The seasonal nitrogen cycle in Wilkinson Basin, Gulf of Maine, as estimated by 1-D biological model optimization

Laurence A. Anderson

Department of Applied Ocean Physics and Engineering, Woods Hole Oceanographic Institution, Woods Hole, MA 02543, U.S.A.

Abstract

The objective of this study was to fit a simple ecosystem model to climatological nitrogen cycle data in the Gulf of Maine, in order to calibrate the biological model for use in future 3-D modelling studies. First depth-dependent monthly climatologies of nitrate, ammonium, chlorophyll, zooplankton, detritus and primary production data from Wilkinson Basin, Gulf of Maine, were created. A 6-box nitrogen-based ecosystem model was objectively fitted to the data through parameter optimization. Optimization was based on weighted least squares with model-data misfits nondimensionalized by assigned uncertainties in the monthly climatological estimates. These uncertainties were estimated as the standard deviations of the raw data from the 6-meter and monthly bin averages. On average the model fits the monthly means almost within their assigned uncertainties.

Several statistics are examined to assess model-data misfit. Pattern statistics such as the correlation coefficient lack practical significance when data errors are large relative to the signal, as in this application. Thus Taylor diagrams were not found to be useful. The RMSE and model bias normalized by the data error were found to be the most useful skill metrics as they indicate whether the model fits the data within its estimated error.

The optimal simulated nitrogen cycle budgets are presented, as an estimate of the seasonal nitrogen cycle in Wilkinson Basin, and discussed in context of the available data. Wilkinson Basin has spring and fall phytoplankton blooms, and strong summer stratification with a deep chlorophyll maximum near 21 m, just above the nitracline. The mean euphotic zone depth is estimated to be 25 m, relatively constant with season. The model estimates annual primary production as 176 g C m^{-2} yr^{-1}, annual new production as 71 g C m^{-2} yr^{-1} and sinking PON fluxes of 9.7 and 4.7 g N m^{-2} yr^{-1} at 24 and 198 m respectively.

Areas for improvement in the biological model, the model optimization method, and significant data gaps are identified.

Key words:

1 Introduction

The nitrogen cycle and phytoplankton biomass in the Gulf of Maine are of interest with regards to assessing the impact of wastewater discharge (Hydroqual, 1995; Libby et al., 2001; Werme and Hunt, 2004), the ecosystem dynamics in a region of significant commercial fisheries (Franks and Chen, 1996, 2001; Steele et al., 2007) and the occurance of harmful algal blooms and eutrophication (Anderson, 1997; Anderson et al., 2002). Diagnosis of the mechanisms responsible for observed plankton and nutrient distributions is a difficult task however, due to the interaction of multiple physical and biological processes. That is, it is difficult to obtain all the relevant types of measurements on sufficient space and time scales to infer with certainty the dynamics that cause observed distributions.

Here the time- and depth-dependent climatological seasonal cycles of nitrogen and phytoplankton biomass in Wilkinson Basin are estimated through model optimization. The biological model equations are a hypothesis about how different nitrogen cycle components are dynamically related. As the ocean ecosystem is enormously complex, any biological model is a simplification; thus the primary hypothesis is that the proposed model is sufficient to simulate the variables of interest. Optimization of biological model parameter values, within observed parameter limits, is used to obtain simultaneous agreement with various data types. Comparison with both assimilated and unassimilated data is used for model skill assessment and validation. If sufficiently validated, a numerical model's field estimates are a dynamically-consistent interpolation and extrapolation of the data that facilitates a consistent budget analysis.

Previous models of phytoplankton biomass in the Gulf of Maine include a 2-box P-Z model for Georges Bank (Wallhead et al., 2009), 3-box N-P-Z models for Georges Bank (Klein, 1987; Lewis et al., 1994; Franks and Chen, 1996) and the Gulf of Maine (Campbell, 1986; Townsend et al., 1994; Franks and Chen, 2001; Tian and Chen, 2006), a 4-box N-P-Z-D model for the Gulf of Maine (Ji et al., 2008), a 6-box model used in Massachusetts Bay (Besiktepe et al., 2003), a 9-box model used on Georges Bank (Ji et al., 2006a,b), 16-box models for Georges Bank (Steele et al., 2007; Steele , 2009) and Cobscook Bay (Campbell, 2004), a 24-box model for the Gulf of Maine (Zhang and Chen, 2007) and a 23- to 25-box water quality model for Massachusetts Bay (Hydroqual, 1995, 2000, 2003; Jiang and Zhou, 2003, 2004, 2006, 2007). Of these, only a few examine the seasonally-resolved annual cycle (Klein, 1987; Steele et al., 2007; Steele , 2009; Hydroqual, 1995, 2000, 2003; Jiang and Zhou, 2003, 2004, 2006; Ji et al., 2008), of which only Steele et al. (2007) and Steele (2009) have used data for model optimization, though under an assumption of steady state.

Here a 6-box ecosystem model is used based on those of Anderson et al. (2000) and Besiktepe et al. (2003). The biological model parameter values are calibrated by objectively fitting the model to depth-dependent monthly data of nitrate, ammonium, chlorophyll, zooplankton, detritus and primary production from Wilkinson Basin. This is the first study that optimizes agreement with several different data types to estimate the depth-dependent seasonal cycles of nitrate, phytoplankton, zooplankton and detritus in Wilkinson Basin.

The focus is on Wilkinson Basin because, of the various regions of the Gulf of Maine, the deep basins are closest to being closed systems, minimizing the impact of lateral advection that is not included in this 1-D model. The shallow areas along the coast have the strong throughflow of the coastal current and are intermittently influenced by coastal upwelling and downwelling, while for Georges Bank on-bank fluxes of nutrients are important on both tidal and annual time scales (Loder et al., 1982; Franks and Chen, 1996; Steele et al., 2007; Ji et al., 2008). One spatial dimension (depth) makes parameter optimization more tractable and facilitates estimation of the seasonal cycles. This Wilkinson Basin study provides an interesting comparison with similar studies for nearby regions (Sec. 5.7).

The model and optimization method are presented in Sec. 2, and the data in Sec. 3. Skill assessment is a vital component of any modeling exercise; this is presented in Sec. 4. The model fits the data almost within its assigned errors on average (Sec. 4); this gives us sufficient confidence to present the simulated budgets (Sec. 5). In Sec. 5 the model validation is further examined through comparison with data-based estimates that were not assimilated. Primary conclusions are summarized in Sec. 6.

2 Model

2.1 Biological model

The model code used is the Water Column Biogeochemical Modelling Workbench developed by G. Evans (Fasham and Evans, 1995; Evans, 1999) with slight modification as noted here. It is a time- and depth-dependent model in which physical processes (mixed-layer depth, vertical diffusivities and vertical velocity) and biological initial conditions are specified, and biological parameter values are optimized by minimizing a cost function to fit observations.

A biological model was desired that explicitly modeled bottom-up (light and nutrient) and top-down (predator) controls on phytoplankton biomass, particularly as the time-dependence of the latter is thought to significantly influence the spring bloom (Keller et al., 2001). It was also desired to include (a) ammonium, so that new production could be separated from recycled production, (b) sinking detritus, which causes the total nitrogen content of a water parcel not to be conserved (e.g. Klein, 1987), and (c) chlorophyll, as

all phytoplankton observations are in units chlorophyll, and phytoplankton nitrogen-to-chlorophyll ratios are not constant. The proposed model thus has six components: nitrate (N), ammonium (A), phytoplankton biomass (P), chlorophyll (Chl), zooplankton (Z) and detritus (D). Bacteria and DON are also important ecosystem components, but since no data could be found on these in Wilkinson Basin with which to constrain the model, it was attempted to model these implicitly. The model is based on Anderson et al. (2000) and similar to that of Besiktepe et al. (2003); the Chl evolution equation was inspired directly by Besiktepe et al. (2003), though slightly modified.

The biological model equations, variables and parameters are described in Tables 1, 2 and 3. Upper-case roman letters denote variables in time and space (i.e. depth); lower-case roman and greek letters denote constants. Note that in the central run some of the parameter values are zero, as determined by optimization, which disables some parameterizations.

The biological model equations are fairly standard. For phytoplankton growth, the minimum of light- and nutrient-limitation factors is used in keeping with Liebig's Law of the Minimum (Ondercin et al., 1995; Hurtt and Armstrong, 1996). The equilibrium carbon-to-chlorophyll ratio (R_o) is a linear function of light (Geider, 1987; Geider et al., 1996, 1997; Zonneveld, 1998; Christian et al., 2002; Besiktepe et al., 2003). Chlorophyll is subject to all the physical and biological sources and sinks that affect phytoplankton, plus an adjustment toward its equilibrium concentration.

Zooplankton graze phytoplankton, detritus and other zooplankton (McCreary et al., 1996). In practice, Z grazing on Z allows grazing rates, assimilation efficiencies and total biomass to be more consistent with experiment-based estimates. The squared terms in the Z grazing parameterizations imitate a threshold at low prey concentrations (Steele and Henderson, 1992). Phytoplankton and zooplankton mortality were parameterized as quadratics; it was decided to allow the optimization to determine whether the linear or squared terms allow better agreement with the data. Detritus remineralization was also parameterized as a quadratic; for a constant sinking rate, the squared term is more consistent with the observation that the sinking particle flux typically decreases inversely ($1/z$) with depth (Martin et al., 1987). Nitrification is light inhibited (Ward, 2000); a quadratic nitrification term is included to explore whether it is better to assume nitrifying bacteria biomass is constant or proportional to A concentration.

Sinking rates for P and D are concentration-dependent, with a minimum value at a threshold concentration, and a maximum value. For P, this represents the negligible sinking rate of small phytoplankton (which dominate when P biomass is low), and the fast sinking rate of large phytoplankton (which dominate when P biomass is high), plus aggregation. For D, this allows the background D concentration (near 1 μM N) to have a low sinking rate, while pulses of high D will have a faster sinking rate. Chl sinks at the same rate as P.

The time-dependent mixed-layer depth was prescribed based on data as

described in Sec. 3. Vertical diffusivity was specified within the mixed layer as 100 cm^2 sec^{-1} and below it as 0.3 cm^2 sec^{-1} (Townsend, 1991; Benitez-Nelson et al., 2000). Vertical velocity was set to zero.

Daily mean photosynthetically-available radiation (PAR) at the sea surface was set as 43% of the daily mean solar radiation, which in turn was determined as a continuous function of yearday from astronomical considerations for 43°N using a mean atmospheric transmission of 51.76% computed from shortwave radiation data at Woods Hole, Massachusetts, from the spring of 1993. The diurnal light cycle is not included, as the daily primary productivity rates and other data used to constrain the model do not resolve diurnal variability.

According to table 2-8 in Hydroqual (1995), wet plus dry atmospheric deposition for Massachusetts Bay (adjacent to Wilkinson Basin) is 69.9 mmol NO_3 m^{-2} yr^{-1} and 17.9 mmol NH_4 m^{-2} yr^{-1} (based on 1.15 m yr^{-1} precipitation, from their table 2-9). These mean fluxes are applied at the surface boundary of the model. While atmospheric nitrogen deposition is actually episodic, it would not be appropriate to force this climatological simulation with submonthly variability; it would be appropriate to apply monthly variability, but no seasonality is apparent (Jordan and Talbot, 2000). At the bottom boundary, vertical diffusive fluxes are set to zero i.e. concentrations "below" the boundary are assumed identical to those just above, though D and P are allowed to sink out through the bottom boundary. Because of this open boundary condition, the model is not required to conserve total nitrogen.

The vertical resolution of the model is 6 meters, to a total depth of 198 meters. Timestep is determined internally in the code, based on the time rate of change of the biological state variables, and varied between 1 hour and 2 days.

2.2 Cost function

To fit the model to the data, the minimum in the following cost function was sought

$$Cost = \left[\frac{1}{n} \sum_{i=1}^{n} \left(\frac{m_i - d_i}{\hat{\sigma}_i} \right)^2 \right]^{1/2} \tag{1}$$

where d_i is the assimilated data (primarily monthly and 6-meter bin averages; Sec. 3), m_i is the model estimate, $\hat{\sigma}_i$ is an estimate of the uncertainty in that data value and n is the number of observations. The assimilated data types are nitrate (NO_3), ammonium (NH_4), Chlorophyll a, vertically-integrated zooplankton biomass, detritus and primary production rates. The optimization method is thus weighted least squares. A weight is necessary for nondimensionalization because multiple data types with different units are used. These weights should be determined by objective means. Previous studies have used weights based on global data means (Evans, 2003), global data

variances (Friedrichs et al., 2007) or observational errors (Matear, 1995). (The term "global" in this context refers to an average over all data of that data type.) Here this last view is taken—that, to measure skill, model-data misfit should be weighed against the uncertainty in that data point (see Chapter 14 in Press et al., 1986; Tarantola, 1987). In this climatological simulation, the assigned uncertainty includes both observational error and variability unresolved by the model (see below). Minimization of Eq. 1 asks the model to fit each data point relative to its individual uncertainty, with a target misfit of zero. The outer part of Eq. 1 is designed to give the Cost a meaningful value. If the model-data misfit is equal to the uncertainty at every point, then the Cost is 1.0; if the misfit is twice the uncertainty everywhere then the Cost is 2.0, and so on.

In this study of the climatological seasonal cycle of Wilkinson Basin, the raw data is not a true time series, but collected over many years and includes spatial and interannual variability, particularly differences in the timing of the spring bloom due to interannual differences in mixed layer depth. Conseqently it is not sensible to fit the model to all the original data points, but rather the model is fit to climatological monthly means computed from the original data. The construction of monthly climatologies was in fact the first objective of this study. These monthly means are the d_i in Eq. 1. It is next considered how to assign the uncertainties in these means, including the degree to which one should expect a model to fit a climatology.

In a model-data comparison, the assigned data error should include not only the measurement error (σ_m), estimated as the standard deviation of replicate measurements, but also the "representativeness error" (Ivanov and Palamarchuk, 2007), which is the error associated with comparing a measurement made on a small space and time scale with output from a model that only resolves a much larger space and time scale, particularly when a significant amount of variability is known to occur in nature on the model's spatiotemporal subgridscale. For submonthly data with a standard deviation of $\hat{\sigma}$, the representativeness error of a single data point (i.e. the error estimate for it being an accurate estimate of the monthly mean) is $\hat{\sigma}$. Thus an estimate of the true monthly mean computed from n submonthly data points has a representativeness error of $\hat{\sigma}/\sqrt{n}$, which is the sampling (standard) error. This suggests that the monthly mean estimates d_i should be assigned uncertainties computed as the standard errors of the data within each monthly-depth bin ($\hat{\sigma}_i/\sqrt{n_i}$), which is generally greater than σ_m.

The computed standard errors however were considered lower limit estimates of the uncertainty for a few reasons. The first is the correlation between data points. The value of n_i in the denominator should be reduced if the data in a given year are within the dominant submonthly spatiotemporal correlation scales i.e. are not independent. Interannual correlations also need to be considered e.g. decadal trends and the North Atlantic Oscillation. To compute these correlation scales requires much greater data density than this dataset has. Thus while the monthly means were simply computed from all the data

points, it was recognized that the error should be normalized by something significantly less than $\sqrt{n_i}$. The second reason is the issue of unknown inter-cruise calibration offsets. Some cruise data show offsets from other cruises, and the standard deviation (rather than the standard error) better estimates such calibration error, if these are calibration errors and not natural variability. Furthermore there was a lack of confidence that fitting a climatological mean within its standard error is phenomenologically correct (see Section 4.6 in Wunsch, 1996). For example, some years the nitrate drawdown occurs a month earlier than in other years; the climatological mean is a slow drawdown over two months. Yet when forcing the model with climatological mean mixed-layer depths, the bloom will initiate once the "critical depth" is reached (Sverdrup, 1953), such that, if the uptake rates are correct, the slow climatological nitrate drawdown should not be closely reproduced. This suggests leniency should be granted in fitting the climatological means when the observed variability is high.

Consequently $\hat{\sigma}_i$ in Eq. 1 was based on the standard deviations (rather than standard errors) of the original data in each monthly-depth bin. The standard deviations can be considered upper limit estimates of the uncertainty; they are an accurate estimate of the error if the cruise offsets (mentioned above) are calibration errors, though there is no way of knowing if this is always the case. Normalizing the misfit by the standard deviation makes the method similar to maximum likelihood estimation and Bayesian linear regression, discussed below. In effect the model is asked to fit the climatology relative to the observed variability or likelihood. It was decided to first see if the model could fit the data within these upper limit error bounds; if so, the assigned errors could afterwards be reduced to affect a closer agreement to climatological means.

The value of $\hat{\sigma}_i$ was computed as the mean of the standard deviations computed for each month and depth data bin (Sec. 3), which was 1.47 μM for NO_3, 0.27 μM for NH_4, 0.22 mg m^{-3} for Chlorophyll a, 10.8 mmol N m^{-2} for vertically-integrated zooplankton and 0.86 μM N for detritus. For the primary production data, the standard deviations were found to decrease significantly with depth, so $\hat{\sigma}_i$ was allowed to be depth-dependent, computed as 0.27 mmol N m^{-3} d^{-1} above 18 m, 0.16 at 21 m, 0.08 at 27 m, 0.04 at 33 m and 0.03 mmol N m^{-3} d^{-1} below 36 m. In the general case $\hat{\sigma}_i$ varies for each data point (each data bin, in this study), as the standard deviation of the data within that bin (Press et al., 1986). However in this study many bins contain only one data point, and those that contain two have widely-varying estimates of $\hat{\sigma}_i$. Thus due to lack of data, lack of clear spatial and temporal trends in $\hat{\sigma}_i$, and for simplicity of application and interpretation, a single $\hat{\sigma}_i$ is used for each data type (except for PP).

It will be seen that the error estimates based on submonthly variability are fairly large relative to the monthly means, particularly for NH_4. Some of this NH_4 variability may be natural, though some may be due to measurement inaccuracy (Brzezinski, 1988). The large NH_4 $\hat{\sigma}_i$ gives the NH_4 model-data

misfit little influence in the Cost. It is still desirable to assimilate these data however, to avoid the pitfall revealed by Armstrong et al. (1995), that optimization can drive unconstrained state variables to unrealistic values. These large uncertainties affect the choice of skill metrics (Sec. 4.3).

This method of weighted least squares with weights estimated from the deviation of the original data from the large-scale trend is similar to maximum likelihood estimation (MLE)(Wunsch, 1996; Bennet, 2002; Evensen, 2007). However in many applications of MLE the deviation of the data from the trend is considered to be a globally-constant observational error that is estimated, whereas here the model is fit to the monthly means rather than the original data and the data errors are estimated prior to optimization. Also MLE generally involves parameter optimization based on assumed probability distributions of parameter values, which are not assumed here (Sec. 2.3). Thus this method is a variant or subset of MLE. The MLE context reveals that Gaussian-distributed errors are assumed and correlations between errors ignored (i.e. a diagonal covariance matrix assumed).

Similarly, this method can be seen as a subset of the even broader approach of Bayesian linear regression (Gelman, 2004; Lee, 2004; Bolstad, 2004). Bayesian estimation can include in the Cost function a mathematical version of Occam's razor, which penalizes the use of more parameters that do not statistically improve fit over the use of fewer parameters. This approach is not considered for three reasons. First, model formulas were preferred that were based on scientific understanding of the mechanisms, rather than simplicity, which should yield better predictions under varied forcing, and which permits evaluation of the proposed formulas. Secondly, what was considered to be the fewest acceptable number of state variables (N-P-Z-D-A-Chl) and simple formulas (linear and quadratic terms) for poorly-understood processes are already used. Third, the model actually does not fit all the data adequately, which suggests the model is still incomplete. Had the model been able to fit all the data within its (leniently) assigned errors, then one would be in a position to apply Occam's razor to find the simplest model that still explains the data.

Some studies have used Cost functions that further inversely weight each data type by the number of observations of that type, in order to give equal influence of each data type in the Cost (e.g. Friedrichs et al., 2007). This was tried, and not surprisingly the few Z data were better fit at the expense of fit to the more numerous NO_3, Chl and PP data. Since the primary objective is to estimate the nitrogen and phytoplankton cycles, this result was not considered preferable. The decision is probably best decided based on one's data and application viz. whether the variable of interest is the most or least well-sampled.

2.3 Parameter optimization method

For clarity, a "run" is defined as a single optimization procedure run to completion (convergence). A run is comprised of many "iterations", each of

which is a 1-year forward simulation starting on January 1 and conducted with a specific set of parameter values, and for which the Cost is computed. The "results" of a run are that of the final best-fit iteration.

In the Water Column Biogeochemical Modelling Workbench code, Powell's conjugate direction method is used for optimization. The parameter values start at initial estimates, and sequentially each parameter value is perturbed until the local minimum in the Cost is found; then perturbations are made in the vector direction of the combined parameter modifications to further reduce the Cost more efficiently. The process is repeated until convergence is achieved i.e. further iterations do not decrease the Cost by a prescribed amount. The parameters are given initial values as well as maximum and minimum limits, which govern the amount the parameters are perturbed between iterations.

Unlike Evans (1999), the misfits of the parameters from their initial values are not included in the cost function, for two reasons. The first is that the parameter probability distributions are not known. For example, a maximum growth rate of 1.5 day^{-1} seems just as acceptable as 1.0 day^{-1}. In essence, the probability distributions are considered flat between minimum and maximum limits. Perhaps in the future enough measurements will have been made in Wilkinson Basin to estimate probability distributions of these model parameters. Upper and lower limits are set however, based on observations (though not exclusively from the Gulf of Maine). The second reason is that, for an optimal nitrogen cycle budget, the best possible fit to the data is desired. Including parameter misfit in the Cost sacrifices fit to the data for fit to parameter initial guesses. Since the model does not yet fit all the data adequately, I was not yet willing to make that sacrifice.

3 Data

The first objective of the study was to create a depth-dependent monthly climatology of the available data regarding the nitrogen cycle in Wilkinson Basin. The model is merely an attempt to interpolate this data and combine it into a consistent mass-conserving budget.

Chlorophyll estimates were based on a monthly version of O'Reilly and Zetlin (1998) bi-monthly Chl data for the Wilkinson Basin "tiles". The original data is binned in 11 layers from 0 to 113 meters; these were linearly interpolated to the model levels. The data (Fig. 1a) show a major spring bloom, a minor fall bloom, and a summer subsurface maximum.

NO_3 and NH_4 data were from the Garside et al. (1996) dataset (post-1965 data only), augmented with additional data (up to 1991) retrieved in March 2002 from the NODC online archive, 1996 cruise data (Graziano et al., 2000), 1997-1999 Georges Bank cruise data (Townsend and Thomas, 2001, 2002) and 1998-2001 ECOHAB data (Townsend et al., 2001, 2005; Love et al., 2005). Stations were used that were located within the Wilkinson Basin tiles defined by O'Reilly and Zetlin (1998). Obvious outliers were removed; the NODC data

in particular required significant quality control, as it also contained misnavigations (decimal degrees misinterpreted as minutes) and misidentified data (data with the wrong column header), most of which were found by comparison with the Garside dataset. The data were binned monthly and into 6-meter vertical intervals for NO_3 and 18-meter vertical intervals for NH_4 (due to less available NH_4 data). Means for each bin were computed from the $\log_{10}$ of concentration, because distributions were non-normal. The resultant estimates (Fig. 1b,d) show some short time- and length-scale variance, presumably due to a combination of measurement error and natural variability. Consequently, smoother NO_3 and NH_4 estimates (Fig. 1c,e) were computed at mid-month from all data in each depth bin using a 30-day e-folding scale ($x(t) = \sum_{i=1} x_i \exp(-(t-t_i)^2/(30\text{ days})^2)$, smoothed vertically with a 1-2-1 filter ($x_k^{new} = 0.25x_{k-1} + 0.5x_k + 0.25x_{k+1}$). The smoothed estimate was subsampled to have data at the same depths and times as the binned dataset.

No Sep-Dec NH_4 data could be found. The NH_4 data were therefore interpolated to December 15, to constrain concentrations at the end of the year to be similar to observed January values.

Zooplankton biomass (in mmol N m^{-2}) was estimated as two-thirds copepods (dominated by *Calanus finmarchicus*) and one-third microheterotrophs (Schlitz and Cohen, 1984; Davis, 1987; Sherman et al., 1987; Cohen and Grosslein, 1987; Meise and O'Reilly, 1996). The copepod biomass estimate was based on the Meise and O'Reilly (1996) bimonthly *Calanus finmarchicus* atlas for the Wilkinson Basin tiles defined by O'Reilly and Zetlin (1998). Vertically-integrated abundance (# m^{-2}) was converted to biomass (mmol N m^{-2}) using ratios of 11.07 mg N per individual fifth-stage copepodite (C5) and 23.39 mg N per adult female (C6F) (Durbin et al., 1995, their table 1), assuming primarily C5s in May-Dec, C6Fs in Mar-Apr, and 50% of each in Jan-Feb (Meise and O'Reilly, 1996, their fig. 4). Microheterotroph biomass was assumed proportional to the vertically-integrated Chl estimate i.e. with spring and fall maxima (e.g. Montagnes et al., 1988). The resultant total zooplankton biomass estimate (Fig. 1f) show a maximum in May-June and a winter minimum.

Detritus (non-living PON) was estimated as PON $- r_{N:Chl}$Chl, where $r_{N:Chl}$ is a conversion factor assuming 50 g C (g Chl)$^{-1}$ and 6.625 mol C (mol N)$^{-1}$. The PON and Chl data were taken from the Massachusetts Water Resources Authority's (MWRA) 1992-2002 Massachusetts Bay "Boundary" and "Offshore" stations (see Libby et al., 2001, for station locations). The data was binned temporally at the mean times of the bimonthly cruises, and vertically based on availability. The data show significant seasonality, with a maximum in June, but little relationship with depth (Fig. 1g). As with NO_3 and NH_4, a smoothed version was also computed (Fig. 1h) by smoothing once vertically with a 1-2-1 filter; temporal smoothing was not necessary. Note the effect of the smoothing is generally much less than the error estimate (0.86 μM N for D).

Primary productivity data were obtained from the IMCS Primary Pro-

ductivity Database (*http://marine.rutgers.edu/opp/Database/DB.html*) and Graziano et al. (2000), converted with Redfield's C:N ratio. The data were put into monthly and 6-m bins (Fig. 1j); a smoothed estimate was also computed using a 30-day e-folding timescale and applying a 1-2-1 vertical filter (Fig. 1k). The data show spring and fall maxima, and often a subsurface maximum at 9 m.

Mixed-layer depth (MLD) was determined as follows. Potential density was computed from T and S observations associated with the nutrient dataset, and estimated at bi-weekly and 6-m vertical intervals using a 30-day temporal e-folding scale and a 15-meter vertical e-folding scale. MLD was then computed as the depth at which potential density exceeded surface density by 0.125 kg m^{-3} (Levitus, 1982), though adjusted in fall and winter to agree with the NO_3 data, and is shown in Fig. 1i. (The NO_3 data does not give a better estimate, but as the NO_3 data is assimilated, a mismatch would be problematic.) The very shallow summer MLD (6 m) is confirmed from individual CTD profiles. In nature sub-monthly MLD variability occurs; this is not an important omission in this model in summer (as then the upper 24 m are nutrient-limited) but it is a relevant concern in winter.

Initial conditions for the biological state variables were created by temporally interpolating the data to January 1. P was initialized based on the Chl data, using for conversion R_o from Table 1 and the January 1 light field estimate.

Only data between 0 and 99 meters were assimilated. This approach is based on the observation that biological rates generally decrease with depth, such that the variability in the data below 100 meters is likely more influenced by physical processes (mixed layer depth, isopycnal displacement, advection and diffusion), while the data above 100 m is better suited for optimizing the biological model parameter values. For example, the individual cruise data have measurements (from Niskin bottles) typically every 10 to 50 m, such that the small-scale variability in N below 100 m (Fig. 1b) is primarily due to adjacent depth bins containing data from different cruises; requiring the model to fit this deep variability may result in unusual biological parameter values. The smoothed data estimates were assimilated (viz. Fig. 1c, e, h and k) to minimize the impact of high-frequency variations (attributed to undersampling and unresolved variance) on the parameter optimization.

4 Results

Table 4 lists some of the runs performed. Run 1 is the "central" run, with the lowest cost; the other runs are sensitivity tests in which a parameter of Run 1 was held to a constant value or constrained differently and the full optimization procedure repeated. Runs 10, 18, 32, 35, 40 and 71 actually have slightly lower costs than Run 1, but by an amount smaller than the convergence criterion to stop optimization (0.003), such that the costs are

effectively indistinguishable. The total number of Runs was 197; Table 4 only lists the Runs that surround the final solution (Run 1) in parameter space.

4.1 Reduction in Cost

Model misfit to the data for a typical run (viz. Run 1) is shown in Fig. 2. By the gradual refinement of parameter values, the cost (Eq. 1) decreased by a factor of 2.0 from the first iteration (2.15) to the last iteration (1.08), approaching a value of 1.0 where on average the misfit equals the uncertainty in the data. In all, 2218 iterations (1-year simulations) were conducted before convergence was achieved i.e. until further improvement in cost was considered negligible. Note most of the progress was made in the first 120 iterations. The computational requirement of this Run was 100 minutes on a Dell Precision Workstation 650 with a 2.4 GHz processor.

4.2 Initial and final parameter values

Table 5 lists the parameter initial values, maximum and minimum limits, and final (optimized) values for Run 1. Note several parameter values are held constant or set to zero. This was determined as follows.

Experimentation showed that a lower Cost could be obtained by optimizing only some of the parameters rather than all of them. This is because some of the parameters are largely redundant over the primary data ranges, such that they cannot be determined simultaneously with confidence (Friedrichs et al., 2007). To determine which parameter to fix and which to optimize, a pragmatic approach was taken. In each Run in Table 4 a certain parameter was fixed to its initial value, its previous optimal value or a limit (e.g. zero), or optimized. If fixing a parameter to a constant value resulted in an equal or lower final cost, the parameter was subsequently kept fixed. If fixing a parameter resulted in a higher final cost, the parameter was allowed to vary. By cycling through the list of parameters repeatedly, and comparing the final cost of different optimization Runs, it was determined both which parameters were largely interdependent, and which parameterizations led to optimal agreement with the data (Tables 4 and 5). Values for parameters not determined by optimization are given in Table 5.

The optimized parameter value uncertainties in Table 5 were computed as the square root of the diagonal elements of the inverse of the Hessian matrix, where the Hessian is $d^2F/dp_i dp_j$ (Matear, 1995; Fennel et al., 2001), where p_i is an optimized parameter, $F = 0.5 * n * Cost^2$ and the number of observations n=531 for Run 1. The finite difference computation of $d^2F/dp_i dp_j$ used dp_i=0.05p_i, as smaller dp_i values did not ensure $F(p + dp_i) - 2F(p) + F(p - dp_i) > 0$ apparently due to round off error in the temporal integration or cost computation. The computed uncertainties in Table 5 range between 3-29%, suggesting all the optimized parameters are well-constrained. Following Fennel et al. (2001), eigenvalues and eigenvectors of the Hessian were

computed to diagnose parameter dependency. The condition number (ratio of largest to smallest eigenvalue) is 1.9e+7, suggesting significant dependency remains between some of the optimized parameters. The smallest eigenvalue is associated with nitrification parameter b_5, suggesting it is the least well-determined, probably through correlation with nitrification parameter b_3. The next two smallest are for detritus sinking parameters w_{d1} and w_{dn}, suggesting interdependency. Yet fixing any of these to constant values, even their Run 1 optimized value, results in higher cost (Runs 12, 49, 50, 19, 20, 63-66, 70).

Tables 4 and 5 indicate that quadratic P and Z loss terms give better agreement with the data than linear loss terms. This is probably because a quadratic allows higher loss at high concentrations and lower loss at low concentrations than a linear term. For example, Run 1 has lower variance in Chl than Run 5. For Z, the misfit is primarily due to underestimation (Sec 5.4), and the quadratic loss term allows a higher mean Z (Run 1 vs. Run 15). In contrast, linear A and D decay terms are preferred. Finally, concentration-dependent P and D sinking rates are preferred over constant sinking rates (Runs 19, 69 and 70). Another way to capture this would be constant sinking rates but two or more P and D classes.

4.3 Simulation Skill Assessment

Before discussing the meaning of the model results, let us assess the goodness-of-fit of the model to the observations. Table 6 shows various statistics for Run 1. Definitions are given in the Table 6 footnote; see Jolliff et al. (2009) and Stow et al. (2009) for more thorough descriptions. A variety of statistics are examined to see which statistics are most relevant to this application, given the large data errors.

The correlation coefficient, r, and the model standard deviation normalized by the data standard deviation, σ_m/σ_d, ideally would be 1.0. Note that they are "pattern" statistics which do not take into consideration data error. The RMSE, Bias, and unbiased RMSE ($\text{RMSE}^2 = \text{Bias}^2 + \text{uRMSE}^2$; Jolliff et al., 2009) are dimensional, with ideal values of 0 (Table 6). Here we use the sign convention for uRMSE of Jolliff et al. (2009), who further normalize these by σ_d, also given in Table 6 (nRMSE, nBias and nuRMSE).

A Taylor diagram (Taylor, 2001) is constructed from the values in Table 6 (Fig. 3a). It is a polar plot with σ_m/σ_d as distance from the origin, and $\text{acos}(r)$ as the angle from the x-axis; thus the ideal value is (x=1,y=0). Model results with radius greater (less) than the 1.0 circle have more (less) variance than the data; model results close to the x-axis (y-axis) have high (low) pattern correlation with the data. Fig. 3a illustrates that NO_3 and Chl have higher correlation coefficients and σ_m/σ_d close to 1.0. Thus pattern agreement between model and data NO_3 and Chl is good. The standard deviation of model A is 3 times that of the (monthly mean) data, while those of Z and D are only half. Model D and Z have very low correlation coefficients, followed by A. None of this however tells whether the model fits the data within its

prescribed uncertainty.

A Target diagram (Jolliff et al., 2009) is constructed (Fig. 3b) from the values of nBias and nuRMSE in Table 6. By the equation above, one can see that the radial distance from the origin is nRMSE such that the 1.0 circle indictates the RMSE equals σ_d. The Target diagram shows that A has the greatest nRMSE while Z and D have the greatest nBias. Again however these metrics do not tell whether the model fits the data within its uncertainty, which in this study are quite large, and for some data types larger than σ_d.

It is helpful to distinguish the relationship between the standard deviation of the monthly-binned data σ_d and the assigned data uncertainty $\hat{\sigma}_i$. The original, unbinned data contains both "signal" and "noise", and the binned data values are our best estimate of the signal we want and expect the model to fit. The σ_d in Table 6 is computed from the monthly-binned values, as these are the data assimilated; σ_d contains the variance of the large-scale spatial and temporal trends, i.e. the signal. The $\hat{\sigma}_i$ is the average standard deviation of the unbinned data within each time-depth bin (Sec. 2), i.e. the variability of unresolved submonthly and interannual processes. Thus σ_d is the standard deviation of the signal, while $\hat{\sigma}_i$ is the standard deviation of the noise, as defined by scale separation through binning. This separation allows us to keep the model from "fitting noise" (Lynch et al., 2006). In our application, the NH_4 data showed very large submonthly variability—either due to measurement error or natural variability—which the model was not expected or desired to fit. Monthly binning prevented fitting this noise. Yet the σ_d computed from the NH_4 monthly means is much smaller than our error estimates of those means, $\hat{\sigma}_i$ (Table 6). That is, the NH_4 σ_d is not practically significant. This suggested that normalizing the quantities in Table 6 by $\hat{\sigma}_i$ rather than σ_d might provide informative metrics. Note $\hat{\sigma}_i$ can also be thought of as approximately the standard deviation of the original data from the mean trend, i.e. the deviation of the original data when the large-scale trend is removed. Normalization in Table 6 by $\hat{\sigma}_i$ rather than σ_d therefore bears resemblance to maximum likelihood metrics.

Considering the case in which RMSE, Bias and uRMSE in Table 6 are normalized by $\hat{\sigma}_i$, values less than 1.0 mean that on average the model is within $\hat{\sigma}_i$ of the data, which is useful information. Comparisons between $\hat{\sigma}_i$ and RMSE can be easily made in Table 6, but difficulty arises when different data points have different $\hat{\sigma}_i$ (e.g. as with PP). The desired effect can be obtained by pre-normalizing each data point d_i and corresponding model value m_i by the corresponding $\hat{\sigma}_i$. An additional benefit of doing this is that different data types, now nondimensionalized, can be included in the same summation, to yield an overall model score for inter-simulation comparison.

Doing this normalization of d_i and m_i by $\hat{\sigma}_i$, and then running through the same equations in Table 6, yields Table 7 and Fig. 3c,d. The values for r and σ_m/σ_d in Table 7 using $d_i/\hat{\sigma}_i$ and $m_i/\hat{\sigma}_i$ are identical to r and σ_m/σ_d in Table 6, because $\hat{\sigma}_i$ drops out when constant; the exception is PP due to its variable $\hat{\sigma}_i$. So this normalization does not cause changes in the Taylor diagram

(when $\hat{\sigma}_i$ is constant) nor allowed it to take into account data uncertainty. The overall r and σ_m/σ_d are good, primarily because of the N and Chl data, which make up 65% of the data points (Fig. 3c). The modified Target diagram (Fig. 3d) however has changed significantly, both in the results and their interpretation. Bias and RMSE are now measured against data uncertainty, with the 1.0 circle indicating whether on average the model-data misfit is less than the uncertainty. The values of RMSE and Bias in Table 7 and Fig. 3d show that the model is on average within 1.34 times the uncertainty for all data types, with an overall RMSE of 1.08. Model A and D on average fit the data within their assigned errors. Modeled Z needs the most improvement; since the parameter values have been optimized, this can only be achieved by changes in model equations or state variables. Under this normalization minimizing Eq. 1 is identical to minimizing the RMSE. Given the Cost function, Fig. 3d also indicates that the Cost is currently putting less effort (per datum) into fitting D and A than the others.

In Table 6 r, σ_m/σ_d, nRMSE, nBias and nuRMSE largely tell the same story; that model N and Chl fit the data best, followed by PP (except for nBias). Yet Table 7 RMSE indicates that A and D are on average within the $\hat{\sigma}_i$ of the data, related to the fact that A and D have large relative uncertainties (Fig. 4; $\hat{\sigma}_i/\bar{m}$ in Table 6).

Other statistics that have been used as skill metrics include the Reliability Index (RI), the Modelling Efficiency (MEF) and binary discriminators such as the Receiver Operator Characteristic (Stow et al., 2009). However the RI does not take into account data errors, even when substituting $d_i/\hat{\sigma}_i$ and $m_i/\hat{\sigma}_i$ for d_i and m_i, and the MEF simply equals 1-nuRMSE^2. Binary discriminators were developed for situations where a binary decision has to be made, unlike this application, and are not easily reduced to a single number for model intercomparison. As such these are not treated here.

A skill metric should take three things into account: the model output, the data, and the uncertainty in the data. Model-data misfit is acceptable if it is within the observational error of the data. To be concerned with driving the model-data misfit well below the data uncertainty is considered "overfitting" or "fitting noise" (Lynch et al., 2006). None of the statistics in Table 6 include a variable for data uncertainty. By substituting $d_i/\hat{\sigma}_i$ and $m_i/\hat{\sigma}_i$ for d_i and m_i in the equations for RMSE and Bias, making them $\sqrt{\sum_{i=1}^{n}((m_i - d_i)/\hat{\sigma}_i)^2/n}$ and $\sum_{i=1}^{n}((m_i - d_i)/\hat{\sigma}_i)/n$, these terms quantify misfit relative to data error, and thus are useful skill metrics when data errors are significant, and as model-data misfits approach data errors.

5 Discussion

The central simulation (Run 1) results are shown in Figs. 4 through 7. Because the simulation is not a repeating annual cycle, the values on Jan 1 are not identical to those on Dec 31; nevertheless agreement is generally very

close.

A repeating annual cycle was not demanded, firstly, because Wilkinson Basin is not truly closed. Horizontal transport can be significant on annual timescales (Brown and Irish, 1993). Adjacent Massachusetts Bay annually imports DIN and exports organic nitrogen (fig. 6-2 in Hydroqual, 2000), such that Wilkinson Basin likely imports organic nitrogen which is converted to DIN. Thus nitrate and detritus are probably not in periodic steady states. Furthermore, the sinking PON flux removes nitrogen from top 200 m at a rate several times greater than the air-sea nitrogen input (Charette et al., 2001). Nitrate has a subsuface maximum around 200 m (Townsend, 1998) due to horizontal import of Slope Water; at this maximum, the vertical diffusive NO_3 flux must be zero. Thus the sinking PON flux at 200 m must either be balanced by a net upward advective flux of DIN at 200 m, horizontal import of DIN (between 50-200 m), or horizontal import of DON or PON (between 0-30 m). This issue of long-timescale total nitrogen conservation remains to be addressed. In effect, prescription of the observed initial conditions at the beginning of every iteration makes up for annual imbalances in the modeled nitrogen cycle.

Secondly, conducting parameter optimization simulations with a repeating annual cycle would mean that the final year (iteration) of each sensitivity run (e.g. Table 4) starts from a different January 1 initial condition (IC). This makes it unclear whether differences in Costs between runs are due to the difference a parameter has on the seasonal evolution or due to the difference in IC i.e. the decadal-timescale feedback of the parameter on the winter nutrient distribution. It may not make sense to compare parameter sensitivites starting from different IC, or to evaluate parameter sentivities starting from IC in poor agreement with the data due to long-term model drift. Therefore here every iteration (and run in Table 4) start from the same January 1st IC, which is based on the data; consequently we are only examining sensitivity to parameters on a seasonal timescale. Such limited-timescale parameter optimization is commonly used in operational (weekly to monthly) ocean forecasting.

Ideally horizontal transport of state variables into and out of Wilkinson Basin would be included as input to the 1-D model. The observations are too sparse in space and time to make accurate estimates of these time- and depth-dependent flux divergences, though such circulation-driven biogeochemical fluxes could be estimated from a 3-D physical-biological model. This suggests an iterative approach, with the 1-D model run to optimize the model parameters to data (since the 1-D model can be run thousands of times), a 3-D model run to compute horizontal fluxes, these fluxes used as input to improve of the 1-D model optimization, and so on. The horizontal flux estimates also would allow the 1-D model to be run longer than one year, to assess the impact of parameter values on longer-term budgets, which should further constrain the parameter values. The 1-D simulation presented here can be considered the first step in such a process.

The previous section on Skill examined model validation with regards to

the data that were assimilated. Below we present the model results and model validation based on comparison with information that was not assimilated. This includes the unsmoothed data (Fig. 1) where it probably gives more accurate estimates than the smoothed data at the bloom peaks. In some cases the model results bear more resemblence to the unsmoothed data than the assimilated smoothed data. It is always a challenge to find the right balance between too much smoothing (losing sharp peaks) and too little smoothing (allowing outliers). This intercomparison provides further insight into where model improvement is needed.

5.1 Nitrate.

Surface N agrees reasonably with observations (Figs. 4a, 5a). In November and December, simulated N has "steps" due to the vertical grid resolution; N increases rapidly when the deepening mixed layer entrains another level (Fig. 4a). During October most of the entrained N is consumed by P (Fig. 4b, 5b) i.e. the fall bloom.

In spring, the simulated N drawdown is twice as rapid as suggested by the data (Fig. 4a), namely the simulated drawdown occurs primarily in April, while the data suggest a decline over March and April. The simulated N drawdown is probably influenced by the assimilated Chl data, which constrains the bloom to begin in April. This reveals a discrepancy in the data; it is difficult to reconcile the observed March N drawdown without an increase in Chl (or Z or D). Satellite measurements (Thomas et al., 2003) however suggest surface Chl in Wilkinson Basin does increase from February to March, with typical March concentrations over 1 mg Chl m^{-3} .

Annual budgets from the simulation (Fig. 6) estimate that 90% of the NO_3 input into the top 24 m is supplied by vertical mixing (primarily in winter), 8% by (wet and dry) atmospheric deposition and 2% by *in situ* nitrification. Annual total nitrogen input into the top 24 m is 1302 mmol N m^{-2} yr^{-1}. Of this, 7% is due to the atmospheric N and A fluxes; interestingly 23% and 10% are due to upward mixing of A and D respectively, which have maxima just below the euphotic zone, with only 60% due to upward mixing of NO_3 itself.

5.2 Phytoplankton.

Simulated Chl also agrees well with the data (Figs. 4b-c, 5b-c) with a strong spring bloom in April, a summer deep Chl maximum (DCM), and a smaller fall bloom in October. Note satellite data support a November Chl peak (Thomas et al., 2003).

Chl starts increasing rapidly in the second half of March (Fig. 4b), due to decreasing light-limitation caused primarily by the rapidly shoaling MLD (Fig. 5b-c) though increasing PAR also contributes (Fig. 4f). During this rapid growth phase P losses are still significant, the net growth being due to a relatively small imbalance between sources and sinks (Fig. 7a); this is consistent

with the observation that during spring P increases only by 1 μM even though 7 μM of N is drawn down (Figs. 4a-b, 5b). During the bloom, P losses to grazing, sinking, losses to D ("aggregation") and vertical mixing (while the MLD is > 24 m) are all significant (Fig. 7b). The sharp decline in P occurs when N suddenly becomes depleted, allowing P growth to drop below the loss rates (Fig. 7a).

The simulation shows that temporally-decreasing light limitation in spring, coupled with the fact that P growth rates are faster than Z growth rates, allows growth to stay just ahead of grazing, until nutrients are depleted. The close balance between daily primary production and losses allows the spring bloom to be simulated in models without zooplankton (Sverdrup, 1953; Hydroqual, 1995; McGillicuddy et al., 2003).

In summer the simulation develops a DCM at a depth of 21 m, in agreement with observations (Fig. 5c). The simulation suggests the DCM is also a biomass maximum (Fig. 5b), in agreement with Holligan et al. (1984). The simulation matches observed Chl at 3 m well in summer, but overestimates Chl at 21 m in July and August, though it is unclear why observed Chl is lower in July-August than in June or September, given the similarity in physical conditions in all these months. Perhaps Z grazing is highest then, because Z biomass is maximum. The simulated DCM is at the 8% light level, though observations indicate it resides at the 1-5% light level (Holligan et al., 1984) and the 3% light level (O'Reilly et al., 1987). This suggests k_w or k_c may be too low, though higher values fit the data less well (Runs 51-54 in Table 4). A variable Chl:C ratio was not needed to simulate the DCM, as the Chl:C ratio does not change sharply enough with depth to account for the increase in Chl. The DCM is at the top of the nitracline, the primary cause being nutrient limitation of biomass.

In summer Z grazing dominates P loss (Fig. 7b); the Z excrete A which fuels recycled production. If it were true that in summer most phytoplankton production is exported by P sinking rather than Z grazing (Walsh et al., 1987), such an export of 12 mmol N m^{-2} d^{-1} from the euphotic zone would rapidly drive P and PP far below observed. The sum of the atmospheric N flux (0.24 mmol N m^{-2} d^{-1}), the vertical diffusive N flux (0.77 mmol N m^{-2} d^{-1} at 24 m), and a maximum estimate of the decline in the D+P+Z standing stocks (3 mmol N m^{-3} $\times$ 24 m $\div$ 100 days = 0.72 mmol N m^{-2} d^{-1}) are not enough to keep up with such an export. Thus while P sinking is significant in the annual average (Fig. 6; Walsh et al., 1987), it is not in summer.

5.3 Primary Production.

Simulated primary production at 3 m peaks in April and October (Fig. 4d), in agreement with the unsmoothed PP data (Fig. 1j). Annual mean primary production is 2.22 mol N m^{-2} yr^{-1} (Fig. 6), slightly lower than observation-based estimates (Table 8). Annual mean new production is 0.890 mol N m^{-2} yr^{-1}, within the range of observation-based estimates. The f-ratio is 0.6 in

winter and 0.2 in summer (Fig. 7d).

Simulated primary production (Fig. 7c) has a subsurface maximum at the DCM during the stratified season (May-Sep). The PP data indicate this is incorrect (Fig. 1, where the max is at 9 m from Feb-Sep), though it is sometimes observed (figs. 7 and 10 in Holligan et al., 1984; fig. 21.10 in O'Reilly et al., 1987; fig. 8 in Charette et al., 2001). More precisely, model PP at 21 m is in good agreement with observations (Fig. 4e), but model PP at 3 m is too high in spring and too low in summer (Fig. 4d, 5e), as often occurs in simple ecosystem models (Fasham et al., 1993; Lefevre et al., 2003). Observed PP at 3 m (Fig. 4d) shows surprisingly little seasonality, given the large seasonality in Chl, N and PAR (Figs. 4a,b,f). A contributing factor may be the use of smoothed PP data, as the binned PP data (Fig. 1j) suggests similar PP in April and October at 3 m (0.732 and 0.814 mmol N m^{-3} d^{-1}), twice that of August (0.349 mmol N m^{-3} d^{-1}). In addition, the model may need to include DON seasonality or more complex treatment of C:Chl:N ratios (Lefevre et al., 2003). Silicate limitation may also affect the evolution of the spring bloom (Townsend and Thomas, 2002).

The mean depth of the euphotic zone (defined as the zero line between net biological source and sink, i.e. excluding sinking and mixing) is 25 m for P and 28 m for N in summer, just below the DCM (21 m) and summer nitracline (24 m). Thus for estimates of fluxes into the euphotic zone, we use 24 m (the bottom of level 4).

In the top 24 m, 51% of P losses are due to Z grazing, 10% to sinking, 11% to loss to D (i.e. senescence to phytodetritus; Turner, 2002; Cuny et al., 2002), and 28% to vertical mixing in winter (Fig. 6). This generally agrees with estimates that a significant portion of PP is lost as phytodetritus (Walsh et al., 1987). However Dagg and Turner (1982) estimate that 50% of primary production is grazed by mesozooplankton alone. Similarly, Cohen and Grosslein (1987) estimate macrozooplankton production as 8.07% of annual PP and microzooplankton production as 14.30% of PP; assuming assimilation efficiencies of 20% for macrozooplankton (Anderson and Hessen, 1995) and 33% for microzooplankton (Peligri et al., 1999), this suggests macrozooplankton graze 40% of PP and microzooplankton graze 43%. Although not included in our model, P excretion of DOM is estimated at 16% of PP (Walsh et al., 1987).

5.4 Zooplankton.

Simulated Z peaks in spring rather than mid-summer (Fig. 4h). At 3 m Z closely follows P, with a spring and fall maxima (Fig. 4g). Thus Z is behaving more like microzooplankton than macrozooplankton, due to the high grazing rate g and assimilation efficiency f_a (Table 5). This shows an inherent difficulty in using one Z state variable; copepods dominate biomass but microheterotrophs dominate grazing. In summer, model Z biomass is highest at the DCM (which has been observed at times; Townsend et al., 1984; Malkiel et al.,

2006), and there are significant concentrations below 30 m feeding primarily on D (Fig. 5f).

Vertically-integrated Z biomass was assimilated because the observed vertical distribution of Z biomass is not well known, largely because of diel vertical migration, which is sometimes observed for some species though not always (Durbin et al., 1995). Because of this irregularity, vertical Z migration was not included in the standard model.

In the euphotic zone, Z graze primarily on P, though grazing on Z is significant (Fig. 6). In the aphotic zone, grazing on P, D and Z is more closely balanced. Of Z losses, 22% goes to D and the rest to A. In Run 1, Z are net producers of D in both the euphotic zone and aphotic zone.

Secondary production in the model, viz. $f_a(G_p + G_d + G_z)Z$, is 53 g C m^{-2} yr^{-1}, similar to the Cohen and Grosslein (1987) estimate of 57 g C m^{-2} yr^{-1} (assuming 1 g C = 10 kcal for Z). Note however Cohen and Grosslein's estimate is based in part on the assumption that Z only graze P, and thus is likely an underestimate; the model eqivalent of $f_a G_p Z$ is 43 g C m^{-2} yr^{-1}.

5.5 Detritus.

Simulated D predicts a peak at the time of the spring bloom; unfortunately there is a lack of D data between April 12 and June 9 for comparison. As a large amount of P are estimated to sink out of the euphotic zone rather than be grazed by mesozooplankton (Walsh et al., 1987), and only a fraction of what Z graze would go into D, an April peak seems likely. Charette et al. (2001, fig. 13) show POC export at 50 m to be higher in March than in June in both Wilkinson and Jordan Basins.

The data suggest the D peak occurs simultaneously at 27 and 99 m depth (Fig. 4i,j), suggesting a fast sinking rate. Yet D concentrations never get below 1 μM at 27 or 99 m, suggesting a slowly-sinking, refractory component. The model is roughly able to fit the data by means of a concentration-dependent sinking rate and a slow remineralization rate (Table 5). With constant sinking rates, probably two D classes would be needed.

The MWRA data show uniformity in D with depth, rather than a factor of 2 decline over 100 m (Martin et al., 1987). Since this data is from the edge of Massachusetts Bay (<110 m water depth) it may be an overestimate for Wilkinson Basin. Charette et al. (2001) estimate POM concentrations for Wilkinson Basin generally a factor of 2 lower than the MWRA-based estimates, though this is generally within the estimated D error bounds (Fig. 4i,j). Charette et al. (2001) also found general uniformity in D with depth and season, with deep (> 70 m) POC of 3-5 μM and POC fluxes of 14-18 mmol m^{-2} d^{-1} which indicate sinking rates of 2-4 m/d. The model yields similar sinking rates (0.5-4 m/d), with annual sinking PON (=P+D) fluxes of 1.91 and 0.92 mmol N m^{-2} d^{-1} at 24 and 198 m respectively.

5.6 Ammonium.

Simulated ammonium at 9 m (Fig. 4k) is generally lower than observed, though within the uncertainty based on data variability. Some of the data variability may be due to measurement imprecision (Brzezinski, 1988), but it is also possible that natural variability is high, due to its production by Z (which are spatially patchy), rapid consumption by P and bacteria, and sporadic sources from the atmosphere and the coast. At 45 m (Fig. 4l), simulated A has an Apr-May peak much larger than observed, associated with the grazing of sinking P by Z (Figs. 5, 6).

The simulated A maximum is just below the DCM, similar to typically observed (Holligan et al., 1984). The NODC A data (Fig. 1d,e) show an A maximum between 50 and 100 m, deeper than the P and D maxima; the fact that it occurs in almost every month in Fig. 1d suggests it is not an artefact of undersampling. Perhaps it is the result of vertical Z migration, which is not included in the model, or horizontal advection of seafloor-regenerated A from shallow areas. Deep A maxima are generally not observed in the ocean, due to nitrification, suggesting horizontal advection; in any case, if such a deep A maxima does exist in Wilkinson Basin, it is below the depth for utilization by P, such that its presence (and absence in the simulation) does not impact P production or biomass.

Run 1 estimates 0.03-0.06 μM NO_3 and 0.002-0.003 μM NH_4 at the sea surface in summer, lower than observed (0.08-0.17 μM NO_3 and 0.1-0.2 μM NH_4; Fig. 1) though within the assigned data uncertainty. The high observed values remain unexplained by the model, and are probably due to unresolved processes (e.g. submesoscale upwelling, vertical Z migration) assuming the data are not contaminated.

5.7 Comparison with other regions

The data show that the seasonal cycle of phytoplankton in Wilkinson Basin is similar to the open ocean at this latitude of 42°N (Strass and Woods, 1991; Ducklow and Harris, 1993; Marra and Ho, 1993; Harrison et al., 1993). Chlorophyll and primary production peak in spring (April) and fall (October). Strong stratification occurs in summer, and the model suggests primary production in the mixed layer then is essentially nutrient-limited (rather than light-limited), despite detectable nutrients. There are differences with the open ocean however. CTD profiles in summer often show extremely thin mixed layers (e.g. 1 m), presumably due to strong light absorption, which in turn are due to high Chl and organic matter concentrations (Sosik et al., 2001). The deep chlorophyll maximum is at about 21 m depth, and the euphotic zone depth is estimated at 25 m. Detritus, NH_4 and Chl concentrations are higher, and atmospheric nitrogen inputs are non-negligible. These higher concentrations are ultimately related to horizontal transport from nearby shallow areas.

An interesting comparison can be made between our annual budgets (Fig. 6)

and those for Georges Bank just to the east by Steele et al. (2007, their fig. 4). Our annual primary and new production estimates (Fig. 6) are 49% and 56% theirs, though annual f-ratios are similar (0.29 vs. their 0.26). According to our model, 82% and 18% of primary production is lost to zooplankton and detritus, respectively; their analysis indicates 92% and 8% for Georges Bank. Wilkinson Basin shows spring and fall chlorophyll blooms, while ocean color data generally only shows a spring bloom on Georges Bank (Thomas et al., 2003). Field measurements of Chl on the shoals of Georges Bank do show enhanced Chl in fall, although this may be due to horizontal advection (O'Reilly and Zetlin, 1998). In Wilkinson Basin, nitrate is replenished annually by vertical mixing, while Georges Bank relies on horizontal advection. Hence the necessity of taking into account horizontal fluxes in an annual budget of Georges Bank.

Massachusetts Bay (< 110 m deep), to the west of Wilkinson Basin, also generally exhibits spring and fall blooms, though interannual variability is significant (Libby et al., 2006). For example the spring bloom is very weak in some years, and the fall bloom has been observed to occur in September, October or November (fig. 4-32 in Libby et al., 2001), which may be expected for Wilkinson Basin. The spring bloom occurs on average earlier in Massachusetts Bay than in Wilkinson Basin, and even earlier in Cape Cod Bay, as shallow areas warm and stratify earlier, due to their bottom-limited winter MLD. Observations of the summer DCM and nitracline depth in Massachusetts Bay at a given station in a given year are rather erratic (Figs. 3.5-3.11 in Jiang and Zhou, 2006), attributed to mesoscale eddies and filaments. Massachusetts Bay imports 33 kilotons of DIN per year across its open boundary and exports an equal amount of organic nitrogen (fig. 6-2 in Hydroqual, 2000). Its coastal and atmospheric nitrogen loading of 26 kilotons per year is balanced by local burial and sedimentary denitrification. Note these numbers are about 10% of its annual primary production of 350-500 g C m^{-2} yr^{-1} (fig. 5-25 in Libby et al., 2001, 2006; Kelly and Doering, 1997), which is 2-3 times that of Wilkinson Basin, and similar to Georges Bank.

5.8 Areas for improvement

Model PP and Z are most outside the prescribed uncertainties in the data (Fig. 4). As these misfits could not be fixed through parameter optimization, changes to the model equations or state variables are required. They perhaps can be solved by adding a second P class (small phytoplankton, to maintain a nominal PP—also supported by the preference for a concentration-dependent P sinking rate) and a second Z class (to differentiate microzooplankton, which dominate grazing, from macrozooplankton, which dominate biomass.) Macrozooplankton have lower growth rates, assimilation efficiencies and mortality rates which should help obtain the June Z maximum that lags the P maximum by two months. The concentration-dependent D sinking rate also suggests a second D class.

The optimization scheme assimilates the (time-averaged) data at single points in time, which is not a direct comparison. Comparing time-averaged model output with the data would avoid phase (timing) errors and allow the inclusion of the daily light cycle and diel zooplankton migration.

This study also has identified key gaps in the extant observational database needed to constrain the nitrogen budget in Wilkinson Basin. In particular, microzooplankton biomass, bacterial biomass, particulate and dissolved organic matter, and particle flux observations are extremely sparse. In our combined dataset, the time period Sep-Dec has few observations, even of nutrients. The A, D and Z data estimates have large uncertainties; more high quality data is needed to better constrain these estimates and thus the annual cycle.

6 Conclusions

A 6-box ecosystem model was fit to data through parameter optimization in order to provide a dynamically-consistent best estimate of the seasonal and annual nitrogen budgets in Wilkinson Basin. A summary of the observations are given in Fig. 1, and best estimates of the annual nitrogen cycle budget are given Fig. 6. The model estimates annual primary production as 176 g C m^{-2} yr^{-1}, annual new production as 71 g C m^{-2} yr^{-1} and sinking PON fluxes of 9.7 and 4.7 g N m^{-2} yr^{-1} at 24 and 198 m respectively. The model does not agree with the data in all instances, and does not include horizontal transports; as such this is merely a first estimate, and not the final word.

Model optimization was based on weighted least squares, with model-data misfits normalized by data uncertainty. The original data were reduced to monthly means, in order to separate the large-scale temporal and spatial "signal" from the submonthly and interannual "noise" that could not be reproduced by the model. The data uncertainties were computed as the standard deviations of the original data in monthly and 6-m bins (i.e. the noise), making the method similar to maximum likelihood estimation. These estimated uncertainties ($\hat{\sigma}_i$) are generally large relative to the seasonal cycle (σ_d), such that the assimilated data have a low "signal-to-noise" ratio. On average, the model fits the data at 1.08 times the estimated uncertainties.

A variety of statistics were examined. Pattern statistics (such as the Taylor diagram) do not explicitly take into account the data uncertainties, and therefore are difficult to interpret when the signal-to-noise ratio is low. Using uncertainty-normalized data and model estimates $d_i/\hat{\sigma}_i$ and $m_i/\hat{\sigma}_i$ in the equations for RMSE and Bias tell whether the model is fitting the data within its assigned error, and thus were found to be the most appropriate metrics in this low signal-to-noise application.

While the 6-box biological model is successful in reproducing most of the data within its assigned uncertainty, it has difficulty reproducing the seasonal trends in observed primary production and zooplankton biomass. The latter is probably because macrozooplankton dominate biomass while microzooplank-

ton dominate grazing, and the two have very different growth rates, assimilation efficiencies and mortality rates. This suggests the addition of a second phytoplankton class and second zooplankton class may be needed to better reproduce the primary production and zooplankton biomass seasonality.

This study illustrates how phytoplankton abundance and primary productivity interact with NH_4, sinking detritus, DOM and zooplankton in a web of bottom-up and top-down controls. More high-quality observations of all these fields are needed to better constrain our understanding of the mechanics of these interactions.

The optimized model estimates the typical depth-dependent seasonal cycles of nitrogen and phytoplankton in Wilkinson Basin (Figs. 5 and 6). As such the calibrated biological model can be used as a starting point in 3-D simulations examining nitrogen budgets and ecosystem variability. The model could be readily applied to the carbon cycle (assuming Redfield C:N ratios), though for e.g. air-sea CO_2 flux estimates more model components are needed. More extensize model development would also be needed for application to fisheries (viz. explicitly resolve important Z prey species) and wastewater discharge eutrophication (viz. add sedimentary nitrogen-cycle processes).

Acknowledgements

This work was supported by ONR, NSF, and NOAA grant to Dennis McGillicuddy. I would like to thank Dennis McGillicuddy, Geoff Evans for the model code, and Chris Garside, Wendy Leo and Olga Kosnyrev for help in collecting the data sets.

References

Allen, J.I., Holt, J.T., Blackford, J., Proctor, R., 2007. Error quantification of a high-resolution coupled hydrodynamic-ecosystem coastal-ocean model: Part 2. Chlorophyll-a, nutrients and SPM. J. Mar. Syst. 68, 381-404.

Anderson, D.M., 1997. Bloom dynamics of toxic Alexandrium species in the northeastern U.S., Limnol. Oceanogr. 42, 1009-1022.

Anderson, D.M., Glibert, P.M., Burkholder, J.M., 2002. Harmful algal blooms and eutrophication: nutrient sources, composition, and consequences, Estuaries 25(4b), 704-726.

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 Res. I 47, 1787-1827.

Anderson, T.R., Hessen, D.O., 1995. Carbon or nitrogen limitation in marine copepods? J. Plankton Res. 17, 317-331.

Armstrong, R.A., Sarmiento, J.L., Slater, R.D., 1995. Monitoring ocean productivity by assimilating satellite chlorophyll into ecosystem models. In:

975 Ecological Time Series, T.M. Powell and J.H. Steele, eds. Chapmal & Hall,
976 p 371-390.
977 Benitez-Nelson, C.R., Buesseler, K.O., Crossin, G., 2000. Upper ocean carbon
978 export, horizontal transport, and vertical eddy diffusivity in the southwest-
979 ern Gulf of Maine. Cont. Shelf Res. 20, 707-736.
980 Bennet, A.F., 2002. Inverse Modeling of the Ocean and Atmosphere. Cam-
981 bridge University Press, Cambridge, 234 pp.
982 Besiktepe, S.T., Lermusiaux, P.F.J., Robinson, A.R., 2003. Coupled physical
983 and biogeochemical data-driven simulations of Massachusetts Bay in late
984 summer: real-time and postcruise data assimilation. J. Mar. Syst. 40-41,
985 171-212.
986 Bisagni, J.J., 2003. Seasonal variability of nitrate supply and potential new
987 production in the Gulf of Maine and Georges Bank regions. J. Geo-
988 phys. Res. 108 (C11), 8015, doi:10.1029/2001JC001136.
989 Bolstad, W.M., 2004. Introduction to Bayesian Statistics. Wiley-IEEE, 376
990 pp.
991 Brown, W.S., Irish, J.D, 1993. The annual variation of water mass structure
992 in the Gulf of Maine: 1986-1987. J. Mar. Res. 51, 53-107.
993 Brzezinski, M.A., 1988. Vertical distribution of ammonium in stratified olig-
994 otrophic waters. Limnol. Oceanogr. 33, 1176-1182.
995 Campbell, D.E., 1986. Process variability in the Gulf of Maine—A macroes-
996 tuarine environment, in: Estuarine Variability, D.A. Wolfe (ed.), p 261-275,
997 Academic Press, Orlando, FL.
998 Campbell, D.E., 2004. Evaluation and energy analysis of the Cobscook Bay
999 ecosystem. Northeastern Naturalist 11 (SP2), 355-424.
1000 Charette, M.A., Moran, S.B., Pike, S.M., 2001. Investigating the carbon cy-
1001 cle in the Gulf of Maine using the natural tracer thorium 234. J. Geo-
1002 phys. Res. 106, 11553-11579.
1003 Christian, J.R., Verschell, M.A., Murtugudde, R., Busalacchi, A.J., McClain,
1004 C.R., 2002. Biogeochemical modelling of the tropical Pacific Ocean. I: Sea-
1005 sonal and interannual variability. Deep-Sea Res. II 49, 509-543.
1006 Cohen, E.B., Grosslein, M.D., 1987. Production on Georges Bank compared
1007 with other shelf ecosystems, in: Georges Bank, R. H. Backus (ed.), p 383-
1008 391, MIT Press, Cambridge, Mass.
1009 Cuny, P., Marty, J.-C., Chiaverini, J., Vescovali, I., Raphel, D., Rontani, J.-F.,
1010 2002. One-year seasonal survey of the chlorophyll photodegradation process
1011 in the northwestern Mediterranean Sea. Deep-Sea Res. II 49, 1987-2005.
1012 Dagg, M.J., Turner, J.T., 1982. The impact of copepod grazing on the phy-
1013 toplankton of Georges Bank and the New York Bight. Can. J. Fish. Aquat.
1014 Sci. 39, 979-990.
1015 Davis, C.S., 1987. Zooplankton life cycles, in: Georges Bank, R. H. Backus
1016 (ed.), p 256-267, MIT Press, Cambridge, Mass.
1017 Ducklow, H.W., Harris, R.P., 1993. Introduction to the JGOFS North Atlantic
1018 Bloom Experiment. Deep-Sea Res. II 40, 1-8.
1019 Durbin. E.G., Gilman, S.L., Campbell, R.G., Durbin, A.G., 1995. Abundance,

biomass, vertical migration and estimated development rate of the copepod Calanus finmarchicus in the southern Gulf of Maine during late spring. Cont. Shelf Res. 15, 571-591.

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

Evans, G.T., 2003. Defining misfit between biogeochemical models and data sets. J. Mar. Syst. 40-41, 49-54.

Evensen, G., 2007. Data Assimilation. Springer, Berlin. 279 pp.

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, Phil. Trans. R. Soc. Lond. B 348, 203-209.

Fasham, M.J.R., Sarmiento, J.L., Slater, R.D., Ducklow, H.W., Williams, R., 1993. Ecosytem behaviour at Bermuda Station "S" and Ocean Weather Station "India": a general circulation model and observational analysis. Global Biogeochemical Cycles 7, 379-415.

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

Franks, P.J.S., Chen, C., 1996. Plankton production in tidal fronts: A model of Georges Bank in summer. Journal of Marine Research 54, 631-651.

Franks, P.J.S., Chen, C., 2001. A 3-D prognostic numerical study of the Georges bank ecosystem. Part II: biological-physical model. Deep-Sea Res. II 48, 457-482.

Friedrichs, M.A.M., Dusenberry, J.A., Anderson, L.A., Armstrong, R.A., Chai, F., Christian, J.R., Doney, S.C., Dunne, J., Fujii, M., Hood, R., McGillicuddy, D.J., Moore, J.K., Schartau, M., Spitz, Y.H., Wiggert, J.D., 2007. Assessment of skill and portability in regional marine biogeochemical models: Role of multiple planktonic groups. J. Geophys. Res. 112, C08001, doi:10.1029/2006JC003852.

Garside, C., Garside, J.C., Keller, M.D., Sieracki, M.E., 1996. The formation of high-nutrient-low salinity water in the Gulf of Maine: A nutrient trap? Estuarine, Coastal and Shelf Science 42, 617-628.

Geider, R.J., 1987. Light and temperature dependence of the carbon to chlorophyll a ratio in microalgae and cyanobacteria: Implications for physiology and growth of phytoplankton. New Phytol. 106, 1-34.

Geider, R.J., MacIntyre, H.L., Kana, T.M., 1996. A dynamic model of photoadaptation in phytoplankton. Limnol. Oceanogr. 41, 1-15.

Geider, R.J., MacIntyre, H.L., Kana, T.M., 1997. Dynamic model of phytoplankton growth and acclimation: responses of the balanced growth rate and the chlorophyll a:carbon ratio to light, nutrient-limitation and temperature. Mar. Ecol. Prog. Ser. 148, 187-200.

Gelman, A., 2004. Bayesian Data Analysis. CRC Press, 668 pp.

Graziano, L.M., Balch, W.M., Drapeau, D., Bowler, B.C., Vaillancourt, R., Dunford, S., 2000. Organic and inorganic carbon production in the Gulf of Maine. Cont. Shelf Res. 20, 685-705.

Harrison, W.G., Head, E.J.H., Horne, E.P.W., Irwin, B., Li, W.K.W., Longhurst, A.R., Paranjape, M.A., Platt, T., 1993. The western North Atlantic Bloom Experiment. Deep-Sea Res. II 40, 279-305.

Holligan, P.M., Balch, W.M., Yentsch, C.M., 1984. The significance of subsurface chlorophyll, nitrite and ammonium maxima in relation to nitrogen for phytoplankton growth in stratified waters of the Gulf of Maine, Journal of Marine Research 42, 1051-1073.

Hurtt, G.C., Armstrong, R.A., 1996. A pelagic ecosystem model calibrated with BATS data. Deep-Sea Res. II 43, 653-683.

HydroQual, 1995. A water quality model for Massachusetts and Cape Cod Bays: Calibration of the Bays Eutrophication Model (BEM), EnQuaD Tech. Report 95-8, Massachusetts Water Resources Authority, Boston, MA, 402 pp.

HydroQual, 2000. Bays Eutrophication Model (BEM): modeling analysis for the period 1992-1994, Boston: Massachusetts Water Resources Authority, Report ENQUAD 2000-02, 158 pp.

Hydroqual, 2003. Bays Eutrophication Model (BEM): Model verification for the period 1998-1999. Boston: Massachusetts Water Resources Authority, Report ENQUAD 2003-03, 318 pp.

Ivanov, S., Palamarchuk, J., 2007. Fine-scale precipitation structure of a cold front and the problem of the representativeness error. Adv. Geosci. 10, 3-8.

Ji, R., Chen, C., Franks, P.J.S., Townsend, D.W., Durbin, E.G., Beardsley, R.C., Lough, R.G., Houghton, R.W., 2006a. Spring phytoplankton bloom and associated lower trophic level food web dynamics on Georges Bank: 1-D and 2-D model studies. Deep-Sea Res. II 53, 2656-2683.

Ji, R., Chen, C., Franks, P.J.S., Townsend, D.W., Durbin, E.G., Beardsley, R.C., Lough, R.G., Houghton, R.W., 2006b. The impact of Scotian Shelf water "cross-over" in the plankton dynamics on Georges Bank: A 3-D experiment for the 1999 spring bloom. Deep-Sea Res. II 53, 2684-2707.

Ji, R., Davis, C., Chen, C., Beardsley, R., 2008. Influence of local and external processes on the annual nitrogen cycle and primary productivity on Georges Bank: A 3-D biological-physical modeling study. Journal of Marine Systems 73, 31-47.

Jiang, M., Zhou, M., 2003. Massachusetts Bay Hydrodynamic Model and Water Quality Model results in 1998-1999: Comparison report between HydroQual and University of Massachusetts Boston runs. Boston: Massachusetts Water Resources Authority, Report ENQUAD 2003-10, 42 pp.

Jiang, M., Zhou, M., 2004. Bays Eutrophication Model (BEM) model verification for the period 2000-2001. Boston: Massachusetts Water Resources Authority, Report ENQUAD 2004-09, 90 pp.

Jiang, M., Zhou, M., 2006. Massachusetts Bay Eutrophication Model: 2002-2004 simulation. Boston: Massachusetts Water Resources Authority, Report ENQUAD 2006-13, 126 pp.

Jiang, M., Zhou, M., 2007. User's guide to the water-quality part if the Bays Eutrophication Model (BEM). Boston: MWRA Report 2007-09, 36 pp.

Jolliff, J.K., Kindle, J.C., Shulman, I., Penta, B., Friedrichs, M.A.M., Helber, R., Arnone, R.A., 2009. Summary diagrams for coupled hydrodynamic-ecosystem model skill assessment. J. Mar. Syst. 76, 64-82.

Jordan, C.E., Talbot, R.W., 2000. Direct atmospheric deposition of water-soluble nitrogen to the Gulf of Maine. Global Biogeochem. Cycles 14, 1315-1329.

Keller, A.A., Taylor, C., Oviatt, C., Dorrington, T., Holcombe, G., Reed, L., 2001. Phytoplankton production patters on Massachusetts Bay and the absence of the 1998 winter-spring bloom. Marine Biology 138, 1051-1062.

Kelly, J.R., Doering, P.H., 1997. Monitoring and modeling primary production in coastal waters: studies in Massachusetts Bay 1992-1994. Mar. Ecol. Prog. Ser. 148, 155-168.

Klein, P., 1987. A simulation of some physical and biological interactions. In: Georges Bank, R. H. Backus (ed.), p 395-402, MIT Press, Cambridge, Mass.

Lee, P.M., 2004. Bayesian Statistics: An Introduction. Arnold Press, 368 pp.

Lefevre, N., Taylor, A.H., Gilbert, F.J., Geider, R.J., 2003. Modeling carbon to nitrogen and carbon to chlorophyll a ratios in the ocean at low latitudes: Evaluation of the role of physiological plasticity. Limnol. Oceanogr. 48, 1796-1807.

Levitus, S., 1982. Climatological Atlas of the World Ocean, NOAA Professional Paper 13, U.S. Government Printing Office, Washington, DC, 190 pp.

Lewis, C.V.W., Davis, C.S., Gawarkiewicz, G., 1994. Wind forced biological-physical interactions on an isolated offshore bank, Deep-Sea Res. II 41, 51-73.

Libby, P.S., Hunt, C.D., McLeod, L.A., Geyer, W.R., Keller, A.A., Oviatt, C.A., Borkman, D., Turner, J.T., 2001. 2000 Annual Water Column Monitoring Report, Boston: Massachusetts Water Resources Authority. Report ENQUAD 2001-17, 196 p.

Libby, P.S., Geyer, W.R., Keller, A.A., Mansfield, A.D., Turner, J.T., Anderson, D.M., Borkman, D., Rust, S.W., Hyde, K., Oviatt, C.A., 2006. 2005 Annual water column monitoring report. Boston: Massachusetts Water Resources Authority. Report ENQUAD 2006-20, 182 p.

Loder, J.W., Wright, D.G., Garrett, C., Juszko, B.-A., 1982. Horizontal exchange on central Georges Bank. Can. J. Fish. Aquat. Sci. 39, 1130-1137.

Love, R.C., Loder, T.C., Keafer, B.A., 2005. Nutrient conditions during Alexandrium fundyense blooms in the western Gulf of Maine, USA, Deep-Sea Res. II 52, 2450-2466.

Lynch, D.R., McGillicuddy, D.J., Werner, F.E., 2006. Skill vocabulary - a starting point. Unpublished manuscript, available at http://www-nml.thayer.dartmouth.edu/Publications/internal_reports/NML-06-Skill/vocab1.pdf.

Malkiel, E., Abras, J.N., Widder, E.A., Katz, J., 2006. On the spatial distribution and nearest neighbor distance between particles in the water column determined from *in situ* holographic measurements. J. Plankton Res. 28,

149-170.
Marra, J., Ho, C., 1993. Initiation of the spring bloom in the northeast Atlantic (47°N, 20°W): a numerical simulation. Deep-Sea Res. II 40, 55-73.
Martin, J.H., Knauer, G.A., Karl, D.M., Broenkow, W.W., 1987. VERTEX: carbon cycling in the northeast Pacific. Deep-Sea Res. 34, 267-285.
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.
McCreary, J.P., Jr., Kohler, K.E., Hood, R.R., Olson, D.B., 1996. A four-component ecosystem model of biological activity in the Arabian Sea. Prog. Oceanog. 37, 193-240.
McGillicuddy, D.J., Jr., Anderson, L.A., Doney, S.C., Maltrud, M.E., 2003. Eddy-driven sources and sinks of nutrients in the upper ocean: Results from a 0.1 resolution model of the North Atlantic. Global Biogeochem. Cycles, 17(2), 1035, doi:10.1029/2002GB001987.
Meise, C.J., O'Reilly, J.E., 1996. Spatial and seasonal patterns in abundance and age-composition of Calanus finmarchicus in the Gulf of Maine and on Georges Bank: 1977-1987. Deep-Sea Res. II 43, 1473-1501.
Montagnes, D.J.S., Lynn, D.H., Roff, J.C., Taylor, W.D., 1988. The annual cycle of heterotrophic planktonic ciliates in the waters surrounding the Isles of Shoals, Gulf of Maine: an assessment of their trophic role. Mar. Biol. 99, 21-30.
Ondercin, D.G., Atkinson, C.A., Kiefer, D.A., 1995. The distribution of bioluminesence and chlorophyll during the late summer in the North Atlantic: Maps and a predictive model. J. Geophys. Res. 100, 6575-6590.
O'Reilly, J.E., Evans-Zetlin, C., Busch, D.A., 1987. Primary production, in: Georges Bank, R. H. Backus (ed.), p 220-233, MIT Press, Cambridge, Mass.
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 Tech. Report NMFS 139, Seattle, WA, 120 p.
Parsons, T.R., Takahashi, M., Hargrave, B., 1984. *Biological Oceanographic Processes*, Pergamon Press, Elmsford, New York, 330 pp.
Pelegri, S.P., Christaki, U., Dolan, J., Rassoulzadegan, F., 1999. Particulate and dissolved organic carbon production by the heterotrophic nanoflagellate Pteridomonas danica Patterson and Fenchel, Microb. Ecol. 37, 276-284.
Press, W.H., Flannery, B.P., Teukolsky, S.A., Vetterling, W.T., 1986. *Numerical Recipes*. Cambridge University Press, Cambridge, UK. 818 pp.
Schlitz, R.J., Cohen, E.B., 1984. A nitrogen budget for the Gulf of Maine and Georges Bank. Biological Oceanography 3(2), 203-222.
Sherman, K., Smith, W.G., Green, J.R., Cohen, E.B., Berman, M.S., Marti, K.A., Goulet, J.R., 1987. Zooplankton production and the fisheries of the Northeastern Shelf, in: Georges Bank, R. H. Backus (ed.), p 268-282, MIT Press, Cambridge, Mass.
Sosik, H.M., Green, R.E., Pegau, W.S., Roesler, C.S., 2001. Temporal and vertical variability in optical properties of New England shelf waters during

late summer and spring. J. Geophys. Res. 106, 9455-9472.
Steele, J., 2009. Assessment of some linear food web methods. J. Mar. Syst. 76, 186-194.
Steele, J.H., Henderson, E.W., 1992. A simple model for plankton patchiness. Journal of Plankton Research 14, 1397-1403.
Steele, J.H., Collie, J.S., Bisagni, J.J., Gifford, D.J., Fogarty, M.J., Link, J.S., Sullivan, B.K., Sieracki, M.E., Beet, A.R., Mountain, D.G., Durbin, E.G., Palka, D., Stockhausen, W.T., 2007. Balancing end-to-end budgets of the Georges Bank ecosystem. Progress in Oceanography 74, 423-448.
Stow, C.A., Jolliff, J., McGillicuddy, D.J., Doney, S.C., Allen, J.I., Friedrichs, M.A.M., Rose, K.A., Wallhead, P., 2009. Skill assessment for coupled biological/physical models of marine systems. J. Mar. Syst. 76, 4-15.
Strass, V.H., Woods, J.D., 1991. New production in the summer revealed by the meridional slope of the deep chlorophyll maximum. Deep=Sea Res. A 38, 35-56.
Sverdrup, H.U., 1953. On conditions for the vernal blooming of phytoplankton. J. Cons. Explor. Mer. 18, 287-295.
Tarantola, A., 1987. Inverse Problem Theory: Methods for Data Fitting and Model Parameter Estimation. Elsevier, New York.
Taylor, K.E., 2001. Summarizing multiple aspects of model performance in a single diagram. J. Geophys. Res. 106 (D7), 7183-7192.
Thomas, A.C., Townsend, D.W., Weatherbee, R., 2003. Satellite-measured phytoplankton variability in the Gulf of Maine. Cont. Shelf. Res. 23, 971-989.
Tian, R., Chen, C., 2006. Influence of model geometrical fitting and turbulence parameterization on phytoplankton simulation in the Gulf of Maine. Deep-Sea Res. II 53, 2808-2832.
Townsend, D.W., 1991. Influences of oceanographic processes on the biological productivity of the Gulf of Maine. Reviews in Aquatic Sciences 5, 211-230.
Townsend, D.W., 1998. Sources and cycling of nitrogen in the Gulf of Maine. Journal of Marine Systems 16, 283-295.
Townsend, D.W., Cammen, L.M., Holligan, P.M., Campbell, D.E., Pettigrew, N.R., 1994. Causes and consequences of variability in the timing of spring phytoplankton blooms. Deep-Sea Res. I 41, 747-765.
Townsend, D.W., Cucci, T.L., Berman, T., 1984. Subsurface chlorophyll maxima and vertical distribution of zooplankton in the Gulf of Maine. J. Plankton Res. 6, 793-802.
Townsend, D.W., Pettigrew, N.R., Thomas, A.C., 2001. Offshore blooms of the red tide dinoflagellate, Alexandrium sp., in the Gulf of Maine. Cont. Shelf Res. 21, 347-369.
Townsend, D.W., Pettigrew, N.R., Thomas, A.C., 2005. On the nature of Alexandrium fundyense blooms in the Gulf of Maine. Deep-Sea Res. II 52, 2603-2630.
Townsend, D.W., Thomas, A.C., 2001. Winter-Spring transition of phytoplankton chlorophyll and inorganic nutrients on Georges Bank. Deep-Sea

Res. II. 48, 199-214.
Townsend, D.W., Thomas, M., 2002. Springtime nutrient and phytoplankton dynamics on Georges Bank. Marine Ecology Progress Series 228, 57-74.
Turner, J.T., 2002. Zooplankton fecal pellets, marine snow and sinking phytoplankton blooms. Aquat. Microb. Ecol. 27, 57-102.
Wallhead, P.J., Martin, A.P., Srokosz, M.A., Franks, P.J.S., 2009. Skill assessment via cross-validation and Monte Carlo simulation: An application to Georges Bank plankton models. J. Mar. Syst. 76, 134-150.
Walsh, J.J., Whitledge, T.E., O'Reilly, J.E., Phoel, W.C., Draxler, A.F., 1987. Nitrogen cycling on Georges Bank and the New York Shelf: A comparison between well-mixed and seasonally stratified waters, in: Georges Bank, R. H. Backus (ed.), p 234-246, MIT Press, Cambridge, Mass.
Ward, B.B., 2000. Nitrification and the marine nitrogen cycle, in: Microbial Ecology of the Oceans, D. L. Kirchman (ed.), p 427-453, Wiley-Liss, New York, NY.
Werme, C., Hunt, C.D., 2004. 2003 outfall monitoring overview. Boston: Massachusetts Water Resources Authority. Report ENQUAD 2004-13. 97 pp.
Wunsch, C., 1996. The Ocean Circulation Inverse Problem. Cambridge University Press, Cambridge. 442 pp.
Zhang, Y., Chen, Y., 2007. Modeling and evaluating ecosystem in 1980s and 1990s for American lobster (Homarus americanus) in the Gulf of Maine. Ecological Modeling 203, 475-489.
Zonneveld, C., 1998. A cell-based model for the chlorophyll a to carbon ratio on phytoplankton. Ecological Modelling 113, 55-70.

Figure Captions

Fig. 1. The data. The Chl, NO_3, NH_4, Z and PP data are monthly mean estimates; thus e.g. the third data point is March. The Chl, NO_3 and PP data are 6-m vertical bin averages; NH_4 data are 18-m bin averages. Data gaps appear when no data is in a 6-m monthly bin. December NH_4 "data" are actually interpolated from August and January data to keep the model on track. Note that the colorbar ranges differ in panels (b) vs. (c), (d) vs. (e), (g) vs. (h) and (j) vs. (k). Mixed layer depth (i) is computed primarily from biweekly mean T and S data. The data (a), (c), (e), (f), (h) and (k) to 100 m are assimilated into the model, with (i) prescribed.

Fig. 2. Cost versus iteration for Run 1.

Fig. 3. (a) Taylor diagram for Run 1. The ideal point is (x=1,y=0). (b) Target diagram for Run 1. The ideal point is (x=0,y=0). (c) Modified Taylor diagram for Run 1. The ideal point is (x=1,y=0). (d) Modified Target diagram for Run 1. The ideal point is (x=0,y=0). N=Nitrate; A=Ammonium; C=Chl; D=Detritus; P=Primary Production; Z=Zooplankton; O=Overall.

Fig. 4. Run 1 model-data comparison at selected levels. The circles are the assimilated data, with the $\hat{\sigma}_i$ error estimates shown as vertical lines. Note the N, Chl, PP, Z and A data are monthly estimates; thus e.g. the third data point is March. The solid black line is the model output. In (c), the dashed line is model Chl at 3 m, for comparison.

Fig. 5. Run 1 model-data comparison. The N, Chl, PP, Z and A data are binned monthly. The colors within the circles are the data; the background is the model. The white line is the mixed-layer depth.

Fig. 6. Run 1 annual budgets, in mmol N m^{-2} yr^{-1}. The left number is the 0-24 m integral, which approximates the euphotic zone. The right number is the 0-198 m integral, which is the entire water column. Sinking fluxes of D and P at 24 and 198 m are shown as outward-pointing lines. "Dt" is the vertically-integrated change over one year. "Mix" is the diffusive flux at at 24 and 198 m.

Fig. 7. Run 1 output. (a) Time series of all phytoplankton sources (N to P, A to P) and sinks (P to Z, P to D, P sinking, vertical mixing) as shown in Fig. 6 for 0-24 m. The total sum equals the time rate of change $\partial P/\partial t$. (b) Time series of the individual phytoplankton source and sink terms shown in Fig. 6, as they contribute to $\partial P/\partial t$. (c) Bi-monthly averaged primary production. (d) F-ratio, calculated as NO_3 uptake divided by NO_3 + NH_4 uptake by phytoplankton.

Table 1
Biological Model Equations

$$dN/dt = -UQ_nP + (b_3A + b_4A^2)\exp(-I/b_5)$$

$$dP/dt = U(Q_n + Q_a)P - G_pZ - M_p + W_p dP/dz$$

$$dChl/dt = (Chl/P)(dP/dt) + (1 - R_o/R)Chl/\tau$$

$$dZ/dt = (G_p + G_d + G_z)Z - M_a - M_d - G_zZ$$

$$dD/dt = M_p + M_d - b_1T_aD - b_2T_aD^2 - G_dZ + W_d dD/dz$$

$$dA/dt = M_a + b_1T_aD + b_2T_aD^2 - UQ_aP - (b_3A + b_4A^2)\exp(-I/b_5)$$

$$K = (k_w + k_cChl)\Delta z$$

$$I_{bot} = I_{top}\exp(-K)$$

$$I = (I_{top} - I_{bot})/K$$

$$R_o = \min(r_{mx}, r_{mn} + r_1I)$$

$$R = r_nP/Chl$$

$$L = 1 - \exp(-\alpha I/(\mu R))$$

$$Q_a = (A/n_a)/(1 + (A/n_a) + (N/n_n))$$

$$Q_n = (N/n_n)/(1 + (A/n_a) + (N/n_n))$$

$$U = \mu T_a \min(1, L/(Q_a + Q_n))$$

$$T_a = 2^{(T-16)/10}$$

$$G_p = gT_a s_p^2 P^2/(1 + s_p^2P^2 + s_d^2D^2 + s_z^2Z^2)$$

$$G_d = gT_a s_d^2 D^2/(1 + s_p^2P^2 + s_d^2D^2 + s_z^2Z^2)$$

$$G_z = gT_a s_z^2 Z^2/(1 + s_p^2P^2 + s_d^2D^2 + s_z^2Z^2)$$

$$M_p = a_1P + a_2P^2$$

$$M_a = f_n(G_p + G_d + G_z)Z + (1 - f_d)(a_3Z + a_4Z^2)$$

$$M_d = (1 - f_a - f_n)(G_p + G_d + G_z)Z + f_d(a_3Z + a_4Z^2)$$

$$W_p = \max(w_{pn}, \min(w_{mx}, w_{pn} + w_{p1}(P - w_{p2})))$$

$$W_d = \max(w_{dn}, \min(w_{mx}, w_{dn} + w_{d1}(D - w_{d2})))$$

Table 2
Biological Model Variables

Variable	Description
A	ammonium concentration (μM)
Chl	chlorophyll concentration (mg Chl m^{-3})
D	detritus concentration (μM N)
G_d	zooplankton grazing rate of detritus (1/day)
G_p	zooplankton grazing rate of phytoplankton (1/day)
G_z	zooplankton grazing rate of zooplankton (1/day)
I	average PAR intensity in the model level (W m^{-2})
I_{bot}	PAR intensity at the bottom of the model level (W m^{-2})
I_{top}	PAR intensity at the top of the model level (W m^{-2})
K	optical attenuation factor (nondimensional)
L	light limitation factor (nondimensional)
M_a	zooplankton excretion and mortality to A (μM N/day)
M_d	zooplankton egestion and mortality to D (μM N/day)
M_p	phytoplankton excretion and mortality to D (μM N/day)
N	nitrate concentration (μM)
P	phytoplankton concentration (μM N)
Q_a	ammonium uptake factor (nondimensional)
Q_n	nitrate uptake factor (nondimensional)
R	instantaneous carbon-to-chlorophyll ratio (g C (g Chl)$^{-1}$)
R_o	equilibrium carbon-to-chlorophyll ratio (g C (g Chl)$^{-1}$)
T	temperature (C)
T_a	temperature dependence factor (nondimensional)
U	phytoplankton growth rate (day^{-1})
W_d	detritus sinking rate (m day^{-1})
W_p	phytoplankton sinking rate (m day^{-1})
Z	zooplankton concentration (μM N)

Table 3
Biological Model Parameters

Parameter	Description
α	light-dependent growth rate (g C (g Chl day)$^{-1}$ m^2 W^{-1})
Δz	vertical thickness of the model level = 6 meters
μ	maximum phytoplankton growth rate (day^{-1})
τ	chlorophyll adjustment timescale = 6 days
a_1	linear phytoplankton mortality rate (day^{-1})
a_2	quadratic phytoplankton mortality rate (day^{-1} μM^{-1})
a_3	linear zooplankton mortality rate (day^{-1})
a_4	quadratic zooplankton mortality rate (day^{-1} μM^{-1})
b_1	linear detritus remineralization rate (day^{-1})
b_2	quadratic detritus remineralization rate (day^{-1} μM^{-1})
b_3	linear nitrification rate (day^{-1})
b_4	quadratic nitrification rate (day^{-1} μM^{-1})
b_5	nitrification light-inhibition factor (W m^{-2})
f_a	fraction of zooplankton grazing assimilated (nondim.)
f_d	fraction of zooplankton mortality lost to D (nondim.)
f_n	fraction of zooplankton grazing excreted to A (nondim.)
g	maximum zooplankton grazing rate (day^{-1})
k_c	light attenuation coefficient of chlorophyll (m^2 (mg Chl)$^{-1}$)
k_w	light attenuation coefficient of seawater (m^{-1})
n_a	ammonium uptake half-saturation constant (μM)
n_n	nitrate uptake half-saturation constant (μM)
r_1	C:Chl ratio light-dependence factor = 1 g C (g Chl)$^{-1}$ m^2 W^{-1}
r_{mn}	minimum carbon-to-chlorophyll ratio = 25 g C (g Chl)$^{-1}$
r_{mx}	maximum carbon-to-chlorophyll ratio = 100 g C (g Chl)$^{-1}$
r_n	nitrogen-to-carbon conversion factor = 79.5 g C (mol N)$^{-1}$
s_d	zooplankton grazing half-saturation constant for D ((μM N)$^{-1}$)
s_p	zooplankton grazing half-saturation constant for P ((μM N)$^{-1}$)
s_z	zooplankton grazing half-saturation constant for Z ((μM N)$^{-1}$)
w_{mx}	maximum sinking rate = 5 m day^{-1}

Table 3
Biological Model Parameters (continued)

Parameter	Description
w_{dn}	minimum detritus sinking rate (m day^{-1})
w_{d1}	linear detritus sinking rate factor (m^4 (mmol day)$^{-1}$)
w_{d2}	linear detritus sinking rate threshold (μM)
w_{pn}	minimum phytoplankton sinking rate (m day^{-1})
w_{p1}	linear phytoplankton sinking rate factor (m^4 (mmol day)$^{-1}$)
w_{p2}	linear phytoplankton sinking rate threshold (μM)

All molar units refer to moles of nitrogen per liter of seawater.

Table 4
Model Runs with Cost and Description

Run	Cost	Description
1	1.077	Central Run
2	1.098	$b_1 = 0, b_2 = 0.01$
3	1.124	$b_1 = 0, b_2 = 0.02$
4	1.091	$s_z = 2$
5	1.119	$a_2 = 0, 0 < a_1 \approx 0.1 < 0.4$
6	1.145	$s_d = 1$
7	1.082	$s_d = 0$
8	1.090	$s_z = 1$
9	1.083	$s_z = 0$
10	1.077	$w_{p2} = 0$
11	1.096	$w_{d2} = 0$
12	1.092	$b_5 = 15$
13	1.092	$b_3 = 0.05$
14	1.093	$b_3 = 0.1$
15	1.097	$a_4 = 0, 0 < a_3 \approx 0.1 < 0.5$
16	1.091	$0 < s_d \approx 1 < 4$
17	1.091	$w_{d2} = 0, 0 < w_{dn} \approx 1 < 5$
18	1.077	$w_{pn} = 0, 0 < w_{p2} \approx 0.15 < 0.3$
19	1.101	$w_{d1} = 0, 0 < w_{dn} \approx 1 < 5$
20	1.081	$w_{dn} = 0$
21	1.094	$w_{pn} = 0$
22	1.081	$w_{pn} = 0, w_{p2} = 0$
23	1.088	$0.01 < n_n \approx 0.1 < 1$
24	1.083	$0.01 < n_a \approx 0.1 < 1$
25	1.115	$1 < \mu \approx 2 < 3$
26	1.081	$\alpha = 4$
27	1.087	$g = 1$
28	1.090	$s_p = 3.8$
29	1.097	$s_p = 3$
30	1.078	$s_d = 0.2$

Table 4
Model Runs with Cost and Description (continued)

Run	Cost	Description
31	1.082	$\alpha = 4.9$
32	1.077	$g = 1.3$
33	1.091	$a_4 = 0.27$
34	1.079	$\alpha = 5$
35	1.077	$a_2 = 0.11$
36	1.080	$a_2 = 0.1$
37	1.094	$0.4 < f_n \approx 0.5 < 0.7$
38	1.086	$0.2 < f_a \approx 0.25 < 0.3$
39	1.090	$a_4 = 0.25$
40	1.077	$a_2 = 0.12$
41	1.110	$g = 1.4$
42	1.090	$f_d = 0.4$
43	1.090	$f_d = 0.5$
44	1.092	$b_1 = 0.01$
45	1.096	$0.02 < b_1 \approx 0.03 < 0.04$
46	1.084	$b_1 = 0, 0.01 < b_2 \approx 0.02 < 0.04$
47	1.091	$b_3 = 0.085$
48	1.093	$b_3 = 0, 0 < b_4 \approx 0.05 < 0.1$
49	1.089	$b_5 = 12.459$
50	1.091	$b_5 = 10$
51	1.093	$k_c = 0.0182$
52	1.090	$k_c = 0.0170$
53	1.093	$k_w = 0.114$
54	1.090	$k_w = 0.112$
55	1.084	$a_2 = 0.15$
56	1.079	$0.005 < b_1 \approx 0.02 < 0.04$
57	1.092	$w_{pn} = 0.3$
58	1.091	$w_{pn} = 0.1$
59	1.092	$w_{pn} = 0.5$
60	1.078	$w_{p1} = 3.5$

Table 4
Model Runs with Cost and Description (continued)

Run	Cost	Description
61	1.083	$w_{p1} = 3$
62	1.092	$0 < w_{p1} \approx 1 < 3$
63	1.088	$w_{dn} = 0.53$
64	1.090	$w_{dn} = 0.5$
65	1.078	$w_{d1} = 3.36$
66	1.080	$w_{d1} = 3$
67	1.090	$0 < w_{d2} \approx 0.5 < 1$
68	1.089	$0 < w_{d1} \approx 1 < 3$
69	1.112	$w_{p1} = 0$
70	1.100	$w_{d1} = 0$
71	1.077	$a_2 = 0.13$
72	1.0802	$a_2 = 0.14$
73	1.0844	$w_{p2} = 0.16$
74	1.0837	$w_{p2} = 0.17$
75	1.0844	$w_{p2} = 0.15$

Table 5
Run 1 Parameter Values

parameter	optimized value	initial value	min. limit	max. limit
n_n	1.0[a]	—	—	—
n_a	0.01[b]	—	—	—
μ	3.0[a]	—	—	—
α	4.89 ± 0.18	4.0	3.0	5.0[c]
g	1.286 ± 0.040	1.0	0.3[d]	1.5[d]
s_p	3.76 ± 0.12	2.0	0.2	4.0
s_d	0.200 ± 0.040	1.0	0.2	4.0
s_z	1.96 ± 0.11	2.0	0.2	4.0
a_1	0.0	—	—	—
a_2	0.111 ± 0.011	0.1	0.0	0.4
f_n	0.6[d]	—	—	—
f_a	0.3[d]	—	—	—
a_3	0.0	—	—	—
a_4	0.272 ± 0.015	0.1	0.0	0.3
f_d	0.499 ± 0.063	0.4	0.3[d]	0.5[d]
b_1	0.0101 ± 0.0018	0.02	0.01	0.04
b_2	0.0	—	—	—
b_3	0.085 ± 0.017	0.05	0.0	0.1
b_4	0.0	—	—	—
b_5	12.5 ± 3.5	15.0	10.0	20.0
k_c	0.0182 ± 0.0013	0.0303[e]	0.0170[f]	0.0562[g]
k_w	0.1137 ± 0.0038	0.134[e]	0.112[g]	0.160[f]
w_{pn}	0.306 ± 0.076	0.5	0.0	1.0
w_{p1}	3.49 ± 0.39	1.0	0.0	4.0
w_{p2}	0.3[h]	—	—	—

Table 5
Run 1 Parameter Values (continued)

parameter	optimized value	initial value	min. limit	max. limit
w_{dn}	0.53 ± 0.15	1.0	0.5	5.0[i]
w_{d1}	3.36 ± 0.99	1.0	0.0	4.0
w_{d2}	1.0[h]	—	—	—

[a]maximum limit.
[b]minimum limit.
[c]based on Table 16 in Parsons et al. (1984).
[d]assuming g=0.3, f_a=0.2 and f_n=0.5 for copepods (Anderson and Hessen, 1995) and g=1.0, f_a=0.33 and f_n=0.33 for microheterotrophs (Pelegri et al., 1999)
[e]based on a regression of $K_d(443)$ on Chl from data in Table 1 in Sosik et al. (2001).
[f]HydroQual (1995); minimum k_w estimate for Massachusetts Bay, p 5-18.
[g]based on a regression of Chl on $K_d(443)$ from data in Table 1 in Sosik et al. (2001).
[h]based on lowest observed surface concentrations.
[i]based on Charette et al. (2001).

Table 6
Run 1 Model and Data Statistics

	NO_3	NH_4	Chl	PP	Det	Z
units	μM	μM	mg m^{-3}	μM N d^{-1}	μM N	mmol N m^{-2}
n_i	151	54	196	84	34	12
$\bar{m}$ (u)	6.6730	0.1346	0.5052	0.1391	1.0277	20.0604
$\bar{d}$ (u)	6.0856	0.1341	0.5444	0.2581	1.3837	28.5425
σ_m (u)	3.7325	0.1236	0.4524	0.2023	0.2487	5.2751
σ_d (u)	3.5281	0.0386	0.4714	0.2088	0.4844	11.7677
$\hat{\sigma}_i$ (u)	1.47	0.27	0.22	0.03-0.27	0.86	10.8
r	0.9070	0.4169	0.8428	0.5450	0.1786	0.2503
σ_m/σ_d	1.0579	3.2021	0.9598	0.9686	0.5134	0.4483
RMSE (u)	1.6841	0.1131	0.2625	0.2294	0.6166	14.3935
Bias (u)	0.5873	0.0004	-0.0392	-0.1190	-0.3560	-8.4821
uRMSE (u)	1.5784	0.1131	-0.2596	-0.1962	-0.5034	-11.6287
nRMSE	0.4773	2.9298	0.5570	1.0986	1.2729	1.2231
nBias	0.1665	0.0108	-0.0831	-0.5696	-0.7350	-0.7208
nuRMSE	0.4474	2.9298	-0.5507	-0.9393	-1.0393	-0.9882

(u) signifies the quantity has units (see units). Other quantities are dimensionless.
n_i is the number of observations of this type (month-depth bins).
$\bar{m}$ and $\bar{d}$ are model and data means, respectively.
σ_m and σ_d are model and data standard deviations, respectively.
$\hat{\sigma}_i$ is an estimate of the error in the data values (Sec. 2).
$r = (\sum_{i=1}^{n}(m_i - \bar{m})(d_i - \bar{d}))/(n\sigma_m\sigma_d)$, correlation coefficient.
RMSE = $\sqrt{\sum_{i=1}^{n}(m_i - d_i)^2/n}$, Root-Mean-Square Error.
Bias = $\bar{m}$ - $\bar{d}$
uRMSE = sign(σ_m- σ_d) $\sqrt{\sum_{i=1}^{n}((m_i - \bar{m}) - (d_i - \bar{d})))^2/n}$, unbiased RMSE,
using the sign convention of Jolliff et al. (2009)
nRMSE = RMSE/σ_d, normalized RMSE (Jolliff et al., 2009)
nBias = Bias/σ_d, normalized Bias (Jolliff et al., 2009)
nuRMSE = uRMSE/σ_d, normalized unbiased RMSE (Jolliff et al., 2009)

Table 7
Run 1 Model Statistics, using $m_i/\hat{\sigma}_i$ and $d_i/\hat{\sigma}_i$

	NO_3	NH_4	Chl	PP	Det	Z	All
r	0.9070	0.4169	0.8428	0.4594	0.1786	0.2503	0.8946
σ_m/σ_d	1.0579	3.2021	0.9598	1.1690	0.5134	0.4483	1.0881
RMSE	1.1457	0.4189	1.1933	1.0393	0.7170	1.3327	1.0773
Bias	0.3996	0.0015	-0.1781	-0.7335	-0.4140	-0.7854	-0.1123
uRMSE	1.0737	0.4189	-1.1800	0.7363	-0.5854	-1.0767	1.0715

Table 8
Annual New and Primary Production (g C m^{-2} yr^{-1}) in Wilkinson Basin[a]

New Prod.	Prim. Prod.	Reference
190	270	O'Reilly et al. (1987), NCP table 21.5; fig. 21.7
110-186	162-364	Townsend (1991), table 1
59	290	Townsend (1998)
—	182[b]	Graziano et al. (2000), table 3
27-63	—	Benitez-Nelson et al. (2000), export at 10 m
93[c]	—[d]	Charette et al. (2001), table 3 and sta. 34 in table 4
44	276	Bisagni (2003)
—	322	computed from data in Fig. 1j
97 ± 58	267 ± 67	mean and std of observation-based estimates
71	176	This study, Run 1

[a]Conversions made using a Redfield ratio of 6.625 mol C (mol N)$^{-1}$ where necessary
[b]"assuming negligible production from December-January", thus an underestimate of 0-17%, indicating 182-218 g C m^{-2} yr^{-1}
[c]computed as the mean of POC export at 50 m in March (29 mmol C m^{-2} d^{-1}), June (18 mmol C m^{-2} d^{-1}) and September (16.5 mmol C m^{-2} d^{-1})
[d]Charette et al. (2001) observed a PP of 130 g C m^{-2} yr^{-1} in Sep 1997 in Wilkinson Basin (Station 34 in their table 4), but this probably underestimates the annual mean

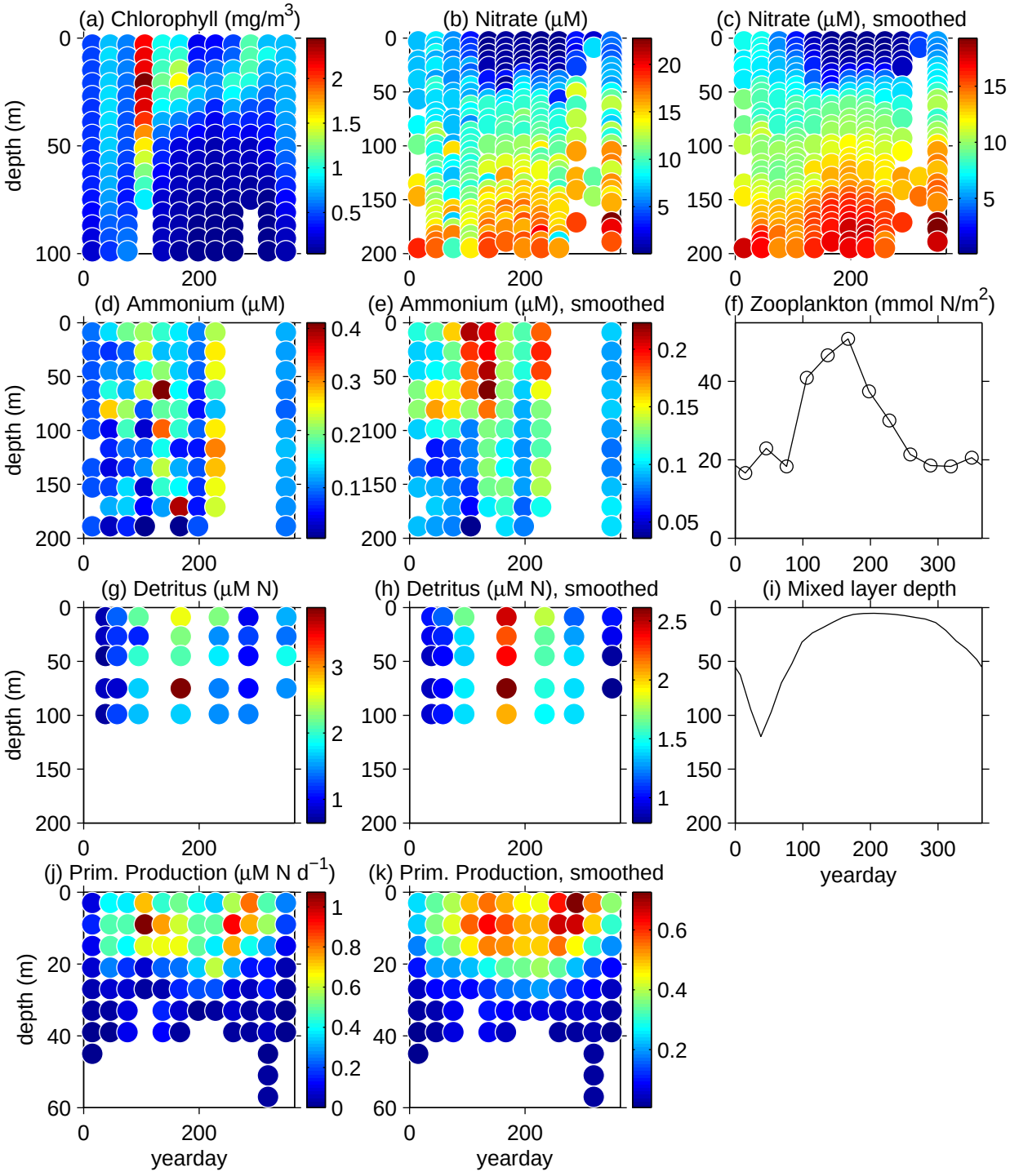
(a) Chlorophyll (mg/m3)
(b) Nitrate (µM)
(c) Nitrate (µM), smoothed
(d) Ammonium (µM)
(e) Ammonium (µM), smoothed
(f) Zooplankton (mmol N/m2)
(g) Detritus (µM N)
(h) Detritus (µM N), smoothed
(i) Mixed layer depth
(j) Prim. Production (µM N d−1)
(k) Prim. Production, smoothed
depth (m)
yearday

Fig. 1.

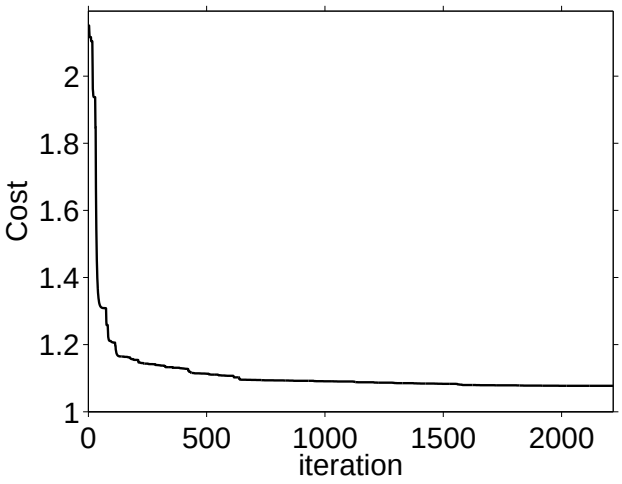

Fig. 2.

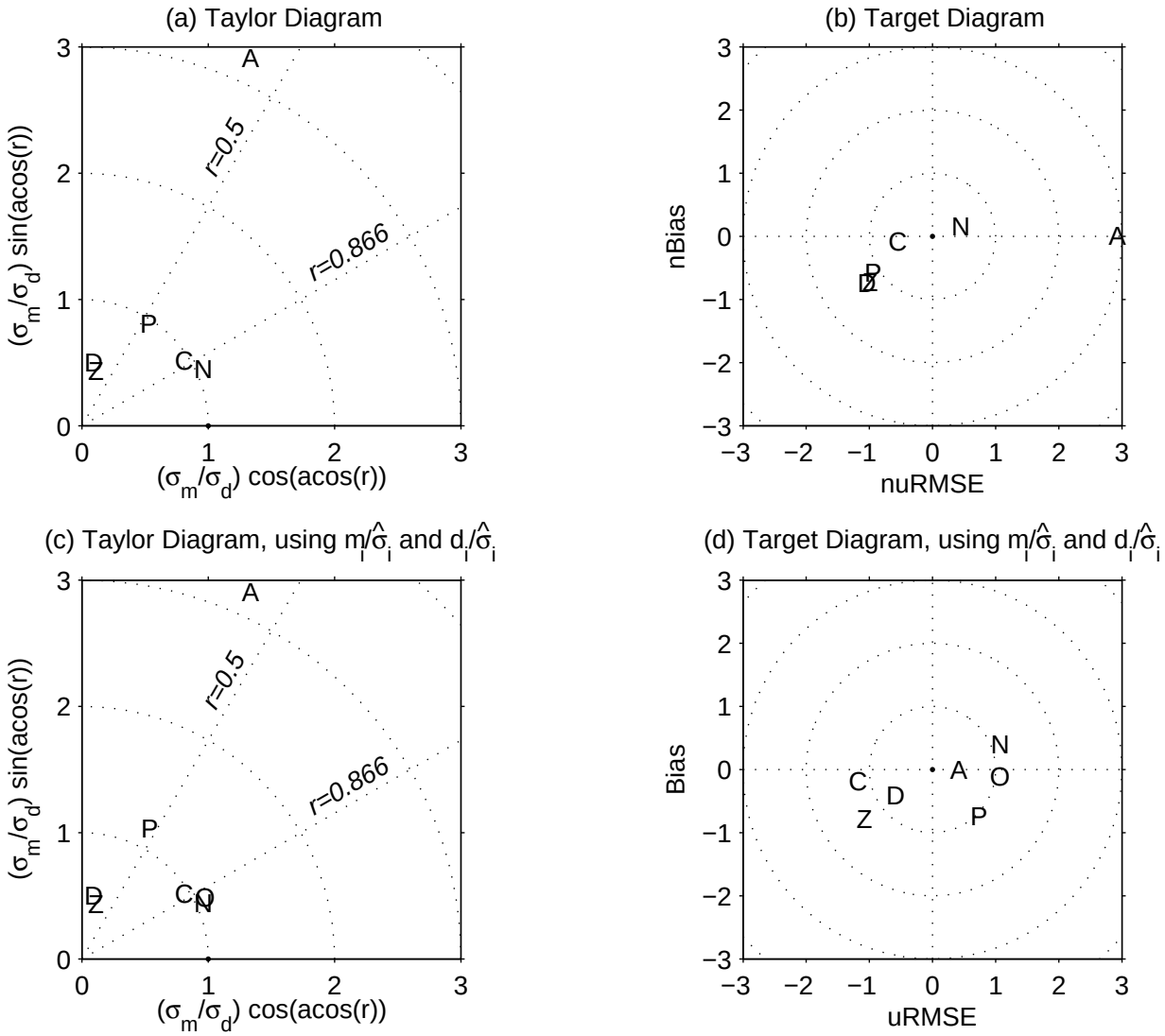
(a) Taylor Diagram
(b) Target Diagram
(c) Taylor Diagram, using $m_i/\hat{\sigma}_i$ and $d_i/\hat{\sigma}_i$
(d) Target Diagram, using $m_i/\hat{\sigma}_i$ and $d_i/\hat{\sigma}_i$
r=0.5
r=0.866
$(\sigma_m/\sigma_d)\sin(\arccos(r))$
$(\sigma_m/\sigma_d)\cos(\arccos(r))$
nBias
nuRMSE
Bias
uRMSE
A
C
N
P
D
Z
O

Fig. 3.

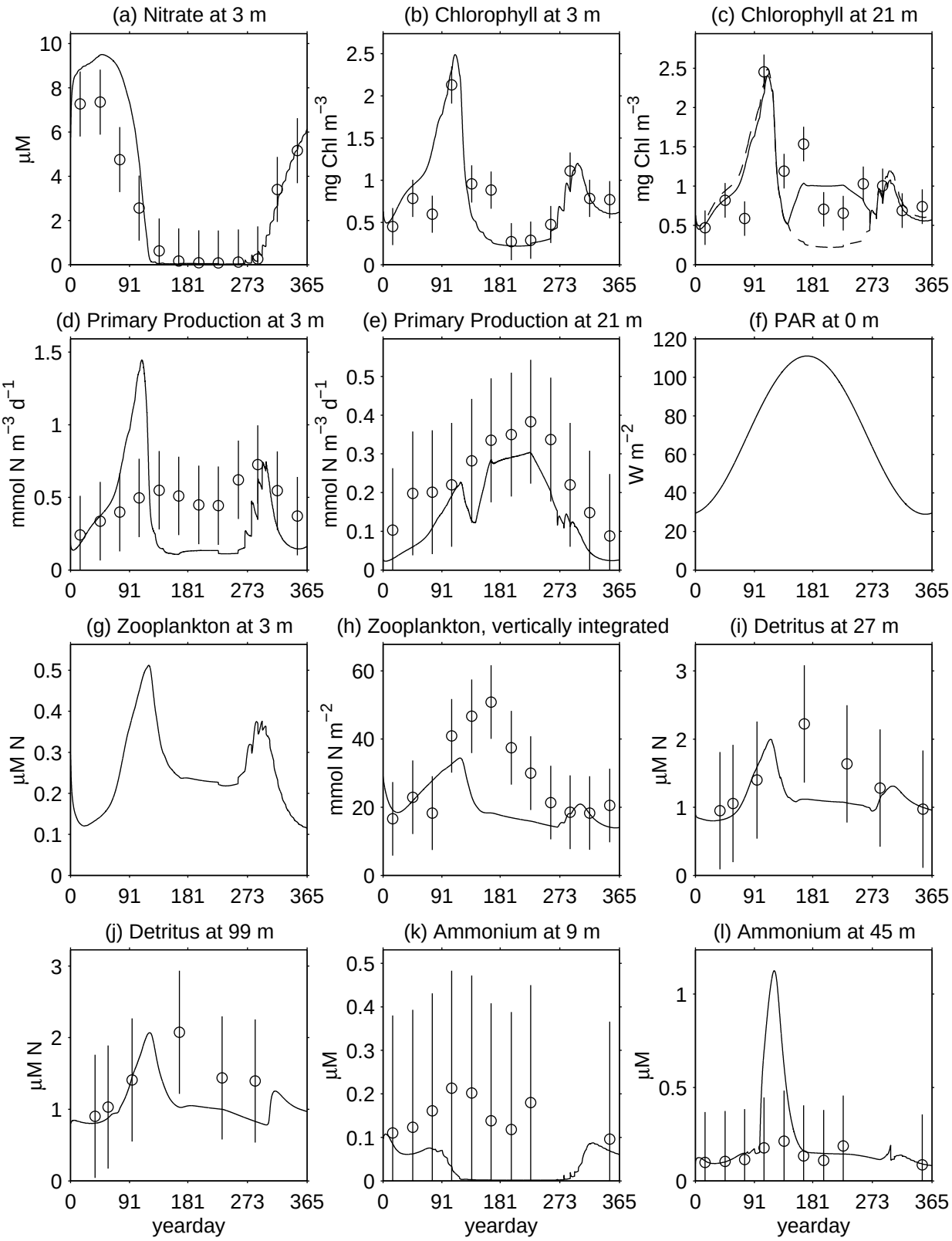
(a) Nitrate at 3 m
µM
(b) Chlorophyll at 3 m
mg Chl m^{-3}
(c) Chlorophyll at 21 m
mg Chl m^{-3}
(d) Primary Production at 3 m
mmol N m^{-3} d^{-1}
(e) Primary Production at 21 m
mmol N m^{-3} d^{-1}
(f) PAR at 0 m
W m^{-2}
(g) Zooplankton at 3 m
µM N
(h) Zooplankton, vertically integrated
mmol N m^{-2}
(i) Detritus at 27 m
µM N
(j) Detritus at 99 m
µM N
(k) Ammonium at 9 m
µM
(l) Ammonium at 45 m
µM
0 91 181 273 365
yearday

Fig. 4.

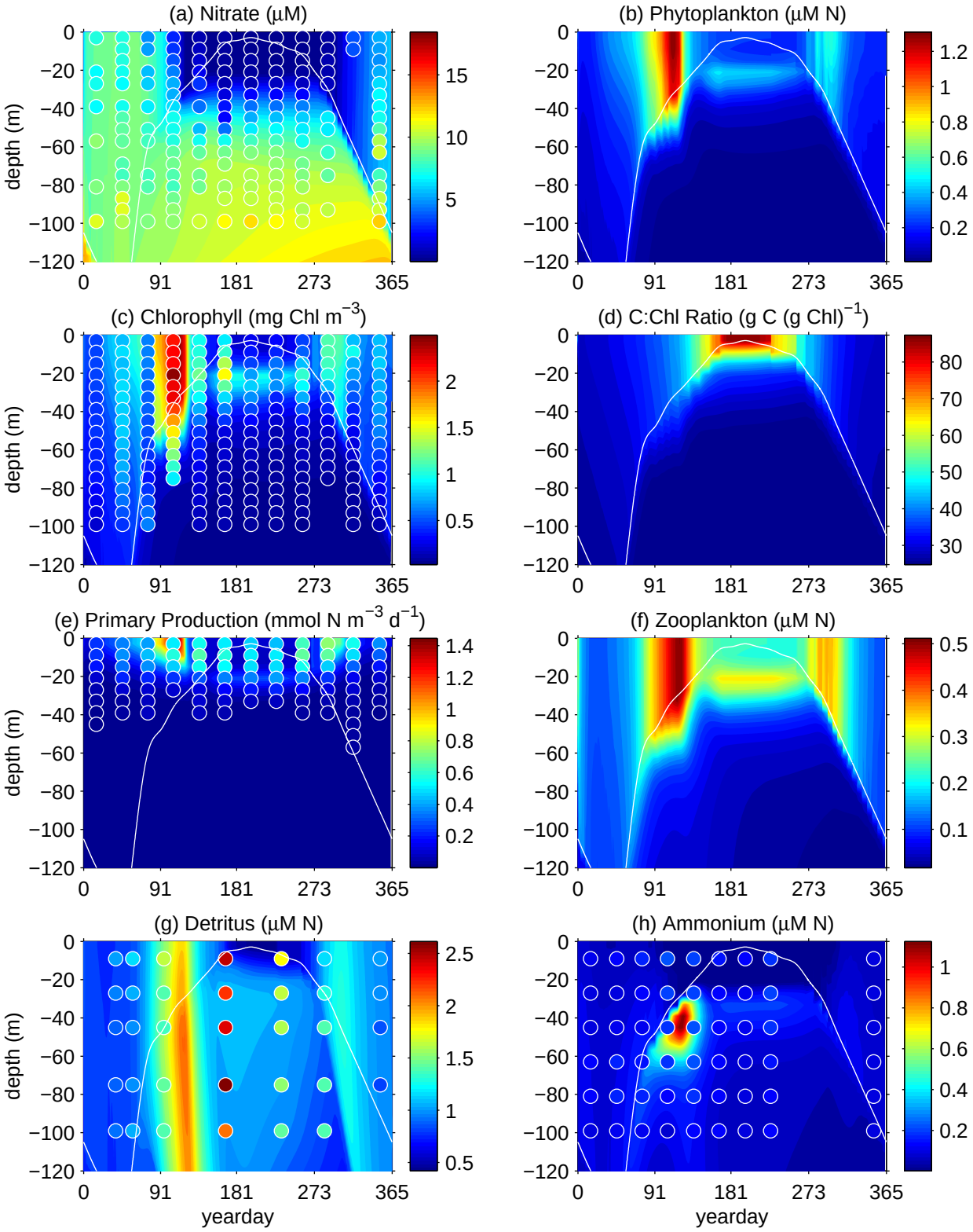
(a) Nitrate (μM)
(b) Phytoplankton (μM N)
(c) Chlorophyll (mg Chl m^{-3})
(d) C:Chl Ratio (g C (g Chl)$^{-1}$)
(e) Primary Production (mmol N m^{-3} d^{-1})
(f) Zooplankton (μM N)
(g) Detritus (μM N)
(h) Ammonium (μM N)
depth (m)
yearday
0
−20
−40
−60
−80
−100
−120
0
91
181
273
365
15
10
5
1.2
1
0.8
0.6
0.4
0.2
2
1.5
1
0.5
80
70
60
50
40
30
1.4
1.2
1
0.8
0.6
0.4
0.2
0.5
0.4
0.3
0.2
0.1
2.5
2
1.5
1
0.5

Fig. 5.

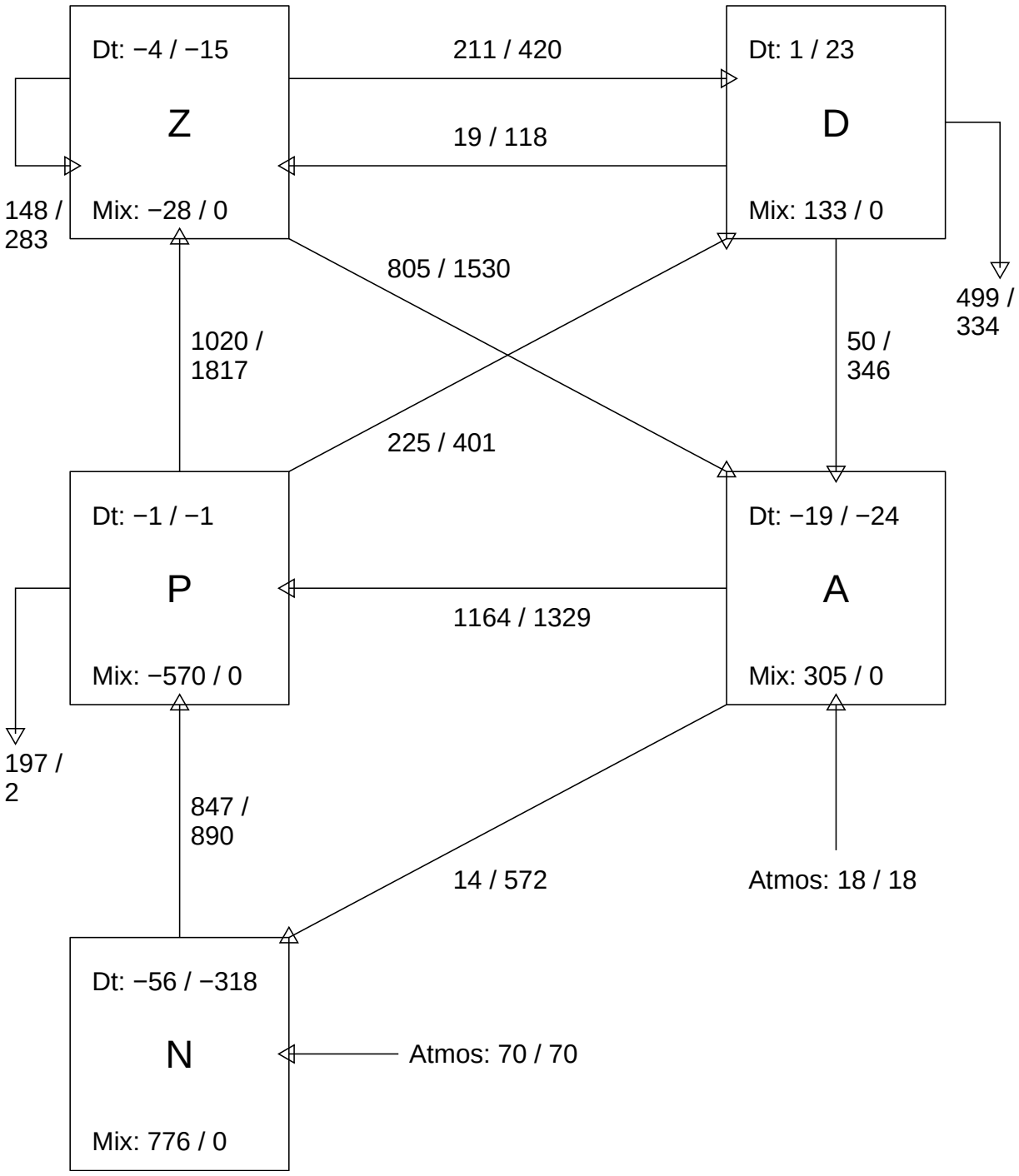
Dt: −4 / −15
Z
Mix: −28 / 0
211 / 420
19 / 118
Dt: 1 / 23
D
Mix: 133 / 0
148 / 283
805 / 1530
499 / 334
1020 / 1817
50 / 346
225 / 401
Dt: −1 / −1
P
Mix: −570 / 0
Dt: −19 / −24
A
Mix: 305 / 0
1164 / 1329
197 / 2
847 / 890
14 / 572
Atmos: 18 / 18
Dt: −56 / −318
N
Mix: 776 / 0
Atmos: 70 / 70

Fig. 6.

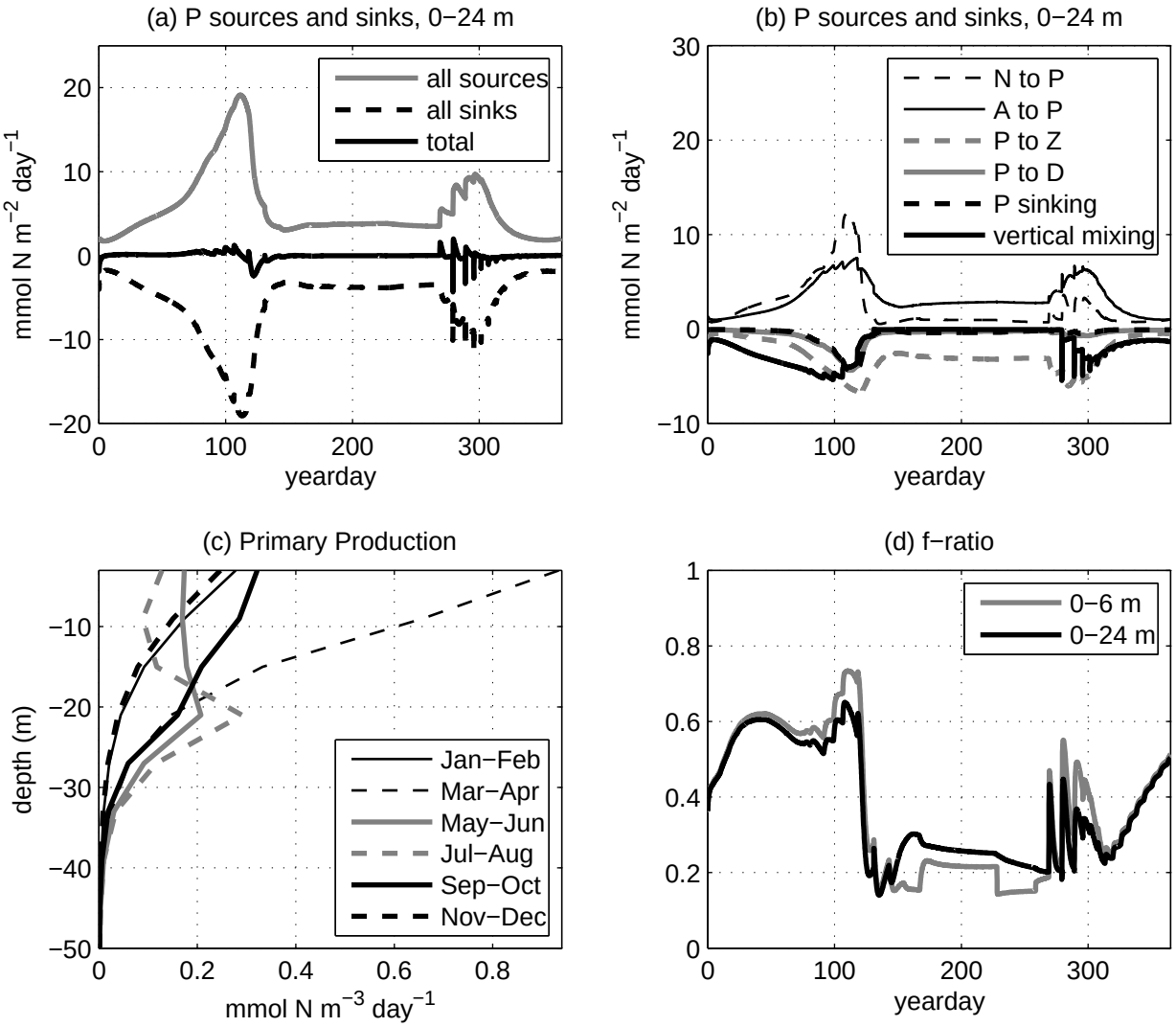
(a) P sources and sinks, 0−24 m
all sources
all sinks
total
mmol N m^{-2} day^{-1}
yearday
(b) P sources and sinks, 0−24 m
N to P
A to P
P to Z
P to D
P sinking
vertical mixing
mmol N m^{-2} day^{-1}
yearday
(c) Primary Production
Jan−Feb
Mar−Apr
May−Jun
Jul−Aug
Sep−Oct
Nov−Dec
depth (m)
mmol N m^{-3} day^{-1}
(d) f−ratio
0−6 m
0−24 m
yearday

Fig. 7.