

Statistical validation of a 3-D bio-physical model of the western North Atlantic

M. K. Lehmann¹, K. Fennel¹, and R. He²

¹Department of Oceanography, Dalhousie University, 1355 Oxford Street, Halifax, Nova Scotia B3H 4J1, Canada

²Department of Marine, Earth and Atmosphere Sciences, North Carolina State University, Raleigh, NC 27695, USA

Received: 21 May 2009 – Published in Biogeosciences Discuss.: 12 June 2009

Revised: 4 September 2009 – Accepted: 10 September 2009 – Published: 2 October 2009

Abstract. High-resolution, physical-biological models of coastal and shelf regions typically use a single functional phytoplankton group, which limits their ability to represent ecological gradients (e.g. highly productive shelf systems adjacent to oligotrophic regions), as these are dominated by different functional phytoplankton groups. We implemented a size-structured ecosystem model in a high-resolution, regional circulation model of the northeast North American shelf and adjacent deep ocean in order to assess whether the added functional complexity of two functional phytoplankton groups improves the model's ability to represent surface chlorophyll concentrations along an ecological gradient encompassing five distinct regions. We used satellite-derived SST and sea-surface chlorophyll for our model assessment, as these allow investigation of spatial variability and temporal variations from monthly to interannual, and analyzed three complimentary statistical measures of model-data agreement: model bias, root mean square error and model efficiency (or skill). All three measures were integrated for the whole domain, for distinct subregions and were calculated in a spatially explicit manner. Comparison with a previously published simulation that used a model with a single phytoplankton functional group indicates that the inclusion of an additional phytoplankton group representing picoplankton markedly improves the model's skill.

for the North Atlantic have been developed in recent years (e.g., Fasham et al., 1993; Dutkiewicz et al., 2001; Oschlies, 2002; Lima and Doney, 2004; Coles and Hood, 2007), but none with sufficient resolution to describe biogeochemical processes on continental shelves. Continental shelves are of global importance because of their high levels of primary production (Longhurst, 2007), enhanced air-sea CO₂ exchange (e.g. Cai et al., 2003; Thomas et al., 2004) and because they are sites of significant sediment denitrification (Seitzinger et al., 2006). Furthermore, marine biological and biogeochemical responses to climate change in coastal and shelf seas will be most relevant to human activities as fisheries, aquaculture, tourism and resource extraction are concentrated in these regions, emphasizing the need for predictive regional models of coastal and shelf seas.

Our area of interest in this study is a coastal segment of the northeastern North American shelf that includes the Gulf of Maine (GOM), Middle Atlantic Bight (MAB) and the adjacent deep ocean. Some regional biogeochemical models for this area (Franks and Chen, 2001; Fennel et al., 2006, 2008; Previdi et al., 2009) and other coastal and shelf seas (e.g. Gruber et al., 2006; Pätsch and Kühn, 2008) have been developed in recent years. While the regional models differ in the details of their biological parameterizations, all of them have relatively simple biological components and describe only one phytoplankton functional group. While single phytoplankton models can be tuned to represent different ecological regimes, they are less likely to capture a range of conditions with one single set of parameters (Friedrichs et al., 2007). Our region of interest encompasses a diverse set of ecological regimes that differ in terms of physical forcing, nutrient supply mechanisms and are composed of taxonomically and biogeochemically diverse phytoplankton groups. For example, the continental shelf includes some of the most strongly stratified coastal waters globally in the central MAB adjacent to regions that are well mixed year round (Nantucket Shoals and Georges Bank). Ecological gradients are

1 Introduction

Coupled physical-biological and biogeochemical models are indispensable tools for advancing our understanding of oceanographic processes and for predicting responses to climate change. A number of basin-scale coupled models



Correspondence to: K. Fennel
(katja.fennel@dal.ca)

also pronounced, with highly productive regions near shore (maximum chlorophyll concentrations of up to 50 mg m^{-3}) to the oligotrophic Sargasso Sea (maximum chlorophyll concentrations of $<1 \text{ mg m}^{-3}$). Capturing this range in ecological regimes with a simple biological model and a single set of model parameters is difficult.

Experience with basin-scale models of the North Atlantic has shown that biological models with only one phytoplankton functional group, e.g. the NPZD-type model of Oschlies (2002) and the slightly more complex model by Fasham et al. (1993), had difficulty resolving the oligotrophic regime. Inclusion of an additional phytoplankton group and multiple nutrients by Lima and Doney (2004) markedly improved the representation of the oligotrophic ocean in their coupled North Atlantic model. A single phytoplankton group model similar to Fasham et al. (1993) model has been used previously in a high-resolution circulation model for the northeast North American shelf (Fennel et al., 2006, 2008) with encouraging levels of model skill in describing temporal and spatial ecological dynamics on the shelf and in the slope waters, although chlorophyll in the oligotrophic ocean was generally underestimated, mirroring the experience with basin-scale models.

Here we assess whether the biological formulation of Lima and Doney (2004) with two functional phytoplankton groups (representing diatoms and picoplankton) yields an improvement over the model simulations by Fennel et al. (2006, 2008). The inclusion of picoplankton with a preference for bacterially recycled ammonium from detritus underpins the microbial loop and improved the model's ability to simulate, with a single set of parameters, oceanographic regimes ranging from the subtropical gyre to high latitudes in the North Atlantic domain. The biological model of Lima and Doney (2004) is here coupled to a high-resolution physical model of the GOM and MAB (He and Chen, 2009).

Validating biological models is difficult because of the general paucity of observational data, especially at the relevant spatial and temporal scales, and because biological model variables are often not directly equivalent to measured quantities. We use satellite-derived sea-surface temperature and chlorophyll concentrations, as these allow us to analyze monthly to interannual variations, longer-term trends and spatial variability, including mesoscale features in the whole model domain. Specifically, we analyze a four-year simulation and focus on a gradient in conditions from a highly productive region in the MAB to the oligotrophic ocean.

Increasing emphasis has recently been placed on formal quantitative metrics of model skill that measure agreement between spatially resolved observations and model predictions (e.g. Allen et al., 2007; Stow et al., 2009) and a number of measures have been proposed. We calculate here the bias, root mean square error, correlation and model efficiency (also referred to as model skill) which give complementary information. We analyze the statistical measures for the whole model domain (collapsed into one scalar mea-

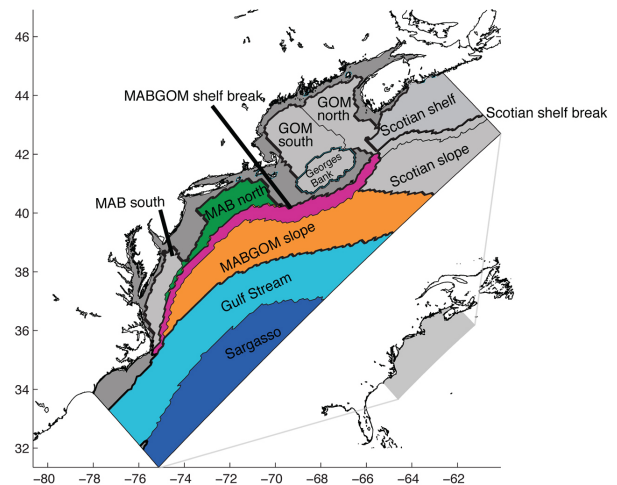


Fig. 1. Study region and subareas used for spatial averaging of model and observ fields.

sure), in a spatially explicit manner (i.e. producing maps of these measures), and for a number of different subregions that were defined in order to delineate distinct, relatively homogeneous subunits within our domain.

2 Methods

2.1 Study area

Our study area includes the Scotian Shelf, GOM, MAB and the adjacent slope sea, the Gulf Stream and the Sargasso Sea (Fig. 1). The shelf and slope region lies at the confluence of the Labrador Current and the Gulf Stream, the western boundary currents of the subpolar and subtropical gyres, respectively (Loder et al., 1998). The Labrador Current is of predominant influence, feeding the equatorward flow of cold, fresh water on the Scotian Shelf, in the Gulf of Maine and the MAB (Chapman and Beardsley, 1989). The shelf water is separated from the warmer, saltier slope water by the shelf-slope front, which lies near the shelf break in the MAB (Linder and Gawarkiewicz, 1998). The Gulf Stream directly influences the Cape Hatteras region, where it branches eastward off the shelf break delineating the slope water to the east. The slope water is a mixture of waters of subtropical and subpolar origin. Shelf-slope exchange of water and constituents occurs through a variety of mechanisms, including frontal instabilities at the shelf-slope front (Houghton et al., 1994; Lozier and Gawarkiewicz, 2001), interactions of Gulf Stream rings with the shelf (Churchill et al., 1986; Ryan et al., 2001), and on-shore flow of dense, nutrient-rich slope water in deep cross-shelf channels like the Northeast Channel (Ramp et al., 1985).

There is significant decadal-scale variability in the volume transport of the Labrador Current, which affects the shelf and

slope ecosystems; this variability is related to large-scale atmospheric variability over the North Atlantic. During periods of enhanced Labrador Current transport, shelf waters in the GOM and MAB freshen (Mountain, 2003; Smith et al., 2001) and an expansion of the volume of slope waters and a southward shift in the position of the Gulf Stream (Rossby and Benway, 2000) have been reported. Associated with these circulation changes are changes in vertical stratification and nutrient supply to the shelf ecosystem (Thomas et al., 2003; Greene and Pershing, 2007) as well as changes in primary production in the slope region (Schollaert et al., 2004). The dominant mode of biological variability is the spring bloom, which starts in the nearshore areas of the MAB in January and progresses offshore and northward following the progressing thermal stratification of the water column. The spring bloom reaches the deep regions of the GOM and the Scotian Shelf by April and is dominated by diatoms larger than $20\ \mu\text{m}$ (Longhurst, 2007; O'Reilly and Zetlin, 1998). The spring bloom terminates with an abrupt drop in chlorophyll concentrations and a shift towards small phytoplankton. Chlorophyll concentrations remain low during the summer months, but increase in September and October as increased wind mixing, convection and decreased solar heating erode density stratification thus inducing the fall bloom (Longhurst, 2007).

2.2 The circulation model

We use the Regional Ocean Modeling System (ROMS, www.myroms.org), a three-dimensional, free-surface, terrain-following numerical circulation model that solves the Navier-Stokes equations using the hydrostatic and Boussinesq approximations (Haidvogel et al., 2008). ROMS is implemented at $\sim 8\ \text{km}$ horizontal resolution for a domain that stretches from the Scotian Shelf to south of Cape Hatteras (horizontal domain dimensions are $1800\ \text{km}$ by $360\ \text{km}$) and includes the slope waters and parts of the Gulf Stream and the Sargasso Sea. The most important features of our physical model implementation are summarized below; details are described in He and Chen (submitted). The model domain has 120 by 160 grid cells in the horizontal direction, ranging in resolution from 6 to 10 km, and 36 terrain-following vertical layers with higher resolution near the surface and bottom.

Physical open-boundary conditions for our domain are specified from the HYCOM NCODA hindcast system, a global $1/12^\circ$ model that assimilates satellite sea surface temperature and surface height and profiling hydrographic data from Argo drifters (Chassignet et al., 2007). Specifically, temperature, salinity and baroclinic velocity from the HYCOM NCODA are prescribed at daily intervals at our open boundary conditions. We use the method of Flather (1976) to specify free-surface and depth-averaged velocity boundary conditions with external values defined by HYCOM NCODA plus M2 tidal harmonics from an ADCIRC simulation of the western Atlantic (Luettich et al., 1992). Our

model simulations were initialized with HYCOM NCODA fields of temperature, salinity, velocity and sea level on 1 November 2003 and ran until 31 December 2007.

Surface forcing comes from the NOAA NCEP North American Reanalysis (NARR) data set and includes cloud coverage, air pressure, freshwater flux, shortwave and longwave radiation, air temperature and surface winds at three hour intervals. Turbulent vertical mixing follows Mellor and Yamada (1982). Fresh water input from rivers is included using a monthly climatology based on data from the US Geological Survey.

2.3 The biological model

Our biological model is only slightly modified from Lima and Doney (2004) and describes the dynamics of diatoms, picoplankton, zooplankton, large detritus, small detritus, and the inorganic nutrients nitrate and ammonium. Here we describe only the salient features of the model and our modifications. We refer the reader to Lima and Doney (2004) for the detailed equations. In the model the nitrogen and carbon content of phytoplankton, zooplankton and detritus is tracked; in addition, the chlorophyll content of both phytoplankton groups is variable. Phytoplankton grow as a function of light, inorganic nitrogen concentrations and temperature. The intracellular ratios of nitrogen, carbon and chlorophyll for phytoplankton are based on Geider et al. (1998) but modified by Lima and Doney (2004) to include both inorganic nitrogen sources, nitrate and ammonium. In addition to being grazed, small and large phytoplankton is lost to small and large detritus, respectively, through a combination of a linear and a quadratic loss term. A combination of linear and quadratic terms also describes the losses from zooplankton to detritus. Decomposition of detritus to ammonium by heterotrophic bacteria is parameterized using a linear remineralization rate. Large detritus sinks at a rate of $10\ \text{m d}^{-1}$, while small detritus and all other biological variables do not sink. We use the same parameter values as Lima and Doney (2004) with the exception of those listed in Table 1. We modified the parameters given in Table 1 from their original values because these modifications improved model-data agreement. For the purpose of parameter tuning we implemented 1-dimensional models for several locations within our model domain, which allowed us to perform many simulations with different parameter sets.

In Lima and Doney's (2004) model, diatoms are distinguished from picoplankton mainly by their requirement for silicate, their higher half-saturation constant for nitrate uptake and by being grazed with lower preference than picoplankton. For the sake of simplicity, our implementation does not include silicate and hence we interpret the diatom group more generally as large phytoplankton. The term picoplankton is usually used to refer to phytoplankton smaller than $2\ \mu\text{m}$ (Sieburth et al., 1978), a practical division based on the pore size of commonly used filters that separates the

Table 1. Parameter values of the biological model. Values in brackets are those of Lima and Doney (2004).

Parameter	Value	Units
Initial slope of P-I curve	0.6 (0.2)	$\text{mmol C (mg chl)}^{-1} \text{d}^{-1} \text{m}^2 \text{W}^{-1}$
Max grazing rate on large phytoplankton	2.0 (2.75)	d^{-1}
Small phytoplankton aggregation rate	0.1 (0.3)	$(\text{mmol N m}^{-3} \text{d})^{-1}$
Large phytoplankton aggregation rate	0.1 (0.3)	$(\text{mmol N m}^{-3} \text{d})^{-1}$
Zooplankton quadratic mortality rate	0.5 (0.25)	$(\text{mmol N m}^{-3} \text{d})^{-1}$
Small detritus remineralization rate	0.35 (0.2)	d^{-1}
Large detritus remineralization rate	0.35 (0.2)	d^{-1}
Large detritus sinking rate	10 (25)	m d^{-1}
Zooplankton egestion allocation factor	0.70 (0.5)	

mostly prokaryotic component of the plankton from the rest. The different maximum grazing rates for the two phytoplankton groups suggest an ecological division based on the size and life histories of their consumers; large phytoplankton are large enough to be subject to filtering appendages of zooplankton such as copepods, while smaller phytoplankton are consumed by unicellular grazers (e.g., Kiorboe, 1993; Strom et al., 2000). For simplicity and model stability, the diverse zooplankton population is parameterized in a single zooplankton compartment (Armstrong, 1999) with an s-shaped grazing function (Holling-type III) and a quadratic mortality term.

We initialized our biological variables using the nitrate, ammonium, phytoplankton, zooplankton and detritus fields for 1 November 2003 from the simulation described in Fennel et al. (2008), dividing the phytoplankton concentration equally into the two phytoplankton groups. Concentrations of the biological variables along the open boundaries were also taken from the simulation of Fennel et al. (2008), which uses a larger domain with similar horizontal resolution.

2.4 Observational data sets

We compare satellite-derived sea-surface chlorophyll concentrations and temperatures with the corresponding model-predicted fields. SeaWiFS ocean color and Advanced Very High Resolution Radiometer (AVHRR) sea-surface temperature (SST) come at a spatial resolution of approximately 1 km and 4 km, respectively, which is higher than the spatial resolution of our model. The temporal resolution is less optimal. While the return period of the satellites is on the order of days, cloud cover usually obstructs at least part of the study region in any given scene and we found that at minimum, a composite of all scenes within a month is required for a gapless view. Hence, we decided to base our model validation exercise on monthly average fields of the surface properties. We obtained remotely sensed chlorophyll concentrations from the SeaWiFS sensor from NASA's Ocean Color Website (<http://oceancolor.gsfc.nasa.gov/>) as Level-2

swath products, and projected each pass onto the model grid by averaging all satellite observations within a 10-km radius from each grid point using a Gaussian weighting function. We then produced monthly average fields of the gridded chlorophyll observations. For SST we used fields from the AVHRR, more specifically the monthly mean fields from the AVHRR reanalysis Version 5.0 available from the National Oceanographic Data Center (Kilpatrick et al., 2001, <http://www.nodc.noaa.gov/sog/pathfinder4km>).

2.5 Subareas

For the purpose of model-data comparison we divided our model domain into a number of subareas that represent meaningful biogeographic units. These divisions are based on the work of O'Reilly and Zetlin (1998), who analyzed more than 50 000 size-fractionated samples of chlorophyll taken in the area between 1977 and 1987. Using a statistical clustering algorithm they divided the MAB and GOM into 25 subareas reflecting regional differences in water depth and in magnitude, timing and size distribution of chlorophyll (see also Fig. 1 in Hofmann et al., 2008). We arrived at a more manageable number of 12 subareas for our model domain and eliminated areas we deemed too small for a meaningful comparison by combining several of the original subareas (Fig. 1). We only report detailed simulation results for five of these: MAB north, MABGOM shelf break, MAB-GOM slope, Gulf Stream and Sargasso Sea which represent an ecological gradient from the shelf sea to the open ocean.

2.6 Statistics of model/observation fit

We calculated three statistics frequently used to quantify model-data misfit: model bias, root mean square error (RMSE) and model efficiency or skill (ME). Here we introduce the equations in their most general form; complete details are given in the Appendix. The model bias measures the mean deviation between model-predicted values (M) and

the observations (O):

$$\text{Bias} = \frac{1}{n} \sum (M - O), \quad (1)$$

where n represents the number of model/observation data pairs. A positive or negative bias reflects a general over- or underestimation of observations by the model, respectively. Note that negative and positive deviations between model and observations will tend to cancel each other; the main purpose of bias is to indicate a persistent error in magnitude of the modeled variable.

RMSE measures the deviation between model and observations in a least-squares sense:

$$\text{RMSE} = \sqrt{\frac{\sum (O - M)^2}{n}}. \quad (2)$$

Taking the square of the deviations ensures that negative as well as positive contributions are added, while the square-root operation restores the units to that of the original variable. The smaller the absolute values of bias and RMSE the better the agreement between model and observations.

ME relates the deviations between model and observations to the variability in the observations:

$$\text{ME} = 1 - \frac{\sum (O - M)^2}{\sum (O - \bar{O})^2}. \quad (3)$$

\bar{O} denotes an average of the observations, i.e. a climatology. ME is always less than or equal to one; ME of one corresponds to a perfect match between model and observations. ME larger than zero indicates that the model is a better predictor than the observational climatology, while ME less than zero implies that the climatology is a better predictor than the model.

In the above equations summations can represent spatial or temporal averaging or both. We calculate bias, RMSE and ME for all three variants, as each can give different insights. Summing over space and time, thus collapsing spatial and temporal variability into scalar values of bias, RMSE and ME (see Eqs. A1–A3) is useful for evaluating first order correspondence between model and observations. We calculate the statistics for the whole model domain as well as for the individual subregions shown in Fig. 1.

Secondly, we derive spatially explicit error statistics by applying only temporal averaging (Eqs. A4–A6). This produces spatial fields for the statistics that quantify the temporal correspondence between model and observations at the spatial resolution of the model. The resulting error maps are useful in pinpointing regions where the model does well and where it needs improvement.

The third option produces time-series of the error statistics by applying spatial averaging only (see Eqs. A7–A9). As for option one, we use the whole domain as well as subareas. The resulting time-series allow evaluation of when the simulation performs well, whether the simulation deteriorates or

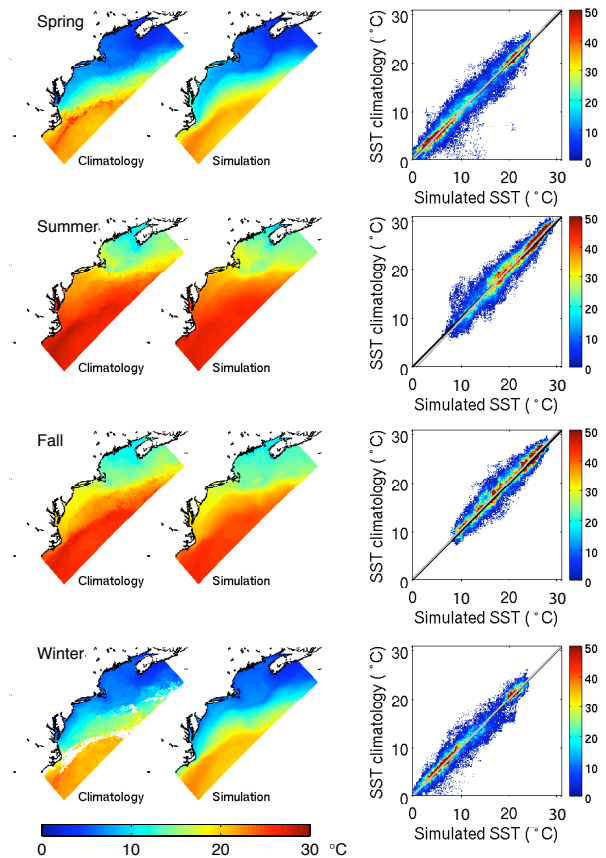


Fig. 2. Mean SST for spring (March–April–May), summer (June–July–August), fall (September–October–November) and winter (December–January–February) for the years 2004 to 2007 from AVHRR observations and the ROMS simulation. The scatterplot shows observations versus model for every pixel with shading according to the density of points.

improves over time, and putting errors into the context of seasonal plankton evolution.

3 Results

3.1 Qualitative assessment of model results

We first describe the agreement between simulated and observed SST and surface chlorophyll concentrations in qualitative terms and evaluate time-series of monthly means of those variables averaged within five subareas that cover a representative range of oceanographic conditions along a gradient from inshore to the oligotrophic ocean.

Simulated, seasonally averaged sea-surface temperatures are shown in comparison with satellite observations in Fig. 2 and agree very well, with only a slight underestimation of the maximum temperature of the Gulf Stream in summer and slightly cooler than observed temperatures in the northern MAB off Long Island in the fall. Monthly averaged SST

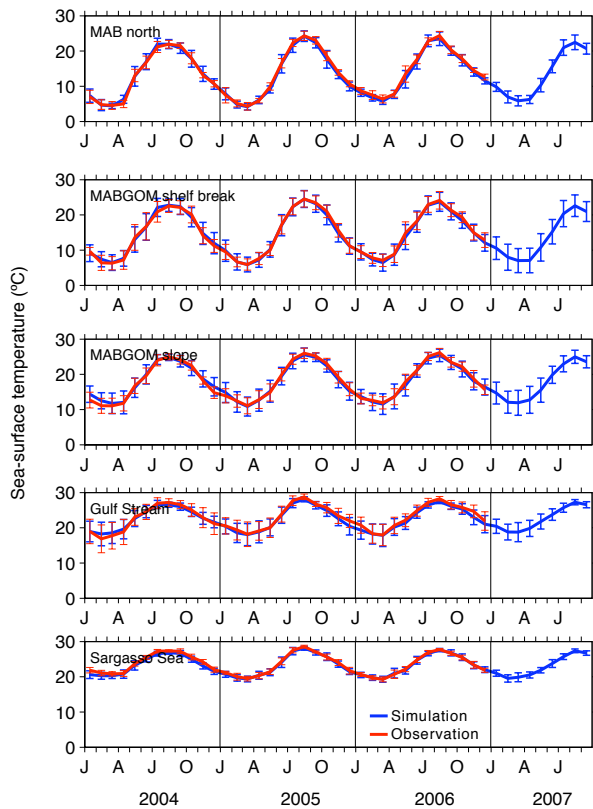


Fig. 3. Monthly SST (solid line) with one standard deviation (error bar) from AVHRR (red) and model simulation (blue) in five subareas.

in the five subareas, MAB north, shelf break, slope, Gulf Stream and Sargasso Sea (Fig. 3), agree very well with the observed values.

Simulated, seasonally averaged surface chlorophyll is shown in comparison with observed fields in Fig. 4. A boundary artifact is noticeable at the southwestern boundary where the Gulf Stream enters the model domain and simulated chlorophyll concentrations are lower than observed. The model captures the inshore-offshore gradient in surface chlorophyll concentrations in the MAB and high chlorophyll concentrations in the tidally mixed areas on Georges Bank and Nantucket Shoals, especially in summer and fall. The model fields deviate from the observations in underestimating the remotely sensed chlorophyll concentrations in the nearshore regions of the MAB and the GOM. However, it should be noted that algorithms deriving chlorophyll from ocean color are known to be problematic in optically complex near-shore waters and may overestimate true chlorophyll concentrations by as much as 100% (O'Reilly et al., 2000; Yoder et al., 2002).

In order to analyze deviations between model-predicted and observed chlorophyll concentrations more closely we discuss regional comparisons for the five subareas (Fig. 5,

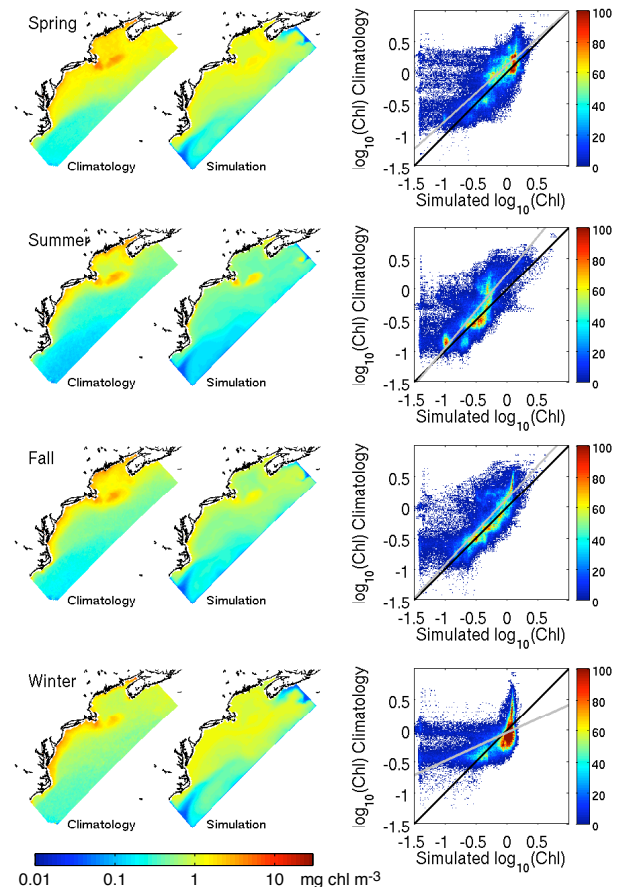


Fig. 4. Mean surface chlorophyll concentration for spring (March–April–May), summer (June–July–August), fall (September–October–November) and winter (December–January–February) for the years 2004 to 2007 from SeaWiFS observations and the ROMS simulation. The scatterplot shows observations versus model for every pixel with shading according to the density of points.

comparisons for the other subareas is given in the Online Supplement, Fig. S1, see <http://www.biogeosciences.net/6/1961/2009/bg-6-1961-2009-supplement.pdf>). It is characteristic of the MAB that chlorophyll concentrations remain high throughout the winter months until March and that the end of the spring bloom is characterized by an abrupt decrease in chlorophyll rather than a distinct peak (O'Reilly and Zetlin, 1998). Chlorophyll concentrations are at a minimum during the summer months and increase during September and October. The MAB is represented by the subregion MAB north in Fig. 5. The simulation captures the annual cycle, with low chlorophyll concentrations in summer, high chlorophyll concentrations in winter and a slight increase from winter levels in March–April. Peak chlorophyll concentrations during the fall blooms and during the winter of 2005–2006 are underestimated by the model, although the model mean is within one standard deviation of the observations.

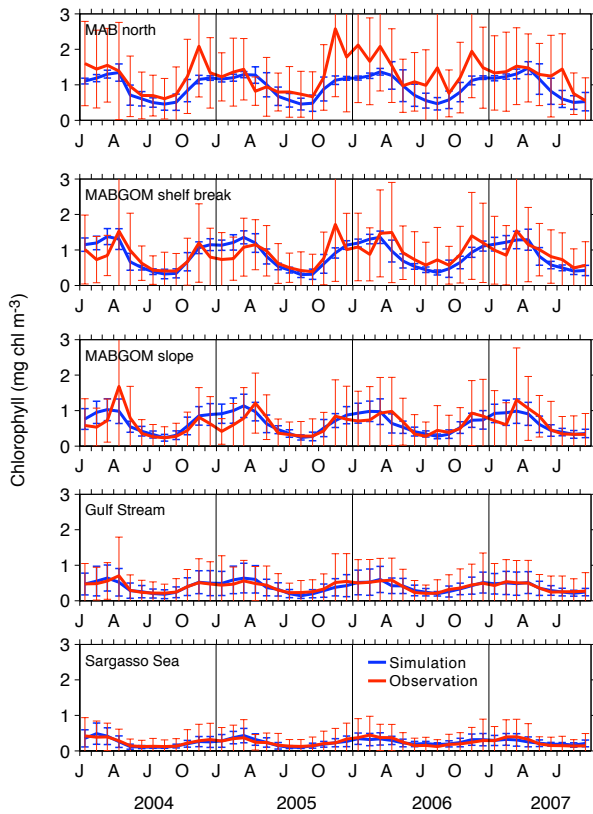


Fig. 5. Monthly mean chlorophyll concentrations (solid lines) with one standard deviation (errorbar) from SeaWiFS (red) and model simulation (blue) in five subareas.

The seaward boundary of the subarea MAB north is marked by the shelf-break front, which surfaces approximately over the 200 m isobath and is included in the narrow subarea MABGOM shelf break. In this area, observed chlorophyll concentrations are generally lower than in MAB north and dip more consistently in winter, resulting in more distinct spring and fall blooms. The simulated surface chlorophyll concentration in MABGOM shelf break (Fig. 5) shows close agreement with the observed timing and magnitude of summertime minima, fall increases and the decline of chlorophyll in April.

Observations for the slope region, delineated by the shelf break front to the west and the Gulf Stream to the east, show distinct fall and spring blooms in November and April. These observations are mirrored well by the simulation (Fig. 5), although simulated chlorophyll concentrations don't show a pronounced dip in winter. Observed chlorophyll levels drop markedly in the Gulf Stream and Sargasso Sea. In both areas a broad winter-spring bloom begins in November with monthly mean surface chlorophyll levels of 0.5 mg m^{-3} in the Gulf Stream and 0.35 mg m^{-3} in the Sargasso Sea. Spring maxima of monthly mean surface chlorophyll are 0.6 mg m^{-3} in the Gulf Stream and 0.4 mg m^{-3} in the Sar-

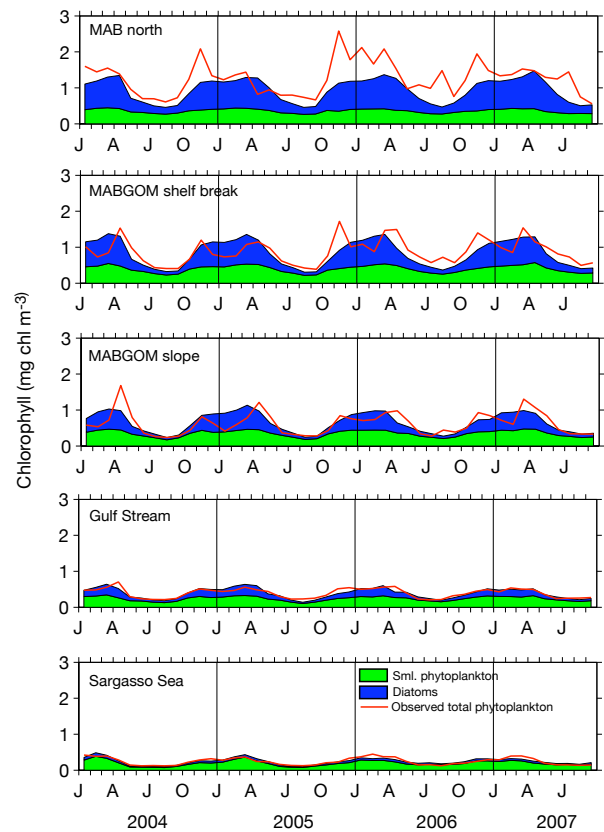


Fig. 6. Monthly mean chlorophyll concentrations of picoplankton (green) and large phytoplankton (blue) compared to SeaWiFS total chlorophyll (red) in five subareas.

gasso Sea between February and March. The simulated chlorophyll time-series in both areas matches the observations very well (Fig. 5).

The spatial gradient in maximum chlorophyll concentrations obvious in Fig. 5 coincides with a shift in dominance of larger phytoplankton on the shelf to picoplankton dominating in the oligotrophic ocean (Fig. 6). In the MAB, shelf break and slope regions the picoplankton chlorophyll concentrations remain relatively constant in time and vary between 0.3 and 0.5 mg m^{-3} only. Increases in total chlorophyll above 0.5 mg m^{-3} are due to larger phytoplankton dominating during fall, winter and spring. This agrees with the succession of phytoplankton size classes observed on the continental shelf by O'Reilly and Zetlin (1998), who report that the end of the spring bloom is marked by a decline of the chlorophyll contribution by large phytoplankton and increasing importance of small phytoplankton. In contrast to shelf waters, picoplankton dominates throughout the year with a slight increase in the contribution of large phytoplankton during the winter-spring bloom in the model for the Gulf Stream and Sargasso Sea regions. (Fig. 6). This dominance of picoplankton and a diatom bloom following winter mixing is a classic

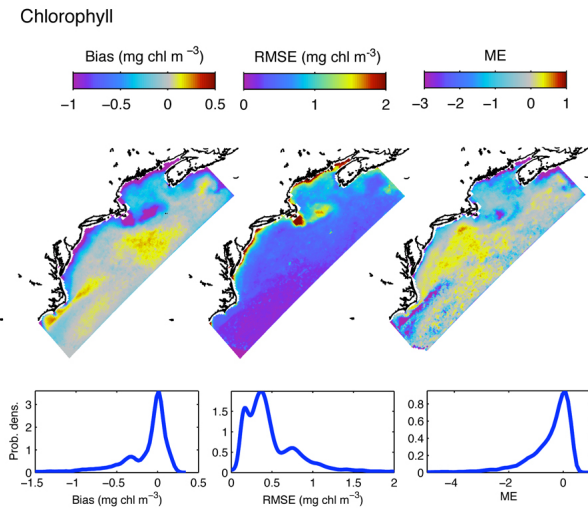


Fig. 7. Maps of bias, RMSE and ME and their probability distributions calculated using monthly mean SeaWiFS chlorophyll from 2004 and 2007.

feature of the open ocean as observed at the Bermuda Atlantic Time Series site (DuRand et al., 2001; Li et al., 2001).

In summary, monthly subarea averages of simulated surface chlorophyll concentrations lie well within one standard deviation of corresponding SeaWiFS observations (with rare exceptions), and timing and magnitude of the monthly surface chlorophyll concentration are generally well matched. Shortcomings of the simulation are an underestimation of chlorophyll levels in MAB north in winter 2005/2006 and a slight overestimation of winter chlorophyll in the shelf break and slope regions.

3.2 Statistical model/data comparisons

Total bias quantifies the average deviation of all monthly model/observation pairs in one scalar (Eq. A1) and is given in Table 2 for the whole domain and for the five subareas. The model has a negative bias when considering the whole domain with slight positive biases in the slope, Gulf Stream and Sargasso Sea, a slight negative bias in the shelf break region and a more pronounced negative bias of $0.3 \text{ mg chl m}^{-3}$ in the MAB north. SST has a slight negative bias, with values between -0.3 and -0.07°C in all subareas including the domain as a whole (Table 3).

The RMSE for surface chlorophyll (Eq. A2) is $\leq 0.64 \text{ mg chl m}^{-3}$ for and, as is the case for model bias, the largest RMSE occurs in MAB north ($0.64 \text{ mg chl m}^{-3}$). RMSE is second largest for the shelf break with $0.46 \text{ mg chl m}^{-3}$ which is an area with small bias. This difference between bias and RMSE is in large part due to the fact that the model overestimates observations at some times, and underestimates observations at other times (see Fig. 5), although spatial variability due to the meandering

shelf break front and mesoscale eddies likely contributes as well. Similarly, the RMSE for SST is relatively large, ranging from 0.7 to 1.4°C (Table 3) despite the excellent agreement evident in Figs. 2 and 3. We attribute this to mismatches due to mesoscale variability that don't lead to systematic over- or underestimation (as indicated by the small biases) but are penalized in the calculation of RMSE.

ME (Eq. A3), evaluates the predictive skill of the model relative to the predictive skill of a climatology; positive values of ME indicate that the model solution represents an improvement over climatology (with ME of one indicating a perfect prediction), while negative values indicate that the climatology is a better predictor of the observations. ME for chlorophyll (Table 2) is positive for all subareas except for the Sargasso Sea. ME for MAB north and the Gulf Stream are very similar, despite the obvious difference in how the model captures variability in observations in these two areas. The low value of ME in the Gulf Stream and the Sargasso Sea is due to the fact that climatology is an excellent predictor of monthly mean chlorophyll in these areas of relatively small seasonality, thus producing a small denominator in (Eq. 3). In MAB north, on the other hand, the numerator is large due to model error, producing a small ME. ME for SST has values near one, indicating that the model predicts observations much better than climatology (Table 3).

3.2.1 Spatial distribution of fit statistics

The temporally integrated, but spatially resolved bias Eq. (A4) is shown in Fig. 7. For surface chlorophyll the slope, Gulf Stream and Sargasso Sea tend to have a small positive bias, i.e. the model overestimates observations slightly. An intrusion of negative bias from the southern boundary hints at lower-than-realistic boundary conditions for chlorophyll in this area, which is advected into the study region by the Gulf Stream. The probability distribution of biases shows that the most likely value is zero, with a second smaller peak at $-0.4 \text{ mg chl m}^{-3}$. The negative tail of the distribution is larger, indicating underestimation is more severe than overestimation; the map shows that this is typically the case in the most nearshore regions. The spatially explicit bias of SST (Fig. 8) shows that the model underestimates temperatures in deep water along the southern boundary, while the simulated shelf and slope waters tend to be warmer than observed. The most likely value of temperature bias is -0.2°C .

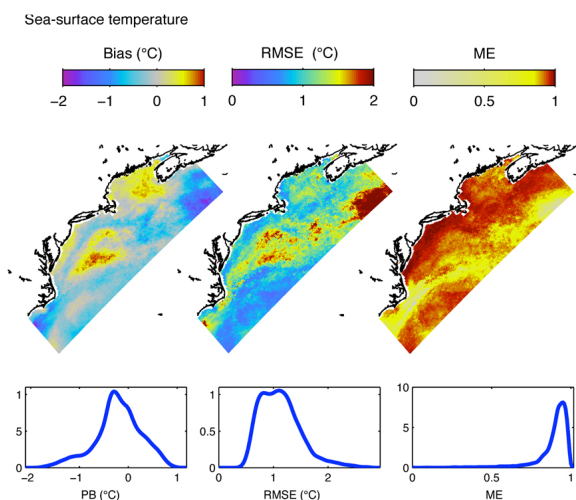
The spatial field of RMSE (Eq. A5) for chlorophyll shows little variability compared to the bias (Fig. 7). The most likely error is about $0.4 \text{ mg chl m}^{-3}$ and almost all grid cells have an RMSE smaller than $0.5 \text{ mg chl m}^{-3}$. The highest errors are found in a narrow band along the coast in the entire study region, on Georges Bank and on the Nantucket Shoals. The error in SST is small in most coastal regions, the Gulf Stream and Sargasso Sea (Fig. 8). The largest temperature errors are found in slope waters throughout the study region.

Table 2. Bias, RMSE and ME for surface chlorophyll.

Domain	Bias (mg chl m^{-3})	RMSE (mg chl m^{-3})	ME (dimensionless)
Whole	-0.18	0.64	0.28
MAB north	-0.32	0.64	0.03
MABGOM shelf break	-0.03	0.46	0.07
MABGOM slope	0.04	0.38	0.16
Gulf Stream	0.01	0.26	0.03
Sargasso Sea	0.01	0.17	-0.05

Table 3. Bias, RMSE and ME for surface temperatures.

Domain	Bias ($^{\circ}\text{C}$)	RMSE ($^{\circ}\text{C}$)	ME (dimensionless)
Whole	-0.26	1.19	0.97
MAB north	-0.25	0.99	0.98
MABGOM shelf break	-0.10	1.28	0.95
MABGOM slope	-0.07	1.38	0.91
Gulf Stream	-0.19	1.02	0.91
Sargasso Sea	-0.29	0.73	0.93

**Fig. 8.** Maps of bias, RMSE and ME and their probability distributions calculated using monthly mean SST (from AVHRR) from 2004 to 2007.

The spatial distribution of ME (Eq. A6) for surface chlorophyll shows that the model predicts the observations better than climatology in a large fraction of the study region (Fig. 7), including the MAB and slope waters. The most likely value for ME is just below zero. Negative values are concentrated in parts of Georges Bank, the Scotian Shelf and the area where the Gulf Stream enters the study region. For SST ME is greater than zero throughout, with lowest values found in the Gulf Stream and the Scotian slope.

3.2.2 Temporal evolution of statistics

The time-series of bias, RMSE and ME (Eqs. A7 to A9) yield information on the temporal patterns of agreement between observations and the simulation for each subarea. The time-series of all statistics and in all subareas for chlorophyll have in common that their values change little between June and October, when biomass is at the summertime low (Fig. 9). Naturally, simulating the timing and magnitude of blooms is the most challenging aspect for ecosystem models, thus most of the excursions in the bias, RMSE and ME occur between fall and spring. Most importantly, however, there is no systematic change in any of the statistics over the course of the four-year simulation, indicating that the model is not drifting from one state to another or adjusting as a consequence of initial conditions. The change of sign of the bias in the slope and shelf break shows the seasonal nature of under- and over-estimation. It is this change of sign that produces a low bias in comparison to RMSE (Table 2). For the MAB, the bias is mostly negative and most strongly so at the time of the fall bloom, illustrating the need for improving the model in this area and time.

The time-series of SST bias (Fig. 10) show that subarea-average deviations from observations are usually between -1 and 1°C in any given month. It is notable that the bias tends to be mostly positive in 2004 and mostly negative by 2006 in all subareas except the Sargasso Sea; this hints at a slight drift in temperatures over the period of simulation. The RMSE time-series for chlorophyll are below 1 mg chl m^{-3} for all regions except for the winter 2005–2006

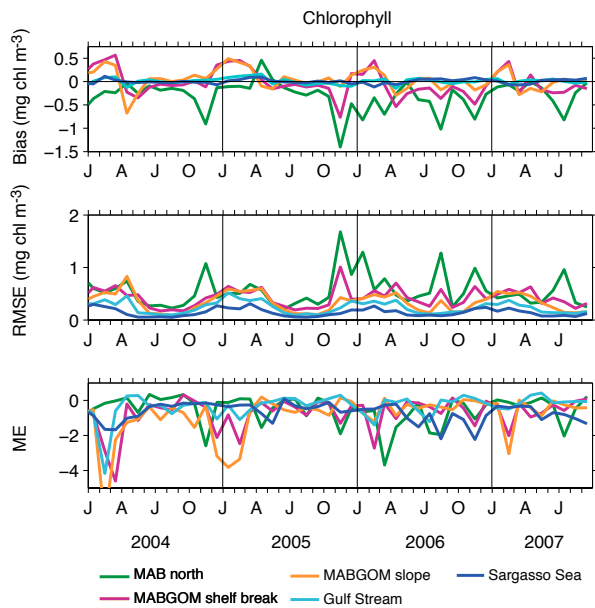


Fig. 9. Time-series of bias, RMSE and ME calculated using monthly mean SeaWiFS chlorophyll.

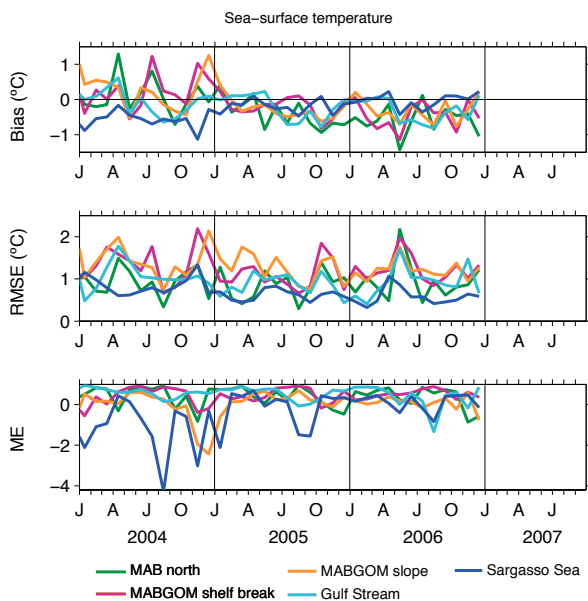


Fig. 10. Time-series of bias, RMSE and ME calculated from monthly mean SST.

bloom in the MAB (Fig. 9). The RMSE of SST appears to vary unsystematically around $\sim 1^\circ\text{C}$ (Fig. 10).

The time-series values of ME (Eq. A9) are near zero most of the time for chlorophyll and typically between zero and one for SST. Negative ME for SST dominates in the Sargasso Sea, which may be surprising considering the good fit between model-predicted and observed temperature in Fig. 3,

but can be explained as follows. The time-series version of ME measures the match of spatial features between model and observations at a specific point in time. Considering that the Sargasso Sea is characterized by mesoscale eddies, which we don't expect to be matched exactly by our non-assimilative model, negative values in ME can result.

4 Discussion

To the best of our knowledge, this is the first application of a high-resolution biological model with two functional phytoplankton groups for northeast North American shelf seas. The model captures well the annual cycle of surface chlorophyll across a physical and ecological gradient. Occasional underestimation of chlorophyll occurs in the MAB (during fall blooms and in winter 2005–2006), but considering all five subareas, from MAB to the Sargasso Sea together, chlorophyll dynamics are matched very well. The general decrease in maximum chlorophyll from the shelf to the oligotrophic ocean coincides with a shift in the dominant phytoplankton functional group from large phytoplankton dominating on the shelf to picoplankton dominating in the Sargasso Sea. This shift results from a change in the dominant nutrient cycling pathway, with large phytoplankton subject to vertical sinking dominating on the shelf and in the slope, and picoplankton fueled by recycled production dominating in the Sargasso Sea. This shift is in line with a well-established paradigm in plankton ecology, reviewed and extended by Cullen et al. (2002), which states that small phytoplankton with high ratios of surface area to volume and, hence, low sinking rates and an advantage in competition for nutrients dominate in stable, nutrient depleted waters like the Sargasso Sea. The microbial loop, which relies on recycled production of small phytoplankton is present everywhere in the ocean, but dominates biomass only in stable, low-nutrient regimes like the Sargasso Sea. The relatively nutrient-rich, coastal waters are characterized by food webs with larger cells that transfer a greater proportion of primary productivity to higher trophic levels and allow for a greater proportion of primary productivity to be exported vertically.

We found that the interpretation of RMSE is problematic, as slight mismatches in timing of blooms or spatial patterns between observations and model are heavily penalized by this measure. For example, a slight shift in timing between the simulated and observed spring bloom would lead to a higher RMSE than if the model completely missed the bloom. Also, if model and observations both displayed mesoscale eddies (which would appear as anomalies in SST and sea surface chlorophyll) but the individual features were offset in space and/or time, RMSE would penalize this more strongly than a situation where the simulation did not predict any of the observed eddies at all. Since we don't expect mesoscale features to be matched exactly in space and time in a simulation that does not assimilate altimeter and other data,

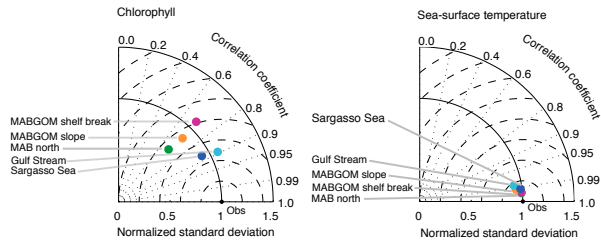


Fig. 11. Taylor diagrams illustrating the match of simulated time-series of subarea average chlorophyll and SST with observations. The full time-series from 2004 to 2007 as shown in Figs. 5 and 3 were used to calculate the statistics.

we interpret the RMSE with caution, especially in areas characterized by mesoscale features (e.g. shelf break front, Gulf Stream, Sargasso Sea). For example, SST and chlorophyll RMSEs for the shelf break and slope sea are relatively high despite small biases, positive ME (Tables 2 and 3) and apparent agreement between model and observations (Figs. 3 and 5). Hence, we put more value in small biases and positive ME.

Spatial maps of the error statistics are useful in illustrating where in space the model performs well and where improvements are most needed. The map of chlorophyll ME indicates that simulated chlorophyll predicts the observations better than climatology in much of the model domain, including the MAB and slope sea. We find that the time-series of bias are useful as well, as they illustrate whether, and if so where, the model drifts in time. We note a slight drift in SST (Fig. 9). The time-series of RMSE and ME are less instructive, as they measure the agreement of spatial patterns at given points in time. As discussed above we don't necessarily expect spatial patterns within relatively homogeneous subregions to agree well in a non-assimilative model.

Correlation, RMSE and the ratio of standard deviations of simulated versus observed variables can be displayed conveniently in a Taylor diagram (Taylor, 2001). In this diagram (Fig. 11), the radial distance of the points from the origin is proportional to the ratio of standard deviations of the model-predicted versus the observed variable. The correlation coefficient between the model-predicted and observed quantity is given along the arch of the diagram and the distance of each point from the reference point (value 1 on the x-axis) indicates the centered RMSE. The RMSE is centered by subtracting the mean from the respective time series, thus removing any existing bias.

The Taylor diagram for the area-averaged time-series of chlorophyll (Fig. 11) shows that simulated chlorophyll concentrations in the Gulf Stream and the Sargasso Sea have the highest correlation coefficient, and thus fit the observations best in terms of timing and magnitude of the evolution of monthly chlorophyll concentrations between 2004 and 2007. MAB north, MABGOM slope and MABGOM shelf break

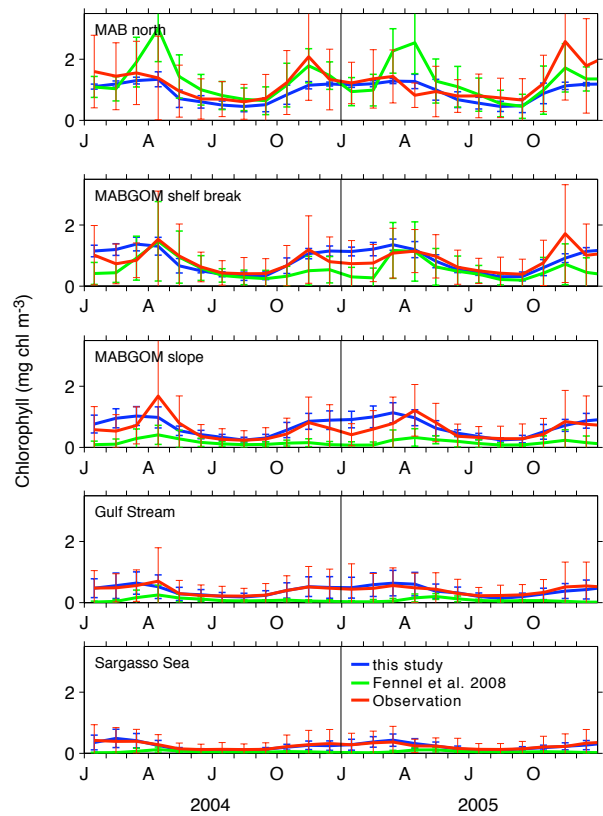


Fig. 12. Time-series of sea surface chlorophyll from Fennel et al. (2008) in contrast with the biological model used here.

have similar correlation coefficients (~ 0.7), which are lower than those of the Sargasso Sea and Gulf Stream (~ 0.9). Of all subareas, MAB north has the greatest distance to the arch of unit standard deviation, i.e. the model underestimates variability of chlorophyll in MAB north (by about 30%), presumably because the fall-winter bloom in 2005–2006 is underestimated by the model.

The model matches observations of SST with a correlation coefficient above 0.99 in all regions indicating excellent timing of the monthly variability (Fig. 11). The standard deviations of the model-predicted and observed time-series are also very similar, with only a slight underestimation in the slope and Gulf Stream.

The statistical summary measures are useful in contrasting two different models or different realizations of the same model. We compare the simulation described here to the simulation in Fennel et al. (2008), which uses a model with only one phytoplankton group, a domain that is larger than the model domain here, but with similar horizontal resolution, and which was run for the years 2004 and 2005 (Figs. 12 to 14). In the MAB, the model with only one phytoplankton group captures fall blooms better than the model described in this study, but overestimates chlorophyll in the spring. In the other four domains (i.e. the shelf break, slope sea, Gulf

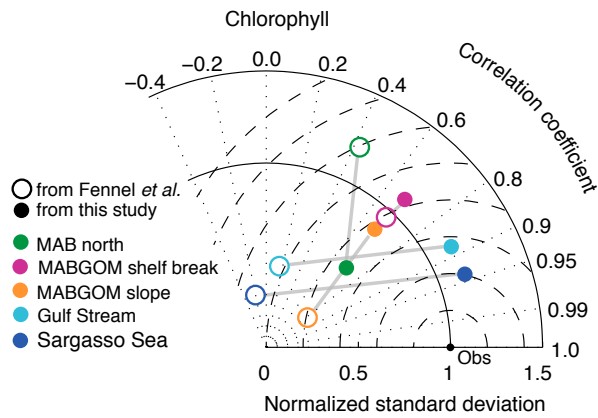


Fig. 13. Taylor diagram contrasting the surface chlorophyll statistics of the model described here and that of Fennel et al. (2008). Only the years 2004 and 2005 are included in this comparison.

Stream and Sargasso Sea) the model described here captures the observed chlorophyll dynamics better. We attribute this directly to the presence of two phytoplankton groups playing different ecological roles. The Taylor diagram (Fig. 13) shows that the fit statistics of the simulated time-series improve in most regions for the model described here compared to the simulation in Fennel et al. (2008).

The values for bias, RMSE, and ME integrated over time and space within subareas and for the entire study region (Eqs. A1, and A3) are shown in Fig. 14 (the statistics for all subareas are given in the Online Supplement in Fig. S2). Again, the model described here, based on the formulation by Lima and Doney (2004), shows generally smaller bias, less error and higher efficiency.

5 Conclusions

The implementation of the size-structured model by Lima and Doney (2004) improved the agreement between model-predicted fields of surface chlorophyll with chlorophyll estimates from the SeaWiFS satellite compared to the model by Fennel et al. (2006, 2008) across an ecological gradient from the productive MAB to the oligotrophic Sargasso Sea. We attribute the improved fit to the model's ability to shift between a foodweb dominated by large phytoplankton that tend to sink and a foodweb dominated by picoplankton and recycled production. The increase in model complexity comes at higher computational cost, however. The single phytoplankton model described in Fennel et al. (2006) has 6 state variables, while Lima and Doney's (2004) model has 17 state variables (in our implementation reduced to 14 by excluding silicate). In Lima and Doney's (2004) simulation the model predicted too small chlorophyll concentrations in the north-east North American shelf region ($<1 \text{ mg m}^{-3}$ in spring and

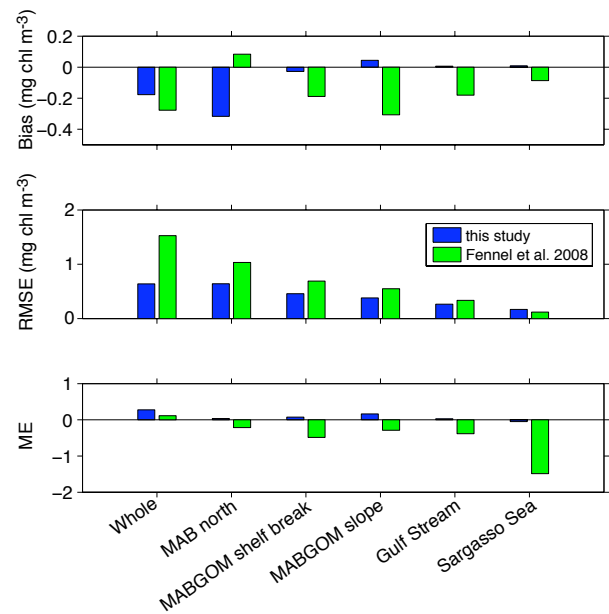


Fig. 14. RMSE, and ME of surface chlorophyll between Fennel et al. (2008) and the model used here. These statistics were computed using simulation from 2004 to 2005 and therefore deviate slightly from those reported in Table 2.

$<0.3 \text{ mg m}^{-3}$ in summer). We attribute the improved prediction in our implementation to higher model resolution.

When assessing model-data agreement using the RMSE, bias and model ME we found that RMSE has to be interpreted with caution. Integrated values and spatial maps of bias and ME were more instructive than RMSE. Time-series of model bias were useful as well.

Appendix A

Detailed equations of error statistics

We calculated three variants of bias, RMSE and ME using (1) spatial and temporal averaging, (2) temporal averaging only, and (3) spatial averaging only. In the first case all spatial and temporal variability is collapsed into scalar values of bias, RMSE and ME calculated as:

$$\text{Bias} = \frac{1}{T \times K} \sum^{k,t} (M(k, t) - O(k, t)) \quad (\text{A1})$$

$$\text{RMSE} = \sqrt{\frac{\sum^{k,t} (O(k, t) - M(k, t))^2}{T \times K}} \quad (\text{A2})$$

$$\text{ME} = 1 - \frac{\sum^{k,t} (O(k, t) - M(k, t))^2}{\sum^{k,t} (O(k, t) - \frac{1}{K} \sum^k \frac{1}{T} \sum^t O(k, t))^2} \quad (\text{A3})$$

Here $M(k, t)$ and $O(k, t)$ correspond to monthly means of model-predicted and observed fields, k and t are spatial and

temporal indices, and K and T are the number of horizontal grid cells and time slices, respectively.

In the temporal averaging case spatially explicit maps of the three error statistics result as follows:

$$\text{Bias}(k) = \frac{1}{T} \sum^t (M(k, t) - O(k, t)) \quad (\text{A4})$$

$$\text{RMSE}(k) = \sqrt{\frac{\sum^t (O(k, t) - M(k, t))^2}{T}} \quad (\text{A5})$$

$$\text{ME}(k) = 1 - \frac{\sum^t (O(k, t) - M(k, t))^2}{\sum^t (O(k, t) - \frac{1}{T} \sum^t O(k, t))^2} \quad (\text{A6})$$

In the spatial averaging case time-evolving error statistics result according to:

$$\text{Bias}(t) = \frac{1}{K} \sum^k (M(k, t) - O(k, t)) \quad (\text{A7})$$

$$\text{RMSE}(t) = \sqrt{\frac{\sum^k (O(k, t) - M(k, t))^2}{K}} \quad (\text{A8})$$

$$\text{ME}(t) = 1 - \frac{\sum^k (O(k, t) - M(k, t))^2}{\sum^k (O(k, t) - \frac{1}{K} \sum^k O(k, t))^2} \quad (\text{A9})$$

Acknowledgements. This work was supported by the ONR MURI grant N00014-06-1-0739 to KF and RH. RH also acknowledges support provided by NASA through grant NNX07AF62G. KF was also supported by NSERC and CFI. We are grateful for the constructive comments from Baris Salihoglu and two anonymous reviewers.

Edited by: C. Heinze

References

- Allen, J. I., Somerfield, P. J., and Gilbert, F. J.: Quantifying uncertainty in high-resolution coupled hydrodynamic-ecosystem models, *J. Marine Syst.*, 64, 3–14, 2007.
- Armstrong, R. A.: Stable model structures for representing biogeochemical diversity and size spectra in plankton communities, *J. Plankton Res.*, 21, 445–464, 1999.
- Cai, W., Wang, Z., and Wang, Y.: The role of marsh-dominated heterotrophic continental margins in transport of CO_2 between the atmosphere, the land-sea interface and the ocean, *Geophys. Res. Lett.*, 30, 1848, doi:10.1029/2003GL017633, 2003.
- Chapman, D. and Beardsley, R.: On the origin of shelf water in the Middle Atlantic Bight, *J. Phys. Oceanogr.*, 19, 384–391, 1989.
- Chassignet, E. P., Hurlburt, H. E., Smedstad, O. M., Halliwell, G. R., Hogan, P. J., Wallcraft, A. J., Baraille, R., and Bleck, R.: The HYCOM (HYbrid Coordinate Ocean Model) data assimilative system, *J. Marine Syst.*, 65, 60–83, 2007.
- Churchill, J., Cornillon, P., and Milkowski, G.: A cyclonic eddy and shelf-slope water exchange associated with a Gulf Stream warm-core ring, *J. Geophys. Res.*, 91, 9615–9623, 1986.
- Coles, V. J. and Hood, R. R.: Modeling the impact of iron and phosphorus limitations on nitrogen fixation in the Atlantic Ocean, *Biogeosciences*, 4, 455–479, 2007, <http://www.biogeosciences.net/4/455/2007/>.
- Cullen, J. J., Franks, P. J. S., Karl, D. M., Longhurst, A.: Chapter 8: Physical influences on marine ecosystem dynamics, in: *The Sea*, edited by: Robinson, A. R., McCarthy, J. J., and Rothschild, B. J., John Wiley and Sons, New York, 12, 297–336, 2002.
- DuRand, M. D., Olson, R. J., Sallie W. Chisholm, S. W.: Phytoplankton population dynamics at the Bermuda Atlantic Time-series station in the Sargasso Sea, *Deep Sea Res. II*, 48, 1983–2003, 2001.
- Dutkiewicz, S., Follows, M., Marshall, J., and Gregg, W.: Interannual variability of phytoplankton abundances in the North Atlantic, *Deep-Sea Res. Pt. II*, 48, 2323–2344, 2001.
- Fasham, M. J. R., Sarmiento, J. L., Slater, R. D., Ducklow, H. W., and Williams, R.: Ecosystem behavior at Bermuda station “S” and ocean weather station “India”: A general-circulation model and observational analysis, *Global Biogeochem. Cy.*, 7, 379–415, 1993.
- Fennel, K., Wilkin, J., Levin, J., Moisan, J., O’Reilly, J., and Haidvogel, D.: Nitrogen cycling in the Middle Atlantic Bight: results from a three-dimensional model and implications for the North Atlantic nitrogen budget, *Global Biogeochem. Cy.*, 20, GB3007, doi:10.1029/2005GB002456, 2006.
- Fennel, K., Wilkin, J., Previdi, M., and Najjar, R.: Denitrification effects on air-sea CO_2 flux in the coastal ocean: simulations for the northwest North Atlantic, *Geophys. Res. Lett.*, 35, L24608, doi:10.1029/2008GL036147, 2008.
- Flather, R. A.: A tidal model of the northwest European continental shelf, *Mem. Soc. R. Sci. Liege*, 10, 141–164, 1976.
- Franks, P. J. S. and Chen, C.: A 3-D prognostic numerical model study of the Georges bank ecosystem. Part II: biological-physical model, *Deep-Sea Res. II*, 48, 457–482, doi:10.1016/S0967-0645(00)00125-9, 2001.
- Friedrichs, M., Dusenberry, J., Anderson, L., Armstrong, R., Chai, F., Christian, J., Doney, S., Dunne, J., Fujii, M., and Hood, R.: Assessment of skill and portability in regional marine biogeochemical models: role of multiple planktonic groups, *J. Geophys. Res.*, 112, 8001, doi:10.1029/2006JC003852, 2007.
- Geider, R., MacIntyre, H., and Kana, T.: A dynamic regulatory model of phytoplankton acclimation to light, nutrients, and temperature, *Limnol. Oceanogr.*, 43, 679–694, 1998.
- Greene, C. and Pershing, A.: Climate drives sea change, *Science*, 315, 1084–1085, 2007.
- Gruber, N., Frenzel, H., Doney, S., Marchesiello, P., McWilliams, J., Moisan, J., Oram, J., Plattner, G., and Stolzenbach, K.: Eddy-resolving simulation of plankton ecosystem dynamics in the California current system, *Deep-Sea Res. I*, 53, 1483–1516, 2006.
- Haidvogel, D. B., Arango, H., Budgell, W. P., Cornuelle, B. D., Curchitser, E., Lorenzo, E. D., Fennel, K., Geyer, W. R., Hermann, A. J., Lanerolle, L., Levin, J., McWilliams, J. C., Miller, A. J., Moore, A. M., Powell, T. M., Shchepetkin, A. F., Sherwood, C. R., Signell, R. P., Warner, J. C., and Wilkin, J.: Ocean forecasting in terrain-following coordinates: formulation and skill assessment of the regional ocean modeling system, *J. Comput. Phys.*, 227, 3595–3624, doi:10.1016/j.jcp.2007.06.016, 2008.

- He, R. and Chen, K.: Investigation of northeastern North America coastal circulation using a regional circulation hindcast model. Part 1: mean circulation, *J. Geophys. Res.*, submitted, 2009.
- Hofmann, E., Druon, J., Fennel, K., Friedrichs, M., Haidvogel, D., Lee, C., Mannino, A., McClain, C., Najjar, R., O'Reilly, J., Pollard, D., Previdi, M., Seitzinger, S., Siewert, S., Signorini, S., and Wilkin, W.: Eastern US continental shelf carbon budget: integrating models, data assimilation and analysis, *Oceanography*, 21, p. 86, 2008.
- Houghton, R., Flagg, C., and Pietrafesa, L.: Shelf-slope water frontal structure, motion and eddy heat flux in the southern Middle Atlantic Bight, *Deep-Sea Res. II*, 41, 273–306, 1994.
- Kilpatrick, K. A., Podesta, G. P., and Evans, R.: Overview of the NOAA/NASA advanced very high resolution radiometer Pathfinder algorithm for sea surface temperature and associated matchup database, *J. Geophys. Res.*, 106(C5), 9179–9198, doi:10.1029/1999JC000065, 2001.
- Kiorboe, T.: Turbulence, phytoplankton cell size and the structure of pelagic food webs, *Adv. Mar. Biol.*, 29, 1–72, 1993.
- Li, W. K. W., Harrison, W. G.: Chlorophyll, bacteria and picoplankton in ecological provinces of the North Atlantic, *Deep Sea Res. II*, 48, 2271–2293, 2001.
- Lima, I. D. and Doney, S. C.: A three-dimensional, multi-nutrient, and size-structured ecosystem model for the North Atlantic, *Global Biogeochem. Cy.*, 18, GB3019, doi:10.1029/2003GB002146, 2004.
- Linder, C. and Gawarkiewicz, G.: A climatology of the shelfbreak front in the Middle Atlantic Bight, *J. Geophys. Res.*, 103, 18405–18423, 1998.
- Loder, J. W., Petrie, B., and Gawarkiewicz, G.: The coastal ocean off northeastern North America: a large-scale view, in: *The Sea*, edited by: Robinson, A. R. and Brink, K. H., 11, 105–133, John Wiley and Sons, Inc., New York, 1998.
- Longhurst, A.: *Ecological Geography of the Sea*, 2nd edn., Academic Press, Amsterdam, 2007.
- Lozier, M. and Gawarkiewicz, G.: Cross-frontal exchange in the Middle Atlantic Bight as evidenced by surface drifters, *J. Phys. Oceanogr.*, 31, 2498–2510, 2001.
- Luettich, R. A., Westerink, J. J., and Scheffner, N. W.: ADCIRC: an advanced three-dimensional circulation model for shelves, coasts, and estuaries. Report 1. Theory and Methodology of ADCIRC-2DDI and ADCIRC-3DL, Tech. Rep. CERC-TR-DRP-92-6, Coastal Engineering Research Center Vicksburg, 1992.
- Mellor, G. L. and Yamada, T.: Development of a turbulence closure-model for geophysical fluid problems, *Rev. Geophys.*, 20, 851–875, 1982.
- Mountain, D.: Variability in the properties of shelf water in the Middle Atlantic Bight, 1977–1999, *J. Geophys. Res.*, 108, 1029–1044, 2003.
- O'Reilly, J. E. and Zetlin, C.: Seasonal, horizontal, and vertical distribution of phytoplankton chlorophyll a in the northeast US continental shelf ecosystem, Tech. Rep. 139, US Department of Commerce, 1998.
- O'Reilly, J. E. and 21 co-authors: Ocean color chlorophyll algorithms for SeaWiFS, OC2, and OC4: Version 4. in: O'Reilly, J. E. and 24 co-authors: SeaWiFS postlaunch calibration and validation analyses, part 3, NASA Tech. Memo., 2000–206892, edited by: Hooker, S. B. and Firestone, E. R., NASA Goddard Space Flight Center, Greenbelt, Maryland, 11, 9–23, 2000.
- Oschlies, A.: Nutrient supply to the surface waters of the North Atlantic: a model study, *J. Geophys. Res.*, 107(C5), 3046, doi:10.1029/2000JC000275, 2002.
- Pätsch, J. and Kühn, W.: Nitrogen and carbon cycling in the North Sea and exchange with the North Atlantic: a model study. Part I. Nitrogen budget and fluxes, *Cont. Shelf Res.*, 28, 767–787, 2008.
- Previdi, M., Fennel, K., Wilkin, L., and Haidvogel, D.: Interannual variability in Atmospheric CO₂ uptake on the northeast US continental shelf, *J. Geophys. Res.*, doi:10.1029/2008JG000881, in press, 2009.
- Ramp, S., Schlitz, R., and Wright, W.: The deep flow through the Northeast Channel, Gulf of Maine, *J. Phys. Oceanogr.*, 15, 1790–1808, 1985.
- Rosby, T. and Benway, R.: Slow variations in mean path of the Gulf Stream east of Cape Hatteras, *Geophys. Res. Lett.*, 27, 117–120, 2000.
- Ryan, J., Yoder, J., and Townsend, D.: Influence of a Gulf Stream warm-core ring on water mass and chlorophyll distributions along the southern flank of Georges Bank, *Deep-Sea Res. II*, 48, 159–178, 2001.
- Schollaert, S., Rossby, T., and Yoder, J.: Gulf Stream cross-frontal exchange: possible mechanisms to explain interannual variations in phytoplankton chlorophyll in the Slope Sea during the SeaWiFS years, *Deep-Sea Res. Pt. II*, 51, 173–188, 2004.
- Seitzinger, S., Harrison, J., Bohlke, J., Bouwman, A., Lowrance, R., Peterson, B., Tobias, C., and Drecht, G.: Denitrification across landscapes and waterscapes: a synthesis, *Ecol. Appl.*, 16, 2064–2090, 2006.
- Sieburth, J. M., Smetacek, V., and Lenz, J.: Pelagic ecosystem structure: heterotrophic compartments of the plankton and their relationship to plankton size fractions, *Limnol. Oceanogr.*, 23, 1256–1263, 1978.
- Smith, P., Houghton, R., Fairbanks, R., and Mountain, D.: Interannual variability of boundary fluxes and water mass properties in the Gulf of Maine and on Georges Bank: 1993–1997, *Deep-Sea Res. II*, 48, 37–70, 2001.
- Stow, C. A., Jolliff, J., McGillicuddy Jr., D. J., Doney, S. C., Allen, J. I., Friedrichs, M. A., Rose, K. A., and Wallhead, P.: Skill assessment for coupled biological/physical models of marine systems, *J. Marine Syst.*, 76, 4–15, 2009.
- Strom, S. L., Miller, C. B., and Frost, B. W.: What sets lower limits to phytoplankton stocks in high-nitrate, low-chlorophyll regions of the open ocean?, *Mar. Ecol.-Prog. Ser.*, 193, 19–31, 2000.
- Taylor, K. E.: Summarizing multiple aspects of model performance in a single diagram, *J. Geophys. Res.-Atmos.*, 106, 7183–7192, 2001.
- Thomas, A., Townsend, D., and Weatherbee, R.: Satellite-measured phytoplankton variability in the Gulf of Maine, *Cont. Shelf Res.*, 23, 971–989, 2003.
- Thomas, H., Bozec, Y., Elkalay, K., and De Baar, H.: Enhanced open ocean storage of CO₂ from shelf sea pumping, *Science*, 304, 1005–1008, 2004.
- Yoder, J. A., Schollaert, S. E., O'Reilly, J. E.: Climatological Phytoplankton Chlorophyll and Sea Surface Temperature Patterns in Continental Shelf and Slope Waters off the Northeast U.S. Coast, *Limnol. Oceanogr.*, 47, 672–682, 2002.