaWiFS data (Fig. 4). Although there was broad agreement between the free-run model and SeaWiFS monthly, the improvement using assimilation was clear. SeaWiFS chlorophyll for the same day as the free run and assimilated models is also shown, but because of cloud obscuration, sun glint, sensor tilt change, and inter-orbit gaps, it is difficult to evaluate the comparison. This illustrates the additional usefulness of assimilation, in providing complete daily coverage.

A more quantitative description of the effectiveness of assimilation is provided using monthly means of the assimilation model and SeaWiFS, and taking the difference (Fig. 5). For March 2001, the overall similarity of the assimilation and SeaWiFS was evident, and largely supported by the difference field. The largest differences occurred in the Arabian Sea, the Congo mouth, and the Mauritanian coast. All of these were by design in the assimilation model, with model weighting factors largest in these areas of low confidence in SeaWiFS. In all cases the differences were underestimates by the assimilation model, which was desired. Other smaller differences occurred in the northern extremities of the North Atlantic,

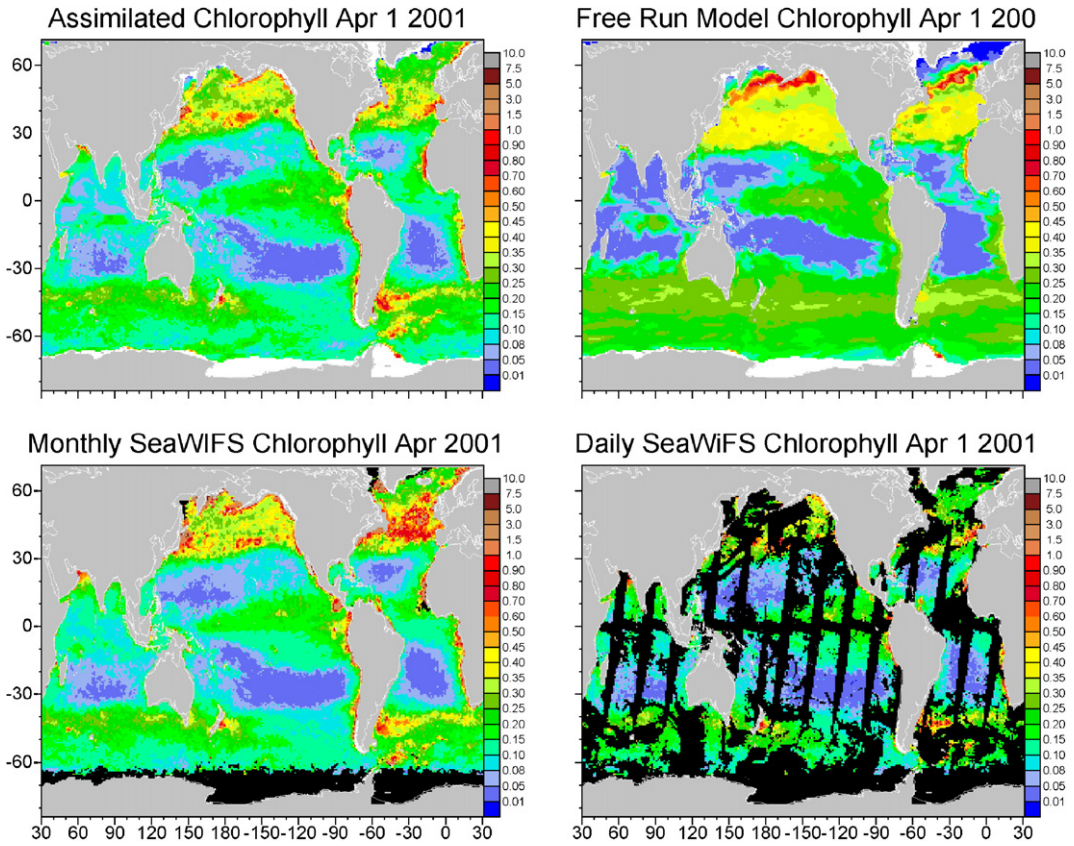


Fig. 4. Comparison of chlorophyll (mg $m^{-3}$) from the assimilation model, the free-run model, and SeaWiFS. The assimilation and free-run chlorophyll distributions represent simulations for April 1, 2001. SeaWiFS data for the same day are shown for comparison, along with the monthly mean. Grey indicates land and coast, black indicates missing data, and white indicates sea ice.

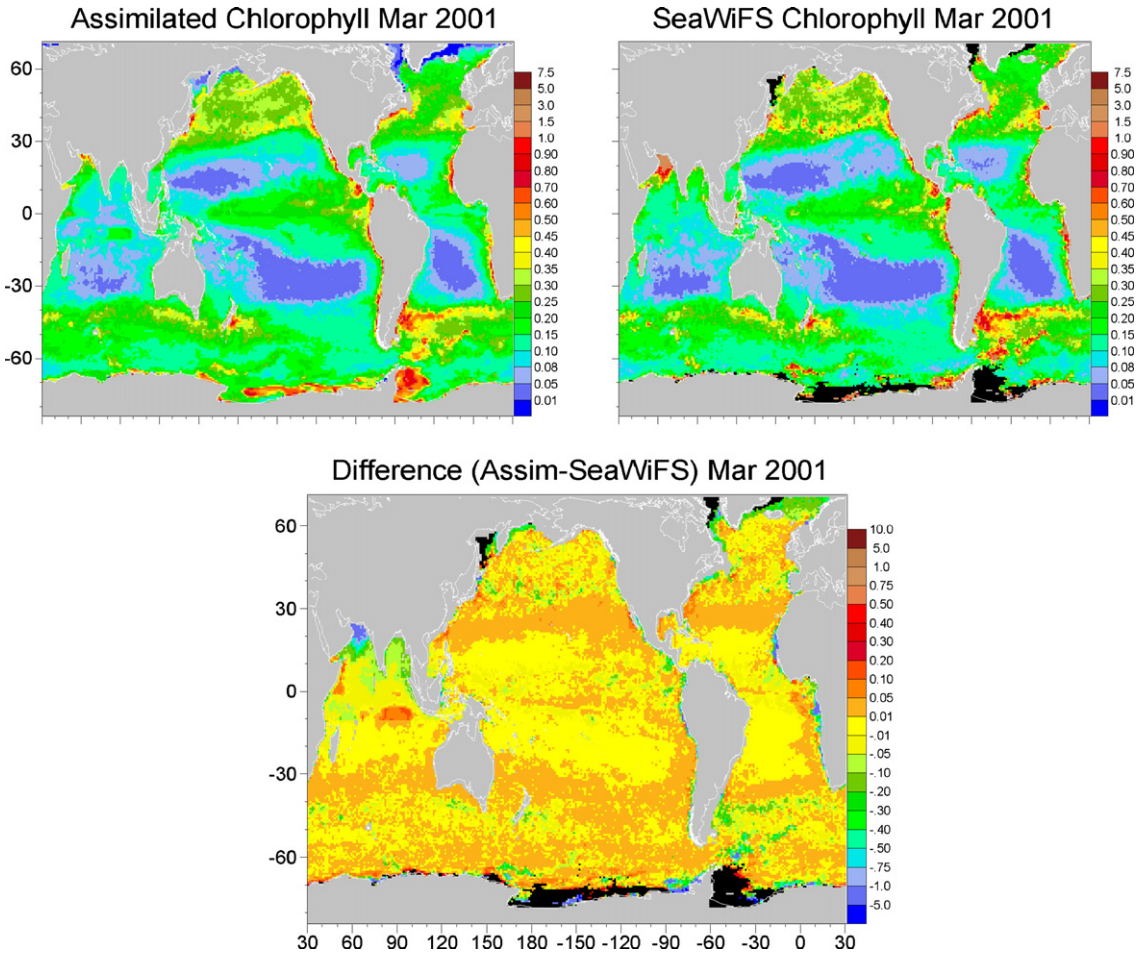


Fig. 5. Assimilation model chlorophyll (mg $m^{-3}$), SeaWiFS mean chlorophyll, and the difference (Assimilation-SeaWiFS) for March 2001.

the Pacific in the extreme western edge of the Bering Sea, and in the Atlantic sector of the Southern Ocean. Again the differences were underestimates by the assimilation. Overestimates by the assimilation were generally small (0.01–0.05 mg $m^{-3}$ chl). A couple of notable exceptions were offshore of the Somalian coast, and the east-central Indian Ocean, where overestimates by the model of 0.05–0.1 mg $m^{-3}$ occurred.

Similar results occurred for September 2001 (Fig. 6). Again the overall agreement between the assimilation model and SeaWiFS was good, with disparities in similar regions, specifically the Congo and Orinoco River outflows, the Arabian Sea, and the upper northern latitudes. There was a band in the Equatorial Atlantic where the assimilation model overestimated SeaWiFS, that did not appear related to the Congo River.

The growth of error as a function of assimilation frequency was tracked using the annual bias and uncertainty (Fig. 7). Using daily assimilation, the annual bias was 5.5% relative to SeaWiFS, which was a very large improvement over the error for the free-run model at 21.0%. The uncertainty improved from 65.3% in the free-run model to 10.1% in the assimilation model. The error grew as the assimilation frequency decreased. The uncertainty was still <30% if the assimilation occurred every 5 days. The bias remained <15% for up to a 6-day assimilation frequency. At the other extreme, very low assimilation frequencies, the annual bias and uncertainty approached the free-run model. The lowest assimilation frequency was once per year (every 183 days) for which the error is indistinguishable from the free-run model.

### 4.3. SeaWiFS assimilation 1997–2003

A long-term run of the free-run model and the assimilation model for 1997–2003 using monthly

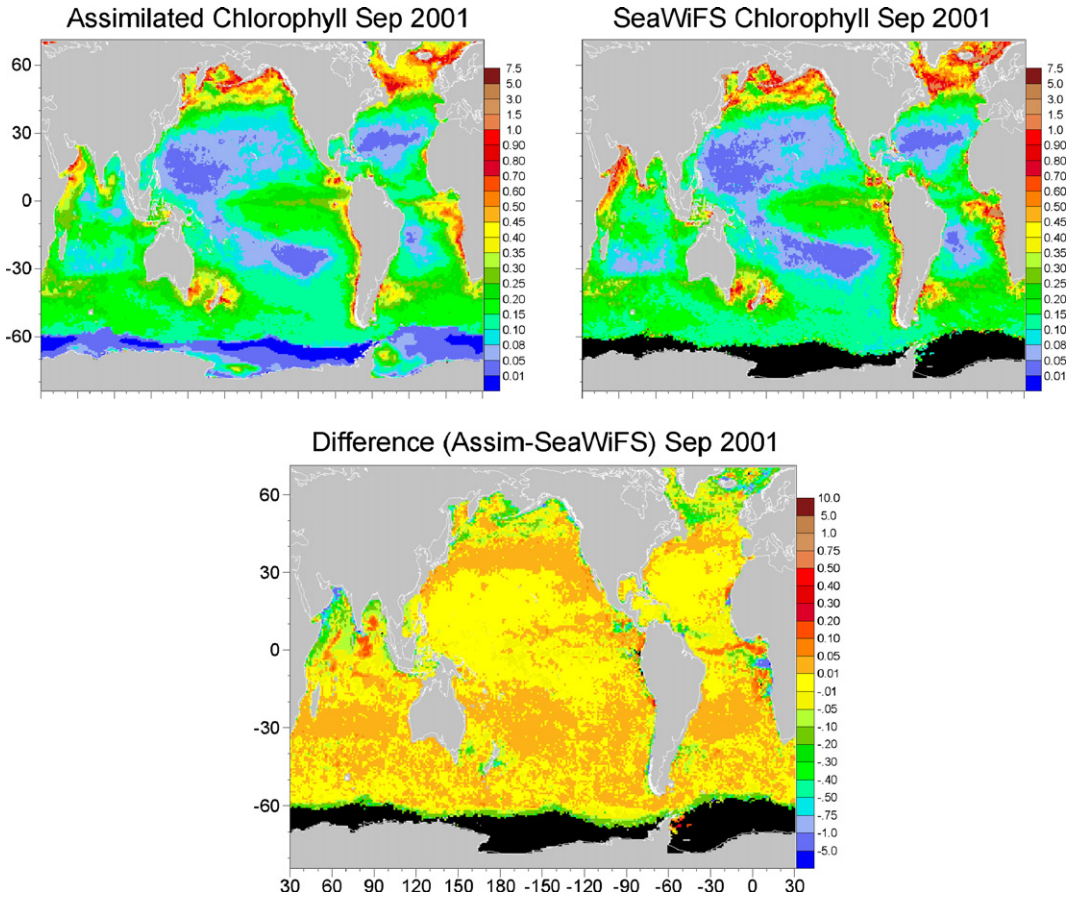


Fig. 6. Assimilation model chlorophyll (mg $m^{-3}$), SeaWiFS mean chlorophyll, and the difference (Assimilation-SeaWiFS) for September 2001.

forcing illustrates the improvement of assimilation in the major oceanographic basins (Fig. 8). The free-run model produced seasonal variability in good agreement with SeaWiFS basin mean chlorophyll, and also good correspondence with low biases in many of the basins, such as the North Central Pacific, North

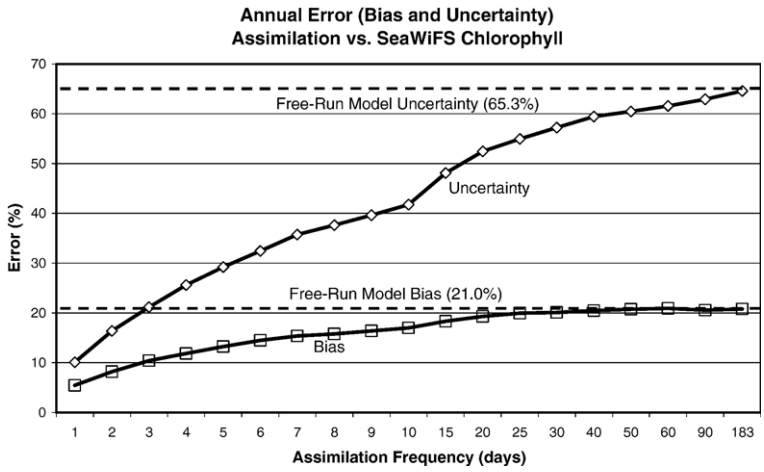


Fig. 7. Annual bias and uncertainty for assimilation as a function of assimilation frequency (1 indicates assimilation performed every day, 2 every 2 days, 3 every 3 days, etc.). The annual bias and uncertainty for the free-run model are shown.

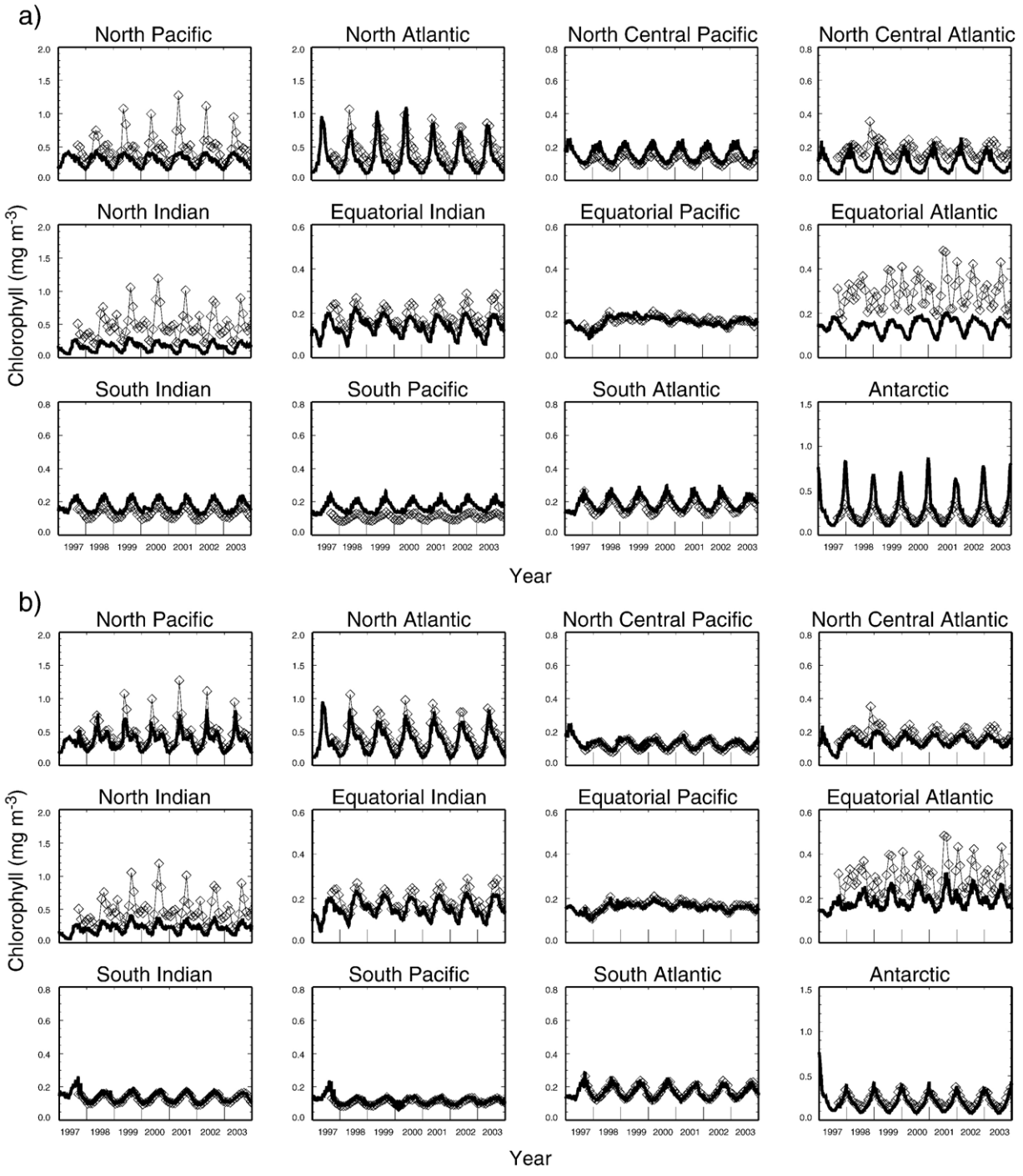


Fig. 8. (a) SeaWiFS monthly mean chlorophyll (diamonds) and daily chlorophyll from the free-run model (solid line) for 1997–2003. (b) SeaWiFS monthly mean chlorophyll (diamonds) and daily chlorophyll from the assimilation model (solid line) for 1997–2003. Assimilation of SeaWiFS chlorophyll did not begin until September 1997 (beginning of SeaWiFS data collection).

Atlantic, Equatorial Pacific, South Indian, and South Atlantic (Fig. 8a). There were several basins where a substantial bias was apparent in the free-run model. This was particularly true in the North Indian and Equatorial Atlantic, where a large underestimate by the model occurred, but also in the spring bloom peaks in the North Pacific and Antarctic basins, with underestimates and overestimates buy the model, respectively.

Table 5
Statistics for the comparison of SeaBASS/NODC chlorophyll data for the period 1998–2003 with coincident, co-located SeaWiFS, free-run model and assimilation model chlorophyll

| | Bias | Uncertainty | *N* |
|---|---|---|---|
| SeaWiFS | −1.3% | 32.7% | 2086 |
| Free-run model | −1.4% | 61.8% | 4465 |
| Assimilation model | 0.1% | 33.4% | 4465 |

*N* indicates the number of points where in situ and satellite/model points were coincident and co-located.

The assimilation model kept the seasonal variability agreement with SeaWiFS that the free-run model demonstrated, but additionally reduced the basin mean biases (Fig. 8b). The Antarctic and North Central Atlantic, which in the free-run model exhibited substantial biases, were now in nearly complete agreement. The large departure in the North Central Atlantic in autumn 1998 corresponded to a massive dust plume arising from northwestern Africa that was apparently undetected by the SeaWiFS processing algorithms (Gregg, 2002b). It was excluded by the assimilation here. The North Indian and Equatorial Atlantic showed improvement in the assimilation model, but still underestimated SeaWiFS. Model weighting factors were very high in these basins because of low confidence in SeaWiFS data (Gregg and Casey, 2004).

The assimilation model continued to underestimate the North Atlantic and North Pacific spring bloom maximum (Fig. 8b). In the North Atlantic, the underestimation appeared to be worse in the assimilation than in the free-run model. However, the distribution of chlorophyll over the North Atlantic was improved by the assimilation (see Fig. 4). The very high spring bloom peaks in SeaWiFS in the North Pacific were not simulated well by either the free-run or assimilation models. Most of the high values were derived from extremely high SeaWiFS chlorophyll in the western Bering Sea, near Kamchatka.

A detailed comparison of the models with in situ data showed major improvement by assimilation. The bias and uncertainty of SeaWiFS against the SeaBASS/NODC in situ were −1.3% and 32.7%, respectively (Table 5). The free-run model performed much more poorly against the in situ data set than SeaWiFS in uncertainty at 61.8%, but comparable bias −1.4%. The assimilation model had similar uncertainty as SeaWiFS compared to the in situ data set (33.4%) but improved and nearly negligible bias (0.1%). The assimilation model (as well as the free-run model) had more than twice the number of coincident, co-located in situ/model matchup data points. This is a consequence of the absence of gaps in the model record in contrast to SeaWiFS.

Global annual primary production estimates from three sources, the VGPM, free-run model, and assimilation model, indicated reasonably good correspondence over the 6-year time series for which SeaWiFS data were used in this effort (Fig. 9). Interannual variability was mimicked among all three estimates. The mean of the 6-year time series indicated that the free-run model overestimated PP as derived from VGPM by nearly 21%. The assimilation model reduced the overestimate by nearly half, producing a more minor overestimate of 10.7%.

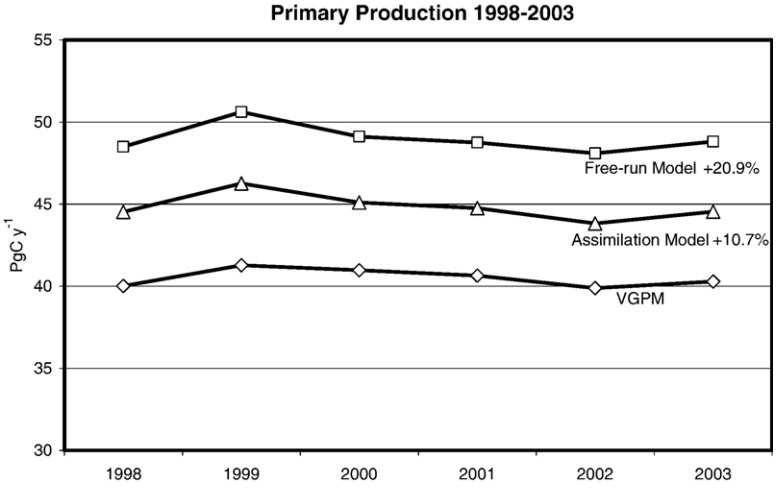


Fig. 9. Annual primary production for the period 1998–2003 from the VGPM, free-run model, and assimilation model. The mean departures over the period for the free-run model and assimilation model are indicated.

## 5. Summary

Assimilation of chlorophyll data from SeaWiFS exhibited substantial improvements over free-run simulations. Biases in basin means were reduced to 5.5% from 21.0%, and mean uncertainties were much lower for the assimilation model (10.1%) than the free-run model (65.3%). This represented a nearly 4-fold improvement in bias and a 6-fold improvement in uncertainty. When compared to in situ data for the 6-year time period from 1998 through 2003, the assimilation model had a bias of 0.1%, with an uncertainty of 33.4% for daily coincident, co-located data. SeaWiFS bias was slightly higher at −1.3% with similar uncertainty at 32.7%. The free-run bias and uncertainty at −1.4% and 61.8%, respectively, indicated how much the assimilation improved model results. Annual primary production indicated a smaller improvement (mean difference from VGPM=10.7% for the assimilation model versus 20.9% for the free-run model), representing an improvement of nearly a factor of 2, assuming the validity of the VGPM. These results suggest promise for assimilation of satellite ocean chlorophyll into global models. But they also point to areas of needed improvement. The fact that the assimilated variable shows the most improvement is not surprising, and is an important attribute for data assimilation. The fact that flux (primary production) exhibited less improvement than biomass (chlorophyll) using assimilation suggests the model continues to trend in the wrong direction despite assimilation. It also suggests that similar results may be expected for other non-assimilated variables, such as phytoplankton group distributions and nutrients. There remains considerable work to be done on assimilation of satellite ocean color, such as better handling of ocean color data errors, utilizing other model variables in a multi-variate solution, accounting for subsurface changes, as well as investigating the potential for using other ocean color products, such as diffuse attenuation coefficient at 490 nm, and potentially new products such as particulate organic carbon and calcite. Nevertheless, there is much potential in ocean color assimilation, and this effort is intended to represent an initial attempt on a global scale.

## Acknowledgements

I would like to acknowledge the NASA OBPG for processing and distributing SeaWiFS V5.1 data and Orbimage Corp. for collecting it. I would also like to thank Nancy Casey, SSAI, for acquiring and providing model forcing data sets from a wide variety of sources and formats. Lars Nerger, NASA/GMAO, provided insights into the assimilation overview. Finally, the comments of two anonymous reviewers are appreciated. This work was supported by NASA EOS and MAP programs.

## Appendix A. Biogeochemical processes model description

NOBM is based on Gregg et al. (2003). There are several new features in the biogeochemical processes model component:

- new maximum phytoplankton growth rates at 20 °C
- full detrital dynamics with 3 components, fully coupled to the OGCM
- a new formulation for the temperature-dependence for grazing
- a new formulation for nitrogen fixation for the cyanobacteria component
- introduction of dissolved iron scavenging and an increase in atmospheric iron solubility
- new nitrogen half-saturation constants for chlorophytes
- new iron half-saturation constants for chlorophytes and cyanobacteria

Other aspects of the biogeochemical processes model are described in Gregg et al. (2003), but are provided here for completeness.

The governing equations of the model are

*Phytoplankton*

$$\frac{\partial}{\partial t}P_i = \nabla(K\nabla P_i) - \nabla\bullet\mathbf{V}P_i - \nabla\bullet(\mathbf{w}_S)_i P_i + \mu_i P_i - \gamma H - \kappa P_i \tag{A1}$$

$i=1$ = diatoms
$i=2$ = chlorophytes
$i=3$ = cyanobacteria
$i=4$ = coccolithophores

*Nutrients*

$$\frac{\partial}{\partial t}N_N = \nabla(K\nabla N_N) - \nabla\bullet\mathbf{V}N_N - b_n[\Sigma_i \mu_i f(NO_3)_i P_i] + R\alpha_C D_{C/}(C:N) \tag{A2}$$

$$\frac{\partial}{\partial t}N_A = \nabla(K\nabla N_A) - \nabla\bullet\mathbf{V}N_A - b_N[\Sigma_i \mu_i \omega(NH_4)_i P_i] + b_N\varepsilon[\gamma H + \eta_2 H^2] \tag{A3}$$

$$\frac{\partial}{\partial t}N_S = \nabla(K\nabla N_S) - \nabla\bullet\mathbf{V}N_S - b_S\mu_1\omega(\mathrm{Si})_1P_1 + R\alpha_S D_S \quad (A4)$$

$$\frac{\partial}{\partial t}N_F = \nabla(K\nabla N_F) - \nabla\bullet\mathbf{V}N_F - b_F[\Sigma_i\mu_i\omega(\mathrm{Fe})_iP_i] + b_F\varepsilon[\gamma H + \eta_2 H^2] + R\alpha_F D_F + \mathrm{A_{Fe}}/L - \theta N_F \quad (A5)$$

$N_N$ nitrate
$N_A$ ammonium
$N_S$ silica
$N_I$ dissolved iron

*Herbivores*

$$\frac{\partial}{\partial t}H = \nabla(K\nabla H) - \nabla\bullet\mathbf{V}H + (1-\varepsilon)\gamma H - \eta_1 H - \eta_2 H^2 \quad (A6)$$

*Detritus*

$$\frac{\partial}{\partial t}D_C = \nabla(K\nabla D_C) - \nabla\bullet\mathbf{V}D_C - \nabla\bullet(\mathbf{w}_d)_C D_C - R\alpha_C D_C + \Phi[\kappa\Sigma_i P_i + \eta_1 H] + \Phi(1-\varepsilon)\eta_2 H^2 \quad (A7)$$

$$\frac{\partial}{\partial t}D_S = \nabla(K\nabla D_S) - \nabla\bullet\mathbf{V}D_S - \nabla\bullet(\mathbf{w}_d)_S D_S - R\alpha_S D_S + b_S[\kappa P_1 + \gamma H] \quad (A8)$$

$$\frac{\partial}{\partial t}D_F = \nabla(K\nabla D_F) - \nabla\bullet\mathbf{V}D_F - \nabla\bullet(\mathbf{w}_d)_I D_F - R\alpha_F D_F + b_F[\kappa\Sigma_i P_i + n_1 H] + b_F(1-\varepsilon)\eta_2 H^2 + \theta N_F \quad (A9)$$

$D_C$ carbon/nitrogen detritus
$D_S$ silica detritus
$D_F$ iron detritus

where the symbols and values are identified in Appendix Table 1. Bold denotes a vector quantity. All biological processes are assumed to cease in the presence of sea ice, which is included as an external forcing field.

### A.1. Phytoplankton

The growth formulation includes dependence on total irradiance ($E_T$), nitrogen as nitrate plus ammonium ($N_T$), silica (Si — for diatoms only), iron (Fe), and temperature (T)

$$\mu_i = \mu_m \min[\omega(E_T)_i, \omega(N_T)_i, \omega(\mathrm{Si})_i, \omega(\mathrm{Fe})_i] R G_i \quad (A10)$$

where $i$ indicates the phytoplankton functional group index (in order, diatoms, chlorophytes, cyanobacteria, and coccolithophores), $\mu$ is the total specific growth rate ($d^{-1}$) of phytoplankton, $\mu_m$ is the maximum growth rate at 20 °C (Appendix Table 1). The term $\omega(E_T)$ represents the fraction of growth that is a function solely of the total irradiance (μmol quanta $m^{-2}$ $s^{-1}$),

$$\omega(E_T) = \frac{E_T}{(E_T + k_E)} \quad (A11)$$

where $k_E$ is the irradiance at which $\mu = 0.5$ μm and equals $0.5I_k$, where $I_k$ is the light saturation parameter. The nutrient-dependent growth fractions are

$$\omega(\mathrm{NO_3})_i = \frac{\mathrm{NO_3}}{[\mathrm{NO_3} + (k_N)_i]} \quad (A12)$$

$$\omega(\mathrm{NH_4})_i = \frac{\mathrm{NH_4}}{[\mathrm{NH_4} + (k_N)_i]} \quad (A13)$$

$$\omega(N_T)_i = \omega(\mathrm{NH_4})_i + f(\mathrm{NO_3})_i \quad (A14)$$

$$f(\mathrm{NO_3}) = \min[\omega(\mathrm{NO_3})_i, 1 - \omega(\mathrm{NH_4})_i] \quad (A15)$$

(Gregg and Walsh, 1992)

$$\omega(\mathrm{Si})_i = \frac{\mathrm{Si}}{[\mathrm{Si} + (k_S)_i]} \quad (A16)$$

$$\omega(\mathrm{Fe})_i = \frac{\mathrm{Fe}}{[\mathrm{Fe} + (k_F)_i]} \quad (A17)$$

Temperature-dependent growth is from Eppley (1972)

$$R = 1.066^{(T-20)} \quad (A18)$$

which produces a temperature-growth factor normalized to 20 °C. The term $G$ in Eq. (A10) is an additional adjustment used for the cyanobacteria component that reduces their growth rate in cold water (<15 °C)

$$G_3 = 0.0294T + 0.558 \quad (A19)$$

$G_i = 1$ for the other three phytoplankton components ($i = 1, 2, 4$). This effect conforms to observations that cyanobacteria are scarce in cold waters (Agawin et al., 1998, 2000). The cyanobacteria component possesses a modest ability to fix nitrogen from the water column, as observed in *Trichodesmium* spp. (Carpenter and

Romans, 1991). The nitrogen fixation is expressed as additional growth occurring when nitrogen availability is $<(k_N)_3$,

$$\mu_{nfix} = 0.25\exp(-75P_3) \quad \text{(A20)}$$

where the index 3 indicates cyanobacteria. The biomass dependence represents a progressive community changeover from non-N-fixing cyanobacteria to N-fixing bacteria as the total population declines under nitrogen-stressed conditions. The total N-limited growth rate plus the additional growth derived from N-fixation is not allowed to exceed the growth rate where total nitrogen $=(k_N)_3$. No accounting for denitrification is made in the model.

Photoadaptation is simulated by stipulating 3 states: 50, 150 and 200 (μmol quanta $m^{-2}$ $s^{-1}$). This is based on laboratory studies which typically divided experiments into low, medium, and high classes of light adaptation. Carbon:chlorophyll ratios ($\Phi$) correspond to the photoadaptation state, to represent the tendency of phytoplankton to preferentially synthesize chlorophyll in low light conditions, to enable more efficient photon capture. The three $\Phi$ states corresponding to the three light states are 25, 50 and 80 g $g^{-1}$. The $\Phi$ results for diatoms in the model closely mimic Anning et al.'s (2000) results for diatoms. For irradiance levels falling between the three light states, the C:chl ratios are linearly interpolated.

Mean irradiance is computed during daylight hours, and then the phytoplankton photoadaptive state is classified accordingly. This calculation is only performed once per day to simulate a delayed photoadaptation response. Light saturation constants for the three light levels are provided in Appendix Table 1.

Phytoplankton vector sinking is treated as additional advection in the $z$-direction, and is given at 31 °C, representing approximately the maximum. It is adjusted by viscosity according to Stokes Law (Csanady, 1986), which is parameterized here by temperature

$$w_s(T) = w_s(31)[0.451 + 0.0178T] \quad \text{(A21)}$$

Coccolithophore sinking rates were allowed to vary as a function of growth rate from 0.3 to 1.4 m $d^{-1}$ based on observations by Fritz and Balch (1996). A linear relationship was assumed

$$w_{s4} = 0.752_{\mu 4}(\text{high}) + 0.225 \quad \text{(A22)}$$

where $w_s$ is the sinking rate of coccolithophores (m $d^{-1}$), $\mu$(high) is the maximum growth rate actually achieved for the previous day, and the subscript 4 represents coccolithophores.

### A.2. Nutrients

The diversity in the processes affecting the four nutrient groups requires elucidation in 4 separate equations, unlike the phytoplankton. All are taken up by phytoplankton growth, with silica subject only to diatom uptake (note the subscript=1 in Eq. (A4) denoting diatoms). For three of the nutrients, nitrate, silica, and dissolved iron, corresponding detrital pools remineralize to return nutrients previously uptaken by phytoplankton. There is no detrital pool for ammonium, which is excreted as a function of herbivore grazing, and as a function of higher order ingestion of herbivores, represented by the term $n_2H^2$ in Eqs. (A3), (A5), (A6), (A7), and (A9). Dissolved iron also has an excretion pathway, but nitrate and silica do not. The nutrient to chlorophyll ratios, denoted $b$ in Eqs. (A2)–(A5), are derived from Redfield ratios, which are constant (Appendix Table 1) and the carbon:chlorophyll ($\Phi$) ratio which is not.

$$b_N = \Phi/\text{C} : \text{N} \quad \text{(A23)}$$

$$b_S = \Phi/\text{C} : \text{S} \quad \text{(A24)}$$

$$b_F = \Phi/\text{C} : \text{Fe} \quad \text{(A25)}$$

This leads to variable nutrient to chlorophyll ratios in the model.

As in Gregg et al. (2003) dust deposition fields are derived from Ginoux et al. (2001). In this model, four dust size fractions are transported, corresponding to clay (smallest) and three increasing fractions of silt. The iron content is assumed to vary among the clay and silt fractions as follows: clay=3.5% iron, silt=1.2% iron (Fung et al., 2000). Iron solubility is assumed at 2% for all fractions, which is toward the low end of current estimates (Fung et al., 2000), but is the same as used by Moore et al. (2004).

### A.3. Herbivores

Grazing uses an Ivlev formulation (McGillicuddy et al., 1995),

$$\gamma(T) = \gamma_m R_H[1-\exp(-\Lambda\Sigma_i P_i)] \quad \text{(A26)}$$

$R_H$ is the maximum grazing rate at 20 °C ($\gamma_m$) adjusted by temperature

$$R_H = 0.06\exp(0.1T) + 0.70 \quad \text{(A27)}$$

The temperature-dependence for grazing is more linear than that for phytoplankton, reflecting the larger size of their overall community. The grazing represents the total loss of phytoplankton to herbivores, as indicated by the summation symbol, but is applied to the individual phytoplankton functional groups proportionately to their relative abundances. This enables herbivore grazing to adapt the prevailing phytoplankton community.

The two loss terms in Eq. (A6) represent the death of herbivores ($n_1H$) and higher order heterotrophic losses ($n_2H^2$). These formulations and parameters (Appendix Table 1) were taken from McGillicuddy et al. (1995).

### A.4. Detritus

Three detrital components represent the three major nutrient elements, carbon/nitrogen, silica, and iron. The nitrogen detritus is kept as carbon in the model, but since the C:N ratio is constant, it is simple to convert when needed. All are subject to advection, diffusion and sinking. Detrital sinking, like phytoplankton sinking, is dependent on viscosity parameterized here in terms of temperature, using the same formulation. Remineralization is also temperature-dependent, but uses the phytoplankton growth-dependence term Eq. (A18). Silica contained in the diatom component of phytoplankton is assumed to pass through herbivores upon grazing directly into the silica detritus pool. No silica remains in the herbivore component at any time.

***Initial conditions***

NOBM underwent a spin-up of a total of 50 years under climatological forcing. For the first 20 years, initial dissolved iron conditions were from Fung et al. (2000), and nitrate and silica distributions were from annual climatologies from the National Oceanographic Data Center (NODC; Conkright et al., 2002b). Ammonium initial conditions were set to 0.5 μM. Initial conditions for all phytoplankton groups and herbivores were set to 0.05 mg $m^{-3}$ chl throughout the entire model domain. Initial conditions for detritus were set to 0. After 20 years, dissolved iron and detritus distributions were retained, while all other fields were reset to their original values. The model was run again for 30 years. This methodology enables dissolved iron to reach steady state without adversely impacting phytoplankton group distributions with excessively low initial values.

## Appendix B. Appendix Table 1

Notation and parameters and variables for NOBM. Values are provided for the parameters and ranges are provided for the variables. When a parameter varies according to temperature, the value at a specified temperature is shown and identified. Nutrient/chlorophyll ratios are variable because of photadaptation-dependence, and only the range is shown, corresponding to low-, and high-light adaptation, and therefore also corresponding to C:chl ratios of 20 and 80 g $g^{-1}$.

| Symbol | Parameter/variable | Value | Units |
|---|---|---|---|
| *General* | | | |
| $K$ | Diffusivity | Variable | $m^2\ s^{-1}$ |
| $\nabla$ | Gradient operator | none | none |
| $\mathbf{V}$ | Vector velocity | Variable | $m\ s^{-1}$ |
| $L$ | Layer thickness | Variable | m |
| *Phytoplankton* | | | |
| $\mathbf{w}_s$ | Vector sinking rate of phytoplankton at 31 °C | | $m\ d^{-1}$ |
| | Diatoms | 1.0 | |
| | Chlorophytes | 0.25 | |
| | Cyanobacteria | 0.0085 | |
| | Coccolithophores | 0.3–1.4 | |
| $\mu$ | Specific growth rate of phytoplankton maximum ($\mu_m$) at 20 °C: | | $d^{-1}$ |
| | Diatoms | 1.50 | |
| | Chlorophytes | 1.26 | |
| | Cyanobacteria | 1.00 | |
| | Coccolithophores | 1.13 | |
| $\omega$ | Fraction of growth due to nutrients, light | 0−1 | none |
| $I_k$ | Light saturation | | μmol quanta $m^{-2}\ s^{-1}$ |
| | Light level: Low (50) / Medium (150) / High (200) | | |
| | Diatoms | 90.0 / 93.0 / 184.0 | |
| | Chlorophytes | 96.9 / 87.0 / 143.7 | |
| | Cyanobacteria | 65.1 / 66.0 / 47.0 | |
| | Coccolithophores | 56.1 / 71.2 / 165.4 | |
| $\kappa$ | Senescence | 0.05 | $d^{-1}$ |
| $k_E$ | Half-saturation for growth as function of quanta | $0.5I_k$ | μmol quanta $m^{-2}\ s^{-1}$ |
| $E_T$ | Total quanta (direct + diffuse) | Variable | μmol quanta $m^{-2}\ s^{-1}$ |
| $R$ | Temperature-dependence for growth | 0.25–9.4 | None |
| $G$ | Temperature-dependence for cyanobacteria growth | 0.5–1.0 | None |
| *Nutrients (N)* | | | |
| $b_{N,S,F}$ | Nutrient:chlorophyll ratio | | μM $(\mu g\ l^{-1})^{-1}$ |
| | Nitrogen | 0.3–1.0 | |
| | Silica | 0.3–1.0 | |
| | Iron | 0.01–0.04 | |
| ε | Nutrient excretion | | $d^{-1}$ |
| | Nitrate | 0.0 | |
| | Ammonium | 0.25 | |
| | Silica | 0.0 | |
| | Iron | 0.25 | |
| $k_{N,S,F}$ | Half-saturation constant | | |
| | Nitrogen | | μM |
| | Diatoms | 1.0 | |

**Appendix B** (*continued*)

| Symbol | Parameter/variable | Value | Units |
|---|---|---|---|
| *Nutrients (N)* | | | |
| | Chlorophytes | 0.67 | |
| | Cyanobacteria | 0.50 | |
| | Coccolithophores | 0.50 | |
| | Silica | | μM |
| | Diatoms | 0.2 | |
| | Iron | | nM |
| | Diatoms | 1.0 | |
| | Chlorophytes | 0.78 | |
| | Cyanobacteria | 0.67 | |
| | Coccolithophores | 0.67 | |
| $\theta$ | Iron scavenging rate | | $d^{-1}$ |
| | Low iron (<0.06 nM) | $2.0\times10^{-4}$ | |
| | High iron (>0.06 nM) | $2.0\times10^{-3}$ | |
| $A_{Fe}$ | Atmospheric deposition of iron | 0.03–967.0 | nmol $m^{-2}$ $d^{-1}$ |
| C:N | Carbon:nitrogen ratio | 79.5 | μg $l^{-1}$ $(\mu M)^{-1}$ |
| C:S | Carbon:silica ratio | 79.5 | μg $l^{-1}$ $(\mu M)^{-1}$ |
| C:Fe | Carbon:iron ratio | 1800 | μg $l^{-1}$ $(nM)^{-1}$ |
| *Herbivores (H)* | | | |
| $\gamma$ | Grazing rate | | |
| | Maximum ($\gamma_m$) at 20 °C | 1.0 | $d^{-1}$ |
| $\Lambda$ | Ivlev constant | 1.0 | $(\mu g\ l^{-1})^{-1}$ |
| $\eta_1$, $\eta_2$ | Heterotrophic loss rates | 0.1, 0.5 | $d^{-1}$ |
| $R_H$ | Temperature-dependence for grazing | 0.75–2.7 | None |
| *Detritus (D)* | | | |
| $\mathbf{w}_d$ | Vector sinking rate of detritus at 31 °C | | m $d^{-1}$ |
| | Carbon/nitrogen detritus | 20.0 | |
| | Silica detritus | 50.0 | |
| | Iron detritus | 20.0 | |
| $\alpha_{C,S,F}$ | Remineralization rate at 20 °C | | $d^{-1}$ |
| | Carbon/nitrate | 0.02 | |
| | Silica | 0.0001 | |
| | Iron | 0.02 | |
| $\Phi$ | Carbon:chlorophyll ratio | Variable | g $g^{-1}$ |

## References


Agawin, N.S.R., Duarte, C.M., Agusti, S., 1998. Growth and abundance of *Synechococcus* sp. in a Mediterranean Bay: seasonality and relationship with temperature. Marine Ecology. Progress Series 170, 45–53.

Agawin, N.S.R., Duarte, C.M., Agusti, S., 2000. Nutrient and temperature control of the contribution of picoplankton to phytoplankton biomass and production. Limnology and Oceanography 45, 591–600.

Allen, J.I., Eknes, M., Evensen, G., 2002. An ensemble Kalman filter with a complex marine ecosystem model: hindcasting phytoplankton in the Cretan Sea. Annales Geophysicae 20, 1–13.

Anderson, L.A., Robinson, A.R., Lozano, C.J., 2000. Physical and biological modeling in the Gulf Stream region: I. Data assimilation methodology. Deep-Sea Research. Part 1. Oceanographic Research Papers 1787–1827.

Anning, T., MacIntyre, H.L., Pratt, S.M., Sammes, P.J., Gibb, S., Geider, R.J., 2000. Photoacclimation in the marine diatom *Skeletonema costatum*. Limnology and Oceanography 45, 1807–1817.

Armstrong, R.A., Sarmiento, J.L., Slater, R.D., 1995. Monitoring ocean productivity by assimilating satellite chlorophyll into ecosystem models. In: Powell, Steele (Eds.), Ecological Time Series. Chapman and Hall, London, pp. 371–390.

Behrenfeld, M.J., Falkowski, P.G., 1997. Photosynthetic rates derived from satellite-based chlorophyll concentrations. Limnology and Oceanography 42, 1–20.

Besiktepe, S.T., Lermusiaux, P.F.J., Robinson, A.R., 2003. Coupled physical and biogeochemical data-driven simulations of Massachusetts Bay in late summer: real-time and postcruise data assimilation. Journal of Marine Systems 40, 171–212.

Campbell, J.W., 1995. The lognormal distribution as a model for bio-optical variability in the sea. Journal of Geophysical Research 100 (C7), 13237–13254.

Carmillet, V., Brankart, J.-M., Brasseur, P., Drange, H., Evensen, G., Verron, J., 2001. A singular evolutive extended Kalman filter to assimilate ocean color data in a coupled physical–biochemical model of the North Atlantic ocean. Ocean Modelling 3, 167–192.

Carpenter, E.J., Romans, K., 1991. Major role of the cyanobacterium *Trichodesmium* in nutrient cycling in the North Atlantic Ocean. Science 254, 1356–1358.

Conkright, M.E., Gregg, W.W., 2003. Comparison of chlorophyll climatologies: in situ, CZCS, blended in situ-CZCS, and SeaWiFS. International Journal of Remote Sensing 24, 969–991.

Conkright, M.E., Antonov, J.I., Baranova, O., Boyer, T.P., Garcia, H.E., Gelfeld, R., Johnson, D., O'Brien, T.D., Smolyar, I., Stephens, C., 2002a. World ocean database 2001, Vol. 1: Introduction. S. Levitus, Ed., NOAA Atlas NESDIS 42, US Govt. Printing Office, Washington, DC, 167 pp.

Conkright, M.E., Garcia, H.E., O'Brien, T.D., Locarnini, R.A., Boyer, T.P., Stephens, C., Antonov, J.I., 2002b. World Ocean Atlas 2001, Volume 4: Nutrients. S. Levitus, Ed., NOAA Atlas NESDIS 52, U.S. Government Printing Office, Wash., D.C., 392 pp., CD-ROMs.

Csanady, G.T., 1986. Mass transfer to and from small particles in the sea. Limnology and Oceanography 31, 237–248.

Eknes, M., Evensen, G., 2002. An ensemble Kalman filter with a 1-D marine ecosystem model. Journal of Marine Systems 36, 75–100.

Eppley, R.W., 1972. Temperature and phytoplankton growth in the sea. Fisheries Bulletin 70, 1063–1085.

Fasham, M.J.R., Evans, G.T., Kiefer, D.A., Creasey, M., Leach, H., 1995. The use of optimization techniques to model marine ecosystem dynamics at the JGOFS station at 47 degrees N 20 degrees W. Philosophical Transactions of the Royal Society of London. B 348, 203–209.

Fasham, M.J.R., Boyd, P.W., Savidge, G., 1999. Modeling the relative contributions of utotrophs and heterotrophs to carbon flow at a Lagrangian JGOFS station in the Northeast Atlantic: the importance of DOC. Limnology and Oceanography 44, 80–94.

Faugeras, B., Levy, M., Memery, L., Verron, J., Blum, J., Charpentier, I., 2003. 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 40–41, 99–125.

Faugeras, B., Bernard, O., Sciandra, A., Levy, M., 2004. A mechanistic modeling and data assimilation approach to estimate the carbon/chlorophyll and carbon/nitrogen ratios in a coupled hydrodynamical–biological model. Nonlinear Processes in Geophysics 11, 515–533.

Fennel, K., Losch, M., Schroter, J., Wenzel, M., 2001. Testing a marine ecosystem model: sensitivity analysis and parameter optimization. Journal of Marine Systems 28, 45–63.

Freidrichs, M.A.M., 2001. A data assimilative marine ecosystem model of the central equatorial Pacific: numerical twin experiments. Journal of Marine Research 59, 859–894.

Freidrichs, M.A.M., 2002. Assimilation of JGOFS EqPac and SeaWiFS data into a marine ecosystem model of the central equatorial Pacific Ocean. Deep-Sea Research. Part 2. Topical Studies in Oceanography 49, 289–320.

Fritz, J.J., Balch, W.M., 1996. A light-limited continuous culture study of *Emiliana huxleyi*: determination of coccolith detachment and its relevance to cell sinking. Journal of Experimental Marine Biology and Ecology 207, 127–147.

Fung, I.Y., Meyn, S.K., Tegen, I., Doney, S.C., John, J.G., Bishop, J.K.B., 2000. Iron supply and demand in the upper ocean. Global Biogeochemical Cycles 14, 281–295.

Garcia-Gorriz, E., Hoepffner, N., Ouberdous, M., 2003. Assimilation of SeaWiFS data in a coupled physical–biological model of the Adriatic Sea. Journal of Marine Systems 40–41, 233–252.

Ginoux, P., Chin, M., Tegen, I., Prospero, J.M., Holben, B., Dubovik, O., Lin, S.-J., 2001. Sources and distributions of dust aerosols simulated with the GOCART model. Journal of Geophysical Research 106, 20255–20273.

Gregg, W.W., 2002a. A coupled ocean-atmosphere radiative model for global ocean biogeochemical models. In: Suarez, M. (Ed.), NASA Global Modeling and Assimilation Series. NASA Technical Memorandum 2002-104606, vol. 22. 33 pp., available using anonymous ftp at nsipp.gsfc.nasa.gov, CABIN/gregg/reprints/gregg_NASATM2002.pdf.

Gregg, W.W., 2002b. Tracking the SeaWiFS record with a coupled physical/biogeochemical/radiative model of the global oceans. Deep-Sea Research. Part 2. Topical Studies in Oceanography 49, 81–105.

Gregg, W.W., Carder, K.L., 1990. A simple spectral solar irradiance model for cloudless maritime atmospheres. Limnology and Oceanography 35, 1657–1675.

Gregg, W.W., Casey, N.W., 2004. Global and regional evaluation of the SeaWiFS chlorophyll data set. Remote Sensing of Environment 93, 463–479.

Gregg, W.W., Conkright, M.E., 2002. Decadal changes in global ocean chlorophyll. Geophysical Research Letters 29 (11). doi:10.1029/2002GL014689.

Gregg, W.W., Conkright, M.E., 2001. Global seasonal climatologies of ocean chlorophyll: blending in situ and satellite data for the Coastal Zone Color Scanner era. Journal of Geophysical Research 106, 2499–2515.

Gregg, W.W., Walsh, J.J., 1992. Simulation of the 1979 spring bloom in the Mid-Atlantic Bight: a coupled physical/biological/optical model. Journal of Geophysical Research 97, 5723–5743.

Gregg, W.W., Ginoux, P., Schopf, P.S., Casey, N.W., 2003. Phytoplankton and iron: validation of a global three-dimensional ocean biogeochemical model. Deep-Sea Research. Part 2. Topical Studies in Oceanography 50, 3143–3169.

Gunson, J., Oschlies, A., Garcon, V., 1999. Sensitivity of ecosystem parameters to simulated satellite ocean color data using a coupled physical–biological model of the North Atlantic. Journal of Marine Research 57, 613–639.

Harmon, R., Challenor, P., 1997. A Markov chain Monte Carlo model for estimation and assimilation into models. Ecological Modelling 101, 41–59.

Hemmings, J.C.P., Srokosz, M.A., Challenor, P., Fasham, M.J.R., 2003. Assimilating satellite ocean-colour observations into oceanic ecosystem models. Philosophical Transactions of the Royal Society of London A-Mathematics, Physics, and Engineering Science 361, 33–39.

Hemmings, J.C.P., Srokosz, M.A., Challenor, P., Fasham, M.J.R., 2004. Split-domain calibration of an ecosystem model using satellite ocean colour data. Journal of Marine Systems 50, 141–179.

Hooker, S.B., Esaias, W.E., Feldman, G.C., Gregg, W.W., McClain, C.R., 1992. An overview of SeaWiFS and ocean color. NASA Technical Memorandum 104566, vol. 1. 25 pp.

Hoteit, I., Triantafyllou, G., Petihakis, G., Allen, J.I., 2003. A singular evolutive extended Kalman filter to assimilate real in situ data in a 1-D marine ecosystem model. Annales Geophysicae 21, 389–397.

Hurtt, G.C., Armstrong, R.A., 1996. A pelagic ecosystem model calibrated with BATS data. Deep-Sea Research. Part 2. Topical Studies in Oceanography 43, 653–683.

Hurtt, G.C., Armstrong, R.A., 1999. A pelagic ecosystem model calibrated with BATS and OWSI data. Deep-Sea Research. Part 2. Topical Studies in Oceanography 46, 27–61.

Ibrahim, H., George, T., George, P., 2004. Towards a data assimilation system for the Cretan Sea ecosystem using a simplified Kalman filter. Journal of Marine Systems 45, 159–171.

Ishizaka, J., 1990. Coupling of Coastal Zone Color Scanner data to a physical–biological model of the southeastern United-States continental-shelf ecosystem. 3. Nutrient and phytoplankton fluxes and CZCS data assimilation. Journal of Geophysical Research 95, 20201–20212.

Kuroda, H., Kishi, M.J., 2004. A data assimilation technique applied to estimate parameters for the NEMURO marine ecosystem model. Ecological Modelling 172, 69–85.

Lawson, L.M., Hofmann, E.E., Spitz, Y.H., 1996. Time series sampling and data assimilation in a simple marine ecosystem model. Deep-Sea Research. Part 2. Topical Studies in Oceanography 43, 625–651.

Leredde, Y., Lauer-Leredde, C., Diaz, C., 2005. On the variational data assimilation by a marine ecosystem model of NPZ type. Comptes Rendus Geoscience 337, 1055–1064.

Losa, S.N., Kivman, G.A., Schroter, J., Wenzel, M., 2003. Sequential weak constraint parameter estimation in an ecosystem model. Journal of Marine Systems 43, 31–49.

Losa, S.N., Kivman, G.A., Ryabchenko, V.A., 2004. Weak constraint parameter estimation for a simple ocean ecosystem model: what can we learn about the model and data? Journal of Marine Systems 45, 1–20.

Magri, S., Brasseur, P., Lacroix, G., 2005. Data assimilation in a marine ecosystem model of the Ligurian Sea. Comptes Rendus Geoscience 337, 1065–1074.

Matear, R.J., 1995. Parameter estimation and analysis of ecosystem models using simulated annealing: a case study at Station P. Journal of Marine Research 53, 571–607.

McGillicuddy, D.J., McCarthy, J.J., Robinson, A.R., 1995. Coupled physical and biological modeling of the spring bloom in the North Atlantic (I): model formulation and one dimensional bloom processes. Deep-Sea Research. Part 1. Oceanographic Research Papers 42, 1313–1357.

McGillicuddy, D.J., Lynch, D.R., Moore, A.M., Gentleman, W.C., Davis, C.S., Meise, C.J., 1998. An adjoint data assimilation approach to diagnosis of physical and biological controls on *Pseudocalanus* spp. in the Gulf of Maine–Georges Bank region. Fisheries Oceanography 7, 205–218.

Miller, A.J., DiLorenzo, E., Neilson, D.J., Cornuelle, B.D., Moisan, J.R., 2000. Modeling CalCOFI observations during El Nino: fitting physics and biology. CalCOFI Report, vol. 41. 11 pp.

Moore, J.K., Doney, S.C., Lindsay, K., 2004. Upper ocean ecosystem dynamics and iron cycling in a global three-dimensional model. Global Biogeochemical Cycles 18. doi:10.1029/2004GB002220.

Natvik, L.J., Evensen, G., 2003a. Assimilation of ocean colour data into a biochemical model of the North Atlantic. Part 1. Data assimilation experiments. Journal of Marine Systems 40–41, 127–153.

Natvik, L.J., Evensen, G., 2003b. Assimilation of ocean colour data into a biochemical model of the North Atlantic. Part 2. Statistical analysis. Journal of Marine Systems 40–41, 155–169.

Natvik, L.-J., Eknes, M., Evensen, G., 2001. A weak constraint inverse for zero-dimensional marine ecosystem model. Journal of Marine Systems 28, 19–44.

Oort, A.H., 1983. Global atmospheric circulation statistics, 1958–1973. NOAA Professional Paper 14 180 pp.

Oschlies, A., Schartau, M., 2005. Basin-scale performance of a locally optimized marine ecosystem model. Journal of Marine Research 63, 335–358.

Popova, E.E., Lozano, C.J., Srokosz, M.A., Fasham, M.J.R., Haley, P.J., Robinson, A.R., 2002. Coupled 3D physical and biological modeling of the mesoscale variability observed in North-East Atlantic in Spring 1997: biological processes. Deep-Sea Research. Part 1. Oceanographic Research Papers 49, 1741–1768.

Prunet, P., Minster, J.F., RuizPino, D., Dadou, I., 1996. Assimilation of surface data in a one-dimensional physical–biogeochemical model of the surface ocean. 1. Method and preliminary results. Global Biogeochemical Cycles 10, 111–138.

Reynolds, R.W., 1988. A real-time global sea surface temperature analysis. Journal of Climate 1, 75–86.

Reynolds, R.W., Smith, T.M., 1994. Improved global sea surface temperature analyses using optimum interpolation. Journal of Climate 7, 75–86.

Schartau, M., Oschlies, A., 2003. Simultaneous data-based optimization of a 2D-ecosystem model at three locations in the North Atlantic: Part I — method and parameter estimates. Journal of Marine Research 61, 765–793.

Schartau, M., Oschlies, A., Willebrand, J., 2001. Parameter estimates of a zero-dimensional ecosystem model applying the adjoint method. Deep-Sea Research. Part 2. Topical Studies in Oceanography 48, 1769–1800.

Schlitzer, R., 2002. Carbon export fluxes in the Southern Ocean: results from inverse modeling and comparison with satellite-based estimates. Deep-Sea Research. Part 2. Topical Studies in Oceanography 49, 1623–1644.

Schopf, P.S., Loughe, A., 1995. A reduced gravity isopycnal ocean model: hindcasts of El Nino. Monthly Weather Review 123, 2839–2863.

Spitz, Y.H., Moisan, J.R., Abbott, M.R., Richman, J.G., 1998. Data assimilation and a pelagic ecosystem model: parameterization using time series observations. Journal of Marine Systems 16, 51–68.

Spitz, Y.H., Moisan, J.R., Abbott, M.R., 2001. Configuring an ecosystem model using data from the Bermuda Atlantic Time Series (BATS). Deep-Sea Research. Part 2. Topical Studies in Oceanography 48, 1733–1768.

Triantafyllou, G., Hoteit, I., Petihakis, G., 2003. A singular evolutive interpolated Kalman filter for efficient data assimilation in a 3-D complex physical–biogeochemical model of the Cretan Sea. Journal of Marine Systems 40, 213–231.

Vallino, J.J., 2000. Improving marine ecosystem models: use of data assimilation and mesocosm experiments. Journal of Marine Research 58, 117–164.

Weber, L., Volker, C., Schartau, M., Wolf-Gladrow, D.A., 2005. Modeling the speciation and biogeochemistry of iron at the Bermuda Atlantic Time-Series study site. Global Biogeochemical Cycles 19, GB1019. doi:10.1029/2004GBC002340.

Werdell, P.J., Bailey, S.W., 2002. The SeaWiFS Bio-Optical Archive and Storage System (SeaBASS): Current Architecture and Implementation. In: Fargion, G.S., McClain, C.R. (Eds.), NASA Technical Memorandum 2002-211617. NASA Goddard Space Flight Center, Greenbelt, MD.