Available online at www.sciencedirect.com
SCIENCE DIRECT®
Journal of Marine Systems 40–41 (2003) 171–212
JOURNAL OF MARINE SYSTEMS
www.elsevier.com/locate/jmarsys

Coupled physical and biogeochemical data-driven simulations of Massachusetts Bay in late summer: real-time and postcruise data assimilation

Şukru T. Beşiktepe[a,*], Pierre F.J. Lermusiaux[b], Allan R. Robinson[b]

[a] *Institute of Marine Sciences, Middle East Technical University, PO Box 28, Erdemli, Icel 33731, Turkey*
[b] *Harvard University, Division of Engineering and Applied Sciences, Cambridge, MA 02138, USA*

Received 17 June 2002; received in revised form 11 December 2002

Abstract

Data-driven forecasts and simulations for Massachusetts Bay based on in situ observations collected during August–September 1998 and on coupled four-dimensional (4-D) physical and biogeochemical models are carried out, evaluated, and studied. The real-time forecasting and adaptive sampling took place from August 17 to October 5, 1998. Simultaneous synoptic physical and biogeochemical data sets were obtained over a range of scales. For the real-time forecasts, the physical model was initialized using hydrographic data from August 1998 and the new biogeochemical model using historical data. The models were forced with real-time meteorological fields and the physical data were assimilated. The resulting interdisciplinary forecasts were robust and the Bay-scale biogeochemical variability was qualitatively well represented. For the postcruise simulations, the August–September 1998 biogeochemical data are utilized. Extensive comparisons of the coupled model fields with data allowed significant improvements of the biogeochemical model. All physical and biogeochemical data are assimilated using an optimal interpolation scheme. Within this scheme, an approximate biogeochemical balance and dynamical adjustments are utilized to derive the non-observed ecosystem variables from the observed ones.

Several processes occurring in the lower trophic levels of Massachusetts Bay during the summer–autumn period over different spatial and temporal scales are described. The coupled dynamics is found to be more vigorous and diverse than previously thought to be the case in this period. For the biogeochemical dynamics, multiscale patchiness occurs. The locations of the patches are mainly defined by physical processes, but their strengths are mainly controlled by biogeochemical processes. The fluxes of nutrients into the euphotic zone are episodic and induced in part by atmospheric forcing. The quasi-weekly passage of storms gradually deepened the mixed layer and often altered the Bay-scale circulation and induced internal submesoscale variability. The physical variability increased the transfer of biogeochemical materials between the surface and deeper layers and modulated the biological processes.

Keywords: Massachusetts Bay; Real time forecasting; Predictive capability; Interdisciplinary data assimilation; Coastal ecosystem; Summer to fall blooms

1. Introduction

The Massachusetts Bay Sea Trial I (MBST-98) was conducted in Massachusetts Bay from August 17 to

* Corresponding author. Tel.: +90-324-521-2406; fax: +90-324-521-2327.
E-mail addresses: sukru@ims.metu.edu.tr (Ş.T. Beşiktepe), pierrel@pacific.harvard.edu (P.F.J. Lermusiaux).

0924-7963/03/$ - see front matter © 2003 Elsevier Science B.V. All rights reserved.
doi:10.1016/S0924-7963(03)00018-6

October 5, 1998 within the framework of the Littoral Ocean Observing and Predicting System (LOOPS) project. The central scientific focus was phytoplankton and zooplankton patchiness, in particular their relationships to physical variabilities. Synoptic physical and biogeochemical data sets were obtained simultaneously over a range of spatial and temporal scales. Real-time interdisciplinary nowcasting and forecasting were carried out, including data assimilation via optimal interpolation (OI) and error subspace statistical estimation. Field and error forecasts were utilized for adaptive sampling with three research vessels and two fleets of autonomous underwater vehicles (AUVs). The present study is part of this LOOPS project. Details on LOOPS and a summary of MBST-98 are given in Robinson et al. (1999) and Lermusiaux (2001).

Because of the existence of a variety of processes with multiple interactive scales (Section 2), Massachusetts Bay is an excellent test arena for coupled interdisciplinary models. Due to its unique characteristics involving nearly complete enclosure and limited size, it is an ideal basin for real-time forecast studies. Furthermore, an intensive observation and modeling program has been initiated by the Massachusetts Water Resources Authority (MWRA) to assess the effect of the Boston sewage discharge on the regional ecosystem (MWRA, 2001 and references cited therein). One component of this program is the Water Column Monitoring Program, which deals with the variations in physical water properties, nutrient concentrations, dissolved oxygen (DO), phytoplankton biomass, and phytoplankton and zooplankton community composition in Massachusetts Bay. Since 1992, the MWRA has collected a large amount of useful physical and biogeochemical data, which allow the testing and calibration of dynamical models. The results of this program are published by the MWRA's Environmental Quality Department (ENQUAD) in a series of technical reports (MWRA, 2001).

Real-time forecasting of coupled physical and biogeochemical dynamics, and coupled data assimilation are novel. The compatible physical and biogeochemical data sets that were collected during LOOPS provided an opportunity to develop a methodology for real-time coupled modeling and coupled data assimilation, as well as to explore coupled dynamics in Massachusetts Bay. Both methodological and dynamical results are presented here.

The melding of models and data via data assimilation provides estimates of nature, which are better estimates than can be obtained by using the data or the model alone. Importantly for ecosystem studies, dynamical hypotheses can be tested quantitatively. Data assimilation techniques and methodologies are well developed in physical oceanography, and general references include Bennett (1992), Malanotte-Rizzoli (1996), Wunsch (1996), Robinson et al. (1998), and Robinson and Lermusiaux (2001). However, data assimilation in biology and coupled assimilation are new and under development. Robinson and Lermusiaux (2002) provide a comprehensive overview and, additionally, several recent developments are reported in this special volume. Dickey (2002) discusses observational capabilities for interdisciplinary data assimilation. Important early ecosystem studies that assimilate natural ocean data involve flux estimations (V'ezina and Pace, 1994), parameter estimations (Fasham and Evans, 1995; Matear, 1995; Prunet et al., 1996; Spitz et al., 1998; Evans, 1999; Hurtt and Armstrong, 1999; Vallino, 2000; Friedrichs and Hofmann, 2001; Spitz et al., 2001; Friedrichs, 2002), and field estimations (Armstrong et al., 1995; McGillicuddy et al., 1998; Lynch et al., 1998). Anderson et al. (2000) and Anderson and Robinson (2000) have carried out coupled physical–biogeochemical hindcasting studies in the Gulf Stream region with data assimilation. They found that the assimilation of both physical and compatible biogeochemical fields was necessary for the success of their model simulations.

The main technical objective here is to describe and evaluate real-time and postcruise data-assimilative numerical simulations, which are carried out in Massachusetts Bay based upon observations collected during August–October 1998 and on four-dimensional (4-D) physical–biogeochemical models. The methodological advances and the dynamical results are described as they are obtained. Specifically, the use of a priori constraints, including approximate dynamical balances and data bounds on model structures and parameters, is shown to be important, both during the model calibration and within the assimilation scheme itself. The skill of the new ecosystem model is evaluated before and

after data assimilation. For the case without data assimilation, it is found that, for good-enough initial data coverage and external atmospheric forcing analyses, the internal predictive capability remains significant after a couple weeks. For the dynamics, the goal is to illustrate and investigate some of the buoyancy-driven and atmospheric-driven ecosystem variability. During MBST-98, the buoyancy and atmospheric forcings, and their interactions, are found to strongly influence the ecosystem of the Bay, including the transition from summer to fall conditions. On days-to-weeks average, the Bay-scale buoyancy-driven circulation is cyclonic, but mesoscale and submesoscale variabilities are found to be vigorous and diverse, which modulates the biological scales and processes. Successive storm systems pass over the Bay on quasi-weekly time scales and these weather events also lead to rapid coupled physical–biogeochemical variations.

In what follows, the Massachusetts Bay ecosystem is reviewed (Section 2). The coupled dynamical models, data assimilation approach, and data sets are then outlined (Section 3). The specifics and several technical advances are provided in the appendices. The real-time forecasts and their results are presented and discussed in Section 4.1, and the postcruise data-driven simulations (hindcasts) and their results in Section 4.2. In the latter, special emphases is put on the initialization of model fields and parameters, and on data assimilation procedures. Based on these real-time and postcruise simulations, the ecosystem evolution and selected coupled processes are described. A summary of results is provided in Section 5.

2. Massachusetts Bay ecosystem: scales, processes, and dynamics

Massachusetts Bay (hereafter "the Bay") is a small semienclosed embayment (about 50×100 km) bounded on three sides by coasts and located at the western boundary of the Gulf of Maine (Fig. 1). Stellwagen Bank, which rises to within 25–30 m of the sea surface, limits the connections to the Gulf of Maine. The average depth of Massachusetts Bay is 35 m; the maximum depth of 80 m is reached in the Stellwagen Basin. The southern part of the Bay, Cape Cod Bay, is shallow (less than 50 m). In the northwestern part, the small embayment east of Boston, called Boston Harbor, has a surface area of 125 km^2 and an average depth of 5 m.

The predominant horizontal flow (Fig. 1) in Massachusetts Bay is a Bay-scale cyclonic circulation mainly driven by the Maine coastal current and atmospheric forcings over the Gulf of Maine (Geyer et al., 1992; Signell et al., 1993; Brown, 1998). However, as was observed and analyzed during MBST-98 (Robinson et al., 1999; Lermusiaux, 2001), this cyclonic circulation is far from permanent (e.g., Signell and List, 1997). It is altered by the local wind forcing and is influenced by buoyancy forcings associated with the variability of the Gulf of Maine (Wallace and Braasch, 1997; Werme and Hunt, 2000; Gangopadhyay et al., 2002). For example, during several days of the summer, strong winds from the south usually occur (e.g., Libby et al., 1999). These southerly winds can reverse the direction of the flow because coastal upwelling creates thermal-wind-driven rim currents. Such events were forecast and observed during MBST-98. It was also found that three branches of the Gulf of Maine coastal current could occur (Fig. 1): one is the coastal current around the whole Massachusetts Bay, one enters the northern half of the Bay but not Cape Cod Bay, and one flows along Stellwagen Bank, without entering the Bay at all. Each of these branches may or may not be present at any time. Two gyres are often present, one in Cape Cod Bay and another to the north of Stellwagen Basin, but their size, position, and sense of rotation are variable. The mesoscale and submesoscale variabilities of these circulation features were also found to be more energetic than previously described.

Nutrients are advected into the Bay by the Gulf of Maine coastal current and its river inputs. The local sources of nutrients are the rivers discharging into Boston Harbor. The total nitrogen load in the harbor is about 6.6×10^5 kmol $year^{-1}$ (Kelly, 1998). A small portion (0.2×10^5 kmol $year^{-1}$) of this load is removed via denitrification to the sediments and burial (Nowicki et al., 1997), but most of the nitrogen entering the harbor is exported to the Bay (Kelly, 1997) because the high rates of anthropogenic nutrient inputs to the harbor exceed the capacity of the sedi-

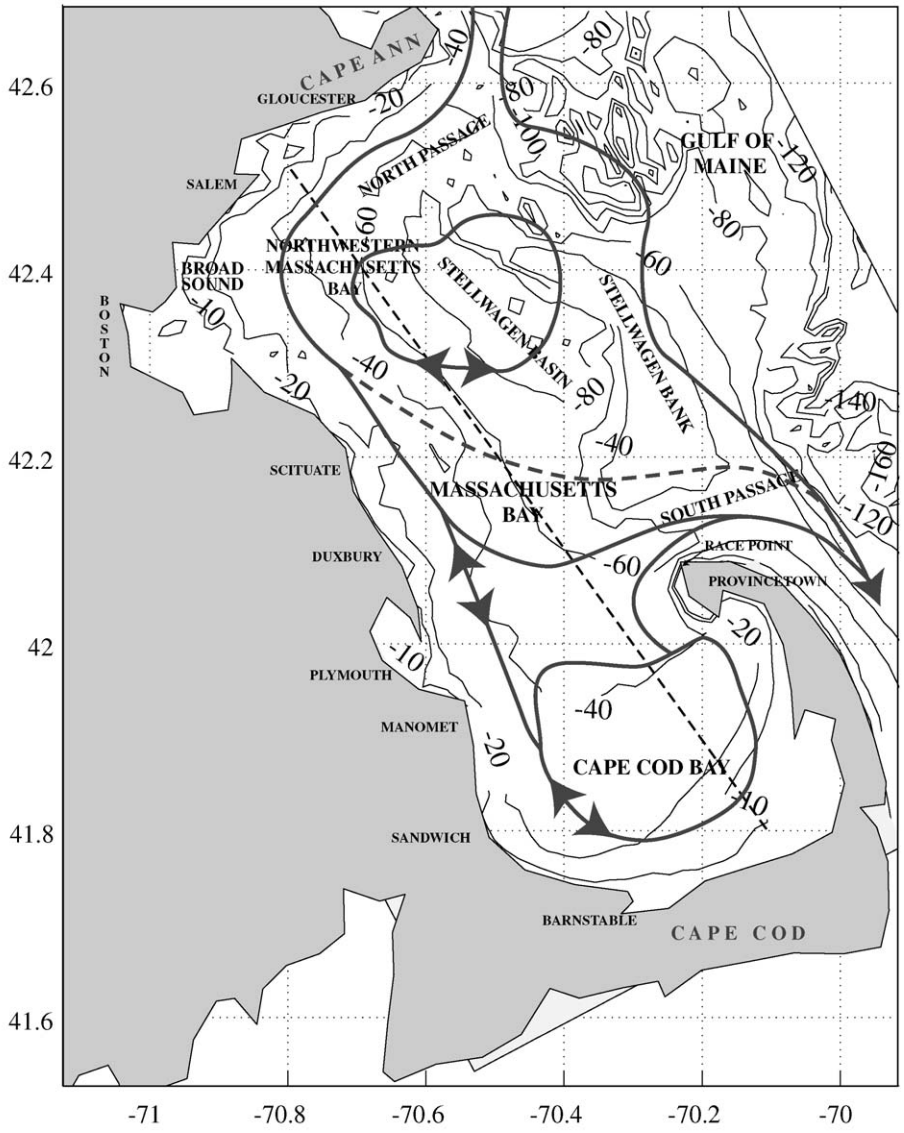

Fig. 1. Cartoon of horizontal circulation patterns for stratified conditions in Massachusetts Bay, overlying topography in meters (thin lines), after Lermusiaux (2001) and Robinson and Lermusiaux (2001). The patterns drawn correspond to main currents in the upper layers of the pycnocline where the buoyancy-driven component of the horizontal flow is often the largest. These patterns are not present at all times. The most common patterns are in solid lines, the less common are dashed. The black dashed line from south to north along the Bay is the position of all vertical cross-sections shown in the text. Geographical locations and topographic features mentioned in the text are also shown.

ment to remove nutrients through denitrification and/or burial (Giblin et al., 1997). The harbor's water residence time is short, on the order of days (Kelly, 1998). Considering annual scales, Massachusetts Bay is relatively well mixed from November to April and the whole water column is then rich in nutrient. Following the spring bloom, the water is stratified. The summer–early autumn season is thus usually characterized by low nutrient concentrations in the surface layers. During May through October, nutrients below the thermocline steadily increase due to regeneration processes, sinking of detritus, and lack of

vertical mixing. The stratification of the water column also leads to a decrease of dissolved oxygen in bottom waters. However, the rate of oxygen decline is about 0.025 mg l^{-1} day^{-1} and the minimum dissolved oxygen measurements are around 4–6 mg l^{-1} (Kelly and Doering, 1999): the decrease in DO is not sufficient to cause hypoxia (Werme and Hunt, 2000).

The annual cycle of phytoplankton in the Bay includes biomass peaks in winter–spring and fall, which is characteristic of temperate midlatitude coastal waters (Keller et al., 2001; Kelly and Doering, 1997). Variable production occurs in summer, in part due to southerly wind events, which trigger coastal upwelling blooms (Libby et al., 1999; Robinson et al., 1999). Annual primary production ranged from a low of 160 g C m^{-2} $year^{-1}$ to a high of 787 g C m^{-2} $year^{-1}$ from 1992 to 1997 (Keller et al., 2001). Diatoms and microflagellates are numerically the most abundant phytoplankton assemblages throughout the year (Libby et al., 1999). Diatoms are the major fraction of the biomass. The winter–spring bloom is dominated by diatoms. Since the whole water column is relatively rich in nutrients from November through April, the bloom is light-limited. However, grazing pressures on phytoplankton can suppress this winter–spring bloom as observed during 1998 (Libby et al., 1999; Werme and Hunt, 2000; Keller et al., 2001). After the spring bloom, high chlorophyll *a* (Chl-*a*) concentrations are restricted to subsurface layers throughout most of the summer. Microflagellates reach their peak abundance during the summer. Small athecate microflagellates are usu-

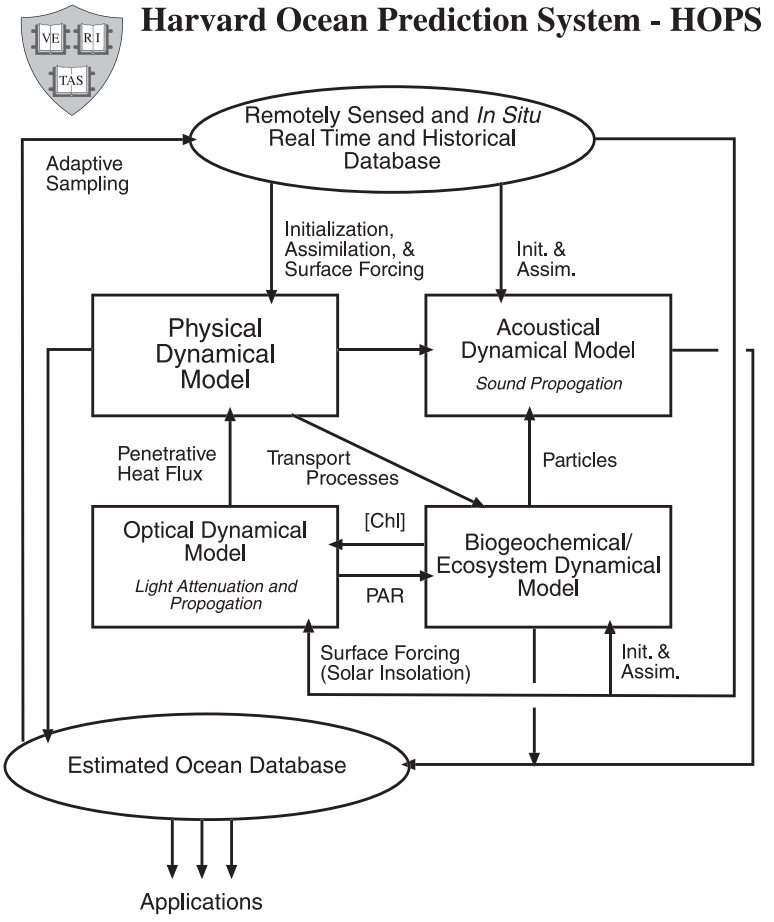

Fig. 2. Schematic of the HOPS.

ally the dominant phytoplankton during August. In October, diatoms become again dominant, along with microflagellates and cryptomonads (Turner, 1992). The MBST-98 experiment occurred during this summer-to-fall transition.

The zooplankton assemblage is dominated by copepods and copepod nauplii (Turner, 1992). These zooplankton communities are typical of North Atlantic coastal assemblages, which are advected into the Bay by the Gulf of Maine coastal current. However, local zooplankton populations are possible (Turner, 1992). During 1998, the zooplankton abundance peaked in May near the coast and remained moderately high until December, but did not reach the high values measured on many other years (Werme and Hunt, 2000).

3. Estimation approach

3.1. Models and data assimilation methodology

The Harvard Ocean Prediction System (HOPS) was used for the physical–biogeochemical forecasting during MBST-98. Its main dynamical components are illustrated on Fig. 2. The system consists of coupled dynamical models, statistical models, initialization procedures, data assimilation schemes, and various visualization and postprocessing tools (Robinson, 1996; Lozano et al., 1996; Haley, 1996; Lermusiaux, 1997; Djurcilov et al., 2002). HOPS is a portable and generic system for interdisciplinary nowcasting, forecasting, and data-driven simulations. It has been applied to many regions including shipboard forecast experiments with verification of skill (Robinson, 1999; Lermusiaux et al., 1999; Robinson et al., 2002).

Presently, the physical model is based on 4-D primitive equation (PE) dynamics (see Appendix A.1 for details). The biogeochemical model coupled to the physical model consists of 4-D advection–diffusion-reaction (ADR) equations for seven state variables governing phytoplankton, zooplankton, detritus, nitrate, ammonium, and chlorophyll *a*. It was, in large part, designed for the MBST-98 experiment and it is schematized in Fig. 3. The development of the new coupled modeling system, including methodologies for the calibration of the biogeochemical model structures and parameters, is an essential component of the present study. Several of the modeling

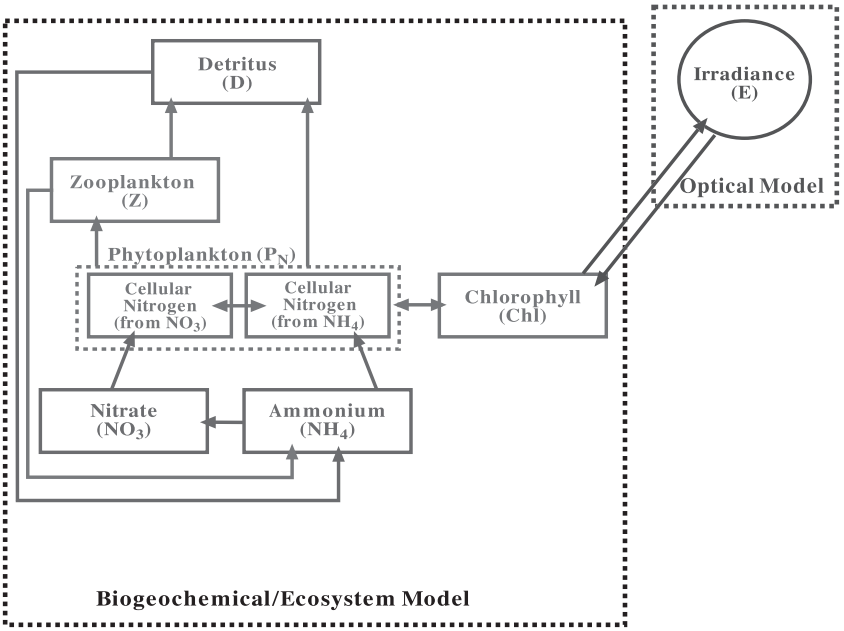

Fig. 3. Schematic representation of the seven-compartment ecosystem model.

advances and improvements are mentioned in Section 4, but critical details on the technical results are presented in Appendix A. These includes the new biogeochemical equations and their parameters (Appendix A.2, Tables 1–2), and the new vertical interdisciplinary boundary conditions (Appendix A.3).

Data assimilation strategies in a forecast system provide the means for model initialization and update, melding of model fields and primary data, tuning of model parameters, and provision of error estimates. Here the optimal interpolation scheme of HOPS is employed to assimilate the synoptic data (Section 3.2) in the ecosystem simulations. The data forecast melding step of this OI scheme consists of a two-scale objective analysis (OA) of the observations, followed by a blending of the forecast with the OA fields (Robinson, 1996; Lozano et al., 1996; Lermusiaux, 1999). Further details are provided in Appendix A.4.

This OI scheme is directly applicable to the state variables for which there are data. However, since not all variables are observed (Appendix B), not all variables can be directly objectively analyzed and then assimilated in accord with error fields. It is nonetheless necessary to infer estimates for the non-observed variables that are compatible with the observed ones, in part to avoid dynamical shocks. The new interdisciplinary procedure (Section 4) that was developed for this study computes a first guess at

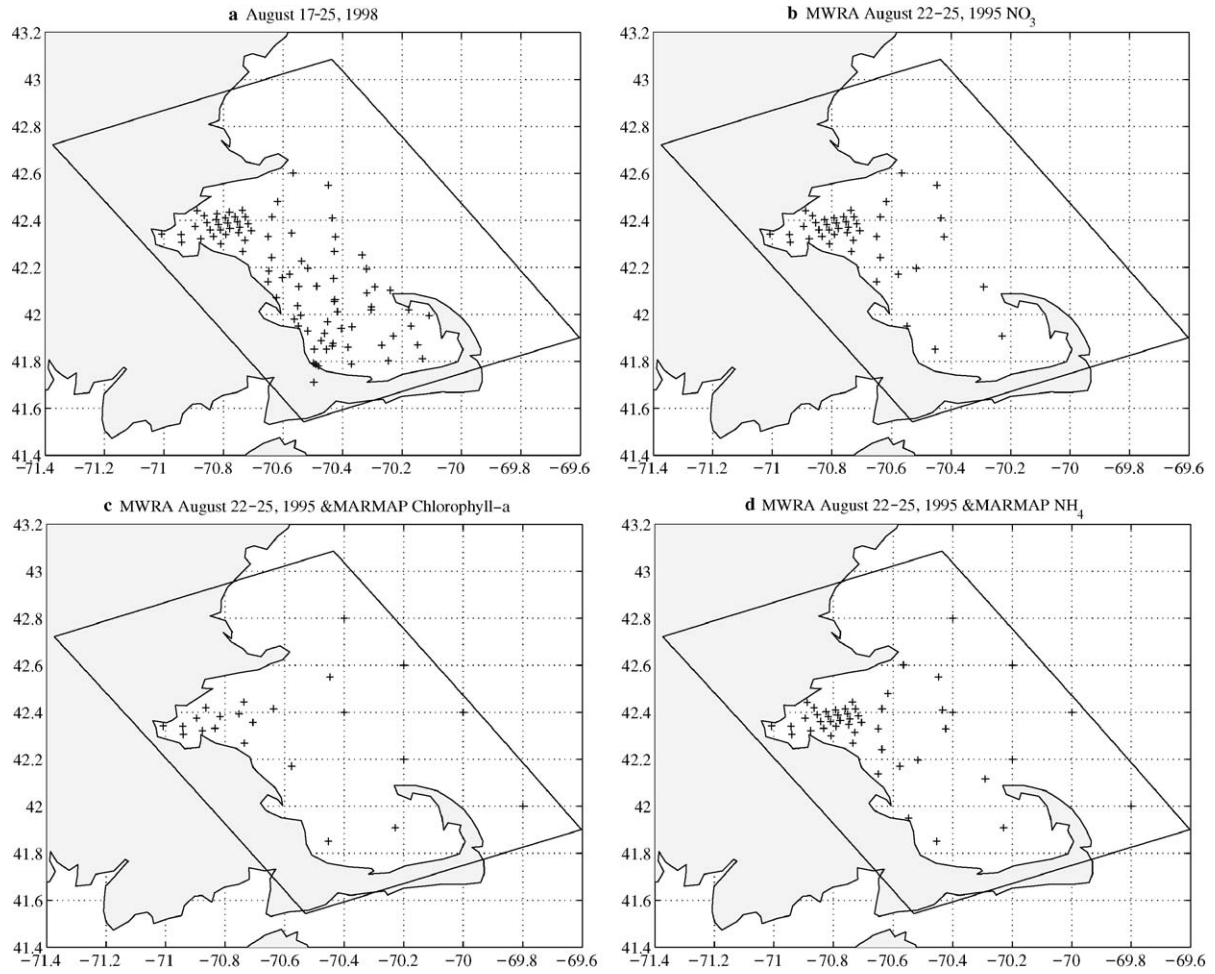

Fig. 4. Model domain and position of sampling stations used to initialize the model. (a) Hydrographic data were collected in real time during MBST-98 (August 17–21, 1998) and MWRA cruises (August 18–25, 1998) (b–d). Biochemical data are MWRA data collected in August 1995.

a non-observed three-dimensional (3-D) field based on approximate relationships between the observed and non-observed variables: the observed variables are first objectively analyzed and the relationships are then used to deduce the first guess at the non-observed 3-D fields. These first-guess fields are then dynamically adjusted to yield initial conditions or assimilation fields, which are consistent with all available data and dynamics.

3.2. Data sets: real-time and postcruise

There are important differences between the data sets used for the real-time and postcruise simulations. The processing and calibration of the biogeochemical components of the MBST-98 and August–October 1998 MWRA data sets were only completed after the real-time experiments. In the real-time simulations (Section 4.1), no biogeochemical data were thus assimilated. However, the August–October 1998 physical data were utilized and assimilated in the PE model, which forced the biogeochemical model. The calibration and initialization of this real-time biogeochemical model were based on the August 1995 MWRA data. In the postcruise simulations (Section 4.2), both physical and biogeochemical data are assimilated. The postcruise biogeochemical model is calibrated and initialized based on these August–October 1998 data. Some historical data are also used to initialize offshore regions.

A summary of all historical and synoptic data sets employed and the references for details on these data sets are provided in Appendix B. The data utilized for the real-time initialization are described next. For the postcruise simulations, the initialization data are outlined in Section 4.2.1 (Fig. 11) and the assimilation data for September 2–4 are outlined in Sections 4.2.2 and 4.2.3 (Fig. 15).

The positions of the hydrographic stations and the corresponding composite profiles, which were used to initialize the physical model in real-time, are shown on Figs. 4a and 5a and b, respectively. These *T* and *S* data were collected during the August 17–21 MBST-98 and August 18–25 MWRA surveys. The positions of the MWRA August 1995 biogeochemical stations are given in Fig. 4b–d; the composite vertical profiles are shown on Fig. 5c–e. The rationale used to choose this August 1995 data set for the initialization of the

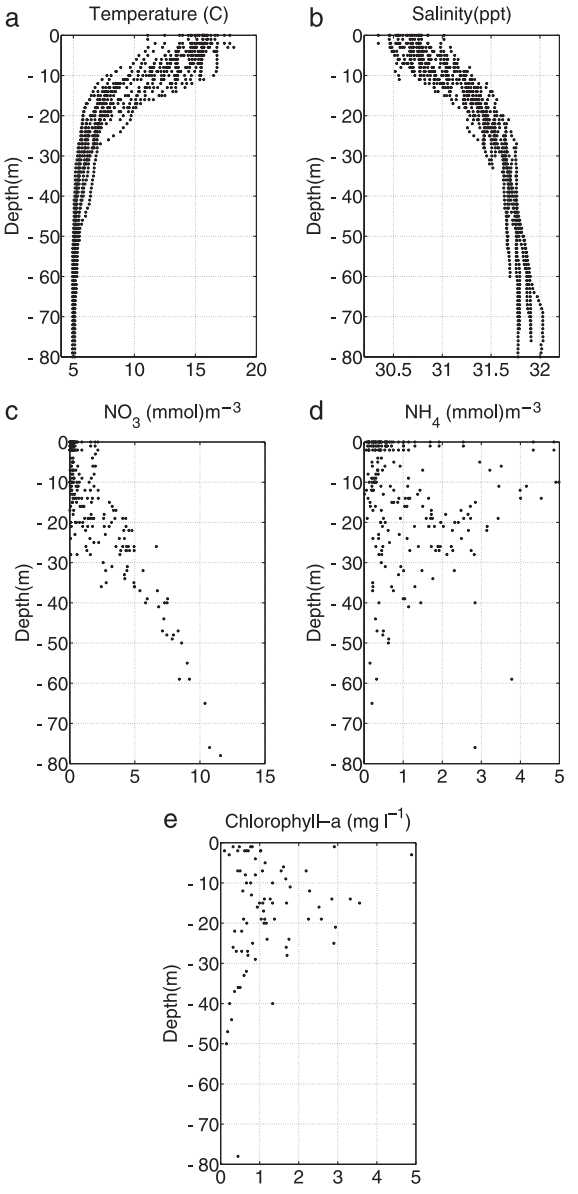

Fig. 5. Composite profiles of (a) temperature (°C); (b) salinity (ppt); (c) nitrate (μmol N l^{-1}); (d) ammonium (μmol N l^{-1}); (e) chlorophyll *a* (mg Chl m^{-3}) used in the real-time model initialization (station positions shown on Fig. 4).

August 1998 biogeochemical state is important. It is based on our approach, which uses historical synoptic data that most likely reflect the anticipated real-time situation. In the present case, the year 1995 was selected because the average primary production dur-

ing February–April 1995 was low and similar to that of February–April 1998 (based on a preliminary analysis of MWRA data for that spring). One could thus expect that the properties of the real-time data to be collected during late August 1998 were going to be similar to those of the August 1995 data, at least on the Bay scales. Therefore, the August 1995 data were used for the initialization. Based on a recent detailed study by Keller et al. (2001), this real-time choice can be confirmed as being relatively good.

During late summer, the water column is well stratified (Fig. 5a and b). This stratification was especially pronounced in 1998 (Libby et al., 1999; Keller et al., 2001). Although there are spatial variabilities in the data (Fig. 5), vertical profiles have similarities over the whole basin. On average, in late summer 1998, the mixed layer extended to a depth of 2–10 m and the thermocline and halocline are located between 2 and 20 m. The nutricline starts around 15–20 m, with nitrate concentration increasing from 1–2 to 12 μmol N l^{-1} near the bottom. The Chl-*a* maxima are located around 15 m and the ammonium maxima around 20 m. This picture is in agreement with the expected biogeochemical processes occurring in the water column. During this late summer period, Chl-*a* concentrations were relatively low in surface waters. Subsurface Chl-*a* maxima occurred near the middle of the thermocline, usually in synchronization with physical submesoscale and mesoscale variabilities. At some locations, relatively pronounced maxima were observed. Along the coastline and in the shallows of Cape Cod Bay, Chl-*a* maxima were near or at the bottom. Overall, all biogeochemical concentrations are low near the surface, except around Boston Harbor where ammonium reaches its maximum.

4. Results

4.1. Real-time forecasts and evaluations

Based upon data assimilation, the coupled physical–biogeochemical models were used to construct realistic physical fields and robust biogeochemical fields. For the physical predictions, detailed skill was an objective, from the mesoscales to the Bay scales. It was found that the accuracy of the physical forecasts and the duration of this accuracy (i.e., limit of predictive capability) varied with the strength and accuracy of the atmospheric forcings. In summary, for medium to strong atmospheric forcings, the accuracy of the ocean forecasts was largely dependent on the accuracy of the US Navy Fleet Numerical Meteorologie and Oceanographic Center (FNMOC) fields. Most of the time, the 1- to 2-day FNMOC forecasts were then good and so were the ocean forecasts. Beyond 2 days, errors in the predicted atmospheric forcings led to ocean predictions of low to zero skill, or phase errors occurred in the prediction. For the periods of weaker atmospheric forcings, which usually occurred in between wind events and lasted for about a week, the situation was different. Because of the longer internal predictability limit and PE model accuracy, the physical ocean forecasts were then good for 5–10 days. More details are presented in Robinson et al. (1999), Lermusiaux et al. (1999), and Lermusiaux (2001).

For the ecosystem predictions, detailed skill was not a real-time objective because calibrated biogeochemical data were not available in real-time, not even for initialization (Section 3.2), and because the four-dimensional dynamics of the summer-to-fall transition was relatively unknown. The goals were thus to: (i) simulate the overall behavior of Bay-scale biogeochemical responses to physical forcings and to produce robust biogeochemical forecasts (Section 4.1.3); and (ii) provide an operational methodology for rapid ecosystem field and parameter estimates based on a relatively simple initialization procedure (Sections 4.1.1 and 4.1.2).

4.1.1. Real-time calibration and selection of model parameters

The biogeochemical model (Appendix A.2) requires the specification of 30 parameters (Table 2 in Appendix A.2), which define the transfer of nitrogen between the state variables. For the real-time simulations, values for primary production parameters (PAR:total surface irradiance, photosynthesis rate, initial slope of *P*/*I* curve) and Chl-*a* equation parameters (Chl:C ratio, N:C ratio of phytoplankton) were obtained from past measurements in Massachusetts Bay (Kelly and Doering, 1997; Murray et al., 1997; Werme and Hunt, 2000). Pure water light attenuation scale is a constant and has a value of 0.038 m^{-1} (Wroblewski, 1977). The value

of the Chl-*a* (phytoplankton) light attenuation scale is given within 0.02–0.04 m^2 (mg Chl)$^{-1}$ by Fasham et al. (1990) but can be as high as 0.07 m^2 (mg Chl)$^{-1}$ in coastal regions (Wroblewski, 1977). The rest of the model parameters listed in Table 2 (see Appendix A.2) were obtained from modeling studies in the eastern coast of the US (Kremer and Nixon, 1978; Hofmann and Ambler, 1988) and the open ocean modeling study of Fasham et al. (1990). The parameters related to the life cycles of phytoplankton and zooplankton are highly variable and depend on the species history and environmental conditions. Ranges of values commonly used in the literature and summarized in Fasham et al. (1990) and Kremer and Nixon (1978) are given in Table 2. The values employed here were chosen based on these ranges, on the analyses described in (Kremer and Nixon, 1978), and on the sensitivity studies described below. The values of the settling (sinking) rates for phytoplankton (detritus) and the parameters involved in the exchanges between the water column and the sediments are usually difficult to find in the literature. For the real-time simulations, these parameter values were determined based on simple sensitivity studies (in the postcruise simulations, the formulations of the bottom boundary conditions were improved).

Real-time simulations were carried out to select some ecosystem model parameters and to assess the sensitivity of the model to biogeochemical factors. To do so, the biogeochemical model was initialized over the whole domain based on the available profiles (Section 3.2) but only one-dimensional (z, t) versions of the model were run forward, for 1 month. During such runs with different parameters, the model was only forced with short-wave radiation. The aim was to determine a set of parameters adequate for the 1998 summer time biota and interactions. The parameters without reference in Table 2 were obtained from these calibrations.

4.1.2. Real-time initialization

The real-time initialization is carried out in two phases. For a first-guess estimate of the initial physical and biological fields, the observed variables are mapped into 3-D fields by objective analyses (Appendix A.4) and the non-observed 3-D fields are computed from these objective maps, using approximate relationships between observed and non-observed variables. These first-guess fields are then dynamically adjusted to yield initial conditions with balanced velocities and biogeochemical fields.

Table 3
Real-time biogeochemical model initialization

	Fields	Initialization procedure
Observed variables (historical data)	Chl(x,y,z)	(i) Calibrate fluorometer data to Chl
		(ii) OA of resulting Chl data
	NO_3(x,y,z)	OA of observed NO_3 data
	NH_4(x,y,z)	OA of observed NH_4 data
Non-observed variables (computed from observed ones)	P(x,y,z)	$\mathrm{Chl}(x,y,z)\theta_{\mathrm{Chl}}^{\mathrm{C}}\theta_{\mathrm{C}}^{\mathrm{N}}/12$
	P_{NO_3}	$P \times NO_3(x,y,z)/(NO_3(x,y,z) + NH_4(x,y,z))$
	P_{NH_4}	$P \times NH_4(x,y,z)/(NO_3(x,y,z) + NH_4(x,y,z))$
	Z(x,y,z)	$0.5 \times P$
	D(x,y,z)	$0.05 \times P$

The first guess at the initial physical fields was obtained for August 17 by assuming velocities in geostrophic balance with the objective analyses of the T and S data (Section 3.2) and by imposing a barotropic transport streamfunction along open boundaries (Lermusiaux, 2001). The first-guess at the initial NO_3, NH_4, and Chl fields (Table 3) was computed by objective analyses of historical data (Section 3.2). The remaining biological fields (Table 3) were derived from these objectively analyzed data. For the real-time simulations, phytoplankton biomass in terms of nitrogen was computed from Chl-*a* using carbon-to-chlorophyll (C:Chl) and carbon-to-nitrogen (C:N) ratios. The C:Chl ratio is highly variable in the literature and an average estimate of 40 mg C (mg Chl)$^{-1}$ was used in our real-time calculations (Table 2). This value was chosen according to past POC and Chl-*a* measurements in the basin (Libby et al., 1999).

The C:N ratio ($1/\theta_{\mathrm{C}}^{\mathrm{N}}$ in Table 2) was taken as 6.625 μmol C (μmol N)$^{-1}$ (Redfield et al., 1963). The zooplankton data available were sparse and in the form of abundance. A first guess at the initial zooplankton field was thus simply chosen to be a fraction (50%) of the phytoplankton biomass field. There were no direct measurements of detritus. Since, in the model (Appendix A.2), detritus is produced as a result

Table 4
Real-time objective analyses scales

Correlation scale	Time τ (day)	Zero crossing D_0 (km)	E-folding D_ℓ (km)
Physical fields			
Bay scale	∞	60	25
Mesoscale	7	20	6.5
Biogeochemical fields			
Large scale	∞	600	400
Bay/large mesoscales	∞	30	15

of biological activity, the initial detritus distribution was assumed to be similar to those of the phytoplankton and zooplankton. With this approximation, the first-guess initial detritus field was set to be proportional to the phytoplankton field. Its magnitude was tuned based on comparing PON and Chl-*a* measurements (Libby et al., 1999) and was set to 5% of the phytoplankton biomass.

The time and space decorrelation scales (Appendix A.4) used to grid the physical (T and S) and biogeochemical fields are given in Table 4. In the real-time experiment, the scales of the biogeochemical OAs were chosen to be larger than those of the physical OAs. This is to accommodate for the lower resolution of the historical biogeochemical data (see Section 3.2 and Fig. 4) and for the fact that these data are not synoptic to August 1998 (i.e., their mesoscale and submesoscale biogeochemical features are uncertainties). Basically, the initial biogeochemical fields aimed to represent the Bay scales and some of the larger mesoscale features.

To illustrate the resulting first guess at the initial fields, salinity, chlorophyll *a*, nitrate, and ammonium cross-sections along the main axis of Massachusetts Bay are shown on Fig. 6. The other biogeochemical fields are all proportional to Chl-*a* (Table 3) and are not shown. Due to the Bay-scale/large-mesoscale scales used in the objective analysis (Table 4), the biogeochemical fields are relatively smooth, but due to the uneven distribution of the stations over the basin (see Fig. 4), they are not dynamically consistent with each other everywhere. For example, Chl-*a* peaks in the central part of the basin (Fig. 6b) while NH_4 is more patchy and is largest at the periphery of the basin. The Chl-*a* maxima in this section are near 10 m, but extend from 2 to 20 m depth.

The initial physical and biogeochemical model fields should be dynamically consistent to avoid shocks at the start of the forecast. The first guess at these initial fields (Fig. 6) was thus dynamically adjusted in two steps as follows. The physical model, initialized with the first-guess physical fields, was first run for 5 days (from August 17 to August 22) to adjust the vertical velocities. During this physical adjustment, ageostrophic vertical velocities arise, which is important for biogeochemical forcings. Due to the limited sampling in August 1995 (Fig. 4b–d), errors in the measurements, and very approximate initialization formulas (Table 3), the first guess at the initial biogeochemical fields (Fig. 6) is also not compatible everywhere, nor is it adjusted to the dynamical model structures and parameters. To adjust these fields, the 3-D biogeochemical model was run for 10 days without advection. The only physical processes allowed were the vertical diffusion and the daily light forcing. This second adjustment aimed to obtain biogeochemical fields and processes that are in dynamical equilibrium among themselves.

The above two dynamical runs are said to occur in the "adjustment space" (Anderson and Robinson, 2000), as opposed to "data space" and "simulation space." The final states of these biogeochemical and physical adjustments were used as the initial conditions of the coupled model simulations. Their spatial distributions are illustrated on Fig. 7. As expected for late summer conditions, the water column is well stratified. On average, the mixed layer extends to a depth of 2–8 m, with a thermocline near 10 m. By construction, the initial biogeochemical fields are smooth. The Chl-*a* maximum (reaching about 2 mg m^{-3} Chl) is generally around 20–25 m, near the lower pycnocline. The nutricline is typically just below the Chl-*a* maximum and nitrate increases towards the bottom. The ammonium distribution shows trends similar to those of nitrate. Both nutrients are depleted within the mixed layer and the biological production in the euphotic zone is nutrient-limited. During August to early September, the production is mainly controlled by physical processes at the thermocline, which bring nutrients to the euphotic zone. The nutrient inputs to the euphotic zone by benthic flux are limited due to the strong pycnocline.

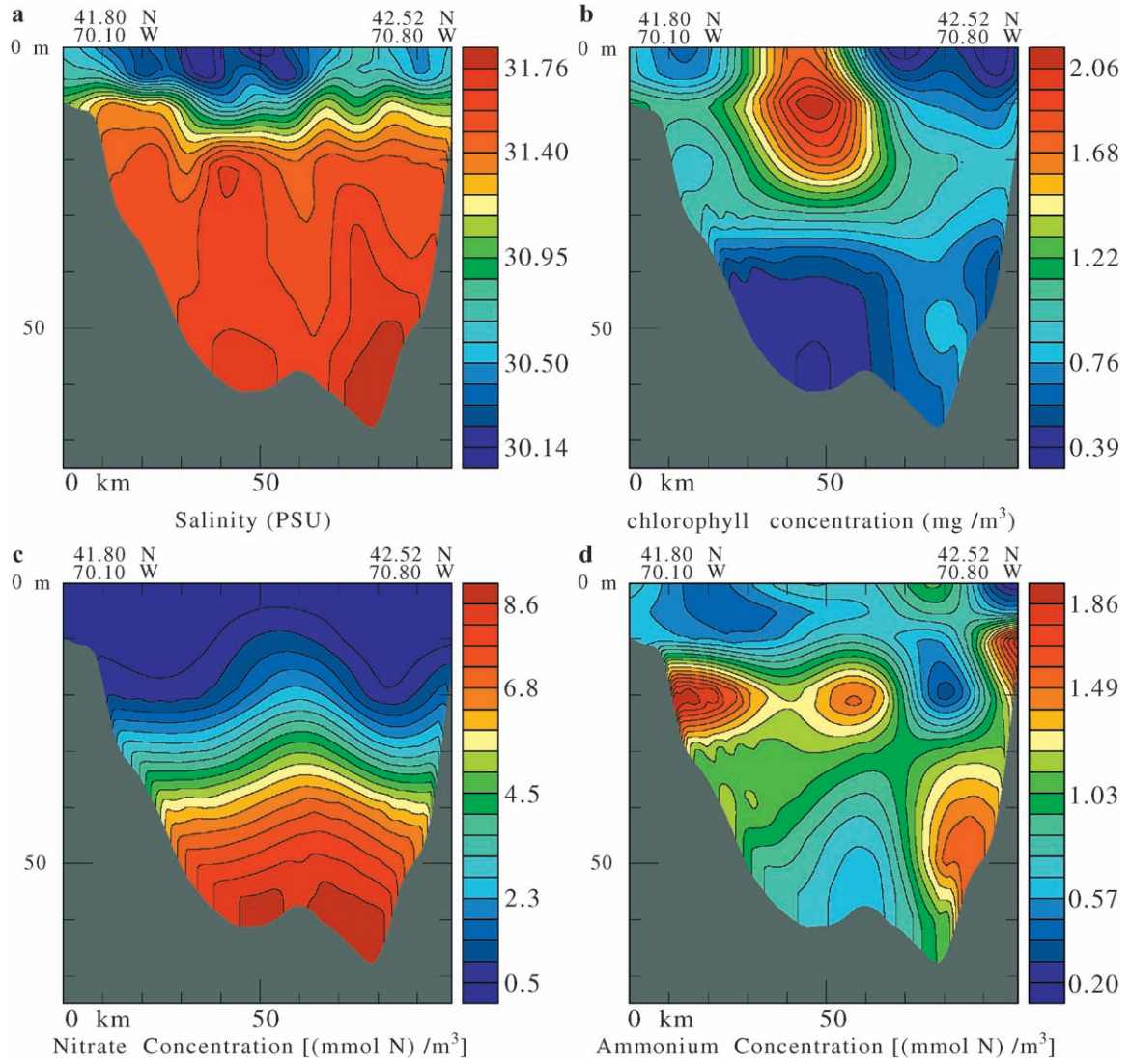

Fig. 6. Vertical cross-sections from south to north along Massachusetts Bay in the objective analyses of the measured variables: (a) salinity (ppt); (b) chlorophyll *a* (mg Chl m^{-3}); (c) nitrate (μmol N l^{-1}); (d) ammonium (μmol N l^{-1}). These fields illustrate the (nonadjusted) first guess at the initial conditions.

4.1.3. Real-time simulations: ecosystem evolution and processes

The coupled model simulation reported here was the final one being carried out in real time. It is conducted nominally for the period August 17–October 5, 1998. The adjustment of the initial biogeochemical fields to the physical advection, surface atmospheric forcings, and assimilation of the August 17–21 physical data takes place within the first 5–10 days of the simulation. During this period, biogeochemical features become tighter and submesoscale and mesoscale variability is created. Once this initial period is passed, biogeochemical variables reach a more realistic dynamical equilibrium, which consists of biogeochemical fields that slowly vary in accord with the physical dynamics and fluctuate in accord with the daily light variations. In this presentation, results from days 10 and 30 are discussed to illustrate

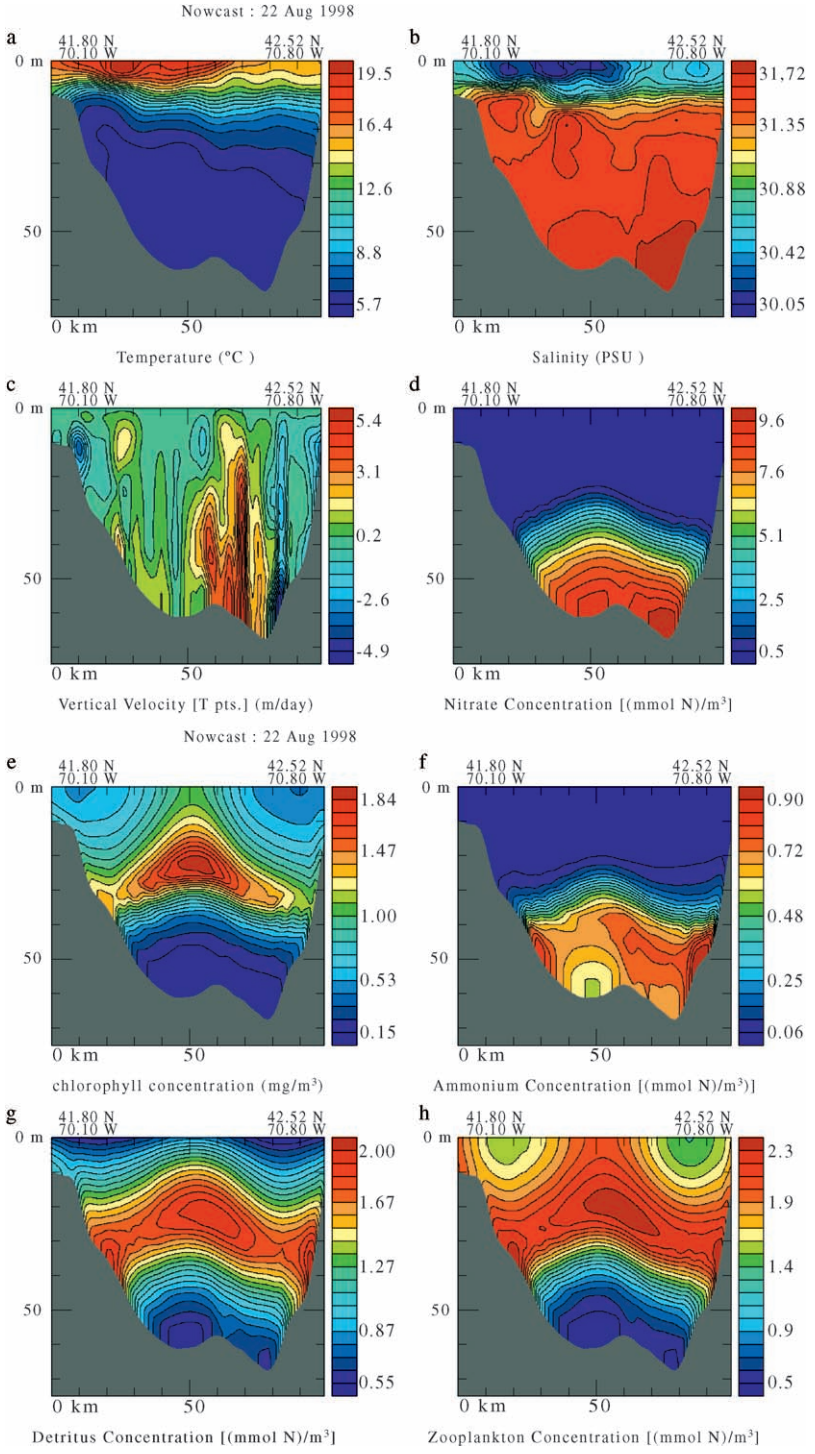

Fig. 7. Vertical-cross sections from south to north along Massachusetts Bay in the (adjusted) fields used to initialize the model in real time: (a) temperature (°C); (b) salinity (ppt); (c) vertical velocity (m day^{-1}); (d) nitrate (μmol l^{-1} N); (e) chlorophyll *a* (mg m^{-3} Chl); (f) ammonium (μmol l^{-1} N); (g) detritus (μmol l^{-1} N); (h) zooplankton (μmol l^{-1} N).

the evolution of the biogeochemical variables (Figs. 8 and 9).

A comparison between the initial and simulated biogeochemical fields on day 10 (Figs. 7 and 8) shows some remarkable differences. The simulated fields display smaller-scale structures and their local maxima are larger. The Chl-*a* maxima were moved up from 25 to 15 m, and the magnitudes of these maxima increased from about 2 to 3 mg m^{-3} Chl. The increase in the magnitude of local maxima and upward movement of the other variables is also evident except for nitrate. The vertical distribution of the nitrate field on day 10 generally reflects initial conditions and its modification is due to vertical advections, sinking of detritus, settling of phytoplankton, and ammonium nitrification. The upward and downward motions in the water column exert strong effects on the biological activity as upward motions can bring nutrient-rich water in the euphotic zone while downward motions can move the nutricline and phytoplankton out of the euphotic zone.

The simulated evolution of the biogeochemical fields during days 0–10 shows that even though (sub)mesoscale and mesoscale structures were not present in the initial conditions, the real-time models with physical data assimilation were capable of creating and maintaining spatial and temporal scales as those expected in Massachusetts Bay based on local measurements (e.g., Libby et al., 1999). The model thus had statistical skill. However, the model had deterministic skill (i.e., forecast the right feature at the right time and place) only at the Bay scales. This is because no in situ synoptic biogeochemical data were utilized in real time and such data are necessary to control the uncertainties at smaller scales.

On day 10, the scales of Chl-*a* are relatively similar to those of the zooplankton (the value of the Ivlev constant has logically a strong influence on this similarity; Tables 1 and 2 in Appendix A.2). These Chl-*a* and *Z* scales are found to be consistent with the scales of physical processes (Fig. 8a–c), mainly imposed by vertical and horizontal velocities. At the base and below the euphotic zone, the scales of NO_3, NH_4, and detritus seem, however, not as synchronized with the physical scales. Biogeochemical processes are there important in controlling the distribution of nutrients through nutrient uptakes, excretion by phytoplankton, and regenerative processes.

During most of the MBST-98, in the real-time simulations, the depth of the euphotic zone was on average deeper than the mixed layer depth, which agrees with the August–September 1998 MWRA data (Libby et al., 1999, 2000; Keller et al., 2001) even though these data were not utilized (Section 3.2). In between these depths, where light and nutrient inputs are sufficient, photosynthesis exceeds phytoplankton loss, which leads to patches of phytoplankton growth. The nutrient enrichments within these depths were here mainly controlled by physical processes. The nutrient variations and patches thus usually had physical scales at these depths and so did the corresponding phytoplankton patches (e.g., Fig. 8). Below the mixed layer, ammonium concentrations increase due to typical regenerative processes. By day 10 (Fig. 8f), the ammonium field indicates that the concentrations in the productive surface layers (euphotic zone) are around 0.1–0.3 μmol l^{-1} N. Below the euphotic zone, they reach 1.5–3 μmol l^{-1} N. Strong vertical stratification favors the accumulation of phytoplankton at depths optimal for growth. This optimal depth depends on the availability of nutrients and PAR and, on day 10, it is near 15 m. At the northern boundary (right sides of figures), upwelled water brings cold and nutrient-rich water to the surface. Consequent blooms in phytoplankton and zooplankton growth are thus observed.

At the end of the simulation, autumn conditions start to dominate, but on September 21 (Fig. 9), not enough strong wind events have yet occurred, and the water column is only partially mixed. The autumn bloom has not yet fully developed everywhere. During September, the mixed layer gradually deepens and, on September 21, it displays a downward inclination from north to south (Fig. 9a and b). The nitrate concentration has increased below the pycnocline due to nitrification and sinking of PON to the bottom (Fig. 9d). In the northern region, high ammonium is observed just below the high Chl-*a* patch.

4.2. Postcruise simulations and analyses

After the real-time experiments, the August–October 1998 raw biogeochemical data were processed and calibrated (Section 3.2). More than 500 postcruise simulations were then carried out based on these synoptic data. These model calibrations involved

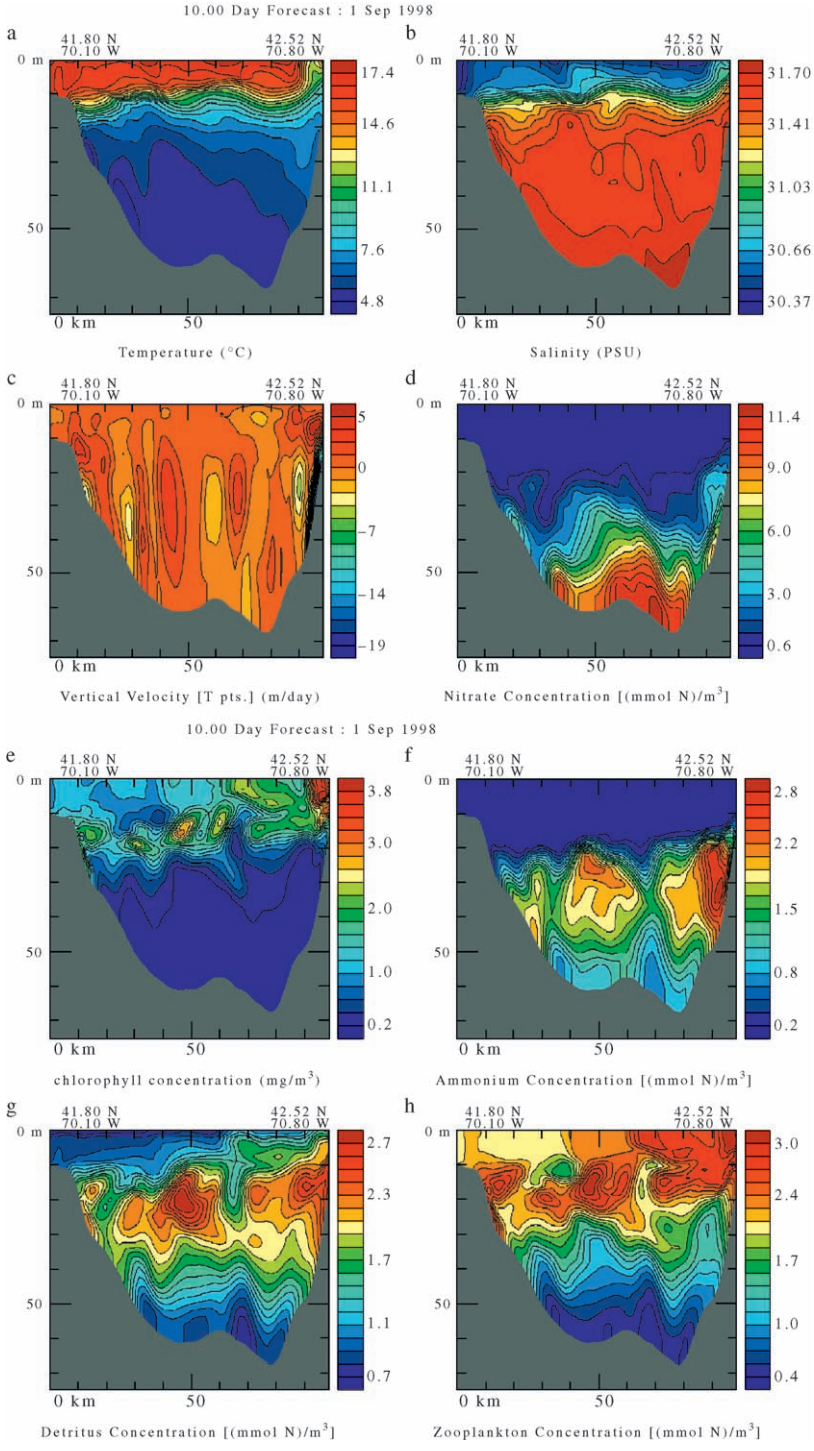

Fig. 8. As in Fig. 7, but after 10 days of simulation.

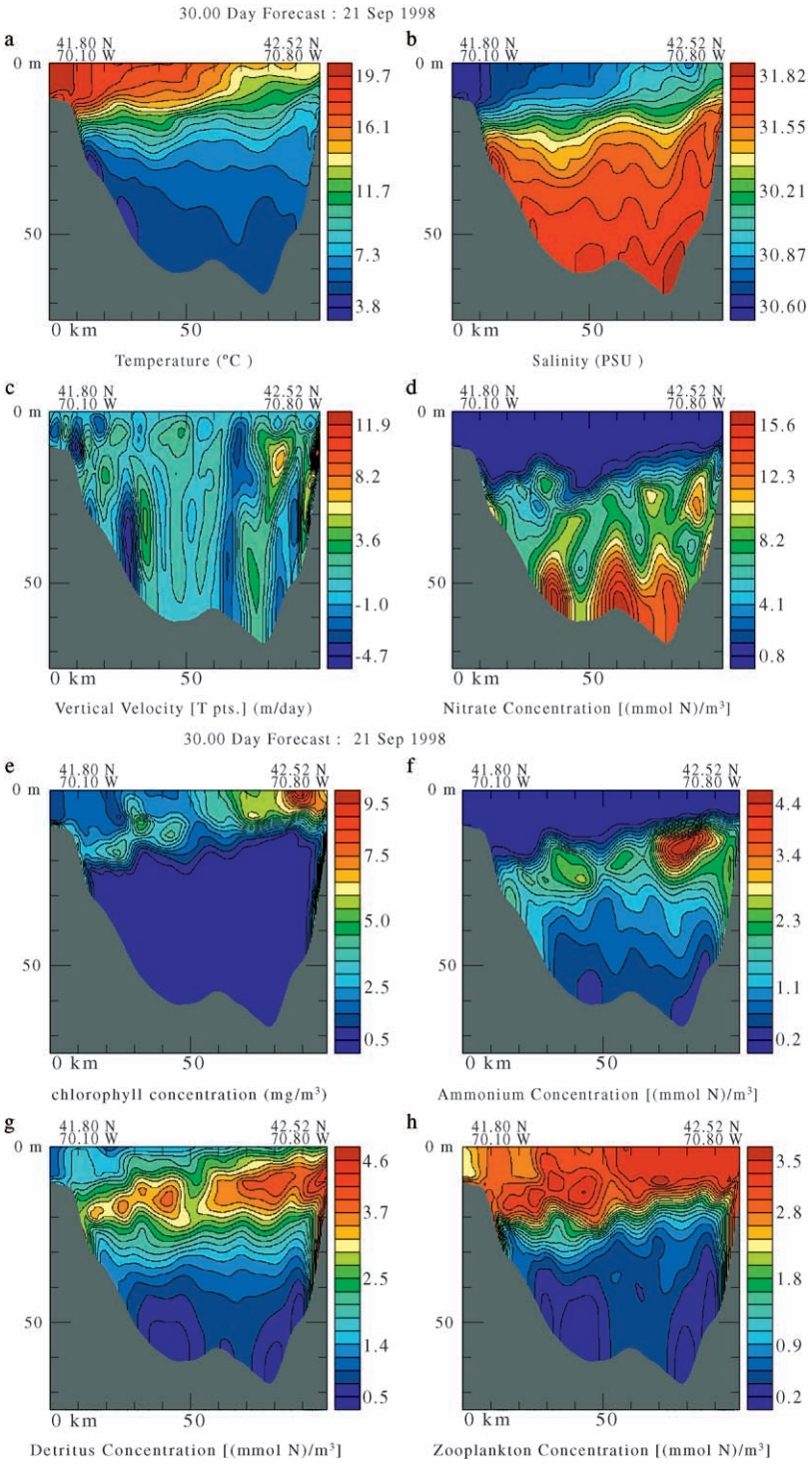

Fig. 9. As in Fig. 7, but after 30 days of simulation.

dynamical adjustments, forecasts with data assimilation, and quantitative data–model comparisons. Their purpose was to improve the real-time simulations, including the initialization procedures, the biogeochemical model structures and parameters, and the biogeochemical data assimilation.

4.2.1. Postcruise model calibration and initialization

The biogeochemical model structures, parameters, and initial fields are now selected simultaneously. This estimation can be computed automatically within the framework of an inverse scheme (e.g., Matear, 1995; Spitz et al., 2001). However, an issue in the current application of such schemes (Robinson and Lermusiaux, 2002) is the possibly strong dependence of the result on the a priori hypotheses (e.g., the first guess used to start the iterations), the constraints employed (weak vs. strong, many vs. few), and even the scheme utilized to minimize the cost function (Vallino, 2000). As a first step, the goal here is thus to search for adequate a priori hypotheses. The automated inversion is not carried out. By successive trials, hypotheses are modified and improved as a function of data–model misfits.

The a priori constraints employed included diverse data bounds and approximate dominant (dynamical) balances in the biogeochemical model. The latter were mainly used to determine a first guess at the initial fields. For given constraints, the fit of numerical simulations to data was measured, and the model structures and parameters as well as the constraints were then modified accordingly. In measuring the skill of model simulations (e.g., Section 4.2.2), the relative data uncertainties were taken into account. This procedure was repeated until the fit was judged to be adequate. The objective of such a research was to obtain a biogeochemical model with constraints and initial conditions that were good enough for assimilation studies. Even though many iterations were carried out, for clarity, only a summary on the parameter bounds, model properties, and initial fields obtained at the end of the iterations is presented.

4.2.1.1. Parameter bounds. The synoptic data bounds that were utilized for the model calibration and initialization were based on the MBST-98 Chl, NO_3, NH_4, T, and S data as well as other August–October 1998 MWRA data (Section 3.2). The amplitude of the photoinhibition process (β in Table 2, Appendix A.2) was, for example, bounded based on surface NH_4 data and production-vs.-light MWRA data. Values published in the literature (e.g., Libby et al., 1999; Keller et al., 2001) and previous surveys were also employed, as is described in Table 2 and Section 4.1 for the real-time case. For example, the sinking velocity for "detritus," which is here an aggregation of detritus, dissolved organic nitrogen, and bacteria, was constrained by literature surveys (Parsons et al., 1984; L.A. Anderson, personal communication). Bounds for the photosynthetic parameters (Table 2) were obtained from independent analyses of the 1998 MWRA data (Keller et al., 2001). Global properties were also utilized (e.g., the averaged total nitrogen content is important for the stability of the biogeochemical system).

4.2.1.2. Model formulations. Model structures (see Appendix A), including boundary conditions, were further optimized. For the physical model, open-boundary conditions, atmospheric forcings, and some numerical parameters were improved. For example, a weekly relaxation of the Gulf of Maine buoyancy forcings was combined with an Orlanski radiation scheme at the northern and eastern boundaries; surface heating computations and the time management of the atmospheric forcings were improved; and several wind-mixing coefficients were further calibrated. For the biogeochemical model, the structures of the surface and bottom boundary condition models were modified and the acclimation of Chl-*a* to light was added (see Tables 1 and 2 in Appendix A.2 for details). Special attention was paid to the daily light regime model and cloud coverage data, and the surface short-wave radiation forcing was improved. In postcruise simulations, the corresponding intraday time scales are taken into account. In particular, biogeochemical fields are now output exactly at noon, Boston time.

4.2.1.3. Model initialization and approximate dynamical balance. The initialization of the postcruise simulations is carried out in two phases, as in real-time: first, a first guess is computed, and second, it is dynamically adjusted. A time line of these initialization steps and subsequent data assimilations (Section

4.2.3) is given on Fig. 10. For the physical fields, the initialization procedure was similar to that carried out in real time (Section 4.1.2). The first-guess physical state is computed for August 17, using the OAs of the initial T and S data, geostrophy, and an imposed barotropic transport streamfunction along open boundaries. The positions of the MBST-98 and MWRA T and S profiles, which are used in the mesoscale stage of the initial OAs, are plotted on Fig. 11a and b. To adjust velocities, the momentum equations are then integrated for 4 days (Fig. 10), keeping T and S fixed.

To compute the first guess at the initial biogeochemical fields, approximate dominant (dynamical) balances in the biogeochemical equations (Table 1, Eqs. (1)–(7)) were researched. The model equations were nondimensionalized using an estimate of the dominant physical and biogeochemical scales. Note that these scales were computed from the hypothesized parameter and state variable ranges. Following a perturbation method, the biogeochemical variables were then expanded into power series: zeroth-order terms, first-order terms, etc. The zeroth-order terms were then balanced against each other. Of course, this zeroth-order balance depended on the parameter bounds chosen, and vice versa.

Based on the extensive simulations that were carried out, for the parameter bounds that led to largest skill, it was found that at zeroth order, the daily averaged effects of the fast intraday light time scale are approximately balanced against the other dominant but slower terms such that at the end of the day, at zeroth order, biogeochemical variables return to their values at the beginning of the day. Without entering in the specifics (to be reported elsewhere), for each equation and at each point in space, the daily average of the production terms is balanced against the other biogeochemical terms and dominant physical terms, if they are judged to be of the same order. For example, the zooplankton zeroth-order balance consists of a local grazing–mortality equilibrium while the detritus balance also involves the vertical sinking and surface mixing layer terms.

This new concept of circadian (daily) zeroth-order biogeochemical balance (Fig. 10), when combined with the data available for initialization, importantly allows to compute a first guess at the initial values of the non-observed state variables. To do so, the observed fields are inserted in the circadian balance and the balance is then (numerically) solved to esti-

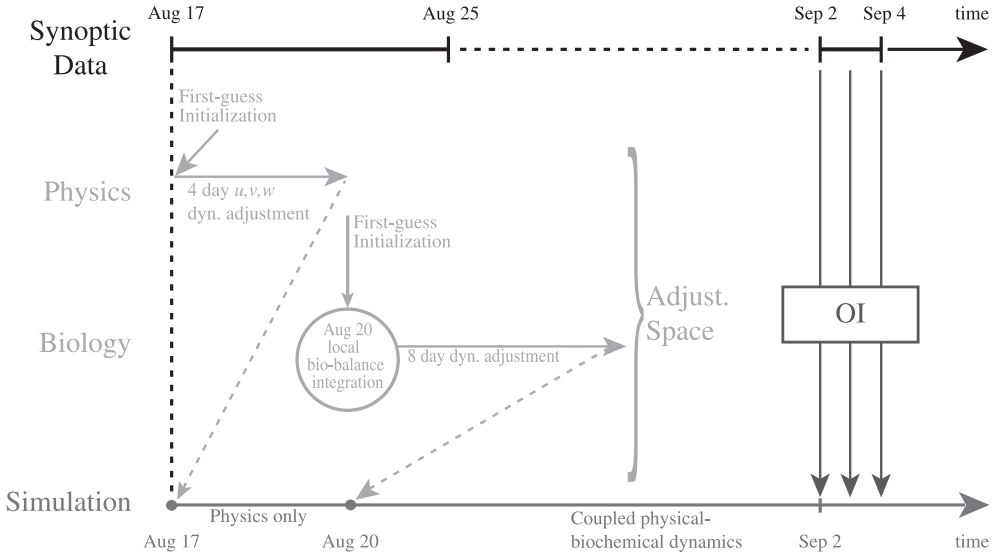

Fig. 10. Time-line diagram for the biogeochemical–physical synoptic data, dynamical adjustments, and simulations with data assimilation. Adjustments are carried out offline. Final adjusted states are utilized to initialize simulations (backward dashed arrows) and to correct the forecasts in a data assimilation (the OI box contains dynamical adjustments).

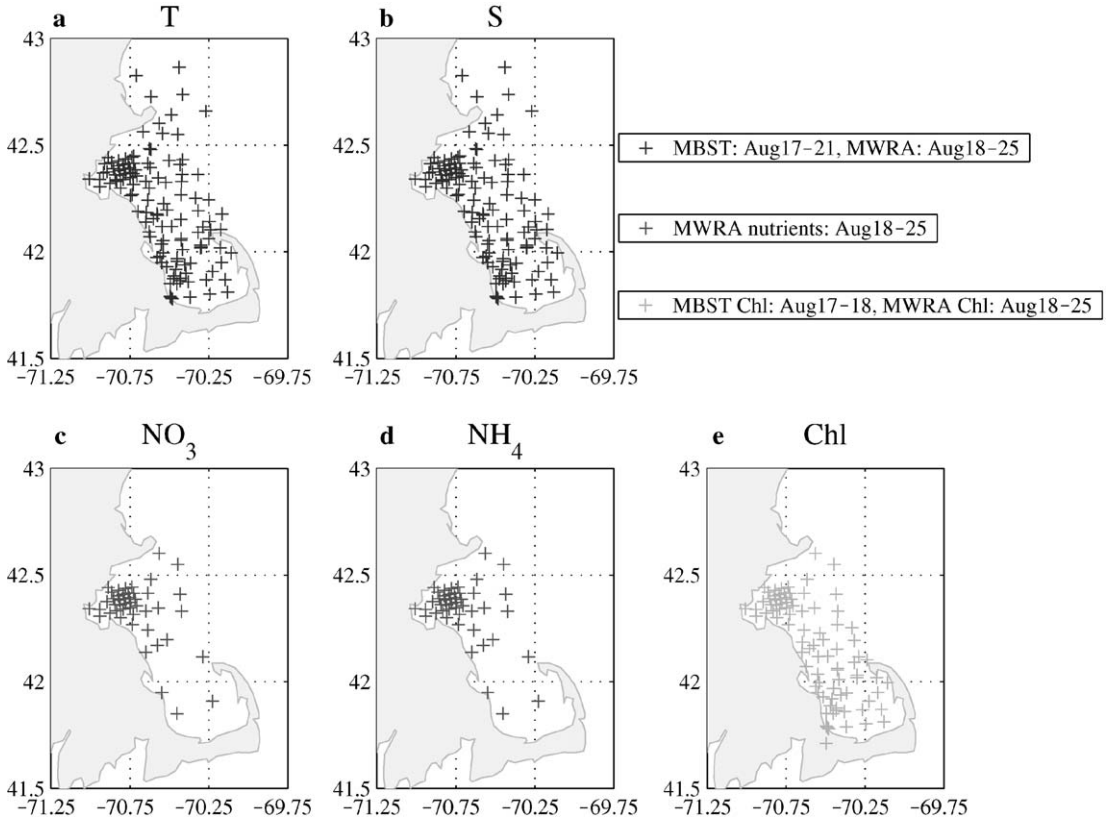

Fig. 11. Positions of the synoptic data used in the initialization of the postcruise simulations.

mate the unknowns, the non-observed fields. This computation replaces Table 3 used in real time.

Presently, a first guess at the observed Chl, NO_3, and NH_4 fields is obtained for August 20 by objective analyses (Appendix A.4). The August 17–25 MBST-98 and MWRA initialization data (station positions on Fig. 11c–e) are used in both stages of the OA for August 20 (central data date). For the Bay-scale OA, some historical data (Appendix B) are also used. The biogeochemical OA scales differ from those used in real time (Section 4.2): 60 km zero crossing, 25 km e-folding, and ∞ time scale are used for the Bay scale; 20 km zero crossing, 6.5 km e-folding, and 7-day time scale are used for the mesoscale. The biogeochemical OA scales are as those of the physics because the biogeochemical data are now synoptic and physical–biogeochemical interactions lead to common scales. Note that these objectively analyzed fields constrain the approximate circadian local balance and the parameters of the Chl, NO_3, and NH_4 equations (Eqs. (4), (5), and (7) in Table 1, Appendix A.2): the three of them must be compatible. The non-observed P, Z, and D first-guess fields for August 20 are ultimately obtained by solving the remaining equations in the circadian local balance (Fig. 10).

Cross-sections along Massachusetts Bay in these first-guess Chl, Z, NO_3, and NH_4 fields are shown on Fig. 12. Focusing on the objectively analyzed fields, based on the August 17–25 data (Fig. 11c–e), the Chl-a section (Fig. 12a) is relatively accurate at the mesoscales, but the NO_3 and NH_4 sections (Fig. 12c and d) are less accurate. This is because most of the nutrient stations are located near Boston Harbor; there are only a few profiles in Cape Cod Bay (on the left in Fig. 12). In addition, the nutrient data are bottle data which are of limited vertical resolution and are here simply linearly interpolated in the vertical. The Z section (Fig. 12b) results from the circadian local

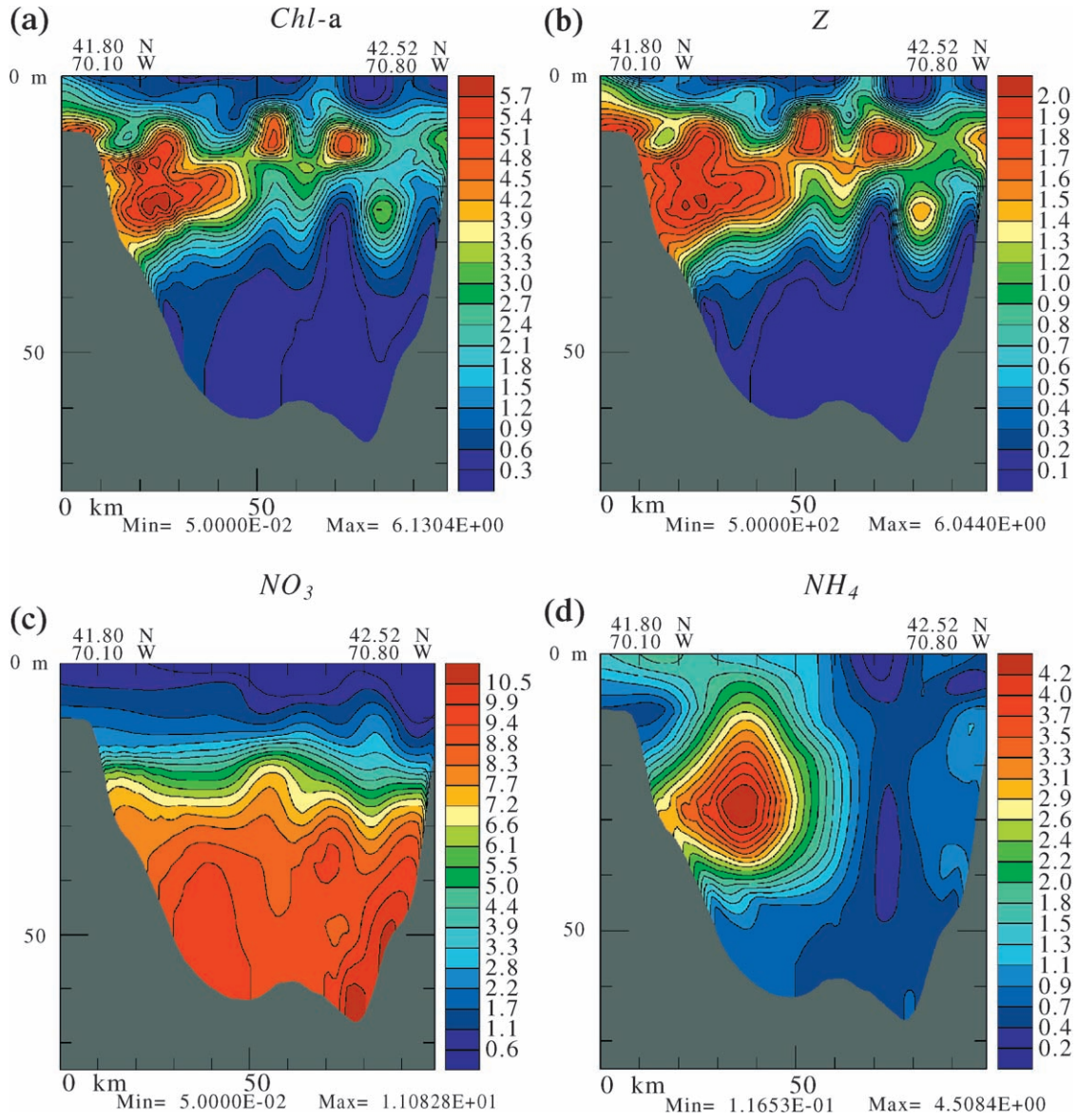

Fig. 12. Postcruise biochemical field initialization: first guess at the initial fields. For each panel, the four cross-sections are along Massachusetts Bay, from south to north.

balance. It is in zeroth-order biogeochemical balance with the *P* field.

With the above initialization scheme, for every set of model structures and parameters, there is a set of first-guess biogeochemical fields. The advantage is that, up to data uncertainties, these initial fields are at zeroth-order compatible with the model structures and parameters, and with the data available. Therefore, the duration of the costly second step of the initialization, the dynamical adjustment (Fig. 10), was reduced. It involved the time rate of change, diffusion, bio-sources/sinks, and light-forcing terms on August 20 (the forcings were daily periodic). The results are illustrated on Fig. 13. Comparing Fig. 13 with Fig. 12, the zeroth-order fields remain relatively unchanged; by construction, only higher-order terms are adjusted. The

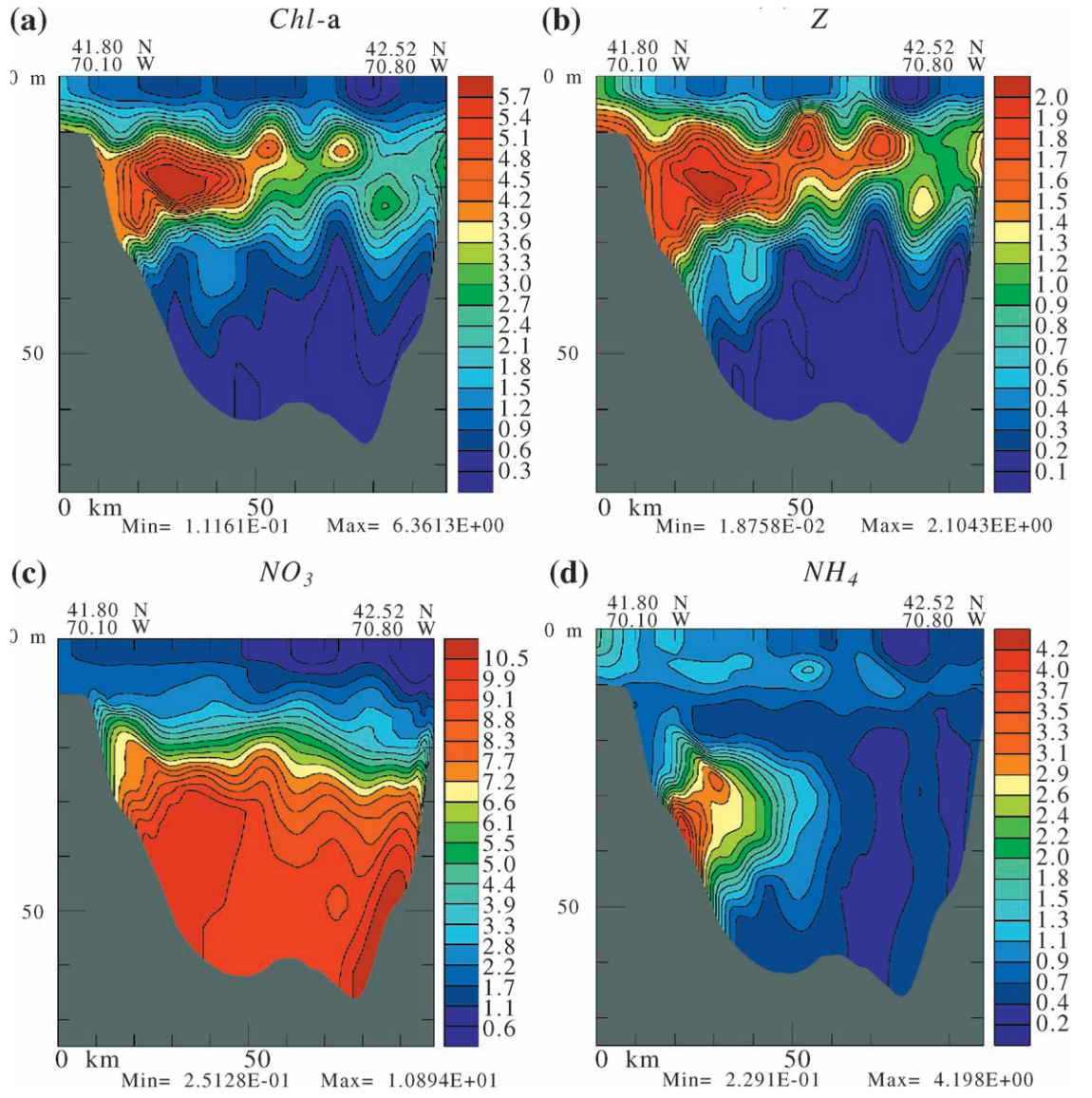

Fig. 13. Postcruise biochemical field initialization: dynamically adjusted initial fields. The cross-sections are as the first guess ones (same color scaling as on Fig. 12), but after dynamical adjustment.

largest changes are in the regions where data are too scarce for the objective analyses (Fig. 11) (e.g., in Cape Cod Bay, the NH_4 first guess is adjusted to the mesoscale resolution Chl and Z fields). Elsewhere, adjustments mainly involve diffusion and biogeochemical terms, which adapt the first-guess fields to the topography and remove numerical interpolation effects and objective analyses overshooting. The limited amplitude of the dynamical adjustment is an achievement. In the real-time procedure (Section 4.1.2), biogeochemical fields were drifting substantially during this adjustment because the model parameters and nitrogen contents of the first-guess fields were not compatible at zeroth order.

4.2.1.4. Postcruise and real-time fields. It is in fact only at the Bay scales that the 1998 initial conditions (Fig. 13; see also Fig. 17 in Section 4.2.4) have some similarities with the real-time initial fields (Fig. 7), which are based on the 1995 data. Differences are considerable in the smaller scales, for several reasons. First, the mesoscale physical forcings were altered and the data resolution and coverage are better in 1998 (Fig. 11) than in 1995 (Fig. 5). Second, the scales of the real-time biogeochemical OAs (Table 4) were enlarged to reflect the largest uncertainties due to the use of historical data. In the postcruise simulations, these scales were reduced, which explains why the postcruise fields display pronounced (sub)mesoscale variabilities as observed in the 1998 data. Third, the first guess at the real-time initial fields (Fig. 6) was unbalanced, which led to a model drift in the dynamical adjustments.

Finally, differences are also due to interannual variabilities. During the winter of 1998, the region experienced exceptional weather events, which can be attributed to the 1997 El Niño (Libby et al., 2000). As a result, the winter water temperatures were warmer and river outflows were higher. These conditions led to an early onset of the stratification in the spring and to a stronger summer stratification. Due to these extreme conditions and related higher grazing pressures, a classic winter–spring bloom was not really observed in 1998 (Keller et al., 2001). Instead, the phytoplankton abundance and chlorophyll *a* values observed by the MWRA monotonically increased from winter to fall. This continued production and strong stratification resulted in accumulation of elevated organic and inorganic nutrients below the pycnocline (Figs. 12 and 13) as compared to 1995 (Figs. 6 and 7).

4.2.1.5. Simulation time line and iterative calibration based on forecast skill. Once fields and parameters were initialized, the full physical model was integrated from the August 17 initial conditions to August 20 (Fig. 10), hence building the internal response to atmospheric forcings. The biogeochemical model was then started, using the August 20 initial conditions, and, from there on, the coupled physical–biogeochemical model was run forward. At subsequent data times (e.g., on September 2; Fig. 10), the skill of the model was evaluated, before and after data assimilation (Sections 4.2.2 and 4.2.3). If the skill was not sufficient, a revised model calibration and initialization were carried out, and the whole process (Fig. 10) was repeated.

4.2.2. Postcruise skill evaluation

The quality of a particular simulation was measured quantitatively by comparing model fields to data values at data points and to objectively analyzed data fields. To do so, all synoptic data until the end of September 1998 were employed. The update surveys (September 2–4) played a special role since they corresponded to the first new data collected after the initialization surveys (Appendix B). Once the skill was evaluated on a given day, the data collected on that day were assimilated (Section 4.2.3) and the simulation continued up to the next data time and skill evaluation. The model parameters of the resulting best postcruise simulation are listed in Table 2 (last column, Appendix A.2). Most parameter values differ from the real-time ones. For some parameters, relatively small differences can have a significant impact on the model solution (e.g., modify the averaged depth of the Chl-*a* maximum or the properties of surface nutrients). Such parameters include the C:Chl ratio, the zooplankton grazing and mortality parameters, and the phytoplankton growth parameters (e.g., half-saturation constants for nutrients uptakes). However, within our extensive set of simulations, if parameter values differed by less than 20% from the present best ones, the results remained qualitatively identical. The final results are therefore qualitatively robust. Qualitatively here means that patches, blooms, and in general biogeochemical gradients occur at similar times and places. Quantitatively (i.e., based on strict comparisons of model fields with data), the postcruise values listed in Table 2 led to the best results. Note that since a numerical inverse scheme was not utilized, a set of parameters somewhat different from those in Table 2 could give even better quantitative results.

The skill evaluations are illustrated by Fig. 14, for the case of the best postcruise simulation. The ecosystem forecast integrated from August 20 (using analyzed atmospheric forcings if available; see Appendix A.3) is compared to the data collected during September 2–4. Fig. 15 shows the station positions for these update surveys. Since Chl-*a* in the biogeochemical field is the most densely observed in the vertical and horizontal, it is chosen to illustrate

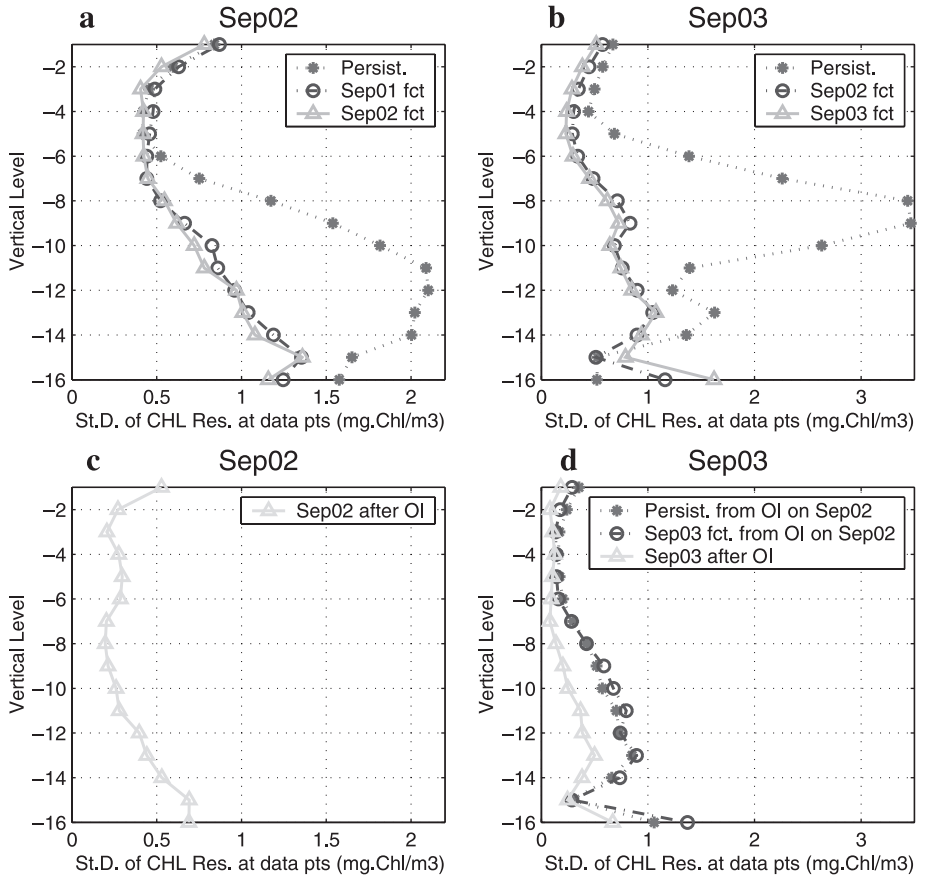

Fig. 14. Skill evaluation for the postcruise model structures and parameters. The skill metric shown is the sigma level by sigma level root mean square of data forecast residuals (bias not shown). Top two panels (a–b): Pure forecast skill (no assimilation). Bottom two panels (c–d): Skill after data assimilation. Fct = forecast.

the model skill (nutrient data lead to similar results). On Fig. 14, the top two panels illustrate the pure forecast skill (no assimilation) and the bottom two panels, the skill after data assimilation. The skill measure shown is the root mean square (RMS) of the differences between model fields and data values at data points, averaged over each sigma level. The mean of these differences, which up to data errors is the bias of the model fields, was also computed. It led to similar skill results. Considering first the pure forecasts (Fig. 14a and b), a goal is to find out if forecasts have better skill than the initial conditions (i.e., Does the forecast "beat persistence"?). Ideally, this should be the case, especially in dynamically active regions where model impacts should be significant. Based on the Chl-*a* data collected on September 2 (Fig. 14a), the September 2 forecast is largely superior to the persistence forecast (initial conditions on August 20), on average by about 40%. The largest impact is around level 11. This is because, on September 2, this level is near the averaged depth of the Chl-*a* maxima in Cape Cod Bay (see Chl-*a* data coverage on Fig. 15e), hence near the depths of largest spatial variability. The September 2 forecast also improves the September 1 forecast (Fig. 14a), on average by about 7%, which suggests little phase error. Now comparing the model forecasts to the September 3 Chl-*a* data (Fig. 14b), results are similar. Briefly, the September 3 forecast is better than persistence on average by 55% and the largest impacts remain near the most dynamically active depths (i.e., the depth of the Chl-*a* maxima; on September 3, it was

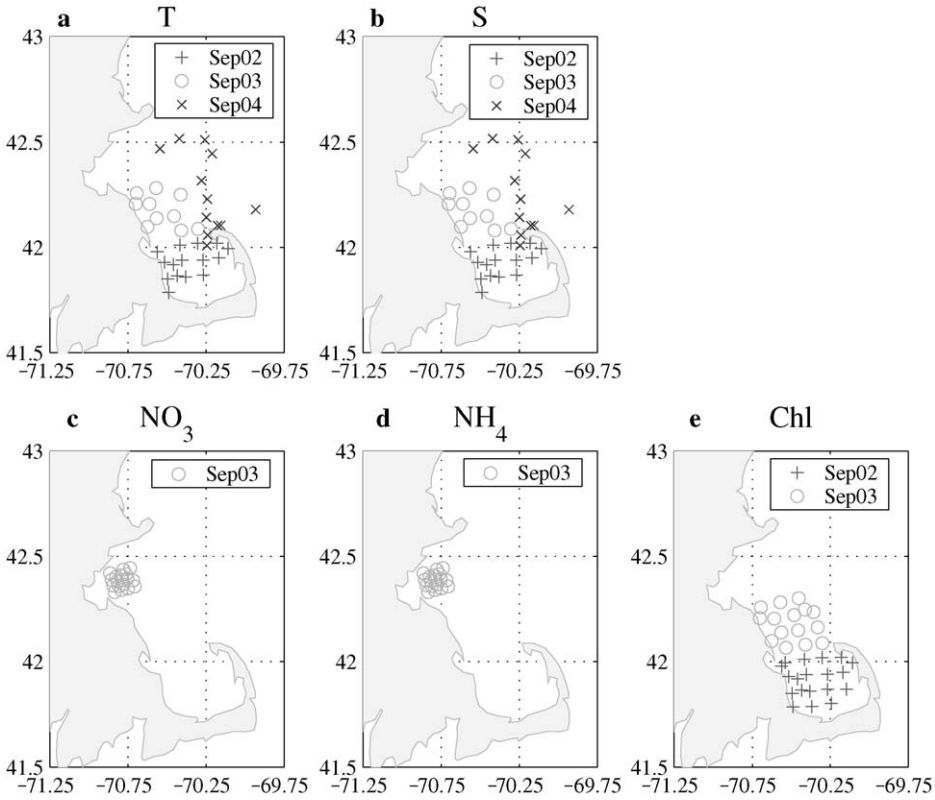

Fig. 15. Positions of the data assimilated in the postcruise simulations during the September 2–4 update surveys.

around level 8 in the center of the Bay, location of the data; see Fig. 15e). The skill after data assimilation is discussed next.

4.2.3. Postcruise physical and biogeochemical data assimilation

Once sufficient iterations of the calibration, initialization, and evaluation of skill for the September 2–4 forecasts were completed, different biogeochemical data assimilation strategies based on the OI scheme (Section 3.1) were compared. The best strategy was determined by computing the forecast skill of the different data-assimilative simulations for September 2–30. The time line (Fig. 10) only schematizes the assimilation of the first new data, the September 2–4 update surveys (see Fig. 15), because it is the focus of this subsection. Subsequently, for the September 17–30 data (Appendix B), assimilation strategies remained the same.

In the postcruise simulations, both physical and biogeochemical data are assimilated. For the physical OI, the T and S data available on a given day t_k are objectively analyzed using the same parameters as for the initialization (Section 4.2.1), except the time decay, which is reduced to 1 day. Internal velocities in geostrophic equilibrium with these analyses are then computed. These velocity and hydrographic fields are ultimately blended (Appendix A.4) with the forecast on days $t_k - 0.25$, t_k, and $t_k + 0.25$, with relative data weights of 70%, 90%, and 60%, respectively. This time ramping allows the dynamics to adjust to the new local data.

The approach for the biogeochemical OI is as for the physical OI. The parameters of the OA and blending steps are identical. Differences mainly arise because the non-observed velocities are replaced by the non-observed biogeochemical variables and because geostrophy and momentum adjustments (Sec-

tion 4.2.1) are replaced by circadian balance and biogeochemical adjustments.

To study the impact of the circadian balance and biogeochemical adjustments, a total of eight different biogeochemical OI strategies was investigated. Without entering into details (to be reported elsewhere), the simplest one consisted of assimilating the OAs of the observed biogeochemical variables only, without any adjustment and only at data times t_k (no time ramping). The most advanced one consisted of assimilating both the observed and non-observed variables at every point where a variable was observed and with the same time ramping as for the physical OI. In this latter case, a first guess at the non-observed variables (P, Z, D) is first computed using the circadian local balance equations (Section 4.2.1) and OAs of the observed variables at t_k (no historical data are used in these OAs for assimilation). The resulting complete biogeochemical fields are then dynamically adjusted, as in the initialization (Fig. 10), except that the observed variables are here reassimilated by OI during the adjustment. This is to maintain the local nitrogen content of the observed fields while high-order terms adjust in the non-observed fields. It is these dynamically adjusted fields (here seven fields) which are assimilated, with a non-dimensional error field equal to the minimum of the nondimensional observed error fields and with a time ramping equal to the physical one. Measuring the forecast skill by comparing model fields with data, the latter strategy was overall the best one during the September 2–30 period. Note that the same strategy, but without the dynamical adjustment step, was as good on average. This indicates that the local circadian balance is adequate at zeroth order and efficient.

The OI skill was measured by comparing model fields with the September 2–4 and September 17–30 data (Appendix B) at data points and with objective analyses of these data. It is illustrated by the two bottom panels of Fig. 14. The Chl-*a* skill is shown because the September 2–3 Chl-*a* data cover about 3/4 of the Bay while the nutrient data are only collected near Boston Harbor, at a few bottle depths (Fig. 15c–e). The assimilation on September 2 (Fig. 14c) logically reduces the averaged data residuals, at all depths. On Fig. 14d, three model Chl-*a* fields are compared to the September 3 data (Fig. 15e). First, for all three curves, assimilating data on September 2 improved the pure forecasts to September 3 (Fig. 14b) even though the Chl-*a* data coverage on September 3 differ from that on September 2. Comparing curves, the September 3 forecast from the September 2 OI fields (assimilates all September 2 but no September 3 data) is better than the persisted OI field of September 2 in the top eight levels but is worse below: on average, they perform similarly. The OI on September 3 (third curve on Fig. 14d) again logically reduces data–model residuals. Interestingly, the a posteriori residuals on September 3 are smaller than these on September 2 (Fig. 14c). The same remark applies to the amplitude of the data corrections. In addition to a possible convergence of the OI, this may suggest that the simulation is better in the center of the Bay than in Cape Cod Bay, reflecting the coverage of the initial biogeochemical samplings (Fig. 11c–e).

4.2.4. Postcruise simulations: ecosystem evolution and processes

For most of the simulation period, on days-to-weeks average, the Bay-scale buoyancy circulation was cyclonic. However, mesoscale and submesoscale circulation features were highly variable during this period. Atmospheric forcings also altered the Bay-scale circulation. For example, in response to wind events, the Cape Cod Bay circulation occasionally changed from cyclonic to anticyclonic. A main focus here is on the responses of the evolving biological activity to such highly variable buoyancy-driven and atmospheric-driven features.

During MBST-98, successive storm systems passed over the Bay on August 24, August 29, September 5–6, September 9–10, September 13, September 16, September 23, and September 28. On the event scales, these quasi-weekly weather systems had considerable impacts on the ecosystem variabilities of the Bay. For conciseness, the model results illustrated here are chosen according to three of these storm events (August 29, September 5–6, and September 23). Of course, to discuss important characteristics of the overall ecosystem evolution, several features that occurred before or after these events are also described.

On Fig. 16, the chlorophyll *a* and velocity fields at 20 m are illustrated on four snapshots, in each case at noon local time: on August 20 for the initial values and on August 29, September 7, and September 24 for the three selected storm events. On August 20 (Fig. 16a), all biogeochemical fields dis-

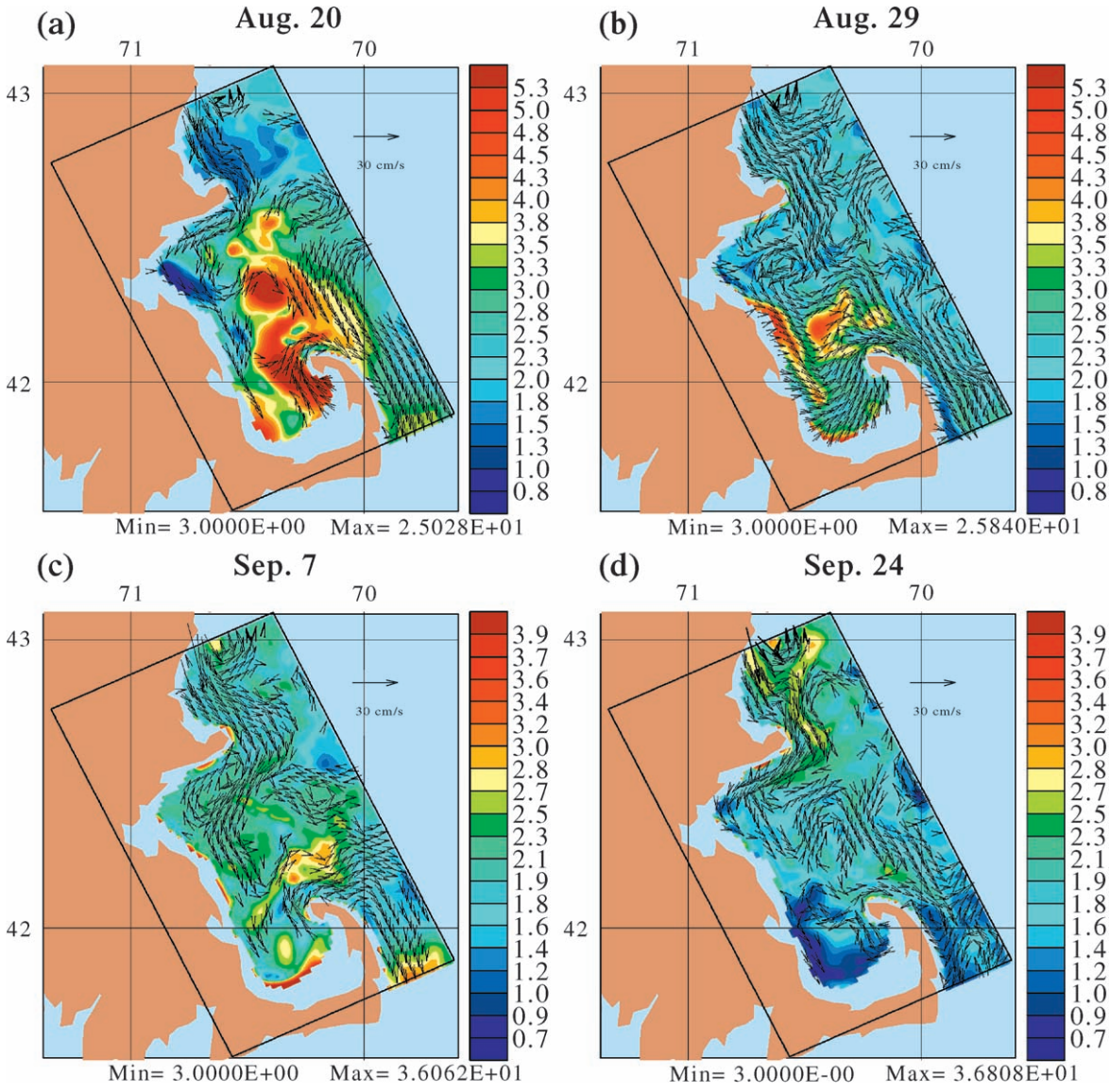

Fig. 16. Horizontal distributions of velocities overlaid on chlorophyll *a* (mg m^{-3} Chl) at 20 m and at noon time, on August 20 (initialization), August 29, September 7, and September 24, as estimated by the best postcruise simulation. Note that the color scale used in August differs for that used in September.

play maximum values west of Race Point. This is due to the advection of nutrient-rich bottom waters upwelled into the euphotic zone from the shallows of Cape Cod Bay (e.g., Billingsgate Shoal). Such Chl-*a* values were near the maximum ones measured during the summer–autumn of 1998 across the basin (Libby et al., 1999). The upwelling and mixing were influenced by the northeasterly winds that occurred before and during the initial sampling (Appendix B), but also reflect the local prevailing buoyancy circulation in this period. High Chl-*a* values also occur north of Race Point, on the shallow Stellwagen Bank, especially on its southwestern flank. This is in part due to topographic upwelling, both wind-

driven and buoyancy-driven by branches of the Maine coastal current meandering around and above the bank. On August 29 (Fig. 16b), elevated Chl-*a* values offshore between Boston and Plymouth are interestingly due to a downwelling event driven by the strong northwesterly winds of August 29. In accord with the physical findings of Lermusiaux (2001), by Ekman transport, nutrient-rich bottom waters at the coast are advected offshore and, by turbulent wind mixing and vertical advections, they lead to offshore production in the euphotic zone. This bloom lasted for 3 days. Chlorophyll *a*, phytoplankton, and productivity measurements confirm that a bloom occurred in this area (Libby et al., 2000). On September 7 (Fig. 16c), after 2 days of medium-to-strong northerly winds, Chl-*a* values at 20 m were varying between 2 and 3 mg m^{-3} Chl, which is lower than on earlier days. This is actually due to a deepening and vertical mixing of the Chl-*a* maximum. The 20-m Chl *a* values on the southern flank of Stellwagen Bank have been substantially reduced. The largest Chl-*a* at 20 m is now at the bottom, along the coastline. The strongest storm occurred on September 23 and the circulation responded quickly. By September 24 (Fig. 16d), the strong winds from the north had increased the depth of the surface mixed layer to 10–15 m and had eroded the thermocline, which, by then, extended from 10 to 50 m. In Cape Cod Bay, the local buoyancy circulation had become mostly cyclonic (Fig. 16d), which allowed to maintain a low Chl-*a* maximum near the center of the Bay at 20 m.

Figs. 17–20 show vertical cross-sections in the temperature, zonal velocity, nitrate, ammonium, chlorophyll *a*, zooplankton, detritus, and nitrate production rate fields, from south to north along Massachusetts Bay and at the same dates as on Fig. 16.

The model initialization data (Fig. 17) show that stratified waters in Massachusetts Bay are characterized by a subsurface layer of Chl-*a* maxima. The distributions of nitrate, ammonium, detritus, and zooplankton follow this structure, which is typical of summer conditions. Nutrient concentrations are generally low in the surface waters and increase near the pycnocline. The degradation of organic matter, remineralization, and nitrification contribute to the nitrate increase in bottom waters. The Chl-*a* maxima are around 10–25 m in depth and vary between 3 and 6 μg CH l^{-1} in magnitude (Fig. 17e). High Chl-*a* values are in the southern part of the basin where a bloom occurs (Fig. 16a). Although there is a continuous utilization of nutrients by phytoplankton, NO_3 concentrations are relatively high ($\sim$ 2–7 μmol N l^{-1}) at the depths of the Chl-*a* maxima. This NO_3, together with the locally high values of the regenerated NH_4, is able to maintain the bloom for the first 10 days of the simulation. All of these local features are mainly the result of a sustained advection from the shallow and nutrient-rich Cape Cod Bay. By day 10 (Fig. 16b), the bloom is dispersed and partly transported outside of the Bay. The local uptake of nitrate by phytoplankton (nitrate production rate) is shown on Fig. 17h. As a proxy for new production (Lalli and Parsons, 1997), it indicates that the new production is deeper in the northern half of Bay than in the southern half. This is in a large part due to the spatial variability of the euphotic zone in the Bay. This variability itself mainly results from the spatial variability of the incident radiation and light attenuation by the Chl-*a* field (Appendices B.2 and B.3). As is characteristic of late summer conditions, the new production is low and does not follow the detailed structure of Chl-*a*. The phytoplankton biomass is indeed sustained by the ammonium production rate (not shown), i.e., a proxi for regenerated production.

On August 29 (Fig. 18), the high Chl-*a* observed at initialization still exists in Cape Cod Bay. The nitrate production rate (Fig. 18h) is still higher in this area than in other regions (note that the high rate is localized near the sea floor as shown in the cross-section, but it actually occurs all around Cape Cod Bay). Detritus, nitrate, and ammonium concentrations below the high Chl-*a* have increased since August 20 (Fig. 17). This is due to the sinking of higher detritus concentrations produced by the high biological activity in this area, and to nutrient regeneration and remineralization processes at depth. Comparing Fig. 17f with Fig. 18f, the zooplankton biomass remains relatively unchanged, but the spatial distribution is modified due to advection and mixing.

Towards autumn (Fig. 19), the mixed layer gradually deepens and upper layers are becoming cooler due to stronger episodic winds and lower heat input. A few storms forced relatively deep mixing and induced eddies in the Bay. The anticyclonic eddy in the north (Fig. 19a and b) lowers the thermocline and

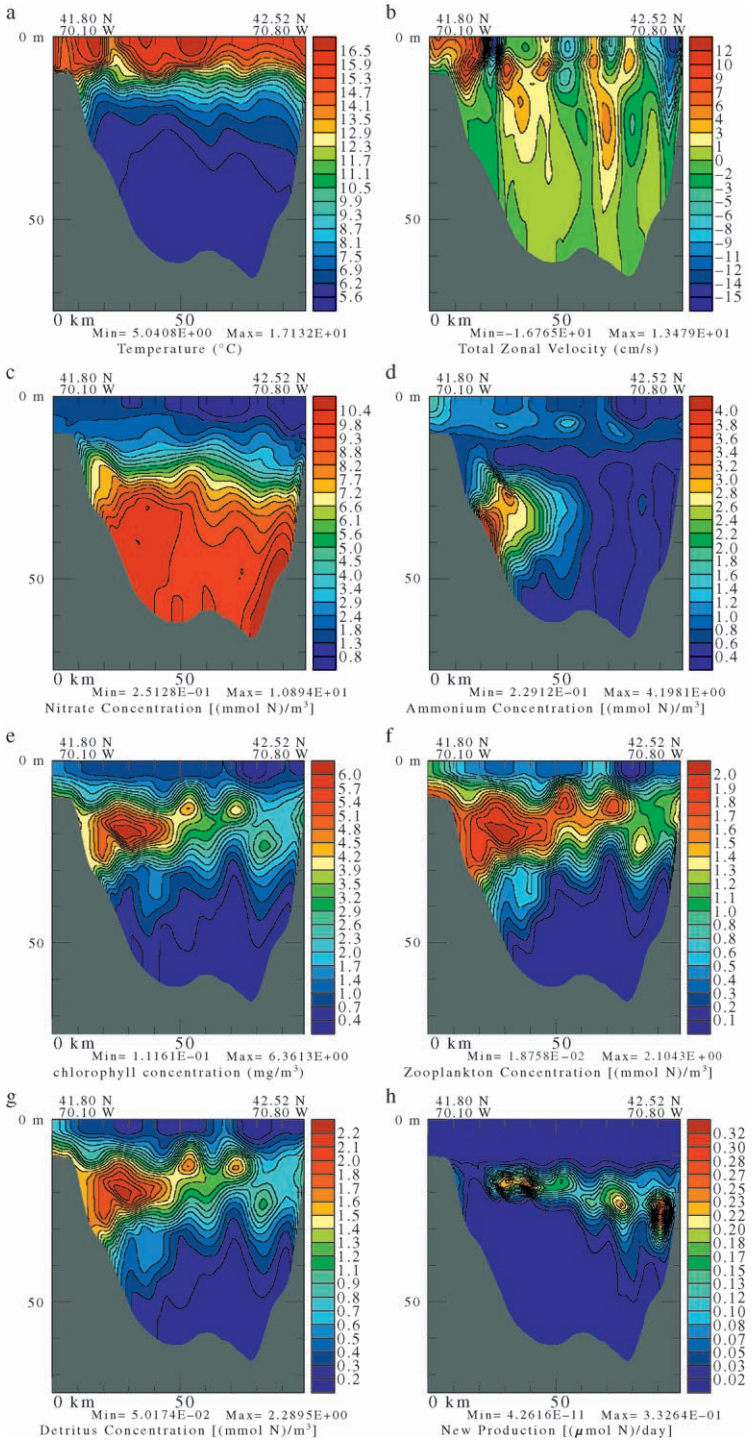

Fig. 17. Vertical cross-sections from south to north along Massachusetts Bay on August 20, as used to initialize the best postcruise data-driven simulation: (a) temperature (°C); (b) zonal velocity (cm s^{-1}); (c) nitrate (μmol l^{-1} N); (d) ammonium (μmol l^{-1} N); (e) chlorophyll *a* (mg m^{-3} Chl); (f) zooplankton (μmol l^{-1} N); (g) detritus (μmol l^{-1} N); and (h) nitrate production rate (μmol day^{-1} N), which is a proxy for new production.

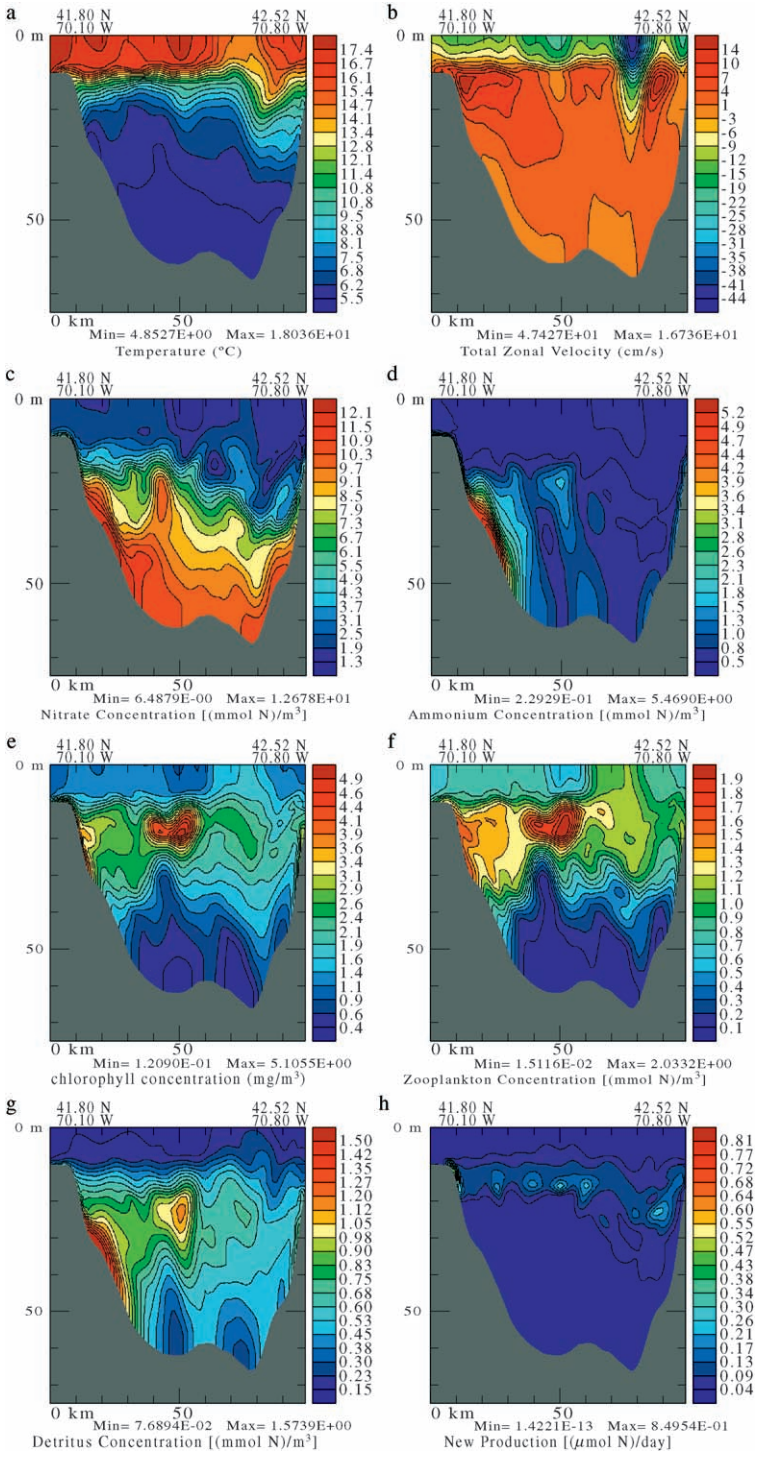

Fig. 18. As in Fig. 17, but for 29 August (day 9).

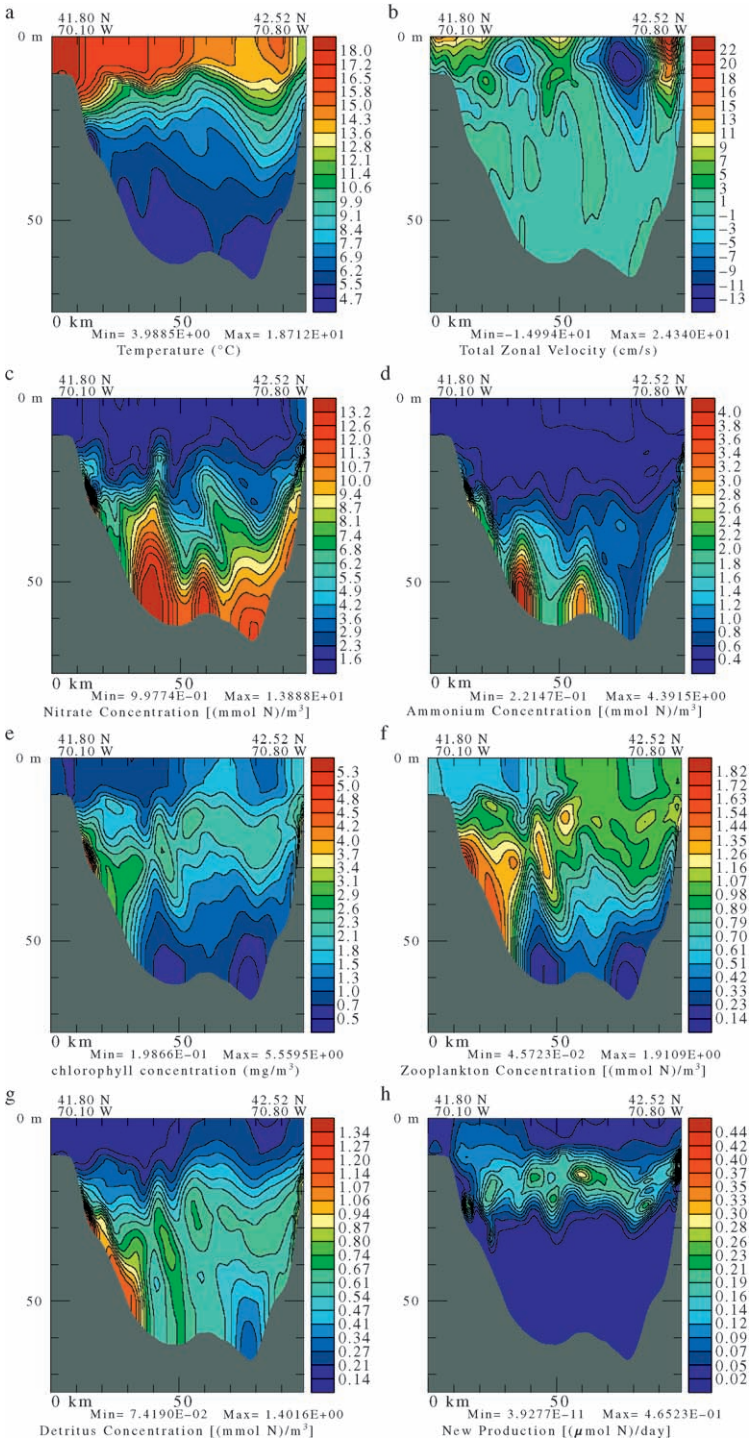

Fig. 19. As in Fig. 17, but for September 7.

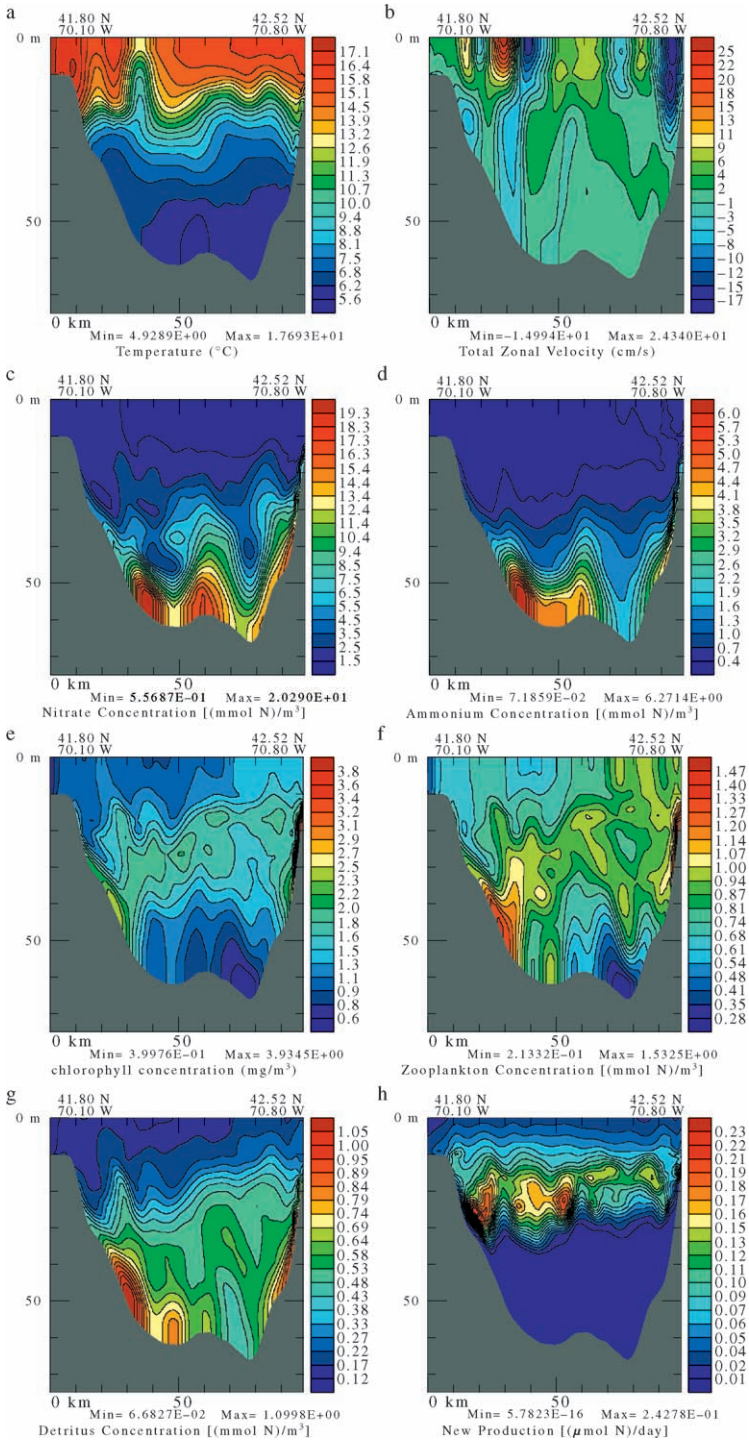

Fig. 20. As in Fig. 17, but for September 24.

pumps biogeochemical material out of the euphotic zone in its center. However, along its boundary, upwelling occurs (Lermusiaux, 2001), which brings nutrients in the euphotic zone and leads to small local blooms. In particular, such blooms are observed in the surface mixed layer and within a thin layer along the bottom at the northern coastline. The nitrate production rate displays maxima in this area (see dark red spots on Fig. 19h). The scales of the patches in this proxy for new production are smaller than the scales of the Chl-*a* patches. This is because primary production is mainly maintained by ammonium and the corresponding regenerated production. After the depletion of nutrients imported in the euphotic zone, the standing stock of phytoplankton is sustained for a couple of days with the utilization of the ammonium from remineralization and fractions of zooplankton mortality and excretion and also eddy diffusion of nutrients at the base of thermocline. Detritus sinks below the thermocline, where its remineralization generates ammonium peaks.

From August through September (Figs. 17–20), a gradual increase of nitrate, ammonium, and detritus below the pycnocline is observed in the simulations. This occurs even though there is a continued supply of nutrients to the euphotic zone during this period. From August 20 (Fig. 17a) to September 24 (Fig. 20a), the upper thermocline deepened approximately 10 m and the thermocline extension eroded down to 50 m. This deepening is not sufficient to overturn the water column and so enter winter conditions. Due to the maintained stratified conditions, detritus and nutrient concentrations on September 24 (Fig. 20) are likely reaching their maximum values below the pycnocline.

The nitrate concentrations on September 24 (Fig. 20c) are still low in the surface mixed layer but are much higher than in August. They reach 2–3.5 μmol l^{-1} N in Cape Cod Bay (south). The nutricline deepened following the deepening of the mixed layer. The ammonium and detritus patterns in the deeper layers are similar to those of nitrate. However, the ammonium is now relatively depleted at the surface because of nitrification and because the phytoplankton preferentially takes up ammonia over nitrate (Appendix A.2). The Chl-*a* maximum (Fig. 20e) also moved down to about 30 m and is inclined from north to south. A similar tendency was observed in the real-time simulations (e.g., Fig. 9). The Chl-*a* maximum is now broader than in August. The surface Chl-*a* values have increased a bit to about 1–1.5 mg Chl m^{-3}. This is likely the surface signal of the start of the autumn bloom. Overall, the zooplankton patterns follow the Chl-*a* patterns. Differences in patterns are largest at the surface and at depth. This is in part due to the saturation of the zooplankton grazing and Ivlev constant (Section 3.2). Compared to other biogeochemical variables, the zooplankton biomass varied the least on average during the August–October period. However, the zooplankton variations are important since zooplankton mortality and grazing mediate the phytoplankton growth in the present regenerated production cycle. In particular, note that during the 47 days of simulation, the zooplankton patches deepened, their scales became larger, and their maximum values decreased from about 2 μmol N l^{-1} (Fig. 17f) to 1.3 μmol N l^{-1}.

An example of eddy-induced biological production is observed on September 24 and corresponds to an upper layer cyclone in southern Cape Cod Bay. The vertical distributions of temperature and zonal velocities show the vertical extent of the eddy in this southern region (Fig. 20a and b). The wind-induced eddy is mostly confined to the upper 20 m of the water column and is able to entrain waters from 20 to 25 m to the upper layers. Nutrients and detritus are raised from deep layers, and the regenerated production (not shown) and to a lesser extent the new production (Fig. 20h) are then enhanced in this area, especially along the high-shear edges of the eddy where vertical cells occur (Lermusiaux, 2001). In situ observations show that the autumn bloom started by the end of September. From our postcruise simulations, it can be hypothesized that the autumn bloom in 1998 was likely initiated by the successive storms of September, in particular the very strong one on September 23, and by the corresponding deepening of the mixed-layer and subsequent eddy-induced mixing.

5. Summary and conclusions

Real-time and postcruise data-assimilative numerical simulations were carried out and evaluated for the Massachusetts Bay marine ecosystem during August–October 1998 using physical, chemical, and biological data and coupled four-dimensional (4D) physical–

biogeochemical models. The biogeochemical model was in part developed for this study. Other technical advances included methods for the initialization and calibration of the structures and parameters of the biogeochemical model, and for the assimilation via optimal interpolation of dynamically adjusted fields. All of these techniques should be useful in other midlatitude coastal ecosystems.

During the real-time forecasting, the physical model was initialized with hydrographic data from the August 1998 surveys and the biogeochemical model with historical data. This was carried out based on simple linear relations between observed and non-observed state variables, with subsequent dynamical adjustments. Initial physical–biogeochemical compatibility was achieved offline of the main forecast: velocities were adjusted for 5 days and biogeochemical fields for 10 days with frozen physics but no advection. During the ecosystem simulation, the coupled models were forced with real-time atmospheric fluxes. Physical data collected during the experiment were assimilated, but biogeochemical data were not.

The skill of the physical model was found to depend on the strength and accuracy of the atmospheric forecasts. For strong winds, the model skill was usually limited to about 2 days mainly because of atmospheric forcing uncertainties. For weak winds, the physical predictive capability was extended to about 1 week. For the ecosystem simulations, robust, reasonable fields were achieved in real time, with a full set of parameters estimated from historical and real-time data. After about 1 week, biogeochemical fields appeared to be evolving in coupled dynamical equilibrium. They exhibited qualitative forecast skill only at the Bay scales, but included (sub)mesoscale features generated within the model. This skill at the Bay scales in part resulted from the forcing by the real-time 4-D physical forecasts with physical data assimilation, which created and maintained spatial and temporal scales.

After the calibration and analysis of the 1998 biogeochemical data, an extensive number of post-cruise data-driven simulations (hindcasts) were carried out and studied. These simulations assimilated both physical and biogeochemical data. Based on sensitivity studies and data–model misfits, the parameterizations of several biogeochemical and optical processes, the boundary condition formulations, and the initialization and assimilation procedures were improved. The final initialization procedure included bounded parameter ranges and dominant dynamical balances. Starting from the observed biogeochemical state variables (Chl, NO_3, NH_4), which are in dynamical equilibrium with the physics, the initial conditions for the non-observed variables (P, Z, D) are obtained by imposing a constraint of zeroth-order dynamical periodicity over a daily cycle. Full compatibility of the physics and biology is then easily achieved through a short 4-D dynamical integration with frozen physical fields. This procedure, based on dominant biogeochemical balance, is also used to update the non-observed variables at data assimilation steps. The skill of the coupled models was evaluated before and after data assimilation, showing substantial improvement over the real-time case, on all scales. For good enough initial data and atmospheric analyses, the predictive capability of the internal ecosystem remained significant after a couple of weeks.

Overall, the ecosystem of the Bay, including the transition from summer to fall conditions, was found to depend crucially on the successive atmospheric events, (sub)mesoscale buoyancy features, and interactions of these two highly variable physical forcings. Storm systems passed over the Bay quasi-weekly and, in response to their variable winds, coastal upwelling blooms occurred at various locations around the Bay. The storms altered the Bay-scale circulation and induced mesoscale variability, often modifying existing eddies or creating new ones. The rapid (sub)mesoscale physical injections of nutrients into the euphotic zone were found to produce subsurface phytoplankton patches on the physical (sub)mesoscales. Mesoscale eddies, jets, and meanders were able to maintain above-average (anticyclonic) or below-average (cyclonic) biological activity on the mesoscales, but regardless of the direction of curvature, higher production, supported mainly by submesoscale vertical cells, always occurred along high-shear edges of these eddies, jets, and meanders.

In late August and early September, all biogeochemical fields displayed maximum values west of Race Point because of advection of nutrient-rich bottom waters upwelled into the euphotic zone from the shallows of Cape Cod Bay. High biogeochemical activity also occurred on the southwestern flank of

Stellwagen Bank in part due to topographic upwelling, both wind-driven and buoyancy-driven by branches of the Maine coastal current meandering around and above the bank. A few downwelling events driven by strong northwesterly winds also occurred. These events advect the coastal, nutrient-rich, bottom waters offshore and, by wind mixing and vertical advections, can lead to offshore biological production.

During September, wind events strengthened and became more frequent, which deepened the mixed layer, and eroded the stratification, which led to deeper mixing of biogeochemical fields and the onset of autumn blooms. These blooms start on both nitrate and ammonium, but are usually maintained by regenerated production for several days to about a week. As the mixed layer deepens, the depths of the subsurface phytoplankton maxima increase while the biomass itself decreases. At the same time, the detritus and nutrient concentrations increase below these maxima and below the pycnocline.

Acknowledgements

We thank the individuals who participated in the LOOPS experiment for their helpful collaboration; M.J. Mickelson and the other members of the MWRA for their data and data analysis; and the NMFS Marine Monitoring and Prediction (MARMAP) program for historical data. We thank J.A. Dusenberry for providing the basis of the biogeochemical model. P.F.J.L. thanks L.A. Anderson, J.A. Dusenberry, J.J. McCarthy, and R.C. Tian for several biogeochemical modeling discussions; G. Strout for discussions on the calibration of fluorometer data to Chl-*a*; and P.J. Haley and W.G. Leslie for software and data expertise. We thank P. Moreno for comments on the manuscript and M. Armstrong for preparing some of the figures. Sukru Besiktepe was a visiting scientist at Harvard University while conducting the real-time biogeochemical forecasts. He was initially supported by the AFMIS project, and by NSF and TUBITAK through the joint NSF-TUBITAK program. P.F.J.L. and A.R.R. were supported, in part, by the Office of Naval Research under grants N00014-95-1-0371, N00014-97-1-0239, and N00014-97-1-1018 and the National Science Foundation under grant NSF-5710001319.

Appendix A. Coupled models and data assimilation scheme

A.1. Physical model

The physical dynamical model employed here is the 4-D primitive equation model of HOPS, which is based on the GFDL integration algorithm (Robinson, 1996; Spall and Robinson, 1990). The model grid covers the area shown in Fig. 4 at a resolution of 1.668 km with 16 terrains—following levels in the vertical. Horizontal subgridscale processes are parameterized using Shapiro filters, which are applied on the submesoscale components of the total PE fields (Lermusiaux, 1997). The filter is scale-sensitive and, after calibration of its parameters, can remove numerical noise while keeping dynamical scales mostly unchanged. The bulk vertical diffusion is parameterized based on the local Richardson number. The transfer of atmospheric fluxes (see Appendix A.3) to the water column involves a vertical mixing–length turbulent model based on a locally computed ‘‘Ekman depth’’ (L.A. Anderson and C.J. Lozano, personal communication). A bottom boundary layer and coastal friction parameterizations are also incorporated (Lermusiaux, 1997). At open boundaries, conditions based on an Orlanski radiation scheme (Orlanski, 1976) and on a weekly relaxation of the Gulf of Maine buoyancy forcings are employed. Across coastlines, the normal flow and tracer fluxes are set to zero. Details on the physical model and on its calibration for the mesoscale dynamics of Massachusetts Bay are given in Lermusiaux (2001) and references therein.

A.2. Biogeochemical model

The biogeochemical model coupled to the physical model includes phytoplankton P, zooplankton Z, detritus D, nitrate NO_3, ammonium NH_4, and chlorophyll a Chl-a. The core of this new model was mainly developed by J.A. Dusenberry (personal communication), based on previous HOPS models. P.F.J. Lermusiaux later modified and improved several portions of the model. The resulting biogeochemical model and its rationale are now described.

In comparison with previous synoptic data sets, the MBST-98 interdisciplinary data are unprecedented in synoptic coverage and breadth of type (K. Sherman,

Table 1
HOPS biogeochemical model equations for Massachusetts Bay

Generic ADR equation:

$$\frac{\partial \phi_i}{\partial t} + \mathbf{u}\cdot\nabla\phi_i - \nabla_h(A_i\nabla\phi_i) - \frac{\partial K_i \partial \phi_i/\partial z}{\partial z} = \mathscr{B}_i(\phi_1,\ldots,\phi_i,\ldots,\phi_7) \quad (1-7)$$

Phytoplankton dynamics ($P \doteq P_{NO_3} + P_{NH_4}$):

$$\mathscr{B}_P = V^{NO_3} + V^{NH_4} - gZ - n_3 P - n_4 P^2 - v_P \frac{dP}{dz} \quad (1-2)$$

where:

$$V^{NH_4} = \frac{P_P}{\theta^C_{Chl}} \frac{NH_4}{k_{NH_4} + NH_4} P; \quad V^{NO_3} = \frac{P_P}{\theta^C_{Chl}} \frac{NO_3 e^{-\psi NH_4}}{k_{NO_3} + NO_3} P; \quad P_P = P_m(1 - e^{-\alpha E/P_m})e^{-\beta E/P_m}$$

Zooplankton dynamics:

$$\mathscr{B}_Z = (1-\gamma_1-\gamma_2)gZ - n_1 Z - n_2 Z^2, \qquad \text{where} \quad g = R_m(1 - e^{-\Lambda P}) \quad (3)$$

Nitrate dynamics:

$$\mathscr{B}_{NO_3} = -V^{NO_3} + k_N NH_4 \quad (4)$$

Ammonium dynamics:

$$\mathscr{B}_{NH_4} = -V^{NH_4} - k_N NH_4 + \gamma_1 gZ + (1-\epsilon_1)n_1 Z + (1-\epsilon_2)n_2 Z^2 + k_D D \quad (5)$$

Detritus dynamics:

$$\mathscr{B}_D = -k_D D + \gamma_2 gZ + \epsilon_1 n_1 Z + \epsilon_2 n_2 Z^2 + n_3 P + n_4 P^2 - v_D \frac{dD}{dz} \quad (6)$$

Chlorophyll *a* dynamics ($Chl \doteq \theta^{Chl}_N P$):

$$\frac{dChl}{dt} = \theta^{Chl}_N \frac{dP}{dt} + P\frac{d\theta^{Chl}_N}{dt} = \theta^{Chl}_{dt}\frac{dP}{dt} + \gamma Chl\left(1 - \frac{\theta^{Chl}_N}{\theta^{Chl}_{N_\infty}}\right) \quad (7)$$

where:

$d\theta^{Chl}_N/dt = \gamma\theta^{Chl}_N(1 - \theta^{Chl}_N/\theta^{Chl}_{N_\infty})$ is used with $\theta^{Chl}_{N_\infty} = 1/(\theta^N_{Chl_\infty}(0) + \delta_\infty\ E(x,y,z,t))$

Optical model:

$$E(x,y,z,t) = E(x,y,0,t)e^{-k_w z + k_c \int_0^z Chl dz}$$

Bottom boundary conditions:

$$\kappa_v \frac{dP}{dz} = f^r_P v_P P; \quad \kappa_v \frac{dChl}{dz} = f^r_P v_P Chl; \quad \kappa_v \frac{dD}{dz} = f^r_P v_D D; \quad \kappa_v \frac{dZ}{dz} = 0;$$

$$\kappa_v \frac{dNH_4}{dz} = -f^r f_{NH_4}(f^r_D v_D D + f^r_P v_P P); \quad \kappa_v \frac{dNO_3}{dz} = -f^r(1 - f_{NH_4})(f^r_D v_D D + f^r_P v_P P)$$

Table 2
Model parameters: real time, real-time alternate (from literature), and postcruise values

Parameter	Symbol	Units	Real time value	Alternate value	Postcruise value
Phytoplankton light attenuation scale	k_c	m^2 (mg Chl)$^{-1}$	0.0480	0.02–0.07[a,b]	0.031
PAR over total surface irradiance	–	–	0.45[e]		0.45
H_2O light attenuation scale	k	m^{-1}	0.04		0.04
"Maximum" photosynthesis rate	P_m	mg C (mg Chl)$^{-1}$ s^{-1}	2.8e – 3[e]	4.0e – 4	9.26e – 4[d]
Initial slope of P/I curve	α	mg C (mg Chl)$^{-1}$ s^{-1}/ (μmol photons m^{-2} s^{-1})	1.6e – 5[e]	3.19e – 5	1.5e – 5[d]
Photoinhibition parameter	β	mg C (mg Chl)$^{-1}$ s^{-1}/ (μmol photons m^{-2} s^{-1})	0.0		0.12e – 5
Half-saturation for NO_3 uptake	k_{NO_3}	μmol l^{-1} N	0.5	0.1–10.3[a]	0.7
Half-saturation for NH_4 uptake	k_{NH_4}	μmol l^{-1} N	0.2	0.1–5.7[a]	0.6
NH_4^+ inhibition parameter	ψ	(μmol l^{-1} N)$^{-1}$	3.5	1.5–3.5[a]	5.5
Linear phytoplankton mortality rate	n_3	day^{-1}	0.05	0.045–0.11[a]	0.032
Quadratic phytoplankton mortality rate	n_4	day^{-1} (μmol l^{-1} N)$^{-1}$	0.0		0
Maximum zooplanktongrazing rate	R_m	day^{-1}	0.5	0.1–2.5[c]	0.47
Ivlev constant	Λ	(μmol l^{-1} N)$^{-1}$	0.75		0.24
Percent of zooplanktongrazing excr. as NH_4^+	γ_1		0.16	0.75	0.27
Percent of zooplanktongrazing excr. as detritus	γ_2		0.04		0.16
Linear z grazing loss rate	n_1	day^{-1}	0.05		0.029
Quadratic z grazing loss rate	n_2	day^{-1} (μmol l^{-1} N)$^{-1}$	0.1		0.096
Percent of linear z grazing loss to detritus	ε_1		0.3	0.33	0.3
Percent of quadratic z grazing loss to detritus	ε_2		0.2		0.2
Phytoplankton settling velocity	v_P	m day^{-1}	0.0		0.3
Detritus sinking velocity	v_D	m day^{-1}	1.0	10.0	3.0
Deposited fraction of settling phytoplankton	f_P^r		1.0		0.6
Deposited fraction of sinking detritus	f_D^r		1.0		0.6
Buried fraction of total remineralization	$1-f^r$		0.0		0.0
Fraction of sediment flux that is NH_4	f_{NH_4}		0.0		0.65
Nitrification time scale	k_N	day^{-1}	0.15		0.06
Detritus remineralization time scale	k_D	day^{-1}	0.1		0.19
N:C ratio of phytoplankton	θ_C^N	μmol N (μmol C)$^{-1}$	0.15[e]		0.15
C:Chl ratio $\left(\theta_{Chl}^C \doteq \frac{12}{\theta_N^{Chl}\theta_C^N}\right)$	θ_{Chl}^C	mg C (mg Chl)$^{-1}$	40[e]	10–160[e]	dynamic (100–120)
Photoacclimation linear time scale	γ	day^{-1}	0.0		1/6
$(P{:}Chl)_\infty$ ratio at $E=0$	$\theta_{Chl\infty}^N(0)$	mol N (g Chl)$^{-1}$	–		$100\times\theta_C^N/12=1.25$
Slope of $(P{:}Chl)_\infty$	δ_∞	(mol N (g Chl)$^{-1}$)/ (μmol photons m^{-2} s^{-1})	–		1.2078e – 4

[a] Fasham et al. (1990).
[b] Wroblewski (1977).
[c] Based on MWRA data collected in 1995.
[d] Parsons et al. (1984) and Keller et al. (2001).
[e] Kremer and Nixon (1978) and Hofmann and Ambler (1988).

N.M.F.S., personal communication). However, the biogeochemical data remain relatively limited, which constrains the number of state variables that can be included in the biogeochemical model. The explicit modeling of chlorophyll *a* is nonetheless important as it allows the relatively direct use of satellite images (sea surface color) and in situ fluorometer profiles for model validation and data assimilation. It also allows the parameterization of photoacclimation in terms of a light-dependent carbon-to-chlorophyll *a* ratio. The model equations and conceptual design of the food web are given in Table 1 and Fig. 3, respectively. In the model, fluxes and state variables are expressed in terms of nitrogen (e.g., state variables are in μmol N l^{-1}; except chlorophyll *a*, which is in mg Chl m^{-3}). For all biogeochemical state variables ϕ_i, a four-

dimensional advection–diffusion-reaction (ADR) equation is employed (Table 1). Presently, the ϕ_i values are: [P_{NO_3}, P_{NH_4}, NO_3, Z, NH_4, D, Chl]. The model structures (i.e., the biogeochemical dynamics or reaction term $\mathscr{B}_i$ for each ϕ_i) are defined in Table 1 and described below. The model parameters and their values in the real-time and postcruise simulations are listed in Table 2 (units of nondimensional parameters are left blank).

A.2.1. Biogeochemical model equations

Although fluxes are total biomass, phytoplankton is divided into two state variables in the model implementation: one P_{NO_3} corresponds to new production, or cellular nitrogen resulting from NO_3 uptake, and the other P_{NH_4} to regenerated production, or cellular nitrogen resulting from NH_4 uptake. This division allows a tracking of new and regenerated primary productions and is mainly implemented for the purpose of understanding. Note that it leads to a biogeochemical model with seven instead of six state variables. The variability of the phytoplankton biomass (Eqs. (1) and (2)) is governed by primary production and losses due to grazing and mortality. Phytoplankton growth is controlled by light and nutrient limitation terms. Uptake rates of dissolved inorganic nutrients (V^{NO3}/P and V^{NH4}/P) involve the light-limited growth rate and Michaelis–Menten kinetics for nutrient uptake as proposed by Wroblewski (1977). Light-limited growth is implemented following a three-parameter exponential model (Platt et al., 1980), taking into account the self-shading by Chl-*a* in addition to the water attenuation. The loss of nitrogen from the phytoplankton biomass is represented in the form of linear and quadratic loss rates. The use of a quadratic term can account for higher loss when the phytoplankton concentration is high. The loss of phytoplankton due to settling is implemented as an advection term: it is proportional to the vertical gradient of phytoplankton concentration with a constant settling velocity here.

The modeled zooplankton (Eq. (3)) grazes on phytoplankton only and this is governed by the Ivlev grazing function (Ivlev, 1955). Zooplankton grazing of detritus is not included. The zooplankton mortality is implemented as linear and quadratic loss rates. A quadratic mortality term mainly accounts for higher predators that are not explicitly modeled in the present formulation. The detritus pool in the model (Eq. (6)) is formed from phytoplankton mortality and from fractions of zooplankton mortality and excretion (egestion). Detritus is subject to sinking and, as for the phytoplankton, this is represented as being proportional to the vertical gradient of its concentration. Detritus is directly remineralized to ammonium without explicitly modeling the microbial loop or food web (Azam et al., 1983; Ducklow, 1983; Sherr and Sherr, 1988). Detritus should thus be understood as an aggregation of detritus, dissolved organic nitrogen, and bacteria, which has implications on the values of model parameters (Section 4.2.1). Ammonium (Eq. (5)) is generated by remineralization and fractions of zooplankton mortality and excretion. Ammonium losses are due to phytoplankton uptake and a conversion to nitrate using a nitrification time scale. The nitrate equation (Eq. (4)) involves the new production and nitrification terms. The chlorophyll *a* concentration (Eq. (7)) is assumed to be proportional to phytoplankton biomass in nitrogen units (i.e., $\mathrm{Chl} \doteq \theta_N^{Chl} P$), but the chlorophyll *a*-to-nitrogen ratio of phytoplankton (θ_N^{Chl}) is allowed to vary owing to photoacclimation. The first term in Eq. (7) is simply computed from the phytoplankton equation. The second term parameterizes the hysteresis in photoacclimation dynamics (Cullen and Lewis, 1988; Dusenberry, 2000). It models the delayed tendency of phytoplankton to acclimate to a fully adapted ratio, $\theta_{N\infty}^{Chl}$. This fully adapted ratio is here assumed to be the inverse of a linear function of the local light intensity $E(x,y,z,t)$. The parameters of this function $(P{:}\mathrm{Chl})_\infty$ are the intersect at zero light $\theta_{Chl\,\infty}^{N}$ (0) and the slope δ_∞ [they correspond to 100 (120) mg C (mg Chl)$^{-1}$ at 0 (1000) W m^{-2}; see Table 2].

A.2.2. Biogeochemical model parameters

The estimation of biogeochemical model parameters is challenging, mainly because of the lack of sufficient data, model nonlinearities, and imperfect model structures (e.g., Robinson and Lermusiaux, 2002). For this study, parameters were estimated from a combination of in situ data, literature surveys, and approximate dynamical constraints. In Table 2, the real-time column corresponds to the values employed in forecasts (Section 4.1). The alternate values were also utilized in real-time: they are parameter bounds obtained from the literature and from MWRA data

collected in 1995. After the cruise, a much wider range of values (not reported here) was investigated for all parameters. These postcruise sensitivity studies involved a total of more than 500 different numerical simulations (Section 4.2). The present optimal set of postcruise parameters is listed in the last column of Table 2. Note that based on data–model misfits, several portions of the biogeochemical model were improved or added during the postcruise analysis (e.g., initialization, photoacclimation, surface and bottom boundary conditions, etc.).

A.3. Vertical boundary conditions: atmospheric and sediment fluxes

The physical PE model is forced with surface atmospheric fields [wind stress, fresh water flux (E–P), and net heat flux] computed from the US Navy Fleet Numerical Meteorologic and Oceanographic Center (FNMOC) daily meteorological fields. These fluxes are interpolated in time and then applied as boundary conditions on the surface model level. For the biogeochemical model, short-wave radiation is calculated from the FNMOC daily cloud cover fields. To compute the photosynthetically available radiation (PAR) in the production terms of the P and Chl equations, the daily short-wave radiation is interpolated to hourly values assuming a half-sine curve between sunrise and sunset, giving a maximum at midday. This fast-varying short-wave radiation field is applied at the surface, and the available irradiance in the water column, $E(x,y,z,t)$ (Tables 1 and 2), is computed using an exponential decay optical model. For each real-time forecast simulations, the atmospheric forcings were analyzed in atmospheric fields for all days prior to the start of the forecast and forecast atmospheric fields for all days thereafter. Presently, the last analysis is on September 23. For the postcruise simulations, all atmospheric fields are analysed, if available (they are forecast if not).

At the bottom, a dynamic stress balance involving a drag coefficient is applied to the momentum equations and a Rayleigh friction parameterizes a simple physical boundary layer (Lermusiaux, 1997). Salinity and temperature fluxes are set to zero. For each biogeochemical state variable, the vertical eddy flux at the bottom is set to be balanced with the flux from/to the sediments (Tables 1 and 2). The zooplankton flux is set to zero. As the settling phytoplankton (v_P) and sinking detritus (v_D) reach the bottom, a portion of it remains at the ocean bottom ($1-f_P^r$, $1-f_D^r$) while the remainder is deposited in the sediments (f_P^r, f_D^r). In these sediments, remineralization occurs and a flux of nutrients (ammonium and nitrate) is released from the sediments to the water column. This upward flux of nutrients can be smaller than the downward flux of material due to burial ($1-f^r$). In our model, we assume a local instantaneous equilibrium between the upward nutrient flux (nutrients not buried) and the portions of the downward phytoplankton and detritus fluxes that are deposited but not buried in the sediments. Specifically (Table 2), f_P^r (f_D^r) is the fraction of phytoplankton (detritus) sinking at the bottom that is deposited in the sediments [the rest remain as phytoplankton (detritus) at the bottom], f^r is the fraction of total nitrogen entering the sediments that is not buried but remineralizes at the bottom level, and f_{NH_4} is the fraction of the total dissolved inorganic nitrogen flux from sediments (i.e., ammonium; the nitrate fraction is $1-f_{NH_4}$). The values of f_{NH_4} in coastal regions and in the open ocean differ: in coastal regions, the ammonium flux is usually larger than the nitrate flux ($f_{NH_4}>0.5$, as estimated in Section 4.2).

A.4. Data assimilation scheme

The optimal interpolation scheme of HOPS is employed. Its data forecast melding consists of a two-scale objective analysis of the observations, followed by a blending of the forecast with the OA fields, in accord with the errors of the objectively analyzed data (Robinson, 1996; Lozano et al., 1996; Lermusiaux, 1999).

The OA scheme, which is also employed in the initialization of model fields, is based on the Gauss–Markov or minimum error variance criterion (e.g., Carter and Robinson, 1987). The mapping of observations is carried out by successive corrections on horizontal grids and, presently, it is a two-scale OA. A field ψ is estimated by $\hat{\psi}(\mathbf{r})=\bar{\psi}(\mathbf{r})+\psi'(\mathbf{r})$, where $\mathbf{r}=(x,y)$. The first OA is used to estimate the (spatially variable) mean or largest dynamical scales considered, $\bar{\psi}$ (e.g., the large-scale field). The second OA determines the main scales of interest, ψ' (e.g., the mesoscale). In the first stage, the large-scale departure $\psi_L(\mathbf{r})$ from a horizontally uniform field ψ_U is

estimated, which leads to an estimate of the mean field $\bar{\psi}(\mathbf{r}) = \psi_U + \psi_L(\mathbf{r})$. In the second stage, the mesoscale departure $\psi'(\mathbf{r})$ from this mean field $\bar{\psi}(\mathbf{r})$ is estimated. Most of the time, in the first stage, both historical and synoptic data are utilized, while in the second stage, only synoptic data are employed. This process is repeated for several depths. The form of the correlation function in the two stages of the OA is identical, but its parameters are not. The scalable correlation C between two variations at the points $\mathbf{r}_1$, $\mathbf{r}_2$ occurring at the times t_1, t_2, respectively, is modeled by:

$$C(\mathbf{r}_1, \mathbf{r}_2) = \left(1 - \frac{s^2}{D_0^2}\right) \exp\left[-\frac{s^2}{2D_\ell^2} - \frac{(t_1 - t_2)^2}{2\tau^2}\right],$$

where $s^2 = (\mathbf{r}_1 - \mathbf{r}_2)^T(\mathbf{r}_1 - \mathbf{r}_2)$ with superscript T denoting the transpose of a vector; τ and D_ℓ are the time and space decorrelation scales, respectively; and D_0 is the spatial zero crossing of the correlation function C. Once the observations are mapped into 3-D fields, they are linearly combined (blended) with the corresponding forecast fields, in accord with the a posteriori error variance field of the objectively analyzed data.

Appendix B. Historical and synoptic data sets and their utilization

The data used in this study are divided into: (i) historical data used for model preparation and calibration, (ii) data used for initialization of the real-time simulations, and (iii) data collected during August–September 1998 (MBST-98 period).

B.1. Data for model calibration

The tuning of the dynamical and numerical parameters of the physical model started more than a year prior to MBST-98 (P.J. Haley, personal communication; Lermusiaux, 2001) and involved Observation System Simulation Experiments (e.g., GLOBEC, 1994). The successive physical calibrations were based on previous experience (e.g., Geyer et al., 1992; Signell and List, 1997), historical MWRA (Murray et al., 1997) and UMASS data (G.B. Gardner, personal communication), climatological data (LOC; Lozier et al., 1996), satellite data (Porter, 1998; Rothschild et al., 1999), and FNMOC data for surface atmospheric forcing. The biogeochemical data used for the ecosystem model calibration were the same as for its initialization.

B.2. Data for real-time model initialization

The physical data (Figs. 4a and 5a and b) used to initialize the PE model consisted of real-time MBST-98 hydrographic data collected during August 17–21, 1998 and MWRA data collected during August 1998, and of historical MWRA and climatological LOC hydrographic data to initialize the Outer Cape region (east of Cape Cod; Fig. 1), which was not sampled during the initialization surveys (see Lermusiaux, 2001).

For the calibration and initialization of the real-time ecosystem simulations (Figs. 4b–d and 5c–e), a subset of the MWRA August 1995 data (Murray et al., 1997; Coniaris, 1997) was used. NO_3 data from NODC (http://www.nodc.noaa.gov/JOPI/) and Chl-*a* data from MARMAP (O'Reilly and Zetlin, 1998) were also employed to initialize the offshore part of the model region, which was not sampled by the MWRA.

B.3. Data collected during August–September 1998 and used for assimilation

During MBST-98, from August 17 until October 5, observations were gathered on multiple scales using ships (*R/V Lucky Lady*, UMASS; *R/V Oceanus*, WHOI) and several AUVs (Yu et al., 2002). The gathering occurred in three phases (Rothschild et al., 1999): the initialization surveys (August 17–21), update surveys (September 2–4), and 2 weeks of intensive operations (September 17–October 5). These data were collected based on an adaptive sampling methodology carried out in real time (Robinson et al., 1999; Lermusiaux, 2001). The adaptive sampling plans were designed based on forecasts of fields and on forecasts of dominant error or variability covariances, both of which assimilated the prior data. The physical data consisted of shipboard CTD (conductivity–temperature–depth) data providing temperature and salinity profiles at Bay scale, mesoscale, and submesoscale resolutions, and of AUV data collected during September 17–October 5 at

submesoscale and turbulence scale resolutions but only in restricted areas of interest.

For the ecosystem, the *R/V Lucky Lady* sampled the mesoscales and submesoscales, mainly in Cape Cod Bay. It collected Chl-*a* fluorescence and PAR data with sensors attached to its CTD, and bottle data for NO_3, NH_4, Chl-*a*, and pheaopigment. The measure of zooplankton using acoustic sensors (TAPS; Holliday et al., 1989) was tested and a few uncalibrated *Z* profiles were obtained on October 4–5, 1998. The *R/V Oceanus* sampled the mesoscales, north of Cape Cod Bay and in the open boundary regions. It collected Chl-*a* fluorescence data measured by a fluorometer attached to its CTD. The fluorometer measurements were calibrated to Chl-*a* using bottle data afterwards (G. Strout, personal communication). Further details on the MBST-98 data are reported in Rothschild et al. (1999).

During MBST-98, the MWRA conducted four surveys, which are also used in this study. In two of these surveys (August 18–25 and October 5–16, 1998), 33 stations were visited throughout the Bay, including Boston Harbor. The other two surveys, conducted on September 3 and 24, 1998, were limited to the vicinity of Boston Harbor. In addition to hydrographic profiles, the MWRA collected water column nutrient, Chl-*a*, dissolved oxygen, productivity, respiration, and plankton profiles. These MWRA data are described in Libby et al. (1999).

These MBST-98 and MWRA data sets are assimilated in the postcruise simulations. The station positions for the first series of assimilations are given on Fig. 15 (Fig. 11 gives the station positions for the postcruise initialization). In the real-time simulations, only the physical data sets are assimilated.

References

Anderson, L.A., Robinson, A.R., 2000. Physical and biological modeling in the Gulf Stream region: Part II. Physical and biological processes. Deep-Sea Research 48, 1139–1168.

Anderson, L.A., Robinson, A.R., Lozano, C.J., 2000. Physical and biological modeling in the Gulf Stream region: I. Data assimilation methodology. Deep-Sea Research 47, 1787–1827.

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

Azam, F., Fenchel, T., Field, J.G., Gray, J.S., Meyer-Reil, L.A., Thingstad, F., 1983. The ecological role of water-column microbes in the sea. Marine Ecology Progress Series 10, 257–263.

Bennett, A.F., 1992. Inverse Methods in Physical Oceanography Cambridge Univ. Press, Cambridge. 386 pp.

Brown, W.S., 1998. Wind-forced pressure response of the Gulf of Maine. Journal of Geophysical Research 103, 30661–30678.

Carter, E.F., Robinson, A.R., 1987. Analysis models for the estimation of oceanic fields. Journal of Atmospheric and Oceanic Technology 4, 49–74.

Coniaris, C., 1997. Physical and biological processes controlling nutrient dynamics in Cape Cod Bay. In: Wallace, G.T., Braasch, E.F. (Eds.), Proceedings of the Gulf of Maine Ecosystem Dynamics Scientific Symposium and Workshop, RARGOM Report 97-1. Regional Association for Research on the Gulf of Maine, Hanover, NH, p. 288.

Cullen, J.J., Lewis, M.R., 1988. The kinetics of algal photoadaptation in the context of vertical mixing. Journal of Plankton Research 10, 1039–1063.

Dickey, T.D., 2002. Future ocean observations for interdisciplinary data assimilation systems. Journal of Marine Systems 40–41C, 5–48 (this issue).

Djurcilov, S., Kim, K., Lermusiaux, P.F.J., Pang, A., 2002. Visualizing scalar volumetric data with uncertainty. Computers and Graphics 26 (2), 239–248.

Ducklow, H.W., 1983. Production and fate of bacteria in the ocean. Bioscience 33, 494–501.

Dusenberry, J.A., 2000. Steady-state single cell model simulations of photoacclimation in a vertically mixed layer: implications for biological tracer studies and primary productivity. Journal of Marine Systems 24, 201–220.

Evans, G.T., 1999. The role of local models and data sets in the Joint Global Ocean Flux Study. Deep-Sea Research 46, 1369–1389.

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

Fasham, M.J.R., Ducklow, H.W., McKelvie, S.M., 1990. A nitrogen-based model of plankton dynamics in the oceanic mixed layer. Journal of Marine Research 48, 591–639.

Friedrichs, M.A.M., 2002. Assimilation of JGOFS EqPac and SeaWiFs data into a marine ecosystem model of the central equatorial Pacific Ocean. Deep-Sea Research, Part 2 49, 289–319.

Friedrichs, M.A.M., Hofmann, E.E., 2001. Physical control of biological processes in the central equatorial Pacific. Deep-Sea Research, Part 2 48, 1023–1069.

Gangopadhyay, A., Robinson, A.R., Haley, P.J., Leslie, W.G., Lozano, C.J., Bisagni, J.J., 2002. Feature Oriented Regional Modeling and Simulations (FORMS) in the Gulf of Maine and Georges Bank. Continental Shelf Research 103, 30661–30678.

Geyer, W.R., Gardner, G.B., Brown, W.S., Irish, J., Dutman, B., Loder, T., Signell, R.P., 1992. Physical oceanographic investigation of Massachusetts and Cape Cod Bays. Report to the Massachusetts Bays Program, MBP-92-03. Massachusetts Water Resources Authority, Boston. 497 pp.

Giblin, A.E., Hopkinson, C.S., Tucker, J., 1997. Benthic metabo-

lism and nutrient cycling in Boston Harbor, Massachusetts. Estuaries 20, 346–364.

GLOBEC, 1994. An Advanced Modeling/Observation System (AMOS) for physical–biological–chemical ecosystem research and monitoring (concepts and methodology). GLOBEC Special Contributions, vol. 2. Harvard University, Cambridge, MA. 149 pp.

Haley Jr., P.J., 1996. GRIDS. Harvard Open Ocean Model Reports, vol. 54. Harvard University, Cambridge, MA. 13 pp.

Hofmann, E.E., Ambler, J.W., 1988. Plankton dynamics on the outer southeastern U.S. continental shelf: Part II. A time-dependent biological model. Journal of Marine Research 46, 883–917.

Holliday, D.V., Pieper, R.E., Kleppel, G.S., 1989. Determination of zooplankton size and distribution with multi-frequency acoustic technology. Journal du Conseil—Conseil International pour l'Exploration de la Mer 46, 52–61.

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

Ivlev, V.S., 1955. Experimental Ecology of the Feeding of Fishes. Yale Univ. Press, New Haven. 302 pp.

Keller, A.A., Taylor, C., Oviatt, C., Dorrington, T., Holkombe, G., Reed, L., 2001. Phytoplankton production patterns in Massachusetts Bay and the absence of the 1998 winter–spring bloom. Marine Biology 138, 1051–1062.

Kelly, J.R., 1997. Nitrogen flow and the interaction of Boston Harbor with Massachusetts Bay. Estuaries 20, 365–380.

Kelly, J.R., 1998. Quantification and potential role of ocean nutrient loading to Boston Harbor, Massachusetts, USA. Marine Ecology Progress Series 173, 53–65.

Kelly, J.R., Doering, P.H., 1997. Monitoring and modeling primary production in coastal waters: studies in Massachusetts Bay 1992–1994. Marine Ecology Progress Series 148, 155–168.

Kelly, J.R., Doering, P.H., 1999. Seasonal deepening of the pycnocline in a shallow shelf ecosystem and its influence on near-bottom dissolved oxygen. Marine Ecology Progress Series 178, 151–168.

Kremer, J.N., Nixon, S.W., 1978. A coastal marine ecosystem simulation and analysis. Ecological Studies, vol. 24. Springer-Verlag, Berlin. 217 pp.

Lalli, C.M., Parsons, T.R., 1997. Biological Oceanography: An Introduction. Butterworth, Heinemann. 314 pp.

Lermusiaux, P.F.J., 1997. Error subspace data assimilation methods for ocean field estimation: theory, validation and applications. PhD Thesis. Harvard University.

Lermusiaux, P.F.J., 1999. Data assimilation via error subspace statistical estimation: Part II. Middle Atlantic Bight shelfbreak front simulations and ESSE validation. Monthly Weather Review 127 (7), 1408–1432.

Lermusiaux, P.F.J., 2001. Evolving the subspace of the three-dimensional multiscale ocean variability: Massachusetts Bay. Journal of Marine Systems 29, 385–422.

Lermusiaux, P.F.J., Haley, P.J., Leslie, W.G., Kroujiline, D.K., Robinson, A.R., 1999. Forecasting the physical circulation and variability of Mass. Bay in MBST-98. Appendix VI of AFMIS Report NAG13-48 (Rothschild, B. and the AFMIS Group). NASA.

Libby, P.S., McLeod, L.A., Albro, C.S., Hunt, C.D., Keller, A.A., Oviant, C.A., Turner, J.T., 1999. Semi-Annual Water Column Monitoring Report: August–December 1998, Report ENQUAD, 1999-11. Massachusetts Water Resources Authority, Boston. 131 pp.

Libby, P.S., Albro, C.S., Hunt, C.D., Geyer, W.R., Keller, A.A., Oviant, C.A., Turner, J.T., 2000. 1998 Annual Water Column Monitoring Report, Report ENQUAD, 1999-16. Massachusetts Water Resources Authority, Boston. 173 pp.

Lozano, C.J., Robinson, A.R., Arango, H.G., Gangopadhyay, A., Sloan, N.Q., Haley, P.J., Leslie, W.G., 1996. An interdisciplinary ocean prediction system: assimilation strategies and structured data models. In: Malanotte-Rizzoli, P. (Ed.), Modern Approaches to Data Assimilation in Ocean Modelling. Elsevier Oceanography Series, Elsevier, Cambridge, MA, pp. 413–452.

Lozier, M.S., McCartney, M., Owens, W.B., 1996. The climatology of the North Atlantic. Progress in Oceanography 36, 1–44.

Lynch, D.R., Gentlemen, W.C., McGillicuddy Jr., D.J., Davis, C.S., 1998. Biological/physical simulations of *Calanus finmarchicus* population dynamics in the Gulf of Maine. Marine Ecology Progress Series 169, 189–210.

Malanotte-Rizzoli, P. (Ed.), 1996. Approaches to Data Assimilation in Ocean Modeling. Elsevier Oceanography Series. 455 pp.

Massachusetts Water Resources Authority (MWRA), 2001. Massachusetts Water Resources Authority Contingency Plan Revision 1, Report ENQUAD, ma-071. Massachusetts Water Resources Authority, Boston. See also: http://www.mwra.state.ma.us/harbor/enquad/. 47 pp.

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

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

Murray, P.M., Cibik, S.J., Lemieux, K.B., Zavistoski, R.A., Morton, J.E., Howes, B.L., Taylor, C.D., Loder, T.C., 1997. Semi-Annual Water Column Monitoring Report: August–December 1995, Report ENQUAD, 1997-07 Massachusetts Water Resources Authority, Boston. 280 pp.

Nowicki, B.L., Kelly, J.R., Requintina, E., Van Keuren, D., 1997. Nitrogen losses through sediment denitrification in Boston Harbor and Massachusetts Bay. Estuaries 20, 626–639.

O'Reilly, J.E., Zetlin, C., 1998. Seasonal, horizontal, and vertical distribution of phytoplankton chlorophyll *a* in the northeast U.S. continental shelf ecosystem. NOAA Technical Report NMFS 139. Scientific Publications Office, National Marine Fisheries Service, Seattle, WA. 154 pp.

Orlanski, I., 1976. A simple boundary condition for unbounded hyperbolic flows. Journal of Computatonal Physics 41, 251–269.

Parsons, T.R., Takahashi, M., Hargrave, B., 1984. Biological Oceanographic Processes, 3rd ed. Pergamon, Oxford. 330 pp.

Platt, T., Gallegos, C.L., Harrison, W.G., 1980. Photoinhibition of photosynthesis in natural assemblages of marine phytoplankton. Journal of Marine Research 38, 687–701.

Porter, D.L., 1998. Ocean Remote Sensing. Johns Hopkins University, http://fermi.jhuapl.edu/index.html.

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

Redfield, A.C., Ketchum, B.H., Richards, F.A., 1963. The influence of organisms on the composition of sea-water. In: Hill, M.N. (Ed.), The Sea, vol. 2. Interscience, New York, pp. 26–77.

Robinson, A.R., 1996. Physical processes, field estimation and an approach to interdisciplinary ocean modeling. Earth Science Reviews 39, 3–54.

Robinson, A.R., 1999. Forecasting and simulating coastal ocean processes and variabilities with the Harvard Ocean Prediction System. In: Mooers, C.N.K. (Ed.), Coastal Ocean Prediction. American Geophysical Union, Washington, DC, pp. 77–100.

Robinson, A.R., Lermusiaux, P.F.J., 2001. Data assimilation in models. Encyclopedia of Ocean Sciences. Academic Press, London, pp. 623–634.

Robinson, A.R., Lermusiaux, P.F.J., 2002. Data assimilation for modeling and predicting coupled physical–biological interactions in the sea. In: Robinson, A.R., McCarthy, J.J., Rothschild, B.J. (Eds.), The SeaBiological–Physical Interactions in the Ocean, vol. 12. Wiley, New York, pp. 475–536.

Robinson, A.R., and the LOOPS Group, 1999. Real-time forecasting of the multidisciplinary coastal ocean with the Littoral Ocean Observing and Predicting System (LOOPS). Preprint Volume of the Third Conference on Coastal Atmospheric and Oceanic Prediction and Processes, New Orleans, LA, November 3–5, 1999. American Meteorological Society, Boston, MA.

Robinson, A.R., Lermusiaux, P.F.J., Sloan, N.Q., 1998. Data assimilation. In: Robinson, A.R., Brink, K. (Eds.), The Sea. The Global Coastal Ocean: Processes and Methods, vol. 10. Wiley, New York, pp. 541–594.

Robinson, A.R., Sellschopp, J., Leslie, W.G., Onken, R., Alvarez, A., Baldasserini, G., Haley Jr., P.J., Lermusiaux, P.J.F., Lozano, C.J., Nacini, E., Stoner, R., Zanasca, P., 2002. Real-time forecasting of synoptic transients in the Eastern Ligurian Sea. Journal of Marine Systems (in review).

Rothschild, B., and the AFMIS Group, 1999. Advanced Fisheries Management Information System, AFMIS Report NAG13-48 NASA, Dartmouth, MA.

Sherr, E., Sherr, B., 1988. Role of microbes in pelagic food webs: a revised concept. Limnology and Oceanography 33, 1225–1227.

Signell, R.P., List, J.H., 1997. Effect of wave-enhanced bottom friction on storm-driven circulation in Massachusetts Bay. Journal of the Waterway, Port, Coastal and Ocean Engineering-ASCE 123, 233–239.

Signell, R.P., Jenter, H.L., Blumberg, A.F., 1993. Modeling the seasonal circulation in Massachusetts Bay. Proceedings of the 3rd International Conference, Sponsored by Waterway, September, 1993. Estuarine and Coastal Modeling, vol. III. American Society of Civil Engineers, Oak Brook, IL.

Spall, M.A., Robinson, A.R., 1990. Regional primitive equation studies of the gulf stream meander and ring formation region. Journal of Physical Oceanography 20, 985–1016.

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

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

Turner, J.T., 1992. Planktonic copepods of Boston Harbor, Massachusetts Bay and Cape Cod Bay. Hydrobiologia 293, 405–413.

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

V'ezina, A.F., Pace, M.L., 1994. An inverse model analysis of planktonic food webs in experimental lake. Canadian Journal of Fisheries and Aquatic Sciences 51 (9), 2034–2044.

Wallace, G.T., Braasch, E.F. (Eds.), 1997. Proceedings of the Gulf of Maine Ecosystem Dynamics, A Scientific Symposium and Workshop, RARGOM Report 97-1. Regional Association for Research on the Gulf of Maine, Hanover, NH.

Werme, C., Hunt, C.D., 2000. 1998 Outfall Monitoring Overview, Report ENQUAD, 2000-04. Massachusetts Water Resources Authority, Boston. 66 pp.

Wroblewski, J.S., 1977. A model of phytoplankton plume formation during variable Oregon upwelling. Journal of Marine Research 35, 357–394.

Wunsch, C., 1996. The Ocean Circulation Inverse Problem. Cambridge Univ. Press, Cambridge. 456 pp.

Yu, X., Dickey, T., Bellingham, J., Manov, D., Streilein, K., 2002. The application of autonomous underwater vehicles for interdisciplinary measurements in Massachusetts and Cape Cod Bays. Continental Shelf Research 22, 2225–2245.