

Geophysical Research Letters

Supporting Information for

Diatom Hotspots Driven by Western Boundary Current Instability

Hilde Oliver¹, Weifeng G. Zhang¹, Walker O. Smith, Jr.,^{2,3} Philip Alatalo¹, P. Dreux Chappell⁴, Andrew J. Hirzel¹, Corday R. Selden⁴, Heidi M. Sosik¹, Rachel H. R. Stanley⁵, Yifan Zhu⁴, and Dennis J. McGillicuddy, Jr.¹

¹Woods Hole Oceanographic Institution, Woods Hole, MA, USA

²Virginia Institute of Marine Science, College of William & Mary, Gloucester Point, VA, USA

³School of Oceanography, Shanghai Jiao Tong University, Shanghai, People's Republic of China

⁴Department of Ocean and Earth Sciences, Old Dominion University, Norfolk, VA USA

⁵Department of Chemistry, Wellesley College, Wellesley, MA, USA

Contents of this file

Text S1 to S12
Figures S1 to S10
Tables S1 to S3

Additional Supporting Information (Files uploaded separately)

Caption for Dataset S1

Introduction

The Supporting Information text primarily details the collection and processing of the data presented in the main text, in addition to describing the methodology used for the bio-optical model and the 1D physical-biological model. The included Figures and Tables provide additional visualizations to complement the Main Text, as well as to support the Supporting Information text.

Supplementary Information Text

Text S1. CTD profiling and bottle sampling

The CTD rosette was equipped with a SeaBird 911 CTD system, a WetLabs ECO-AFL/FL fluorometer, and a photosynthetically active radiation (PAR) sensor. Surface PAR levels were also measured. PAR, temperature, salinity, and fluorescence were measured on all CTD casts.

Twenty-four 10 L Niskin bottles fitted with Teflon-coated external closures were used for water column sampling. At each station, samples were typically collected at 12 discrete depths for assessment of nutrient concentrations. These samples were syringe-filtered and stored at -20°C until analysis at the WHOI Nutrient Analytical Facility. Nitrate and silicate were measured using standard AutoAnalyzer techniques. To measure ammonium concentrations, site water was cartridge-filtered (0.1 µm, Pall Co.) directly from Niskin bottles using a peristaltic pump. Filtrate was collected in Falcon™ tubes that were pre-treated with orthophthaldialdehyde (OPA) and measured on-board via the OPA method (Holmes et al., 1999) with a detection limit of 10 nM.

To measure particulate organic carbon and nitrogen, water was collected from the Niskin bottles and filtered through combusted 0.7 µm glass fiber filters (Whatman GF/F), rinsed with a weak acid (0.01 N HCl in seawater) to remove carbonates, then dried in combusted glass vials at 60 °C. Diatom biomass was assessed by sampling for biogenic silica. Samples were filtered through 0.6 µm polycarbonate filters, dried at 60°C in plastic Petri dishes, and dissolved in strong acid.

Fluorescence was converted into chlorophyll *a* concentrations using a regression between fluorescence values and extracted chlorophyll *a* measurements from Niskin bottles. CTD fluorescence (FL) was converted to chlorophyll *a* concentration with the following regression:

$$\mathbf{Chl\ a = 0.50FL + 0.13} \quad \mathbf{(1)}$$

Which has $R^2 = 0.88$.

Text S2. Video Plankton Recorder Surveys

To assess the large microplankton and small mesoplankton community alongside fine-scale measurements of temperature, salinity, and fluorescence, a Video Plankton Recorder II (VPR, from SeaScan Inc.) was towed behind the ship. The VPR consists of a towed body, containing a Seabird Electronics Inc. CTD (SBE 49 FastCat), oxygen sensor (SBE 43), fluorometer (ECO FLNTU-4050), ECO Triplet (ECO BBFL2-123), PAR (photosynthetically active radiation; Biospherical Instruments Inc. QCP-200L), and a synchronized video camera and xenon strobe (Davis et al., 2005). The VPR was towed at 10 knots (5.1 m s⁻¹), undulating between depths as shallow as 5m and as deep as 120m approximately every 6 minutes. This provided a minimum horizontal resolution of 1.8 km throughout the tow. For this study, we show data from VPR surveys 3, 7, and 8, which consisted of 234, 126, and 110 complete transits from the surface to 120 m.

Concentrations of chlorophyll *a* were determined from fluorescence by regressing chlorophyll values from the CTD casts immediately following the three VPR transects,

using the fluorescence (FL) from the nearest VPR oscillation to the CTD profile. The regression used was

$$Chl\ a = 0.59FL + 0.13 \quad (2)$$

which has $R^2 = 0.81$.

Text S3. REMUS 600 hotspot survey

On July 12-13, the autonomous underwater vehicle (AUV) REMUS 600 was deployed to survey the hotspots (Fig. S9). The AUV was equipped with a CTD, a WET Labs - ECO Triplet sensor, and an optical nitrate sensor (Submersible Ultraviolet Nitrate Analyzer-SUNA). REMUS chlorophyll *a* and nitrate concentrations were determined from AUV measurements by regressing chlorophyll and nitrate values from the CTD casts and bottle nutrient measurements immediately preceding the REMUS survey, applying values from the nearest REMUS oscillation to the CTD profile. For chlorophyll, we use the regression

$$Chl\ a = 0.70FL + 0.30 \quad (3)$$

which has $R^2 = 0.55$. For nitrate, we use the regression

$$NO_3 = 0.75[SUNA\ NO_3] - 2.87 \quad (4)$$

which has $R^2 = 0.88$.

Text S4. Imaging FlowCytobot

The Imaging FlowCytobot (IFCB) records high-resolution ($\sim 1\ \mu\text{m}$) images of a wide variety of phytoplankton taxa, from those $< 10\ \mu\text{m}$ to those hundreds of microns long (Sosik & Olson, 2007). The IFCB is equipped with a red diode laser; the laser light is scattered as particles pass through. Frame capture from a CCD camera was triggered by both fluorescence and scattering. The high-quality images allow the identification of many taxa to the genus level. IFCB images were recorded from Niskin bottles taken both within and outside the biomass peaks at hotspot stations. Images were processed using the methods described in (Sosik et al., 2020), with biovolume calculated with the method described in Moberg & Sosik, 2012. Results in Fig. 3 are based on cell, chain, and colony biovolumes summed over the designated taxonomic groupings. Due to the lower size range detection limit for IFCB, only targets $> 5\ \mu\text{m}$ (in terms of either equivalent spherical diameter or maximum Feret diameter) were included in the biovolume sums. For species and genus level information about diatom contributions (Fig. 3b, taxa that contributed 8% or more to total diatom biovolume at any depth are shown explicitly; taxa consistently below that threshold are aggregated as "other diatoms").

Text S5. VPR automated classification of diatom rods

Diatom distributions were measured in situ using the VPR. Plankton video was collected at 30 Hz. Individual 1380 x 1034 pixel video frames ($\sim 20\ \text{mm} \times 15\ \text{mm} \times 23\ \text{mm}$

volume imaged) were passed through object-identification software to identify “regions of interest” (ROIs), which were then saved to disk with a time-stamp naming convention. ROIs were automatically classified using a convolutional neural network (CNN) specifically designed for plankton (González et al., 2019). The CNN was trained using approximately 1000 manually annotated ROIs for 29 categories (Table S1). The training dataset was composed of ROIs from two cruises, the March 2018 *R/V Neil Armstrong* cruise AR29 and the July 2019 *R/V Thomas G. Thompson* cruise TN368. Most categories were trained with ROIs solely from TN368, but three (echinoderm larvae, bloom conditions, and spherical colonial plankton) were from AR29. Bloom conditions were defined as ROIs that contained spherical colonial plankton (*Phaeocystis pouchetti* in the case of AR29), diatom chains, and marine snow. Diatom hotspot conditions within TN368 were composed of *Thalassiosira diporocyclus*, diatom chains, and marine snow. The globular morphology of *T. diporocyclus* was similar enough to *P. pouchetti* that the CNN associated the diatom hotspot observed in TN368 with the bloom condition category, enabling the use of the bloom condition category as a proxy for diatom hotspot conditions.

CNN accuracy by taxon was determined by F1 score ($F1 = 2 * (\text{precision} * \text{recall}) / (\text{precision} + \text{recall})$). Precision (true positives / (true positives + false positives)) and recall (true positives / (true positives + false negatives)) were determined by withholding 20% of the original training set ROIs for each taxon and running the CNN on the withheld ROIs. The overall CNN F1 score was 86.64%.

To quantify the accuracy of the diatom distributions, the first 100 ROIs from each hour classified as diatom chains were checked for false positives. This yielded a mean false positive rate of approximately 17.4%. False negative rates were computed by counting the number of false negatives for the first 250 ROIs of each hour within VPR Tow 2 in the following categories: bloom conditions (51.8%), bubbles (0.6%), marine snow (2.3%), out-of-focus (2.4%), and unknown (6.8%). These categories were chosen because they were the most abundant within the tow, each containing more than 10,000 ROIs. Out-of-focus was the most abundant category, containing approximately four times as many classified ROIs (294,809) as that of diatom chains (61,153). The other categories contained approximately the same number of ROIs as that of diatom chains (bubbles = 55,505; marine snow = 45,417; unknown = 89,389), except bloom conditions, which contained 13,049 ROIs.

To account for the impact of false negatives on the diatom distributions, the counts were augmented with an estimate computed by multiplying the rate of false negatives for each hour by the taxon concentrations within that hour (i.e., $\text{diatom}_{\text{xyz hour}} = \text{diatom}_{\text{original xyz hour}} + (\text{non-diatom taxon}_{\text{xyz hour}} * \text{false negative rate}_{\text{hour}})$). The inclusion of diatom chain false negatives from other categories increases the diatom chain absolute abundance by approximately a factor of three, without changing their geographical distribution (Fig. S4).

Precise quantification of diatom concentrations is further challenged by the fact that there were highly variable numbers of diatom chains per ROI in the cores of the hotspots, in addition to colonial forms of *T. diporocyclus* and marine snow. The density of material in the bloom conditions category was such that in-focus objects were often

occluded by out-of-focus objects, making the enumeration of diatoms within individual ROIs difficult, time-consuming, and subjective. As such, ascertaining the peak concentration of diatoms within the hotspots is impractical.

We therefore decided to present our results in terms of relative diatom concentration using the “diatoms” category, which consists primarily of single in-focus diatom chains. The resulting maps are a reliable proxy for the overall diatom distribution, recognizing of course that the peak concentrations in the inner cores of the diatom hotspots are underestimated.

Text S6. DNA Sequencing and Analysis

Water samples were collected in acid-washed (10% HCl) 4 L amber polypropylene bottles from Niskin bottles and gently filtered (1 to 5 L) onto 0.22 μm polyethersulfone membrane filters (Sterivex™, MilliporeSigma, USA) using a peristaltic pump. Invitrogen RNA_{later} stabilizing solution (ThermoFisher Scientific, USA) was applied to filters following manufacturer specifications. Following the overnight preservation at 4°C, samples were stored in liquid nitrogen (during the cruise) or at -80°C (in the lab) until processing. Before nucleic acid extractions, samples were thawed and the stabilizing solution was vacated. The filter was then removed from the Sterivex™ frame and submerged in RLTPlus buffer (1% β -mercaptoethanol by volume; Qiagen, Germany). Filters were bead-beaten for 2 min using 0.1 and 0.5 μm silica beads to facilitate cell break-down. Finally, the lysate was homogenized using a QIAshredder Spin column (Qiagen, Germany). DNA and RNA were then co-extracted with the AllPrep DNA/RNA Mini Kit (Qiagen, Germany).

The V4 region of 18S rDNA was amplified in triplicate using a polymerase chain reaction (PCR) method (Chappell et al., 2019). The primers employed were designed specifically for diatom DNA barcoding (Zimmermann et al., 2011) modified for two-step amplicon sequencing with Illumina (USA) technology. Replicate PCR products were pooled, verified via gel electrophoresis, and purified (GeneJet PCR Purification Kit; ThermoFisher Scientific, USA). The Nextera DNA Library Preparation Kit (Illumina, USA) was used to prepare the DNA library. In brief, purified PCR products were diluted to < 55 ng DNA μl^{-1} and indexed in a 50 μl reaction following the manufacturer’s protocol. Phusion High-Fidelity PCR Master Mix with HF Buffer (ThermoFisher Scientific, USA) was used in this reaction. Second-round PCR products were cleaned using Ampure XP paramagnetic beads following manufacturer’s protocol (Beckman Coulter, USA), quantified fluorometrically (Invitrogen Qubit 2.0, ThermoFisher Scientific, USA), and diluted to 4 nM. The quality of a subset of samples was evaluated with a Fragment Analyzer System (Agilent Technologies Inc., USA). High-quality samples were then pooled. To improve the quality of sequencing reads, the multiplexed amplicon library was combined 85:15 with PhiX Control v3 Library (4nM, Illumina, USA). The amplicon library was sequenced on an Illumina MiSeq platform using the Illumina v3 2 \times 300 cycle kit.

Diatom 18S rDNA amplicon sequences were de-multiplexed and analyzed using the DADA2 pipeline (version 1.16.0, Callahan et al., 2016). At the start of the pipeline, primers were removed using cutadapt (version 2.10, Martin, 2011). The following

parameters were used to filter reads using the DADA2 filterAndTrim command: maxN = 0, truncLen = c(230,190), maxEE = c(2,2), truncQ = 2, minLen = 50, rm.phix = TRUE. After the DADA2 algorithm generated initial amplicon sequence variants (ASVs), ASVs shorter than 395 or longer than 410 were filtered out. Additionally, chimeric ASVs were filtered out using the removeBimeraDenovo command with the following parameters: method = "consensus", minFoldParentOverAbundance = 8.

The taxonomic assignment of the ASVs followed the protocol of Chappell et al. 2019. The ASV FASTA file was used as an input for a nucleotide BLAST (Altschul et al., 1990) against an in-house database that combined Stramenopile 18S sequences from NCBI (downloaded as of 23 June 2020) and the SILVA eukaryote 18S database. ASVs that did not have a diatom as a top hit were not included in subsequent analyses. ASVs were classified to the species level if they had a >99% identity to a diatom sequence. ASVs were classified as species-like (akin to the cf. designation in morphology) if they had 98-99% identity to a diatom sequence. ASVs with a <98 % identity to a diatom sequence were only classified to the genus level. PRIMER v.7 (Clarke & Gorley, 2018) was used to log transform ASV counts and generate a Bray-Curtis similarity matrix (Bray & Curtis, 1957) that was visualized in a cluster dendrogram generated using the group average clustering mode.

Text S7. Incubation-based primary productivity

Water samples were taken from Niskin bottles at known isolumes, then placed in sterile 285 mL Qorpak bottles, then ~20 μCi $\text{NaH}^{14}\text{CO}_3$ was added. An on-deck incubator holding the bottles had surface seawater flowing through it, with irradiance attenuated by neutral density filters to the light levels at the isolumes sampled. Blue filters were used for isolumes below 30% E_0 . After 24 h, samples were filtered through GFF filters and placed in 7 mL scintillation vials. Size fractionations were conducted at all stations using 20 μm Poretics filters on subsamples from each bottle. 100 μL 1N HCl was added to volatilize absorbed inorganic ^{14}C . Ecolume (5 mL) was then added to each vial, and all vials were counted after 24 h on a liquid scintillation counter. Total activity was measured by counting 100 μL of non-acidified sample in β -phenethylamine.

Text S8. Bio-optical model

Productivity was estimated using a bio-optical model (Fig. S6c, d, g). The model was based on the formulation of Behrenfeld & Falkowski (1997a, 1997b), where depth-resolved productivity is a function of temperature, irradiance (PAR), and chlorophyll concentrations (Eq. 5).

$$PP = C_z \times P_{opt}^B \times f(E_0) \quad (5)$$

where **PP** is primary productivity, **C_z** chlorophyll concentration, **P_{opt}^B** the maximum photosynthetic rate within the water column, and **f(E₀)** an irradiance function within the water column, where **E₀** is surface PAR. A photosynthesis-irradiance response (Jassby & Platt, 1976) was used:

$$P_z^B = P_{opt}^B \times \tanh \left[E_z / E_k \right] \quad (6)$$

At station SLP in the slope sea cyclonic feature (CTD Cast 100), the chlorophyll-normalized carbon fixation rate of hotspot phytoplankton was highest at 15% of surface irradiance levels (45 m depth, Fig S6a), so $E_k = 0.15 \times$ (surface PAR) where PAR < 100 $\mu\text{mol photons m}^{-2} \text{s}^{-1}$, and $E_k = 0.25 \times$ (surface PAR) when PAR > 100 $\mu\text{mol photons m}^{-2} \text{s}^{-1}$.

P_{opt}^B was derived by the 7th order polynomial that was based on 1,698 radioisotope measurements made throughout the ocean (Behrenfeld & Falkowski, 1997b). A photoinhibition multiplier based on the same data set was used to reduce productivity due to photoinhibition when irradiance was > 3 mol photons $\text{m}^{-2} \text{d}^{-1}$ (Eq. 7) (Behrenfeld & Falkowski, 1997b):

$$P_{inh} = -0.0204E_0^2 + 2.515E_0 - 6.675 \quad (7)$$

E_z values were derived from the relationship described by Platt & Jassby, (1976).

To generate irradiance attenuation profiles, the relationship between chlorophyll and attenuation in daytime casts (Morel, 1974) was used to correct for an offset that was observed. We believe this offset was due to dissolved organic carbon that originated from the continental shelf. To compare all stations, we used a constant surface irradiance, which was equal to that of a clear sky irradiance value. Productivity was estimated from the above equations at 1-m intervals from the surface to the 1% isolume, defined as the absolute compensation irradiance 1.3 $\mu\text{mol photons m}^{-2} \text{d}^{-1}$ estimated for the North Atlantic (Siegel et al., 2002). Note that the bio-optical model accounts for declining carbon fixed per unit chlorophyll with declining light (Fig. S6e, jumps in the profiles 10-30 m above the absolute compensation irradiance indicate the transition from $E_k = 0.25 \times$ surface PAR to $E_k = 0.15 \times$ surface PAR). At the surface, the carbon fixed per unit chlorophyll is 7 ± 3 times greater than at the hotspots of maximum productivity (Fig. S6e). Measured ratios of POC to chlorophyll vary widely at the hotspot productivity maximum layers, suggesting potentially high concentrations of marine snow (Fig. S6f), and therefore were not applicable for constraining C:Chl ratios within hotspot algal biomass.

Model performance was also assessed for non-hotspots casts conducted for TN368. The model was run for all primary productivity casts (Fig. S6g, h), and then compared against the incubation productivity (Fig. S6h, i). The model was likely to either over- or underestimate the incubation productivity at most stations, excepting inner shelf stations 41, 43, and 80 (shown by magenta circles on map in Fig. S7), where modeled productivity was much lower than measured (Fig. S6i). Chlorophyll was higher than other shelf stations in the top 20-30 m at these stations (Fig. S6g), but the measured productivity is enhanced as shallow as the sea surface. The model photoinhibition term thus may not be applicable when modeling these stations.

The incubation integrated productivity suggests higher productivity on the shelf and lower productivity within the hotspots, excepting cast 84 where the chlorophyll

maximum productivity was measured (Fig. S7a). The modeled productivity is similar between the hotspots and the shelf, however (Fig S7b), and predicts higher integrated productivity at all hotspot stations relative to the incubation integrated productivity (Fig S7c), despite closely reproducing profiles of measured productivity at the hotspot casts (Fig S6d). While the model typically underestimates productivity for non-hotspot casts, with the inter-quartile range of model-incubation productivity differences for non-hotspot casts being less than 0, the model projects higher productivity for all hotspot casts (Fig. S7e).

To assess model sensitivity and uncertainty, we conducted an error perturbation analysis on the bio-optical model, modifying key parameters over their uncertainties: upper water column light attenuation coefficient k_d (± 0.01 , unitless), measured chlorophyll a concentration ($\pm 5\%$), and absolute compensation irradiance ($\pm 0.3 \mu\text{mol photons m}^{-2} \text{d}^{-1}$). We reran the bio-optical model for each parameter mean, minimum, maximum combination, for a total of 27 runs, as done in the error perturbation analysis in Friedrichs et al. (2009). The maximum perturbation to integrated productivity while varying each parameter one at a time, and all parameters simultaneously are reported in Table S2. Error sums in quadrature are also reported in Table S2. We use the maximum perturbation when varying all parameters simultaneously when reporting the modeled vs. measured difference in integrated productivity.

Text S9. Gross Primary Production estimate from triple oxygen isotopes

Triple oxygen isotopes of dissolved oxygen quantify the fraction of oxygen-derived from photosynthesis rather than from atmospheric oxygen. Profiles of triple oxygen isotopes (Fig. S8) show that photosynthetic oxygen was enhanced in the hotspots compared to samples collected at similar depths but non-hotspot locations on the cruise. Typically, triple oxygen isotopes are used in surface waters to estimate mixed layer gross primary production (GPP). In this case, however, we are interested in GPP below the mixed layer – at a depth of ~50 m in the diatom hotspot. We also do not have a time-series (as in Juranek & Quay, 2013) and thus do not know the amount of time it takes for the triple oxygen isotopic signature to change. As a result, we cannot exactly quantify GPP. However, we can estimate GPP by assuming the deep water (100 m depth) is the source water for the diatom hotspot – or at least has analogous triple oxygen isotopic composition - and that the transit time was about 2 weeks (Fig. 1). The current best practice is to calculate GPP from $\delta^{17}\text{O}$ and $\delta^{18}\text{O}$. However, in this case, since we are calculating GPP below the surface and do not have a regular time series, the equations in Prokopenko et al. (2011) and Kaiser (2011) are not suitable. We thus perform a rough calculation based on the $^{17}\Delta$ where $^{17}\Delta$ is defined as

$$^{17}\Delta = 10^6 \times \left(\left(\frac{\delta^{17}\text{O}}{1000} + 1 \right) - \lambda \left(\frac{\delta^{18}\text{O}}{1000} + 1 \right) \right)$$

(8)

where $\delta^{17}\text{O} = (^{17}\text{O}_{\text{meas}}/^{17}\text{O}_{\text{air}} - 1) \times 1000$, and $\lambda = 0.5179$ (Luz & Barkan, 2005), corresponding to the fractionation associated with respiration (Fig. S8). As described by Hendricks et al. (2004) the ratio R of photosynthetic oxygen to atmospheric oxygen in a water sample is

$$R = \frac{{}^{17}\Delta_{eq} - {}^{17}\Delta_{meas}}{{}^{17}\Delta_{meas} - {}^{17}\Delta_{eq} - {}^{17}\Delta_P} \quad (9)$$

where ${}^{17}\Delta_{eq}$ is the ${}^{17}\Delta$ signature of water equilibrated with atmospheric oxygen = 8 per meg (Stanley et al., 2010), ${}^{17}\Delta_{meas}$ is the ${}^{17}\Delta$ signature measured in a particular water sample, and ${}^{17}\Delta_P$ is the ${}^{17}\Delta$ signature of oxygen produced completely from photosynthesis = 249 per meg (Luz & Barkan, 2000). Therefore, the fraction of photosynthetic oxygen in the water, f_P , is given by:

$$f_P = \frac{R}{R + 1} \quad (10)$$

assuming the only sources of oxygen are photosynthetic or atmospheric.

By combining the fraction of photosynthetic oxygen measured at the diatom hotspot depth (47 m at Station 71) and at the deep source water depth (denoted $f_{P,hs}$ and $f_{P,d}$ respectively) with the oxygen concentration at the hotspot and deep depths ($[\text{O}_2]_{hs}$ and $[\text{O}_2]_d$ respectively), it is possible to calculate the total amount of photosynthetic oxygen, $[\text{O}_2]_{photo}$, added to the water during the transit time when the 100 m water (or source water with analogous oxygen characteristics) was transported to the diatom hotspot:

$$[\text{O}_2]_{photo} = \frac{(f_{P,hs}[\text{O}_2]_{hs} - f_{P,d}[\text{O}_2]_d - \alpha([\text{O}_2]_{hs} - [\text{O}_2]_d))}{(1 - \alpha)} \quad (11)$$

where α is the fraction of oxygen that has been respired that was photosynthetic in origin. If all the respiration happened at the beginning of the transit, then α would equal $f_{P,d}$. If all the respiration happened at the end of the transit, α would equal $f_{P,hs}$. We assume that α is an average of those two possibilities but we also calculate $[\text{O}_2]_{photo}$ using both extremes.

Finally, the rate of GPP, in O_2 units, is given by:

$$GPP = \frac{[\text{O}_2]_{photo}}{\Delta t} \quad (12)$$

where Δt is the assumed time required for transporting the source water to the hotspot.

The triple oxygen isotopic signature, ${}^{17}\Delta_{meas}$, was measured on an isotope ratio mass spectrometer after the samples had been purified to remove all gases other than O_2 and Ar from the samples (Stanley et al., 2010). Δt is estimated to be 14 days (Fig. 1). Oxygen concentration was measured by an SBE O_2 sensor on the CTD. The O_2 sensor has been calibrated in the factory within six months. It was not calibrated on the cruise but since

the difference in hotspot and deep O₂ concentrations is used in this calculation, and this calculation is only a rough estimate due to large unknown in other parameters, a lack of onboard calibration should not lead to significant errors. The measurement uncertainty of ¹⁷Δ based on equilibrated water samples measured during the same period is 5 per meg. The largest source of uncertainty in the calculation comes from the estimates of α and Δt.

The best estimate of GPP, as calculated from comparing the 47 m diatom hotspot sample on cast 71 to the 100 m estimated source water sample from the same cast is 5.7 mmol O₂ m⁻³ d⁻¹, which equals 4.0 mmol C m⁻³ d⁻¹ using a photosynthetic quotient of 1.4. With α equal to the two extremes of f_{P,d} or f_{P,hs}, then the GPP rate ranges from 4.5 mmol O₂ m⁻³ d⁻¹ = 3.2 mmol C m⁻³ d⁻¹ to 7.5 mmol O₂ m⁻³ d⁻¹ = 5.3 mmol C m⁻³ d⁻¹. Given GPP is expected to be roughly a factor of 2.6 larger than NPP, this estimate agrees well with the completely independent estimate of NPP based on the ¹⁴C incubations and bio-optical model.

Text S10. Euphotic depth calculation for Fig. 4

For Fig. 4, the euphotic depth is calculated as the depth of 1% of surface irradiance. Irradiance (*I*) at each depth of the VPR and REMUS 600 data is determined by:

$$\frac{\partial I}{\partial z} = -k_d I \quad (13)$$

where *k_d* is the light attenuation coefficient, and

$$k_d = k_z + k_p [\text{Chl } a] \quad (14)$$

where *k_z* is light attenuation of clear seawater (*k_z* = 0.04 m⁻¹), and *k_p* is the light attenuation from phytoplankton biomass (*k_p* = 0.13 ± 0.04 (μg Chl a L⁻¹)⁻¹ m⁻¹). We derived *k_p* using CTD profiles of photosynthetically active radiation (PAR) from daytime casts from the slope and the shelf, determining for each cast the value of *k_p* that results in the closest 1% light depth to that determined by the ratio of PAR on the rosette to surface PAR sensor values. The two 1% isolumes shown in Fig. 4 contain the 1% depth when applying the *k_p* range of the mean ± 1 standard deviation (*k_p* = 0.09 – 0.16 (μg Chl a L⁻¹)⁻¹ m⁻¹).

Text S11. One-dimensional model

We selected the model initial temperature and salinity profiles using the 1^o summer objectively analyzed climatological means from the National Centers for Environmental Information (NCEI) Northwest Atlantic Regional Ocean Climatology (Seidov, 2016), and nitrate profiles using the 2018 NCEI World Ocean Atlas (Garcia et al., 2019). We initialize the 1D model with the climatological profile south of the MAB most closely satisfying the definition of the GS northern edge (temperature at 200 m = 15°C (Joyce et al., 2000); location shown by white star in Fig. 1A; profiles shown in gray in Fig. 5A).

To simulate macronutrient drawdown, we couple a dissolved inorganic nitrogen (DIN), silicic acid (dSi), and light-limited nitrogen-phytoplankton-zooplankton-detritus (NPZD) biogeochemical model to the physical model (adapted from the Powell NPZD model (Powell et al., 2006); see Text S12 for model equations and parameters). We derive the slope water light-attenuation coefficient using profiles of photosynthetically active radiation (PAR) from daytime casts during TN368. We assume a POC:Chl ratio of 50 and a C:N ratio of 106:16 (Redfield, 1963).

We ran the model for one month (1 Jul – 31 Jul 2019), with a time step of 300 s, and a constant vertical resolution of 10 cm. We assume that the GS upwelling and phytoplankton bloom would take no longer than one month to develop because the large features characterizing the MAB slope region generally take 1-3 weeks to evolve, such as the intrusion of the meander onto the shelfbreak and the development of the WCR streamer (Fig. 1).

The upward vertical velocity was held constant over the entire model period and is applied to the NCEI GS endmembers. To prevent the upward velocities resulting in shallower mixed layers, we impose a constant mixed layer depth of 17 m, taken as the mean of all TN368 CTD mixed layer depths from the slope sea (defined where salinity at 10 m is > 34.6 , with the mixed layer depth defined as the depth where the potential density change from the surface value reaches $\Delta\rho = 0.04 \frac{kg}{m^3}$). To account for fluxes from below, we apply a gradient boundary condition at the base of the model for temperature, salinity, nitrate, and silicate. We use a vertically uniform diffusivity $K_z = 10^{-5} m^2 s^{-1}$.

Model surface forcing was downloaded from National Centers for Environmental Prediction – National Center for Atmospheric Research (NCEP/NCAR) July 2019 mean reanalysis fields of surface air temperature, latent heat flux, net longwave radiation, net shortwave radiation, precipitation rate, surface air pressure, and wind stress from 72°W – 69°W and 37°N – 39°N (Kalnay et al., 1996).

We tested the model with a range of maximum vertical velocities based on those likely to occur at the edge of a WCR (Zhang & Partida, 2018). A vertical velocity of $10^{-4} m s^{-1}$ at 150 m results in the best matchups between temperature, salinity, nitrate, and chlorophyll between the model and observations at station SLP. We note that because the biogeochemical parameters and initial biomass concentrations can be tuned to result in smaller model-data differences with a given vertical velocity, our chosen upwelling rate is merely one out of many possible values.

Text S12. Biogeochemical model details and equations.

State variables in the biogeochemical model nitrogen and silicon budgets are phytoplankton biomass (**P**), zooplankton biomass (**Z**), small detritus (**SD**), and large detritus (**LD**). Fluxes include nutrient uptake (**U**), grazing (**G**), excretion, egestion, mortality, aggregation, and remineralization. Phytoplankton growth (= net nutrient uptake, **U**) is the product of the maximum growth rate, light availability, and nutrient availability. Nutrient availability is the availability of the scarcer nutrient (Liebig's law of the minimum). We also incorporate elements of the model developed by Fennel et al.,

(2006) by using a Holling-type formulation for grazing, and the second-order aggregation of phytoplankton and small detritus into large detritus. Our model includes zooplankton assimilation efficiency (β_N), excretion (γ_n), mortality (ζ_d), remineralization (δ), and aggregation (τ) (Table S3).

The governing nitrogen model state equations are as follows (terms in units of μM or $\mu\text{M d}^{-1}$, with parameters as defined in Table S3):

$$\underbrace{\frac{\partial \text{DIN}}{\partial t}}_{\text{dissolved inorganic N}} = \underbrace{-UP_N}_{\text{uptake by phytoplankton}} + \underbrace{+\gamma_n GZ_N}_{\text{zooplankton excretion}} + \underbrace{+\delta SD_N + \delta_{LDN} LD_N}_{\text{remineralization}} \quad (15)$$

$$\underbrace{\frac{\partial P_N}{\partial t}}_{\text{phytoplankton}} = \underbrace{+UP_N}_{\text{uptake}} + \underbrace{-GZ_N}_{\text{grazing}} - \underbrace{\sigma_{P1} P1_N}_{\text{mortality}} - \underbrace{\tau (SD_N + P_N) P_N}_{\text{aggregation}} - \underbrace{w_P \frac{\partial P_N}{\partial z}}_{\text{sinking}} \quad (16)$$

$$\underbrace{\frac{\partial Z_N}{\partial t}}_{\text{zooplankton}} = \underbrace{\beta GZ_N}_{\text{grazing}} - \underbrace{\gamma_n GZ_N}_{\text{excretion}} - \underbrace{\zeta_d Z_N}_{\text{mortality}} \quad (17)$$

$$\underbrace{\frac{\partial SD_N}{\partial t}}_{\text{small detritus}} = \underbrace{+\sigma_P P_N}_{\text{phytoplankton mortality}} - \underbrace{\tau (SD_N + P_N) SD_N}_{\text{aggregation}} - \underbrace{\delta SD_N}_{\text{remineralization}} - \underbrace{w_{SD} \frac{\partial SD_N}{\partial z}}_{\text{sinking}} \quad (18)$$

$$\underbrace{\frac{\partial LD_N}{\partial t}}_{\text{large detritus}} = \underbrace{+G(1-\beta) Z_N}_{\text{zooplankton egestion}} + \underbrace{+\zeta_d Z_N}_{\text{zooplankton mortality}} + \underbrace{+\tau (SD_N + P_N)^2}_{\text{aggregation}} - \underbrace{\delta LD_N}_{\text{remineralization}} - \underbrace{w_{LD} \frac{\partial LD_N}{\partial z}}_{\text{sinking}} \quad (19)$$

The governing silicon model state equations (with silica uptake converted from nitrogen uptake equations using an Si:N ratio, R) are as follows (terms in units of μM or $\mu\text{M d}^{-1}$, with parameter definitions in Table S3):

$$\underbrace{\frac{\partial dSi}{\partial t}}_{\text{silicate}} = \underbrace{-R \times UP_N}_{\text{uptake by phytoplankton}} + \underbrace{+\delta SD_{Si} + \delta LD_{Si}}_{\text{remineralization}} \quad (20)$$

$$\underbrace{\frac{\partial P_{Si}}{\partial t}}_{\text{phytoplankton}} = \underbrace{+R \times UP_N}_{\text{uptake}} - \underbrace{P_{Si} \left(\frac{G Z_N}{P_N} \right)}_{\text{grazing}} - \underbrace{\sigma_P P_{Si}}_{\text{mortality}} - \underbrace{\tau (SD_N + P_N) P_{Si}}_{\text{aggregation}} - \underbrace{w_P \frac{\partial P_{Si}}{\partial z}}_{\text{sinking}} \quad (21)$$

$$\underbrace{\frac{\partial SD_{Si}}{\partial t}}_{\text{small detritus}} = \underbrace{+\sigma_P P_{Si}}_{\text{phytoplankton mortality}} - \underbrace{\tau (SD_N + P_N) SD_{Si}}_{\text{aggregation}} - \underbrace{\delta SD_{Si}}_{\text{reminerlization}} - \underbrace{w_{SD} \frac{\partial SD_{Si}}{\partial z}}_{\text{sinking}} \quad (22)$$

$$\underbrace{\frac{\partial LD_{Si}}{\partial t}}_{\text{large detritus}} = \underbrace{+P_{Si} \left(\frac{G Z_N}{P_N} \right)}_{\text{zooplankton egestion}} + \underbrace{\tau (SD_N + P_N) (SD_{Si} + P_{Si})}_{\text{aggregation}} - \underbrace{\delta LD_{Si}}_{\text{reminerlization}} - \underbrace{w_{LD} \frac{\partial LD_{Si}}{\partial z}}_{\text{sinking}} \quad (23)$$

The biogeochemical model uses the following functions:

Specific growth (uptake) rates for phytoplankton (U):

$$U = P_{max} \times \underbrace{\min \left(\frac{dSi}{k_{Si} + dSi}, \frac{DIN}{k_N + DIN} \right)}_{\text{limitation by most scarce nutrient}} \underbrace{\frac{\alpha I}{\sqrt{\alpha^2 I^2 + P_{max}^2}}}_{\text{light limitation}} \quad (24)$$

Photosynthetically active radiation (I , in $W m^{-2}$):

$$\frac{\partial I}{\partial z} = -k_d I \quad (25)$$

Light attenuation (k_d , in m^{-1}) from seawater ($k_z = 0.04$) and from phytoplankton biomass (k_p) (Fasham et al., 1990):

$$k_d = k_z + k_p P_N \quad (26)$$

Zooplankton grazing rates (Holling-type formulation, Fennel et al., 2006) on phytoplankton (G) in day^{-1} :

$$G = R_{max} \frac{P_N^2}{k_{phy} + P_N^2} \quad (27)$$

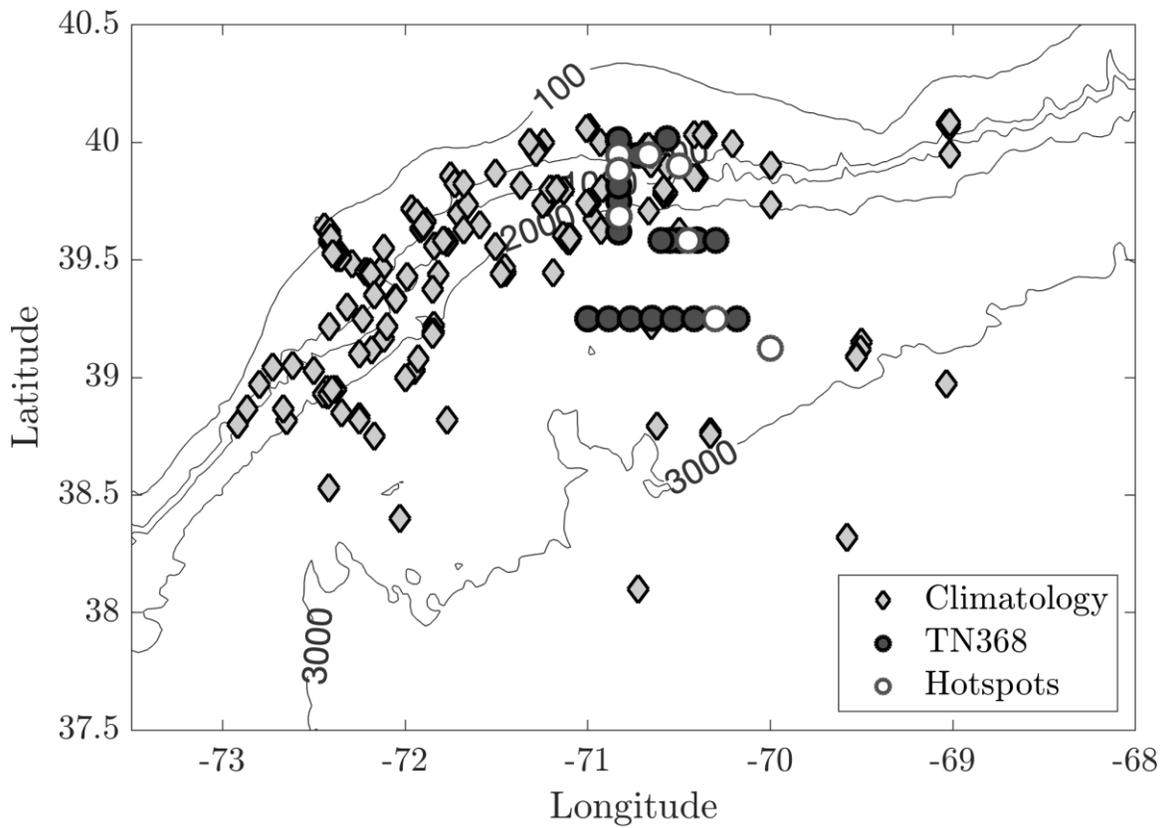


Fig. S1. Locations of June-September slope sea (bottom depth > 200 m) profiles from the MAB climatology (diamonds; Zhang et al., 2013) and TN368 (circles) used for the histogram in Figure 1D. Hotspot profiles are shown by white circles, and all other TN368 profiles are shown as dark gray circles.

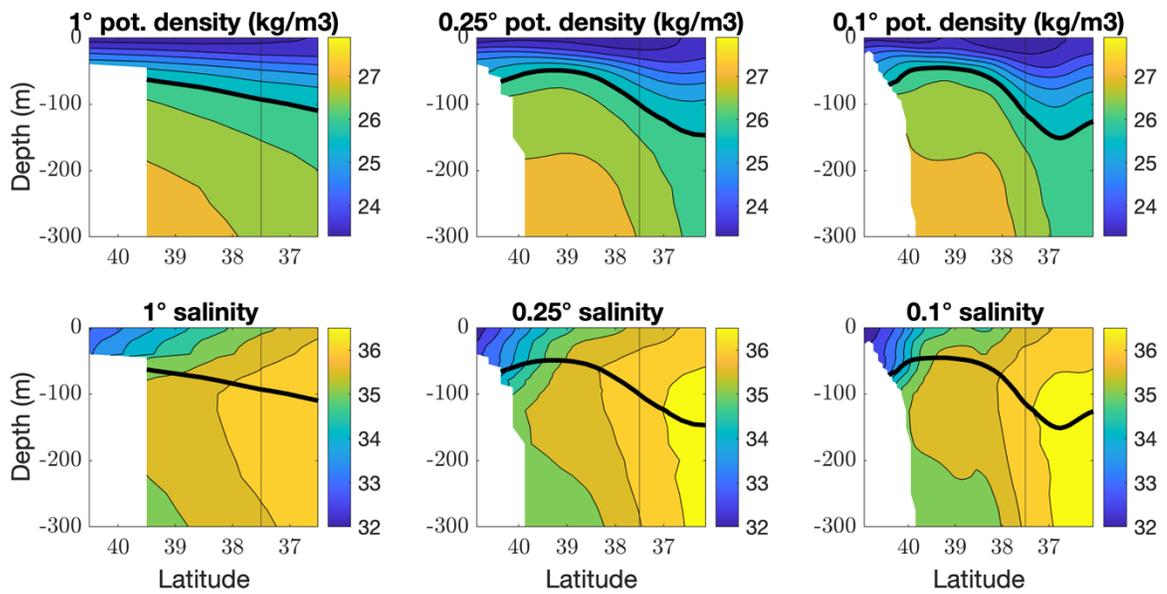


Fig. S2. North-south density and salinity cross-sections from NCEI NWA Summer Climatology at 70°W, for 1°, 0.25°, and 0.1° data product resolutions. Vertical lines at 37.5 show the location of the GS endmember profile, where $T(200\text{ m})$ is closest to 15°C in 1° climatology.

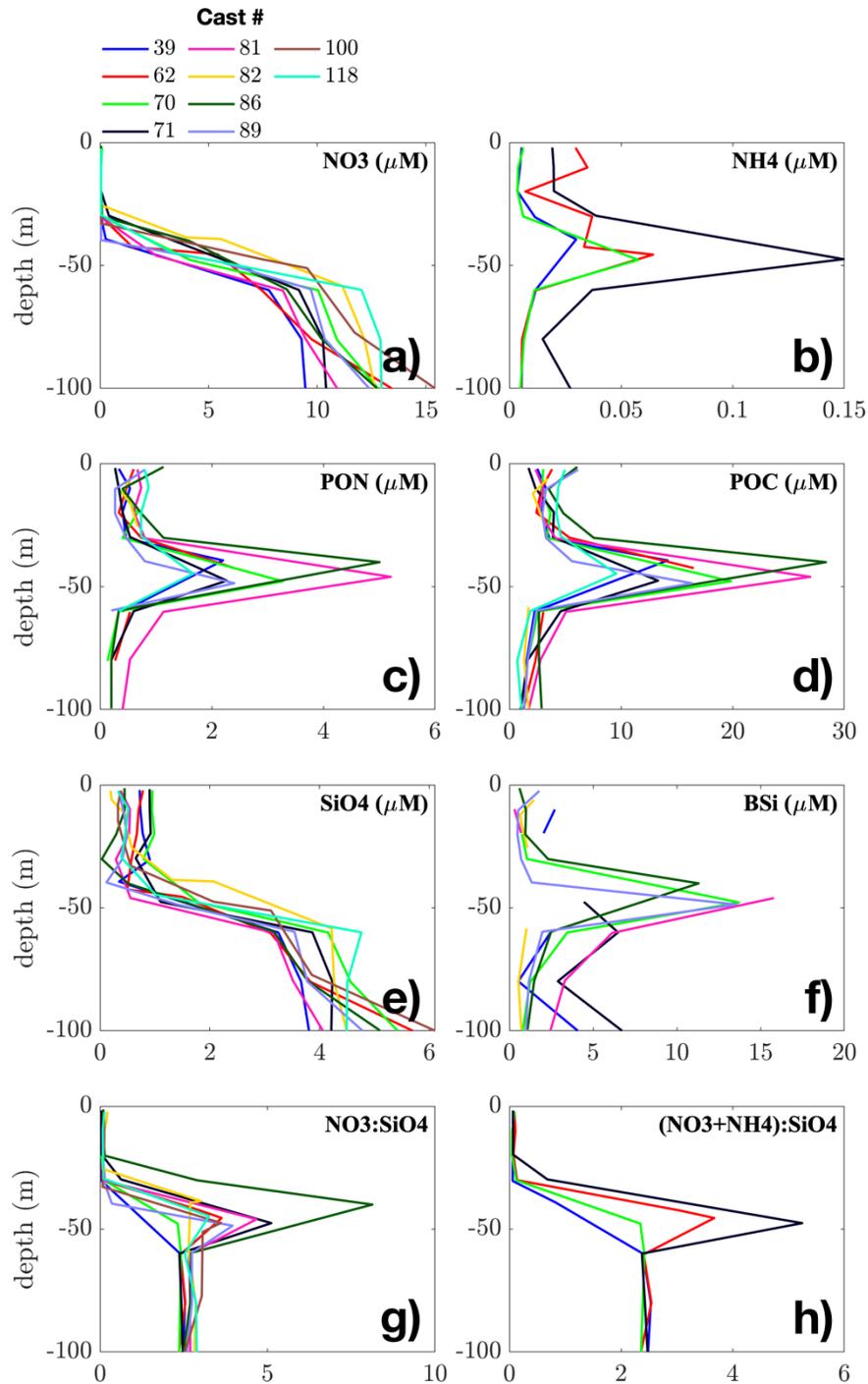


Fig. S3. Bottle file vertical profiles from hotspot CTD casts. Profiles include a) nitrate (NO₃), b) ammonium (NH₄), c) particulate organic nitrogen (PON), d) particulate organic carbon (POC), e) silicate (SiO₄), f) biogenic silica (BSi), g) nitrate to silicate ratios, and h) nitrate + ammonium to silicate ratios. Hotspot CTD casts are defined as casts where the chlorophyll maximum was between 35 and 50 m, the chlorophyll maximum was > 2.5 $\mu\text{g/L}$, and at the chlorophyll maximum salinity > 35.6.

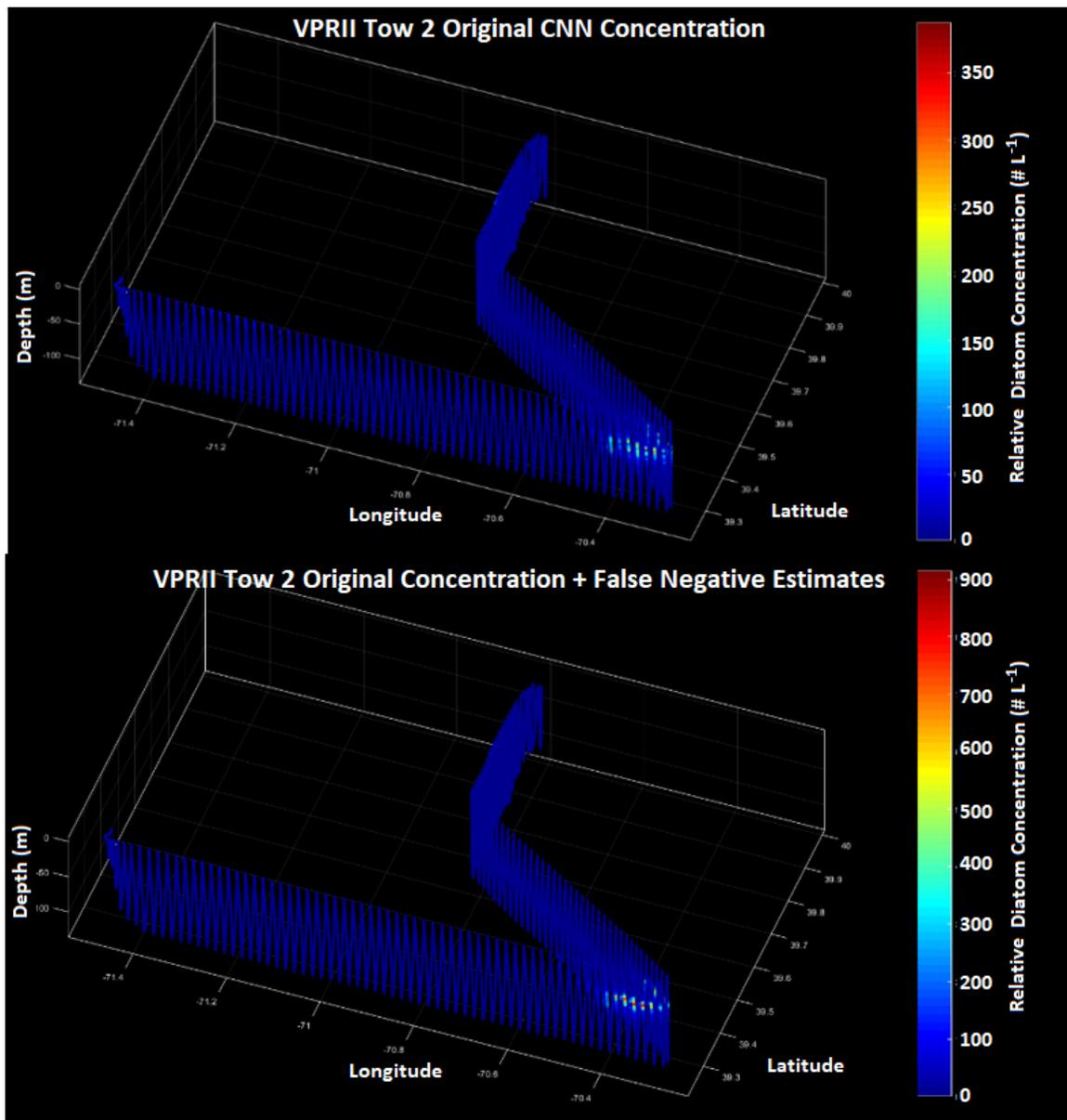


Fig. S4 Towed VPR diatom distributions. Top: original diatom distribution obtained from CNN classification without considering false positives and false negatives. Bottom: diatom distribution corrected for a theoretical maximum diatom concentration (original+false negatives) to determine an upper bound on the diatom hotspot concentration.

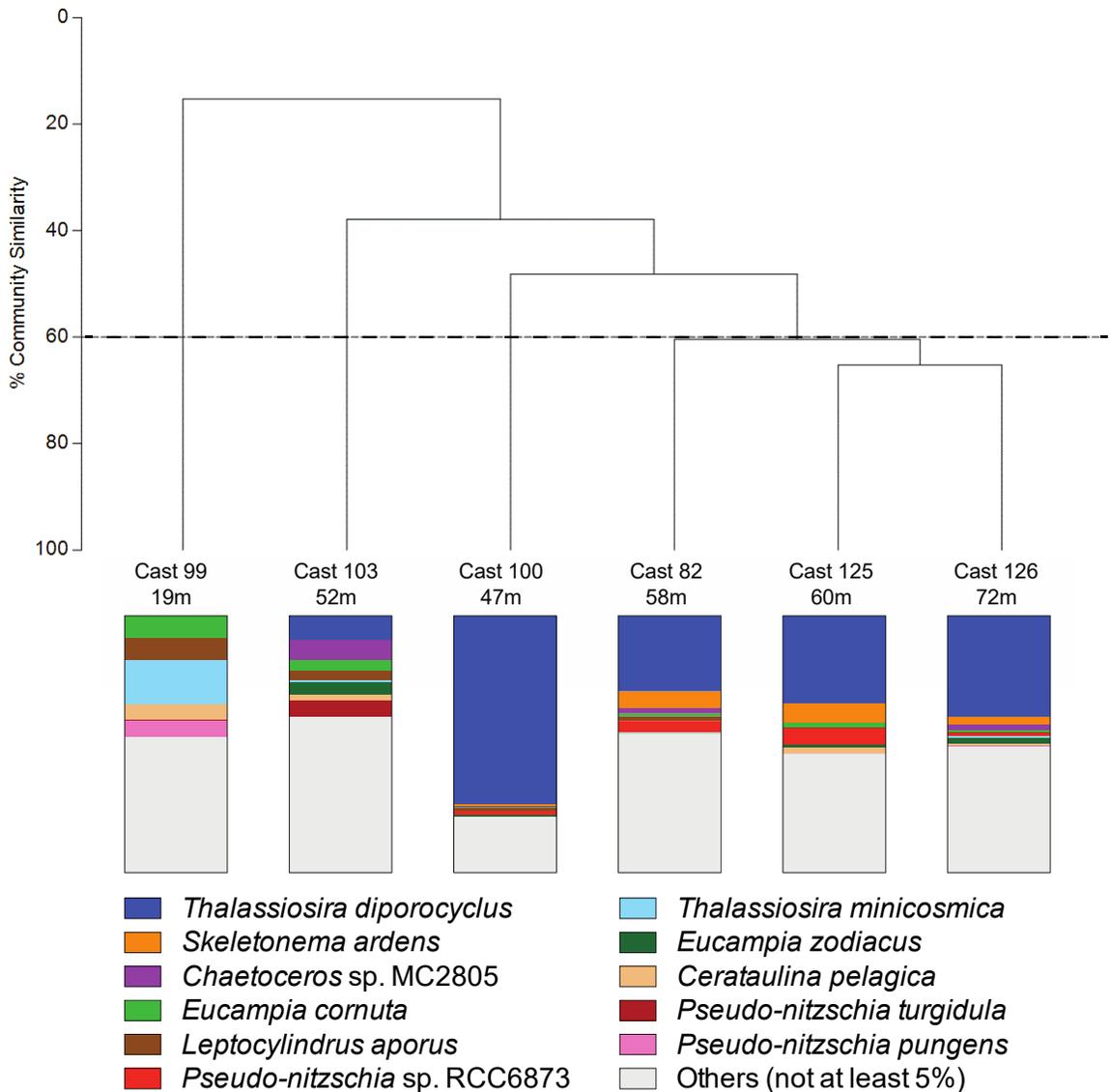


Fig. S5. Bray-Curtis similarity of the diatom 18S sequencing results from samples collected at the chlorophyll maximum (depth indicated below cast number) from a non-hotspot station on the shelf (Cast 99), a non-hotspot station on the slope (Cast 103), and 4 stations associated with the hotspot (Casts 82, 100, 125, and 126). The dashed line corresponds to 60% community similarity between samples. Before Bray-Curtis analysis, raw ASV count data was log-transformed. Below the dendrogram, the relative abundance of diatom species in each sample is shown. For graphing purposes, ASVs that were >99 % similar were grouped. Only diatoms that accounted for at least 5% of sequences at one of the stations are identified. The remaining diatom species are grouped in the “Others (not at least 5%)” category.

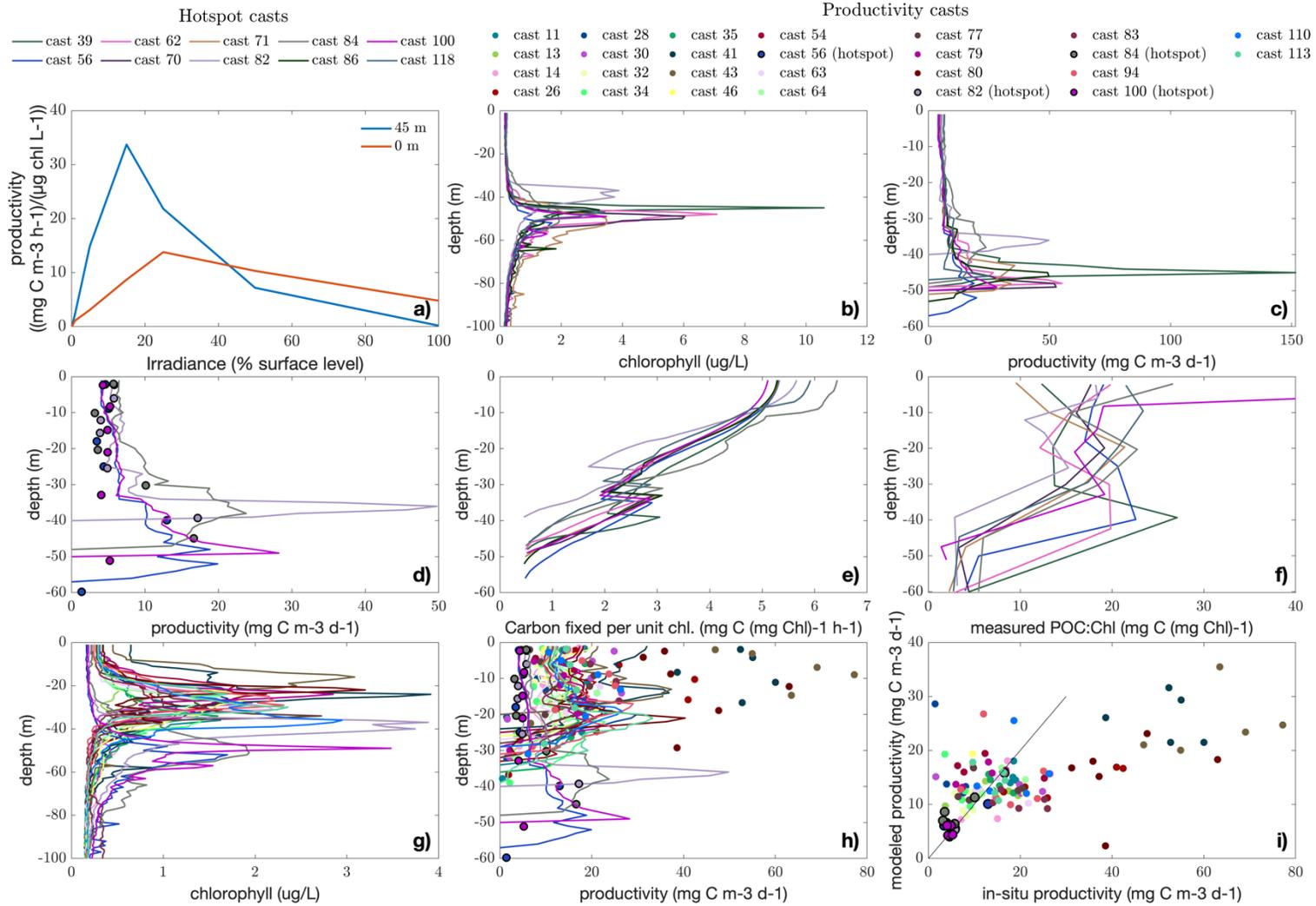


Fig. S6. Primary productivity: photosynthesis-irradiance experiments, from a bio-optical model, and measured *in situ*. a) Photosynthesis-irradiance experiment results from 45 m and 0 m at Station SLP (cast 100), chlorophyll-normalized; b) profiles of

chlorophyll *a* for hotspot casts derived from fluorescence measurements, c) bio-optical modeled productivity, down to the 1.3 $\mu\text{mol photons m}^{-2} \text{d}^{-1}$ isolume (the absolute compensation irradiance, Siegel et al., 2002); d) same as B, but only for casts for which *in situ* rates were measured, with in-situ rates shown as dots; e) profiles of modeled carbon fixed per unit chlorophyll down to the 1.3 $\mu\text{mol photons m}^{-2} \text{d}^{-1}$ isolume, with sharp transitions indicating the 100 $\mu\text{mol photons m}^{-2} \text{d}^{-1}$ isolume below which $E_k = 0.15 \times$ surface PAR ($E_k = 0.25 \times$ surface PAR above, see Text S8); f) ratio of POC (Fig. S3) to chlorophyll, showing high POC at many hotspot chlorophyll maxima, suggesting high concentrations of marine snow; g) profiles of chlorophyll for all primary productivity casts conducted during TN368; h) measured and modeled productivity for all TN368 productivity casts; i) measured vs. modeled productivity for all productivity casts, with black one-to-one line shown. Note the model underestimation of productivity at inner shelf stations 41, 43, and 80 (shown in Fig. S7).

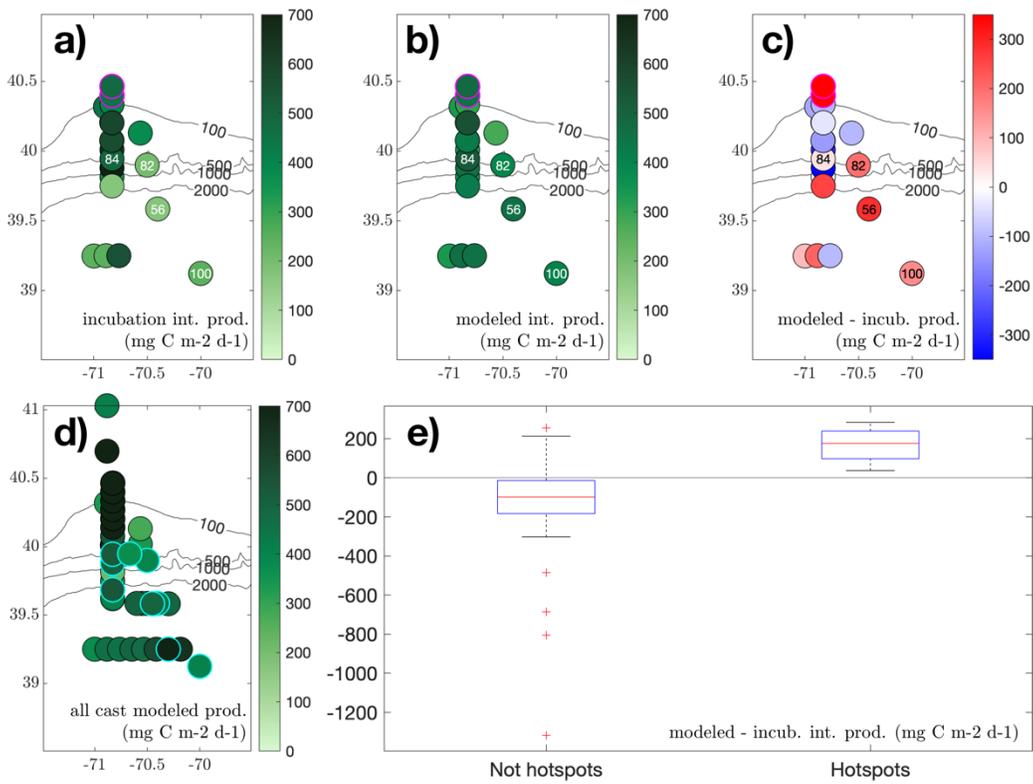


Fig. S7. Modeled and measured integrated primary productivity. A) Integrated primary productivity for all TN368 productivity incubation casts, with inner shelf stations 41, 43, and 80 highlighted in magenta, and hotspot casts labeled; B) modeled integrated primary productivity for the same set of casts; C) the difference between modeled and incubation productivity; D) modeled productivity for all TN368 casts, with hotspot casts highlighted in cyan; E) boxplot of the difference between modeled and incubation productivity, for non-hotspot and hotspot productivity casts.

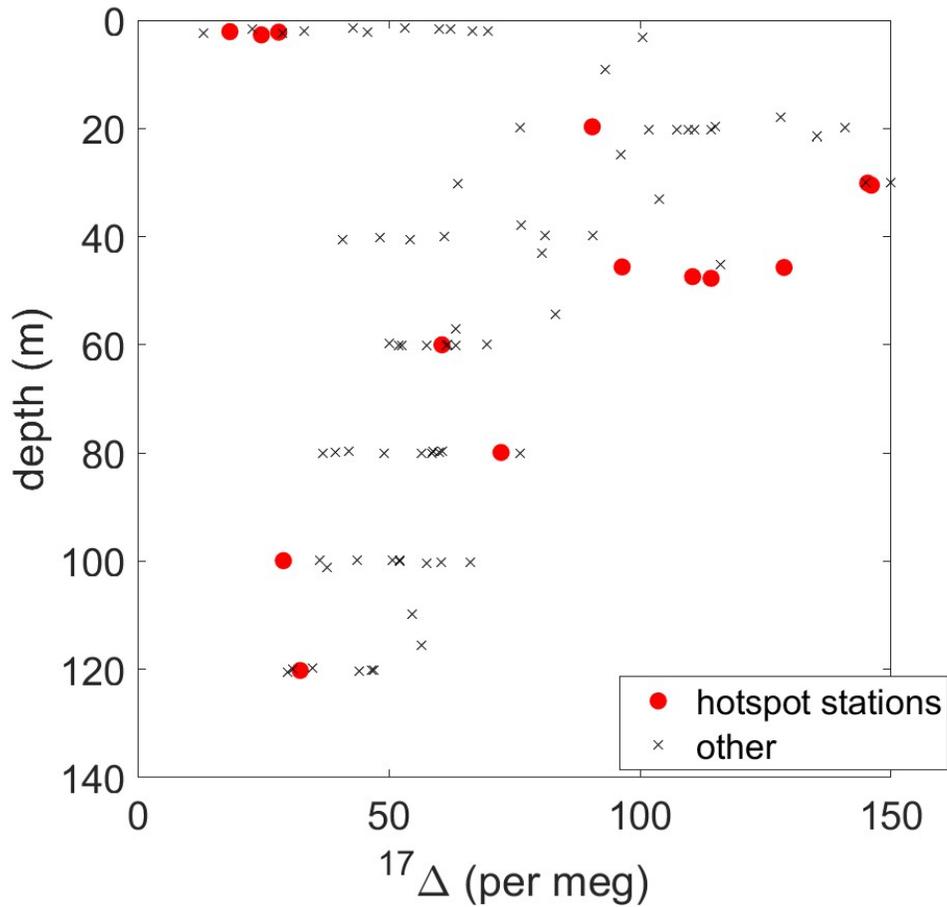


Fig. S8. Triple Oxygen Isotope signature, $^{17}\Delta$ (as defined in Text S9), of samples collected at stations designated as diatom hotspots (red circles) or other, non-hotspot stations (black x). A larger $^{17}\Delta$ means that a larger fraction of the dissolved oxygen present in the sample stems from photosynthesis as compared to atmospheric sources. Note that at the depths of the diatom hotspots (around 50 m), the samples in the hotspot stations have larger $^{17}\Delta$ and thus a higher fraction of photosynthetic oxygen than the most of the non-hotspot stations.

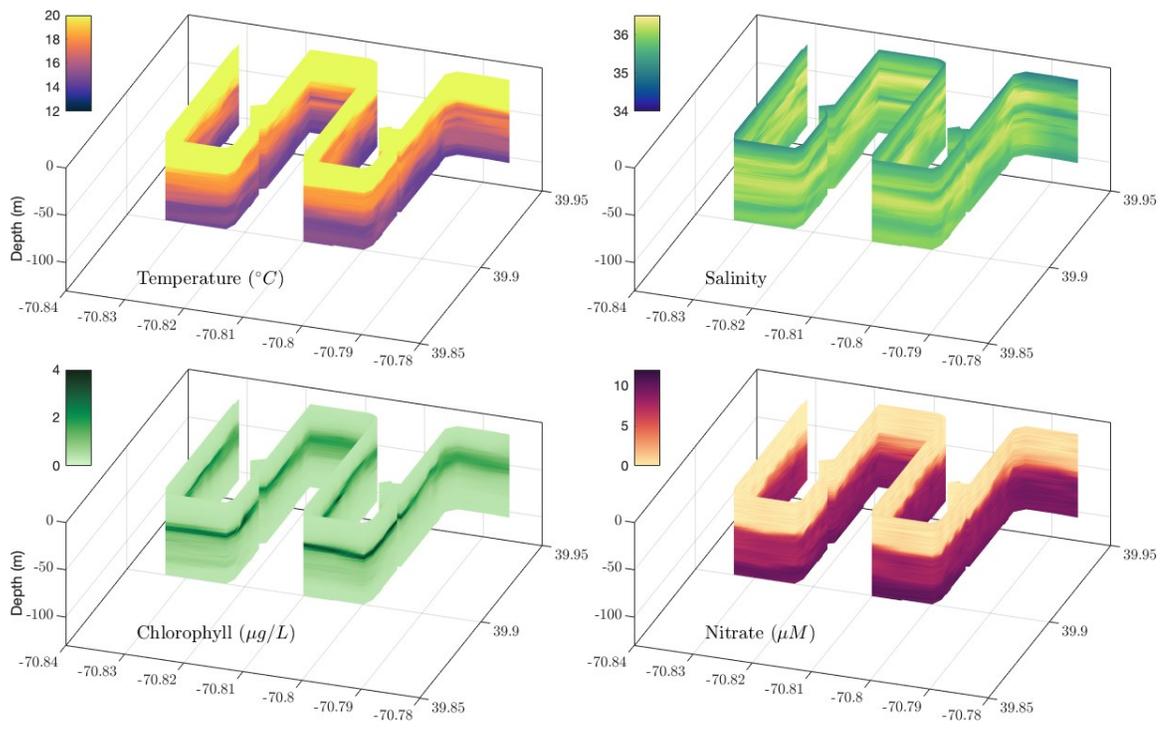


Fig. S9. REMUS 600 Mission 4A measurements of salinity, temperature, chlorophyll, and nitrate.

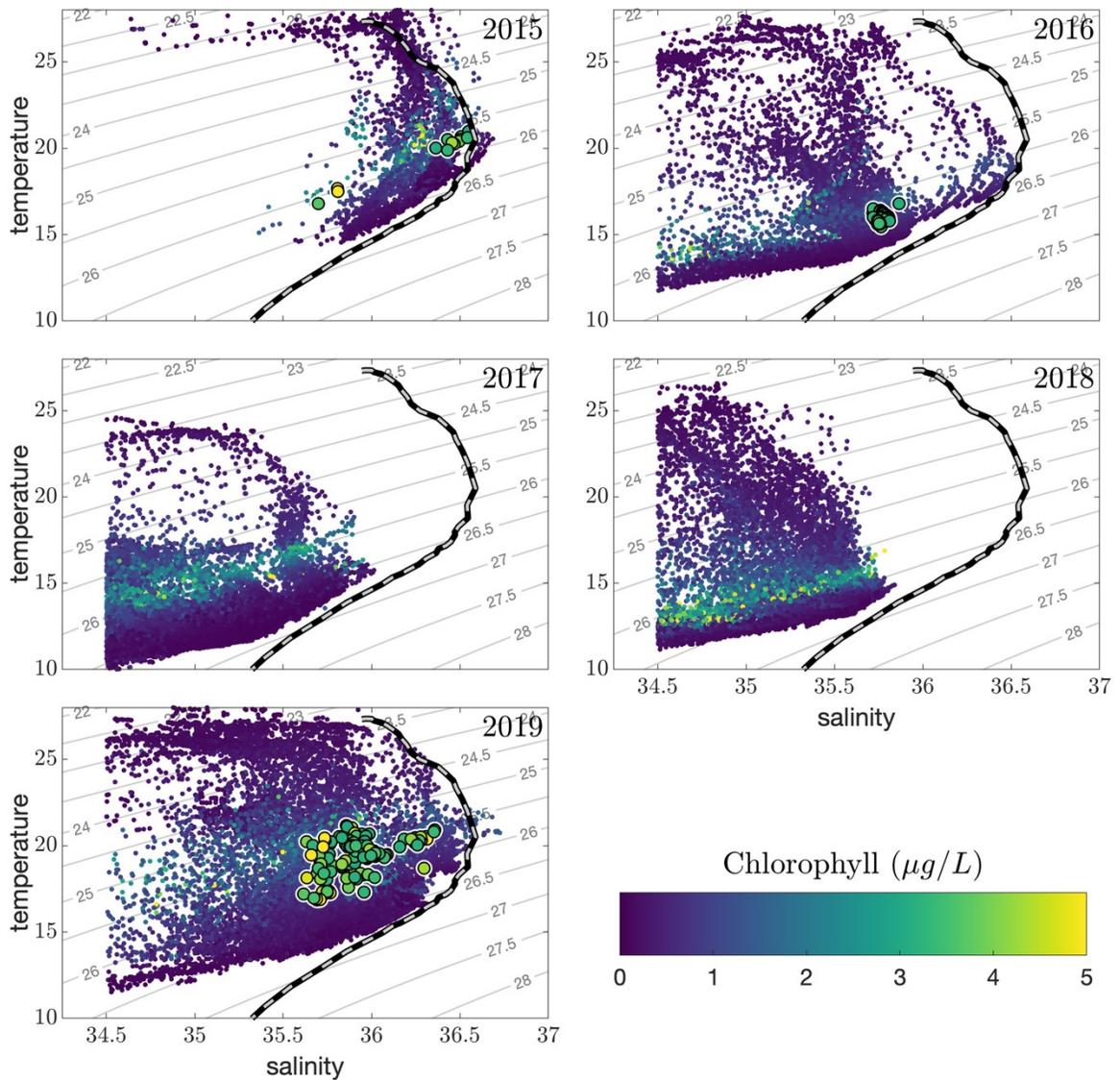


Fig. S10. OOI summer (June-August) slope sea glider measurements of temperature, salinity, and chlorophyll from 2015 – 2019. Large points show where 1) salinity > 35.6, 2) chlorophyll > 3 $\mu\text{g/L}$, and 3) depth of the chlorophyll max between 40 and 60 m. Dashed black/white lines show Gulf Stream T-S profile from 69.5 E, 37 N from 1/10° summer objectively analyzed climatological means from the National Centers for Environmental Information (NCEI) Northwest Atlantic Regional Ocean Climatology (Seidov, 2016).

Table S1. Classification categories, source of their training sets, and accuracy of the CNN classifications. The F1 score was calculated using 20% of the original CNN training withheld for evaluation purposes. Categories presented in this study are marked with an asterisk. AR29 is the March 2018 cruise of the R/V *Neil Armstrong* to the MAB, and TN368 is the July 2019 cruise of the R/V *Thomas G. Thompson*.

Category:	Training ROI Source:	F1 Score:
Bloom Conditions *	AR29	99%
Bubbles (Air) *	TN368	79%
<i>Ceratium</i> sp.	TN368	44%
Chaetognaths	TN368	81%
<i>Calanus finmarchicus</i>	TN368	52%
<i>Chaetoceros</i> sp. (combined with diatoms) *	TN368	12%
Copepods (small, unable to identify)	TN368	78%
<i>Coscinodiscus</i> sp. (combined with diatoms) *	TN368	86%
Decapod Larvae	TN368	57%
Diatoms (unidentified diatom chains) *	TN368	86%
Diatoms (Bloom, combined with diatoms) *	TN368	80%
Diatoms (Coils, mostly <i>Guinardia</i> sp., combined with diatoms) *	TN368	82%
Echinoderm Larvae	AR29	97%
Fecal Strings and Pellets	TN368	60%
<i>Foraminifera</i>	TN368	94%
Gelatinous Zooplankton	TN368	63%
Krill	TN368	0%
Krill Nauplii	TN368	78%
Larvaceans	TN368	79%
Marine Snow *	TN368	98%
<i>Oithona</i> sp. with visible egg sacs	TN368	25%
Out-of-Focus *	TN368	87%
<i>Phaeocystida</i> Protozoans	TN368	100%
<i>Pseudocalanus</i> sp. with visible egg sacs	TN368	67%
Pteropods	TN368	53%
Radiolarians	TN368	81%
Spherical Colonial Plankton (<i>Phaeocystis pouchetii</i>)	AR29	97%
<i>Thalassiosira diporocyclus</i> (combined with diatoms) *	TN368	86%
<i>Trichodesmium</i> sp.	TN368	62%
Unknown *	N/A	N/A

Table S2. Bio-optical modeled integrated productivity sensitivities and comparison with estimates of integrated productivity made with measurements made *in situ*

Cast #	Original modeled integrated productivity (g C m ⁻² d ⁻¹)	Upper k _d perturb. (g C m ⁻² d ⁻¹)	Chl a perturb. (g C m ⁻² d ⁻¹)	Absolute compensation irradi. perturb. (g C m ⁻² d ⁻¹)	Maximum perturb. (27 combinations in total) (g C m ⁻² d ⁻¹)	Uncertainty sum in quadrature (g C m ⁻² d ⁻¹)	Estimated prod. from <i>in situ</i> measurements (g C m ⁻² d ⁻¹)	Modeled productivity : measured productivity
39	0.7	± 0.2	< ± 0.05	< ± 0.05	± 0.3	± 0.2	-	-
56	0.5	± 0.2	< ± 0.05	< ± 0.05	± 0.2	± 0.2	0.2	2.6 ± 0.9
62	0.5	± 0.2	± 0.1	< ± 0.05	± 0.3	± 0.3	-	-
70	0.5	± 0.3	± 0.1	< ± 0.05	± 0.4	± 0.3	-	-
71	0.5	± 0.2	< ± 0.05	< ± 0.05	± 0.3	± 0.2	-	-
82	0.4	± 0.1	< ± 0.05	< ± 0.05	± 0.2	± 0.1	0.2	2.0 ± 1.0
84	0.5	± 0.1	< ± 0.05	< ± 0.05	± 0.1	± 0.1	0.5	1.1 ± 0.3
86	0.5	± 0.2	< ± 0.05	< ± 0.05	± 0.3	± 0.2	-	-
100	0.4	± 0.1	< ± 0.05	< ± 0.05	± 0.2	± 0.1	0.2	1.7 ± 0.5
118	0.3	± 0.1	< ± 0.05	< ± 0.05	± 0.2	± 0.2	-	-

Table S3. Parameters used in the biogeochemical model.

Symbol	Description	Value	Units
P_{max}	Phytoplankton max growth rate	3.0 ^a	day ⁻¹
α	Phytoplankton initial slope P-E curve	0.1 ^a	(W m ⁻²) ⁻¹ day ⁻¹
k_N	Phytoplankton nitrate uptake half-saturation	0.2	mmol N m ⁻³
k_{Si}	Phytoplankton silicate uptake half-saturation	0.2	mmol Si m ⁻³
R_{max}	Max grazing rate on phytoplankton	0.5	day ⁻¹
k_{phy}	Grazing on phytoplankton half-saturation	2.0	(mmol N m ⁻³) ²
k_p	Phytoplankton light attenuation coefficient	0.08 ^b	(mmol N) ⁻¹
R	Si:N ratio for phytoplankton uptake	$0.13 \times \frac{106}{16}$ ^c	mmol Si (mmol N) ⁻¹
γ_n	Excretion rate of nitrogen by zooplankton	0.1	dimensionless
σ_P	Mortality of phytoplankton	0.2	day ⁻¹
δ	Remineralization rate	0.05	day ⁻¹
ζ_d	Mortality of zooplankton	0.1	day ⁻¹
τ	Aggregation rate of phytoplankton and small detritus	0.01	(mmol N m ⁻³) ⁻¹ day ⁻¹
w_P	Sinking rate of phytoplankton	2.5	m day ⁻¹
w_{SD}	Sinking rate of small detritus	2.5	m day ⁻¹
w_{LD}	Sinking rate of large detritus	5.0	m day ⁻¹
β	Zooplankton assimilation efficiency	0.75	dimensionless

a. This study's bio-optical model.

b. Derived from TN368 CTD casts (see Text S10) and converted to (mmol N)⁻¹ by assuming a POC:Chl ratio of 50 and a C:N ratio of 106:16.

c. Brzezinski, (2004)

Dataset S1 (separate excel file). Details are provided for the read counts of ASVs in each of the samples used for DNA sequencing (Text S6).

References

- Altschul, S. F., Gish, W., Miller, W., Myers, E. W., & Lipman, D. J. (1990). Basic local alignment search tool. *Journal of Molecular Biology*, *215*(3), 403–410. Retrieved from [https://doi.org/10.1016/S0022-2836\(05\)80360-2](https://doi.org/10.1016/S0022-2836(05)80360-2)
- Behrenfeld, M. J., & Falkowski, P. G. (1997a). A consumer's guide to phytoplankton primary productivity models. *Limnology and Oceanography*, *42*(7), 1479–1491. <https://doi.org/10.4319/lo.1997.42.7.1479>
- Behrenfeld, M. J., & Falkowski, P. G. (1997b). Photosynthetic rates derived from satellite-based chlorophyll concentration. *Limnology and Oceanography*, *42*(1), 1–20. <https://doi.org/10.4319/lo.1997.42.1.0001>
- Bray, J. R., & Curtis, J. T. (1957). An Ordination of the Upland Forest Communities of Southern Wisconsin. *Ecological Monographs*, *27*(4), 325–349. <https://doi.org/10.2307/1942268>
- Brzezinski, M. A. (2004). The Si:C:N Ratio of Marine Diatoms: Interspecific Variability and the Effect of some Environmental Variables. *Journal of Phycology*, *21*(3), 347–357. <https://doi.org/10.1111/j.0022-3646.1985.00347.x>
- Callahan, B. J., McMurdie, P. J., Rosen, M. J., Han, A. W., Johnson, A. J. A., & Holmes, S. P. (2016). DADA2: High-resolution sample inference from Illumina amplicon data. *Nature Methods*, *13*(7), 581–583. <https://doi.org/10.1038/nmeth.3869>
- Chappell, P. D., Virginia Armbrust, E., Barbeau, K. A., Bundy, R. M., Moffett, J. W., Vedamati, J., & Jenkins, B. D. (2019). Patterns of diatom diversity correlate with dissolved trace metal concentrations and longitudinal position in the northeast Pacific coastal–offshore transition zone. *Marine Ecology Progress Series*, *609*, 69–86. <https://doi.org/10.3354/meps12810>
- Clarke, K. R., & Gorley, R. N. (2018). Getting started with PRIMER v7. *Primer-E*, (1), 20. <https://doi.org/10.22201/ib.20078706e.2018.3.2409>
- Davis, C. S., Thwaites, F. T., Gallager, S. M., Hu, Q., Naiman, M., Sutton, T., et al. (2005). A three-axis fast-tow digital Video Plankton Recorder for rapid surveys of plankton taxa and hydrography. *Limnology and Oceanography: Methods*, *3*, 59–74. <https://doi.org/10.4319/lom.2005.3.59>
- Fasham, M. J. R., Ducklow, H. W., & McKelvie, S. M. (1990). A nitrogen-based model of plankton dynamics in the ocean mixed layer. *Journal of Marine Research*, *48*(3), 591–639.
- Fennel, K., Wilkin, J., Levin, J., Moisan, J., O'Reilly, J., & Haidvogel, D. (2006). Nitrogen cycling in the Middle Atlantic Bight: Results from a three-dimensional model and implications for the North Atlantic nitrogen budget. *Global Biogeochemical Cycles*, *20*(3), 1–14. <https://doi.org/10.1029/2005GB002456>
- Friedrichs, M. A. M., Carr, M. E., Barber, R. T., Scardi, M., Antoine, D., Armstrong, R. A., et al. (2009). Assessing the uncertainties of model estimates of primary productivity in

- the tropical Pacific Ocean. *Journal of Marine Systems*, 76(1–2), 113–133.
<https://doi.org/10.1016/j.jmarsys.2008.05.010>
- Garcia, H., Weathers, K. W., Paver, C. R., Smolyar, I., Boyer, T. P., Locarnini, R. A., et al. (2019). World Ocean Atlas 2018. Vol. 4: Dissolved Inorganic Nutrients (phosphate, nitrate and nitrate+nitrite, silicate). *NOAA Atlas NESDIS 82*, 4(July), 35.
- González, P., Castaño, A., Peacock, E. E., Díez, J., del Coz, J. J., & Sosik, H. M. (2019). Automatic plankton quantification using deep features. *Journal of Plankton Research*, 41(4), 449–463. <https://doi.org/10.1093/plankt/fbz023>
- Hendricks, M. B., Bender, M. L., & Barnett, B. A. (2004). Net and gross O₂ production in the southern ocean from measurements of biological O₂ saturation and its triple isotope composition. *Deep-Sea Research Part I: Oceanographic Research Papers*, 51(11), 1541–1561. <https://doi.org/10.1016/j.dsr.2004.06.006>
- Holmes, R. M., Aminot, A., Kérouel, R., Hooker, B. A., & Peterson, B. J. (1999). A simple and precise method for measuring ammonium in marine and freshwater ecosystems. *Canadian Journal of Fisheries and Aquatic Sciences*, 56(10), 1801–1808.
<https://doi.org/10.1139/f99-128>
- Jassby, A. D., & Platt, T. (1976). Mathematical formulation of the relationship between photosynthesis and light for phytoplankton. *Limnology and Oceanography*, 21(4), 540–547. <https://doi.org/10.4319/lo.1976.21.4.0540>
- Joyce, T. M., Deser, C., & Spall, M. A. (2000). The relation between decadal variability of subtropical mode water and the North Atlantic Oscillation. *Journal of Climate*, 13(14), 2550–2569. [https://doi.org/10.1175/1520-0442\(2000\)013<2550:TRBDVO>2.0.CO;2](https://doi.org/10.1175/1520-0442(2000)013<2550:TRBDVO>2.0.CO;2)
- Juranek, L. W., & Quay, P. D. (2013). Using Triple Isotopes of Dissolved Oxygen to Evaluate Global Marine Productivity. *Annual Review of Marine Science*, 5(1), 503–524.
<https://doi.org/10.1146/annurev-marine-121211-172430>
- Kaiser, J. (2011). Technical note: Consistent calculation of aquatic gross production from oxygen triple isotope measurements. *Biogeosciences*, 8(7), 1793–1811.
<https://doi.org/10.5194/bg-8-1793-2011>
- Kalnay, E., Kanamitsu, M., Kistler, R., Collins, W., Deaven, D., Gandin, L., et al. (1996). The NCEP/NCAR 40-year reanalysis project. *Bulletin of the American Meteorological Society*, 77(3), 437–471.
- Luz, B., & Barkan, E. (2000). Assessment of oceanic productivity with the triple-isotope composition of dissolved oxygen. *Science*, 288(5473), 2028–2031.
<https://doi.org/10.1126/science.288.5473.2028>
- Luz, B., & Barkan, E. (2005). The isotopic ratios ¹⁷O/¹⁶O and ¹⁸O/¹⁶O in molecular oxygen and their significance in biogeochemistry. *Geochimica et Cosmochimica Acta*, 69(5), 1099–1110. <https://doi.org/10.1016/j.gca.2004.09.001>
- Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.Journal*, 17(1), 10. <https://doi.org/10.14806/ej.17.1.200>
- Moberg, E. A., & Sosik, H. M. (2012). Distance maps to estimate cell volume from two-dimensional plankton images. *Limnology and Oceanography: Methods*, 10(4), 278–288. <https://doi.org/10.4319/lom.2012.10.278>

- Morel, A. (1974). Optical properties of pure water and pure sea water. In *Optical Aspects of Oceanography* (pp. 1–24).
- Platt, T., & Jassby, A. D. (1976). The Relationship Between Photosynthesis and Light for Natural Assemblages of Coastal Marine Phytoplankton. *Journal of Phycology*, 12(4), 421–430. <https://doi.org/10.1111/j.1529-8817.1976.tb02866.x>
- Powell, T. M., Lewis, C. V. W. W., Curchitser, E. N., Haidvogel, D. B., Hermann, A. J., & Dobbins, E. L. (2006). Results from a three-dimensional, nested biological-physical model of the California Current System and comparisons with statistics from satellite imagery. *Journal of Geophysical Research: Oceans*, 111(7), 1–14. <https://doi.org/10.1029/2004JC002506>
- Prokopenko, M. G., Pauluis, O. M., Granger, J., & Yeung, L. Y. (2011). Exact evaluation of gross photosynthetic production from the oxygen triple-isotope composition of O₂: Implications for the net-to-gross primary production ratios. *Geophysical Research Letters*, 38(14), 1–5. <https://doi.org/10.1029/2011GL047652>
- Redfield, A. C. (1963). The influence of organisms on the composition of sea-water. *The Sea*, 26–77.
- Seidov, D. (2016). Northwest Atlantic Regional Climatology (NCEI Accession 0155889), 1–56. <https://doi.org/10.7289/V5RF5S2Q>
- Siegel, D. A., Doney, S. C., & Yoder, J. A. (2002). The North Atlantic spring phytoplankton bloom and Sverdrup's critical depth hypothesis. *Science*, 296(5568), 730–3. <https://doi.org/10.1126/science.1069174>
- Sosik, Heidi M., & Olson, R. J. (2007). Automated taxonomic classification of phytoplankton sampled with imaging-in-flow cytometry. *Limnology and Oceanography: Methods*, 5(6), 204–216. <https://doi.org/10.4319/lom.2007.5.204>
- Sosik, H.M., Peacock, E., & Santos, M. (2020). Abundance and biovolume of taxonomically-resolved phytoplankton and microzooplankton imaged continuously underway with an Imaging FlowCytobot along the NES-LTER Transect in winter 2018 ver 1. <https://doi.org/https://doi.org/10.6073/pasta/74775c4af51c237f2a20e4a8c011bc53>
- Stanley, R. H.R., Doney, S. C., Jenkins, W. J., & Lott, D. E. (2012). Apparent oxygen utilization rates calculated from tritium and helium-3 profiles at the Bermuda Atlantic Time-series Study site. *Biogeosciences*, 9(6), 1969–1983. <https://doi.org/10.5194/bg-9-1969-2012>
- Stanley, Rachel H.R., Kirkpatrick, J. B., Cassar, N., Barnett, B. A., & Bender, M. L. (2010). Net community production and gross primary production rates in the western equatorial Pacific. *Global Biogeochemical Cycles*, 24(4), 1–17. <https://doi.org/10.1029/2009GB003651>
- Zhang, W. G., & Partida, J. (2018). Frontal subduction of the Mid-Atlantic Bight Shelf Water at the onshore edge of a warm-core ring. *Journal of Geophysical Research: Oceans*, 123(11), 7795–7818. <https://doi.org/10.1029/2018JC013794>
- Zhang, W. G., McGillicuddy, D. J., & Gawarkiewicz, G. G. (2013). Is biological productivity enhanced at the New England shelfbreak front? *Journal of Geophysical Research: Oceans*, 118(1), 517–535. <https://doi.org/10.1002/jgrc.20068>

Zimmermann, J., Jahn, R., & Gemeinholzer, B. (2011). Barcoding diatoms: evaluation of the V4 subregion on the 18S rRNA gene, including new primers and protocols. *Organisms Diversity & Evolution*, 11(3), 173–192. <https://doi.org/10.1007/s13127-011-0050-6>