2002

# Comparison of algorithms for estimating ocean primary production from surface chlorophyll, temperature, and irradiance

Janet Campbell

David Antoine

Robert Armstrong

Kevin Arrigo

William Balch

*See next page for additional authors*



Citation/Publisher Attribution
Campbell, J. W., D. Antoine, R. Armstrong, K. Arrigo, W. Balch, R. Barber, M. Behrenfeld, R. Bidigare, J. Bishop, M.-E. Carr, W. Esaias, P. Falkowski, N. Hoepffner, R. Iverson, D. Kiefer, S. Lohrenz, J. Marra, A. Morel, J. Ryan, V. Vedernikov, K. Waters, C. Yentsch, and J. Yoder, Comparison of algorithms for estimating ocean primary production from surface chlorophyll, temperature, and irradiance, *Global Biogeochem. Cycles*, 16(3), doi: 10.1029/2001GB001444, 2002.
Available at: https://doi.org/10.1029/2001GB001444

## Authors

Janet Campbell, David Antoine, Robert Armstrong, Kevin Arrigo, William Balch, Richard Barber, Michael Behrenfeld, Robert Bidigare, James Bishop, Mary-Elena Carr, Wayne Esaias, Paul Falkowski, Nicholas Hoepffner, Richard Iverson, David Keifer, Steven Lohrenz, John Marra, Andre Morel, John Ryan, Vladimir Vederinkov, Kirk Waters, Charles Yentsch, and James A. Yoder

GLOBAL BIOGEOCHEMICAL CYCLES, VOL. 16, NO. 3, 1035, 10.1029/2001GB001444, 2002

# Comparison of algorithms for estimating ocean primary production from surface chlorophyll, temperature, and irradiance

Janet Campbell,[1] David Antoine,[2] Robert Armstrong,[3] Kevin Arrigo,[4] William Balch,[5] Richard Barber,[6] Michael Behrenfeld,[7] Robert Bidigare,[8] James Bishop,[9] Mary-Elena Carr,[10] Wayne Esaias,[7] Paul Falkowski,[11] Nicolas Hoepffner,[12] Richard Iverson,[13] Dale Kiefer,[14] Steven Lohrenz,[15] John Marra,[16] Andre Morel,[2] John Ryan,[17] Vladimir Vedernikov,[18] Kirk Waters,[19] Charles Yentsch,[5] and James Yoder[20]

Received 21 May 2001; revised 8 March 2002; accepted 8 March 2002; published 17 July 2002.

[1] Results of a single-blind round-robin comparison of satellite primary productivity algorithms are presented. The goal of the round-robin exercise was to determine the accuracy of the algorithms in predicting depth-integrated primary production from information amenable to remote sensing. Twelve algorithms, developed by 10 teams, were evaluated by comparing their ability to estimate depth-integrated daily production (IP, mg C $m^{-2}$) at 89 stations in geographically diverse provinces. Algorithms were furnished information about the surface chlorophyll concentration, temperature, photosynthetic available radiation, latitude, longitude, and day of the year. Algorithm results were then compared with IP estimates derived from $^{14}C$ uptake measurements at the same stations. Estimates from the best-performing algorithms were generally within a factor of 2 of the $^{14}C$-derived estimates. Many algorithms had systematic biases that can possibly be eliminated by reparameterizing underlying relationships. The performance of the algorithms and degree of correlation with each other were independent of the algorithms' complexity. *INDEX TERMS:* 4894 Oceanography: Biological and Chemical: Instruments and techniques; 4275 Oceanography: General: Remote sensing and electromagnetic processes (0689); 4805 Oceanography: Biological and Chemical: Biogeochemical cycles (1615); 4806 Oceanography: Biological and Chemical: Carbon cycling; 4853 Oceanography: Biological and Chemical: Photosynthesis; *KEYWORDS:* primary productivity, algorithms, ocean color, remote sensing, satellite, chlorophyll

## 1. Introduction

[2] Global maps of the upper-ocean chlorophyll concentration are now being generated routinely by satellite ocean color sensors. These multispectral sensors are able to map the chlorophyll concentration, a measure of phytoplankton biomass, by detecting spectral shifts in upwelling radiance. As the chlorophyll concentration increases, blue light is increasingly absorbed, and thus less is scattered back into space. Although global coverage can nominally be achieved every 1–2 days, the actual temporal resolution is reduced to

[1]Ocean Process Analysis Laboratory, Institute for Study of Earth, Oceans, and Space, University of New Hampshire, Durham, New Hampshire, USA.
[2]Laboratoire d'Océanographie de Villefranche, CNRS/INSU and Universite Pierre et Marie Curie, Villefranche sur Mer, France.
[3]Marine Sciences Research Center, Stony Brook University, Stony Brook, New York, USA.
[4]Department of Geophysics, Stanford University, Stanford, California, USA.
[5]Bigelow Laboratory for Ocean Sciences, West Boothbay Harbor, Maine, USA.
[6]Marine Laboratory, Nicholas School of the Environment and Earth Sciences, Duke University, Beaufort, North Carolina, USA.
[7]NASA Goddard Space Flight Center, Greenbelt, Maryland, USA.
[8]Department of Oceanography, University of Hawaii, Honolulu, Hawaii, USA.


0886-6236/02/2001GB001444

[9]Lawrence Berkeley National Laboratory, Berkeley, California, USA.
[10]Jet Propulsion Laboratory, Pasadena, California, USA.
[11]Department of Geology and Institute of Marine and Coastal Sciences, Rutgers University, New Brunswick, New Jersey, USA.
[12]Institute for Environment and Sustainability, Joint Research Centre of the European Commission, Ispra, Italy.
[13]Department of Oceanography, Florida State University, Tallahassee, Florida, USA.
[14]Department of Biological Sciences, University of Southern California, Los Angeles, California, USA.
[15]Department of Marine Science, University of Southern Mississippi, Stennis Space Center, Mississippi, USA.
[16]Lamont-Doherty Earth Observatory, Columbia University, Palisades, New York, USA.
[17]Monterey Bay Aquarium and Research Institute, Moss Landing, California, USA.
[18]P. P. Shirshov Institute of Oceanology, Moscow, Russia.
[19]NOAA Coastal Services Center, Charleston, South Carolina, USA.
[20]Graduate School of Oceanography, University of Rhode Island, Narragansett, Rhode Island, USA.

**Table 1.** Algorithm Testing Subcommittee of NASA's Ocean Primary Productivity Working Group[a]

| Participant | Affiliation |
|---|---|
| Robert Armstrong | Stony Brook University |
| Richard T. Barber | Duke University |
| James Bishop | Lawrence Berkeley National Laboratory |
| Janet W. Campbell | University of New Hampshire |
| Mary-Elena Carr | Jet Propulsion Laboratory |
| Wayne E. Esaias | NASA Goddard Space Flight Center |
| Richard Iverson | Florida State University |
| Charles S. Yentsch | Bigelow Laboratory for Ocean Sciences |

[a] These individuals were responsible for conducting the primary productivity algorithm round-robin experiment. They agreed not to participate by testing algorithms of their own.

~5–10 days because of cloud cover. Nevertheless, the coverage afforded by satellite remote sensing is vastly greater than that obtainable by any other means.

[3] A principal use of the global ocean chlorophyll data is to estimate oceanic primary production [*Behrenfeld et al.*, 2001]. The mathematical models or procedures for estimating primary production from satellite data are known as primary productivity algorithms. In the early days of the Coastal Zone Color Scanner (CZCS), simple statistical relationships were proposed for calculating primary production from the surface chlorophyll concentration [e.g., *Smith and Baker*, 1978; *Eppley et al.*, 1985]. Such empirically derived algorithms are still considered useful when applied to annually averaged data [*Iverson et al.*, 2000], but they are not sufficiently accurate to estimate production at seasonal timescales. The surface chlorophyll concentration explains only ~30% of the variance in primary production at the scale of a single station [*Balch et al.*, 1992; *Campbell and O'Reilly*, 1988].

[4] Over the past 2 decades, scientists have sought to improve algorithms by combining the satellite-derived chlorophyll data with other remotely sensed fields, such as sea surface temperature (SST) and photosynthetic available radiation (PAR). These algorithms incorporate models of the photosynthetic response of phytoplankton to light, temperature, and other environmental variables, and some also incorporate models of the vertical distribution of these properties within the euphotic zone [*Balch et al.*, 1989; *Morel*, 1991; *Platt and Sathyendranath*, 1993; *Howard*, 1995; *Antoine and Morel*, 1996a; *Behrenfeld and Falkowski*, 1997a; *Ondrusek et al.*, 2001]. Algorithms have been used to estimate global oceanic primary production from CZCS data [*Antoine and Morel*, 1996b; *Longhurst et al.*, 1995; *Behrenfeld and Falkowski*, 1997a; *Howard and Yoder*, 1997], and more recently from Sea-viewing Wide Field-of-view Sensor (SeaWiFS) data [*Behrenfeld et al.*, 2001]. Global maps of the average daily primary production for varying periods (weeks, months, and years) are now being produced from Moderate Resolution Imaging Spectroradiometer (MODIS) data.

[5] While many of the photosynthetic responses (to light, temperature, etc.) are commonly represented, model-based algorithms differ with respect to structure and computational complexity [*Behrenfeld and Falkowski*, 1997b]. Models may be similar in structure but require different parameters depending on whether they describe daily, hourly, or instantaneous production, and even where these aspects are similar, algorithms often yield different results because of differences in their parameterization. *Balch et al.* [1992] evaluated a variety of algorithms (both empirical and model based), using in situ productivity measurements from a large globally distributed data set, and found that they generally accounted for <50% of the variance in primary production.

[6] In January 1994 the National Aeronautics and Space Administration (NASA) convened an Ocean Primary Productivity Working Group with the goal of developing one or more "consensus" algorithms to be applied to satellite ocean color data. The working group initiated a series of round-robin experiments to evaluate and compare primary productivity algorithms. The approach was to use in situ data to test the ability of algorithms to predict depth-integrated daily production (IP, mg C $m^{-2}$) based on information amenable to remote sensing. It was decided to compare algorithm performances with one another and with estimates based on $^{14}C$ incubations.

[7] Our understanding of primary productivity in the ocean is largely based on the assimilation of inorganic carbon from $^{14}C$ techniques [*Longhurst et al.*, 1995], and thus it was considered appropriate to compare the algorithm estimates with $^{14}C$-based estimates. However, it was recognized that the $^{14}C$-based estimates are themselves subject to error [*Peterson*, 1980; *Fitzwater et al.*, 1982; *Richardson*, 1991]. The $^{14}C$ incubation technique measures photosynthetic carbon fixation within a confined volume of seawater, and there are no methods for absolute calibration of bottle incubations [*Balch*, 1997]. Furthermore, there is no universally

**Table 2.** Participant Teams Whose Algorithms Were Tested in Round Robin

| Participant | Affiliation |
|---|---|
| David Antoine and Andre Morel | Laboratoire de Physique et Chimie Marines |
| William Balch and Bruce Bowler | Bigelow Laboratory for Ocean Sciences |
| Michael Behrenfeld and Paul Falkowski | NASA Goddard Space Flight Center/Rutgers University |
| Nicolas Hoepffner | Joint Research Centre of the European Commission |
| Dale Kiefer | University of Southern California |
| Steven Lohrenz | University of Southern Mississippi |
| John Marra | Lamont-Doherty Earth Observatory |
| Vladimir Vedernikov | P.P. Shirshov Institute of Oceanology |
| Kirk Waters and Bob Bidigare | NOAA Coastal Services Center/University of Hawaii |
| James Yoder and John Ryan | University of Rhode Island/Monterey Bay Aquarium and Research Institute |

**Table 3.** Data Sets Used to Test Algorithms[a]

| Data Set | Region | $n$ | IP | Latitude | | Longitude | | Months | Years |
|---|---|---|---|---|---|---|---|---|---|
| AMERIEZ | Antarctica | 7 | 315 | −66 | −58 | −51 | −38 | March–Nov. | 1983–1986 |
| SUPER | North Pacific | 10 | 688 | 50 | 53 | −145 | −145 | June–Aug. | 1987–1988 |
| EqPac nonequator | Tropical Pacific | 13 | 724 | −12 | 12 | −140 | −140 | Feb.–Sept. | 1996–1996 |
| NABE | Northeast Atlantic | 12 | 1027 | 46 | 47 | −20 | −19 | April–May | 1989–1989 |
| EqPac equator | Equatorial Pacific | 8 | 1170 | 0 | 0 | −140 | −140 | Feb.–Oct. | 1996–1996 |
| Arabian Sea | Arabian Sea | 12 | 1192 | 10 | 19 | 57 | 67 | March–Aug. | 1995–1995 |
| PROBES | Bering Sea | 9 | 1203 | 55 | 58 | −167 | −164 | April–June | 1979–1981 |
| MARMAP | Northwest Atlantic | 8 | 1560 | 40 | 43 | −71 | −67 | Aug.–Sept. | 1981–1981 |
| Palmer LTER | Antarctica | 10 | 1795 | −65 | −65 | −64 | −64 | Feb.–Dec. | 1991–1992 |

[a] Columns are the region, number of stations ($n$), average depth-integrated daily production (IP, mg C $m^{-2}$), and ranges in latitude, longitude, months, and years.

accepted method for measuring and verifying vertically integrated production derived from discrete bottle measurements. Despite this fact, here we treat the $^{14}C$-based estimates as "truth" and refer to the differences between algorithm-derived and $^{14}C$-derived estimates as "errors." In all statistical analyses, however, the two are recognized as being subject to error.

[8] Participation in the round robin was solicited through a widely distributed "Dear Colleague" letter. A central ground rule was that the algorithms tested would be identified only by code numbers. The first round-robin experiment involved data from only 25 stations and was thus limited in scope. It was decided that a more comprehensive second round was needed. In this paper, we present results of the second round-robin experiment involving data from 89 stations with wide geographic coverage. Round two was open to all participants of round one, as well as to others who had responded positively to the initial invitation.

[9] The following questions were addressed: (1) How do algorithm estimates of primary production derived strictly from surface information compare with estimates derived from $^{14}C$ incubation methods? (2) How does the error in satellite-derived chlorophyll concentration affect the accuracy of the primary productivity algorithms? (3) Are there regional differences in the performance of algorithms? (4) How do algorithms compare with each other in terms of complexity vis-a-vis performance?

## 2. Methods

[10] A subcommittee of NASA's Ocean Primary Productivity Working Group was formed to administer the round-robin experiments (Table 1), and there were 10 participant teams (Table 2) who volunteered to test their algorithms. A test data set was assembled to be used for evaluating algorithms. Algorithm developers were provided with only the information accessible to spaceborne sensors, and they subsequently returned predictions of integral production at each station. Results were compared with estimates derived from the $^{14}C$ incubations and with the results of other algorithms.

### 2.1. Test Data

[11] Data from 89 stations were obtained from nine sources (Table 3), representing diverse geographic regions and a variety of measurement techniques. The acquired station data included the downwelling photosynthetic available radiation incident on the water surface (daily PAR between 400 and 700 nm, in mol photons $m^{-2}$) and measurements of the chlorophyll concentration, temperature, and PAR at discrete depths in the upper water column. From the profile data, we determined surface chlorophyll ($B_{sfc}$, mg Chl $m^{-3}$), sea surface temperature (SST, °C), and the euphotic depth or 1% light level ($Z_m$, m).

[12] In addition, we were provided $^{14}C$-based estimates of the daily primary production ($P_i$, mg C $m^{-3}$) at discrete depths $Z_i$ ranging from the surface to the 1% light depth. "Measured" integral production, $IP_{meas}$, was computed for each station by trapezoidal integration, using the formula

$$IP_{meas} = P_1 Z_1 + \sum_{i=2}^{m} 0.5(P_{i-1} + P_i)(Z_i - Z_{i-1}), \qquad (1)$$

where the number of depths ($m$) varied among stations. Integral chlorophyll (IB, mg Chl $m^{-2}$) was also computed over the same layer by a similar formula. The surface information provided to the algorithm developers and other information not provided (e.g., $IP_{meas}$, IB, and $Z_m$) are listed in Table 4.

[13] The measurement methods were consistent within each data set but differed between data sets. The equatorial Pacific (EqPac [*Barber et al.*, 1996]), North Atlantic (NABE [*Ducklow and Harris*, 1993]), and Arabian Sea [*Barber et al.*, 2001] data were from the Joint Global Ocean Flux Study (JGOFS) [*Knudson et al.*, 1989; *Chipman et al.*, 1993] process studies. Primary production measurements from these campaigns were based on 24-hour, in situ incubations, in accordance with JGOFS protocols. The SUPER data set [*Welschmeyer et al.*, 1993] also used 24-hour, in situ incubations. Simulated in situ incubations were used to produce the Antarctic Marine Ecosystem Research at the Ice Edge Zone (AMERIEZ) data (24-hour incubations [*Smith and Nelson*, 1990]), the PROBES data (dawn-to-dusk incubations [*Codispoti et al.*, 1982]), and the Marine Resources Monitoring, Assessment, and Prediction (MARMAP) data (6-hour incubations scaled by daily PAR measurements [*O'Reilly et al.*, 1987]). The Palmer data from the Long-Term Ecological Research (LTER) site [*Moline and Prezelin*, 1997] were based on 90-min incubations in photosynthetrons [e.g., *Prezelin and Glover*, 1991] that were then scaled to estimate daily rates. This methodological diversity

**Table 4.** In Situ Data Used to Test Primary Productivity Algorithms[a]

| Station | Region | Lat. | Long. | Date | SST | PAR | Sfc. Chl | $Z_m$ | $IB_{meas}$ | $IP_{meas}$ |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | EqPac nonequator | 12 | −140 | 5 Feb. 1996 | 25.9 | 25.2 | 0.114 | 85 | 14.2 | 309 |
| 2 | | 5 | −140 | 13 Feb. 1996 | 28.5 | 30.7 | 0.179 | 71 | 14.0 | 427 |
| 3 | | 3 | −140 | 15 Feb. 1996 | 28.4 | 36.2 | 0.210 | 71 | 16.1 | 670 |
| 4 | | −2 | −140 | 1 March 1996 | 28.6 | 27.0 | 0.132 | 106 | 16.5 | 586 |
| 5 | | −5 | −140 | 4 March 1996 | 28.7 | 36.4 | 0.170 | 106 | 19.9 | 675 |
| 6 | | 12 | −140 | 11 Aug. 96 | 28.4 | 41.6 | 0.063 | 79 | 10.0 | 319 |
| 7 | | 5 | −140 | 19 Aug. 1996 | 27.5 | 33.7 | 0.284 | 79 | 22.1 | 561 |
| 8 | | 3 | −140 | 22 Aug. 1996 | 27.0 | 36.9 | 0.230 | 79 | 17.3 | 637 |
| 9 | | 2 | −140 | 24 Aug. 1996 | 23.8 | 38.9 | 0.355 | 79 | 27.7 | 1630 |
| 10 | | −2 | −140 | 3 Sept. 1996 | 25.5 | 35.2 | 0.223 | 79 | 20.4 | 1047 |
| 11 | | −3 | −140 | 6 Sept. 1996 | 25.3 | 36.1 | 0.274 | 79 | 26.2 | 1362 |
| 12 | | −5 | −140 | 8 Sept. 1996 | 25.9 | 30.8 | 0.225 | 79 | 21.7 | 861 |
| 13 | | −12 | −140 | 13 Sept. 1996 | 26.4 | 30.9 | 0.135 | 79 | 14.6 | 323 |
| 14 | EqPac equator | 0 | −140 | 23 Feb. 1996 | 28.3 | 14.6 | 0.227 | 71 | 18.9 | 513 |
| 15 | | 0 | −140 | 24 Feb. 1996 | 28.3 | 43.8 | 0.247 | 71 | 17.6 | 867 |
| 16 | | 0 | −140 | 29 Aug. 1996 | 24.8 | 36.6 | 0.372 | 79 | 28.8 | 1399 |
| 17 | | 0 | −140 | 30 Aug. 1996 | 24.9 | 33.2 | 0.257 | 79 | 25.9 | 1041 |
| 18 | | 0 | −140 | 14 Oct. 1996 | 25.0 | 33.5 | 0.240 | 82 | 27.9 | 1573 |
| 19 | | 0 | −140 | 16 Oct. 1996 | 25.3 | 34.6 | 0.218 | 82 | 25.3 | 1373 |
| 20 | | 0 | −140 | 18 Oct. 1996 | 25.2 | 32.7 | 0.247 | 82 | 27.5 | 1432 |
| 21 | | 0 | −140 | 20 Oct. 1996 | 25.1 | 32.7 | 0.242 | 82 | 23.1 | 1163 |
| 22 | PROBES | 55 | −165 | 17 April 1979 | 3.5 | 26.3 | 2.21 | 32 | 78.7 | 1090 |
| 23 | | 56 | −167 | 18 April 1979 | 4.0 | 28.8 | 1.92 | 35 | 63.2 | 891 |
| 24 | | 57 | −166 | 22 April 1979 | 3.4 | 32.4 | 7.04 | 27 | 134.1 | 1576 |
| 25 | | 55 | −165 | 6 May 1979 | 5.5 | 27.3 | 8.86 | 14 | 207.8 | 2306 |
| 26 | | 55 | −166 | 20 May 1979 | 5.0 | 33.4 | 4.05 | 28 | 113.1 | 2022 |
| 27 | | 58 | −164 | 8 June 1979 | 8.4 | 50.1 | 1.06 | 32 | 111.9 | 1309 |
| 28 | | 55 | −167 | 15 April 1981 | 3.9 | 28.2 | 0.96 | 33 | 29.5 | 364 |
| 29 | | 55 | −167 | 1 June 1981 | 6.9 | 36.9 | 4.34 | 18 | 82.1 | 1053 |
| 30 | | 56 | −166 | 5 June 1981 | 7.5 | 34.3 | 0.68 | 28 | 33.9 | 214 |
| 31 | SUPER | 50 | −145 | 3 June 1987 | 7.2 | 36.7 | 0.282 | 59 | 21.2 | 595 |
| 32 | | 53 | −145 | 9 June 1987 | 6.9 | 26.4 | 0.659 | 57 | 36.8 | 671 |
| 33 | | 50 | −145 | 15 June 1987 | 7.7 | 48.2 | 0.344 | 60 | 24.4 | 913 |
| 34 | | 50 | −145 | 18 June 1987 | 7.4 | 21.5 | 0.531 | 60 | 27.0 | 1541 |
| 35 | | 50 | −145 | 20 Sept. 1987 | 11.7 | 35.5 | 0.402 | 56 | 18.9 | 887 |
| 36 | | 50 | −145 | 8 May 1988 | 5.6 | 29.1 | 0.270 | 89 | 26.5 | 366 |
| 37 | | 50 | −145 | 27 May 1988 | 7.0 | 36.5 | 0.166 | 69 | 12.1 | 446 |
| 38 | | 53 | −145 | 5 Aug. 1988 | 11.5 | 23.5 | 0.129 | 81 | 15.8 | 360 |
| 39 | | 53 | −145 | 19 Aug. 1988 | 11.8 | 20.2 | 0.191 | 68 | 14.8 | 479 |
| 40 | | 50 | −145 | 25 Aug. 1988 | 12.0 | 39.8 | 0.297 | 60 | 16.0 | 621 |
| 41 | AMERIEZ | −58 | −38 | 18 Nov. 1983 | −0.1 | 19.9 | 2.97 | 57 | 146.3 | 502 |
| 42 | | −60 | −38 | 21 Nov. 1983 | −1.4 | 40.2 | 0.43 | 65 | 29.5 | 457 |
| 43 | | −60 | −38 | 23 Nov. 1983 | −1.3 | 50.4 | 0.41 | 63 | 44.7 | 273 |
| 44 | | −60 | −40 | 27 Nov. 1983 | −0.7 | 13.9 | 4.70 | 22 | 123.1 | 633 |
| 45 | | −65 | −48 | 11 March 1986 | −1.7 | 9.4 | 0.11 | 70 | 7.1 | 148 |
| 46 | | −66 | −49 | 16 March 1986 | −1.8 | 7.3 | 0.08 | 70 | 7.9 | 105 |
| 47 | | −65 | −51 | 23 March 1986 | −1.8 | 7.6 | 0.07 | 70 | 5.7 | 88 |
| 48 | Palmer LTER | −65 | −64 | 10 Dec. 1991 | −0.4 | 45.2 | 0.72 | 40 | 55.3 | 1259 |
| 49 | | −65 | −64 | 16 Dec. 1991 | −0.1 | 64.2 | 1.17 | 25 | 92.7 | 3207 |
| 50 | | −65 | −64 | 28 Dec. 1991 | 2.2 | 67.1 | 2.52 | 19 | 184.8 | 6308 |
| 51 | | −65 | −64 | 4 Jan. 1992 | 0.7 | 42.9 | 11.59 | 12 | 157.2 | 3894 |
| 52 | | −65 | −64 | 16 Jan. 1992 | 0.5 | 33.1 | 0.84 | 35 | 51.5 | 994 |
| 53 | | −65 | −64 | 24 Jan. 1992 | −0.2 | 14.1 | 0.73 | 29 | 19.9 | 187 |
| 54 | | −65 | −64 | 3 Feb. 1992 | 0.4 | 35.5 | 1.06 | 26 | 20.2 | 220 |
| 55 | | −65 | −64 | 10 Feb. 1992 | 0.3 | 24.1 | 0.69 | 47 | 34.0 | 673 |
| 56 | | −65 | −64 | 17 Feb. 1992 | 0.4 | 17.7 | 3.45 | 69 | 63.2 | 340 |
| 57 | | −65 | −64 | 27 Feb. 1992 | 0.2 | 40.8 | 2.43 | 35 | 58.8 | 868 |
| 58 | Arabian Sea | 19 | 67 | 19 March 1995 | 25.5 | 53.4 | 0.541 | 40 | 20.1 | 1327 |
| 59 | | 10 | 65 | 24 March 1995 | 29.0 | 51.4 | 0.077 | 73 | 11.1 | 602 |
| 60 | | 14 | 65 | 27 March 1995 | 27.8 | 52.3 | 0.078 | 64 | 10.3 | 679 |
| 61 | | 16 | 62 | 31 March 1995 | 27.1 | 51.2 | 0.188 | 40 | 14.6 | 841 |
| 62 | | 17 | 60 | 3 April 1995 | 26.9 | 56.3 | 0.218 | 39 | 21.8 | 1145 |
| 63 | | 18 | 58 | 6 April 1995 | 26.9 | 57.3 | 0.164 | 38 | 10.8 | 651 |
| 64 | | 19 | 67 | 22 July 1995 | 28.0 | 40.2 | 0.308 | 61 | 20.8 | 886 |
| 65 | | 14 | 65 | 28 July 1995 | 27.4 | 52.5 | 0.521 | 46 | 23.8 | 1455 |
| 66 | | 10 | 65 | 31 July 1995 | 28.0 | 48.8 | 0.583 | 48 | 25.7 | 1542 |
| 67 | | 16 | 62 | 4 Aug. 1995 | 25.8 | 53.7 | 0.436 | 48 | 21.1 | 1522 |

**Table 4.** (continued)

| Station | Region | Lat. | Long. | Date | SST | PAR | Sfc. Chl | $Z_m$ | $IB_{meas}$ | $IP_{meas}$ |
|---|---|---|---|---|---|---|---|---|---|---|
| 68 | | 18 | 58 | 11 Aug. 1995 | 23.2 | 49.6 | 1.358 | 27 | 26.0 | 2141 |
| 69 | | 18 | 57 | 12 Aug. 1995 | 20.7 | 51.4 | 0.569 | 27 | 17.8 | 1518 |
| 70 | NABE | 47 | −20 | 25 April 1989 | 12.6 | 49.9 | 0.565 | 59 | 34.4 | 944 |
| 71 | | 47 | −20 | 26 April 1989 | 12.6 | 19.6 | 0.908 | 61 | 57.9 | 876 |
| 72 | | 47 | −19 | 27 April 1989 | 12.6 | 13.8 | 0.748 | 65 | 52.3 | 682 |
| 73 | | 47 | −20 | 29 April 1989 | 12.6 | 30.4 | 1.061 | 53 | 47.3 | 910 |
| 74 | | 46 | −20 | 30 April 1989 | 12.7 | 45.9 | 0.807 | 55 | 45.1 | 1286 |
| 75 | | 46 | −20 | 1 May 1989 | 12.7 | 13.9 | 0.879 | 59 | 36.5 | 781 |
| 76 | | 47 | −20 | 2 May 1989 | 12.5 | 53.6 | 1.066 | 50 | 48.6 | 1387 |
| 77 | | 47 | −20 | 3 May 1989 | 12.7 | 24.8 | 1.135 | 49 | 32.9 | 915 |
| 78 | | 47 | −20 | 4 May 1989 | 12.5 | 13.7 | 1.107 | 50 | 64.2 | 852 |
| 79 | | 47 | −20 | 5 May 1989 | 12.4 | 50.3 | 1.274 | 49 | 61.6 | 1402 |
| 80 | | 46 | −19 | 6 May 1989 | 13.1 | 27.1 | 0.710 | 49 | 83.8 | 1031 |
| 81 | | 46 | −19 | 8 May 1989 | 13.0 | 21.8 | 1.724 | 40 | 82.2 | 1253 |
| 82 | MARMAP | 41 | −71 | 27 Aug. 1981 | 19.6 | 35.0 | 3.62 | 15 | 47.4 | 1716 |
| 83 | | 43 | −71 | 28 Aug. 1981 | 13.9 | 48.6 | 7.23 | 14 | 68.5 | 3482 |
| 84 | | 43 | −70 | 28 Aug. 1981 | 14.9 | 48.6 | 1.72 | 21 | 51.5 | 1161 |
| 85 | | 42 | −67 | 29 Aug. 1981 | 16.1 | 40.6 | 1.29 | 25 | 37.5 | 691 |
| 86 | | 40 | −69 | 31 Aug. 1981 | 19.4 | 33.2 | 0.52 | 37 | 51.4 | 1248 |
| 87 | | 41 | −70 | 2 Sept. 1981 | 14.4 | 40.3 | 3.76 | 22 | 63.0 | 1864 |
| 88 | | 41 | −70 | 2 Sept. 1981 | 18.7 | 40.3 | 0.64 | 40 | 32.7 | 1412 |
| 89 | | 41 | −71 | 3 Sept. 1981 | 18.0 | 21.1 | 0.59 | 40 | 40.7 | 904 |

[a] Surface and column-integrated data for the 89 stations that were used to test algorithms are given. Values are listed with the number of significant figures provided in the original data sets. Units are SST, °C; PAR, mol photons $m^{-2}$; Surface (Sfc.) Chl, mg Chl $m^{-3}$; 1% light level $Z_m$, m; integral measured chlorophyll $IB_{meas}$, mg Chl $m^{-2}$; and depth-integrated daily production $IP_{meas}$, mg C $m^{-2}$.

introduced a source of variance in the test data, and consequently in the algorithm performance statistics, that was largely confounded with regional effects. However, we accepted the diversity under the premise that a similar diversity might have existed in the data sets used to parameterize algorithms (see section 5).

[14] Integral primary production and surface chlorophyll spanned 2 orders of magnitude in the test data set, while SST and PAR varied over the wide ranges found globally (Figure 1). The symbols shown in Figure 1 denote the various data sets and regions, and these will be used consistently in subsequent figures. The widest ranges in production, biomass, and irradiance and the lowest temperatures were found in the Antarctic data (Palmer LTER and AMERIEZ). There was a general positive correspondence between IP and $B_{sfc}$, and between IP and PAR, but there were no simple empirical relationships useful for algorithm purposes. There was no apparent relationship between IP and SST at temperatures below 20°C, but in the equatorial Pacific and Arabian Sea, where surface temperatures were above 20°C, production decreased with increasing surface temperature.

## 2.2. Evaluation Procedures

[15] The data analysis and evaluation of algorithm results were carried out at the University of New Hampshire (UNH) under the direction of the first author with input from other members of the algorithm testing subcommittee (ATS) (Table 1). The complete test data set was assembled and resident on computers at UNH, but algorithm codes were not exchanged. That is, each participant team was responsible for running its own algorithm code based on input data furnished by the ATS.

[16] The information provided for each station included (1) latitude and longitude to the nearest 0.1°, (2) day of the year, (3) incident daily PAR (mole photons $m^{-2}$), (4) SST (°C), and (5) two values for the surface chlorophyll concentration (mg Chl $m^{-3}$). One of the chlorophyll values (randomly assigned) was the measured surface chlorophyll ($B_{sfc}$), and the other was a simulated satellite-derived chlorophyll ($B_{sat}$) computed as

$$B_{sat} = B_{sfc} 10^{\Delta B}, \quad (2)$$

where $\Delta_B$ was a pseudorandom normal (Gaussian) error with zero mean and standard deviation equal to 0.3. This error represents a factor-of-2 uncertainty in satellite-derived chlorophyll that has been reported for open ocean (Case 1) waters [*O'Reilly et al.*, 1998; *Gordon et al.*, 1985]. The 89 values of $\Delta_B$ were statistically independent.

[17] Participants did not know which chlorophyll was the measured surface chlorophyll, $B_{sfc}$, and which was the corrupted "satellite" chlorophyll, $B_{sat}$. They were asked to return two algorithm estimates of integral production for each station, one for each chlorophyll value. Their results were then "unscrambled" to identify the integral production estimate based on measured chlorophyll, $IP_{alg}$, and that based on $B_{sat}$, $IP_{sat}$.

### 2.2.1. Performance Indices

[18] The performance of each algorithm was based on a log-difference error (Δ) defined as

$$\Delta = \log(IP_{alg}) - \log(IP_{meas}), \quad (3)$$

which is a measure of relative error. Performance indices were the mean ($M$), standard deviation ($S$), and root-mean-square (RMS) of the 89 log-difference errors. Since the units of these indices are decades of log, and not easily

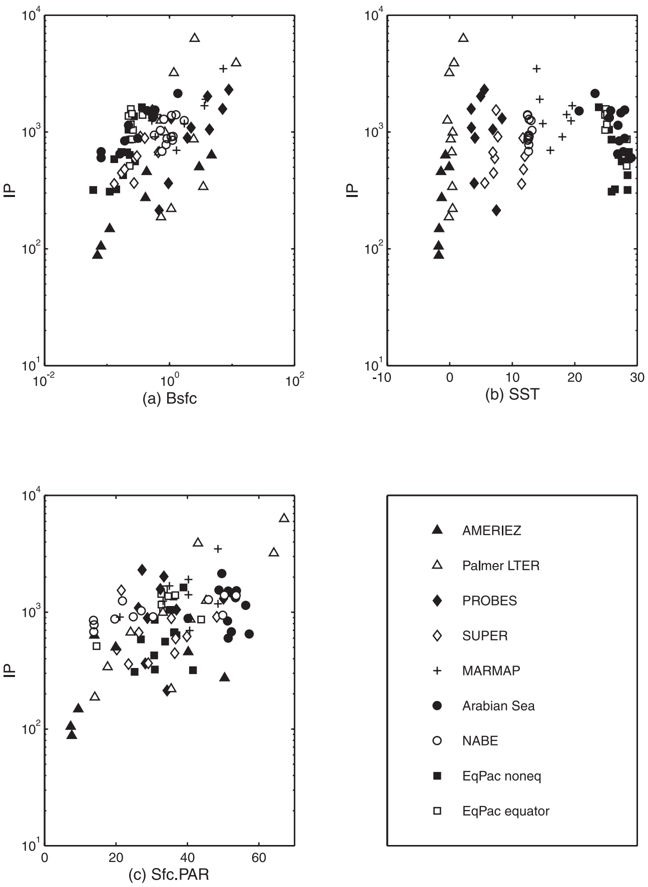


**Figure 1.** Relationships found in in situ data between daily depth-integrated primary production (IP, mg C m$^{-2}$) and properties amenable to remote sensing. (a) Surface chlorophyll concentration ($B_{sfc}$, mg Chl m$^{-3}$). (b) Sea surface temperature (SST, °C). (c) Above-water daily photosynthetic available radiation (PAR, mol photons m$^{-2}$). Symbols shown here will be used consistently in all figures.

translated into absolute terms, we also present three inverse-transformed values:

$$F_{med} = 10^{M}, \tag{4}$$

$$F_{min} = 10^{M-S}, \quad F_{max} = 10^{M+S}. \tag{5}$$

[19] Log-difference errors ($\Delta$) for each algorithm tended to be symmetrically distributed about their mean and approximately normally distributed. Assuming an underlying normal distribution for $\Delta$, $F_{med}$ would be the median value of the ratio

$$F = \frac{IP_{alg}}{IP_{meas}} = 10^{\Delta}, \tag{6}$$

and 68% of the $F$ values would lie within the "one-sigma" range ($F_{min}$ to $F_{max}$).

### 2.2.2. Effect of Errors in the Satellite Chlorophyll

[20] The $IP_{sat}$ estimate based on the simulated satellite chlorophyll, $B_{sat}$, was subject to two errors: the relative error $\Delta$ defined in equation (3) and an error due to the satellite chlorophyll error, $\Delta_B$, which is

$$\Delta_{sat} = \log(IP_{sat}) - \log(IP_{alg}). \tag{7}$$

[21] To investigate the effect of errors in the satellite chlorophyll, $\Delta_{sat}$ was regressed against $\Delta_B$. The slope of this regression yields information about the sensitivity of the IP algorithm to errors in the satellite chlorophyll retrieval. A slope of 1 would indicate that the resulting error in IP is directly proportional to the error in $B_{sat}$, whereas a slope less (greater) than 1 shows less (greater) sensitivity.

### 2.2.3. Regional Analyses

[22] To investigate regional differences, performance indices were computed for each data set separately. Although there were methodological differences between data sets, we treated the different data sets as "regions" for the purpose of this analysis. A two-way analysis of variance (ANOVA) was performed on the $\Delta$ data to determine whether there

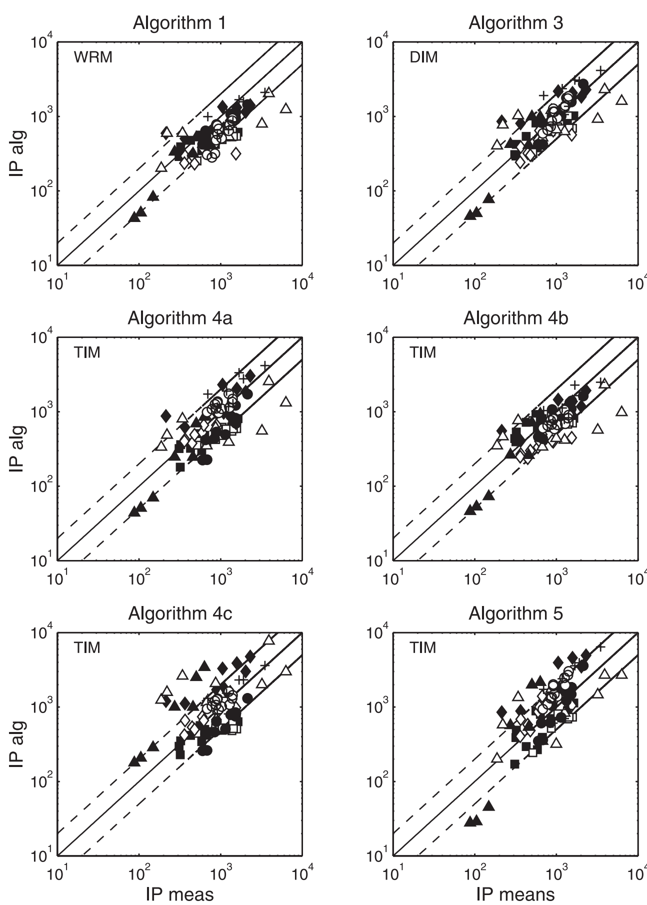


**Figure 2.** Scatterplots of algorithm-derived primary production ($IP_{alg}$, mg C m$^{-2}$) versus production measured in situ ($IP_{meas}$, mg C m$^{-2}$) for 12 algorithms tested. Solid line represents perfect agreement, and dashed lines represent factor-of-2 relative errors. Algorithm category [*Behrenfeld and Falkowski*, 1997b] is shown in upper left corner of each plot.

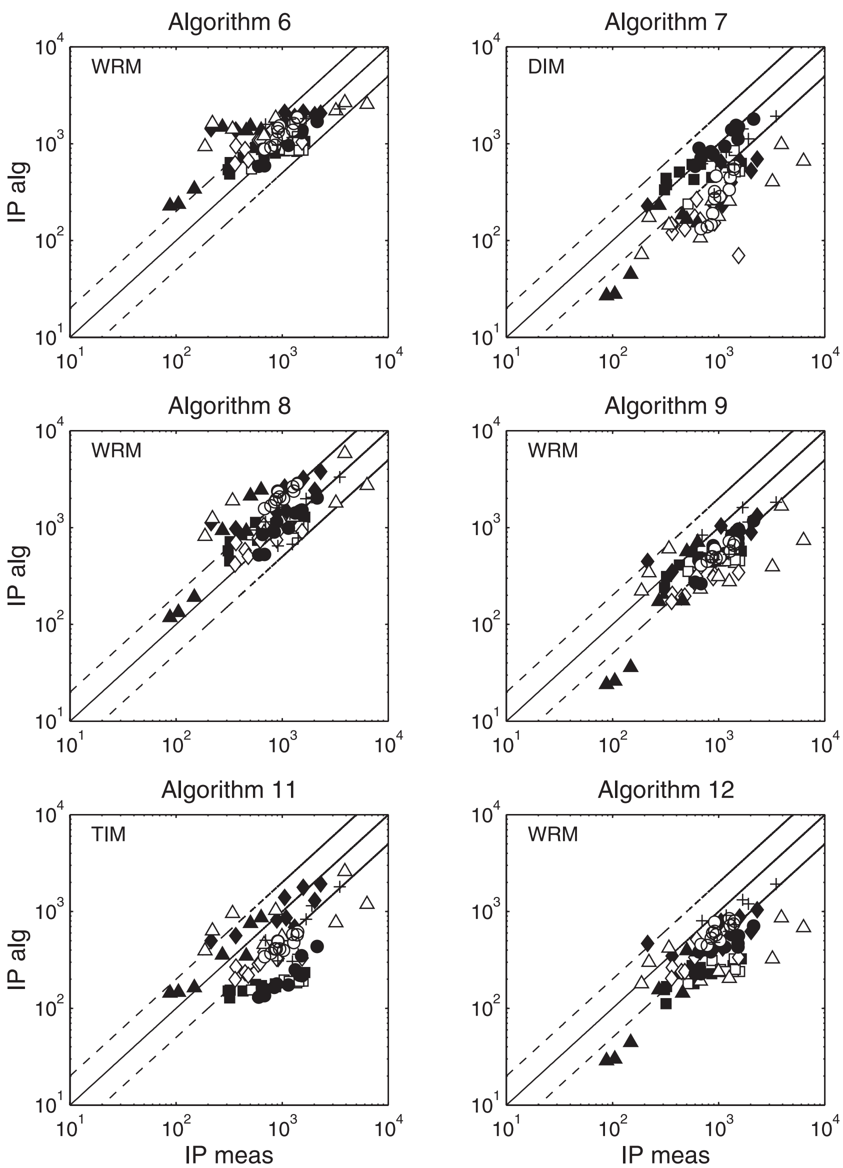


**Figure 2.** (continued)

**Table 5.** Performance Indices for Relative Errors in Algorithms as Compared With Measured IP[a]

| Algorithm | Type | $M$ | $S$ | RMS | $F_{med}$ | $F_{min}$ | $F_{max}$ |
|---|---|---|---|---|---|---|---|
| 1 | WRM | −0.18 | 0.19 | 0.26 | 0.66 | 0.42 | 1.04 |
| 3 | DIM | −0.01 | 0.22 | 0.22 | 0.98 | 0.59 | 1.61 |
| 4a | TIM | −0.09 | 0.25 | 0.26 | 0.81 | 0.46 | 1.43 |
| 4b | TIM | −0.11 | 0.21 | 0.24 | 0.77 | 0.47 | 1.26 |
| 4c | TIM | 0.03 | 0.32 | 0.32 | 1.08 | 0.52 | 2.24 |
| 5 | TIM | 0.05 | 0.28 | 0.28 | 1.11 | 0.58 | 2.12 |
| 6 | WRM | 0.13 | 0.24 | 0.27 | 1.34 | 0.78 | 2.33 |
| 7 | DIM | −0.36 | 0.29 | 0.46 | 0.44 | 0.23 | 0.85 |
| 8 | WRM | 0.15 | 0.23 | 0.27 | 1.40 | 0.82 | 2.38 |
| 9 | WRM | −0.27 | 0.21 | 0.35 | 0.53 | 0.33 | 0.86 |
| 11 | TIM | −0.37 | 0.34 | 0.50 | 0.42 | 0.20 | 0.92 |
| 12 | WRM | −0.36 | 0.24 | 0.43 | 0.44 | 0.25 | 0.75 |

[a] Columns are the mean ($M$), standard deviation ($S$), and root mean square (RMS) of the log-difference error ($\Delta$). The geometric mean and one-sigma range of the ratio ($F = IP_{alg}/IP_{meas}$) are given by $F_{med}$, $F_{min}$, and $F_{max}$, respectively. The algorithm type is based on the categories defined by *Behrenfeld and Falkowski* [1997b].

**Table 6.** Performance Indices When Algorithms Used "Corrupted" Satellite Chlorophyll ($B_{sat}$) Instead of Measured Chlorophyll[a]

| Algorithm | RMS, % | $F_{min}$, % | $F_{max}$, % | Correlation | Slope | Intercept |
|---|---|---|---|---|---|---|
| 1 | 0.31 (+17) | 0.37 (−13) | 1.16 (+12) | 0.99 | 0.50 | −0.003 |
| 3 | 0.28 (+27) | 0.50 (−15) | 1.80 (+11) | 0.98 | 0.62 | −0.008 |
| 4a | 0.30 (+17) | 0.41 (−11) | 1.57 (+10) | 0.99 | 0.57 | −0.001 |
| 4b | 0.29 (+21) | 0.42 (−12) | 1.41 (+12) | 0.99 | 0.56 | −0.001 |
| 4c | 0.35 (+10) | 0.48 (− 8) | 2.40 (+7) | 0.99 | 0.56 | 0.000 |
| 5 | 0.35 (+22) | 0.49 (−16) | 2.42 (+14) | 0.99 | 0.76 | −0.003 |
| 6 | 0.28 (+3) | 0.74 (−5) | 2.35 (+1) | 0.94 | 0.30 | −0.006 |
| 7 | 0.47 (+2) | 0.22 (−4) | 0.95 (+13) | 0.85 | 0.37 | 0.018 |
| 8 | 0.30 (+11) | 0.74 (−10) | 2.56 (+7) | 0.98 | 0.56 | −0.004 |
| 9 | 0.39 (+12) | 0.28 (−13) | 0.96 (+11) | 0.99 | 0.58 | −0.005 |
| 11 | 0.51 (+2) | 0.19 (−3) | 0.96 (+4) | 0.96 | 0.42 | 0.004 |
| 12 | 0.49 (+13) | 0.23 (−10) | 0.82 (+9) | 0.97 | 0.53 | −0.001 |

[a] Columns are the resulting values of RMS, $F_{min}$ and $F_{max}$, and (in parentheses) the percentage change relative to the values in Table 4. The correlation, slope, and intercept are based on regressions of $\Delta_{sat}$ versus $\Delta_B$.

were significant differences in algorithms, regions, and "interactions" between algorithms and regions.

### 2.2.4. Comparing algorithms

[23] To compare algorithms, a correlation coefficient ($r$) was calculated from the log-transformed results, log($IP_{alg}$), for each pair of algorithms. We did not use correlations (or $r^2$ as a measure "percent variance explained") to measure the performance of the algorithms themselves, because high $r^2$ is not a sufficient condition for good agreement. That is, high linear correlation can exist despite systematic errors. In the case of two algorithms, however, we computed correlations and average ratios, and we also examined all pairwise plots to determine the degree of agreement. Results were considered in the context of differences in algorithm structure and complexity.

## 3. Algorithms

[24] Twelve algorithms by 10 teams were tested in the round robin. Each participant team was assigned a code number to identify its algorithm's results in subsequent comparisons. Code numbers ranged from 1 to 12. (Teams 2 and 10 dropped out after receiving code numbers). Team 4 submitted results for three algorithms, which were identified by letters (e.g., 4a, 4b, and 4c).

[25] Although the identity of the algorithm developers is not revealed, in accordance with ground rules, paragraphs describing each of the algorithms are provided in Appendix A. Participation in the round robin was voluntary, and thus the nature of the algorithms tested was purely fortuitous. As it turned out, the algorithms belonged to three of the four major categories of complexity described by *Behrenfeld and Falkowski* [1997b]. The category not represented was that of wavelength-integrated models (WIMs). In this category, time and depth are resolved, but light is not spectrally resolved.

[26] Five algorithms (numbers 1, 6, 8, 9, and 12) were from the wavelength-resolved model (WRM) class, which is the most detailed and highly resolved of all algorithm types. In algorithms of this category, the photosynthetic rate is computed at each depth and at various times throughout the day, based on a spectrally resolved underwater light field. In some cases the vertical profile of chlorophyll was modeled (1, 8, and 9), and in others it was assumed to be uniform (6 and 12). Four of the algorithms (1, 8, 9, and 12) applied a photosynthesis-irradiance relationship to calculate the chlorophyll-specific productivity. They required parameterizations of the maximum light-saturated rate of photosynthesis ($P^B_{max}$) and photosynthetic efficiency ($\alpha$, the slope of the $P^B$ versus $E$ curve in low light). An alternative approach, used by algorithm 6, was to calculate the radiant energy absorbed by the phytoplankton and then apply a quantum yield ($\varphi$, moles carbon fixed per mole photons absorbed) to derive productivity. Temperature is generally used in the parameterization of $P^B_{max}$ or $\varphi_{max}$.

[27] Five algorithms (4a, 4b, 4c, 5, and 11) belonged to the class of time-integrated models (TIMs), in which depth is resolved but both time and wavelength are integrated. These algorithms employed models of the daily production normalized to chlorophyll ($P_z/B_z$) as functions of the daily irradiance $E_z$ at depth $Z$. Such models might resemble the photosynthesis-irradiance models ($P^B$ versus $E$), as was the case for the 4a–4c algorithms, or they might be based on a

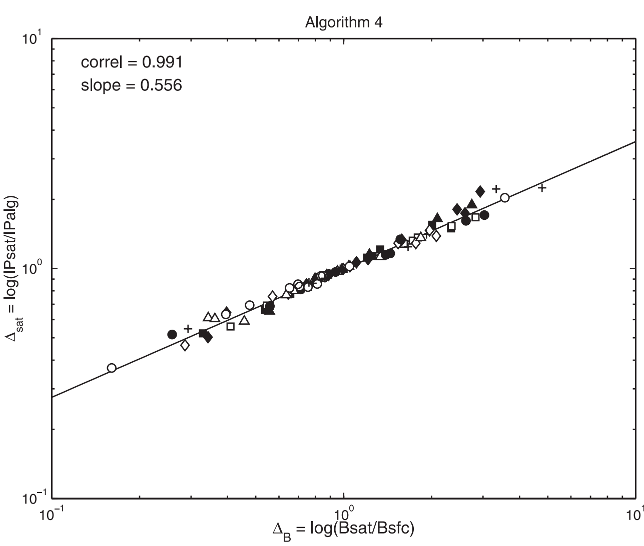


**Figure 3.** Error in satellite-derived production [$\Delta_{sat}$ = log($IP_{sat}/IP_{alg}$)] associated with a simulated error in satellite chlorophyll [$\Delta_B$ = log($B_{sat}/B_{sfc}$)] for algorithm 4b. This is typical of relationships seen for other algorithms (Table 6).

**Table 7.** Performance Indices for Pooled Data Within Each Region (Data Set)

| Region | $n$ | $M$ | $S$ | RMS | $F_{med}$ | $F_{min}$ | $F_{max}$ |
|---|---|---|---|---|---|---|---|
| AMERIEZ | 7 | −0.02 | 0.40 | 0.40 | 0.95 | 0.65 | 1.47 |
| SUPER | 10 | −0.16 | 0.23 | 0.28 | 0.70 | 0.68 | 1.08 |
| EqPac nonequator | 13 | −0.14 | 0.28 | 0.31 | 0.72 | 0.65 | 1.15 |
| NABE | 12 | −0.07 | 0.30 | 0.31 | 0.84 | 0.69 | 1.25 |
| EqPac equator | 8 | −0.32 | 0.38 | 0.50 | 0.48 | 0.49 | 1.06 |
| Arabian Sea | 12 | −0.19 | 0.26 | 0.32 | 0.65 | 0.64 | 1.08 |
| PROBES | 9 | 0.06 | 0.23 | 0.23 | 1.14 | 0.84 | 1.33 |
| MARMAP | 8 | −0.06 | 0.40 | 0.40 | 0.88 | 0.64 | 1.40 |
| Palmer LTER | 10 | −0.12 | 0.25 | 0.27 | 0.76 | 0.69 | 1.13 |
| Total | 89 | −0.12 | 0.31 | 0.33 | 0.76 | 0.65 | 1.22 |

quantum yield approach (5 and 11). Algorithms 5 and 11 modeled the vertical distribution of chlorophyll, whereas the 4a–4c algorithms assumed a uniform chlorophyll profile. Chlorophyll was multiplied by the modeled $P_z/B_z$ and then integrated over the water column to estimate depth-integrated production.

[28] Two algorithms (3 and 7) belonged to the Depth Integrated Model (DIM) category. In these models, there was no vertical resolution of chlorophyll, light, or other properties, but rather IP was derived from integrated (IB) or average euphotic zone chlorophyll, surface PAR, and SST. Details of the individual algorithms are provided in Appendix A.

## 4. Results

### 4.1. Comparisons With $^{14}$C-Based Estimates

[29] The 12 algorithms varied widely in performance (Figure 2 and Table 5). Estimates falling within a factor of 2 of the $^{14}$C-based estimates are points bounded by the dashed lines in Figure 2. Many of the estimates fell within this factor-of-2 range, with the most notable exceptions occurring at Antarctic stations (open and solid triangles) and at PROBES stations in the Bering Sea (solid diamonds).

[30] Performance indices are listed in Table 5. As a benchmark, RMS values of <0.3 indicate agreement within a factor of 2. The RMS values comprise a random error (indexed by $S$) and a systematic error or bias, $M$. Most algorithms exhibited large biases as indicated by nonzero values of $M$, which translated to median ratios, $F_{med}$, ranging from 0.42 (algorithm 11) to 1.4 (algorithm 8). If the biases could be eliminated, the RMS error would equal $S$, in which case 10 of the 12 algorithms would be within a factor of 2 ($S < 0.3$). It may be possible to eliminate biases by reparameterizing the underlying relationships between production, chlorophyll, and light. The sensitivity of the algorithms to model parameterization may be seen by comparing results for algorithms 4a, 4b, and 4c, which differed only in their parameterization of $P^B_{opt}$ (the TIM equivalent of $P^B_{max}$).

### 4.2. Effect of $B_{sat}$ Error

[31] Errors in the satellite chlorophyll algorithm were simulated using a random number generator that introduced a factor-of-2 error in $B_{sat}$. Considering the magnitude of the $B_{sat}$ errors, the resulting increases in $IP_{sat}$ errors (Table 6) were remarkably small. Numbers in parentheses are the relative changes in algorithm performance compared with their corresponding values in Table 5. The RMS differences increased between 3 and 27%, chiefly due to increases in $S$, as reflected by the expanded one-sigma range of $F$. Algorithm 6 seemed the least sensitive to the chlorophyll errors, while algorithms 3 and 5 showed the greatest sensitivity.

[32] Regressions of $\Delta_{sat} = \log(IP_{sat}/IP_{alg})$ versus $\Delta_B = \log(B_{sat}/B_{sfc})$ provided additional insight concerning the sensitivity of IP estimates to errors in $B_{sat}$. The results for algorithm 4b shown in Figure 3 are typical. In all but three cases, correlations between $\Delta_{sat}$ and $\Delta_B$ were >0.97; all were >0.85. These high correlations reflect the deterministic nature of the relationships between $IP_{alg}$ and $B_{sfc}$. For any algorithm, if $IP_{alg}$ were directly proportional to $B_{sfc}$, then the regression of $\Delta_{sat}$ versus $\Delta_B$ would have a slope of 1. Instead, the slopes ranged from 0.30 to 0.76, indicating that errors in $B_{sfc}$ produced less-than-proportionate errors in IP. This is due, in part, to the nonlinearity of IP with respect to chlorophyll, which affects the depth of integration as well as the light-harvesting capacity of the phytoplankton. Most slopes fell between 0.5 and 0.6, which is consistent with several studies [*Eppley et al.*, 1985; *Morel and Berthon*, 1989; *Morel*, 1988], which found that IP varies approximately as $\sqrt{B_{sfc}}$. The algorithm having the smallest slope (algorithm 6) was the one least affected by errors in $B_{sfc}$ based on changes in its performance indices. Likewise, the two highest slopes correspond to the two algorithms (3 and 5) that showed the greatest sensitivity.

### 4.3. Regional Comparisons

[33] Performance indices for pooled results from the nine regions are listed in Table 7. The PROBES (Bering Sea) region was the only case where the algorithms, on average, overestimated the $^{14}$C-based estimates (by 14%). In all other regions the algorithms underestimated the $^{14}$C-based estimates (between 5 and 52%). The PROBES (Bering Sea) region had the lowest pooled RMS (0.23), whereas the EqPac equator had the highest value (0.50), largely because of a high negative bias ($M = -0.32$). No region was uniformly better or worse for all algorithms. Individual algorithm RMS values (not shown) ranged from 0.19 to 0.47 in the PROBES region and from 0.13 to 0.79 in the EqPac equator region. The Arabian Sea had the three lowest RMS values (0.07, 0.07, and 0.06, which were for algorithms 6, 7, and 8, respectively).

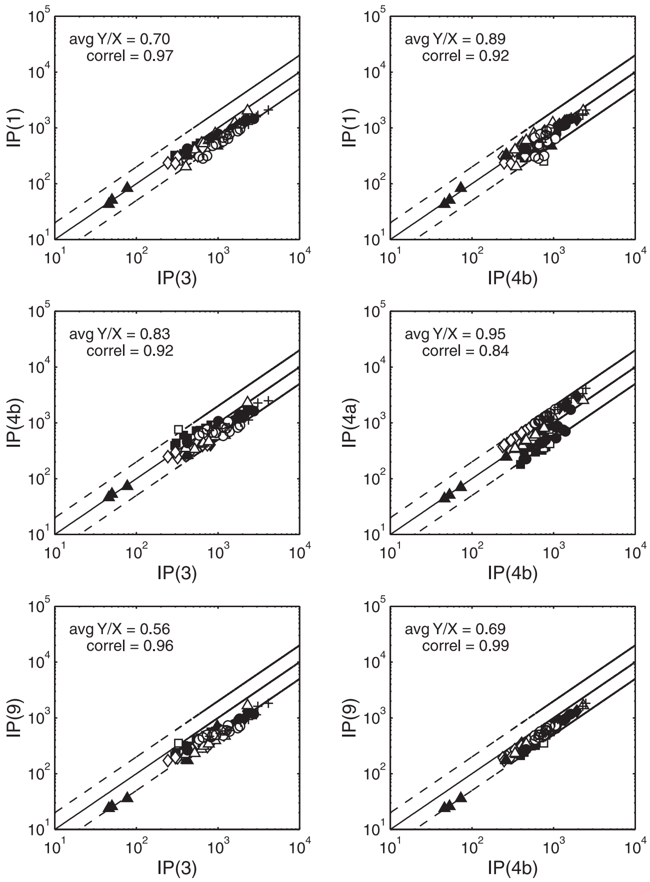


**Figure 4.** Scatterplots comparing algorithms 1 (WRM), 3 (DIM), 4a (TIM), 4b (TIM), and 9 (WRM).

[34] The two Antarctica regions (Palmer and AMERIEZ) appeared to have the worst results, judging from outliers in Figure 2, and yet pooled errors from these regions did not have especially high values of RMS or *M*. These regions included the lowest and highest values of primary production, as well as extremes in other variables. The apparent poor performance might actually be a result of the range in the Antarctic data causing high and low values to be more conspicuous.

[35] The two-way (regions and algorithms) ANOVA confirmed what was obvious from Figure 2, namely, that there were highly significant differences among regions [$F(8,960) = 30.4$; $p \ll 0.0005$] and among algorithms [$F(11,960) = 80.9$; $p \ll 0.0005$]. The algorithm-region interaction was also significant [$F(88,960) = 5.7$; $p \ll 0.0005$], indicating that algorithm performances were region dependent.

### 4.4. Algorithm Comparisons

[36] Although there were significant differences among algorithms, pairwise comparisons revealed high correlations (>0.9) in many cases. In general, the degree of correlation was unrelated to the algorithm complexity or category [*Behrenfeld and Falkowski*, 1997b]. Examples of several highly correlated pairs are illustrated in Figure 4. The two algorithms most highly correlated were 9 (WRM) and 4b (TIM). The three best-performing algorithms of each type [1 (WRM), 3 (DIM), and 4b (TIM)] were strongly correlated with one another. From these results it is clear that the structure or complexity of an algorithm seems to have no relationship to its performance.

## 5. Discussion

[37] Some of the variance in performance is likely due to methodological differences within the test data set itself, particularly in the diversity of $^{14}$C incubation methods. Short-term incubations (e.g., the Palmer LTER data) generally approximate gross primary production, whereas longer-term (e.g., 24-hour) incubations more closely approximate net primary production. If the algorithms were parameterized to yield net primary production, then they should consistently underestimate the Palmer LTER data. From Table 7, we see that the algorithms did, in fact, underestimate Palmer data, but to a lesser extent than the JGOFS data sets (EqPac and Arabian Sea), which were 24-hour incubations. There appeared to be no trends relative to the type of incubation (whether in situ, simulated in situ, or in a photosynthetron).

[38] The data furnished to participant teams did not include the year. In retrospect, we think this might have been a mistake, particularly in areas such as the equatorial Pacific where interannual variability is high and well understood. Another reason the year is important is that the ‘‘clean techniques’’ used for the past 2 decades generally produce higher estimates of primary production [*Fitzwater et al.*, 1982]. These considerations would only apply if algorithms somehow adjusted for interannual variability or the technique used. Since the year of the measurement was not provided, the algorithms made no adjustment for these factors. It is interesting that the oldest data set (PROBES, 1979–1981) was the only one in which algorithms overestimated the $^{14}$C-based production, a result that could be explained if the algorithms had been parameterized for clean techniques. The highest negative bias was in the equatorial Pacific, where, on average, algorithm predictions were only half the measured IP values, but these data (Eqpac equator, February–October 1996) were from a ‘‘normal’’ year, before the onset of the 1997–1998 El Niño. This does not explain why the algorithms would be too low, assuming they were also parameterized for normal conditions.

[39] A comparison of algorithm predictions with measurements made at discrete locations does not account for the real value of the remote sensing measurement, namely, the improved spatial and temporal coverage afforded by satellite observations. The coverage and long-term consistency afforded by remote sensing can compensate to some degree for its lack of accuracy. Ideally, both ship-based and satellite measurements will be used to monitor for changes in primary production at large scales. A robust integrative model, assimilating both in situ and remotely sensed data, will likely be required. A relatively simple technique has been demonstrated with CZCS data whereby global satellite maps are adjusted by blending them with in situ measurements [*Gregg and Conkright*, 2001]. The result is to

remove biases found in the satellite products. We observed significant biases relative to the $^{14}C$ data and also when comparing algorithms with one another. We strongly urge that additional effort be invested to understand why algorithms differ systematically from one another and from the $^{14}C$ data.

[40] The round-robin experiments elicited some debate as to whether computationally complex algorithms are worthwhile. The fact that simpler algorithms (DIMs and TIMs) performed as well as or better than complex algorithms (WRMs) suggests that the computational complexity may be unnecessary and, in fact, may be ill advised given concerns about scaling (see below). However, several participants argued in favor of the highly resolved models. They maintain that computational complexity is not an issue, because of the speed of modern computers, whereas the advantage is that it links algorithms to the experimental methods, carried out at the same scales, which inform our understanding of photosynthesis in the ocean [*Kirk*, 1994; *Falkowski and Raven*, 1997]. Such detailed algorithms have a greater opportunity to incorporate future advances in remote sensing that might provide information on accessory pigments, absorption by dissolved organic matter, or fluorescence yield. So far, however, there has not been a clear demonstration that additional complexity improves the performance of algorithms.

[41] Scaling issues are potentially an important concern that should be addressed with more rigor [*Bidigare et al.*, 1992; *Campbell et al.*, 1995]. Satellite-derived fields represent mean properties at much larger scales than the in situ data used to parameterize the algorithms. Typically, satellite-derived primary production represents the mean production over an area of at least 1 $km^2$ (often much larger) and over timescales of a week or longer. Many of the algorithms employ nonlinear relationships that were derived from measurements made at the spatial scale of an incubation bottle and at the timescale of hours. When the same models are applied to chlorophyll, light, and temperature averaged over the satellite scales, the result is not necessarily the mean primary production at the larger scale. Variance existing within the larger ''averaging bin'' affects the mean IP at the larger scale, but this variance is generally not incorporated into model parameterizations (for a good discussion of this issue see *Trela et al.* [1995]). This points to the importance of matching the scales at which the models are parameterized to the scales of the satellite products.

[42] Satellite-derived primary production is much more difficult to ''validate'' than many of the other derived properties such as chlorophyll or SST. The latter can be validated by obtaining in situ measurements at the time of a satellite overpass. Although the spatial scale of the in situ measurement would not match that of the pixel (1 $km^2$), at least the two would be simultaneous. Because incubations take several hours, the in situ primary production measurement will never match the timescale of a polar-orbiting satellite. At best, one can compare the estimate for a particular day and pixel. A more elaborate validation effort would be to observe diurnal changes in chlorophyll and light and then consider how this variability affects the satellite IP estimate.

[43] Knowledge of the vertical distribution of chlorophyll and light should improve primary productivity algorithms. The vertical distribution of light was represented in all algorithms except the DIMs (3 and 7), but only five algorithms modeled the vertical chlorophyll structure (1, 5, 8, 9, and 11). There was no evidence, however, that modeling the vertical structure was advantageous. One of the DIMs (algorithm 3), with no vertical resolution, did as well as or better than the algorithms with depth-resolved properties.

[44] *Behrenfeld and Falkowski* [1997a] demonstrated that the single most important parameter needed to improve algorithms is information on the maximum light-saturated rate of photosynthesis, $P^B_{\max}$ (or $P^B_{\rm opt}$). In many of the tested algorithms, temperature was used to derive this parameter, but the lack of consistency among available models suggests that temperature alone is not enough [*Behrenfeld and Falkowski*, 1997b]. Recently, a new $P^B_{\max}$ model has been developed [*Behrenfeld et al.*, 2002] that accounts for the effects of nutrient availability and photoacclimation. For this to be applicable to remote sensing, this model still requires the development of methods to assess the nutrient status and the physical structure, but results are promising. An avenue of current research along these lines involves the use of the natural (solar-stimulated) chlorophyll fluorescence, which can be remotely sensed by a sensor with sufficient spectral and radiometric sensitivity [*Letelier and Abbott*, 1996]. The MODIS instrument [*Esaias et al.*, 1998] is currently making such measurements. The fluorescence yield may be inversely related to the quantum yield of photosynthesis [*Falkowski and Kiefer*, 1985; *Kiefer and Reynolds*, 1992], and thus if reliable measures of chlorophyll, PAR, and chlorophyll fluorescence can be made, these may be used to parameterize $P^B_{\max}$. A combination of satellite and in situ measurements will be needed to address these issues.

## 6. Conclusions

[45] Conclusions related to the four questions addresseed by this study are summarized as follows:

[46] 1. How do algorithm estimates of primary production derived strictly from surface information compare with estimates derived from $^{14}C$ incubation methods? The 12 algorithms tested varied widely in performance. The best-performing algorithms agreed with the $^{14}C$-based estimates within a factor of 2. Two of these algorithms have been adapted by NASA for producing primary productivity maps with MODIS data. Most of the algorithms had significant biases causing them to differ systematically from the in situ data. A concerted effort should be made to understand the cause of the biases and to eliminate them if possible.

[47] 2. How does the error in satellite-derived chlorophyll concentration affect the accuracy of the primary productivity algorithms? The relative errors in primary productivity ($\Delta_{\rm sat}$) resulting from the simulated errors in surface chlorophyll concentration ($\Delta_B$) were highly correlated with $\Delta_B$. This fact reflects the deterministic relationship between

production and chlorophyll in the underlying models. The slopes of the regressions ($\Delta_{sat}$ versus $\Delta_B$) ranged between 0.3 and 0.8, indicating that errors in surface chlorophyll produce less-than-proportionate errors in IP.

[48] 3. Are there regional differences in the performance of algorithms? There were significant regional differences, as well as algorithm-region interactions, indicated by the ANOVA results. No one region was uniformly better or worse for all algorithms. The region with the most serious biases was the equatorial Pacific, where algorithms underestimated in situ measurements by a factor of 2.

[49] 4. How do algorithms compare with each other in terms of complexity vis-a-vis performance? Many of the algorithms were highly correlated with one another. This was not surprising, since several are based on the same models, but what was surprising was that the level of agreement had no apparent relationship to the mathematical structure or complexity of the algorithms. In some cases, complex algorithms based on depth-, time- and wavelength-resolved models were highly correlated with simpler algorithms that were time and/or depth integrated. There were distinct systematic differences between algorithms. A future effort to understand systematic differences is strongly recommended.

## 7. Future Considerations

[50] Four of the algorithms tested are now being applied operationally to satellite data, or are planned for use with near-future missions. A third round-robin exercise is currently underway. In the third round robin, algorithms are given global fields of satellite-derived chlorophyll, SST, and PAR, and a detailed comparison of the algorithms is being conducted to determine how and where they differ.

[51] In accordance with our recommendation, future round robins will not be blind. The anonymous nature of the results presented here seriously diminishes their usefulness beyond the participants themselves. A more open approach would have facilitated detailed comparisons between algorithms to investigate, for example, why there were systematic differences (e.g., Figure 4). The only way this could have been done under the ground rules of a blind comparison would have been if the ATS ran the codes instead of the development teams. The level of effort involved on the part of the ATS was not feasible at the time this exercise was conducted.

[52] Comparisons with in situ data are also being made in the ongoing round robin. Algorithms will be compared with over 1,000 in situ measurements, all made according to JGOFS protocols. The number of stations (89) used for evaluating algorithms in the second round robin was much too small to adequately characterize the performance of algorithms. The goal of the algorithms should be net primary production, because that is what both land and ocean satellite products are intended to represent [*Behrenfeld et al.*, 2001]. Thus 24-hour in situ incubations are the preferred method.

## Appendix A: Algorithms

### A1. Algorithm 1

[53] This algorithm employed a photosynthesis-irradiance relationship with physiological $P$ versus $E$ parameters ($\alpha$ and $P_{max}$) taken from the literature. The relationship of *Eppley* [1972] was used to compute $P_{max}$ as a function of temperature. The spectral downwelling irradiance incident at the surface was estimated based on the 5S code [*Tanré et al.*, 1990] and on cloudiness determined as the ratio of the given PAR to clear-sky PAR estimated from the model. Chlorophyll profiles were based on statistical models that were selected based on the upper (surface) chlorophyll concentration as an index of the "trophic level" [*Morel and Berthon*, 1989; *Berthon and Morel*, 1992]. A bio-optical model, based on optical measurements made at sea, was used to propagate the radiative field through the water column. The shapes of the algal absorption spectra were derived from in vitro experiments. The magnitude of the spectra employed a statistical analysis of chlorophyll-specific absorption of algae as a function of the trophic level [*Bricaud et al.*, 1995] and the *Wozniak et al.* [1992] results concerning variations in the quantum yield with trophic level [see *Morel et al.*, 1996].

### A2. Algorithm 3

[54] Chlorophyll concentration was assumed to be uniform over the euphotic layer, and IP was computed as: $IP_{alg} = P^B B_{sfc} Z_m$, where $P^B$ is the daily primary production rate per mg chlorophyll and $Z_m$ is the depth of the 1% light level. A simple $P^B$ versus $E$ model was used to compute $P^B$ as a function of the average PAR within the euphotic layer, $E = E_0/4.6$. The $P^B$ versus $E$ model used a constant value of $\alpha$ = 0.11 mg C (mg Chl)$^{-1}$ h$^{-1}$ (W m$^{-2}$)$^{-1}$ from *Platt et al.* [1991] for Atlantic noncoastal waters and used a relationship in which $P^B_{max}$ depends on SST [*Eppley*, 1972].

### A3. Algorithms 4a–4c

[55] In these algorithms the relationship between daily carbon fixation and daily average PAR at each depth was calculated using a constant slope for the light-limited region of the water column and using various models for the maximum photosynthetic rate ($P^b_{opt}$). The three versions differed with respect to the models used for $P^b_{opt}$. Algorithm 4a employed a seventh-order polynomial fit to empirical data as described by *Behrenfeld and Falkowski* [1997a]; algorithm 4b used a modification of the Eppley model as described by *Antoine and Morel* [1996a]; and algorithm 4c assumed a constant value of $P^b_{opt}$ equal to 4.6 mg C (mg Chl)$^{-1}$ h$^{-1}$. The chlorophyll profile was assumed to be constant and equal to the surface value.

### A4. Algorithm 5

[56] In this algorithm a chlorophyll-specific absorption coefficient for PAR was modeled as a function of time of year, ranging from 0.006 to 0.015 m$^2$ (mg Chl)$^{-1}$, with the maximum occurring in the summer months. The total attenuation coefficient for PAR included phytoplankton absorption, together with water and detrital attenuation, and then PAR irradiance profiles, $E_{par}(z)$, were derived according to Beer's Law. Chlorophyll was assumed to have a Gaussian-shaped subsurface chlorophyll maximum for surface values <0.4 mg m$^{-3}$ and was assumed constant with

depth, otherwise. Production as a function of depth was then calculated using an irradiance-dependent formulation for quantum yield together with phytoplankton absorption and $E_{par}(z)$, and production was then integrated over depth.

## A5. Algorithm 6

[57] This algorithm calculates the spectral radiation absorbed by phytoplankton and multiplies that by a quantum yield to compute the hourly rate of primary production at each depth. The solar irradiance is split into spectral components via the 5S radiative transfer code [*Tanré et al.*, 1990], and the spectral light field is propagated through the ocean using very simple two-stream approximations. The absorption and scattering coefficients required for this were obtained from the literature [*Smith and Baker*, 1981; *Pope and Fry*, 1997; *Bricaud et al.*, 1995; *Gordon and Morel*, 1983; *Petzold*, 1972]. All absorption calculations were carried out spectrally and then integrated (400–700 nm). Quantum yield was calculated using a parameterization based on maximum quantum yield of 0.03 mol C (mol quanta)$^{-1}$ and a light-dependent term. The chlorophyll profile was assumed to be vertically uniform.

## A6. Algorithm 7

[58] This algorithm is based on empirical relationships developed by the author from data obtained on many expeditions in tropical, temperate, and polar regions. The primary production data was from half-day in situ incubations, and chlorophyll was measured by spectrophotometric methods without applying a correction for phaeopigments. Estimation of the daily primary production was obtained using a "psi-based" formulation: IP = $\psi \cdot E_0 \cdot$ DL $\cdot$ IB, where $\psi$ ("psi") was empirically modeled as a function of temperature for three trophic zones determined by the surface chlorophyll level. $E_0$ was the daily incident radiation; DL was the hours of daylight, and IB was empirically modeled from the author's own data, where different models were applied depending on the surface chlorophyll level and the zone (tropic, temperate, or polar).

## A7. Algorithm 8

[59] This algorithm employed a photosynthesis-irradiance relationship whose parameters were determined statistically for the biogeochemical province in which the station is located [*Longhurst et al.*, 1995]. Similarly, the vertical chlorophyll profile was based on statistical models of profiles for each province. Surface incident irradiance was determined based on cloudiness (in a manner similar to that used by algorithm 1). A full radiative transfer code was then used to propagate spectral irradiance downward through the water column.

## A8. Algorithm 9

[60] This algorithm is similar to that described by *Morel* [1991]; solar spectral irradiance was estimated using the *Gregg and Carder* [1990] model with a wind speed of 4 m s$^{-1}$, water vapor of 2 cm, and visibility of 23 km. Clear-sky surface spectral irradiance was scaled to the measured surface PAR. The diffuse downwelling attenuation coefficient was estimated as the sum of the total absorption coefficient plus backscattering coefficient divided by the average cosine. Methods for estimating total absorption and backscattering are from *Morel* [1991, 1988]. The vertical profile of chlorophyll was simulated using the models of *Morel and Berthon* [1989], and the temperature dependence of $P^B_{max}$ was based on an Eppley model as modified by *Antoine and Morel* [1996a]. A constant value of 0.033 m$^2$ (mg Chl)$^{-1}$ was used for the chlorophyll-specific absorption coefficient at 440 nm. Daily primary production was determined by trapezoidal integration in hourly time steps over the photoperiod and at 0.5-m-depth intervals.

## A9. Algorithm 11

[61] This algorithm used input data on surface chlorophyll, temperature, and light and estimated vertical profiles of these three properties over the euphotic zone. The depth distributions of chlorophyll and temperature were estimated using empirical relationships derived from a large globally distributed data set. The daily production at each depth was calculated as the product of chlorophyll, the daily PAR ($E_{par}$), and the chlorophyll-specific light utilization efficiency ($\psi$). The chlorophyll-specific light utilization efficiency, $\psi$, was constrained not to exceed a theoretical maximum based on the ambient temperature.

## A10. Algorithm 12

[62] In this algorithm the light field was spectrally and vertically resolved, but a uniform vertical distribution of chlorophyll was assumed. The algorithm calculation requires knowledge of surface chlorophyll concentration, surface light, temperature, mixed layer depth, and the concentration of a limiting nutrient. From this information, estimates of the *P* versus *E* parameters were made, and thus *P* was determined at each depth and integrated to estimate IP.

[63] **Acknowledgments.** Support for NASA's Ocean Primary Productivity Working Group was provided by NASA. The first author was funded by NASA's Earth Science Enterprise as a member of the MODIS Instrument Team (NAS5-96063) and SeaWiFS Science Team (NSG5-6289). The other U.S. participants were funded by NASA's Earth Science Enterprise through grants to the various individuals (too numerous to mention by number). The authors wish to thank Seung-Hyun Son, for assistance with the graphics, and Mark Dowell, Timothy Moore, and three anonymous reviewers whose helpful suggestions substantially improved the paper. In addition, we are grateful to the organizers of JGOFS, for open access to their data, and to colleagues who furnished data, including Barbara Prezelin and Mark Moline (Palmer LTER), Nick Welschmeyer (SUPER), Walker Smith (AMERIEZ), Jay O'Reilly (MARMAP), and Lou Codispoti (PROBES).

## References

Antoine, D., and A. Morel, Ocean primary production, 1, Adaptation of a spectral light-photosynthesis model in view of application to satellite chlorophyll observations, *Global Biogeochem. Cycles*, *10*, 43–55, 1996a.

Antoine, D., and A. Morel, Ocean primary production, 2, Estimation at global scale from satellite (Coastal Zone Color Scanner) chlorophyll, *Global Biogeochem. Cycles*, *10*, 57–69, 1996b.

Balch, W. M., Accuracy of historical primary production measurements, in International Workshop on Oceanographic Biological and Chemical Data

Management, edited by S. Levitus and L. Oliounine, *NOAA/NESDIS Tech. Rep. 8*, pp. 137–146, 1997.
Balch, W. M., R. W. Eppley, and M. R. Abbott, Remote sensing of primary production, II, A semi-analytical algorithm based on pigments, temperature, and light, *Deep Sea Res.*, *36*, 1201–1217, 1989.
Balch, W. M., R. Evans, J. Brown, G. Feldman, C. McClain, and W. Esaias, The remote sensing of ocean primary productivity: Use of a new data compilation to test satellite algorithms, *J. Geophys. Res.*, *97*(C2), 2279–2293, 1992.
Barber, R. T., M. P. Sanderson, S. T. Lindley, F. Chai, J. Newton, C. C. Trees, D. G. Foley, and F. P. Chavez, Primary productivity and its regulation in the equatorial Pacific during and following the 1991–92 El Niño, *Deep Sea Res., Part II*, *43*, 933–969, 1996.
Barber, R. T., J. Marra, R. C. Bidigare, L. A. Codispoti, D. Halpern, Z. Johnson, M. Latasa, R. Goericke, and S. L. Smith, Primary productivity and its regulation in the Arabian Sea during 1995, *Deep Sea Res., Part II*, *48*, 1127–1172, 2001.
Behrenfeld, M. J., and P. G. Falkowski, Photosynthetic rates derived from satellite-based chlorophyll concentration, *Limnol. Oceanogr.*, *42*, 1–20, 1997a.
Behrenfeld, M. J., and P. G. Falkowski, A consumer's guide to phytoplankton primary productivity models, *Limnol. Oceanogr.*, *42*, 1479–1491, 1997b.
Behrenfeld, M. J., et al., Temporal variations in the photosynthetic biosphere, *Science*, *291*, 2594–2597, 2001.
Behrenfeld, M. J., E. Marañón, D. A. Siegel, and S. B. Hooker, A photoacclimation and nutrient-based model of light-saturated photosynthesis for quantifying oceanic primary production, *Mar. Ecol. Prog. Ser.*, *228*, 103–117, 2002.
Berthon, J.-F., and A. Morel, Validation of a spectral light-photosynthesis model and use of the model in conjunction with remotely sensed pigment observations, *Limnol. Oceanogr.*, *37*, 781–796, 1992.
Bidigare, R. R., B. B. Prezelin, and R. C. Smith, Bio-optical models and the problems of scaling, in *Primary Productivity and Biogeochemical Cycles in the Sea*, edited by P. G. Falkowski and A. D. Woodhead, pp. 175–212, Plenum, New York, 1992.
Bricaud, A., M. Bain, A. Morel, and H. Claustre, Variability in the chlorophyll-specific absorption coefficients of natural phytoplankton: Analysis and parameterization, *J. Geophys. Res.*, *100*(C7), 13,321–13,332, 1995.
Campbell, J. W., and J. E. O'Reilly, Role of satellites in estimating primary productivity on the northwest Atlantic continental shelf, *Cont. Shelf Res.*, *8*(2), 179–204, 1988.
Campbell, J. W., S. R. Gaudreau, and G. M. Weiss, The challenge of scaling primary productivity models to the global ocean: A statistical approach, paper presented at International Conference on Photosynthesis and Remote Sensing, Eur. Assoc. of Remote Sens. Lab., Montpellier, France, 28–30 August 1995.
Chipman, D. W., J. Marra, and T. Takahashi, Primary production at 47N and 20W in the North Atlantic Ocean: A comparison between the 14C incubation method and the mixed layer carbon budget, *Deep Sea Res.*, *40*, 151–169, 1993.
Codispoti, L. A., G. E. Friederich, R. L. Iverson, and D. W. Hood, Temporal changes in the inorganic carbon system of the southeastern Bering Sea during spring 1980, *Nature*, *296*, 242–245, 1982.
Ducklow, H. W., and R. P. Harris, JGOFS: The North Atlantic Bloom Experiment, *Deep Sea Res., Part II*, *40*, 1–8, 1993.
Eppley, R. W., Temperature and phytoplankton growth in the sea, *Fish. Bull.*, *70*, 1063–1085, 1972.
Eppley, R. W., E. Steward, M. R. Abbott, and U. Heyman, Estimating ocean primary production from satellite chlorophyll: Introduction to regional differences and statistics for the Southern California Bight, *J. Plankton Res.*, *7*, 57–70, 1985.
Esaias, W. E., et al., An overview of MODIS capabilities for ocean science observations, *IEEE Trans. Geosci. Remote Sens.*, *36*, 1250–1265, 1998.
Falkowski, P. G., and D. A. Kiefer, Chlorophyll *a* fluorescence and phytoplankton: Relationship to photosynthesis and biomass, *J. Plankton Res.*, *7*, 715–731, 1985.
Falkowski, P. G., and J. A. Raven, *Aquatic Photosynthesis*, Blackwell Sci., Malden, Mass., 1997.
Fitzwater, S. E., G. A. Knauer, and J. H. Martin, Metal contamination and its effect on primary production measurements, *Limnol. Oceanogr.*, *27*, 544–551, 1982.
Gordon, H. R., and A. Y. Morel, *Remote Assessment of Ocean Color for Interpretation of Satellite Visible Imagery*, Springer-Verlag, New York, 1983.
Gordon, H. R., R. W. Austin, D. K. Clark, W. A. Hovis, and C. S. Yentsch, Ocean color measurements, in *Advances in Geophysics*, vol. 27, *Satellite Oceanic Remote Sensing*, edited by G. Salzman, pp. 297–333, Academic, San Diego, Calif., 1985.
Gregg, W. W., and K. L. Carder, A simple spectral solar irradiance model for cloudless maritime atmospheres, *Limnol. Oceanogr.*, *35*, 1657–1675, 1990.
Gregg, W. W., and M. E. Conkright, Global seasonal climatologies of ocean chlorophyll: Blending in situ and satellite data for the Coastal Zone Color Scanner era, *J. Geophys. Res.*, *106*(C2), 2499–2525, 2001.
Howard, K. L., Estimating global ocean primary production using satellite-derived data, M.S. thesis, 98 pp., Univ. of R. I., Kingston, 1995.
Howard, K. L., and J. A. Yoder, Contribution of the subtropical ocean to global primary primary production, in *Space Remote Sensing of the Subtropical Oceans*, edited by C.-T. Liu, pp. 157–168, Pergamon, New York, 1997.
Iverson, R. L., W. E. Esaias, and K. Turpie, Ocean annual phytoplankton carbon and new production, and annual export production estimated with empirical equations and CZCS data, *Global Change Biol.*, *6*, 57–72, 2000.
Kiefer, D. A., and R. A. Reynolds, Advances in understanding phytoplankton fluorescence and photosynthesis, in *Primary Productivity and Biogeochemical Cycles in the Sea*, edited by P. G. Falkowski and A. D. Woodhead, pp. 155–174, Plenum, New York, 1992.
Kirk, J. T. O., *Light and Photosynthesis in Aquatic Ecosystems*, 2nd ed., Cambridge Univ. Press, New York, 1994.
Knudson, C., W. S. Chamberlin, and J. Marra, Primary production and irradiance data for the U.S. JGOFS (leg 2), Atlantis II (Cruise 119-4), *L-DGO Tech. Rep. LDGO-89-4*, Lamont-Doherty Earth Obs., Palisades, New York, 1989.
Letelier, R. M., and M. R. Abbott, An analysis of chlorophyll fluorescence algorithms for the Moderate Resolution Imaging Spectroradiometer (MODIS), *Remote Sens. Environ.*, *58*, 215–223, 1996.
Longhurst, A., S. Sathyendranath, T. Platt, and C. Caverhill, An estimate of global primary production in the ocean from satellite radiometer data, *J. Plankton Res.*, *17*, 1245–1271, 1995.
Moline, M. A., and B. B. Prezelin, High-resolution time-series data for 1991/1992 primary production and related parameters at a Palmer LTER coastal site: Implications for modeling carbon fixation in the Southern Ocean, *Polar Biol.*, *17*(1), 39–53, 1997.
Morel, A., Optical modeling of the upper ocean in relation to its biogenous matter content (case 1 waters), *J. Geophys. Res.*, *93*(C9), 10,749–10,768, 1988.
Morel, A., Light and marine photosynthesis: A spectral model with geochemical and climatological implications, *Prog. Oceanogr.*, *26*, 263–306, 1991.
Morel, A., and J.-F. Berthon, Surface pigments, algal biomass profiles, and potential production of the euphotic layer: Relationships reinvestigated in view of remote-sensing applications, *Limnol. Oceanogr.*, *34*, 1545–1562, 1989.
Morel, A., D. Antoine, M. Babin, and Y. Dandonneau, Measured and modeled primary production in the Northeast Atlantic (EUMELI JGOFS program): The impact of natural variations in photosynthetic parameters on model predictive skill, *Deep Sea Res., Part 1*, *43*, 1273–1304, 1996.
Ondrusek, M. E., R. R. Bidigare, K. Waters, and D. M. Karl, A predictive model for estimating rates of primary production in the subtropical North Pacific Ocean, *Deep Sea Res., Part II*, *48*, 1837–1863, 2001.
O'Reilly, J. E., C. Evans-Zetlin, and D. A. Busch, Primary production, chap. 21, in *Georges Bank*, edited by R. H. Backus, pp. 220–233, MIT Press, Cambridge, Mass., 1987.
O'Reilly, J. E., S. Maritorena, B. G. Mitchell, D. A. Siegel, K. L. Carder, S. A. Garver, M. Kahru, and C. McClain, Ocean color chlorophyll algorithms for SeaWiFS, *J. Geophys. Res.*, *103*(C11), 24,937–24,953, 1998.
Peterson, B. J., Aquatic primary productivity and the $^{14}CO_2$ method: A history of the productivity problem, *Ann. Rev. Ecol. Syst.*, *11*, 369–385, 1980.
Petzold, T., Volume scattering functions for selected ocean waters, *SIO Ref. 72-78*, Scripps Inst. of Oceanogr., San Diego, Calif., 1972.
Platt, T., Primary production of the ocean water column as a function of surface light intensity: Algorithms for remote sensing, *Deep Sea Res.*, *33*, 1–15, 1986.
Platt, T., and S. Sathyendranath, Estimators of primary production for interpretation of remotely sensed data on ocean color, *J. Geophys. Res.*, *98*(C8), 14,561–14,567, 1993.
Platt, T., C. Caverhill, and S. Sathyendranath, Basin-scale estimates of primary production by remote sensing: The North Atlantic, *J. Geophys. Res.*, *96*(C8), 15,147–15,159, 1991.
Pope, R. M., and E. S. Fry, Absorption spectrum (380–700 nm) of pure water, II, Integrating cavity measurements, *Appl. Opt.*, *36*, 8710–8723, 1997.

Prezelin, B. B., and H. E. Glover, Variability in time/space estimates of phytoplankton, biomass, and productivity in the Sargasso Sea, *J. Plankton Res.*, *13*, 45–67, 1991.

Richardson, K., Comparison of $^{14}C$ primary production determinations made by different laboratories, *Mar. Ecol. Prog. Ser.*, *72*, 189–201, 1991.

Smith, R. C., and K. S. Baker, The bio-optical state of ocean waters and remote sensing, *Limnol. Oceanogr.*, *23*, 247–259, 1978.

Smith, R. C., and K. S. Baker, Optical properties of the clearest natural waters (200–800 nm), *Appl. Optics*, *20*, 177–184, 1981.

Smith, W. O., Jr., and D. M. Nelson, Primary productivity and nutrient uptake in an Antarctic marginal ice zone during austral spring and autumn, *Limnol. Oceanogr.*, *35*, 809–821, 1990.

Tanré, D., C. Deroo, P. Duhaut, M. Herman, J. Morcrette, J. Perbos, and P. Deschamps, Description of a computer code to simulate the satellite signal in the solar spectrum: The 5S code, *Int. J. Remote Sens.*, *11*, 659–668, 1990.

Trela, P., S. Sathyendranath, R. M. Moore, and D. E. Kelley, Effect of the nonlinearity of the carbonate system on partial pressure of carbon dioxide in the oceans, *J. Geophys. Res.*, *100*(C4), 6829–6844, 1995.

Welschmeyer, N. A., S. Strom, R. Goericke, G. Ditullio, M. Belvin, and W. Petersen, Primary production in the sub-Arctic Pacific Ocean: Project SUPER, *Mar. Ecol. Prog. Ser.*, *74*, 101–135, 1993.

Wozniak, B., J. Dera, and O. Koblentz-Mischke, Modeling the relationship between primary production, optical properties, and nutrients in the sea, *Proc. SPIE Soc. Opt. Eng.*, *1750*, 246–275, 1992.

D. Antoine and A. Morel, Laboratoire d'Océanographie de Villefranche, CNRS/INSU and Universite Pierre et Marie Curie, Quai de La Darse, BP 8, F-06238 Villefranche sur Mer, France. (david@obs-vlfr.fr; morel@obs-vlfr.fr)

R. Armstrong, Marine Sciences Research Center, Stony Brook University, Stony Brook, NY 11794-5000, USA. (rarmstrong@notes.cc.sunysb.edu)

K. Arrigo, Department of Geophysics, Stanford University, Mitchell Building, Room 360, Stanford, CA 94305-2215, USA. (arrigo@pangea.stanford.edu)

W. Balch and C. Yentsch, Bigelow Laboratory for Ocean Sciences, West Booth Bay Harbor, ME 04575, USA. (bbalch@bigelow.org; cyentsch@bigelow.org)

R. Barber, Marine Laboratory, Duke University, 135 Duke Marine Lab Road, Beaufort, NC 28516-9721, USA. (rbarber@duke.edu)

M. Behrenfeld and W. Esaias, NASA Goddard Space Flight Center, Code 971.1, Greenbelt, MD 20771-0001, USA. (mjb@neptune.gsfc.nasa.gov; wayne.esaias@gsfc.nasa.gov)

R. Bidigare, Department of Oceanography, University of Hawaii, Honolulu, HI 96822, USA. (bidigare@soest.hawaii.edu)

J. Bishop, Earth Sciences Division, Lawrence Berkeley National Laboratory, 1 Cyclotron Road, M/S 90-1116, Berkeley, CA 94720, USA. (JKBishop@lbl.gov)

J. W. Campbell, Ocean Process Analysis Laboratory, Institute for Study of Earth, Oceans, and Space, University of New Hampshire, Morse Hall, Durham, NH 03824-3525, USA. (janet.campbell@unh.edu)

M.-E. Carr, Jet Propulsion Laboratory, Mail Stop 300-323, 4800 Oak Grove Drive, Pasadena, CA 91109, USA. (mec@pacific.jpl.nasa.gov)

P. Falkowski, Department of Geology and Institute of Marine and Coastal Sciences, Rutgers University, New Brunswick, N.J. 08901, USA. (falko@imcs.Rutgers.edu)

N. Hoepffner, Institute for Environment and Sustainability, Inland and Marine Waters Unit, Joint Research Centre of the European Commission, I-21020 Ispra, Italy. (nicolas.hoepffner@jrc.it)

R. Iverson, Department of Oceanography, Florida State University, Tallahassee, FL 32306-3048, USA. (riverson@ocean.fsu.edu)

D. Kiefer, Department of Biological Sciences, University of Southern California, Los Angeles, CA 90089, USA. (kiefer@physics1.usc.edu)

S. Lohrenz, Department of Marine Science, University of Southern Mississippi, 1020 Balch Boulevard, Stennis Space Center, MS 39529, USA. (steven.lohrenz@usm.edu)

J. Marra, Lamont-Doherty Earth Observatory, Columbia University, 61 Route 9W, Palisades, NY 10964, USA. (marra@ldeo.columbia.edu)

J. Ryan, Monterey Bay Aquarium Research Institute, 7700 Sandholdt Road, Moss Landing, CA 95039, USA. (ryjo@mbari.org)

V. Vedernikov, P. P. Shirshov Institute of Oceanology, Russian Academy of Sciences, 36 Nakhimovsky Prospect, Moscow, 117997 Russia. (oleg@manta.sio.rssi.ru)

K. Waters, NOAA Coastal Services Center, 2234 South Hobson Avenue Charleston, SC 29405-2413, USA. (Kirk.Waters@noaa.gov)

J. Yoder, Graduate School of Oceanography, University of Rhode Island, South Ferry Road, Narragansett, RI 02882, USA. (jyoder@gso.uri.edu)