Source of Acquisition
NASA Goddard Space Flight Center

Sampling Biases in MODIS and SeaWiFS Ocean Chlorophyll Data

Authors: Watson W. Gregg (GMAO) – Watson.Gregg@nasa.gov
Nancy W. Casey (SSAI) – nancy.casey@gsfc.nasa.gov

Abstract

Although modern ocean color sensors, such as MODIS and SeaWiFS are often considered global missions, in reality it takes many days, even months, to sample the ocean surface enough to provide complete global coverage. The irregular temporal sampling of ocean color sensors can produce biases in monthly and annual mean chlorophyll estimates. We quantified the biases due to sampling using data assimilation to create a "truth field", which we then sub-sampled using the observational patterns of MODIS and SeaWiFS. Monthly and annual mean chlorophyll estimates from these sub-sampled, incomplete daily fields were constructed and compared to monthly and annual means from the complete daily fields of the assimilation model, at a spatial resolution of 1.25° longitude by 0.67° latitude.

The results showed that global annual mean biases were positive, reaching nearly 8% (MODIS) and >5% (SeaWiFS). For perspective the maximum interannual variability in the SeaWiFS chlorophyll record was about 3%. Annual mean sampling biases were low (<3%) in the mid-latitudes (between -40° and 40°). Low interannual variability in the global annual mean sampling biases suggested that global scale trend analyses were valid.

High latitude biases were much higher than the global annual means, up to 20% as a basin annual mean, and over 80% in some months. This was the result of the high solar zenith angle exclusion in the processing algorithms. Only data where the solar angle is <75° are permitted, in contrast to the assimilation which samples regularly over the entire area and month. High solar zenith angles do not facilitate phytoplankton photosynthesis and consequently low chlorophyll concentrations occurring here are missed by the data sets. Ocean color sensors selectively sample in locations and times of favorable phytoplankton growth, producing overestimates of chlorophyll.
The biases derived from lack of sampling in the high latitudes varied monthly, leading to artifacts in the apparent seasonal cycle from ocean color sensors. A false secondary peak in chlorophyll occurred in May-August, which resulted from lack of sampling in the Antarctic.

Persistent clouds, characteristic in the North Pacific, also produced overestimates, again by selectively sampling only the high growth periods. In contrast, areas characterized by thick aerosols showed chlorophyll underestimates to nearly -30% in basin monthly means. This was the result of selective sampling in lower aerosol thickness periods, which corresponded with lower phytoplankton growth periods.

A combination of MODIS and SeaWiFS sampling was most effective at reducing mid-latitude biases due to inter-orbit gaps, sun glint, and sensor tilt changes. But these biases were low using a single sensor, suggesting multiple sensors had little effect in reducing global and regional monthly and annual mean biases.

Ocean color data are an invaluable source of information about global biological processes. However, these results suggest that sampling errors need to be considered in applications involving global and regional mean chlorophyll biomasses as well as seasonal variability and regional trend analysis.

Sampling Biases in MODIS and SeaWiFS Ocean Chlorophyll Data

Authors: Watson W. Gregg (GMAO) – Watson.Gregg@nasa.gov
Nancy W. Casey (SSAI) – nancy.casey@gsfc.nasa.gov

Popular Summary

Although modern ocean color sensors, such as MODIS and SeaWiFS are often considered global missions, in reality it takes many days, even months, to sample the ocean surface enough to provide complete global coverage. The irregular temporal sampling of ocean color sensors can produce biases in monthly and annual mean chlorophyll estimates. We quantified the biases due to sampling using data assimilation to create a "truth field", which we then sub-sampled using the observational patterns of MODIS and SeaWiFS. Monthly and annual mean chlorophyll estimates from these sub-sampled, incomplete daily fields were constructed and compared to monthly and annual means from the complete daily fields of the assimilation model, at a spatial resolution of 1.25° longitude by 0.67° latitude.

The results showed that global annual mean biases were positive, reaching nearly 8% (MODIS) and >5% (SeaWiFS). For perspective the maximum interannual variability in the SeaWiFS chlorophyll record was about 3%.

High latitude biases were much higher than the global annual means, up to 20% as a basin annual mean, and over 80% in some months. This was the result of the high solar zenith angle exclusion in the processing algorithms. Only data where the solar angle is <75° are permitted, in contrast to the assimilation which samples regularly over the entire area and month. High solar zenith angles do not facilitate phytoplankton photosynthesis and consequently low chlorophyll concentrations occurring here are missed by the data sets. Ocean color sensors selectively sample in locations and times of favorable phytoplankton growth, producing overestimates of chlorophyll.

The biases derived from lack of sampling in the high latitudes varied monthly, leading to artifacts in the apparent seasonal cycle from ocean color sensors. A false secondary peak in chlorophyll occurred in May-August, which resulted from lack of sampling in the Antarctic.

Persistent clouds, characteristic in the North Pacific, also produced overestimates, again by selectively sampling only the high growth periods. In contrast, areas characterized by thick aerosols showed chlorophyll underestimates to nearly -30% in basin monthly means. This was the result of selective sampling in lower aerosol thickness periods, which corresponded with lower phytoplankton growth periods.

Ocean color data are an invaluable source of information about global biological processes. However, these results suggest that sampling errors need to be considered in applications involving global and regional mean chlorophyll biomasses as well as seasonal variability and regional trend analysis.

Sampling Biases in MODIS and SeaWiFS Ocean Chlorophyll Data

Watson W. Gregg
Global Modeling and Assimilation Office
NASA/Goddard Space Flight Center
Greenbelt, MD 20771
watson.gregg@nasa.gov

and

Nancy W. Casey
Science Systems and Applications, Inc.
Lanham, MD 20706
nancy.casey@gsfc.nasa.gov

Abstract

Although modern ocean color sensors, such as MODIS and SeaWiFS are often considered global missions, in reality it takes many days, even months, to sample the ocean surface enough to provide complete global coverage. The irregular temporal sampling of ocean color sensors can produce biases in monthly and annual mean chlorophyll estimates. We quantified the biases due to sampling using data assimilation to create a "truth field", which we then sub-sampled using the observational patterns of MODIS and SeaWiFS. Monthly and annual mean chlorophyll estimates from these sub-sampled, incomplete daily fields were constructed and compared to monthly and annual means from the complete daily fields of the assimilation model, at a spatial resolution of 1.25° longitude by 0.67° latitude.

The results showed that global annual mean biases were positive, reaching nearly 8% (MODIS) and >5% (SeaWiFS). For perspective the maximum interannual variability in the SeaWiFS chlorophyll record was about 3%. Annual mean sampling biases were low (<3%) in the mid-latitudes (between -40° and 40°). Low interannual variability in the global annual mean sampling biases suggested that global scale trend analyses were valid.

High latitude biases were much higher than the global annual means, up to 20% as a basin annual mean, and over 80% in some months. This was the result of the high solar zenith angle exclusion in the processing algorithms. Only data where the solar angle is <75° are permitted, in contrast to the assimilation which samples regularly over the entire area and month. High solar

zenith angles do not facilitate phytoplankton photosynthesis and consequently low chlorophyll concentrations occurring here are missed by the data sets. Ocean color sensors selectively sample in locations and times of favorable phytoplankton growth, producing overestimates of chlorophyll.

The biases derived from lack of sampling in the high latitudes varied monthly, leading to artifacts in the apparent seasonal cycle from ocean color sensors. A false secondary peak in chlorophyll occurred in May-August, which resulted from lack of sampling in the Antarctic.

Persistent clouds, characteristic in the North Pacific, also produced overestimates, again by selectively sampling only the high growth periods. In contrast, areas characterized by thick aerosols showed chlorophyll underestimates to nearly -30% in basin monthly means. This was the result of selective sampling in lower aerosol thickness periods, which corresponded with lower phytoplankton growth periods.

A combination of MODIS and SeaWiFS sampling was most effective at reducing mid-latitude biases due to inter-orbit gaps, sun glint, and sensor tilt changes. But these biases were low using a single sensor, suggesting multiple sensors had little effect in reducing global and regional monthly and annual mean biases.

Ocean color data are an invaluable source of information about global biological processes. However, these results suggest that sampling errors need to be considered in applications involving global and regional mean chlorophyll biomasses as well as seasonal variability and regional trend analysis.

Introduction

Modern ocean color sensors, such as the Moderate Resolution Imaging Spectroradiometer (MODIS) on the Aqua spacecraft and the Sea-viewing Wide Field-of-view Sensor (SeaWiFS) are widely assumed to produce global representations of ocean color. However, truly global representations are achieved only after accumulating data over many days and even months. On a daily basis, clouds (and associated stray light), thick aerosols, inter-orbit gaps, sun glint, and high solar zenith angles prevent complete daily coverage by either obscuration or lack of sampling (Figure 1). SeaWiFS has additional data gaps caused by a sensor tilt change. Kwiatkowska et al. (2007) have shown the daily ocean coverage to be about 16% for SeaWiFS and 13% for MODIS-Aqua at 9 km resolution.

These data gaps nearly disappear in the monthly means. Repeated attempts to view the surface over a month usually result in a clear observation for most of the ocean. Exceptions include persistent clouds and aerosols, as well as the high latitudes in local winter, where processing algorithms exclude data beyond a solar zenith angle limit, typically about 75° (Wang, 2002).

In the monthly fields, the data gaps have been filled with observations taken on different days, often occurring with temporary breaks in cloud or aerosols. High latitude monthly mean observations in local winter comprise only the most equator-ward locations and only the time of the month where solar angles do not exceed the prescribed limit.

What effects do these sampling irregularities have on the monthly mean? Does the monthly mean produce the same global representation of ocean color quantities as regular (complete daily) sampling would? Or does the irregular sampling produce biases in the global and regional monthly means? These questions have implications for estimates of large-scale chlorophyll concentrations and primary production from satellites.

These are the questions under investigation here. Using data assimilation, we produce realistic, complete monthly fields of ocean chlorophyll, and then sub-sample these fields with the observational pattern of modern ocean color sensors, to evaluate the biases in the remote sensing data.

Methods

The key to quantitatively evaluating the errors associated with irregular sampling by satellites is to first obtain complete daily fields. We can then use the satellite observational patterns to sub-sample the complete daily fields, just like the satellite would see them. Monthly and annual averages constructed from the sub-sampled satellite "observations" of the daily fields can then be compared to monthly and annual means derived from the underlying complete daily fields. The difference provides a reliable representation of the sampling bias, and the sampling bias alone, since there are no artifacts associated with sensor, calibration, or algorithms.

There are potentially many ways to construct the underlying complete daily fields that we could use to determine sampling bias. It is a simple problem in gap-filling, for which many approaches have been developed, among them linear and nonlinear extrapolation, objective analysis, optimal interpolation, climatologies, etc., and various combinations. It is a common problem for in situ data applications. The National Oceanographic Data Center (NODC) uses

objective analysis to provide gridded data from the sparsely sampled in situ observations (Levitus, 1982).

We choose a three-dimensional global biogeochemical model to produce the complete daily chlorophyll fields required for this analysis. This option is preferred over statistical or climatological representations because it derives daily fields from temporally consistent dynamical processes. Further, by assimilating satellite chlorophyll data into the model, we minimize local biases of the free-run (unassimilated) model to provide realistic representations of daily global chlorophyll that enables us to understand the scientific importance of the sampling biases. These complete daily global fields of ocean chlorophyll from the assimilation provide a "truth field" to evaluate the sampling biases in satellite observations. The procedure for assimilation is depicted in Figure 2.

We then sample these complete fields using the orbit and scanning characteristics of MODIS-Aqua and SeaWiFS for each day (Figure 3). Essentially we overlay the daily satellite sampling pattern on the assimilation field. It is conceptually like flying a satellite over an ocean that looks like the assimilation field. If the satellite observes an ocean location for a given day, the assimilation values for that grid point are taken. If the satellite viewing is obscured or missing for that location and day, no assimilation values are used. This enables us to provide an unambiguous representation of the sampling error, uncomplicated by sensor data retrieval or derived products.

Three years are analyzed (2003, 2004, 2005) to observe inter-annual variability in the sampling errors. We report the mean error statistics (differences) over these three years. For sampling by MODIS and SeaWiFS, grid points where observations are missing for the entire month (Figure 3) are not included in the basin or annual monthly mean. For the assimilation, all ocean grid points are occupied. Annual means are the arithmetic average of all months, again only where observations exist.

Assimilation Model

The free-run model, the NASA Ocean Biogeochemical Model (NOBM), is fully described in Gregg (2007). Briefly, the model is a 3D representation of global biogeochemical processes, including explicit circulation dynamics, biological processes, and radiative transfer. The biological model contains multiple nutrient and phytoplankton groups, along with detrital, grazing, and carbon components in a fully interactive and dynamical representation of the global

oceans. Sea ice cover is required to modify local dynamics in the presence of ice. The model increments are 1.25° longitude by 0.67° latitude extending from 84°S to 72°N. The model domain includes bottom depths > 200m, so coastal processes are excluded.

Seasonal variability in NOBM is positively correlated with SeaWiFS in all 12 of the major oceanographic basins (Figure 4; Gregg et al., 2003). Annual mean chlorophyll has a bias of -8% with SeaWiFS (model low), +1.1% with Aqua, and -17.1% with in situ data from NODC (Conkright et al., 2002). Global annual primary production estimates are within 1% and 9% of estimates derived from SeaWiFS and Aqua, respectively, using a common primary production algorithm (Behrenfeld and Falkowski, 1997). Point-to-point comparisons of the model with NODC and NASA (Werdell and Bailey, 2002) in situ chlorophyll show a bias of -1.4% and an uncertainty (root-mean-square, RMS) of 61.8% (Gregg, 2007).

The assimilation model uses the Conditional Relaxation Analysis Method (CRAM) in combination with NOBM and remote sensing data (Gregg, 2007). The method is used for bias correction in Optimal Interpolation Sea Surface Temperature (OISST) data (Reynolds, 1988; Reynolds and Smith, 1994), and has been used successfully for ocean color applications (Gregg and Conkright, 2001; 2002; Conkright and Gregg, 2003). Assimilation is performed daily, using log-transformed remote sensing and model chlorophyll. Data errors are explicitly included (Gregg, 2007) based on a global statistical comparison of SeaWiFS and in situ data (Gregg and Casey, 2004).

The assimilation model produces a bias of 0.1% when compared to NODC and NASA in situ data over 6 years (Gregg, 2007). This bias is lower than for SeaWiFS (-1.3% against the same data set). The uncertainty (33.4%) is comparable with SeaWiFS (32.7%).

The model is forced by monthly mean atmospheric data, including wind stress and sea surface temperature for the circulation model, iron and sea ice for the biological model, and atmospheric optical data fields including clouds and aerosols for the radiative model (Gregg, 2007). The monthly mean data are linearly-interpolated to daily forcing. While the use of interpolated monthly mean forcing may influence the variability of our daily results, it is important to note that daily variability is enforced by the assimilation in most areas of the ocean, except in the high latitudes in local winter. Even here the daily variability caused by clouds is small, since the amount of light entering the oceans is always small. In cloud-contaminated areas, the

assimilation of nearby observations is sufficient to influence the model via the CRAM method.

Data

Remote sensing ocean chlorophyll data are 9 km Level-3 Standard Mapped Image daily fields from the NASA Ocean Color Web for both MODIS-Aqua (Version 1.1) and SeaWiFS (Version 5.1). The data are re-gridded from 9 km to the model resolution for assimilation. MODIS-Aqua data became available in July 2002. SeaWiFS data have been available since September 1997. Our analysis begins in 2003 because this is the first full year both MODIS-Aqua and SeaWiFS data are available.

In order to minimize differences associated with high solar zenith angle, grid points containing ice are excluded from the analysis. The major seasonal ice fields develop at solar zenith angles > 75°, exceeding the solar zenith angle processing threshold. In addition, biological dynamics in the ice sheets are not well known. Elimination of any grid point containing ice minimizes the effects of these two issues, and provides a lower bound on the error estimates. Monthly sea ice data are obtained from the NOAA Optimum Interpolation Sea Surface Temperature Analysis (OIv2) data set (Reynolds et al., 2002).

Statistical Assessment

Statistical comparisons primarily emphasize global and basin differences in monthly and annual means. Global and basin means are area-weighted. Differences between area-weighted monthly/annual means are expressed as percent error

$$\text{Basin/Global Percent Error} = \frac{C(sat)-C(assim)}{C(assim)} \times 100 \qquad (1)$$

where C(assim) is the assimilation model mean chlorophyll and C(sat) is the sampling by satellite sensor (MODIS or SeaWiFS) chlorophyll mean. This represents the basin or global bias.

Results

Biases in MODIS Sampling

The global annual mean difference between the assimilation ("truth") and sampling by MODIS is 7.7% (Figure 5). These overestimates by MODIS sampling occur entirely in the high latitudes, i.e., Antarctic, North Pacific, and North Atlantic, where mean errors are >10%. All of the other

basins exhibit mean annual biases due to MODIS sampling <±3% (Figure 5). There is also little interannual variability in the global annual mean biases (data not shown).

Monthly mean biases show that the maxima in the North Atlantic and Antarctic exceed 50%, reaching 84% in the Antarctic in 2004 (Figure 6). The timing of the maximum biases is local winter (December and January for the North Atlantic, and June and July for the Antarctic). The maximum biases in the North Pacific are smaller, but still exceed 10%, and occur in local summer, in contrast to the other two high latitude basins. Monthly minima for the Antarctic and North Atlantic basins are never negative, but occasionally the North Pacific shows negative biases, reaching a minimum of -3.6% in Jan 2004. There is considerable interannual variability in the biases derived from MODIS sampling.

The North Indian basin exhibits very large interannual variability in the monthly sampling biases, ranging from nearly -30% in July 2003, to 20% in August 2004, back to -26% in July 2005 (Figure 6). These magnitudes and trends are in contrast with the basin's low and unvarying annual mean biases.

The location of the biases caused by MODIS sampling provides some indications for the cause. The overestimates in the Antarctic lie along the very southern edge of sampling by MODIS in August 2003 (Figure 7). The errors exceed 50% locally. We also note the overestimation trends in the North Atlantic and Pacific, which are partially counterbalanced by patches of underestimates. There is little sampling of the northern Arabian Sea by MODIS this month. Note also the localized underestimates near the US East and West coasts, near the eastern Asian coasts, and along the western South American coasts. The Equatorial Pacific exhibits an alternating pattern of overestimates and underestimates along the axis of the upwelling.

Very large overestimates are apparent in the extreme northern portion of the North Atlantic in March 2004 (Figure 8). These MODIS sampling biases exceed 100% at the northernmost edge of sampling. We also note underestimates in the central Equatorial Atlantic, occasionally <-25%, interspersed with overestimates toward the African coast.

In June 2003, nearly the entire Arabian Sea chlorophyll is underestimated by MODIS sampling (Figure 9), with values locally <-25% to nearly -50%. This corresponds with low sampling frequency (Figure 9). Overestimates in the North Pacific are apparent, especially in the Okhotsk and Bering Seas. Again the nearshore US, western South American, and eastern Asian coasts exhibit substantially lower estimated chlorophyll from MODIS sampling.

Biases in SeaWiFS Sampling

SeaWiFS-sampling annual mean differences (Figure 10) are very similar to MODIS. The global mean difference is less (5.6% compared to 7.7%), deriving mostly from smaller differences in the Antarctic. SeaWiFS sampling extends about 2° latitude farther south (not shown) than MODIS (Figure 1).

Biases in Combined MODIS-SeaWiFS Sampling

When daily overpasses of MODIS and SeaWiFS are combined, we can evaluate the residual errors of multiple sensors. Using two-sensor combination, we see that there is negligible improvement in annual mean biases compared to the fully sampled assimilation (Figure 10). There are improvements relative to MODIS sampling, to be sure, but nearly all is realized with SeaWiFS sampling alone. Occasionally the two-sensor combination shows higher annual mean biases than SeaWiFS alone, but the magnitude is small.

Biases in Seasonal Variability

The biases in sampling by ocean color sensors are not evenly distributed throughout the year. A mean of the three years of monthly mean biases shows that MODIS sampling and the assimilation provide different representations of seasonal variability (Figure 11). The MODIS sampling provides modestly higher estimates of monthly mean chlorophyll than the assimilation in September through April, but vastly different estimates from May though August. This produces an apparent secondary peak in the seasonal cycle in mid-boreal-summer in contrast to the assimilation.

Biases Due to Clouds

Persistent clouds are a problem for remote sensing in the high latitudes. We observed basin mean biases in the local summer in the North Pacific, North Atlantic, and Antarctic to estimate the magnitude of the cloud problem. The local summer season is shown to avoid excessive solar zenith angle biases. The biases due to clouds are largest in the North Pacific (Table 1), followed by the North Atlantic. They are small in the Antarctic. For all basins, SeaWiFS sampling biases are smaller than MODIS, and the two sensor combination has the smallest biases, although only marginally.

Discussion

The global annual mean bias (nearly 8% for MODIS and 6% for SeaWiFS) must be considered in global analysis of phytoplankton biomass and primary production. For perspective, the maximum difference in global annual area-weighted mean chlorophyll in the SeaWiFS record is 3.3% (0.18 mg m^{-3} over the model domain in 1998; 0.186 mg m^{-3} in 1999). Thus the global sampling bias is larger than natural interannual variability. However, year-to-year global annual mean bias estimates from 2003-2005 are steady, suggesting that global-scale trend analyses (e.g., Gregg et al., 2005, Behrenfeld et al., 2006) are valid. Extended analysis for 1998-2005 using SeaWiFS for assimilation (data not shown) support this finding. Further, biases are small ($<\pm3\%$) in most ocean basins for monthly and annual estimates of large-scale chlorophyll concentrations.

These conclusions apply strictly to the resolution under investigation here, namely monthly-to-annual temporal and $<1^{\circ}$ spatial, which we consider climate scales. Daily sampling is even more biased than the monthlies at global and regional scales. Kwiatkowska et al. (2007) have shown that, at higher spatial resolution, ocean surface coverage is drastically reduced. For example at 1 km daily only 5% of the ocean is observed by MODIS and about $1/10^{th}$ of that by SeaWiFS. Since biases are the result of missing daily data, they are likely to be higher when high resolution data are used for estimates of basin and global means.

Local and regional distributions of the monthly and annual biases can be very large. The largest errors are overestimates and occur in the high latitudes. In the North Atlantic and Antarctic, the largest of the errors occur in the local winter season. These biases are caused by exclusion of data where the solar zenith angle is $>75^{\circ}$. Although both MODIS and SeaWiFS collect data at much higher solar zenith angles, 75° is considered a limit on modern atmospheric correction algorithms (Wang, 2002).

In the Antarctic in August 2003, the errors, which exceed 50% locally, are located at the very edge of sampling of the region (Figure 7). The number of grid points sampled this month is small (Figure 7), and all occur at the end of August, when the solar declination has moved far enough south that this portion is illuminated by solar zenith angles $<75^{\circ}$. The MODIS-sampled monthly mean comprises data only for this small portion of the region, for this short time. This portion and time is the most favorable time for phytoplankton growth during this month because it has the lowest solar angles and therefore the most light for photosynthesis. Conversely, the assimilation model represents the mean of the entire region for the entire month, most of which

was under high solar zenith angles unfavorable for reproduction. This is how the overestimates due to sampling losses at high solar zenith angle occur.

As an example, we can see the how the sampling evolves over the year at 58°S latitude (Figure 12). May-July have no observations, as the latitude experiences solar zenith angles greater than the processing limit. April has observations only for the first half of the month, representing the portion of the month that has the lowest solar angles. August only has a few observations occurring at the very end of the month.

A similar situation is found in the extreme northern portion of the North Atlantic in March 2004 (Figure 8). Here MODIS sampling only occurs at the end of March, under lower solar zenith angles, biasing the monthly mean for times most favorable to growth (i.e., lowest solar angles and most light). These regional sampling biases contribute to the annual global mean (Figure 5). Unless they are accounted for, estimates of chlorophyll biomass and primary production will follow the sampling biases shown here. For primary production the error may be smaller because the solar angle is still high at this time and place, and most irradiance is likely diffuse, which does not penetrate deeply into the ocean, limiting phytoplankton growth rates.

The solar zenith angle bias is reduced for SeaWiFS in the Antarctic because its noon orbit provides overall more favorable solar angles for the duration of its orbits. However, the bias is still substantial.

The solar angle bias also exhibits monthly variability. The largest biases occur in May through August. These biases lead to an apparent secondary peak in the global seasonal trend occurring in these months (Figure 11). It appears in actual MODIS global monthly mean data (Figure 13). This is an artifact arising from lack of data sampling in the vast ocean area of the Antarctic. The biases are not as large when it is the northern high latitudes that are under-sampled, because there is less ocean area. However, our analysis likely underestimates the magnitude of this bias given the domain only extends to 72°N. The asymmetry in the global monthly mean biases leads to a misrepresentation of global seasonal variability by ocean color sensors.

Of the issues that prevent complete sampling by remote ocean color sensors, exclusion of data with high solar zenith angle is the largest source of error. Interannual variability in the biases caused by this source can be large (±10% in the Antarctic, ±20% in the North Atlantic). Thus in addition to the effects on global seasonal variability, these biases can have severe impact on basin-scale trend analysis in the high latitudes.

The solar angle bias occurs in unsampled regions during local winter, where the chlorophyll concentrations are low, according to the model. Are the chlorophyll concentrations really low here in local winter? Using the NODC and NASA in situ archives, we derive a mean value of 0.149 mg m^{-3} for 361 observations in the Antarctic for the months of June-August. The assimilation model mean (not excluding data points containing ice) is 0.141 mg m^{-3} for the same three months averaged over three years. If we assume the unsampled data is 0 mg m^{-3} and we include them in the monthly/annual means, can we remove the sampling bias in the satellite data? When we try this, we find that the seasonal artifact disappears (Figure 13). However, the global mean values are considerably lower than the assimilation.

Clouds, and associated stray light, are the next largest source of sampling error. The most prominent example of this is the North Pacific, which is so impacted by clouds, that even in the prime growing season there can be <5 days of observations per month (Figure 9). The sampling problems are due to persistent cloud cover and lead to biases between 6 and 13% in the growth months of May through July. Again these represent overestimates as a basin mean, although there is considerable variability in the direction of the bias (Figure 9). Interannual variability appears to be relatively small (Figure 6) so this problem may not affect trend analysis here. The overestimates due to clouds also occur in the North Atlantic, but they tend to be smaller (2 to 8%; Table 1). The fact that the biases due to clouds are overestimates suggests that higher chlorophyll concentrations occur during the limited periods of clear skies suitable for remote sensing. This contrasts with the findings by Muller-Karger et al. (1990) and Mitchell et al. (1991), who noted that blooms occurred underneath clouds and when the skies cleared, remote sensors had missed the highest concentrations.

It is interesting that despite persistent clouds, biases in the local summer in the Antarctic are small (Table 1). This is in contrast to the 8-11% errors observed in the North Pacific. A combination of MODIS and SeaWiFS reduced the bias in the North Pacific up to 2.5% over MODIS sampling, but only produces an improvement <1% in the North Atlantic and Antarctic.

Aerosols are the last important contributor to sampling error. Their effects are also localized in both time and space. The North Indian suffers the largest impact, where basin monthly mean biases up to -30% are observed. Again the bias is larger with MODIS sampling than SeaWiFS sampling, due to MODIS' greater tendency to saturate over bright targets. Aerosols in the North Indian can be so thick and persistent that no observations of the ocean can occur for an entire

month (Figure 9). Here the errors are usually underestimates in contrast to sampling operations and clouds. This is because the largest losses of data from aerosol obscuration occur at the same time as the highest chlorophyll concentrations: just before and during the Southwest Monsoon. The mechanism is likely the same: the high winds of the Southwest Monsoon produce upwelling that stimulates phytoplankton blooms, but at the same time transport aerosols from nearby desert regions (Wang et al., 2005).

The remaining sampling issues, namely sun glint, inter-orbit gaps, and sensor tilt changes, occur primarily in the mid-latitudes (Figure 1). The mid-latitudes are the regions that are least impacted by sampling irregularities, in the absence of clouds and aerosols, and routinely produce the highest sampling frequencies (Figures 7 and 8). Remote ocean color sensors provide relatively unbiased large scale estimates of monthly and annual chlorophyll concentrations in these regions (Figures 5 and 10), suggesting that sun glint, inter-orbit gaps, and tilt changes are not problems.

These results have implications for mission strategies. The primary purpose of multiple ocean color satellites, as set forth in IOCCG (1999; 2007), and Gregg et al. (1998), is to improve daily coverage. The improvements in daily coverage result mostly from elimination or reduction of gaps caused by sun glint, inter-orbit gaps, and sensor tilt changes. We report here that these two sampling issues do not contribute to sampling errors at large spatial scales and monthly and annual time scales. We find negligible improvement in annual mean biases using two satellites. Although there is some improvement in areas obscured by clouds and possibly aerosols using multiple observations, it is minimal. The largest biases are caused by exclusion of data associated with high solar zenith angle. These biases cannot be improved substantially by multiple satellites, but rather require extended observations of high scientific quality under high solar zenith angles.

The results here indicate that ocean color sensors produce biases in global and regional mean chlorophyll concentrations, regional interannual variability, and distorted global seasonal variability resulting from irregular sampling. Most of the biases result from data exclusion at high solar angle, although persistent clouds are also a problem. Regionally, the biases can be more important than globally. These sampling biases derive from the sensors' characteristic of selectively sampling during times and locations of favorable growth by phytoplankton.

Ocean color sensors provide a unique source of information on global biological distributions and abundances, but they do not provide an accurate representation by themselves. To obtain representative descriptions of global chlorophyll distributions and abundances, we need additional information to overcome these sampling limitations.

Acknowledgements

We thank the NASA Ocean Biology Processing Group for MODIS and SeaWiFS chlorophyll data. We also thank three anonymous reviewers. This work was supported by the NASA REASoN, EOS, and MAP programs.

References

Behrenfeld, M.J., Falkowski, P.G., 1997. Photosynthetic rates derived from satellite-based chlorophyll concentrations. Limnology and Oceanography, 42: 1-20.

Behrenfeld, M.J., and 9 others, 2006. Climate-driven trends in contemporary ocean productivity. Nature 444: 752-755.

Conkright, M.E., Antonov, J.I., Baranova, O., Boyer, T.P., Garcia, H.E., Gelfeld, R., Johnson, D., O'Brien, T.D., Smolyar, I., Stephens, C., 2002. World ocean database 2001, Vol. 1: Introduction. S. Levitus, Ed., NOAA Atlas NESDIS 42, US Govt. Printing Office, Washington, DC, 167 pp.

Conkright, M.E., Gregg, W.W., 2003. Comparison of chlorophyll climatologies: in situ, CZCS, blended in situ-CZCS, and SeaWiFS. International Journal of Remote Sensing 24: 969-991.

Gregg, W.W., W.E. Esaias, G.C. Feldman, R. Frouin, S.B. Hooker, C.R. McClain, and R.H. Woodward, 1998. Coverage opportunities for global ocean color in a multi-mission era. IEEE Transactions on Geoscience and Remote Sensing 36:1620-1627.

Gregg, W.W., Conkright, M.E., 2001. Global seasonal climatologies of ocean chlorophyll: Blending in situ and satellite data for the Coastal Zone Color Scanner era. Journal of Geophysical Research, 106: 2499-2515.

Gregg, W.W., Conkright, M.E., 2002. Decadal changes in global ocean chlorophyll. Geophysical Research Letters 29 (11): 10.1029/2002GL014689.

Gregg, W.W., P. Ginoux, P.S. Schopf, and N.W. Casey, 2003. Phytoplankton and Iron: Validation of a global three-dimensional ocean biogeochemical model. Deep-Sea Research II 50: 3143-3169.

Gregg, W.W. and N.W. Casey, 2004. Global and regional evaluation of the SeaWiFS chlorophyll data set. Remote Sensing of Environment 93: 463-479.

Gregg, W.W., N.W. Casey, and C.R. McClain, 2005. Recent trends in global ocean chlorophyll. Geophysical Research Letters, 32, L03606, doi:101029/2004GL021808.

Gregg, W.W. and N.W. Casey, 2007. Modeling coccolithophores in the global oceans. Deep-Sea Research II, in press. doi: 10.1016/j.dsr2.2006.12.007 (pre-print available at http://gmao.gsfc.nasa.gov/research/oceanbiology/index.php)

Gregg, W.W., 2007. Assimilation of SeaWiFS Ocean Chlorophyll Data into a Three-

Dimensional Global Ocean Model. Journal of Marine Systems, in press. doi: 10.1016/j.marsys.2006.02.015 (pre-print available at http://gmao.gsfc.nasa.gov/research/oceanbiology/index.php)

IOCCG, 1999. Status and plans for satellite ocean-colour missions: Considerations for complementary missions. Yoder, J.A. (ed.), Reports of the International Ocean-Colour Coordinating Group, No. 2, IOCCG, Dartmouth, Canada No. 2. 43pp.

IOCCG, 2007. Ocean color data merging. Gregg, W.W. (ed.), Reports of the International Ocean-Colour Coordinating Group, No. 6, IOCCG, Dartmouth, Canada, in press.

Kwiatkowska, E., Pinnock, S. and Maritorena, S., 2007. Benefits of merging. In: Ocean color data merging. Gregg, W.W. (ed.), Reports of the International Ocean-Colour Coordinating Group, No. 6, IOCCG, Dartmouth, Canada, in press.

Levitus, S., 1982. Climatological atlas of the world ocean. NOAA Professional Paper 13. US Govt. Printing Office, Washington, DC, 173 pp.

Mitchell, B.G., E.A. Brody, E.-N. Yeh, C. McClain, J. Comiso, and N.G. Maynard, Meridional zonation of the Barents Sea ecosystem inferred from satellite remote sensing and in situ bio-optical observations, 1991. In: Sakshaug, E., Hopkins, C.C.E. and Oritsland, N.A (eds.): Proceedings of the Pro Mare Symposium on Polar Marine Ecology, pp. 147-162, Trondheim, 1990.

Muller-Karger, F.E, C.R. McClain, R.N. Sambrotto, and G.C. Ray, 1990. A comparison of ship and coastal zone color scanner mapped distribution of phytoplankton in the southeastern Bering Sea., Journal of Geophysical Research 95: 11483-11499.

Reynolds, R.W., 1988. A real-time global sea surface temperature analysis. Journal of Climate 1: 75-86.

Reynolds, R.W. Smith, T.M., 1994. Improved global sea surface temperature analyses using optimum interpolation. Journal of Climate 7: 75-86.

Reynolds, R.W., Rayner, N.A., Smith, T.M., Stokes, D.C., and Wang, W., 2002. An improved in situ and satellite SST analysis for climate. Journal of Climate 15: 1609-1625.

Wang, M., 2002. The Rayleigh lookup tables for the SeaWiFS data processing: Accounting for the effects of ocean surface roughness. International Journal of Remote Sensing 23: 2693-2702.

Wang, M., Knobelspiesse, K.D., and McClain, C.R., 2005. Study of the Sea-Viewing Wide Field-of-View Sensor (SeaWiFS) aerosol optical property data over ocean in combination with the ocean color products. Journal of Geophysical Research 110, D10S06, doi:10.1029/2004JD004950, 2005

Werdell, P.J., Bailey, S.W., 2002. The SeaWiFS Bio-Optical Archive and Storage System (SeaBASS): Current Architecture and Implementation. NASA Technical Memorandum 2002-211617, G.S. Fargion and C.R. McClain, eds. NASA Goddard Space Flight Center, Greenbelt, MD.

Table 1. Monthly mean biases due to cloud obscuration in the North Pacific, North Atlantic, and Antarctic basins. The biases shown represent the mean of 3 years (2003-2005).

	N. Pacific (Jun)	N. Atlantic (Jul)	Antarctic (Dec)
MODIS	11.1%	6.1%	3.4%
SeaWiFS	9.4%	5.9%	2.9%
MODIS+SeaWiFS	8.6%	5.0%	2.3%

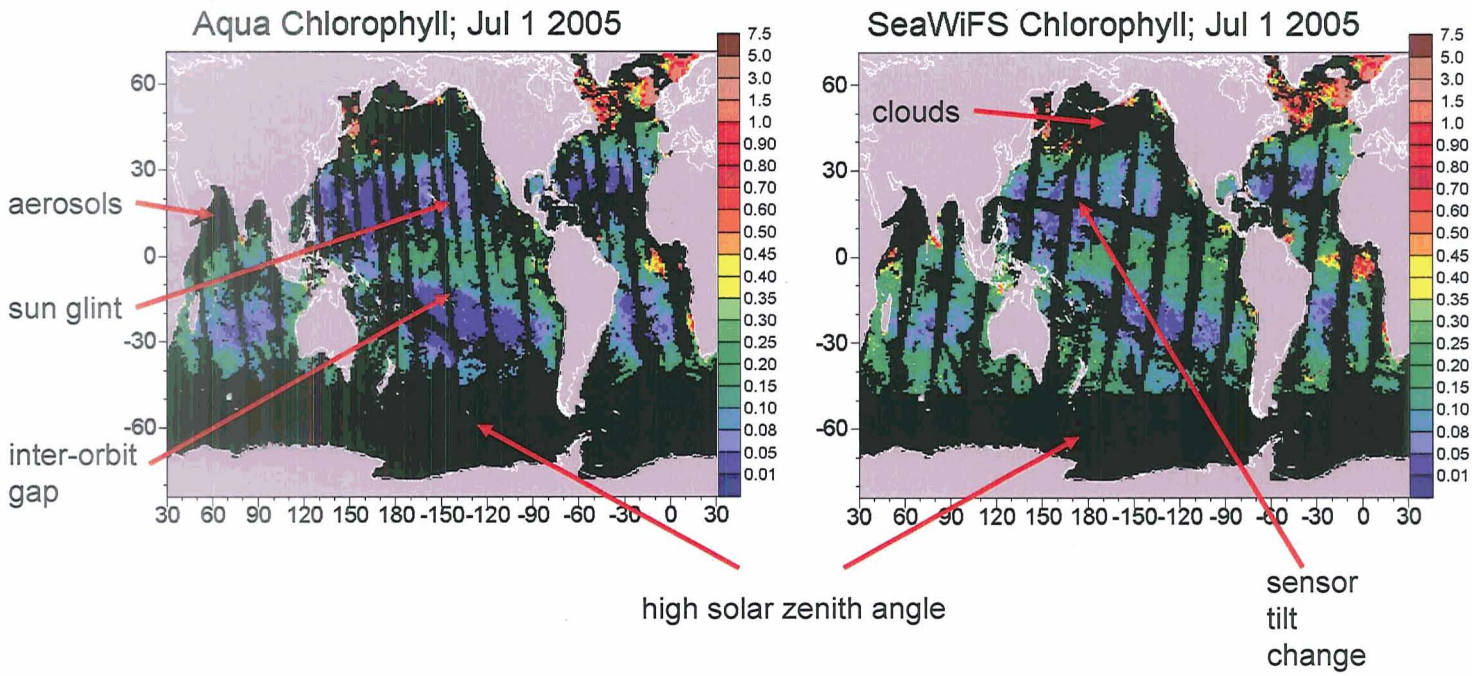


Figure 1. Daily ocean coverage by MODIS-Aqua and SeaWiFS, with sources causing data gaps identified.

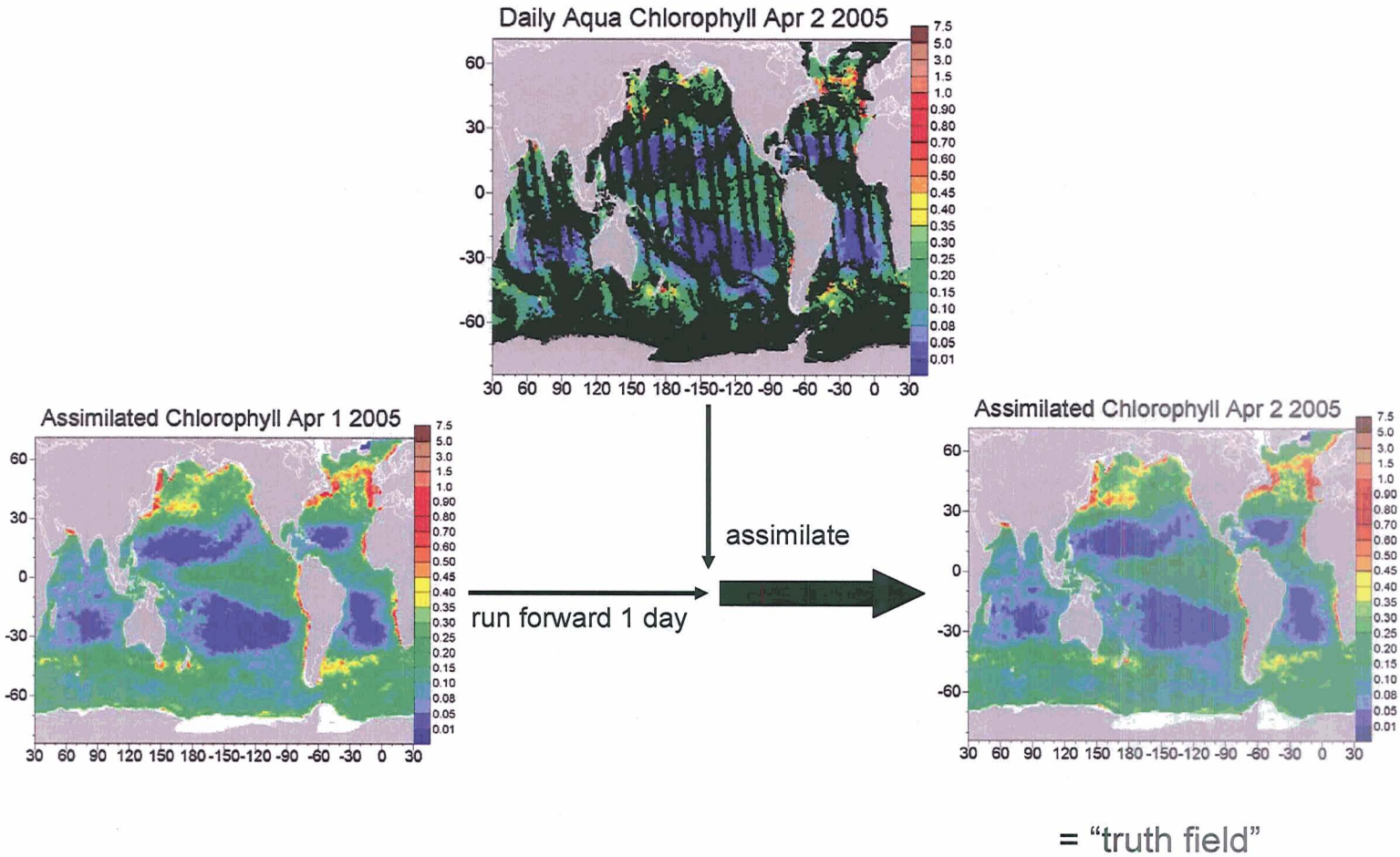


Figure 2. Procedure for assimilating MODIS-Aqua daily data with a global 3D model. The assimilation model from a previous day, April 1, 2005, is integrated forward in "free-run" mode for a day, and then assimilated with the Aqua data. A complete global field is created for each day, representing a model constrained by MODIS-Aqua data, that provides a "truth field" for the analysis of sampling error.

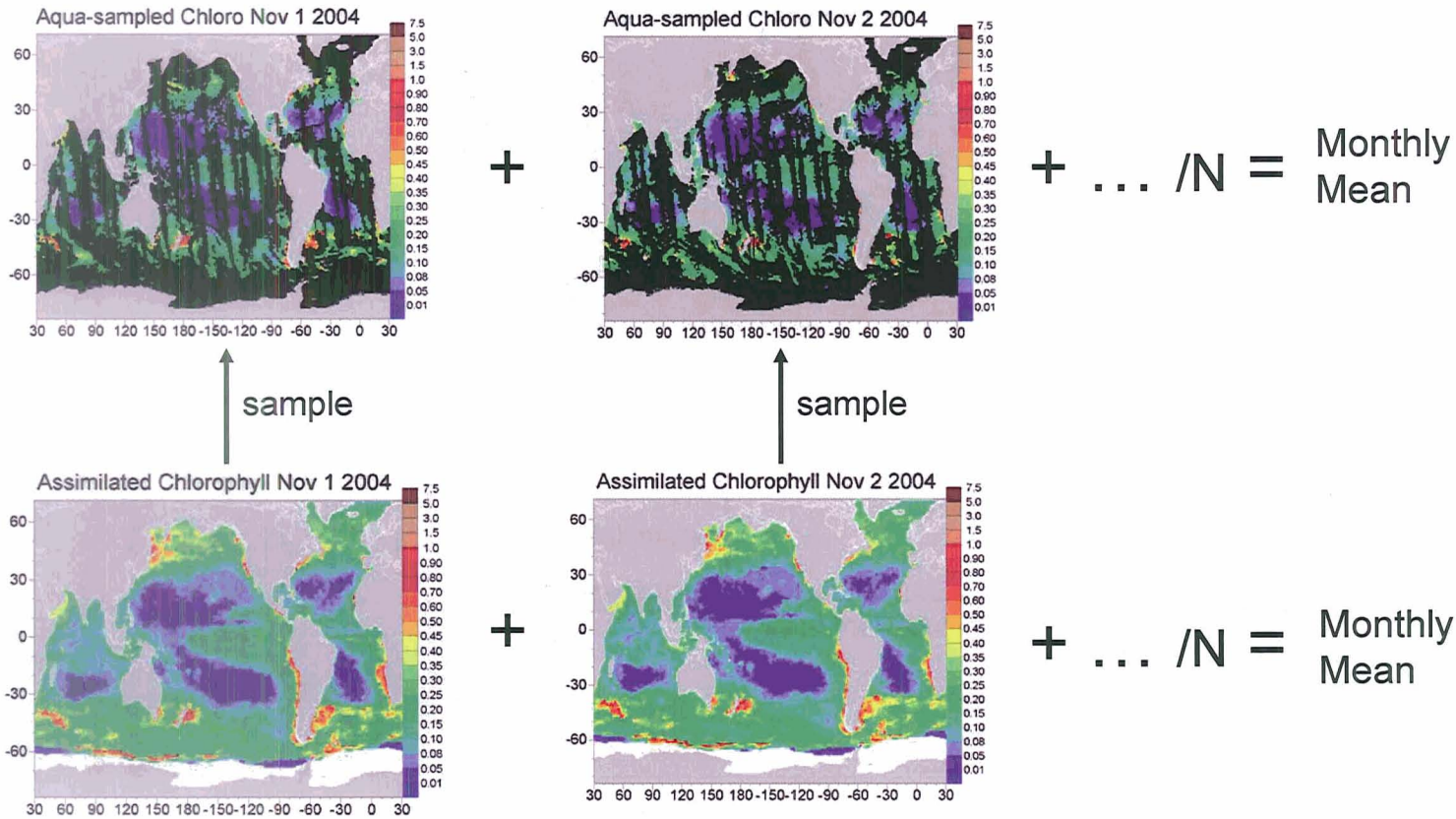


Figure 3. Construction of monthly mean data using complete fields from the assimilation (bottom), and these fields sampled by MODIS-Aqua overpasses (top). N is the number of samples in in grid point per month. Sampling errors are determined by comparing monthly and annual mean fields from the assimilation and the MODIS-Aqua sampling of the assimilation.

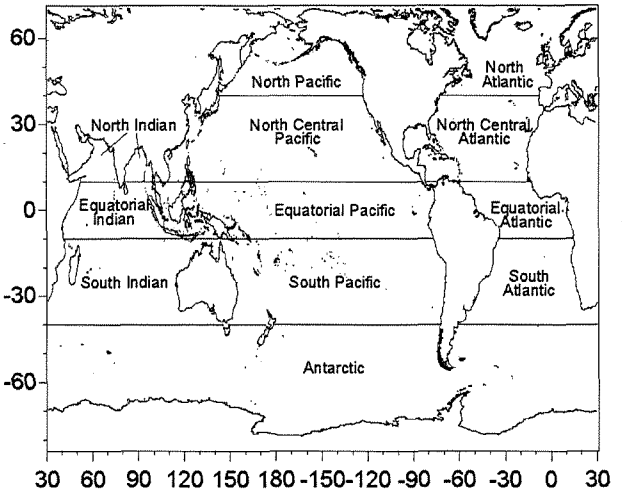


Figure 4. Location of major oceanographic basins.

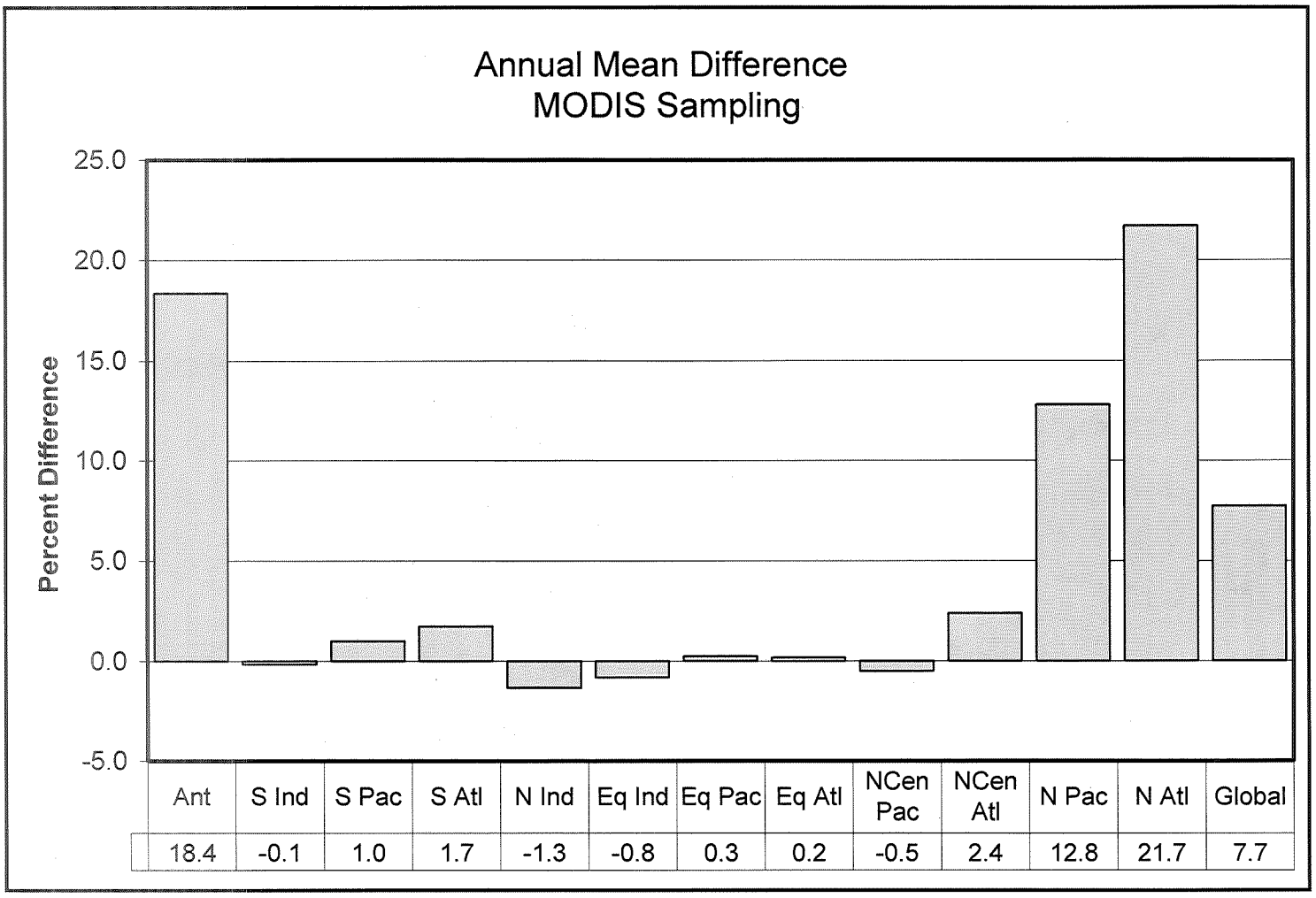


Figure 5. Annual mean difference between the assimilation model and MODIS-Aqua sampling of the assimilation.

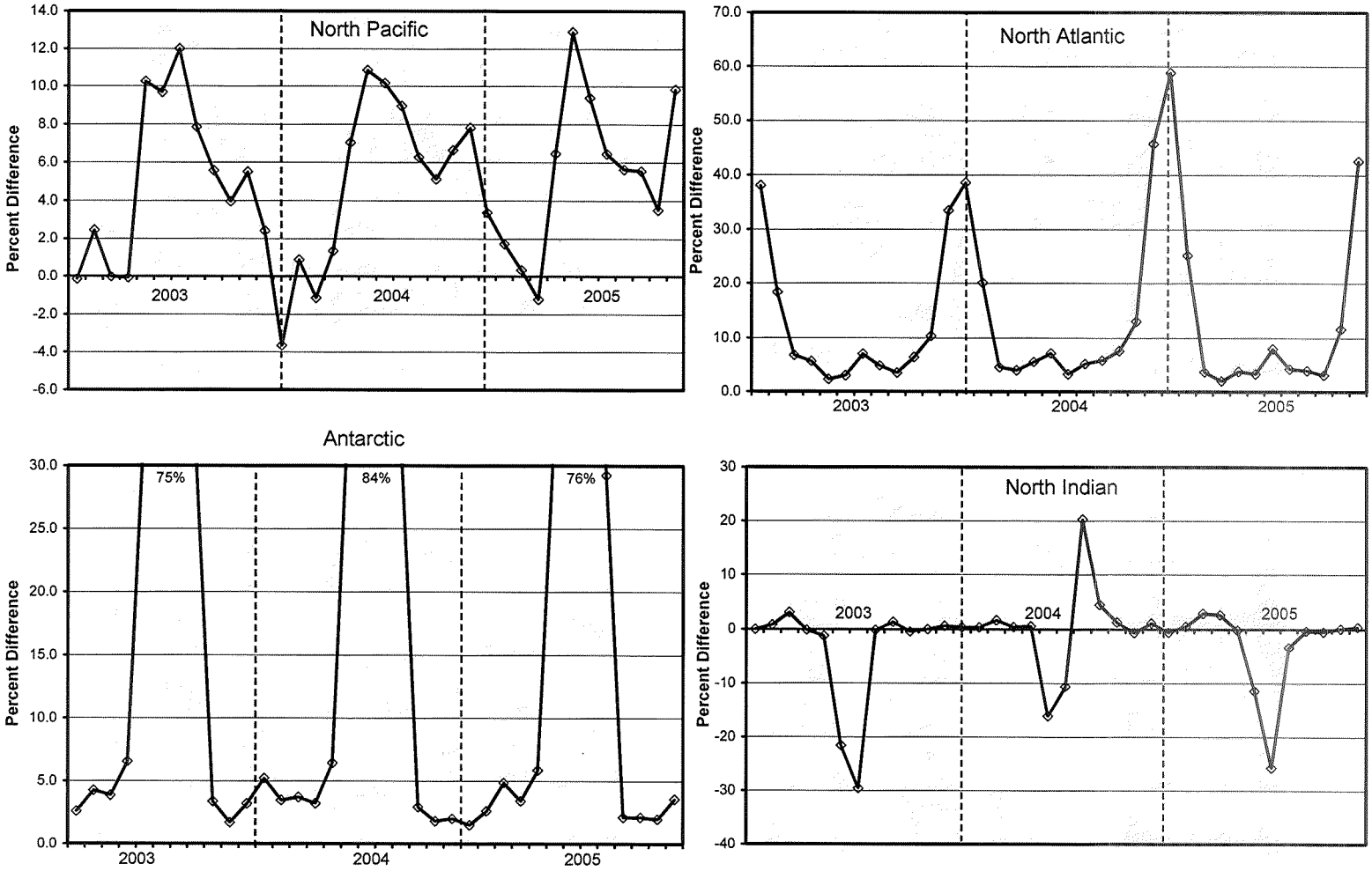


Figure 6. Monthly mean percent difference in chlorophyll between the assimilation and MODIS-Aqua sampling in 4 basins.

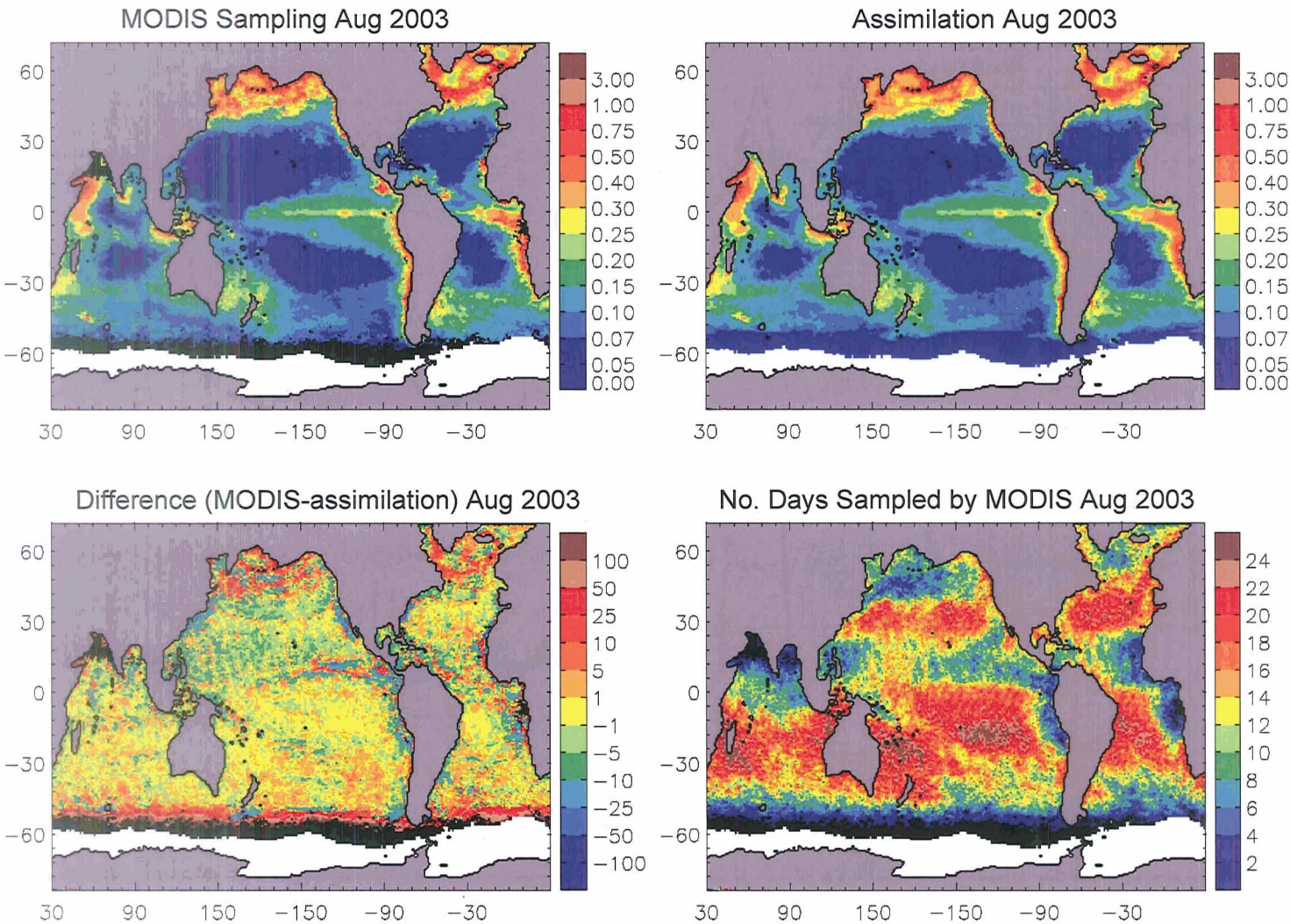


Figure 7. Top right: assimilation model chlorophyll for Aug 2003 (mg m^{-3}). Top left: chlorophyll from assimilation model sampled by MODIS-Aqua (mg m^{-3}). Bottom left: percent difference MODIS sampling – assimilation. Bottom right: number of days sampled by MODIS.

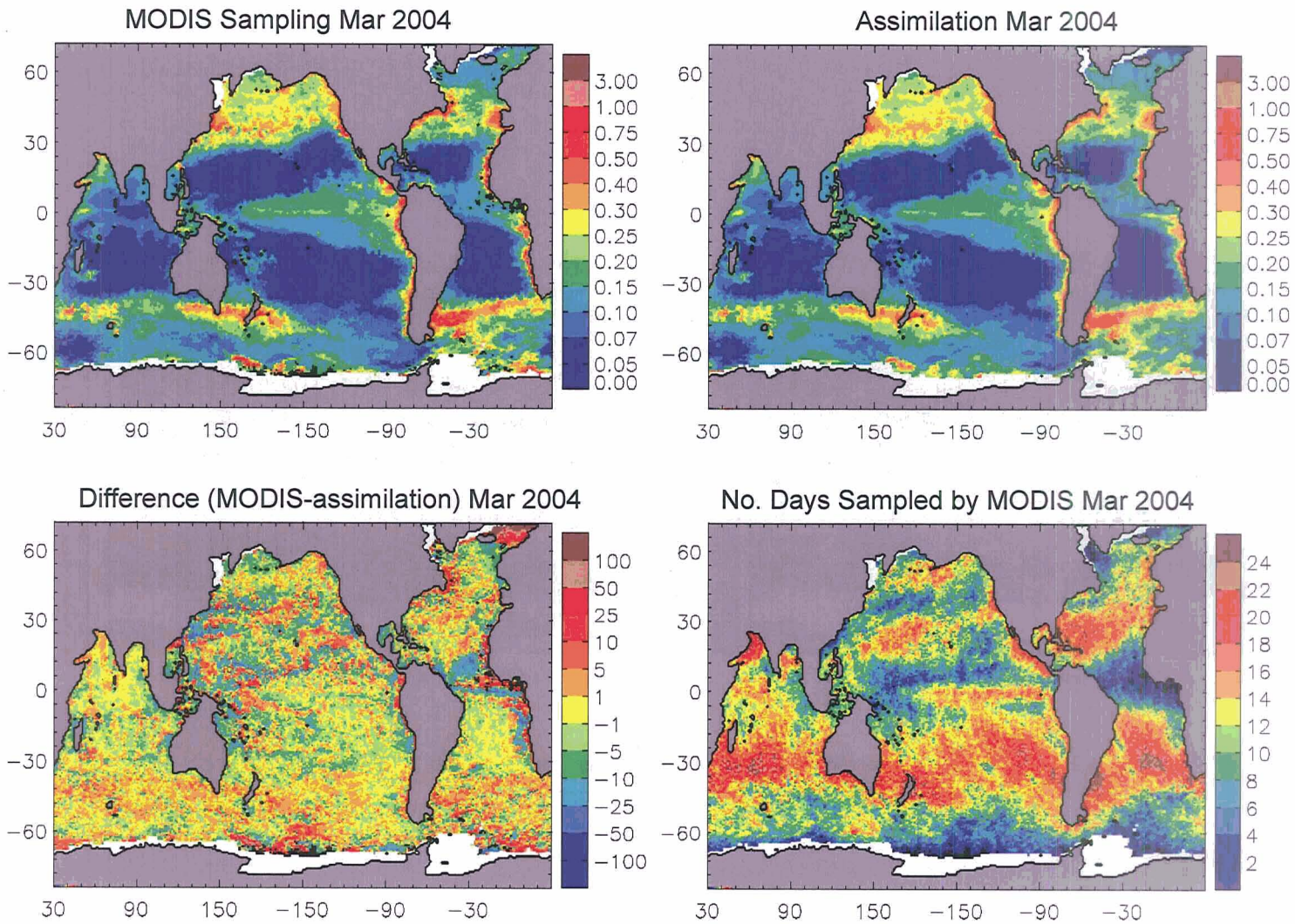


Figure 8. Same as Figure 7 for Mar 2004.

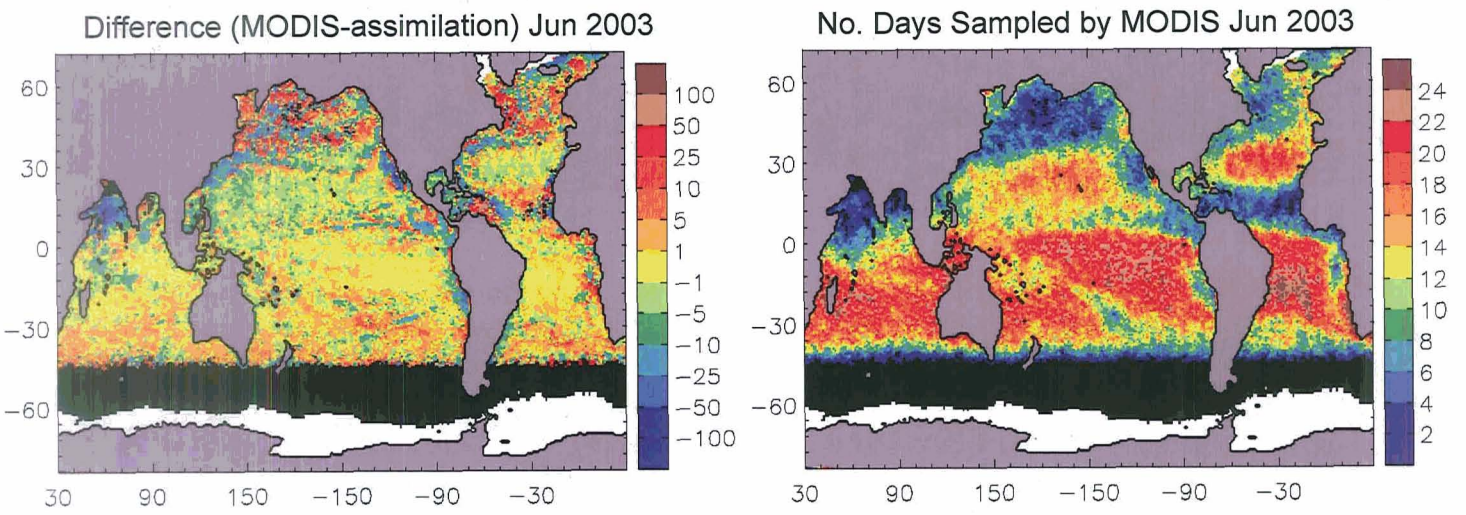
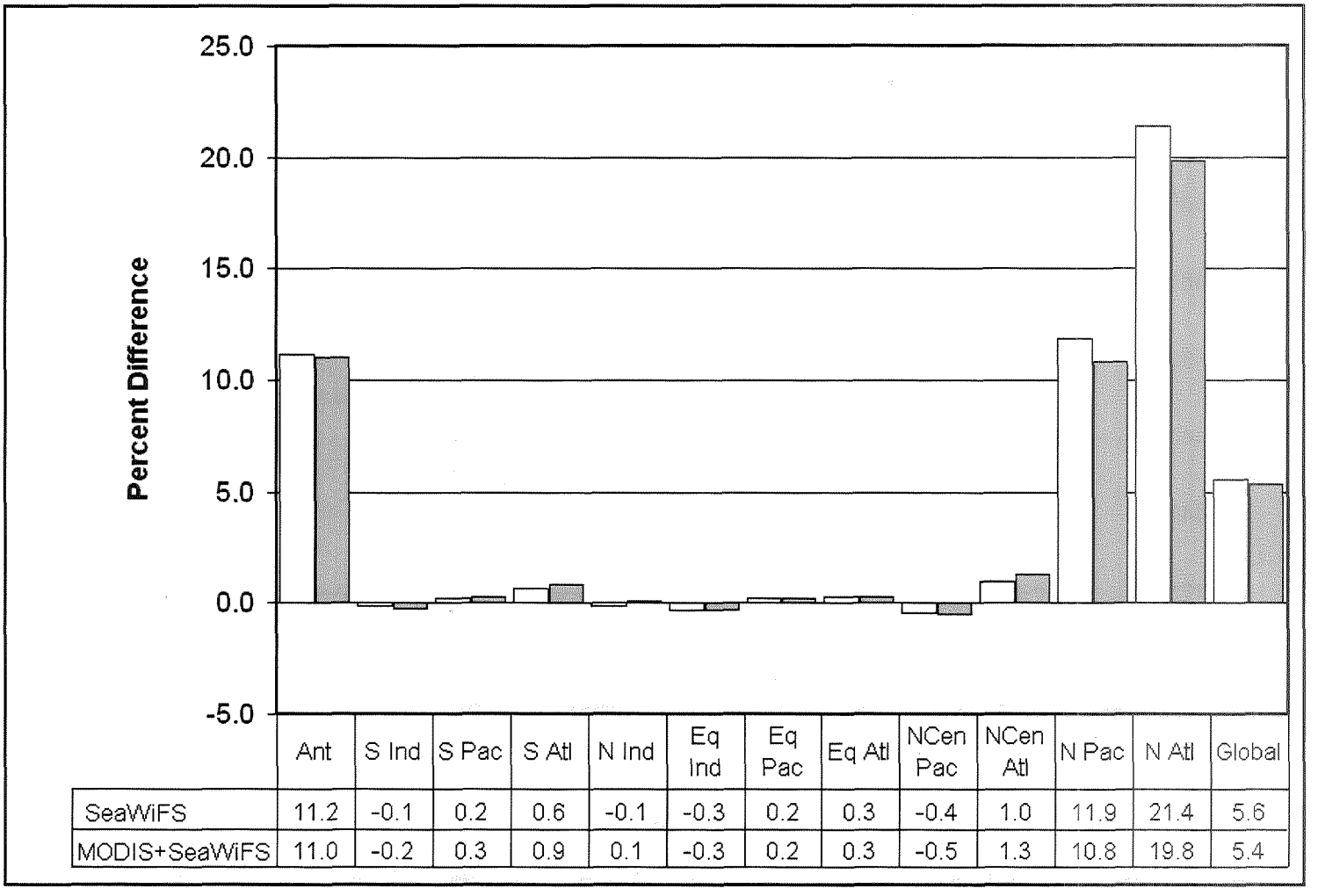


Figure 9. Difference between MODIS sampling and assimilation for Jun 2003 (left). Number of days of MODIS sampling (right).



	Ant	S Ind	S Pac	S Atl	N Ind	Eq Ind	Eq Pac	Eq Atl	NCen Pac	NCen Atl	N Pac	N Atl	Global
SeaWiFS	11.2	-0.1	0.2	0.6	-0.1	-0.3	0.2	0.3	-0.4	1.0	11.9	21.4	5.6
MODIS+SeaWiFS	11.0	-0.2	0.3	0.9	0.1	-0.3	0.2	0.3	-0.5	1.3	10.8	19.8	5.4

Figure 10. Annual mean difference between the assimilation model and SeaWiFS sampling of the assimilation (white bars). The difference between the combination of MODIS-Aqua+SeaWiFS and the assimilation model is shown by the dark bars.

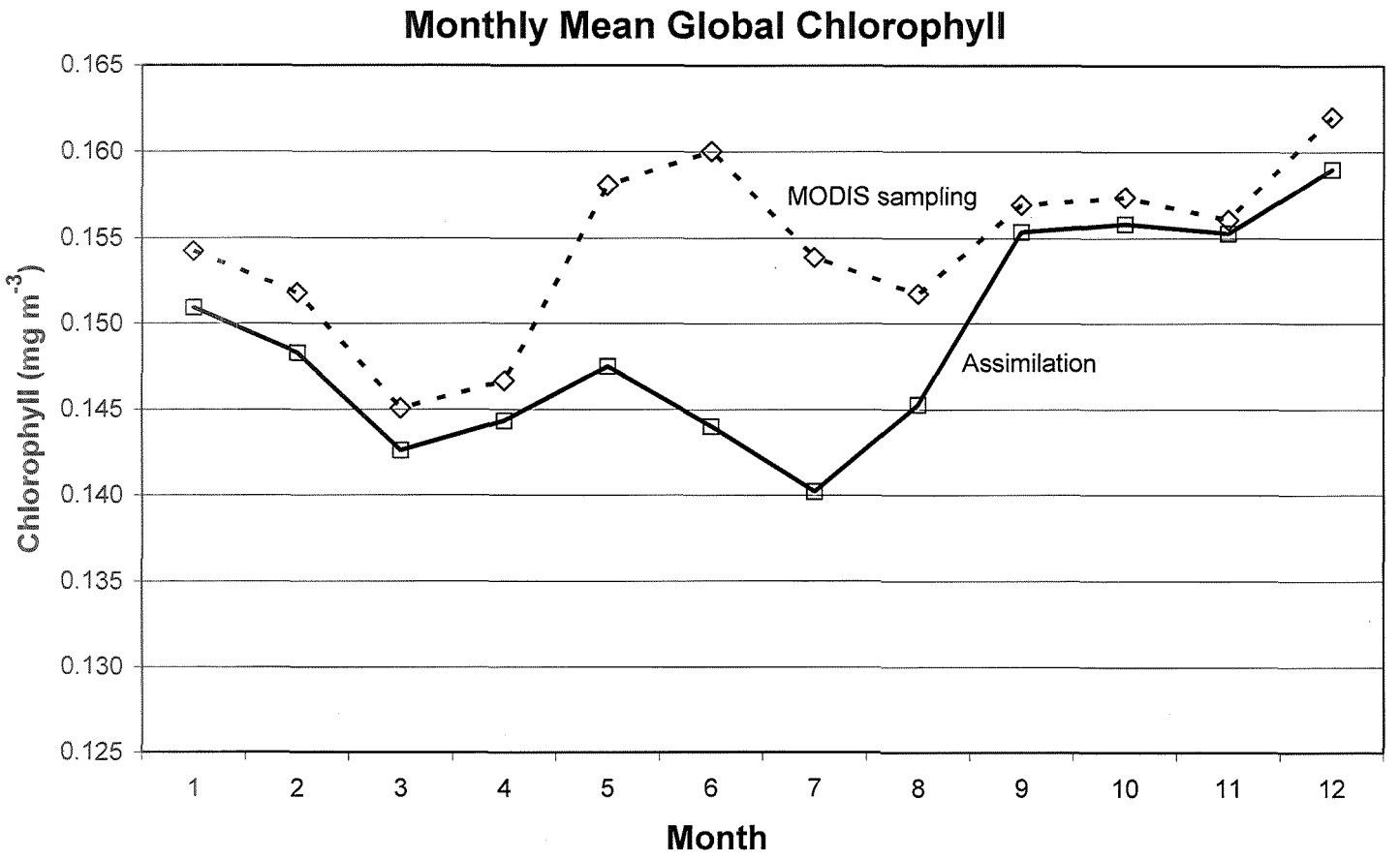


Figure 11. Monthly mean global chlorophyll from the assimilation, and from MODIS-Aqua sampling of the assimilation. The MODIS sampling produces a distorted seasonal trend, suggesting a bimodal seasonal cycle, due to lack of sampling in the Southern Ocean.

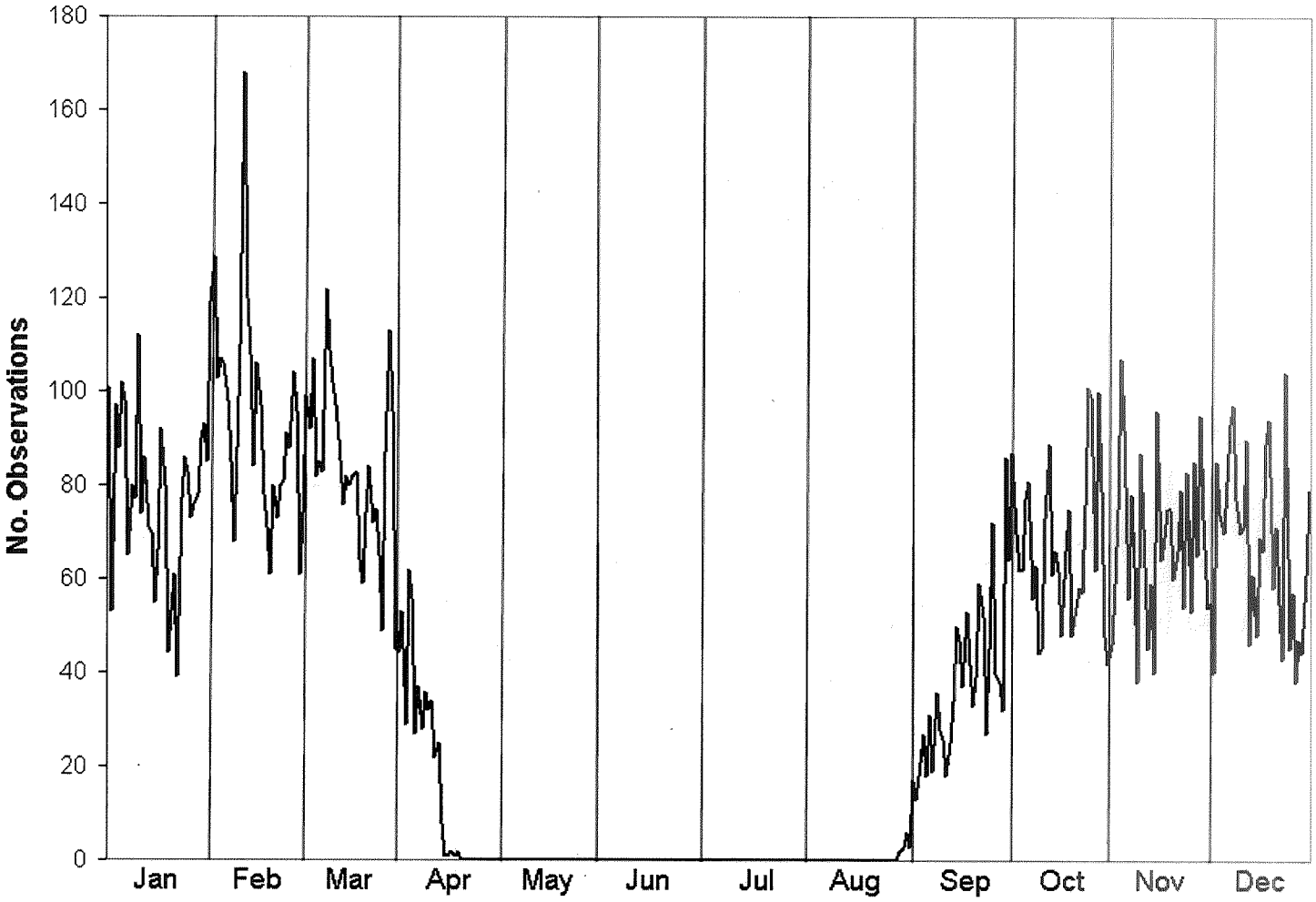


Figure 12. Number of observations per day for MODIS-Aqua at latitude -58°.

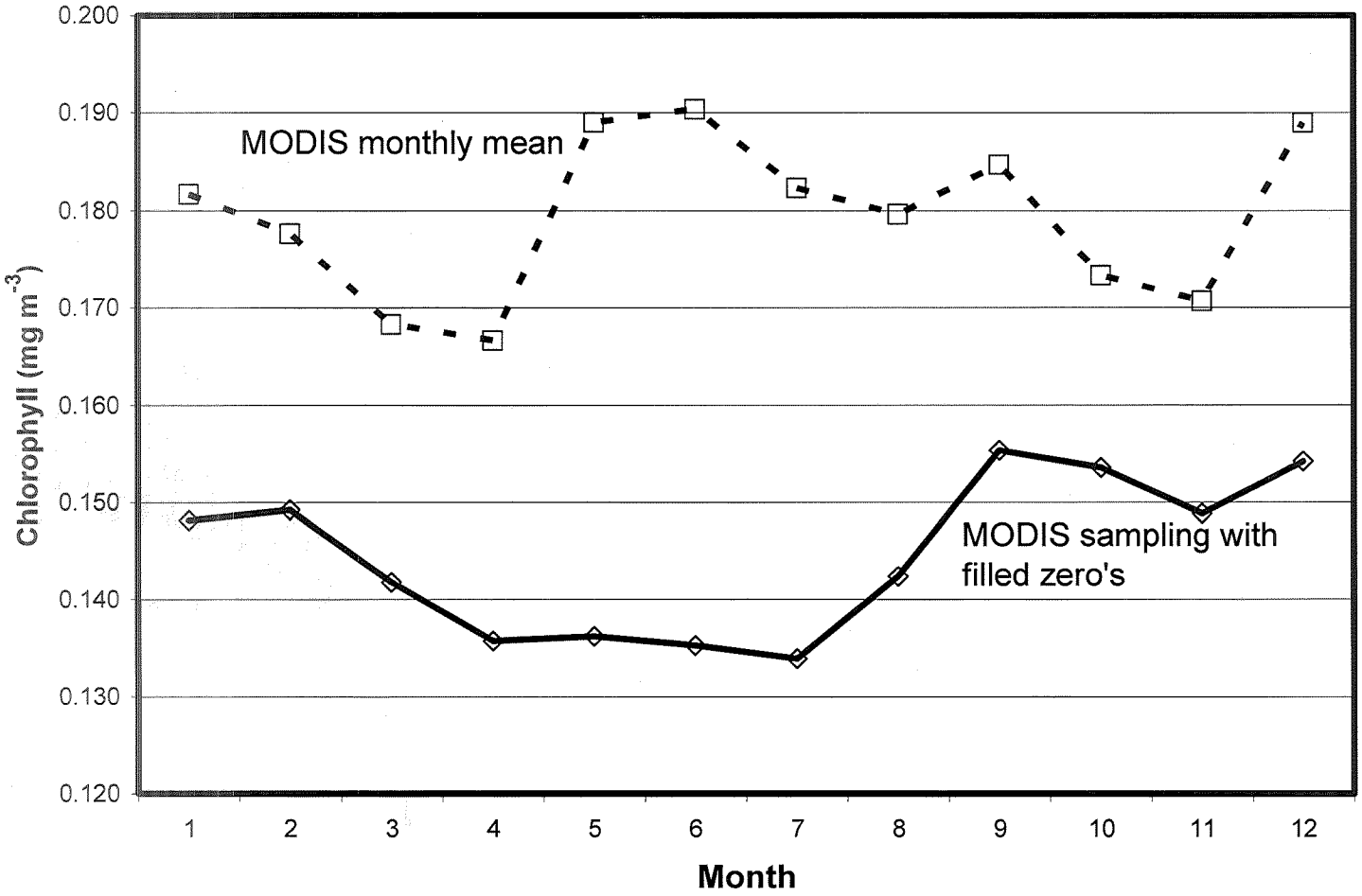


Figure 13. MODIS sampling of the assimilation field with missing data in high latitudes filled with zero's (solid line). Actual MODIS monthly mean data (dashed line), showing similarity with MODIS sampling in seasonal trends,