Impact of bio-optical data assimilation on short-term coupled physical, bio-optical model predictions

Igor Shulman,[1] Sergey Frolov,[2] Stephanie Anderson,[1] Brad Penta,[1] Rick Gould,[1] Peter Sakalaukus,[1] and Sherwin Ladner[1]

Received 31 October 2012; revised 24 March 2013; accepted 25 March 2013; published 30 April 2013.

[1] Data assimilation experiments with the coupled physical, bio-optical model of Monterey Bay are presented. The objective of this study is to investigate whether the assimilation of satellite-derived bio-optical properties can improve the model predictions (phytoplankton population, chlorophyll) in a coastal ocean on time scales of 1–5 days. The Monterey Bay model consists of a physical model based on the Navy Coastal Ocean Model and a biochemical model which includes three nutrients, two phytoplankton groups (diatoms and small phytoplankton), two groups of zooplankton grazers, and two detrital pools. The Navy Coupled Ocean Data Assimilation system is used for the assimilation of physical observations. For the assimilation of bio-optical observations, we used reduced-order Kalman filter with a stationary forecast error covariance. The forecast error covariance is specified in the subspace of the multivariate (bio-optical, physical) empirical orthogonal functions estimated from a monthlong model run. With the assimilation of satellite-derived bio-optical properties (chlorophyll *a* or absorption due to phytoplankton), the model was able to reproduce intensity and tendencies in subsurface chlorophyll distributions observed at water sample locations in the Monterey Bay, CA. Data assimilation also improved agreement between the observed and model-predicted ratios between diatoms and small phytoplankton populations. Model runs with or without assimilation of satellite-derived bio-optical observations show underestimated values of nitrate as compared to the water sample observations. We found that an instantaneous update of nitrate based on statistical relations between temperature and nitrate corrected the model underestimation of the nitrate fields during the multivariate update.

Citation: Shulman, I., S. Frolov, S. Anderson, B. Penta, R. Gould, P. Sakalaukus, and S. Ladner (2013), Impact of bio-optical data assimilation on short-term coupled physical, bio-optical model predictions, *J. Geophys. Res. Oceans*, *118*, 2215–2230, doi:10.1002/jgrc.20177.

1. Introduction

[2] During the last decade, considerable efforts have been made in development and testing approaches for the assimilation of bio-optical properties (especially satellite observations of the ocean color) into biochemical, physical models. Some studies have focused on the optimization of model parameters and parameterizations with regards to observations [see, for example, *Spitz et al.*, 1998; *McGillicuddy et al.*, 1998; *Fennel et al.*, 2001; *Hofmann and Friedrichs*, 2002; *Friedrichs et al.*, 2006; *Smith et al.*, 2009; *Doron et al.*, 2011], while others have focused on the sequential estimation (updating) of model bio-optical and physical state variables based on available observations [for example, *Anderson et al.*, 2000, 2001; *Natvik and Evensen*, 2003; *Besiktepe et al.*, 2003; *Nerger and Gregg*, 2007; *Cossarini et al.*, 2009; *Smith and McGillicuddy*, 2011; *Ciavatta et al.*, 2011; *Ford et al.*, 2012; *Hu et al.*, 2012; *Rousseaux and Gregg*, 2012]. The objectives of many studies were the improvement of seasonal or yearly hindcasts of bio-optical properties. For example, in *Cossarini et al.* [2009], the objective was to investigate the seasonal ecosystem dynamics of the Lagoon of Venice. The objective of *Ciavatta et al.* [2011] was to investigate if a yearlong assimilation of weekly satellite chlorophyll data improves the hindcast of key biogeochemical variables in shelf seas. *Ford et al.* [2012] conducted assimilation of satellite-derived chlorophyll into the global coupled physical, biochemical model. The objective of *Rousseaux and Gregg* [2012] was the study of climate variability and phytoplankton composition in the Pacific Ocean. The impact of yearlong assimilation of SeaWiFS- and MODIS-derived chlorophyll on ecosystem model predictions was investigated in *Hu et al.* [2012]. See *Gregg* [2008], *McClain* [2009], and *Hu et al.* [2012] for a more detail review of data assimilative studies.

[1]Oceanography Division, Naval Research Laboratory, Stennis Space Center, Mississippi, USA.
[2]Naval Research Laboratory, Monterey, California, USA.

Corresponding author: I. Shulman, Oceanography Division, Naval Research Laboratory, Stennis Space Center, MS 39529, USA. (igor.shulman@nrlssc.navy.mil)

2169-9275/13/10.1002/jgrc.20177

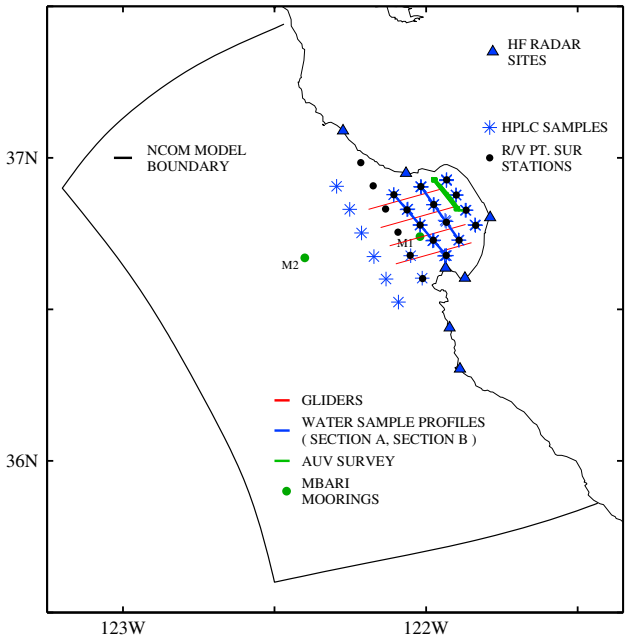

Figure 1. Map of the observational assets during June 2008 field program: MBARI moorings M1 and M2 locations; R/V Point Sur stations and water sample locations; HPLC sample locations; glider tracks (shown schematically); AUV DORADO survey; locations of HF radar sites.

[3] In contrast to the existing studies, the objective of this paper is to investigate whether the assimilation of satellite-derived bio-optical properties (as either chlorophyll *a* (Chl) or absorption coefficient) can improve the ecosystem model predictions of chlorophyll and phytoplankton population in a coastal ocean on time scales of 1–5 days. The specific time scale of 1–5 days is chosen because it is a time scale of availability of the atmospheric model forecast needed to force the oceanic model forecast. The atmospheric model forecast includes predictions of short-wave radiation, which is critical not only for forecasting the heat content and other physical properties of the ocean but also for estimating the photosynthetically active radiation (PAR) which drives photosynthesis of the ecosystem model, and relevant to the forecast of the underwater light. Predictions of optical properties and underwater light are critical for numerous Navy operations, which rely on 1–5 days of forecasts.

[4] We designed our computational experiments to coincide with a large bio-optical field campaign that was conducted in Monterey Bay, California during a sustained wind-driven upwelling event in June 2008. The field program captured the dynamic response of the Bay ecosystem to the continuous supply of nutrients from coastal upwelling. To characterize the dynamics of the system, a combination of field assets and measurements systems was deployed, including ship surveys, buoys, and autonomous underwater vehicles. The experiment was a collaboration between the NRL "Bio-Optical Studies of Predictability and Assimilation for the Coastal Environment (BIOSPACE)" project, Multidisciplinary University Research Initiative (MURI) project "Rapid Environmental Assessment Using an Integrated Coastal Ocean Observation-Modeling System (ESPRESSO)," and the Monterey Bay Aquarium Research Institute (MBARI). The objective of the NRL participation in the experiment was to study the variability and predictability of underwater light and coupled bio-optical and physical properties of the water column on time scales of 1–5 days.

[5] The structure of the paper is as follows: Section 2 describes observations, models, and data assimilation schemes used in this study. The bio-optical physical conditions during the data assimilation experiments are described in section 2.1.3.3. The design of data assimilation experiments is described in section 3. Section 4 presents results of the data assimilation experiments. Section 5 is devoted to discussions and conclusions.

2. Methods

2.1. Observations

2.1.1. Physical Observations

[6] Observations of winds, temperature, and salinity from the Monterey Bay Aquarium Research Institute (MBARI) surface moorings M1 (122.02°W, 36.74°N) and M2 (122.40°W, 36.67°N) are used in this study (Figure 1). Near-surface 3 m wind speed and direction were measured by a MetSys wind monitor. Temperature and salinity were measured by Sea-Bird MicroCAT CTD sensors at 12 depths between 1 and 350 m. According to the manufacturer's stated accuracy, the data are expected to be accurate to within about 0.005°C and 0.006 practical salinity units (psu).

[7] Surface current observations used in this study were derived from the California Coastal Ocean Current Mapping Program's HF radar network (www.cocmp.org). Surface currents were estimated based on inputs from seven HF radar sites (Figure 1). Vector currents were estimated on a Cartesian grid with a horizontal resolution of 3 km by computing the best fit vector velocity components using all radial velocity observations within a radius of 3 km for each grid point each hour [*Paduan et al.*, 2006]. Several studies have investigated the performance of the Monterey Bay HF radar network by comparing the radar-derived currents with in situ velocity observations and by comparing radar-to-radar velocity estimates on the overwater baselines between radar sites [e.g., *Paduan et al.*, 2006]. Consistent uncertainty values emerge in the range of 7–9 cm/s for the remotely estimated velocities.

[8] The R/V Point Sur occupied 25 hydrographic and optical stations from 2 to 13 June 2008 (Figure 1). Temperature and salinity depth profiles with 1 m vertical resolution were derived from Sea-Bird SBE 9+ CTD measurements using standard Sea-Bird processing software. Comparisons of the moored data with adjacent shipboard profiles show agreement to generally be within 0.1°C and 0.01 psu.

[9] Four NRL and two Rutgers University SLOCUM gliders [*Schofield et al.*, 2007] were deployed during a period of 2 weeks of surveys with the R/V Point Sur. The gliders were equipped with a SeaBird CTD and collected temperature and salinity profiles up to 200 m depth mostly inside the Bay because the navigation of gliders outside the bay became difficult due to strong wind–driven currents (~1–2 knots).

[10] Satellite surface temperature data, available in situ temperature, and salinity profiles from the Global Ocean Data Assimilation Experiment (GODAE) data set (http://www.usgodae.org/) are used in this study for the assimilation

into the Monterey Bay model described in section 2.2. The description of the data set, processing, and quality control procedures are described in *Cummings* [2005] and *Cummings et al*. [2009].

2.1.2. Satellite MODIS-Aqua Ocean Color Data: Chlorophyll *a* Concentration and Phytoplankton Absorption Coefficient

[11] The MODIS-Aqua satellite imagery was processed using the NRL Automated Processing System (APS). APS is a complete end-to-end system that includes sensor calibration, atmospheric correction (with near-infrared correction for coastal waters), and bio-optical inversion. APS incorporates, and is consistent with, the latest NASA MODIS code (SeaDAS) [*Gould et al*., 2011; *Martinolich and Scardino*, 2011].

[12] In this study, estimates of the chlorophyll *a* (Chl) and absorption coefficient due to phytoplankton at 488 nm ($a_{\mathrm{ph}}(488)$) from MODIS-Aqua imagery on 5 and 10 June 2008 were assimilated into the bio-optical, physical model described in section 2.2. Chlorophyll data are derived by OC3M algorithm [*O'Reilly et al*., 2000], while $a_{\mathrm{ph}}(488)$ data are derived by using a quasi-analytical algorithm (QAA) [*Lee et al*., 2002] at 1 km pixel resolution. Data are interpolated to the model grid spatially and temporally to 0Z and 12Z (with 12 h data assimilation update cycle (see section 3)).

[13] Errors in satellite derived products as chlorophyll *a* and absorption are usually poorly known. *McClain* [2009] stated that many recent investigations in comparison of satellite derived products with water samples or high-performance liquid chromatography (HPLC) data were inconclusive mostly due to differences in the pigment measurement methodology, i.e., fluorometric for water samples versus high-pressure liquid chromatography (HPLC). In *McClain* [2009]: "The satellite data product accuracy goals generally accepted by the international missions are ±5% for water-leaving radiances and ±35% for chlorophyll in the open ocean." At the same time, it is also stated that errors differ regionally. *Lee et al*. [2010] reported error in estimation of absorption around 10% for values below 0.1 m^{-1}, which is an about average value for the Monterey Bay area.

2.1.3. Bio-Optical Observations Used for Model Predictions Verification

2.1.3.1. Extracted Chlorophyll From the Water Samples

[14] Water was collected at up to 12 depths at each R/V Point Sur station (Figure 1). Samples (280 ml) were taken from the Niskin bottles and filtered through 25 mm Whatman GF/F (glass fiber filters) at 5–7 mm Hg pressure. The filters were then placed into glass scintillation vials with 10 ml of 90% acetone and frozen for 24 h to allow chlorophyll extraction [*Venrick and Hayward*, 1984]. Samples were allowed to warm for several hours in the dark before fluorescence measurements were performed with a Turner 10-AU Fluorometer using standard methods [*Holm-Hansen et al*., 1965; *Lorenzen*, 1966]. To correct for phaeophytin interference, each sample was then acidified with three drops of 5% HCl to convert chlorophyll to phaeophytin. The ratio of these two measurements is directly proportional to chlorophyll concentration.

2.1.3.2. High-Performance Liquid Chromatography Data

[15] Water samples (540 ml) collected from near-surface (~0.5 m) Niskin bottles were filtered onto Whatman glass fiber filters (GF/F). The high-performance liquid chromatography (HPLC) analysis provided pigment indices and proportion factor for microplankton, nanoplankton, and picoplankton [*Vidussi et al*., 2001]. The pigment data indicated that the microplankton fraction was composed predominantly of diatoms (based on the presence of fucoxanthin). For this analysis, the nano- and picoplankton fractions were combined to represent the "small phytoplankton" in our coupled bio-optical physical model (section 2.2). *Claustre et al*. [2004] reported 11.5% uncertainty for fucoxanthin and 7% for chlorophyll *a*.

2.1.3.3. Nitrate Data

[16] Propeller-driven AUV such as the MBARI manufactured DORADO has been described in *Bellingham et al*. [2000] and *Ryan et al*. [2009]. The DORADO was deployed on 3 June 2008 in the Monterey Bay (Figure 1), and instruments on board included in situ ultraviolet spectrophotometer sensor that measured nitrate concentrations [*Johnson and Coletti*, 2002].

2.2. Coupled Physical, Bio-Optical Model of the Monterey Bay

[17] The Monterey Bay model (called the Navy Coastal Ocean Model (NCOM) Innovative Coastal-Ocean Observing Network (ICON)) consists of a physical model [*Shulman et al*., 2007], which is coupled to a biochemical model [*Chai et al*., 2002]. The initial model development started under the National Oceanic Partnership Program ICON project. The physical model of the Monterey Bay is based on the NCOM model, which is a primitive-equation, 3-D, hydrostatic model. It uses the Mellor-Yamada level 2.5 turbulence closure scheme and the Smagorinsky formulation for horizontal mixing [*Martin*, 2006; *Barron et al*., 2006]. The NCOM ICON model is set up on a curvilinear orthogonal grid with resolution ranging from 1 to 4 km. The model domain is shown on Figure 1. The model is forced with surface fluxes from the Coupled Ocean and Atmospheric Mesoscale Prediction System (COAMPS) [*Doyle et al*., 2009] at 3 km horizontal resolution. The 3 km resolution COAMPS grid mesh is centered over Central California and the Monterey Bay. The biochemical model (the Carbon, Silicon, Nitrogen Ecosystem (CoSINE) model) [*Chai et al*., 2002; *Shulman et al*., 2011] of the NCOM ICON simulates the dynamics of two sizes of phytoplankton, small phytoplankton cells (<5 μm in diameter) and diatoms, two zooplankton grazers, nitrate, silicate, ammonium, and two detritus pools (Figure 2). Constituents from the biochemical model are used to estimate chlorophyll and inherent optical properties (IOPs) based on the methodology outlined by *Fujii et al*. [2007]. For example, the model chlorophyll concentration (chl) and absorption due to phytoplankton ($a_{\mathrm{ph}}(\lambda)$) are estimated based on the following:

$$\mathrm{chl} = \mathrm{chl}_1 \cdot P1 + \mathrm{chl}_2 \cdot P2 \tag{1}$$

$$\begin{aligned} a_{\mathrm{ph}}(\lambda) &= a_1^*(\lambda)\cdot\mathrm{chl}_1\cdot P1 + a_2^*(\lambda)\cdot\mathrm{chl}_2\cdot P2 \\ a_i^*(\lambda) &= a^*_{i,(\mathrm{highlight})}(\lambda)\cdot\left(1 - f_{\theta,i}\right) + a^*_{i,(\mathrm{lowlight})}(\lambda)\cdot f_{\theta,i} \\ f_{\theta,i} &= \frac{\mathrm{chl}_i/\mathrm{cn}_i - \theta_{\min}}{\theta_{\max} - \theta_{\min}} \end{aligned} \tag{2}$$

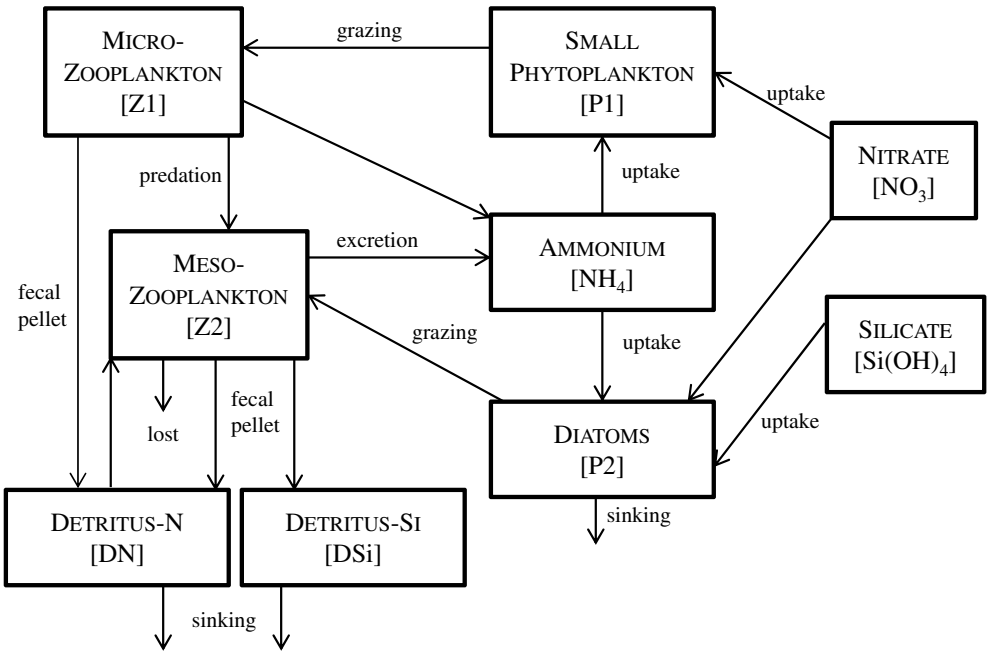

Figure 2. Schematic view and flowchart of the nine-component biochemical model (the Carbon, Silicon, Nitrogen Ecosystem (CoSINE)).

where $P1$ is the small phytoplankton concentration, $P2$ is the diatoms concentration, $a_{\mathrm{ph}}(\lambda)$ is the absorption coefficient due to phytoplankton, λ is the wavelength, $a_1^*(\lambda)$ and $a_2^*(\lambda)$ are chlorophyll-specific absorption coefficients by small phytoplankton and diatoms , $a^*_{i,(\mathrm{highlight})}(\lambda)$ and $a^*_{i,(\mathrm{lowlight})}(\lambda)$ are chlorophyll-specific absorption coefficients at high and low light by each phytoplankton group [*Fujii et al*., 2007], chl_i are chlorophyll to nitrogen conversion coefficients, cn_i are carbon to nitrogen conversion coefficients, $f_{\theta,\,i}$ is the phytoplankton size fraction, and $\theta_{\min}$ and $\theta_{\max}$ are the minimum and maximum phytoplanktonic chlorophyll to carbon ratios [*Fujii et al*., 2007]. Absorption in equation (2) is modeled as a sum of absorptions from small phytoplankton and diatoms. The chlorophyll-specific absorption coefficients for small phytoplankton and diatoms are modeled separately, taking into account their photoadaptive state (e.g., their specific chlorophyll to carbon ratio). This requires specification of high/low light absorption coefficients for each phytoplankton group (small phytoplankton and diatoms). For more details, see *Fujii et al*. [2007]. It is known that phytoplanktonic chlorophyll to carbon ratio is not constant and depends on light, nutrients, temperature, etc. However, to model the ratio as variable will require introduction of more state variables, as well as more highly uncertain model parameters into the biochemical model. Because the objective of the paper is modeling on short-term time scales (1–5 days), we prefer to use (1)–(2) relations rather than to increase a number of the biochemical model state variables and highly uncertain model parameters. Only $P1$ and $P2$ are prognostic variables in (1) and (2).

[18] Phytoplankton photosynthesis in the biochemical model is driven by photosynthetically active radiation (PAR), which is estimated based on the shortwave radiation flux from the COAMPS model. The *Penta et al*. [2008] scheme is used for PAR attenuation with depth.

[19] Open boundary conditions for the NCOM ICON are derived from the regional model of the California Current (NCOM CCS) [*Shulman et al*., 2007]. The NCOM CCS has a horizontal resolution of about 9 km, and the model is forced with atmospheric products derived from the COAMPS [*Doyle et al*., 2009]. As in NCOM ICON model, the biochemical model of the NCOM CCS is also the nine-compartment model of *Chai et al*. [2002].

[20] Open boundary conditions for physical variables (temperature, salinity, velocities) for the regional NCOM CCS model are derived from the NCOM global model [*Rhodes et al*., 2002; *Barron et al*., 2004], which has 1/8° horizontal resolution. The NCOM global model does not have a biochemical model to derive open boundary conditions for the biochemical model of the NCOM CCS. For this reason, biochemical tracers of the NCOM CCS were spun up from the climatological values of the nutrients (nitrate and silicate from The World Atlas) [*Garcia et al*., 2006] and background values for other biochemical variables from October 1998 to June 2008.

2.3. Assimilation of Physical Observations

[21] For the assimilation of physical observations (temperature and salinity), the NCOM ICON model uses the Navy Coupled Ocean Data Assimilation (NCODA) system [*Cummings*, 2005; *Cummings et al*., 2009]. The NCODA is a fully 3-D multivariate optimum interpolation system. Assimilation of temperature and salinity data is performed every 12 h (assimilation cycle). The NCODA assimilates satellite altimeter observations, satellite surface temperature, as well as available in situ vertical temperature and salinity profiles from XBTs, ARGO floats, moored buoys, and gliders from

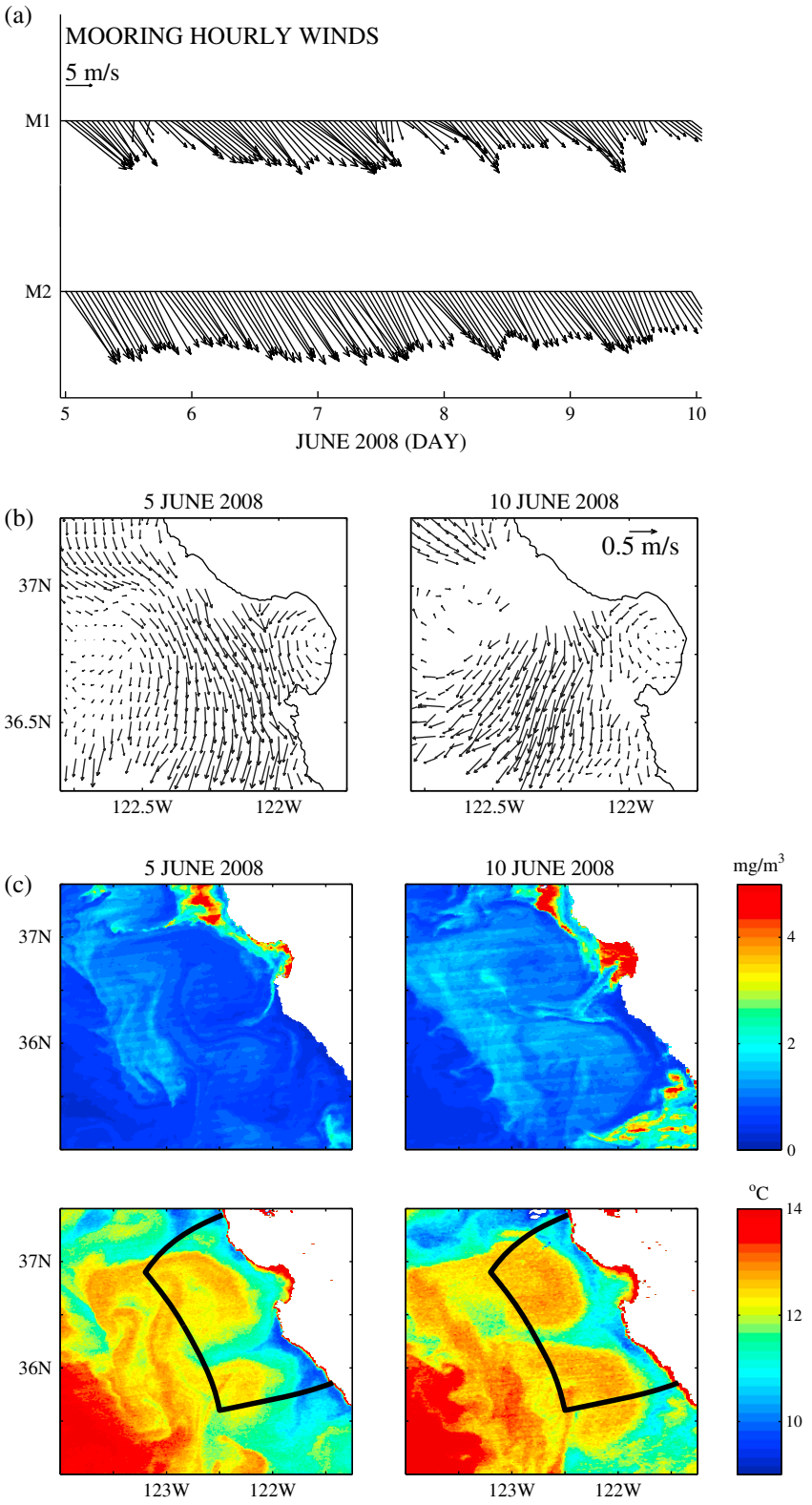

Figure 3. (a) 10 m wind velocity at MBARI moorings M1 and M2. (b) HF radar surface currents. (c) MODIS-Aqua chlorophyll and MODIS-Aqua SSTs (bottom). The modeling domain is shown with black solid line overlay over MODIS-Aqua SST images (bottom).

the Global Ocean Data Assimilation Experiment (GODAE) data set. The description of the data sets, processing, and quality control procedures are described in *Cummings* [2005] and *Cummings et al*. [2009]. Results of glider, ship, and satellite data assimilation into the NCOM ICON model are described in *Shulman et al*. [2009, 2010].

Table 1. Description of the Model Runs

Runs	Assimilation			Multivariate Update		
	Physics (NCODA)	MODIS Chl (BOMA)	MODIS $a_{ph}(488)$ (BOMA)	Small Phytoplankton	Diatoms	Nitrate
Run 1	No	No	No	N/A	N/A	N/A
Run 2	Yes	No	No	N/A	N/A	N/A
Run 3	Yes	Yes	No	Yes	Yes	No
Run 4	Yes	No	Yes	Yes	Yes	No
Run 5	Yes	Yes	No	Yes	Yes	Yes
Run 6[a]	Yes	Yes	No	Yes	Yes	Yes

[a]Adjustment of nitrate based on temperature **versus** nitrate statistical relation (see section 3).

2.4. A Multivariate Data Assimilation of Bio-Optical Properties (BOMA)

2.4.1. Kalman Gain Update

[22] To preserve the robustness of the existing assimilation system for physical fields (NCODA), we decided to decouple updates to the physical fields from the updates to the components of the ecosystem model. To assimilate bio-optical measurements into the ecosystem model, we used reduced-order Kalman filter with a stationary forecast error covariance.

[23] The analysis fields for the bio-optical model state variables were updated using Kalman update equations:

$$x^a = x^f + K\left(y - Hx^f\right) \tag{3}$$

$$K = P_{xy}P_{yy}^{-1} \tag{4}$$

where x^a and x^f are the vector of analyzed and forecasted bio-optical properties, y are available observations, H is the observational operator that maps the model state onto available observations, and K is the Kalman gain matrix. Covariance matrices P_{xy} and P_{yy} in the Kalman gain equation (4) are the cross-covariance between forecast and observation errors and the innovation error covariance matrixes respectively. For a linear measurement operator H, these covariance matrices become:

$$P_{xy} = P^f H^T \tag{5}$$

$$P_{yy} = HP^f H^T + R \tag{6}$$

where P^f is the forecast error covariance matrix, and R is the combined covariance of measurement and representation errors.

2.4.2. Forecast Error Covariance Model

[24] Similar to *Cane et al.* [1996] and *Nerger and Gregg* [2007], we used a stationary form of the error covariance P^f. We specified the forecast error covariance P^f using an ensemble of model states X^{ens} drawn from a historic model run:

$$P^f \approx \alpha P^{\text{ens}} = \alpha E\left[(X^{\text{ens}} - E[X^{\text{ens}}])(X^{\text{ens}} - E[X^{\text{ens}}])^T\right] \tag{7}$$

where α is a scalar that scales the climatological ensemble to be consistent with the statistics of model innovations. Twin data assimilation experiments were conducted, when pseudo- "observations" sampled from the "true" model run were assimilated into the model run with different initial conditions from the "true" run. Optimal value of $\alpha = 0.01$ was determined based on minimization of misfits between "true" and twin data assimilative run.

[25] We drew the ensemble X^{ens} of ~700 model states from a monthlong run of nonassimilated model (see section 3 for details of the run setup). To reduce the storage requirements and because the ensemble approximation P^{ens} was rank deficient, we stored matrix P^{ens} using a truncated series of eigen functions estimated from SVD of X^{ens}:

$$P^{\text{ens}} \approx Z\Lambda Z^T$$

where Z is the matrix of orthonormal 3-D eigen functions (EOFs) and Λ is the diagonal matrix of eigen values. We retained 100 eigen functions that captured 98% of the variance in the ensemble covariance P^{ens}.

[26] In our experiments, we had more observations than ensemble members. Hence, it was more efficient to implement the inverse of covariance P_{yy} in the space of the EOF coefficients instead of the observation space formulation in equation 6. To transform the P_{yy} inverse from observational space to EOF space, and to the form that requires inverse of only R matrix, we used the Sherman-Morrison-Woodbury formula [*Barth et al.*, 2011] as follows:

$$\begin{aligned} P_{yy}^{-1} &= (\alpha HZ\Lambda Z^T H^T + R)^{-1} = (UU^T + R)^{-1} = \\ &= R^{-1} - (R^{-1}U)\left[I + (R^{-1}U)^T U\right](R^{-1}U)^T \end{aligned} \tag{8}$$

where

$$U = \sqrt{\alpha}HZ\sqrt{\Lambda} \tag{9}$$

2.4.3. Observation Error Covariance Model

[27] The combined covariance R of measurement and representation errors is usually poorly known. As we stated in section 2.1.2, "the satellite data product accuracy generally accepted by the international missions are ±5% for water-leaving radiances and ±35% for chlorophyll *a* in the open ocean" [*McClain*, 2009]. However, errors differ regionally. As it is shown in section 4, the coupled physical, bio-optical model (section 2.2) is under productive in the Bay without data assimilation, and it is desirable to increase influence of observations on model predictions. We assumed that covariance R had diagonal structure (uncorrelated errors) and was stationary and proportional to the variance of the observed field. We set the variance of R to be equal to 10% of the field variance. The resulting magnitude of the measurement error was in agreement with uncertainty studies [*Lee et al.*, 2002, 2010] of the QAA satellite retrieval algorithm that was used in our study (section 2.1.2).

2.4.4. Localization

[28] To mitigate for the presence of spurious correlations in our ensemble approximation to the forecast error

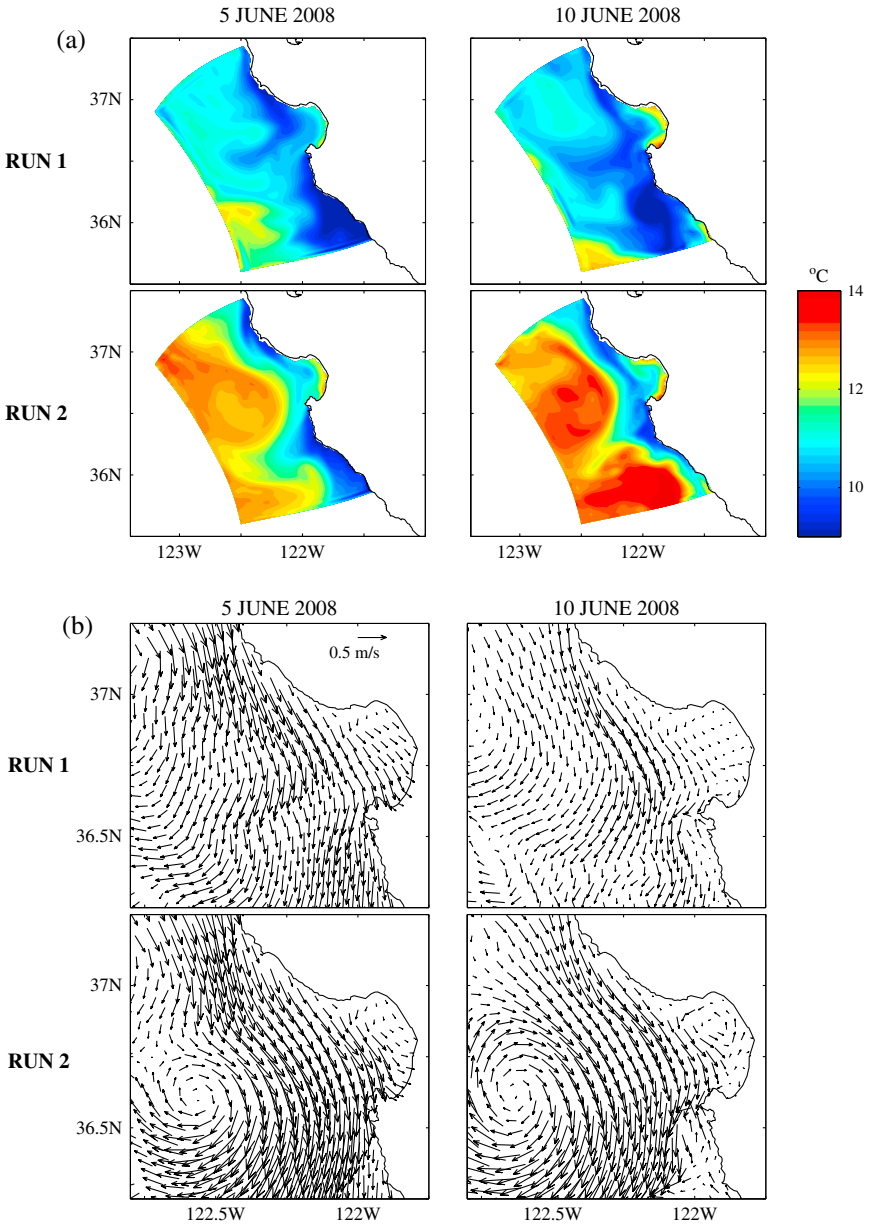

Figure 4. Model-predicted SSTs and surface currents for runs 1 and 2 (see section 3 for model runs design).

covariance equation (7) and to exclude remote observations from the analysis of the local grid point, we localized the forecast error covariance P^{ens} using the box-car localization function:

$$P^f = C_{\text{loc}} * P^{\text{ens}} \tag{10}$$

$$C_{\text{loc}}(x_1,,x_2) = \begin{cases} c(x_1,x_2) = 1 if \|x_1 - x_2\|_2 \leq L_{\text{loc}} \\ c(x_1,x_2) = 0 if \|x_1 - x_2\|_2 > L_{\text{loc}} \end{cases} \tag{11}$$

where L_{loc} is the localization distance. The choice of the localization distance represents a challenge. In *Hu et al.* [2012], for assimilation of satellite-derived chlorophyll observations, the localization distance was set up to 100 km. In our case, this is approximately the size of the modeling domain. Through conducted twin experiments, we established that L_{loc} of 10 km was appropriate for our domain. We only used localization in one of our runs (run 4 in section 3). When localization was used, we implement Kalman filter equations (3–4) as a set of independent filters, with each filter updating a single water column. Because we used the box-car localization function (equation (11)), the update for each water column was equivalent to using nonlocalized filter that only accounted for observations within the localization distance L_{loc}:

$$x^a(i_{\text{wc}}) = x^f(i_{\text{wc}}) + K\left(y_{\text{loc}} - H_{\text{loc}}x^f\right)$$

where i_{wc} are the indices of grid points in a given water column, y_{loc} are observations within the localization radius L_{loc}, and H_{loc} is the observational operator that maps the model state of the updated water column i_{wc} onto observations y_{loc}.

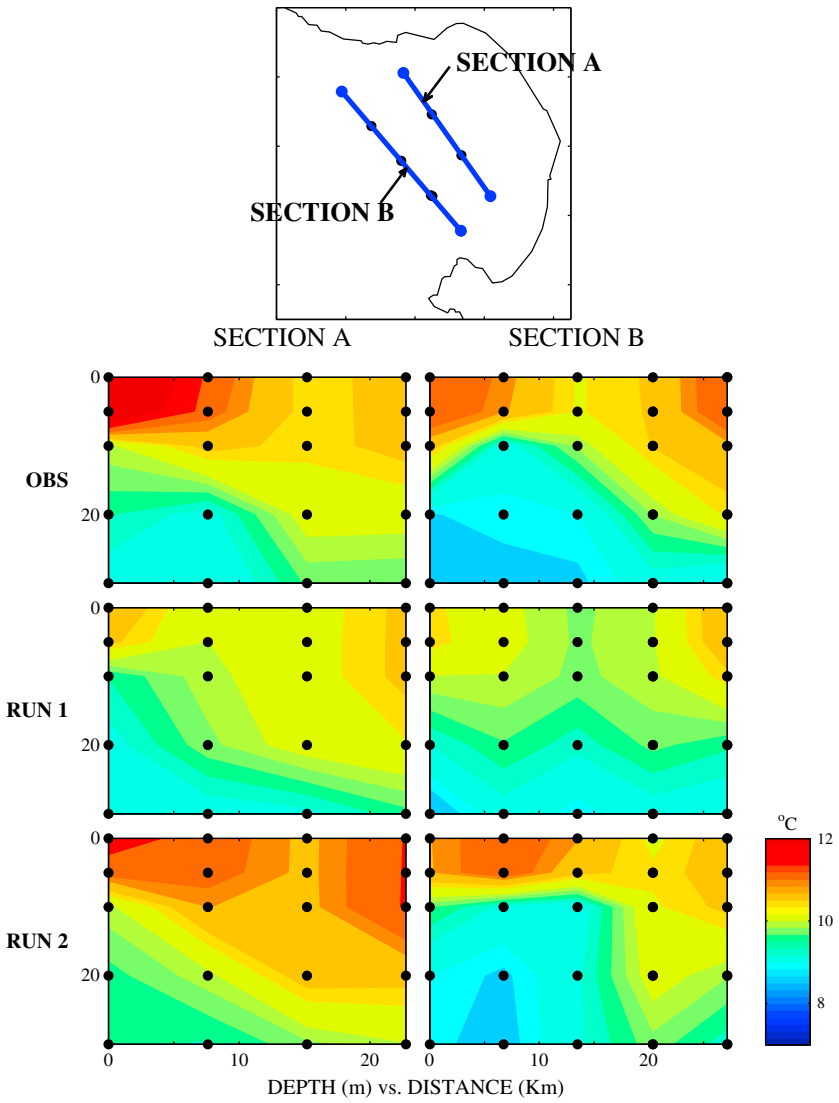

Figure 5. Locations of R/V Point Sur water sample sections A (taken 9 June) and B (taken 10 June) (top insert); observed (second row) and model-predicted (runs 1 and 2) temperature profiles along sections A and B.

3. Bio-Optical, Physical Conditions During Data Assimilation Experiments

[29] Data assimilation experiments described in this study were conducted for the time frame from 5 to 10 June 2008. Observed wind velocities at MBARI moorings (Figure 3) indicate that this period was characterized by steady upwelling winds. At the beginning of the experiment, 33 h low-pass-filtered HF radar surface currents indicate a southward flow along the entrance to the bay that separates a well-defined cyclonic eddy in the Bay and an anticyclonic circulation offshore (Figure 3). Five days later (Figure 3), HF radar data show weakening of the cyclonic circulation. Coincident with this weakening of cyclonic circulation and currents, conditions for phytoplankton growth in the Bay improved as indicated by the increase in surface concentrations of chlorophyll *a* (Figure 3). In accord with Figure 3, the satellite-derived SST images from MODIS-Aqua satellite show development and strengthening of a cold filament along the entrance to the Bay, separating warm, less productive anticyclonic circulation offshore from the more productive waters of the Bay.

4. Design of Data Assimilation Experiments

[30] Table 1 lists the runs and their attributes considered in this study.

[31] Run 1 is the base run of the NCOM ICON model described in section 2.2. The run was initialized from the NCOM CCS model on 22 May 2008 and was run until the end of June without any assimilation of physical or bio-optical observations presented in section 2.1. The output from run 1 (during the month of June) is used to estimate error covariance P^f in accord with section 2.4. All runs described below started from the restart file from run 1 (physical and bio-optical state variables) on 5 June 00Z and were run for 5 days until 10 June 00Z.

[32] Run 2 is the run with the assimilation of physical observations listed in section 2.1.1 with a 12 h data assimilation cycle. Therefore, for each 12 h of the model run,

Table 2. RMSE Between Observed and Model-Predicted Distributions of Temperature and Salinity at Water Sample Sections A and B (Figure 3)[a]

	Temperature		Salinity	
	Section A	Section B	Section A	Section B
Run 1	1.00	1.0	1.00	1.00
Run 2	0.78	0.86	0.35	0.82

[a]RMSE is normalized by the RMSE for the base run 1 (0.9° and 0.06 psu for section A; 0.57° and 0.06 psu for section B).

NCODA assimilated physical observations and created a new restart file (nowcast) with updated (analyzed) temperature and salinity fields. The next segment of the model run was started from this NCODA created nowcast and was run for 12 h until the next model restart file is created. None of the bio-optical data listed in section 2.1.2 were assimilated in run 2. Comparisons of run 2 with the base run 1 highlight the impact of just physical data assimilation on the model predictions of physical, as well as bio-optical properties on time scales of 1–5 days.

[33] Run 3 is the run with the assimilation of physical data as in run 2, but for each 12 h, MODIS-Aqua Chl data (described in section 2.1.2) are assimilated using BOMA (section 2.4). In accord with (3), the only analyzed (updated) bio-optical properties were $P1$ (small phytoplankton) and $P2$ (diatoms). Therefore, for each 12 h of the model run, the NCODA assimilated physical observations and created a new restart file with updated (analyzed) temperature and salinity fields. Using this NCODA created restart file, the BOMA assimilated MODIS-Aqua Chl data and created a new restart file (nowcast) with updated (analyzed) $P1$ and $P2$. The next segment of the model run was started from this BOMA created restart file and was run for 12 h until the next model restart file is created. Comparisons of runs 3 and 1 show the impact of assimilations of physical, as well as MODIS-Aqua Chl data on the model predictions of bio-optical properties. We found that no localization was needed to assimilate MODIS-Aqua Chl data into the model.

[34] Run 4 is a clone of run 3, but the MODIS-Aqua phytoplankton absorption coefficient at 488 nm ($a_{ph}(488)$) data are assimilated in the model instead of the MODIS-Aqua Chl data as in run 3. Unlike run 3, we found that localization was necessary for assimilation of phytoplankton absorption data. Localization distance L_{loc} (in section 2.4.4) was set to 10 km. Comparisons of runs 3 and 4 will provide the impact of the assimilation of surface absorption coefficient versus chlorophyll data on the model predictions of bio-optical properties on time scales 1–5 days.

[35] Run 5 is a clone of run 3. However, the model nitrate is also updated together with the phytoplankton (P1 and P2) through the multivariate data assimilation BOMA in accord with section 2.4. Comparisons of runs 3 and 5 show the impact of also updating nitrate through multivariate assimilation on the model predictions of bio-optical properties.

[36] In the described data assimilative runs 3–5, for each data assimilative cycle (12 h), the assimilation of physical observations (through NCODA) is independent from the assimilation of bio-optical observations (through BOMA). In run 6, we introduced an instantaneous update of the model nitrate based on updated temperature fields (through NCODA). For each data assimilation cycle (12 h), the updated temperature from the NCODA is used to instantaneously update nitrate fields through the observed statistical relations between temperature and nitrate based on the AUV DORADO survey (section 2.1.3) conducted on 3 June prior to the start of the data assimilation experiments (5 June). The updated nitrate field is written into the NCODA-created restart file. Using this NCODA created restart file, the BOMA assimilated MODIS-

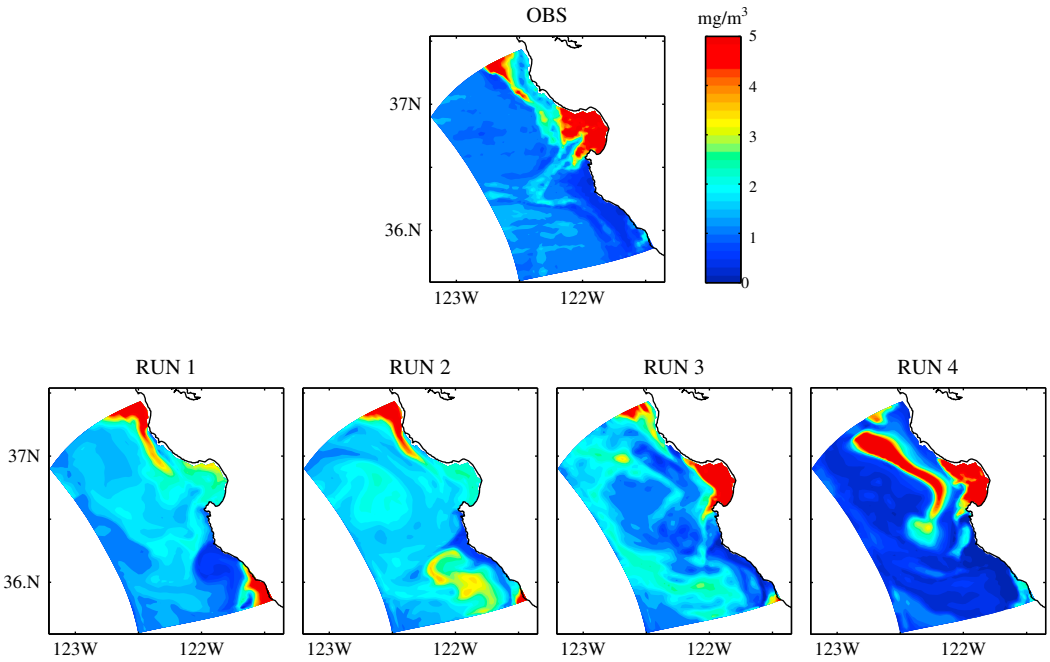

Figure 6. Observed MODIS-Aqua– and model-predicted chlorophyll distributions on 10 June 2008.

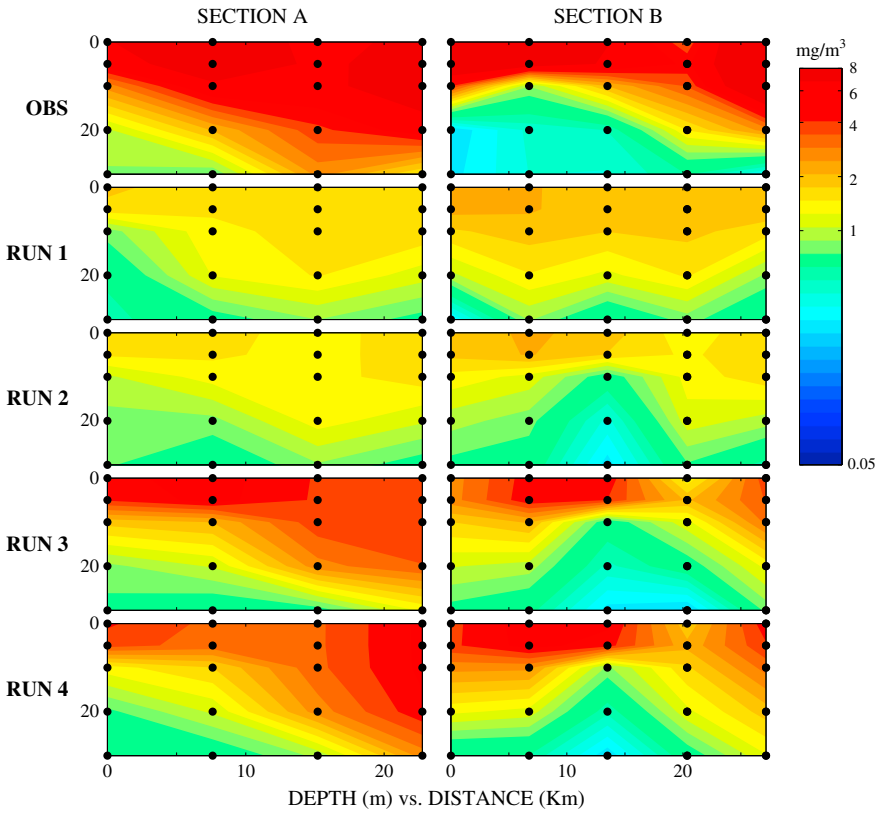

Figure 7. Comparisons of observed (sections A and B, see locations on Figure 3) and model-predicted subsurface chlorophyll distributions at water sample locations.

Aqua Chl data and created restart file (nowcast) with updated (analyzed) P1, P2, and nitrate fields (as in run 5). Comparisons of runs 5 and 6 provide the impact of the instantaneous update of nitrate fields (based on updated physical fields (temperature)) on bio-optical properties predictions.

5. Results

5.1. Assimilation of Physical Data

[37] Figures 4 and 5 provide a comparison of physical properties between runs 1 and 2 (without and with assimilation of physical data, see section 3 and Table 1). There are significant differences in predictions of surface and subsurface physical properties: Run 2 matches much better with observed SSTs (Figure 3), as well as observed subsurface temperature distributions from the water samples (Figure 5). This is also supported by the RMS errors (RMSEs) between observed water samples and model-predicted temperature and salinity fields presented in Table 2. RMSEs for run 2 are reduced by 14%–65% in comparison to the base run 1. Concerning currents, run 2 is also better defined than in run 1 cyclonic circulation in the Bay.

[38] Figure 6 provides a comparison of surface model–predicted chlorophyll distributions for runs 1 and 2. Without the assimilation of MODIS-Aqua Chl, the model predicts much lower chlorophyll values in the Bay for both cases of with (run 2) and without (run 1) assimilation of physical observations.

5.2. Assimilation of Satellite-Derived Bio-Optical Properties

[39] In agreement with satellite observations, the assimilation of MODIS-Aqua Chl increased the model productivity inside the bay and decreased productivity outside the bay for run 3 (Figure 6). The assimilation of $a_{ph}(488)$ (run 4) also increased productivity inside the Bay; however, it also created an artificial tongue of high Chl values offshore from the northern part of the domain along the coast. This might be a result of difficulties in assimilation of offshore values of absorption, which are significantly lower in comparison to the values in the Bay. As stated in section 3, run 4 was done with the localization (see section 2.4.4). This was required to avoid noisy updated fields and to exclude remote $a_{ph}(488)$ observations from the analysis of the local grid point.

Table 3. RMSE Between Observed and Model-Predicted Chlorophyll Distributions at Water Sample Sections A and B (Figure 3)[a]

	Section A	Section B
Run 1	1.00	1.00
Run 2	1.01	1.02
Run 3	0.71	0.95
Run 4	0.65	0.83
Run 5	0.70	0.93
Run 6	0.71	0.94

[a]RMSE is normalized by the RMSE for the base run 1 (5.8 mg/m^3 for section A; 8.6 mg/m^3 for section B).

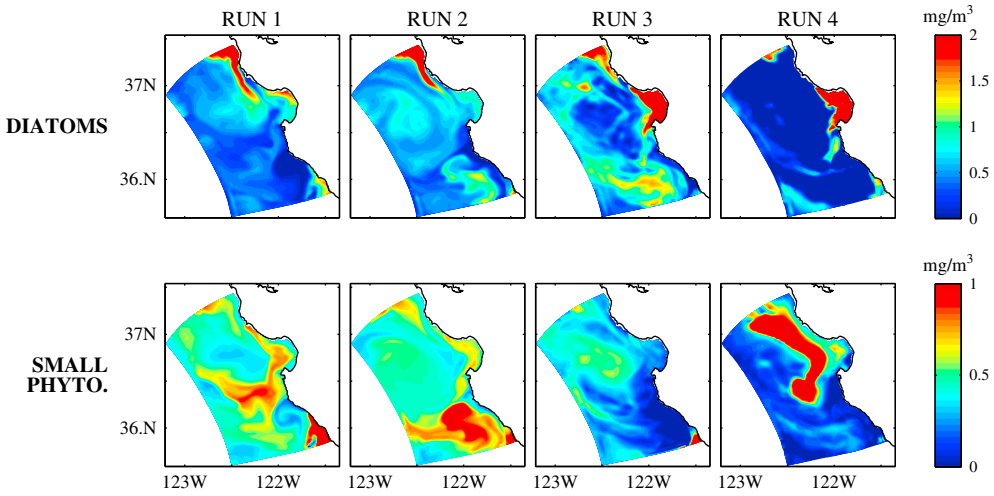

Figure 8. Model-predicted surface diatoms and small phytoplankton distributions.

[40] Figure 7 provides comparisons of the model-predicted subsurface Chl distributions to observed distributions (from water bottle analyses) along sections A and B (recall that chlorophyll data from the water samples were not assimilated). The assimilation of surface MODIS-Aqua Chl improved not only surface (Figure 6) but also subsurface model Chl predictions in the Bay for data assimilative runs 3–4. Quantitatively, this is also reflected in Table 3, where RMSEs between observed Chl from water samples and corresponding model-predicted Chl values (at water sample locations) are presented. All RMSE metrics are normalized by the corresponding RMSE metric for the base run 1 (no assimilation of physical as well as bio-optical properties). Table 3 shows similar values of RMSE metrics for runs 1 and 2. This indicates that while the assimilation of physical observations improved the model predictions of physical properties, the model predictions of Chl are not improved on time scales 1–5 days. Results show that the assimilation of observed surface Chl or $a_{ph}(488)$ provides improvement in subsurface Chl predictions ranging from 5% to 35%. While the assimilation of MODIS-Aqua bio-optical products improved subsurface predictions for runs 3 and 4, the model subsurface predictions of Chl are still underestimated in comparison to the water sample profiles (Figure 7). One of the reasons might be that MODIS-Aqua bio-optical data are assimilated as observed surface values, while satellite data provide an estimate of the average, for example, chlorophyll concentration over the layer between the surface and one attenuation depth. In this case, based on observed profiles on Figure 7, MODIS-Aqua Chl data should somewhat underestimate the "true" surface Chl (this is also illustrated by a comparison of Chl values from the water samples taken at surface and MODIS-Aqua Chl values at water sample locations (comparison is not shown here)). For this reason, assimilation of satellite Chl data (as well as $a_{ph}(488)$) as surface observations should result in underestimated surface and subsurface Chl values in model predictions, which is illustrated in Figure 7.

[41] Assimilation of MODIS-Aqua bio-optical observations increased (decreased) the concentration of diatoms (small phytoplankton) inside the Bay in comparison to nonassimilative runs 1 and 2 (Figure 8). This is supported by comparisons of model predictions with observed fractions of microplankton (analog of diatoms in the model) versus total phytoplankton from HPLC data (section 2.1.3). Comparisons are presented on Figure 9. The HPLC data indicate that there was steady presence of diatoms in the Bay between 5 and 10 June, with the fraction of diatoms to total phytoplankton population in the range of 90%. Runs 1 and 2 show variable fractions of diatoms to the total phytoplankton population ranging from 20% to 80%, but mostly below the observed HPLC fractions.

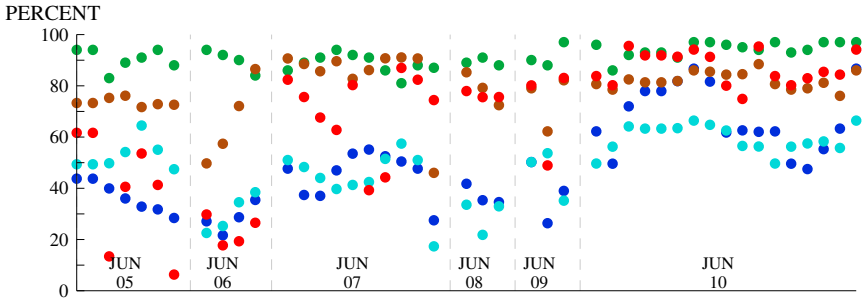

Figure 9. Observed and model-predicted fractions of diatoms to the whole phytoplankton populations at locations of R/V Point Sur water samples. Green, HPLC observed fractions; blue, run 1; light blue, run 2; brown, run 3; red, run 4.

Table 4. RMSE Between HPLC Fractions and Model-Predicted Fractions of Diatoms to Total Phytoplankton Population[a]

	RMSE
Run 1	1.00
Run 2	0.92
Run 3	0.43
Run 4	0.84
Run 5	0.42
Run 6	0.44

[a]RMSE is normalized by the RMSE for the base run 1 (0.52).

However, for run 3 (run with assimilation of MODIS-Aqua surface chlorophyll), the fraction of diatoms increased and partitioning between diatoms and small phytoplankton is in much better agreement with the independent, nonassimilated HPLC observations. This is also reflected in the RMSE metrics presented in Table 4. With the assimilation of MODIS-Aqua Chl data, the RMSE between HPLC observed and model-predicted fractions of diatoms to the total phytoplankton is more than twice smaller for run 3 in comparison to the RMSE for nonassimilative base run 1. There are also improvements in fractions of diatoms to the total phytoplankton predictions for run 4 (assimilation of $a_{ph}(488)$) after a couple days of assimilation (Figure 9 and Table 4).

5.3. Impact on Predictions of Nitrate Distributions

[42] Figure 10 provides comparisons of the observed and model-predicted subsurface nitrate distributions along water sample sections A and B. Runs 1 and 2 without assimilation of MODIS-Aqua Chl data and run 3 with assimilation of MODIS-Aqua Chl data show underestimated values of subsurface nitrate distributions in comparison to water samples. Therefore, while the assimilation of MODIS-Aqua

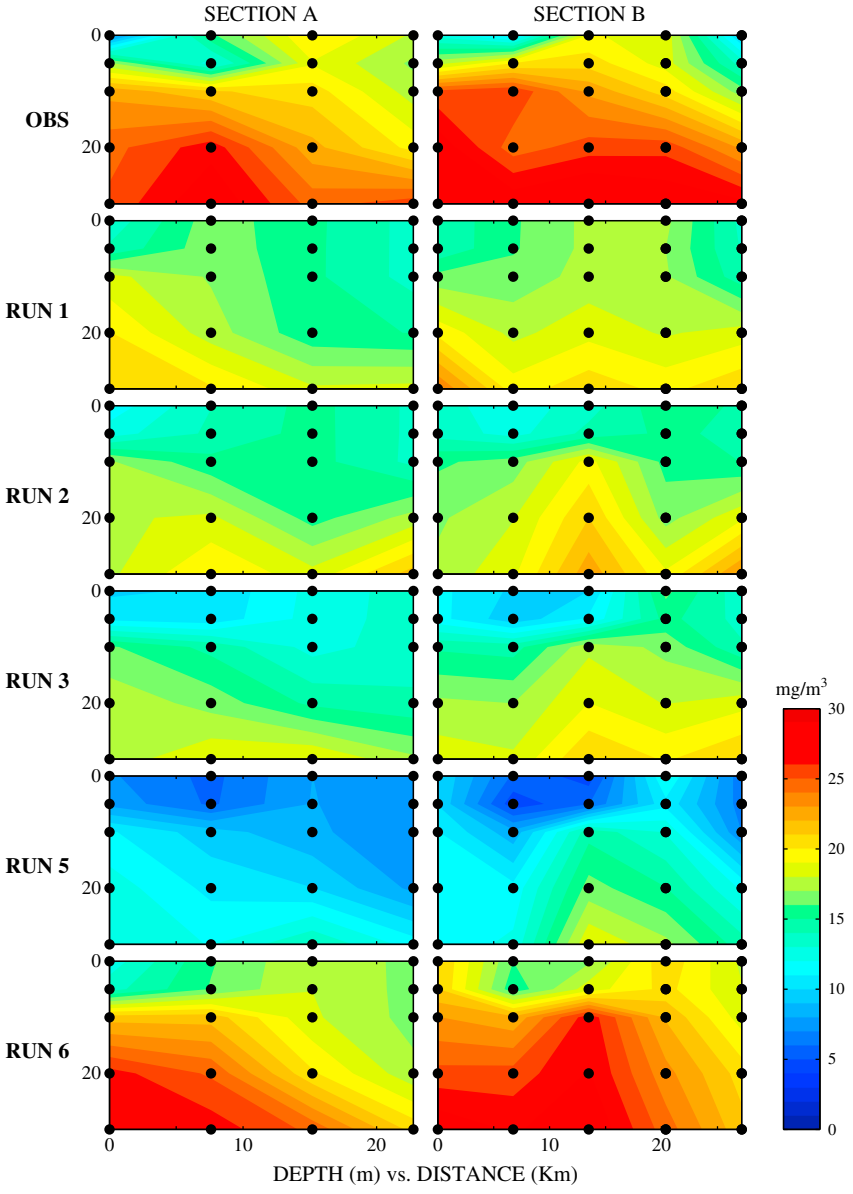

Figure 10. Observed (top row) and model-predicted nitrate distributions for runs 1–6 at water sample sections A and B.

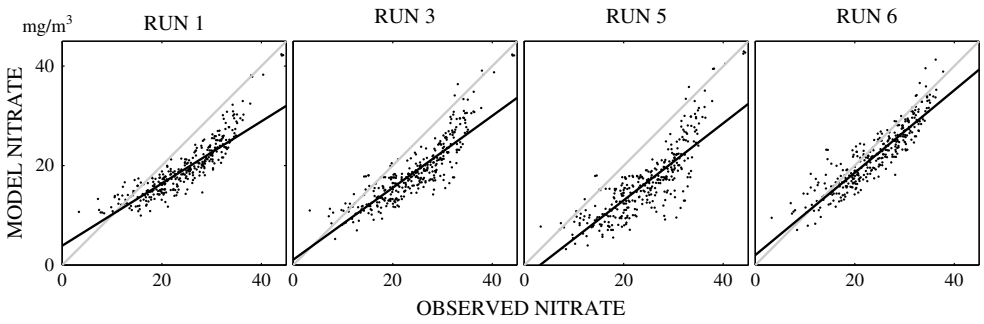

Figure 11. Observed versus model-predicted nitrate for runs 1, 3, 5, and 6 at water sample locations.

Chl improved model subsurface Chl distributions (Figure 7) and partitioning between diatoms and small phytoplankton (Figure 9), it had a minimal impact on nitrate fields in the model. Results are similar for run 4 with the assimilation of $a_{ph}(488)$ (not shown here). For run 5, when phytoplankton (and) and nitrate are updated through the BOMA, the subsurface nitrate distributions are even more underestimated (Figure 10). This is also illustrated by the scatterplots of observed (from water samples) versus the model nitrate fields presented on Figure 11.

[43] As it was demonstrated in section 4.2, the model run 1 without assimilation of MODIS-Aqua Chl underestimates surface and subsurface Chl distributions (Figures 6 and 7). As a result, the assimilation of surface Chl data tends to increase model Chl values and increase phytoplankton population, especially diatom population in the Bay (Figures 8 and 9). However, the increase in the model phytoplankton population results in the decrease of nutrients due to the uptake by phytoplankton for growth, which is statistically inherited in the model multivariate error covariances used

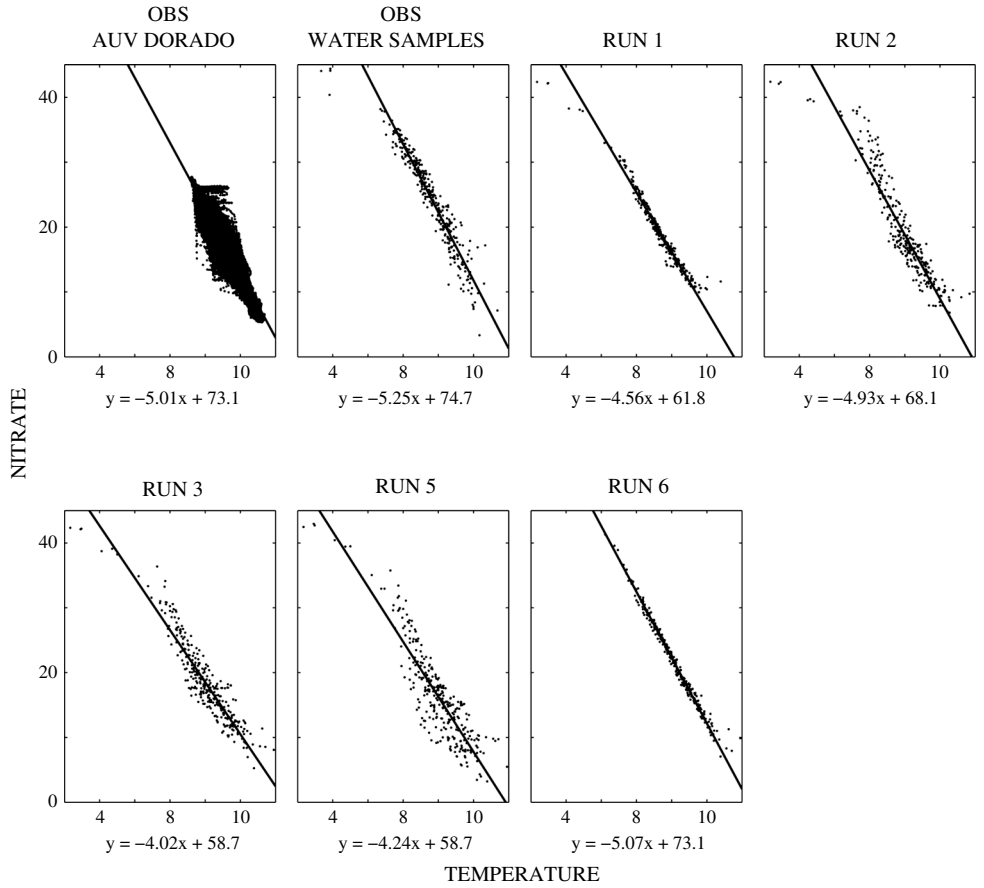

Figure 12. Temperature versus nitrate relations for AUV DORADO, water samples, and model runs.

in the BOMA (section 2.4). This is why the assimilation of MODIS-Aqua Chl increased the model bias in predictions of nitrate values even more for run 5 (Figures 10 and 11), when not only phytoplankton but also nitrate are also updated through the multivariate data assimilation. Note that predictions of Chl and partitioning between diatoms and small phytoplankton for run 5 are similar to run 3 (see Tables 3 and 4).

[44] Figure 12 shows temperature versus nitrate scatterplots of the AUV DORADO survey, from the water samples, and the model runs. The average temperature versus nitrate statistical relation for the AUV survey is very similar to the relation for the water samples, while the AUV survey was taken on 3 June which is about 6 days prior to water samples surveys. This indicates persistence of the same statistical relation between temperature and nitrate on time scales of a week. In run 6 (section 3), for each data assimilation cycle (12 h), the statistical relation between T and nitrate (Figure 12) from the AUV DORADO on 3 June is used to instantaneously update nitrate fields based on the temperature from the NCODA update. The nitrate predictions in run 6 improved significantly and match nitrate observations much better in comparison to other runs (Figures 10 and 11). Therefore, the instantaneous update of nitrate (based on statistical relations between temperature and nitrate) corrected the a priori model underestimation of the nitrate and the reduction of nitrate by the multivariate update. Note that predictions of Chl and partitioning between diatoms and small phytoplankton for run 6 are similar to data assimilation runs 3 and 5 (see Tables 3 and 4).

6. Conclusions and Discussions

[45] Data assimilation experiments were conducted during 5 days of steady upwelling in the Monterey Bay area. The results show that while the assimilation of physical observations improved the model predictions of physical properties, the model underestimates productivity inside the Bay with or without assimilation of physical observations. At the same time, assimilation of MODIS-Aqua–derived optical properties (chlorophyll or absorption due to phytoplankton) significantly improved surface and subsurface agreement between the model and observations. Results show that the reduction in RMSEs between model and independent water samples ranges from 5% to 35% in contrast to the nonassimilative run.

[46] While the assimilation improved the model predictions, the model subsurface Chl distributions retained an underprediction bias as compared to observed profiles from water samples. One of the reasons might be that MODIS-Aqua bio-optical data are assimilated as observed surface values, while satellite data provide an estimate of the average, for example, chlorophyll concentration over the layer between the surface and one attenuation depth. The assimilation of satellite-derived products, not as surface values, but rather as averages over attenuation depth values, is considered as a topic of our future research.

[47] Assimilation of bio-optical data also improved fractionation of phytoplankton biomass between diatoms and small phytoplankton in the model. Without assimilation, the percentage of large diatoms varied during the experiment between 20% and 80%. In contrast, HPLC measurements showed the fraction of diatoms to total phytoplankton population in the range of 90%. However, runs with the assimilation of MODIS-Aqua surface chlorophyll produced much better agreement with the independent, nonassimilated HPLC observations. With the assimilation, the RMSE between HPLC observed and model-predicted fractions of diatoms to the total phytoplankton is less than half smaller than the RMSE for nonassimilative run. There are also improvements in fractions of diatoms to the total phytoplankton predictions for the run with assimilation of $a_{ph}(488)$ after a couple days of assimilation.

[48] To extend of our knowledge, we believe that the present study is the first demonstration of IOP ($a_{ph}(488)$) assimilation into coupled physical, biochemical dynamical model, as well as the first demonstration of a capability to improve the model-predicted fractionation of phytoplankton biomass between diatoms and small phytoplankton.

[49] Model runs with or without assimilation of MODIS-Aqua observations show underestimated values of nitrate distributions in comparison to the water sample observations. The assimilation of MODIS-Aqua observations did not improve the model predictions of nitrate. This can be explained by the fact that multivariate data assimilation tends to increase phytoplankton population in the Bay (due to the underestimated a priori Chl values in the model) and, at the same time, tends to decrease nutrients. The lack of improvements in nitrate distributions in the model suggests deficiencies in the model nitrate initial and open boundary conditions, and the need for nitrate observations for assimilation into the model. These conclusions correlate with results of the *Ourmières et al*. [2009] study. Their goal was an estimation of the basin scale patterns of oceanic primary production and their seasonal variability. *Ourmières et al*. [2009] found that intensive in situ measurements of biogeochemical nutrients are urgently needed at basin scale to improve coupled model predictions. Our results showed that an instantaneous update of nitrate based on statistical relations between temperature and nitrate (derived from the AUV observations taken prior to the data assimilation experiments) corrected the model underestimation of the nitrate fields.

[50] The experiments conducted in this study were limited to a 5-day period during a steady upwelling event. More complicated bio-optical conditions are usually observed during wind weakening and relaxation, when transitions from diatoms to other phytoplankton groups might occur with corresponding drastic changes in bio-optical properties on time scales of days to a week. This might be a combination of changes in physical conditions (for example, dinoflagellates prefer more stable, stratified conditions), as well as changes in nutrient distribution, leading to decreasing diatoms population and replacement by other phytoplankton groups, which are capable of prospering at lower nutrients levels. Also, as demonstrated in *Shulman et al*. [2011, 2012], dinoflagellates play an important role in changes of bio-optical properties during the upwelling events. It was demonstrated that during the upwelling development, dinoflagellates avoided advection and retained their population in the Bay due to their vertical swimming ability. The biochemical model considered here does not include modeling of dinoflagellates dynamics. Inclusion of the dinoflagellates into the biochemical model and conducting data assimilation

experiments during the events influenced by their presence is another topic of our future research.

[51] Our experiments with the ensemble computed from a monthlong model simulation suggest that ensemble methods are very capable at capturing complex multivariate relationships between optical properties, phytoplankton biomass, and ecosystem structure (as represented by small and large phytoplankton pools in the model). Our preliminary experiments encourage further development of ensemble methods for bio-optical data assimilation and uncertainty estimation [*Gould et al*., 2011].

[52] Finally, in the present study, assimilation of physical properties through the NCODA and assimilation of bio-optical properties through BOMA are separated. The adjustment of updated physical and bio-optical variables is achieved through the coupled, bio-optical physical model run during the data assimilation cycle. At the same time, an instantaneous joint update of physical and bio-optical properties is preferred in order to maintain dynamical consistency between the assimilated physical and bio-optical fields [see, for example, *Anderson et al*., 2000, 2001]. The merger of NCODA and BOMA is another topic of our future research.

[53] **Acknowledgments.** This research was funded through the Naval Research Laboratory (NRL) projects, "Bio-optical Studies of Predictability and Assimilation in the Coastal Environment (BIOSPACE)," "Modeling Dynamic Bio-Optical Layers In Coastal Systems (DYaBOLIC)" under program element 61153N, and "Developing Ensemble Methods to Estimate Uncertainties in Remotely Sensed Optical Properties (DEMEN)" under program element 62435. Our thanks also go to Fei Chai and Lei Shi of UMaine for help with the biochemical model. Thanks to all holders of moorings, gliders, AUV data, and to the BIOSPACE members. We thank Ross N. Hoffman of AER and two anonymous reviewers for comments that improved the paper. Computer time for the numerical simulations was provided through a grant from the Department of Defense High Performance Computing Initiative. SF's works was supported by the David and Lucile Packard Foundation and the Office of Naval Research (ONR) under grant N00014-10-1-0424. This manuscript is NRL contribution 7330-12-1495.

References

Anderson, L. A., A. R. Robinson, and C. J. Lozano (2000), Physical and biological modeling in the gulf stream region: I. Data assimilation methodology, *Deep Sea Res., Part 1*, *47*, 1787–1827.

Anderson, L. A., A. R. Robinson, and C. J. Lozano (2001), Physical and biological modeling in the gulf stream region: II. Physical and biological processes, *Deep Sea Res., Part 1*, *48*, 1139–1168.

Barron, C. N., A. B. Kara, P. J. Martin, R. C. Rhodes, and L. F. Smedstad (2006), Formulation, implementation and examination of vertical coordinate choices in the global Navy Coastal Ocean Model (NCOM), *Ocean Modell.*, *11*, 347–375, doi:10.1016/j.ocemod.2005.01.004.

Barron, C. N., A. B. Kara, H. E. Hurlburt, C. Rowley, and L. F. Smedstad (2004), Sea surface height predictions from the Global Navy Coastal Ocean Model (NCOM) during 1998–2001, *J. Atmos. Oceanic Technol.*, *21*, 1876–1893, doi:10.1175/JTECH-1680.1.

Barth A., A. Alvera-Axcarate, J. M. Beckers, J. Staneva, E. V. Stanev, and J. Schulz-Stellenfleth (2011), Correcting surface winds by assimilating high-frequency radar surface currents in the German Bight, *Ocean Dyn.*, *61*, 599–610, doi:10.1007/s10236-010-0369-0.

Bellingham, J. G., K. Streitlien, J. Overland, S. Rajan, P. Stein, J. Stannard, W. Kirkwood, and D. Yoerger (2000), An Arctic basin observational capability using AUVs, *Oceanography*, *13*, 64–71.

Besiktepe, S. T., P. F. J. Lermusiaux, and A. R. Robinson (2003), Coupled physical and biogeochemical data-driven simulations of Massachusetts Bay in late summer: Real-time and postcruise data assimilation, *J. Mar. Syst.*, *40*, 171–212, doi:10.1016/S0924-7963(03)00018-6.

Cane, M. A., A. Kaplan, R. N. Miller, B. Tang, E. Hackert, and A. J. Busalacchi (1996), Mapping tropical Pacific sea level: Data assimilation via a reduced state space Kalman filter, *J. Geophys. Res.*, *101*, 22,599–22,617.

Chai, F., R. C. Dugdale, T.-H. Peng, F. P. Wilkerson, and R. T. Barber (2002), One-dimensional ecosystem model of the equatorial Pacific upwelling system. Part I: Model development and silicon and nitrogen cycle, *Deep Sea Res., Part II*, *49*, 2713–2745, doi:10.1016/S0967-0645(02)00055-3.

Ciavatta, S., R. Torres, S. Saux-Picart, and J. I. Allen (2011), Can ocean color assimilation improve biogeochemical hindcasts in shelf seas?, *J. Geophys. Res.*, *116*, C12043, doi:10.1029/2011JC007219.

Claustre H., et al. (2004), An intercomparison of HPLC phytoplankton pigment methods using in situ samples: Application to remote sensing and database activities, *Mar. Chem.*, *85*(2004), 41–61.

Cossarini G., P. F. J. Lermusiaux, and C. Solidoro (2009), Lagoon of Venice ecosystem: Seasonal dynamics and environmental guidance with uncertainty analyses and error subspace data assimilation, *J. Geophys. Res.*, *114*, C06026, doi:10.1029/2008JC005080.

Cummings, J., et al. (2009), Ocean Data Assimilation Systems for GODAE, *Oceanography*, *22*, 3, 96–109.

Cummings, J. A. (2005), Operational multivariate ocean data assimilation, *Q. J. R. Meteorol. Soc.*, *131*, 3583–3604, doi:10.1256/qj.05.105.

Doron, M., P. Brasseur, and J.-M. Brankart (2011), Stochastic estimation of biogeochemical parameters of a 3D ocean coupled physical-biogeochemical model: Twin experiments, *J. Mar. Syst.*, *87*, 194–207, doi:10.1016/j.jmarsys.2011.04.001.

Doyle, J. D., Q. Jiang, Y. Chao, and J. Farrara (2009), High-resolution real-time modeling of the marine atmospheric boundary layer in support of the AOSN-II field campaign, *Deep Sea Res., Part II*, *56*, 87–99.

Fennel, K., M. Losch, J. Schröter, and M. Wenzel (2001), Testing a marine ecosystem model: Sensitivity analysis and parameter optimization, *J. Mar. Syst.*, *28*, 45–63.

Ford, D. A., K. P. Edwards, D. Lea, R. M. Barciela, M. J. Martin, and J. Demaria (2012), Assimilating GlobColour ocean colour data into a pre-operational physical-biogeochemical model, *Ocean Sci. Discuss.*, *9*, 687–744, doi:10.5194/osd-9-687-2012.

Friedrichs, M. A. M., R. R. Hood, and J. D. Wiggert (2006), Ecosystem model complexity versus physical forcing: Quantification of their relative impact with assimilated Arabian Sea data, *Deep Sea Res., Part II 53*, 576–600.

Fujii, M., E. Boss, and F. Chai (2007), The value of adding optics to ecosystem models: A case study, *Biogeosciences*, *4*, 817–835, doi:10.5194/bg-4-817-2007.

Garcia, H. E., R. A. Locarnini, T. P. Boyer, and J. I. Antonov (2006), *World Ocean Atlas 2005, Volume 4: Nutrients (Phosphate, Nitrate, Silicate)*, edited by S. Levitus, pp. 396, NOAA Atlas NESDIS 64, U.S. Government Printing Office, Washington, D. C.

Gould, Jr., R. W., S. C. McCarthy, I. Shulman, E. Coelho, and J. Richman (2011), Estimating uncertainties in bio-optical products derived from satellite ocean color imagery using an ensemble approach, *Proc. SPIE*, *8175*, 817506–01–817506–10, doi:10.1117/12.897614.

Gregg, W. W., (2008), Assimilation of SeaWiFS ocean chlorophyll data into a three-dimensional global ocean model, *J. Mar. Syst.*, *69*, 205–225.

Hofmann, E. E., and M. A. M. Friedrichs (2002), Predictive modeling for marine ecosystems, *Sea*, *12*, 537–565.

Holm-Hansen, O., C. J. Lorenzen, R. W. Holmes, and J. D. Strickland (1965), Fluorometric determination of chlorophyll, *J. Cons. Cons. Int. Explor. Mer.*, *30*, 3–15.

Hu, J., K. Fennel, J. P. Mattern, and J. Wilkin (2012), Data assimilation with a local Ensemble Kalman Filter applied to a three-dimensional biological model of the Middle Atlantic Bight, *J. Mar. Syst.*, *94*(2012), 145–156.

Johnson, K. S., and L. J. Coletti (2002), In situ ultraviolet spectrophotometry for high resolution and long term monitoring of nitrate, bromide and bisulfide in the ocean, *Deep Sea Res., Part I*, *49*, 1291–1305.

Lee, Z, K. L. Carder, and R. A. Arnone (2002), Deriving inherent optical properties from water color: A multiband quasi-analytical algorithm for optically deep waters, *Appl. Opt. 41*, 5755–5772.

Lee, Z., R. Arnone, C. Hu, P. J. Werdell, and B. Lubac (2010), Uncertainties of optical parameters and their propagations in an analytical ocean color inversion algorithm, *Appl. Opt.*, *49*(3), 369–381.

Lorenzen, C. J. (1966), A method for the continuous measurement of in vivo chlorophyll concentration, *Deep Sea Res.*, *13*, 223–227.

Martinolich, P., and T. Scardino (2011), *Automated Processing System User's Guide Version 4.2*, NRL, Washington, D. C. [Available at http://www7333.nrlssc.navy.mil/docs/aps_v4.2/html/user/aps_chunk/index.xhtml.]

McClain, C. R. (2009), A decade of satellite ocean color observations, *Annu. Rev. Mar. Sci.*, *1*, 19–42.

McGillicuddy, D. J., D. R. Lynch, A. M. Moore, W. C. Gentleman, C. S. Davis, and 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, *Fish. Oceanogr.*, *7*, 205–218.

Natvik, L. J., and G. Evensen (2003), Assimilation of ocean colour data into a biochemical model of the North Atlantic: Part 1. Data assimilation experiments, *J. Mar. Syst.*, *40–41*, 127–153.

Nerger, L., and W. W. Gregg (2007), Assimilation of SeaWiFS data into a global ocean-biogeochemical model using a local SEIK filter, *J. Mar. Syst. 68*, 237–254.

Ourmières, Y., P. Brasseur, M. Lévy, J.-M. Brankart, and J. Verron (2009), On the key role of nutrient data to constrain a coupled physical-biogeochemical assimilative model of the North Atlantic Ocean, *J. Mar. Syst.*, *75*, 100–115, doi:10.1016/j.jmarsys.2008.08.003.

O'Reilly, J. E., et al. (2000), SeaWiFS Postlaunch Calibration and Validation Analyses, Part 3, *NASA Tech. Memo. 2000–206892*, vol. 11, 49 pp., NASA Goddard Space Flight Center, Greenbelt, Md.

Paduan, J. D., K. C. Kim, M. S. Cook, and F. P. Chavez (2006), Calibration and validation of direction-finding high frequency radar ocean surface current observations, *IEEE J. Oceanic Eng.*, *862–875*.

Penta, B., Z. Lee, R. Kudela, S. Palacios, D. Gray, J. Jolliff, and I. Shulman (2008), An underwater light attenuation scheme for marine ecosystem models, *Opt. Express*, *16*, 16,581–16,591.

Rhodes, R. C., et al. (2002), Navy real-time global modeling systems, *Oceanography*, *15*(1), 29–43.

Rousseaux C. S., and W. W. Gregg (2012), Climate variability and phytoplankton composition in the Pacific Ocean, *J. Geophys. Res.*, *117*, C10006, doi:10.1029/2012JC008083.

Ryan, J. P., A. M. Fischer, R. M. Kudela, J. F. R. Gower, S. A. King, R. Marin III, and F. P. Chavez (2009), Influences of upwelling and downwelling winds on red tide bloom dynamics in Monterey Bay, California, *Cont. Shelf Res.*, *29*, 785–795.

Schofield O., J. Kohut, D. Aragon, E. Creed, C. Haldeman, J. Kerfoot, H. Roarty, C. Jones, D. Webb, and S. Glenn (2007), Slocum gliders: Robust and ready, *J. Field Rob.*, *24*(6), 473–485.

Shulman, I., B. Penta, M. A. Moline, S. H. D. Haddock, S. Anderson, M. Oliver, and P. Sakalaukus (2012), Can vertical migrations of dinoflagellates explain observed bioluminescence patterns during an upwelling event? *J. Geophys. Res.*,*117*, C01016, doi:10.1029/2011JC007480.

Shulman, I., M. A. Moline, B. Penta, S. Anderson, M. Oliver, and S. H. D. Haddock (2011), Observed and modeled bio-optical, bioluminescent, and physical properties during a coastal upwelling event in Monterey Bay, California, *J. Geophys. Res.*, *116*, C01018, doi:10.1029/2010JC006525.

Shulman, I., S. Anderson, C. Rowley, S. deRada, J. Doyle, and S. Ramp (2010), Comparisons of upwelling and relaxation events in the Monterey Bay area, *J. Geophys. Res.*, *115*, C06016, doi:10.1029/2009JC005483.

Shulman, I., et al. (2009), Impact of glider data assimilation on the Monterey Bay model, *Deep Sea Res., Part II*, *56*, 128–138.

Shulman, I., J. Kindle, P. Martin, S. deRada, J. Doyle, B. Penta, S. Anderson, F. Chavez, J. Paduan, and S. Ramp (2007), Modeling of upwelling/relaxation events with the Navy Coastal Ocean Model, *J. Geophys. Res.*, *112*, C06023, doi:10.1029/2006JC003946.

Smith, K. W., D. J. McGillicuddy, and D. R. Lynch (2009), Parameter estimation using an ensemble smoother: The effect of the circulation in biological estimation, *J. Mar. Syst.*, *76*(1–2), 162–170.

Smith, K. W., and D. J. McGillicuddy (2011), Dynamical interpolation of surface ocean chlorophyll fields via data assimilation with an iterative ensemble smoother, *J. Mar. Syst.*, *85*, 96–105.

Spitz, Y. H., J. R. Moisan, M. R. Abbott, and J. G. Richman (1998), Data assimilation and a pelagic ecosystem model: Parameterization using time series observations, *J. Mar. Syst.*, *16*, 51–68.

Venrick, E. L., and T. L. Hayward (1984), Determining chlorophyll on the 1984 CalCOFI surveys. *California Coop. Oceanic Fish. Invest. Rep.*, *25*, 74–79.

Vidussi, F., H. Claustre, B. B. Manca, A. Luchetta, and J. C. Marty (2001), Phytoplankton pigment distribution in relation to upper thermocline circulation in the eastern Mediterranean Sea during winter, *J. Geophys. Res.*, *106*, 19,939–19,956.