2010

Challenges of modeling depth-integrated marine primary productivity over multiple decades: A case study at BATS and HOT

VS Saba
Virginia Institute of Marine Science

MAM Friedrichs
Virginia Institute of Marine Science

ME Carr

D Antoine

RA Armstrong

See next page for additional authors

Challenges of modeling depth-integrated marine primary productivity over multiple decades: A case study at BATS and HOT

Vincent S. Saba,[1,2] Marjorie A. M. Friedrichs,[1] Mary-Elena Carr,[3] David Antoine,[4] Robert A. Armstrong,[5] Ichio Asanuma,[6] Olivier Aumont,[7] Nicholas R. Bates,[8] Michael J. Behrenfeld,[9] Val Bennington,[10] Laurent Bopp,[11] Jorn Bruggeman,[12] Erik T. Buitenhuis,[13] Matthew J. Church,[14] Aurea M. Ciotti,[15] Scott C. Doney,[16] Mark Dowell,[17] John Dunne,[18] Stephanie Dutkiewicz,[19] Watson Gregg,[20] Nicolas Hoepffner,[17] Kimberly J. W. Hyde,[21] Joji Ishizaka,[22] Takahiko Kameda,[23] David M. Karl,[14] Ivan Lima,[16] Michael W. Lomas,[8] John Marra,[24] Galen A. McKinley,[10] Frédéric Mélin,[17] J. Keith Moore,[25] André Morel,[4] John O'Reilly,[21] Baris Salihoglu,[26] Michele Scardi,[27] Tim J. Smyth,[28] Shilin Tang,[29] Jerry Tjiputra,[30] Julia Uitz,[31] Marcello Vichi,[32] Kirk Waters,[33] Toby K. Westberry,[9] and Andrew Yool[34]

Received 17 August 2009; revised 9 April 2010; accepted 27 April 2010; published 15 September 2010.

[1] The performance of 36 models (22 ocean color models and 14 biogeochemical ocean circulation models (BOGCMs)) that estimate depth-integrated marine net primary productivity (NPP) was assessed by comparing their output to in situ ^{14}C data at the Bermuda Atlantic Time series Study (BATS) and the Hawaii Ocean Time series (HOT) over nearly two decades. Specifically, skill was assessed based on the models' ability to estimate the observed mean, variability, and trends of NPP. At both sites, more than 90% of the models underestimated mean NPP, with the average bias of the BOGCMs being nearly twice that of the ocean color models. However, the difference in overall skill

[1]Virginia Institute of Marine Science, College of William and Mary, Gloucester Point, Virginia, USA.
[2]Now at Atmospheric and Oceanic Sciences Program, Princeton University, Princeton, New Jersey, USA.
[3]Columbia Climate Center, Earth Institute, Columbia University, New York, New York, USA.
[4]Laboratoire d'Océanographie de Villefranche, UMR 7093, Université Pierre et Marie Curie, Paris 06, CNRS, Villefranche-sur-Mer, France.
[5]School of Marine and Atmospheric Sciences, State University of New York at Stony Brook, Stony Brook, New York, USA.
[6]Tokyo University of Information Sciences, Chiba, Japan.
[7]Laboratoire d'Océanographie: Expérimentation et Approche Numérique, IPSL, UPMC, IRD, CNRS, Plouzané, France.
[8]Bermuda Institute of Ocean Sciences, St. George's, Bermuda.
[9]Department of Botany and Plant Pathology, Oregon State University, Corvallis, Oregon, USA.
[10]Department of Atmospheric and Ocean Sciences, University of Wisconsin-Madison, Madison, Wisconsin, USA.
[11]Laboratoire des Sciences du Climat et de l'Environnement, IPSL, CEA, UVSQ, CNRS, Gif-sur-Yvette, France.
[12]Department of Theoretical Biology, Faculty of Earth and Life Sciences, Vrije University of Amsterdam, Amsterdam, Netherlands.
[13]Laboratory for Global Marine and Atmospheric Chemistry, School of Environmental Sciences, University of East Anglia, Norwich, UK.
[14]Department of Oceanography, School of Ocean and Earth Science and Technology, University of Hawai'i at Mãnoa, Honolulu, Hawaii, USA.
[15]Campus Experimental do Litoral Paulista, UNESP, São Vicente, Brazil.
[16]Department of Marine Chemistry and Geochemistry, Woods Hole Oceanographic Institution, Woods Hole, Massachusetts, USA.
[17]Joint Research Centre, European Commission, Ispra, Italy.
[18]Geophysical Fluid Dynamics Laboratory, NOAA, Princeton, New Jersey, USA.
[19]Earth, Atmospheric and Planetary Sciences, Massachusetts Institute of Technology, Cambridge, Massachusetts, USA.
[20]NASA Global Modeling and Assimilation Office, Goddard Space Flight Center, Greenbelt, Maryland, USA.
[21]National Marine Fisheries Services Narragansett Laboratory, NOAA, Narragansett, Rhode Island, USA.
[22]Hydrospheric Atmospheric Research Center, Nagoya University, Nagoya, Japan.
[23]Group of Oceanography, National Research Institute of Far Seas Fisheries, Shizuoka, Japan.
[24]Geology Department, Brooklyn College of the City University of New York, Brooklyn, New York, USA.
[25]Department of Earth System Science, University of California, Irvine, California, USA.
[26]Institute of Marine Sciences, Middle East Technical University, Erdemli, Turkey.
[27]Department of Biology, University of Rome 'Tor Vergata', Rome, Italy.
[28]Plymouth Marine Laboratory, Plymouth, UK.
[29]Freshwater Institute, Fisheries and Oceans Canada, Winnipeg, Manitoba, Canada.
[30]Geophysical Institute, University of Bergen, Bergen, Norway.
[31]Scripps Institution of Oceanography, University of California, San Diego, La Jolla, California, USA.
[32]Centro Euro-Mediterraneo per i Cambiamenti Climatici, Instituto Nazionale di Geofisica e Vulcanologia, Bologna, Italy.
[33]NOAA Coastal Services Center, Charleston, South Carolina, USA.
[34]National Oceanography Centre, Southampton, Southampton, UK.

0886-6236/10/2009GB003655

between the best BOGCM and the best ocean color model at each site was not significant. Between 1989 and 2007, in situ NPP at BATS and HOT increased by an average of nearly 2% per year and was positively correlated to the North Pacific Gyre Oscillation index. The majority of ocean color models produced in situ NPP trends that were closer to the observed trends when chlorophyll-*a* was derived from high-performance liquid chromatography (HPLC), rather than fluorometric or SeaWiFS data. However, this was a function of time such that average trend magnitude was more accurately estimated over longer time periods. Among BOGCMs, only two individual models successfully produced an increasing NPP trend (one model at each site). We caution against the use of models to assess multiannual changes in NPP over short time periods. Ocean color model estimates of NPP trends could improve if more high quality HPLC chlorophyll-*a* time series were available.

Citation: Saba, V. S., et al. (2010), Challenges of modeling depth-integrated marine primary productivity over multiple decades: A case study at BATS and HOT, *Global Biogeochem. Cycles*, *24*, GB3020, doi:10.1029/2009GB003655.

1. Introduction

[2] Primary productivity is an essential component of both terrestrial and aquatic ecosystems. Primary producers are at the base of food webs and thus drive ecosystem dynamics through bottom-up forcing. Global biogeochemical cycles of major elements, particularly the carbon cycle, are greatly influenced by primary producers. Therefore, understanding the spatial and temporal dynamics of primary productivity is invaluable to multiple disciplines of earth and life sciences.

[3] In the marine environment, in situ measurements of net primary productivity (NPP) (Table 1) are sparse through space and time and can only represent minute fractions of ecosystems. To assess NPP over large areas and annual to decadal time-scales, we must rely on models, some of which use ocean color from satellite sensors while others couple biogeochemistry and ocean circulation. These NPP models have been applied to answer a wide range of scientific questions pertaining to topics such as fisheries management [*Zainuddin et al.*, 2008], sea turtle population dynamics [*Saba et al.*, 2008], the biological pump (export flux) [*Laws et al.*, 2000], oxygen production [*Reuer et al.*, 2007], and contemporary changes in marine phytoplankton [*Behrenfeld et al.*, 2006].

[4] The frequent application and wide variety of model estimates of NPP requires a context in which these models can be evaluated to determine their accuracy; this also facilitates further model development and improvement. The Primary Productivity Algorithm Round Robin (PPARR) provides this framework. Early PPARR studies compared a small number of model estimates to in situ NPP data at ~90 stations from various marine ecosystems [*Campbell et al.*, 2002]. Global fields of NPP estimated by 31 satellite-based ocean color models and coupled biogeochemical ocean general circulation models were contrasted to understand why and where models diverge in their estimates [*Carr et al.*, 2006]. A study comparing the NPP estimates of 30 models to in situ data from ~1000 stations over 13 years in the tropical Pacific Ocean revealed an overall increase in ocean color model skill [*Friedrichs et al.*, 2009] relative to the first PPARR study [*Campbell et al.*, 2002]. Scientists used the comparative results from these PPARR studies, in addition to their own research, to help refine and improve their models, thus demonstrating the success of the PPARR effort.

[5] The aforementioned studies evaluated NPP models at multiple locations through various time periods, primarily because in situ NPP data are typically measured along cruise-ship transects that are on a timescale of days or weeks, or are the result of concentrated sampling in short-term process studies in a specific region. Time series projects, including the Bermuda Atlantic Time series Study (BATS) and the Hawaii Ocean Time series (HOT) located in the subtropical gyres of the North Atlantic and North Pacific Oceans respectively, provide a different sampling scheme. In each time series project, monthly (sometimes inter-monthly) measurements of several oceanographic variables, including NPP, are collected at the same location to produce a data set now spanning almost two decades at these sites. These invaluable data sets [e.g., *Ducklow et al.*, 2009] enable analysis of NPP model skill spanning nearly two decades.

[6] Most ocean color models estimate NPP from concentrations of surface chlorophyll-*a*, which can be derived from satellite sensors or in situ measurements. Since September 1997, the Sea-viewing Wide Field-of-View Sensor (SeaWiFS) has been providing nearly global coverage of ocean color and surface chlorophyll-*a* from space [*McClain et al.*, 2004]. At BATS and HOT, in situ measurements of chlorophyll-*a* are routinely derived from both fluorometry and high-performance liquid chromatography (HPLC). Using these three estimates of surface chlorophyll, we can assess how chlorophyll measurement type affects ocean color NPP estimates.

[7] Given the locations of BATS and HOT, we can also examine how models estimate NPP through multidecadal climate forcing such as the North Atlantic Oscillation (NAO), the North Pacific Gyre Oscillation (NPGO), the El Niño Southern Oscillation (ENSO), and the Pacific Decadal Oscillation (PDO). It has been suggested that these climate indices are associated with trends in NPP in the respective subtropical gyres of the North Atlantic [*Bates*, 2001; *Krause et al.*, 2009; *Lomas et al.*, 2010] and the North Pacific [*Corno et al.*, 2007; *Di Lorenzo et al.*, 2008; *Bidigare et al.*,

Table 1. Acronyms Used Throughout the Paper

Acronym	Definition
1D-ECO	One-dimensional ecosystem model
B	Bias
BATS	Bermuda Atlantic Time series Study
BBP	Particulate backscatter
BOGCM	Biogeochemical ocean general circulation model
CBSAT	Carbon-based satellite ocean color model
Chl-*a*	Chlorophyll-*a*
DI	Depth-integrated
DR	Depth-resolved
ENSO	El Niño Southern Oscillation
HOT	Hawaii Ocean Time series
HPLC	High-performance liquid chromatography
JGOFS	Joint Global Ocean Flux Study
MEI	Multivariate ENSO Index
MLD	Mixed-layer depth
NAO	North Atlantic Oscillation
NPGO	North Pacific Gyre Oscillation
NPP	Net primary productivity
PAR	Photosynthetically active radiation
PDO	Pacific Decadal Oscillation
PPARR	Primary Productivity Algorithm Round Robin
RMSD	Root mean square difference
SAT	Satellite ocean color model
SeaWiFS	Sea-viewing Wide Field-of-view Sensor
SST	Sea surface temperature
STMW	Subtropical Mode Water
uRMSD	Unbiased root mean square difference
WI	Wavelength-integrated
WR	Wavelength-resolved

2009]; thus it is important to assess model performance through these oscillations.

[8] Here we assess the performance of 36 models by comparing estimated NPP to in situ data at BATS and HOT over the course of nearly 20 years. Specifically, we examine the models' ability to estimate the mean, variability, and trends of in situ NPP. We first describe the biological and physical data from the two stations to better understand the observed variability in NPP; we also consider the multi-decadal climate oscillations that may be driving local ecosystem dynamics and influencing NPP variability. Model performance is first assessed in terms of bias and variability and presented using root-mean squared differences and illustrated using target diagrams. Next, linear regression is applied to determine how well models estimate observed NPP trends. This is followed by a sensitivity analysis that compares how ocean color based models estimate NPP when using different measurements of surface chlorophyll-*a* during the nearly 20 year BATS and HOT time series, as well as during the shorter SeaWiFS (post-1997) time series.

2. Methods

2.1. Data

[9] Both BATS (31°40′N, 64°10′W) and HOT (22°45′N, 158°00′W) began as part of the U.S. Joint Global Ocean Flux Study (JGOFS); thus all sampling at the two stations follows JGOFS protocols. Monthly ^{14}C tracer measurements of NPP at these stations were based on dawn to dusk (10 to 16 h) in situ incubations of samples collected at 20 m intervals from 0 to 140 m at BATS and at 20–30 m intervals from 0 to 125 m at HOT (0 to 175 m before the year 2000). At BATS, incubations were performed using both light and dark bottles for the entire time series, whereas at HOT, dark bottles have not been used since 2000. We extracted data at all depths from each project's website (BATS: http://bats.bios.edu; HOT: http://hahana.soest.hawaii.edu/hot/hot_jgofs.html). For the BATS data, we used the mean light values of NPP (N = 3) and subtracted the dark values to remove the carbon produced by non-photoautotrophs. Mean dark values were 13.63% (±8%) of light values at BATS. For the HOT data, we calculated the mean proportion of dark to light bottle values from 1989 to 2000 (5.02% ± 2%) and then used this proportion to calculate NPP for all light bottle samples from 2000 onwards. The trapezoid method was used to integrate daily NPP (mg C m^{-2} day^{-1}) to the base of the euphotic zone (long-term estimated mean = 140 m; [e.g., *Roman et al.*, 2002]).

[10] Additionally, we used surface (average value from 0 to 5 m because the first optical depth is typically between 5 to 10 m in these regions) and integrated chlorophyll-*a* (Chl-*a*) (to 200 m) measured via fluorometry and high-performance liquid chromatography (HPLC), sea surface temperature (SST), temperature at 100 m, salinity at 200 m, and integrated nitrate/nitrite concentration to 150 m. Mixed-layer depths (MLDs) at HOT were derived using the surface offset method ($\Delta\sigma$ = 0.125 kg m^{-3}) based on potential density profiles. The MLDs at BATS were derived from a surface forced mixed-layer model [*Doney*, 1996; *Doney et al.*, 2007] that matches MLD derived using the surface offset method very closely [*Ducklow et al.*, 2009]. To compare ocean color model NPP estimates using in situ measurements of surface Chl-*a* to that of satellite sensors, we used 8-day level-3 SeaWiFS 9 km Chl-*a* data from the Ocean Color Web (http://oceancolor.gsfc.nasa.gov). SeaWiFS Chl-*a* measurements were averaged for the 3 × 3 pixel window (27 × 27 km) that encompassed each of the time series stations. This was done for each 8-day SeaWiFS image that contained the respective date of in situ sampling. Trend analysis of all data was conducted using linear regression (95% confidence interval) of a 12-month moving average of each time series.

[11] We compared time series of the NAO index (http://www.cpc.ncep.noaa.gov/products/precip/CWlink/pna/nao.shtml), NPGO index (http://eros.eas.gatech.edu/npgo), Multivariate ENSO Index (MEI) (http://www.cdc.noaa.gov/people/klaus.wolter/MEI), and PDO index (http://jisao.washington.edu/pdo) to the NPP time series at the two stations. To determine the correlation between the observed data and each of the four multidecadal climate indices, we computed correlation coefficients between a 12-month moving average of each climate index and each station's time series of NPP data in monthly anomaly form (A_m):

$$A_m = O_m - \overline{O_t}$$

where O_m is the observed data during month m and $\overline{O_t}$ is the average monthly value over the entire time series. The 12-month moving average for each A_m was calculated by taking the average A_m over months m to m + 11 months. Correlation coefficients between observed data trends and

Table 2. Contributed Primary Productivity Models[a]

Model	Contributer	Type	Input Variables Used					Reference
			Chl-*a*	SST	PAR	MLD	BBP	
1	Saba	SAT,DI,WI	x					*Eppley et al.* [1985]
2	Saba	SAT,DI,WI	x	x	x	x		*Howard and Yoder* [1997]
3	Saba	SAT,DI,WI	x	x	x			*Carr* [2002]
4	Dowell	SAT,DI,WI	x	x	x	x		M. Dowell et al. (unpublished data, 2008)
5	Scardi	SAT,DI,WI	x	x	x	x		*Scardi* [2001]
6	Ciotti	SAT,DI,WI	x	x	x			*Morel and Maritorena* [2001]
7	Kameda; Ishizaka	SAT,DI,WI	x	x	x			*Kameda and Ishizaka* [2005]
8	Westberry; Behrenfeld	SAT,DI,WI	x	x	x			*Behrenfeld and Falkowski* [1997]
9	Westberry; Behrenfeld	SAT,DI,WI	x	x	x			*Behrenfeld and Falkowski* [1997]
10	Tang	SAT,DI,WI	x	x	x			*Tang et al.* [2008]
11	Tang	SAT,DI,WI	x	x	x			*Tang et al.* [2008]
12	Armstrong	SAT,DR,WI	x	x	x			*Armstrong* [2006]
13	Armstrong	SAT,DR,WI	x	x	x			*Armstrong* [2006]
14	Asanuma	SAT,DR,WI	x	x	x			*Asanuma* [2006]
15	Marra; O'Reilly; Hyde	SAT,DR,WI	x	x	x			*Marra et al.* [2003]
16	Antoine; Morel	SAT,DR,WR	x	x	x	x		*Antoine and Morel* [1996]
17	Uitz	SAT,DR,WR	x		x	x		*Uitz et al.* [2008]
18	Mélin; Hoepffner	SAT,DR,WR	x		x			*Mélin* [2003]
19	Smyth	SAT,DR,WR	x	x	x			*Smyth et al.* [2005]
20	Waters	SAT,DR,WR	x	x	x	x		*Ondrusek et al.* [2001]
21	Westberry; Behrenfeld	CBSAT,DI,WI	x		x	x	x	*Behrenfeld et al.* [2005]
22	Westberry; Behrenfeld	CBSAT,DR,WR	x		x	x	x	*Westberry et al.* [2008]
23	Bennington; McKinley	BOGCM						*Bennington et al.* [2009]
24	Aumont; Bopp	BOGCM						*Aumont and Bopp* [2006]
25	Lima; Moore; Doney	BOGCM						*Moore et al.* [2004]
26	Dutkiewicz	BOGCM						*Dutkiewicz et al.* [2009]
27	Gregg	BOGCM						*Gregg* [2008]
28	Gregg	BOGCM						*Gregg and Casey* [2007]
29	Tjiputra	BOGCM						*Tjiputra et al.* [2010]
30	Bruggeman	BOGCM						J. Bruggeman et al. (unpublished data, 2008)
31	Yool	BOGCM						*Yool et al.*, 2007
32	Buitenhuis	BOGCM						*Buitenhuis et al.* [2010]
33	Dunne	BOGCM						*Dunne et al.* [2005]
34	Vichi	BOGCM						*Vichi et al.* [2007a, 2007b]
35	Salihoglu	1D-ECO		x	x	x		*Salihoglu et al.* [2008]
36	Salihoglu	1D-ECO		x	x	x		*Salihoglu et al.* [2008]

[a]Specific details for each model are described in Text S1.

each of the four multidecadal climate indices were computed using a 12-month moving average of each climate index and each station's time series of NPP data in monthly anomaly form. The 12-month moving average was only applied to the trend analyses (linear regression) and the climate index correlation to filter out noise thus giving a clearer representation of trends. The remaining statistics in this paper were based on the raw data with no moving average.

2.2. Models

[12] Participants in this study contributed output from 36 NPP models (Table 2) including satellite-based ocean color models (SATs; $N = 20$), carbon-based SATs (CBSATs; $N = 2$), biogeochemical ocean general circulation models (BOGCMs; $N = 12$), and one-dimensional ecosystem models (1D-ECOs; $N = 2$). Although the CBSATs are actually a type of satellite-based ocean color model and therefore are technically SATs, here they are separated for this analysis because their behavior is very different from the other SATs. Specific details for each of the 36 models are given in Text S1.[1] Among SATs and CBSATs, model complexity ranged from less complex depth-integrated/wavelength-integrated (DI,WI) and depth-resolved/wavelength-integrated (DR,WI), to more complex depth-resolved/wavelength-resolved (DR,WR). The BOGCMs varied in terms of the number of phytoplankton functional groups, number of nutrients, and forcing fields.

[13] All participants were provided with input data and were asked to estimate integrated NPP to the 1% light-level. We provided BOGCMs with the latitude/longitude and date of the BATS and HOT station samples. In addition to position and date, SATs and CBSATs were also provided with in situ surface fluorometric Chl-*a*, in situ SST, modeled photosynthetically active radiation (PAR) from the National Centers for Environmental Prediction (http://www.cdc.noaa.gov), day length, and MLD. Not all SATs and CBSATs used all input variables provided though the use of surface

[1]Auxiliary materials are available in the HTML. doi:10.1029/2009GB003655.

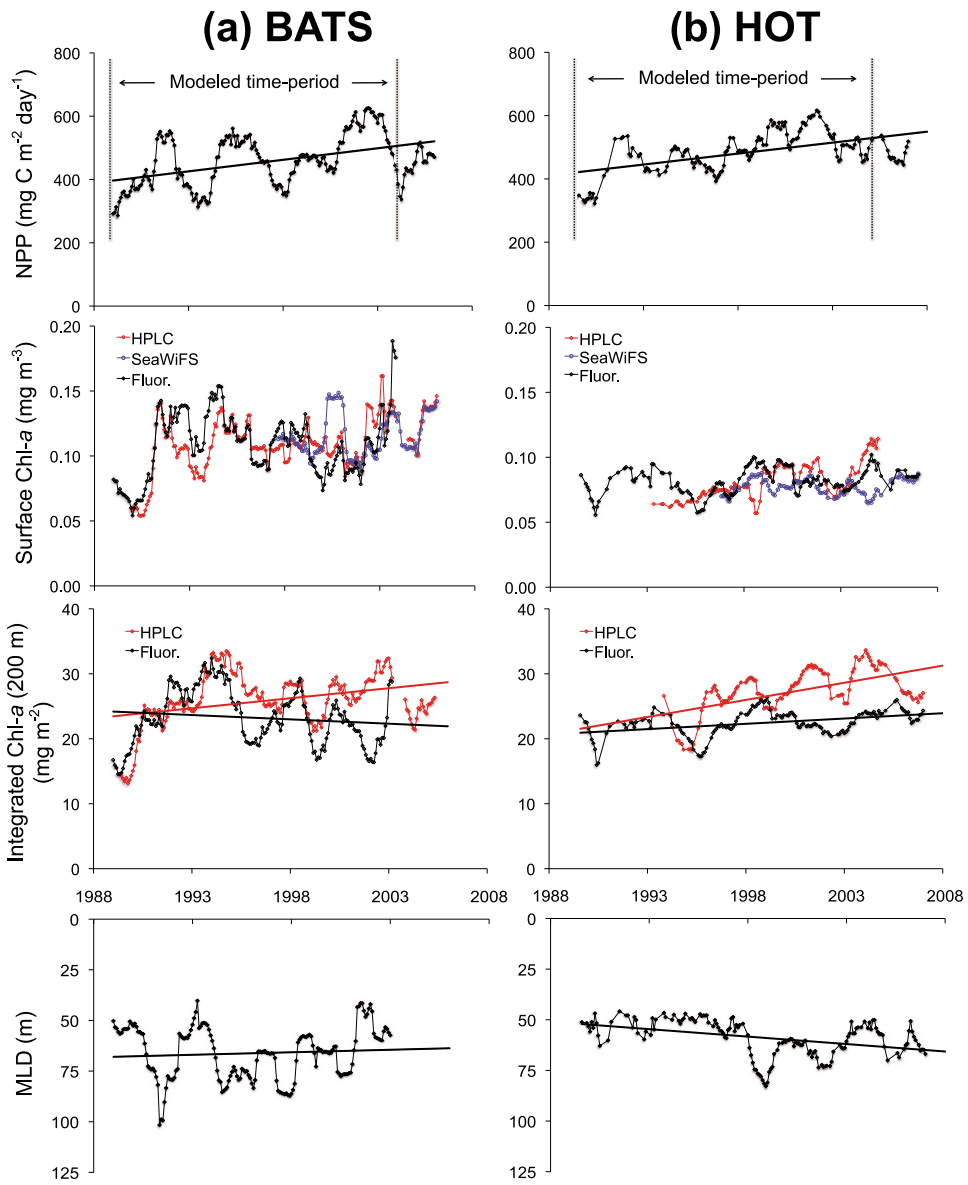

Figure 1. Twelve-month moving average of the observed data at (a) BATS and (b) HOT. All data are in situ with the exception of MLD that is modeled at BATS [*Doney*, 1996; *Doney et al.*, 2007] and calculated based on potential density profiles at HOT. Solid lines represent the fitted trend from linear regression. Statistics for each time series are described in Table 3. Note that the *y* axis scales are different for salinity.

Chl-*a* was universal (Table 2). We also provided SATs and CBSATs with in situ surface Chl-*a* measured via HPLC at BATS and HOT to assess whether Chl-*a* measurement-method affected model estimates. The sample sizes of the two Chl-*a* measurement-methods were comparable at BATS, whereas HPLC samples between 0 and 5 m were not consistently taken at HOT (Figure 1). The CBSATs additionally required particulate backscatter data and these were obtained from the SeaWiFS database (http://oceancolor.gsfc.nasa.gov) using monthly composites. For sample dates prior to the beginning of the SeaWiFS era (September 1997), monthly composites of particulate backscatter were based on the average for each month from September 1997 to September 2007.

[14] The two 1D-ECO models used SST, PAR, and MLD (Table 1) and only varied by the source of MLD. Model 35 used MLD estimated from BATS in situ potential density profiles while Model 36 used MLD from the surface forced model [*Doney*, 1996; *Doney et al.*, 2007; *Ducklow et al.*, 2009] that we provided. Although in situ data in this paper extend to 2007, models were only provided with data extending to late 2003 for BATS and until late 2004 for HOT due to data availability limitations at the start of this study in 2007.

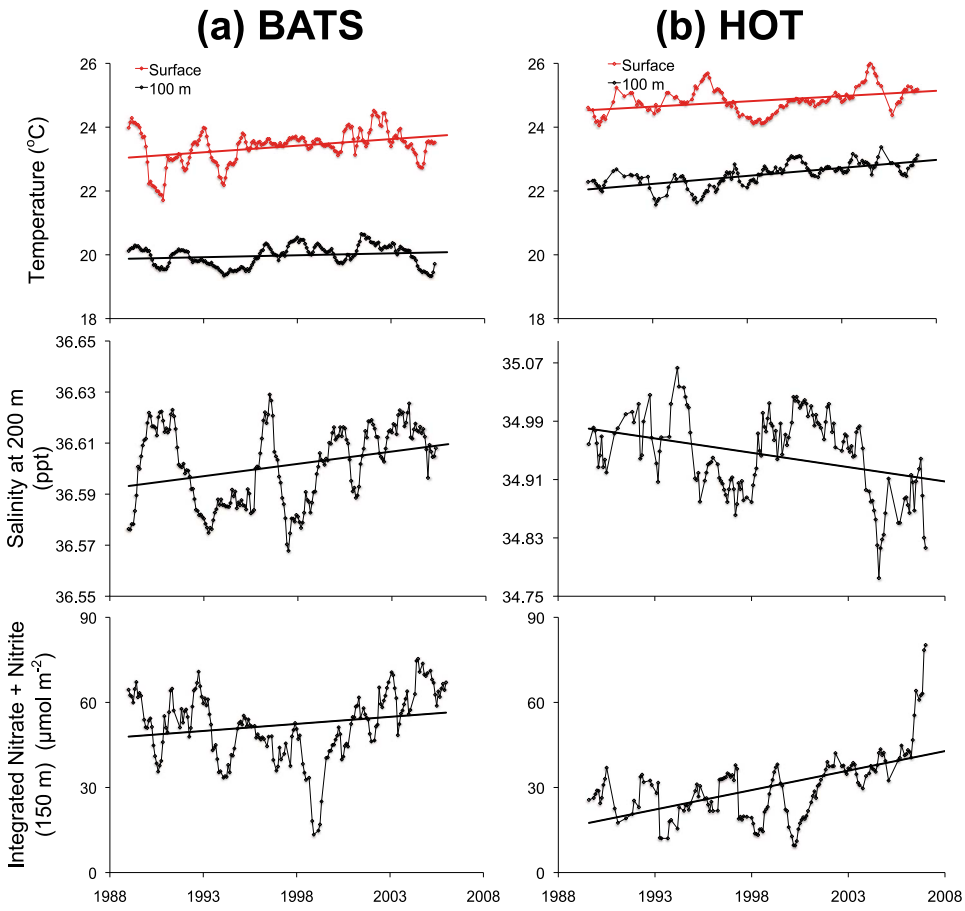

Figure 1. (continued)

[15] An additional analysis was performed for the data sampled during the SeaWiFS era (1997–2004). Ocean color model estimates of NPP were compared using fluorometric, HPLC, and SeaWiFS Chl-*a* at BATS and HOT. The sample sizes of surface fluorometric and HPLC Chl-*a* samples (0 to 5 m) at HOT were more similar during the SeaWiFS era thus we were able to compare model estimates using the different measurement-methods for this time period.

2.3. Model Performance

[16] To assess overall model performance in terms of both bias and variability in a single statistic [*Doney et al.*, 2009; *Stow et al.*, 2009], we used the root mean square difference (RMSD) calculated for each model's N samples of NPP at each site:

$$RMSD = \left(\frac{1}{N}\sum_{i=1}^{N}\Delta(i)^2\right)^{1/2}$$

where model-data misfit in $\log_{10}$ space $\Delta(i)$ is defined as:

$$\Delta(i) = \log(NPP_m(i)) - \log(NPP_d(i))$$

and where $NPP_m(i)$ is modeled NPP and $NPP_d(i)$ represents in situ data at each site. The RMSD statistic assesses model skill such that models with lower values have higher skill. The use of log normalized RMSD to assess overall model performance is consistent with prior PPARR studies [*Campbell et al.*, 2002; *Friedrichs et al.*, 2009]. To assess model skill more specifically (whether a model over- or underestimated NPP), we calculated each model's bias (B) where:

$$B = \overline{\log(NPP_m)} - \overline{\log(NPP_d)}.$$

[17] We determined if certain groups of models had significantly higher skill than others (based on RMSD) by applying an ANOVA method with a 95% confidence interval.

[18] Model performance was also illustrated using target diagrams [*Jolliff et al.*, 2009]. These diagrams are based on the relationship:

$$RMSD^2 = B^2 + uRMSD^2$$

Table 3. Linear Trends of the in Situ Data at BATS and HOT for the Entire Data Sets, Extending Past the Model Range[a]

Station	Parameter	N	Time Series	Trend (12-month Moving Average)	Annual Slope	Overall Change
BATS	NPP	239	1988–2006	Increasing	+10.08 mg C m^{-2} day^{-1} $year^{-1}$	+171.44 mg C m^{-2} day^{-1}
	Fluorometric Chl-*a* (surface)	202	1988–2004	No trend	-	-
	Fluorometric Chl-*a* (integrated 200 m)	190	1988–2004	No trend	-	-
	HPLC Chl-*a* (surface)	218	1990–2006[b]	Increasing	+0.006 mg m^{-3} $year^{-1}$	+0.088 mg m^{-3}
	HPLC Chl-*a* (integrated 200 m)	211	1990–2006[b]	Increasing	+0.740 mg m^{-2} $year^{-1}$	+11.84 mg m^{-2}
	SeaWiFS Chl-*a* (surface)	119	1997–2006	No trend	-	-
	MLD[c]	197	1988–2003	No trend	-	-
	SST	237	1988–2006	No trend	-	-
	Temperature (100 m)	233	1988–2006	No trend	-	-
	Salinity (200 m)	239	1988–2006	Increasing	+0.002 ppt $year^{-1}$	+0.027 ppt
	Nitrate + Nitrite (integrated 150 m)	213	1989–2006	Increasing	+0.05 μmol m^{-2} $year^{-1}$	+0.97 μmol m^{-2}
HOT	NPP	159	1989–2007	Increasing	+10.23 mg C m^{-2} day^{-1} $year^{-1}$	+173.97 mg C m^{-2} day^{-1}
	Fluorometric Chl-*a* (surface)	142	1989–2007	Increasing	+0.001 mg m^{-3} $year^{-1}$	+0.012 mg m^{-3}
	Fluorometric Chl-*a* (integrated 200 m)	159	1989–2007	Increasing	+0.15 mg m^{-2} $year^{-1}$	+2.56 mg m^{-2}
	HPLC Chl-*a* (surface)[d]	72	1995–2005	Increasing	+0.003 mg m^{-3} $year^{-1}$	+0.031 mg m^{-3}
	HPLC Chl-*a* (integrated 200 m)[d]	83	1995–2006	Increasing	+0.28 mg m^{-2} $year^{-1}$	+3.41 mg m^{-2}
	SeaWiFS Chl-*a* (surface)	98	1997–2007	No trend	-	-
	MLD	159	1989–2007	Increasing (deepening)	+0.80 m $year^{-1}$	+13.62 m
	SST	159	1989–2007	Increasing	+0.06°C $year^{-1}$	+0.96°C
	Temperature (100 m)	159	1989–2007	Increasing	+0.06°C $year^{-1}$	+1.00°C
	Salinity (200 m)	159	1989–2007	Decreasing	−0.008 ppt $year^{-1}$	−0.136 ppt
	Nitrate + Nitrite (integrated 150 m)	158	1989–2007	Increasing	+3.28 μmol m^{-2} $year^{-1}$	+55.79 μmol m^{-2}

[a]Linear regression is based on 12-month moving averages of the in situ and estimated fields. Only statistically significant trends are shown ($P < 0.05$; 95% confidence interval) based on linear regression.
[b]HPLC Chl-*a* (surface and integrated) similarly increased from 1990 to 2004.
[c]Modeled MLD [*Doney*, 1996; *Doney et al.*, 2007].
[d]Exludes sample dates when measurements were not taken between 0 and 5 m.

where unbiased RMSD squared ($uRMSD^2$) is defined as:

$$uRMSD^2 = \frac{1}{N}\sum_{i=1}^{N}\left(\left(\log NPP_m(i) - \overline{\log NPP_m}\right) - \left(\log NPP_d(i) - \overline{\log NPP_d}\right)\right)^2.$$

[19] In this way, Target diagrams break down RMSD to show multiple statistics on a single plot: total RMSD as the distance from the origin, bias on the y axis, and the unbiased RMSD on the x axis. By convention, models with standard deviations overestimating the observed standard deviation are plotted on the right side of the plot and models with standard deviations underestimating the observed standard deviation are plotted on the left, i.e., the quantity on the x axis represents signed uRMSD, where:

$$\text{signed } uRMSD = (uRMSD)\, sign(\sigma_m - \sigma_d)$$

and σ_m = standard deviation of $\log NPP_m$ and σ_d = standard deviation of $\log NPP_d$. The target diagram thus enables one to easily visualize whether or not a model over- or under-estimates the mean and variability of NPP. In addition, by plotting a circle with radius equal to the standard deviation of the data, the target diagrams illustrate whether models are performing better than the mean of the observations [*Jolliff et al.*, 2009].

3. Results

3.1. In Situ and Satellite Data Trends

[20] From the late 1980s to the mid- to late 2000s, mean in situ NPP at BATS and HOT was 461 (±201; standard deviation) and 488 (±143) mg C m^{-2} day^{-1} respectively. At both stations, NPP significantly increased ($P < 0.05$) at an average trend of about +10 mg C m^{-2} day^{-1} $year^{-1}$ (Table 3 and Figure 1). During the shorter time series for the model estimates (ending in 2003 and 2004), NPP similarly increased at HOT but increased at a higher rate at BATS (+23 mg C m^{-2} day^{-1} $year^{-1}$; see Table 4). At BATS, the overall increase in NPP occurred primarily in the upper 20 m (35%) (Figure 2a) whereas at HOT, the greatest increase was between 45 and 75 m (36%) (Figure 2b). Other statistically significant trends at BATS included increasing surface and integrated HPLC Chl-*a*, increasing salinity at 200 m, and increasing integrated nitrate plus nitrite (Table 3 and Figure 1a). At HOT, both fluorometric and HPLC surface and integrated Chl-*a* increased, MLD increased (deepened), SST and temperature at 100 m increased, salinity at 200 m decreased, and integrated nitrate plus nitrite

Table 4. Observed and Estimated NPP Among All Model Types During the Modeled Time Period at BATS (1989–2003) and HOT (1989–2004)[a]

Station	NPP Source	Mean NPP (mg C m^{-2} day^{-1})	Trend (12-month Moving Average)	Annual Slope (mg C m^{-2} day^{-1} $year^{-1}$)	Overall Change in NPP (mg C m^{-2} day^{-1})	RMSD
BATS (N = 162; 1989–2003)	In situ (modeled time period)	462.04 (±198)	Increasing	+23.69	+331.73	-
	SAT estimate (fluorometric Chl-*a* based)	327.64 (±170)	Increasing	+5.90	+82.63	0.37 (±0.10)
	SAT estimate (HPLC Chl-*a* based)	334.87 (±160)	Increasing	+12.21	+158.79	0.35 (±0.10)
	CBSAT estimate (fluorometric Chl-*a* based)	216.36 (±170)	No trend	-	-	0.87 (±0.33)
	CBSAT estimate (HPLC Chl-*a* based)	206.61 (±155)	Increasing	+15.41	+200.34	0.86 (±0.37)
	BOGCM estimate	289.98 (±116)	No trend	-	-	0.47 (±0.24)
	1D-ECO estimate	406.60 (±174)	Increasing	+9.30	+130.14	0.23 (±0.00)
	Model 23 estimate	126.50 (±133)	Increasing	+5.87	+82.24	0.81
	Model 36 estimate[b]	440.72 (±234)	Increasing	+11.01	+154.19	0.22
HOT (N = 134; 1989–2004)	In situ (modeled time period)	487.33 (±150)	Increasing	+11.52	+161.33	-
	SAT estimate (fluorometric Chl-*a* based)	325.89 (±74)	Increasing	+0.84	+11.78	0.26 (±0.07)
	CBSAT estimate (fluorometric Chl-*a* based)	336.00 (±87)	No trend	-	-	0.30 (±0.13)
	BOGCM estimate	196.58 (±30)	No trend	-	-	0.62 (±0.38)
	Model 25 estimate	282.48 (±87)	Increasing	+10.66	+149.23	0.29

[a]Linear trends and overall model performance (RMSD ± standard deviation) of the various model types along with select individual models that estimated the correct trends in NPP. Lower RMSD values signify higher model skill.
[b]Model 35 reproduced an increasing trend in NPP that was similar to that of Model 36 although the mean NPP of Model 35 was slightly lower.

increased (Table 3 and Figure 1b). During the SeaWiFS era from 1997 onwards, no significant trends were detected for SeaWiFS Chl-*a* at either station (Table 3 and Figure 1). By contrast HPLC Chl-*a* increased while fluorometric Chl-*a* had no significant trend at both sites during the SeaWiFS era.

[21] In situ NPP at BATS was negatively correlated to the NAO (r = −0.22) while positively correlated to the NPGO (r = 0.35) (Figure 3a). No significant correlations were found between NPP at BATS and the MEI and PDO. NPP at HOT was even more positively correlated to the NPGO (0.61), but was slightly negatively correlated to the MEI (−0.18) and PDO (−0.19), and had no correlation to the NAO (Figure 3b). These correlations will be discussed in detail in section 4.3.

3.2. Overall Model Performance: RMSD

[22] During the modeled time period (1989 to 2003/2004), the average RMSD computed for the SATs was significantly lower (higher skill) than that computed for the BOGCMs at both stations (ANOVA, $P < 0.05$) although some individual BOGCMs outperformed some SATs (Figure 4). Among all models at BATS, the average CBSATs had the lowest skill while the average 1D-ECO models had the highest skill (Figures 4a and 5a). At HOT, however, CBSATs performed as well as the SAT models (Table 4 and Figures 4b and 5b).

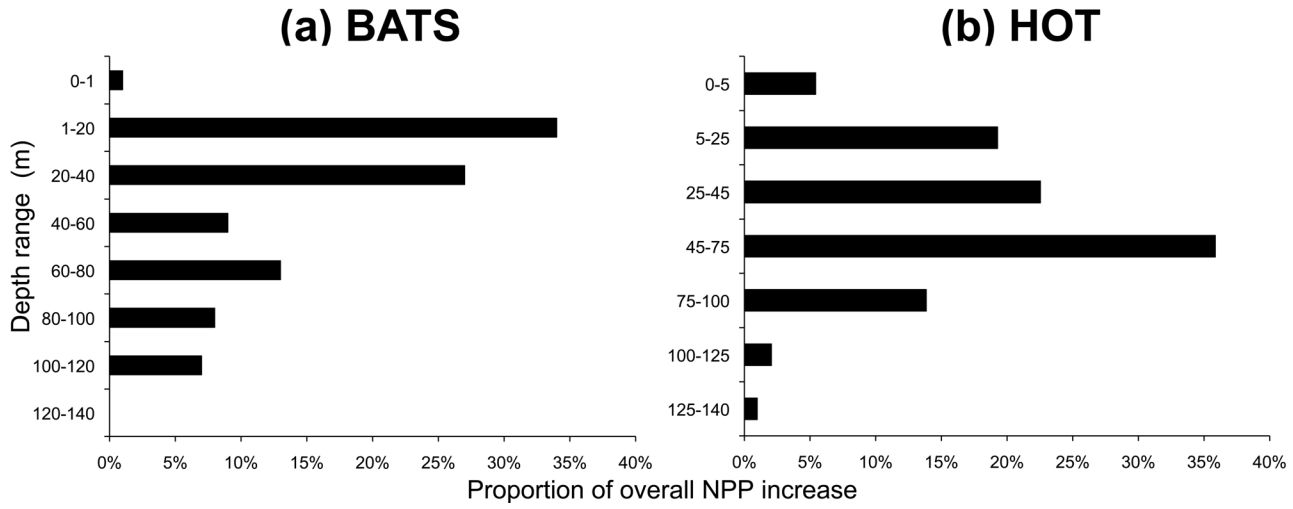

Figure 2. Integrated depth profiles of the overall change in observed NPP at (a) BATS and (b) HOT. The depth ranges for each station are based on where NPP is sampled in the water column. The x axis is the proportion of the overall change in NPP integrated to 140 m.

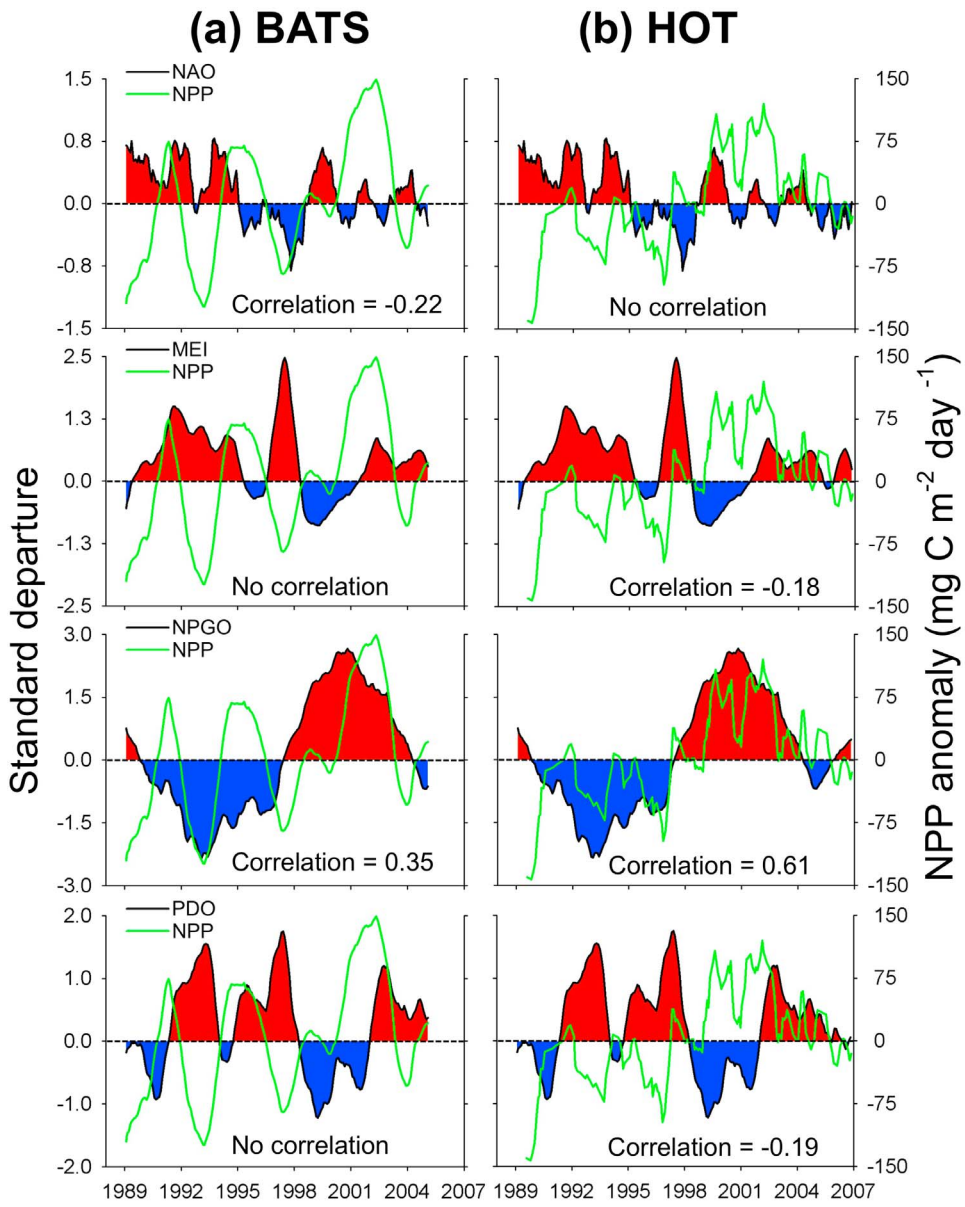

Figure 3. Multidecadal climate indices (12-month moving average) that may be associated with the variability of NPP at (a) BATS and (b) HOT. Observed NPP anomalies are also shown as a 12-month moving average (green line). Correlation coefficients are listed only for significant relationships ($P < 0.05$).

There was no significant difference in model skill between the three types of SATs (Figure 4) and almost all models underestimated NPP at both stations (Table 4 and Figures 5, 6, and 7). During the SeaWiFS time period (1997 to 2003/2004), the skill of the various model types was similar (in terms of RMSD) to that during the entire modeled time period (1989 to 2003/2004) (Tables 4 and 5 and Figure 5). The RMSD of SATs and CBSATs was not significantly affected by the source of Chl-*a* (fluorometric versus HPLC versus SeaWiFS) (Tables 4 and 5 and Figure 5).

3.3. Detailed Model Performance

3.3.1. Model Bias and Variance

[23] The Target diagram (Figure 5) shows bias and signed unbiased RMSD statistics to illustrate how well the models estimate the mean and variability of NPP. At both sites, over 90% of the models underestimated mean NPP, with the BOGCMs on average having a negative bias that was nearly twice that of SATs. Regardless of time period or location, BOGCMs typically estimated the observed variability of NPP more accurately than the mean (i.e., uRMSD < B), thus

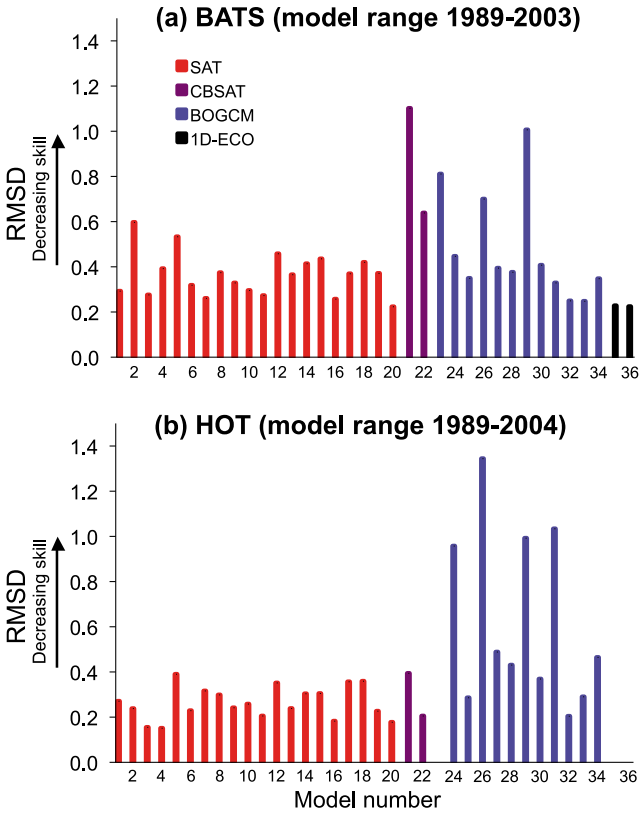

Figure 4. Overall model skill for each model at (a) BATS and (b) HOT during the modeled time period (1989 to 2003/2004). Lower RMSD values correspond to higher model skill. SAT models 1 to 11 are DIWI, 12 to 15 are DRWI, 16 to 20 are DRWR. CBSAT model 21 is DIWI while 22 is DRWR. Models 23, 35, and 36 did not provide output for HOT.

their relatively low skill was driven by bias (Figures 5a–5d). The SATs typically performed more equally in terms of variability and bias (Figures 5a–5d). The CBSATs performed equally in terms of variability and bias at BATS, whereas at HOT they replicated NPP variability very well, and their skill was just limited by bias. The difference in overall skill between the "best" BOGCM and the "best" SAT at each site, i.e., those with the lowest total RMSD, was not significant (Figures 5e–5h). These "best" models, unlike most others, typically overestimated mean NPP and underestimated NPP variance (Figures 5e–5h).

[24] Regardless of time period, the BOGCMs with the lowest RMSD at BATS and HOT were Models 33 and 32 respectively (Figure 4). Among SATs, Model 20 was the overall highest performing model as it had the lowest RMSD at both sites (Figure 4). Another SAT that had relatively high performance at both sites was Model 3 (Figure 4). Although Chl-*a* source did not significantly affect ocean color model performance as measured by total RMSD at either site during either time period, the SATs estimated NPP variability at BATS much better (lower uRMSD) when SeaWiFS Chl-*a* was used as opposed to fluorometric and HPLC (Figure 5b).

3.3.2. Trend Analysis

[25] During the modeled time period at BATS (1989–2003), the average SAT model estimated a statistically significant increasing trend in NPP, regardless of whether HPLC or fluorometric Chl-*a* was used (Table 4 and Figure 6a). However, the magnitude of the NPP increase varied based on Chl-*a* source, with HPLC-based estimates being closer to the observed (Table 4 and Figure 6a). The CBSAT models only estimated an increase in NPP at BATS when HPLC was used. The average BOGCM model did not estimate the increasing trend at either station (Table 4 and Figure 6) although a few individual models were successful. At BATS, Models 23 (BOGCM), 35 and 36 (1D-ECO) produced an increasing NPP trend (Table 4). At HOT, Model 25 was the only BOGCM that produced a significantly increasing NPP trend. Some individual models produced statistically significant decreasing NPP trends, opposite to the observations. At BATS, Model 21 (CBSAT) and Models 26, 27, 28, 32, and 34 (BOGCMs) produced significantly decreasing NPP trends. At HOT, Model 21 (CBSAT) and Models 27, 28, 31, 32, and 33 (BOGCMs) also produced decreasing trends.

[26] The trends of in situ NPP during the SeaWiFS era (post-1997; see Table 5) differed from the trends observed over the entire modeled time period discussed above (Table 4). At BATS, the trend from 1997 to 2003 (+63.18 mg C m^{-2} day^{-1} $year^{-1}$) was nearly three times larger than the trend from 1989 to 2003 (+23.69 mg C m^{-2} day^{-1} $year^{-1}$). On the contrary, at HOT no statistically significant trend was observed from 1997 to 2004, whereas when computed over the entire modeled time period (1989–2004), the average NPP trend was +11.52 mg C m^{-2} day^{-1} $year^{-1}$. At BATS, the ocean color models produced the increasing NPP trend only when HPLC Chl-*a* was used, with the magnitude of the trend being roughly 20% of the observed for the SATs, and roughly 50% of the observed for the CBSATs (Table 5). When fluorometric Chl-*a* was used, opposite (decreasing) trends of NPP were produced by both SATs and CBSATs (Table 5), and when SeaWiFS Chl-*a* was used, neither SATS nor CBSATs produced a significant trend. At HOT, both SATs and CBSATs correctly estimated the lack of a trend when using HPLC Chl-*a*. At this site the CBSATs also correctly estimated the lack of a trend when using SeaWiFS Chl-*a* whereas on average the SATs produced decreasing trends when either fluorometric or SeaWiFS Chl-*a* was used (Table 5).

3.3.3. Trends in Model-Data Misfit

[27] The trends of model-data misfit [$\Delta(i)$] varied with both time period (entire modeled time period versus SeaWiFS time period) and Chl-*a* source. On average, the model-data misfit of BOGCMs significantly increased (became more negative) through time at BATS and HOT during the entire modeled time period, i.e., the models underestimated NPP more as time progressed (Figures 7a and 7b). These BOGCMs had similar trends in model-data misfit through time at BATS during the more recent SeaWiFS years, yet showed no change in misfit at HOT for this time period (Figures 7c and 7d). The average 1D-ECO had no change in model-data misfit at BATS during the entire model range (Figure 7a) but had increasing misfit during the SeaWiFS

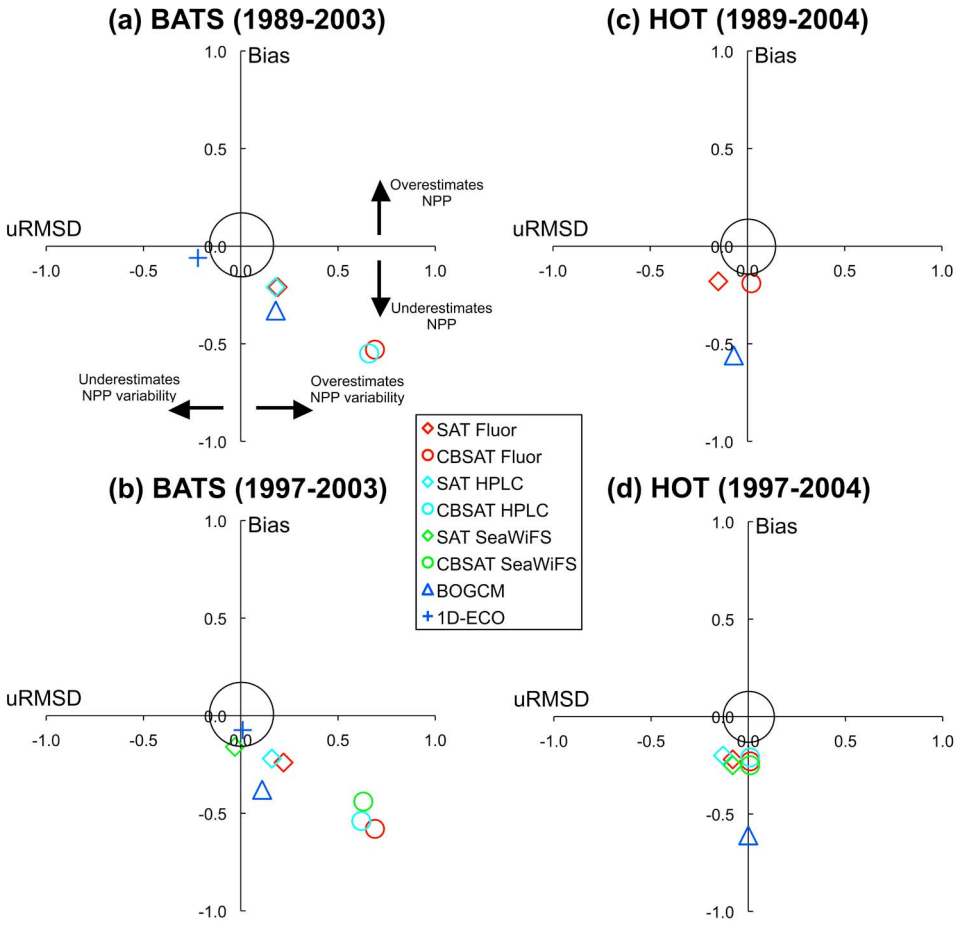

Figure 5. Target diagrams for the various model types and individual models that had high skill at BATS and HOT. (a and e) BATS results for the modeled time period (1989 to 2003) and (b and f) for the SeaWiFS time period (1997–2003). (c and g) HOT results for the modeled time period (1989 to 2004) and (d and h) for the SeaWiFS time period (1997–2004). The solid circle is the standard deviation of the observed data (σ_d). The distance from the origin to each model's symbol is the total RMSD. Refer to section 3.3.1 and Figure 4 for details on individual model performance.

range (Figure 7c). Among SATs, model-data misfit significantly increased at BATS and HOT when the models used fluorometric Chl-*a*; however, when HPLC Chl-*a* was used at BATS, there was no trend in misfit even though the mean misfits were similar (Figures 7a and 7b). This effect of Chl-*a* source on model-data misfit, however, was only evident during the entire modeled time period (Figure 7a). Finally, the CBSATs only had significantly changing misfit during the entire model range at HOT when using fluorometric Chl-*a* (Figure 7b).

4. Discussion

4.1. General Model Performance

[28] The primary goal of this study was to assess how the 36 models estimated the mean, variability, and trends of NPP through multiple decades rather than to strictly critique and compare individual models. A more rigorous assessment of individual model skill at multiple regions, including BATS and HOT, will be presented in a forthcoming paper (V. S. Saba et al., Estimating depth-integrated marine primary productivity in coastal and pelagic regions across the globe: An evaluation of satellite-based ocean color models, manuscript in preparation, 2010). As found in the tropical Pacific study [*Friedrichs et al.*, 2009], there was no relationship between ocean color model skill and model complexity: ocean color models that resolved depth and wavelength did not perform significantly better or worse than the integrated models. These consistent results strengthen the argument that increased model complexity does not always increase model skill.

[29] The performance of the BOGCMs was more variable than that of the SATs, but on average the SATs significantly outperformed the BOGCMs (lower total RMSD) at both sites. The skill of most BOGCMs was limited by their inability to estimate mean NPP rather than NPP variability.

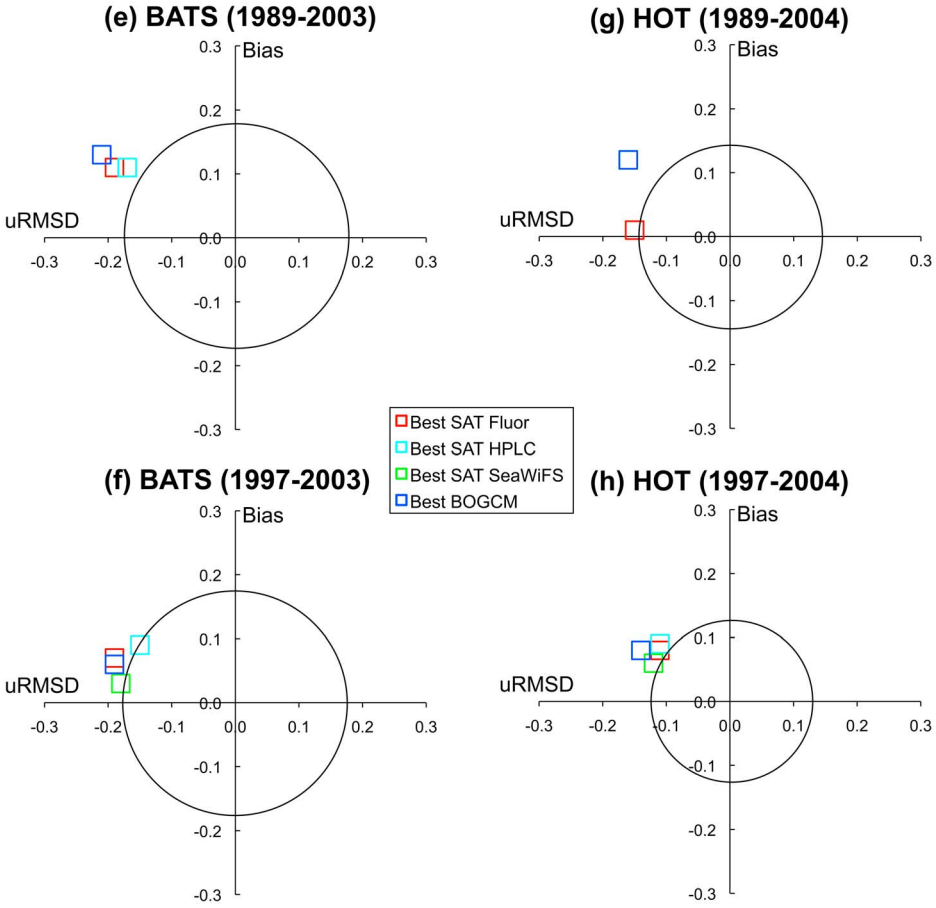

Figure 5. (continued)

In fact, the BOGCMs estimated NPP variability as well or better than the SATs. Interestingly, CBSATs performed significantly better at HOT than at BATS whereas the difference between the two sites for SATs and BOGCMs was not as pronounced. It is not clear whether this is because the CBSATs were particularly tuned for the Pacific, or whether there is something inherent about these carbon-based models that make them much more applicable to HOT than BATS. Model-data misfit among CBSATs is highly correlated to MLD (V. S. Saba et al., unpublished data, 2010) and the sensitivity of Model 21 (DIWI CBSAT) to MLD is a function of season and latitude [*Milutinović et al.*, 2009]. Therefore, the deeper winter-time MLDs at BATS and more temperate nature of the ecosystem may explain why the CBSATs had lower skill there. The 1D-ECO models had high skill at BATS; however, these two models were specifically tuned for the BATS station and are not applicable on a regional or global scale.

[30] Almost all model types consistently underestimated NPP at both BATS and HOT. A model skill assessment done by *Siegel et al.* [2001] using fluorometric derived Chl-*a* measurements as input data showed that 7 of 11 models underestimated NPP at BATS from 1992 to 1997. This is consistent with our results, which show that the majority of the 36 models (as well as the majority of ocean color models) underestimated NPP during this same time period (Figures 6a and 7a). Of the 30 models tested in the recent tropical Pacific comparison exercise [*Friedrichs et al.*, 2009], only 20% underestimated NPP. It is not clear why a greater percentage of models underestimated NPP in our study; we can only surmise that these differences result from the tropical Pacific study covering data from a wider range of marine ecosystems whereas the results from this study are from single station time series sampling representing two individual ecosystems.

[31] High-nitrate low-chlorophyll marine ecosystems, such as the equatorial Pacific, have been considered to be a challenge for ocean color models [*Carr et al.*, 2006]. Therefore it is surprising that the average skill among the 21 ocean color models tested in the tropical Pacific (*Friedrichs et al.* [2009]; mean RMSD = 0.29 ± 0.05) was significantly higher ($P < 0.001$) than the average of the 22 models tested in this study at BATS (0.41 ± 0.19) while similar to those at HOT (0.27 ± 0.07). The poor skill at

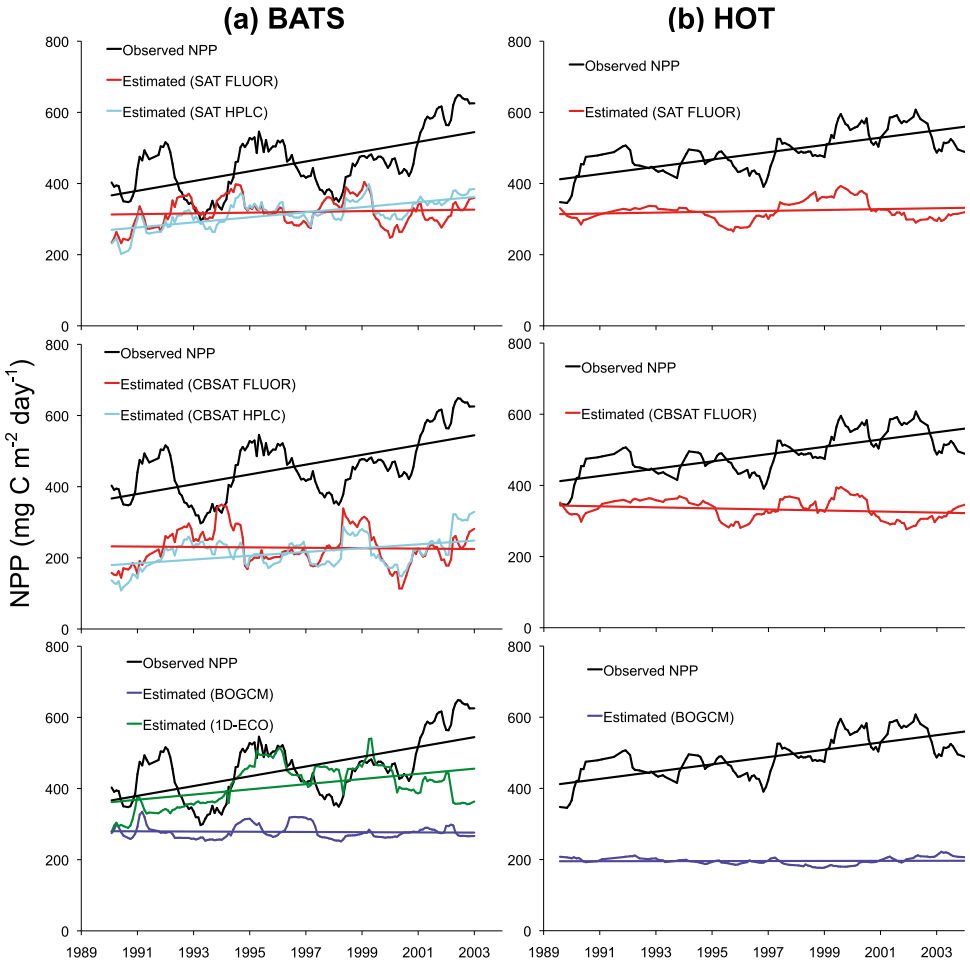

Figure 6. Observed and estimated NPP trends at (a) BATS from 1989 to 2003 and (b) HOT from 1989 to 2004. Trends are shown as a 12-month moving average, and trend lines derive from linear regression. Estimated trends for each model type are based on mean model output for SATs (N = 20), CBSATs (N = 2), BOGCMs, (N = 12 at BATS; N = 11 at HOT), and 1D-ECOs (N = 2 at BATS).

BATS was partially driven by the CBSAT NPP estimates, but even when these models were removed from the mean RMSD, model skill was still significantly higher at HOT and the tropical Pacific.

[32] Among ocean color models, the absolute value of the bias computed by *Friedrichs et al.* [2009] for the tropical Pacific (0.12 ± 0.09) was significantly lower ($P < 0.01$) than at both BATS (0.25 ± 0.14) and HOT (0.20 ± 0.10). In the tropical Pacific exercise, most ocean color models were more limited by their ability to estimate NPP variability rather than mean NPP. This was not true for many of the ocean color models applied to the time series stations discussed here, where the models were equally limited by their ability to estimate mean and variability of NPP.

[33] Mean RMSD from the BOGCMs at HOT (0.62 ± 0.38) was significantly higher ($P < 0.05$) than those found by *Friedrichs et al.* [2009] using the tropical Pacific database (0.34 ± 0.02) but more similar to the RMSD at BATS (0.47 ± 0.24). This was because the absolute value of the bias of BOGCMs was significantly higher at HOT (0.58 ± 0.39) than in the tropical Pacific (0.19 ± 0.12). It is not clear whether this difference is due to something unique about the HOT site, or whether it might result from the fact that different individual BOGCMs participated in this comparison exercise than in the tropical Pacific exercise. As with ocean color models, BOGCMs were also limited by their ability to estimate the variability of NPP in the tropical Pacific when compared to BATS and HOT. This may be due to the higher rate of climate variability in the tropical Pacific derived from multidecadal shifts in ENSO forcing [*Chavez et al.*, 2003] that the BOGCMs were not resolving.

4.2. NPP Trends: Ocean Color Models and Chlorophyll Source

[34] Our results show that for the nearly 15 year BATS and HOT time series analyzed here, the majority of ocean color models estimated the observed increase in NPP at BATS and HOT; however, the ability of ocean color models to successfully estimate these NPP trends depended on time period as well as the type of Chl-a measurement. The majority of SATs and CBSATs estimated the magnitude of the NPP trends to within 50% and 65% respectively, when

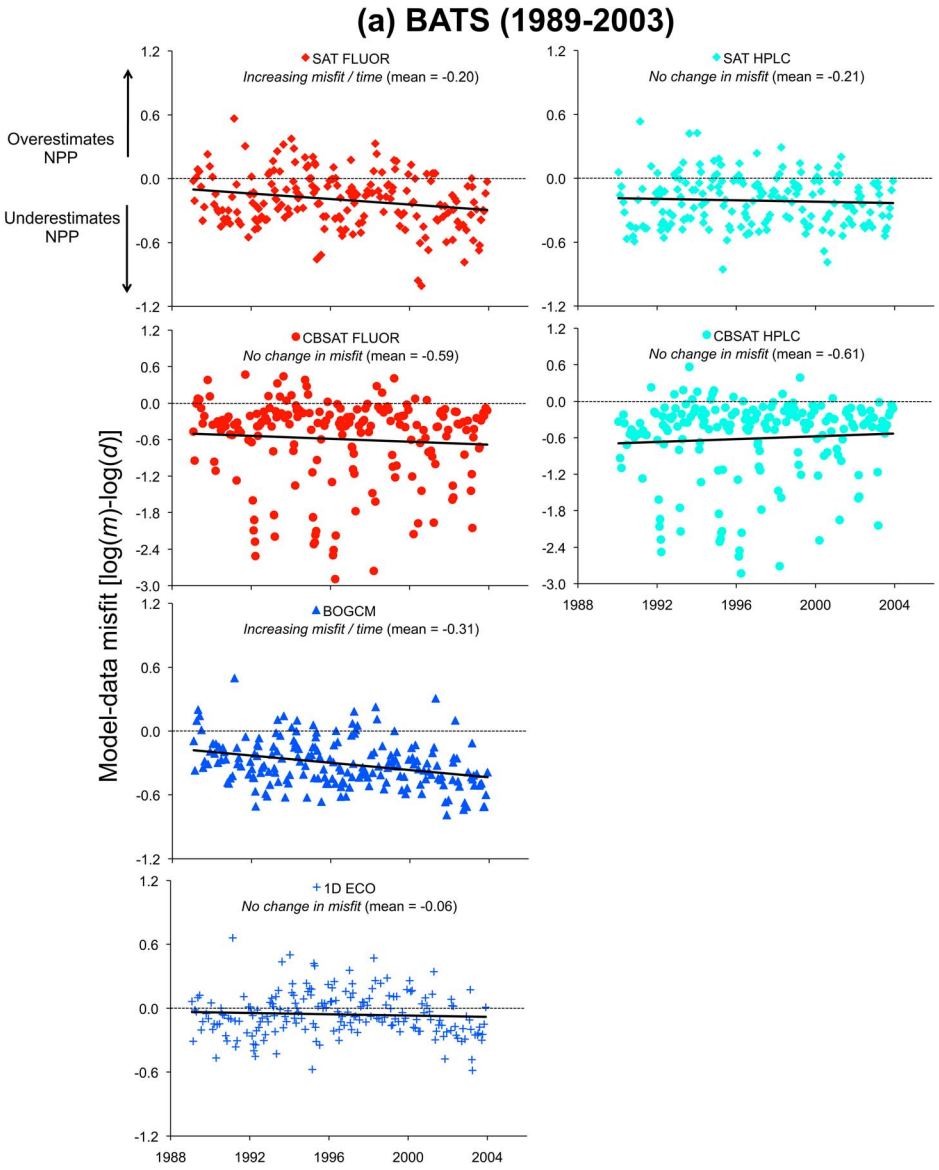

Figure 7. Model-data misfit [$\Delta(i)$] trends for the various model types including ocean color models using the various sources of Chl-*a*. (a) BATS results for the modeled time period (1989 to 2003), (b) HOT results for the modeled time period (1989 to 2004), (c) BATS results for the SeaWiFS time period (1997–2003), and (d) HOT results for the SeaWiFS time period (1997–2004). Solid lines represent the trend lines from linear regression (whether significant or not). Changes in model-data misfit through time are only shown for statistically significant trends ($P < 0.05$; 95% confidence), non-significant trends are noted as "No change in misfit." Note the *y* axis scale difference for CBSATs between BATS and HOT.

HPLC Chl-*a* was used as an input field. In comparison, when fluorometric Chl-*a* was used instead, these models estimated less than 25% of the magnitude of the NPP increase.

[35] During the shorter SeaWiFS time series, the models had much less success in estimating the observed trends. The HPLC-based SATs (CBSATs) estimated the magnitude of the NPP trends at BATS to within 20% (50%), while fluorometric-based SAT and CBSAT estimates produced decreasing trends, opposite of the observed. When SeaWiFS Chl-*a* was used, on average the ocean color models produced no trend at BATS. At HOT, where there was no observed trend, the use of fluorometric Chl-*a* led to a significant decreasing trend. Among SATs, the use of SeaWiFS Chl-*a* did not estimate the correct sign of the observed NPP trends at either site. Although the CBSATs behaved some-

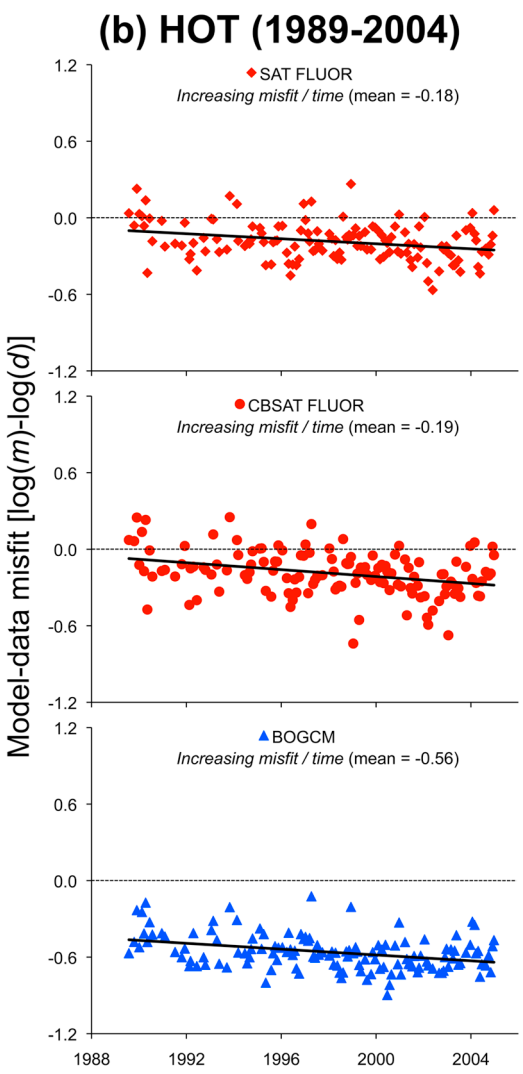

Figure 7. (continued)

what differently than SATs with regard to fluorometric and SeaWiFS derived Chl-*a*, they still consistently estimated NPP trends more accurately when based on HPLC.

[36] The effect of Chl-*a* source on SAT model NPP trend estimates is best illustrated in Figure 7. These model-data misfit plots illustrate why HPLC-based estimates improve the ability of SATs to estimate NPP trends yet do not improve overall model skill in terms of RMSD. During the entire model range at BATS, the model-data misfit of SATs was relatively low in the early portion of the time series but significantly increased as time progressed; when these models used HPLC Chl-*a*, misfit was consistent through time. However, mean model-data misfit did not change using either source of Chl-*a* thus explaining why overall model skill (RMSD) was not affected. This relationship between model-data misfit trends and Chl-*a* source was dependent on time period. During the shortened SeaWiFS range at BATS, model-data misfit of SATs significantly increased, regardless of Chl-*a* source. As mentioned above, this was also the time period during which SATs did not estimate the NPP trends very well, even when using HPLC Chl-*a*.

[37] Time period also affected the trends of model-data misfit for the 1D-ECO models. While no change in misfit was observed among the average 1D-ECO for the entire model range at BATS (Figure 7a), misfit significantly increased during the SeaWiFS range. Consequently, the average 1D-ECO did not estimate the NPP trends during the SeaWiFS time period as it did during the entire modeled time period and this is clearly illustrated in the misfit plots. The limited ability of BOGCMs to estimate NPP trends is also described in the model-data misfit plots such that misfit typically changed through time.

[38] Model-data misfit among CBSATs using fluorometric Chl-*a* at HOT also increased, however, the lack of change observed at BATS using either Chl-*a* source appeared to be due to outliers, although a shift in the direction of the trend lines can still be discerned (Figure 7a).

[39] Therefore, our results suggest the following: 1) a large discrepancy between the observed and estimated magnitude of the NPP trends causes model-data misfit to significantly increase through time; 2) no change in model-data misfit through time exists when models estimate NPP trends more accurately (to within 50%); 3) estimating NPP trends more accurately may not yield greater overall model skill (RMSD), therefore this statistic should be used with caution when assessing a model's ability to estimate NPP trends.

[40] Although the difference between mean fluorometric, HPLC, and SeaWiFS Chl-*a* was not significant at BATS (Figure 1a), ocean color models still estimated the NPP trends more accurately when HPLC Chl-*a* was used (Tables 4 and 5). Ocean color models are extremely sensitive to perturbations in surface Chl-*a* relative to other input variables [*Friedrichs et al.*, 2009]. It has been reported that HPLC measurements of Chl-*a* are more accurate than spectrophotometric and fluorometric measurements [*Pinckney et al.*, 1994]. If satellite-derived surface Chl-*a*, SeaWiFS in this case, can be parameterized using in situ measurements via HPLC, ocean color models might resolve NPP trends more accurately when using satellite data. This, however, would require a large database of in situ HPLC samples spanning multiple regions, which presently does not exist.

4.3. Multidecadal Forcing and NPP Trends

4.3.1. BATS

[41] Previous studies have suggested that variability in marine NPP at BATS is correlated to multidecadal climate indices in the North Atlantic. For example, *Bates* [2001] reported that NPP anomalies at BATS were inversely correlated with NAO variability because negative NAO phases were associated with stronger winds, which resulted in deeper MLDs, cooler SSTs and increases in NPP. Moreover, *Lomas et al.* [2010] reported an increased abundance of prokaryotic phytoplankton and particulate organic carbon export (150 m) that coincided with a negative NAO phase. A weak negative correlation between NPP and NAO was

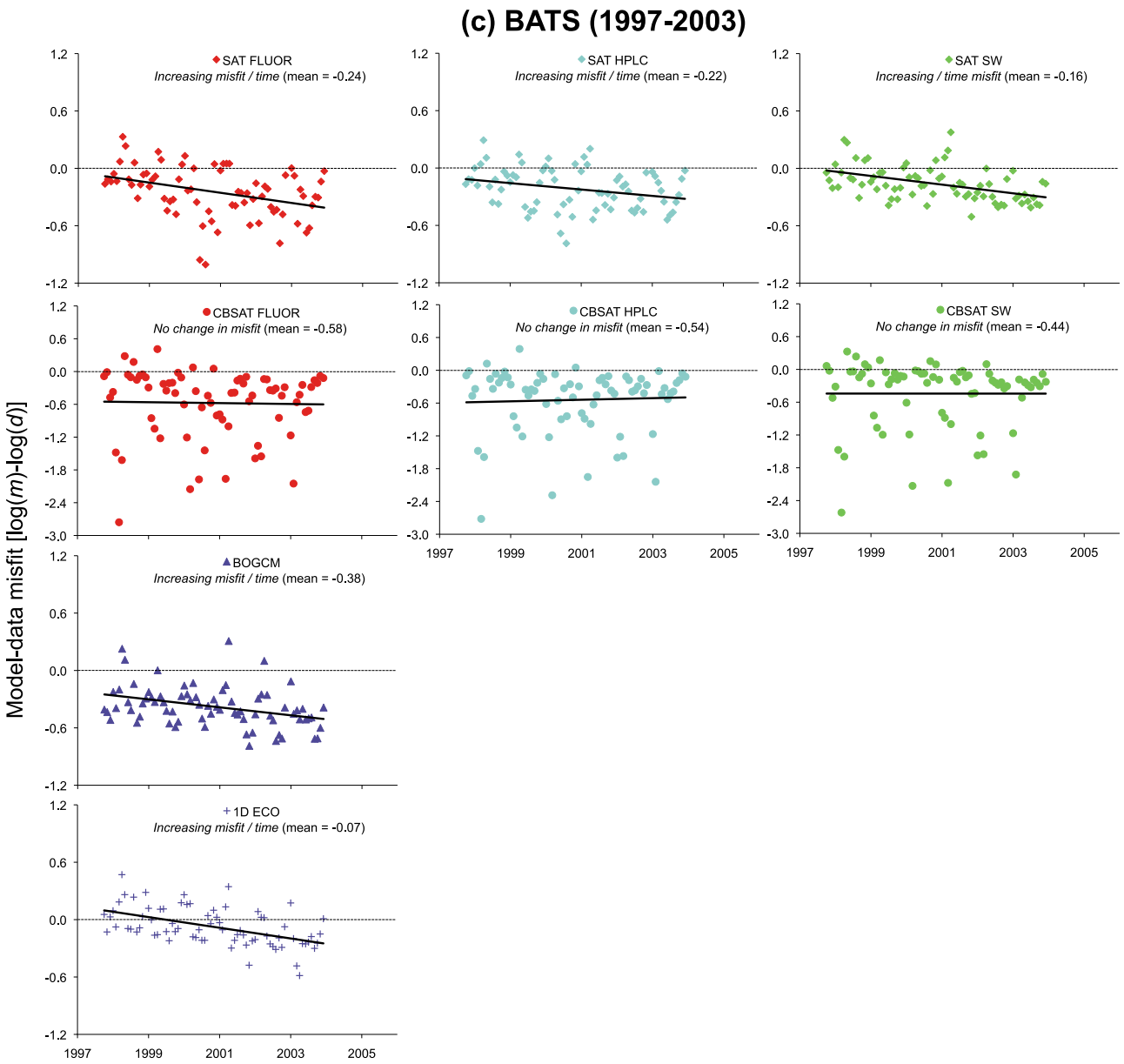

Figure 7. (continued)

also found in the present study. However, while the BATS data time series from 1989 to 2006 showed an increasing trend in NPP, no concomitant trends in SST or MLD occurred. Mesoscale eddies also play a significant role in the enhancement of NPP in the Sargasso Sea where BATS is located [*McGillicuddy et al.*, 2007]. The interannual variability of the winter-time MLD and eddy activity at BATS results in significant interannual changes in new nutrients [*Cianca et al.*, 2007], however, this may not be the case in terms of multidecadal variability.

[42] Other observed parameters at BATS showed concomitant trends with NPP. The increasing trend of both surface and integrated HPLC Chl-*a* was consistent with the increase in NPP. Salinity at 200 m also increased at BATS suggesting a change in water masses in the deeper epipelagic zone. Finally, integrated nitrate plus nitrite (to 150 m) also significantly increased (statistically) although the average annual change was small at only 0.05 μmol m^{-2} year^{-1} (Table 3 and Figure 1).

[43] *Palter et al.* [2005] suggested that the nutrient supply to the North Atlantic Subtropical Gyre is governed by the nutrient concentrations of the underlying Subtropical Mode Water (STMW). The STMW is found in a thick layer of nearly uniform temperature of around 18°C [*Talley and Raymer*, 1982]. During periods of more vigorous STMW production, the mode water layer thickens and the nutrient reservoir of the Gyre becomes depleted; the opposite occurs during periods of reduced STMW formation [*Palter et al.*,

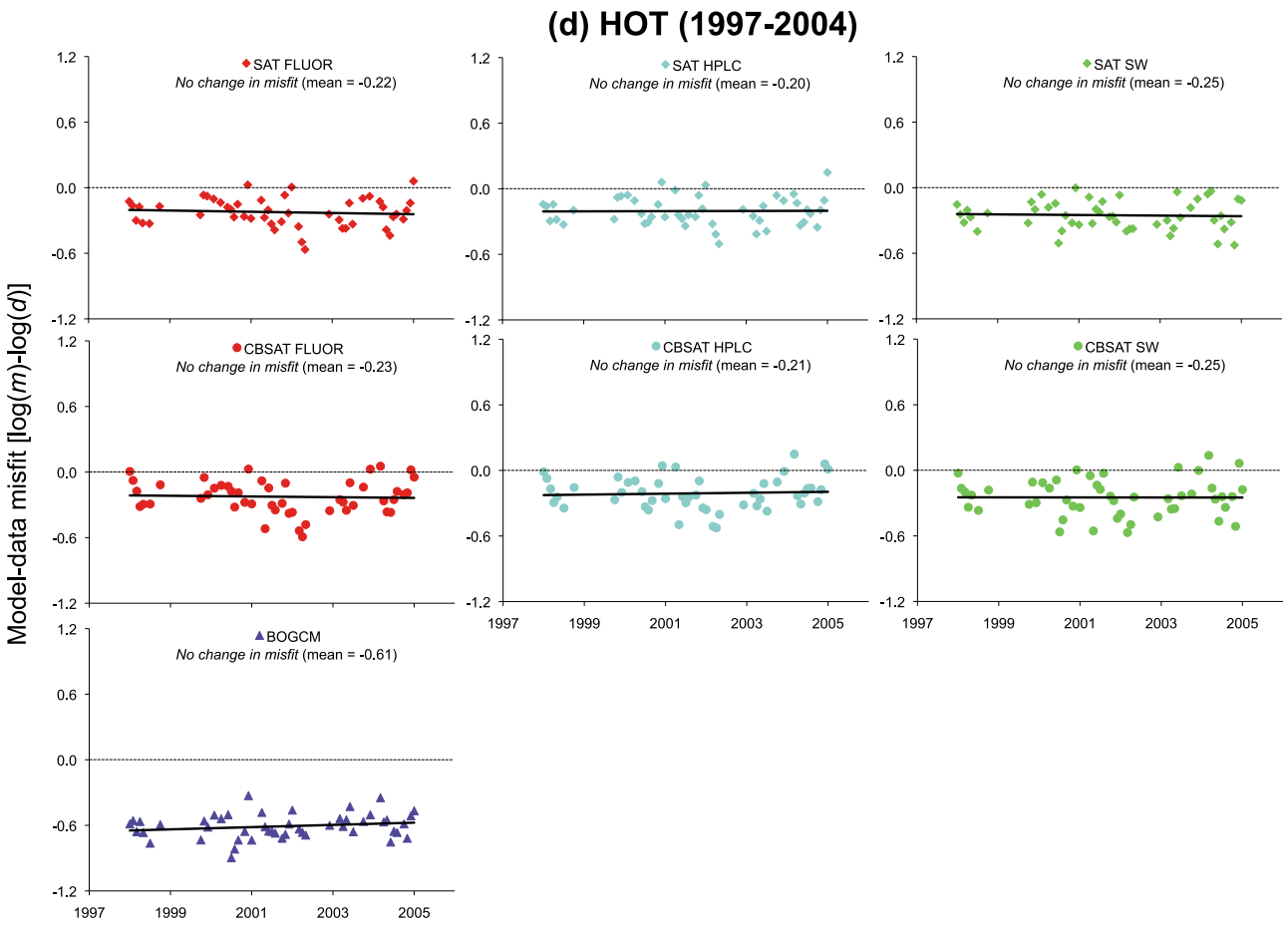

Figure 7. (continued)

2005]. Indeed, *Krause et al.* [2009] reported that the thickness of the STMW layer has been decreasing at BATS since 1997, along with the core of the mode water deepening. However, integrated nitrate plus nitrite (to 150 m) at BATS did not increase substantially (although the increase was still statistically significant) thus the STMW thinning does not appear to result in substantially enhanced nitrate plus nitrite into the euphotic zone on an annual basis to support higher NPP.

[44] *Krause et al.* [2009] reported a 40% decline in integrated biogenic silica and an increase in the abundances of dinoflagellates and prasinophytes in the upper 120 m at BATS; or essentially a reduction in the role of diatoms in the ecosystem. Over the same 15-year time period, the study also reported an increase in density stratification in the upper 200 m. They conclude that a phytoplankton community composition shift occurred which may reflect a change in the biogeochemical processes at BATS [*Krause et al.*, 2009]. Additionally, *Lomas et al.* [2009] showed that more frequent mixing from late-winter storms contributed to pulses of nitrate inputs causing new production that was not accounted for in annual estimates. Indeed, winter storm activity around the BATS station is enhanced during negative phases of NAO [*Bates and Hansell*, 2004]. Therefore, the physical mechanisms that are responsible for the community shift and NPP increase at BATS may be a combination of the thinning STMW and more frequent mixing events during a negative NAO phase. It is also important to note that the increase in prokaryotic phytoplankton (specifically *Synechococcus*) being responsible for the increase in NPP at BATS is counter to conventional theories on the relationship between phytoplankton community composition and productivity [*Lomas et al.*, 2010].

[45] Finally, NPP at BATS was also correlated to the NPGO yet was not correlated to the MEI or the PDO. The absolute magnitude of the correlation coefficients between NPP at BATS and the NAO and NPGO were similar, thus it is likely that either climate index could be used as a proxy for NPP variability at BATS. The fact that the NPGO, a Pacific Ocean climate index, can serve as a proxy for NPP at BATS is not entirely surprising given the teleconnections between Pacific-Atlantic climate shifts [*Müller et al.*, 2008].

4.3.2. HOT

[46] The scenario at HOT may be somewhat more straightforward. A study by *Corno et al.* [2007] reported that ENSO- and PDO-related changes in stratification and vertical

Table 5. Observed and Estimated NPP Among Ocean Color Models at BATS and HOT During the SeaWiFS Time Period at BATS (1997–2003) and HOT (1997–2004)[a]

Station	NPP Source	Mean NPP (mg C m^{-2} day^{-1})	Trend (12-month Moving Average)	Annual Slope (mg C m^{-2} day^{-1} $year^{-1}$)	Overall Change in NPP (mg C m^{-2} day^{-1})	RMSD
BATS (N = 74; 1997–2003)	In situ (SeaWiFS time period)	500.68 (±178)	Increasing	+63.18	+252.71	-
	SAT estimate (fluorometric Chl-*a* based)	323.79 (±207)	Decreasing	−13.88	−55.51	0.39 (±0.10)
	SAT estimate (HPLC Chl-*a* based)	343.01 (±225)	Increasing	+10.78	+43.13	0.33 (±0.09)
	SAT estimate (SeaWiFS Chl-*a* based)	354.79 (±169)	No trend	-	-	0.27 (±0.07)
	CBSAT estimate (fluorometric Chl-*a* based)	231.05 (±214)	Decreasing	−10.97	−43.88	0.90 (±0.30)
	CBSAT estimate (HPLC Chl-*a* based)	237.03 (±201)	Increasing	+30.65	+153.25	0.82 (±0.43)
	CBSAT estimate (SeaWiFS Chl-*a* based)	277.58 (±167)	No trend	-	-	0.77 (±0.39)
HOT (N = 50; 1997–2004)	In situ (SeaWiFS time period)	530.76 (±154)	No trend	-	-	-
	SAT estimate (fluorometric Chl-*a* based)	336.63 (±142)	Decreasing	−9.12	−45.60	0.28 (±0.08)
	SAT estimate (HPLC Chl-*a* based)	346.63 (±70)	No trend	-	-	0.26 (±0.08)
	SAT estimate (SeaWiFS Chl-*a* based)	316.56 (±129)	Decreasing	−4.41	−22.03	0.30 (±0.09)
	CBSAT estimate (fluorometric Chl-*a* based)	336.38 (±150)	Decreasing	−4.09	−24.55	0.30 (±0.15)
	CBSAT estimate (HPLC Chl-*a* based)	351.57 (±70)	No trend	-	-	0.30 (±0.15)
	CBSAT estimate (SeaWiFS Chl-*a* based)	317.10 (±133)	No trend	-	-	0.32 (±0.13)

[a]Linear trends and performance (RMSD ± standard deviation) of SAT and CBSAT models using various measurements of surface Chl-*a*. Lower RMSD values signify higher model skill.

nutrient dynamics were responsible for the increase in NPP at HOT from 1989 to 2004. Their study showed that MLD at HOT was deepening from increased wind-forcing, which in turn caused an increase in the availability of nutrients in the euphotic zone.

[47] In the 1989–2007 HOT data examined here, MLD similarly deepened. Other parameters with concomitant trends as NPP were: increasing fluorometric and HPLC Chl-*a* (both surface and integrated), increasing SST and temperature at 100 m, decreasing salinity at 200 m, and increasing integrated nitrate plus nitrite (to 150 m). The largest increase in NPP at HOT occurred at depths between 45 and 75 m, which is consistent with the integrated NPP increase being caused by subsurface nutrient enhancement. *Bidigare et al.* [2009] suggested that the availability of new nitrogen (in the form of nitrate) was the explanation for the increase in phytoplankton biomass at HOT. Moreover, *Bidigare et al.* [2009] reported that the maximum increase of HPLC measured carotenoid pigments (associated with certain species of chromophyte microalgae) was between 80 and 140 m while concentrations in the upper 40 m were consistently low.

[48] In situ data at HOT was positively correlated to the NPGO and negatively correlated to the PDO and MEI to a much lesser extent ($P < 0.05$) (Figure 3b). *Corno et al.* [2007] suggested that the relationship of NPP at HOT to the ENSO and PDO was somewhat complicated in that it depended on interaction between the two indices. The NPGO is the second mode of SST and sea surface height variability in the North Pacific Gyre whereas PDO is the first mode [*Di Lorenzo et al.*, 2008]. A positive NPGO corresponds to increased open-ocean wind stress curl anomalies in the North Pacific Subtropical Gyre [*Di Lorenzo et al.*, 2008] thus increasing MLD at HOT. Consequently, it is not surprising that the NPGO was positively correlated to NPP at HOT and may serve as a useful indicator of NPP variability.

[49] Climatic regime shifts in the North Pacific Ocean have been reported during the 20th century, occurring most recently around 1989 and then again in 1998 [*McFarlane et al.*, 2000; *Chavez et al.*, 2003; *Kasai and Ono*, 2007]. These regime shifts were identified by changes in SST of the eastern Pacific and were accompanied by changes in the dynamics of key commercial fishery stocks throughout the North Pacific [*McFarlane et al.*, 2000; *Chavez et al.*, 2003]. The shift in 1989 represents the entry into a peak warm period in the eastern Pacific while the shift in 1998 represents the beginning of a cool period. During this cool phase, wind-forcing is stronger and sea surface height would be lower at the HOT station [*Chavez et al.*, 2003; *Di Lorenzo et al.*, 2008]. *Bidigare et al.* [2009] suggested that the rapid ENSO/PDO shift in 1998 caused the upper ocean at the HOT station to become more weakly stratified and more susceptible to wind-forced mixing. Therefore, the regime shift in 1998 was associated with conditions at HOT that favored increased NPP, which was maintained at above average levels until 2006.

5. Summary and Conclusions

[50] In this paper we assessed the performance of 36 models by examining their ability to estimate the mean, variability, and trends of NPP. Our results derived from just two stations in the subtropical North Atlantic and Pacific Oceans; thus our conclusions cannot necessarily be extrapolated to basin scales. Nonetheless, they are likely to be relevant to marine ecosystems that are similar to the BATS and HOT stations.

[51] The majority of the models applied in this study underestimated NPP at BATS and HOT, unlike in the tropical Pacific where NPP was typically overestimated [*Friedrichs et al.*, 2009]. This difference in the models'

ability to estimate mean NPP based on ecosystem type needs to be further explored so that algorithms can be revised to estimate the magnitude of NPP more accurately on a global scale.

[52] Ocean color models applied to BATS and HOT estimated the sign of the observed NPP trend (or lack thereof) when using HPLC-derived measurements of Chl-*a*. To obtain global estimates of productivity, ocean color models must use satellite-derived Chl-*a*; however, our results demonstrate that using HPLC Chl-*a* rather than SeaWiFS or fluorometric Chl-*a* improved the model estimates of the NPP trends for these two locations. The calibration of ocean color satellite sensors to in situ HPLC as opposed to in situ fluorometric Chl-*a* may enhance the accuracy of NPP ocean color models aimed at estimating marine NPP trends, at least in the subtropical gyres where BATS and HOT are located.

[53] In general, surface physical fields, such as SST, did not show as clear a relationship with depth-integrated NPP as did the deeper physical fields. Deep-layer physical dynamics are likely partially responsible for the increase in NPP at BATS. Thus it is not surprising that ocean color models using only surface fields substantially underestimate the increasing NPP trend at this site. When the shorter post-1997 time series (SeaWiFS era) was examined, the models generally had much greater difficulty estimating the NPP trends, despite the fact that at BATS the observed NPP trend was much stronger during this time period. This reiterates the essential requirement for long-term ocean carbon observations and modeling to fully resolve multiannual variability [e.g., *Thomas et al.*, 2008].

[54] Some recent studies [*Behrenfeld et al.*, 2006; *Polovina et al.*, 2008] have reported a decrease in subtropical surface Chl-*a* measured by satellite sensors suggesting that NPP is decreasing. This is clearly not the case at BATS and HOT, both in subtropical gyres, where trends in satellite-derived surface Chl-*a* did not reflect the observed increase in integrated phytoplankton biomass. This discrepancy between satellite-derived surface Chl-*a* and observed trends in NPP at BATS and HOT may not extrapolate to the entire northern subtropical gyres of the Atlantic and Pacific Oceans respectively; a much larger spatial coverage of in situ observations is required to further understand the relationship between satellite-derived surface Chl-*a* and in situ depth-integrated NPP. However, limited temporal and spatial coverage for in situ data remains a major challenge. This suggests the necessity, moving forward over the next decade, for an integrated, multifaceted ocean observing system that incorporates both in situ observations and satellite remote sensing in tandem.

[55] **Acknowledgments.** This research was supported by a grant from the National Aeronautics and Space Agency Ocean Biology and Biogeochemistry program (NNG06GA03G). We are also grateful for the hard work of the many scientists who helped acquire the BATS and HOT database. We also thank Richard Barber and an anonymous reviewer for helpful comments that improved the quality of this paper. This is Contribution 3083 of the Virginia Institute of Marine Science, College of William and Mary.

References

Antoine, D., and A. Morel (1996), Oceanic primary production: I. Adaptation of a spectral light-photosynthesis model in view of application to satellite chlorophyll observations, *Global Biogeochem. Cycles*, *10*, 43–55, doi:10.1029/95GB02831.

Armstrong, R. A. (2006), Optimality-based modeling of nitrogen allocation and photo acclimation in photosynthesis, *Deep Sea Res., Part II*, *53*, 513–531, doi:10.1016/j.dsr2.2006.01.020.

Asanuma, I. (2006), Depth and time resolved primary productivity model examined for optical properties of water, in *Global Climate Change and Response of Carbon Cycle in the Equatorial Pacific and Indian Oceans and Adjacent Landmasses*, *Elsevier Oceanogr. Ser.*, *73*, 89–106.

Aumont, O., and L. Bopp (2006), Globalizing results from ocean in situ iron fertilization studies, *Global Biogeochem. Cycles*, *20*, GB2017, doi:10.1029/2005GB002591.

Bates, N. R. (2001), Interannual variability of oceanic CO_2 and biogeochemical properties in the western North Atlantic subtropical gyre, *Deep Sea Res., Part II*, *48*, 1507–1528, doi:10.1016/S0967-0645(00)00151-X.

Bates, N. R., and D. A. Hansell (2004), Temporal variability of excess nitrate in the subtropical mode water of the North Atlantic Ocean, *Mar. Chem.*, *84*, 225–241, doi:10.1016/j.marchem.2003.08.003.

Behrenfeld, M. J., and P. G. Falkowski (1997), Photosynthetic rates derived from satellite-based chlorophyll concentration, *Limnol. Oceanogr.*, *42*, 1–20, doi:10.4319/lo.1997.42.1.0001.

Behrenfeld, M. J., E. Boss, D. A. Siegel, and D. M. Shea (2005), Carbon-based ocean productivity and phytoplankton physiology from space, *Global Biogeochem. Cycles*, *19*, GB1006, doi:10.1029/2004GB002299.

Behrenfeld, M. J., et al. (2006), Climate-driven trends in contemporary ocean productivity, *Nature*, *444*, 752–755, doi:10.1038/nature05317.

Bennington, V., G. A. McKinley, S. Dutkiewicz, and D. Ullman (2009), What does chlorophyll variability tell us about export and air-sea CO_2 flux variability in the North Atlantic?, *Global Biogeochem. Cycles*, *23*, GB3002, doi:10.1029/2008GB003241.

Bidigare, R. R., et al. (2009), Subtropical ocean ecosystem structure changes forced by North Pacific climate variations, *J. Plankton Res.*, *31*, 1131–1139, doi:10.1093/plankt/fbp064.

Buitenhuis, E., R. Rivkin, S. Sailley, and C. Le Quéré (2010), Global biogeochemical fluxes through microzooplankton, *Global Biogeochem. Cycles*, doi:10.1029/2009GB003601, in press.

Campbell, J., et al. (2002), Comparison of algorithms for estimating ocean primary production from surface chlorophyll, temperature, and irradiance, *Global Biogeochem. Cycles*, *16*(3), 1035, doi:10.1029/2001GB001444.

Carr, M. E. (2002), Estimation of potential productivity in Eastern Boundary Currents using remote sensing, *Deep Sea Res., Part II*, *49*, 59–80, doi:10.1016/S0967-0645(01)00094-7.

Carr, M. E., et al. (2006), A comparison of global estimates of marine primary production from ocean color, *Deep Sea Res., Part II*, *53*, 741–770, doi:10.1016/j.dsr2.2006.01.028.

Chavez, F. P., J. Ryan, S. E. Lluch-Cota, and M. C. Ñiquen (2003), From anchovies to sardines and back: Multidecadal change in the Pacific Ocean, *Science*, *299*, 217–221, doi:10.1126/science.1075880.

Cianca, A., P. Helmke, B. Mourino, M. Rueda, O. Llinas, and S. Neuer (2007), Decadal analysis of hydrography and in situ nutrient budgets in the western and eastern North Atlantic subtropical gyre, *J. Geophys. Res.*, *112*, C07025, doi:10.1029/2006JC003788.

Corno, G., D. M. Karl, M. J. Church, R. M. Letelier, R. Lukas, R. R. Bidigare, and M. R. Abbott (2007), Impact of climate forcing on ecosystem processes in the North Pacific Subtropical Gyre, *J. Geophys. Res.*, *112*, C04021, doi:10.1029/2006JC003730.

Di Lorenzo, E., et al. (2008), North Pacific Gyre Oscillation links ocean climate and ecosystem change, *Geophys. Res. Lett.*, *35*, L08607, doi:10.1029/2007GL032838.

Doney, S. C. (1996), A synoptic atmospheric surface forcing data set and physical upper ocean model for the U.S. JGOFS Bermuda Atlantic Time-Series Study (BATS) site, *J. Geophys. Res.*, *101*, 25,615–25,634, doi:10.1029/96JC01424.

Doney, S. C., S. Yeager, G. Danabasoglu, W. G. Large, and J. C. McWilliams (2007), Mechanisms governing interannual variability of upper-ocean temperature in a global ocean hindcast simulation, *J. Phys. Oceanogr.*, *37*, 1918–1938, doi:10.1175/JPO3089.1.

Doney, S. C., I. Lima, J. K. Moore, K. Lindsay, M. J. Behrenfeld, T. K. Westberry, N. Mahowald, D. M. Glover, and T. Takahashi (2009), Skill metrics for confronting global upper ocean ecosystem-biogeochemistry models against field and remote sensing data, *J. Mar. Syst.*, *76*, 95–112, doi:10.1016/j.jmarsys.2008.05.015.

Ducklow, H., S. C. Doney, and D. K. Steinberg (2009), Contributions of long-term research and time-series observations to marine ecology and biogeochemistry, *Annu. Rev. Mar. Sci.*, *1*, 279–302, doi:10.1146/annurev.marine.010908.163801.

Dunne, J. P., R. A. Armstrong, A. Gnanadesikan, and J. L. Sarmiento (2005), Empirical and mechanistic models for the particle export ratio, *Global Biogeochem. Cycles*, *19*, GB4026, doi:10.1029/2004GB002390.

Dutkiewicz, S., M. Follows, and J. Bragg (2009), Modeling the coupling of ocean ecology and biogeochemistry, *Global Biogeochem. Cycles*, *23*, GB4017, doi:10.1029/2008GB003405.

Eppley, R., E. Steward, M. Abbott, and U. Heyman (1985), Estimating ocean primary production from satellite chlorophyll: Introduction to regional differences and statistics for the southern California Bight, *J. Plankton Res.*, *7*, 57–70, doi:10.1093/plankt/7.1.57.

Friedrichs, M. A. M., et al. (2009), Assessing the uncertainties of model estimates of primary productivity in the tropical Pacific Ocean, *J. Mar. Syst.*, *76*, 113–133, doi:10.1016/j.jmarsys.2008.05.010.

Gregg, W. W. (2008), Assimilation of SeaWiFS ocean chlorophyll data into a three-dimensional global ocean model, *J. Mar. Syst.*, *69*, 205–225, doi:10.1016/j.jmarsys.2006.02.015.

Gregg, W. W., and N. W. Casey (2007), Modeling coccolithophores in the global oceans, *Deep Sea Res., Part II*, *54*, 447–477, doi:10.1016/j.dsr2.2006.12.007.

Howard, K. L., and J. A. Yoder (1997), Contribution of the sub-tropical oceans to global primary production, in *Proceedings of COSPAR Colloquium on Space Remote Sensing of Subtropical Oceans*, edited by C.-T. Liu, pp. 157–168, Pergamon, New York.

Jolliff, J. K., J. C. Kindle, I. Shulman, B. Penta, M. A. M. Friedrichs, R. Helber, and R. A. Arnone (2009), Summary diagrams for coupled hydrodynamic-ecosystem model skill assessment, *J. Mar. Syst.*, *76*, 64–82, doi:10.1016/j.jmarsys.2008.05.014.

Kameda, T., and J. Ishizaka (2005), Size-fractionated primary production estimated by a two-phytoplankton community model applicable to ocean color remote sensing, *J. Oceanogr.*, *61*, 663–672, doi:10.1007/s10872-005-0074-7.

Kasai, H., and T. Ono (2007), Has the 1998 regime shift also occurred in the oceanographic conditions and lower trophic ecosystem of the Oyashio region?, *J. Oceanogr.*, *63*, 661–669, doi:10.1007/s10872-007-0058-x.

Krause, J. W., M. W. Lomas, and D. M. Nelson (2009), Biogenic silica at the Bermuda Atlantic Time-series Study site in the Sargasso Sea: Temporal changes and their inferred controls based on a 15-year record, *Global Biogeochem. Cycles*, *23*, GB3004, doi:10.1029/2008GB003236.

Laws, E., P. Falkowski, W. Smith, H. Ducklow, and J. McCarthy (2000), Temperature effects on export production in the open ocean, *Global Biogeochem. Cycles*, *14*, 1231–1246, doi:10.1029/1999GB001229.

Lomas, M. W., F. Lipshultz, D. M. Nelson, J. W. Krause, and N. R. Bates (2009), Biogeochemical responses to late-winter storms in the Sargasso Sea: I. Pulses of primary and new production, *Deep Sea Res., Part I*, *56*, 843–860, doi:10.1016/j.dsr.2008.09.002.

Lomas, M. W., D. K. Steinberg, T. Dickey, C. A. Carlson, N. B. Nelson, R. H. Condon, and N. R. Bates (2010), Increased ocean carbon export in the Sargasso Sea linked to climate variability is countered by its enhanced mesopelagic attenuation, *Biogeosciences*, *7*, 57–70, doi:10.5194/bg-7-57-2010.

Marra, J., C. Ho, and C. Trees (2003), An alternative algorithm for the calculation of primary productivity from remote sensing data, *LDEO Tech. Rep., LDEO-2003-1*.

McClain, C. R., G. C. Feldman, and S. B. Hooker (2004), An overview of the SeaWiFS project and strategies for producing a climate research quality global ocean bio-optical time series, *Deep Sea Res., Part II*, *51*, 5–42, doi:10.1016/j.dsr2.2003.11.001.

McFarlane, G. A., J. R. King, and R. J. Beamish (2000), Have there been recent changes in climate? Ask the fish, *Prog. Oceanogr.*, *47*, 147–169, doi:10.1016/S0079-6611(00)00034-3.

McGillicuddy, D. J., et al. (2007), Eddy/wind interactions stimulate extraordinary mid-ocean plankton blooms, *Science*, *316*, 1021–1026, doi:10.1126/science.1136256.

Mélin, F. (2003), Potentiel de la télédétection pour l'analyse des propriétésoptiques du systèmeocéan-atmosphèreet application à l'estimation de la photosynthèse phytoplanctonique, Ph.D. diss., Univ. Paul Sabatier, Toulouse, France.

Milutinović, S., M. J. Behrenfeld, J. A. Johannessen, and T. Johannessen (2009), Sensitivity of remote sensing–derived phytoplankton productivity to mixed layer depth: Lessons from the carbon-based productivity model, *Global Biogeochem. Cycles*, *23*, GB4005, doi:10.1029/2008GB003431.

Moore, J. K., S. C. Doney, and K. Lindsay (2004), Upper ocean ecosystem dynamics and iron cycling in a global three-dimensional model, *Global Biogeochem. Cycles*, *18*, GB4028, doi:10.1029/2004GB002220.

Morel, A., and S. Maritorena (2001), Biol.-optical properties of oceanic waters: A reappraisal, *J. Geophys. Res.*, *106*, 7163–7180, doi:10.1029/2000JC000319.

Müller, W. A., C. Frankignoul, and N. Chouaib (2008), Observed decadal tropical Pacific-North Atlantic teleconnections, *Geophys. Res. Lett.*, *35*, L24810, doi:10.1029/2008GL035901.

Ondrusek, M. E., R. R. Bidigare, K. Waters, and D. M. Karl (2001), A predictive model for estimating rates of primary production in the subtropical North Pacific Ocean, *Deep Sea Res., Part II*, *48*, 1837–1863, doi:10.1016/S0967-0645(00)00163-6.

Palter, J., M. Lozier, and R. Barber (2005), The effect of advection on the nutrient reservoir in the North Atlantic subtropical gyre, *Nature*, *437*, 687–692, doi:10.1038/nature03969.

Pinckney, J., R. Papa, and R. Zingmark (1994), Comparison of high-performance liquid chromatographic, spectrophotometric, and fluorometric methods for determining chlorophyll-*a* concentrations in estuarine sediments, *J. Microbiol. Methods*, *19*, 59–66, doi:10.1016/0167-7012(94)90026-4.

Polovina, J. J., E. A. Howell, and M. Abecassis (2008), Ocean's least productive waters are expanding, *Geophys. Res. Lett.*, *35*, L03618, doi:10.1029/2007GL031745.

Reuer, M., B. Barnett, M. L. Bender, P. G. Falkowski, and M. B. Hendricks (2007), New estimates of Southern Ocean biological production rates from O_2/Ar ratios and the triple isotope composition of O_2, *Deep Sea Res., Part I*, *54*, 951–974, doi:10.1016/j.dsr.2007.02.007.

Roman, M., H. Adolf, M. Landry, L. Maldin, D. K. Steinberg, and R. Zhang (2002), Estimates of oceanic mesozooplankton production: A comparison using the Bermuda and Hawaii time-series data, *Deep Sea Res., Part II*, *49*, 175–192, doi:10.1016/S0967-0645(01)00099-6.

Saba, V. S., J. R. Spotila, F. P. Chavez, and J. A. Musick (2008), Bottom-up and climatic forcing on the worldwide population of leatherback turtles, *Ecology*, *89*, 1414–1427, doi:10.1890/07-0364.1.

Salihoglu, B., V. Garçon, A. Oschlies, and M. W. Lomas (2008), Influence of nutrient utilization and remineralization stoichiometry on phytoplankton species and carbon export: A modeling study at BATS, *Deep Sea Res., Part I*, *55*, 73–107, doi:10.1016/j.dsr.2007.09.010.

Scardi, M. (2001), Advances in neural network modeling of phytoplankton primary production, *Ecol. Modell.*, *146*, 33–45, doi:10.1016/S0304-3800(01)00294-0.

Siegel, D. A., et al. (2001), Biol.-optical modeling of primary production on regional scales: The Bermuda BioOptics project, *Deep Sea Res., Part II*, *48*, 1865–1896, doi:10.1016/S0967-0645(00)00167-3.

Smyth, T. J., G. H. Tilstone, and S. B. Groom (2005), Integration of radiative transfer into satellite models of ocean primary production, *J. Geophys. Res.*, *110*, C10014, doi:10.1029/2004JC002784.

Stow, C. A., J. Jolliff, D. J. McGillicuddy Jr., S. C. Doney, J. I. Allen, M. A. M. Friedrichs, K. A. Rose, and P. Wallhead (2009), Skill assessment for coupled biological/physical models of marine systems, *J. Mar. Syst.*, *76*, 4–15, doi:10.1016/j.jmarsys.2008.03.011.

Talley, L. D., and M. E. Raymer (1982), Eighteen degree water variability, *J. Mar. Res.*, *40*, 757–775.

Tang, S., C. Chen, H. Zhan, and T. Zhang (2008), Determination of ocean primary productivity using support vector machines, *Int. J. Remote Sens.*, *29*, 6227–6236, doi:10.1080/01431160802175355.

Thomas, H., A. E. Friederike Prowe, I. D. Lima, S. C. Doney, R. Wanninkhof, R. J. Greatbatch, U. Schuster, and A. Corbière (2008), Changes in the North Atlantic Oscillation influence CO_2 uptake in the North Atlantic over the past 2 decades, *Global Biogeochem. Cycles*, *22*, GB4027, doi:10.1029/2007GB003167.

Tjiputra, J. F., M. Assmann, M. Bentsen, I. Bethke, O. H. Otterå, C. Sturm, and C. Heinze (2010), Bergen Earth system model (BCM-C): Model description and regional climate-carbon cycle feedbacks assessment, *Geosci. Model Dev.*, *3*, 123–141, doi:10.5194/gmd-3-123-2010.

Uitz, J., H. Yannick, F. Bruyant, M. Babin, and H. Caustre (2008), Relating phytoplankton photophysiological properties to community stucture on large scales, *Limnol. Oceanogr.*, *53*, 614–630.

Vichi, M., S. Masina, and N. Pinardi (2007a), A generalized model of pelagic biogeochemistry for the global ocean ecosystem. Part I: Theory, *J. Mar. Syst.*, *64*, 89–109, doi:10.1016/j.jmarsys.2006.03.006.

Vichi, M., S. Masina, and A. Navarra (2007b), A generalized model of pelagic biogeochemistry for the global ocean ecosystem. Part II: Numerical simulations, *J. Mar. Syst.*, *64*, 110–134, doi:10.1016/j.jmarsys.2006.03.014.

Westberry, T. K., M. J. Behrenfeld, E. Siegel, and E. Boss (2008), Carbon-based primary productivity modeling with vertically resolved photoacclimation, *Global Biogeochem. Cycles*, *22*, GB2024, doi:10.1029/2007GB003078.

Yool, A., A. P. Martin, C. Fernandez, and D. R. Clark (2007), The significance of nitrification for oceanic new production, *Nature*, *447*, 999–1002, doi:10.1038/nature05885.

Zainuddin, M., K. Saitoh, and S. Saitoh (2008), Albacore (*Thunnus alalunga*) fishing ground in relation to oceanographic conditions in the western North Pacific Ocean using remotely sensed satellite data, *Fish. Oceanogr.*, *17*, 61–73, doi:10.1111/j.1365-2419.2008.00461.x.

D. Antoine and A. Morel, Laboratoire d'Océanographie de Villefranche, UMR 7093, Université Pierre et Marie Curie, Paris 06, CNRS, Villefranche-sur-Mer, F-06238, France.

R. A. Armstrong, School of Marine and Atmospheric Sciences, State University of New York at Stony Brook, Stony Brook, NY 11794-5000, USA.

I. Asanuma, Tokyo University of Information Sciences, 4-1-1, Onaridai, Wakaba, Chiba, 265-8501, Japan.

O. Aumont, Laboratoire d'Océanographie: Expérimentation et Approche Numérique, IPSL, UPMC, IRD, CNRS, Centre IRD de Bretagne, BP 70, F-29280 Plouzané, France.

N. R. Bates and M. W. Lomas, Bermuda Institute of Ocean Sciences, St. George's GE01, Bermuda.

M. J. Behrenfeld and T. K. Westberry, Department of Botany and Plant Pathology, Oregon State University, Cordley Hall 2082, Corvallis, OR 97331-2902, USA.

V. Bennington and G. A. McKinley, Department of Atmospheric and Ocean Sciences, University of Wisconsin-Madison, 1225 W. Dayton St., Madison, WI 53706, USA.

L. Bopp, Laboratoire des Sciences du Climat et de l'Environnement, IPSL, CEA, UVSQ, CNRS, CEN de Saclay, Bât 701 l'Orme des Merisiers, F-91191 Gif-sur-Yvette, France.

J. Bruggeman, Department of Theoretical Biology, Faculty of Earth and Life Sciences, Vrije University of Amsterdam, de Boelelaan 1085, NL-1081 HV Amsterdam, Netherlands.

E. T. Buitenhuis, Laboratory for Global Marine and Atmospheric Chemistry, School of Environmental Sciences, University of East Anglia, Norwich NR4 7TJ, UK.

M-E. Carr, Columbia Climate Center, Earth Institute, Columbia University, 405 Low Library, MC 4335, 535 West 116th St., New York, NY 10027, USA.

M. J. Church and D. M. Karl, Department of Oceanography, School of Ocean and Earth Science and Technology, University of Hawai'i at Mãnoa, 1000 Pope Rd., MSB 205, Honolulu, HI 96822 USA.

A. M. Ciotti, Campus Experimental do Litoral Paulista, UNESP, PraçaInfante Dom Henrique S/N, São Vicente, São Paulo CEP 11330-900, Brazil.

S. C. Doney and I. Lima, Department of Marine Chemistry and Geochemistry, Woods Hole Oceanographic Institution, 266 Woods Hole Rd., Woods Hole, MA 02543-1050, USA.

M. Dowell, N. Hoepffner, and F. Mélin, Joint Research Centre, European Commission, I-21020 Ispra, Italy.

J. Dunne, Geophysical Fluid Dynamics Laboratory, NOAA, Princeton University Forrestal Campus, 201 Forrestal Rd., Princeton, NJ 08540-6649, USA.

S. Dutkiewicz, Earth, Atmospheric and Planetary Sciences, Massachusetts Institute of Technology, 77 Massachusetts Ave., Cambridge, MA 02139-4307, USA.

M. A. M. Friedrichs, Virginia Institute of Marine Science, College of William and Mary, PO Box 1346, Gloucester Point, VA 23062-1346, USA

W. Gregg, NASA Global Modeling and Assimilation Office, Goddard Space Flight Center, Code 610.1, Greenbelt, MD 20771, USA.

K. J. W. Hyde and J. O'Reilly, National Marine Fisheries Services Narragansett Laboratory, NOAA, 28 Tarzwell Dr., Narragansett, RI 02882, USA.

J. Ishizaka, Hydrospheric Atmospheric Research Center, Nagoya University, F3-1 (200) Furo-cho, Chikusa-ku, Nagoya 464-8601, Japan.

T. Kameda, Group of Oceanography, National Research Institute of Far Seas Fisheries, 5-7-1 Shimizu Orido, Shizuoka 424-8633, Japan.

J. Marra, Geology Department, Brooklyn College of the City University of New York, 2900 Bedford Ave., Brooklyn, NY 11210, USA.

J. K. Moore, Department of Earth System Science, University of California, Croul Hall, Irvine, CA 92697-3100, USA.

V. S. Saba, Atmospheric and Oceanic Sciences Program, Princeton University, 300 Forrestal Rd., Sayre Hall, Princeton, NJ 08544, USA. (vsaba@princeton.edu)

B. Salihoglu, Institute of Marine Sciences, Middle East Technical University, PO Box 28, 33731 Erdemli-Mersin, Turkey.

M. Scardi, Department of Biology, University of Rome 'Tor Vergata', Via della Ricerca Scientifica, I-00133 Roma, Italy.

T. J. Smyth, Plymouth Marine Laboratory, Prospect Place, Plymouth, Devon PL1 3DH, UK.

S. Tang, Freshwater Institute, Fisheries and Oceans Canada, 501 University Crescent, Winnipeg, MB R3T 2N6, Canada.

J. Tjiputra, Geophysical Institute, University of Bergen, Allegaten 70, N-5007 Bergen, Norway.

J. Uitz, Scripps Institution of Oceanography, University of California, San Diego, 9500 Gilman Dr., La Jolla, CA 92093, USA.

M. Vichi, Centro Euro-Mediterraneo per i Cambiamenti Climatici, Istituto Nazionale di Geofisica e Vulcanologia, Via Aldo Moro 44, I-40127 Bologna, Italy.

K. Waters, Coastal Services Center, NOAA, 2234 South Hobson Ave., Charleston, SC 29405-2413, USA.

A. Yool, National Oceanography Centre, Southampton, European Way, Southampton SO14 3ZH, UK.