Biogeosciences, 10, 2711–2724, 2013
www.biogeosciences.net/10/2711/2013/
doi:10.5194/bg-10-2711-2013

Biogeosciences Open Access

Factors challenging our ability to detect long-term trends in ocean chlorophyll

C. Beaulieu[1], S. A. Henson[2], Jorge L. Sarmiento[1], J. P. Dunne[3], S. C. Doney[4], R. R. Rykaczewski[5], and L. Bopp[6]

[1]Program in Atmospheric and Oceanic Sciences, Princeton University, Princeton, New Jersey, USA
[2]National Oceanography Centre, Southampton, UK
[3]Geophysical Fluid Dynamics Laboratory, Princeton, New Jersey, USA
[4]Woods Hole Oceanographic Institution, Woods Hole, Massachusetts, USA
[5]Marine Science Program and Biological Sciences Department, University of South Carolina, Columbia, South Carolina, USA
[6]Laboratoire des Sciences du Climat et de l'Environnement, Gif sur Yvette, France

Correspondence to: C. Beaulieu (beaulieu@princeton.edu)

Received: 17 October 2012 – Published in Biogeosciences Discuss.: 20 November 2012
Revised: 5 March 2013 – Accepted: 20 March 2013 – Published: 23 April 2013

Abstract. Global climate change is expected to affect the ocean's biological productivity. The most comprehensive information available about the global distribution of contemporary ocean primary productivity is derived from satellite data. Large spatial patchiness and interannual to multi-decadal variability in chlorophyll *a* concentration challenges efforts to distinguish a global, secular trend given satellite records which are limited in duration and continuity. The longest ocean color satellite record comes from the Sea-viewing Wide Field-of-view Sensor (SeaWiFS), which failed in December 2010. The Moderate Resolution Imaging Spectroradiometer (MODIS) ocean color sensors are beyond their originally planned operational lifetime. Successful retrieval of a quality signal from the current Visible Infrared Imager Radiometer Suite (VIIRS) instrument, or successful launch of the Ocean and Land Colour Instrument (OLCI) expected in 2014 will hopefully extend the ocean color time series and increase the potential for detecting trends in ocean productivity in the future. Alternatively, a potential discontinuity in the time series of ocean chlorophyll *a*, introduced by a change of instrument without overlap and opportunity for cross-calibration, would make trend detection even more challenging. In this paper, we demonstrate that there are a few regions with statistically significant trends over the ten years of SeaWiFS data, but at a global scale the trend is not large enough to be distinguished from noise. We quantify the degree to which red noise (autocorrelation) especially challenges trend detection in these observational time series. We further demonstrate how discontinuities in the time series at various points would affect our ability to detect trends in ocean chlorophyll *a*. We highlight the importance of maintaining continuous, climate-quality satellite data records for climate-change detection and attribution studies.

1 Introduction

Global climate change is predicted to alter the ocean's biological productivity with implications for fisheries and climate. Results of coupled physical-biogeochemical models are sometimes inconsistent in their estimate of the magnitude and location of changes in marine primary production, depending on the region (e.g. Steinacher et al., 2010). Long-term (100 yr), rapidly declining trends in phytoplankton have been suggested through examination of shipboard measurements (Boyce et al., 2010). However, several authors have contested the methodology and implications of this study, some arguing that the long-term trend detected is an artifact of changes in the measurement techniques (Rykaczewski and Dunne, 2011; Mackas, 2011), and others reporting increases in chlorophyll *a* (hereafter chlorophyll) concentrations for regions that have been studied with consistent sampling methods over multi-decadal scales (e.g. Karl et al., 2001; Corno et al., 2007; Aksnes and Ohman, 2009;

Lomas et al., 2010; Saba et al., 2010; McQuatters-Gollop et al., 2011).

Satellite-derived ocean color and temperature data allow comprehensive estimates of the global distribution of ocean primary productivity, estimated from data provided by the Coastal Zone Color Scanner (CZCS), Ocean Color and Temperature Sensor (OCTS), Sea-viewing Wide Field-of-view Sensor (SeaWiFS), Moderate Resolution Imaging Spectroradiometer (MODIS) and Medium Resolution Imaging Spectrometer (MERIS) ocean color instruments. These data sets have allowed many scientific advances over the past decades as illustrated, for example, in McClain (2009). The CZCS sensor generated the first satellite ocean color data from November 1978 to June 1986; although focused primarily on coastal regions, CZCS also provided a picture of global patterns. However, it is not possible to determine trends from the CZCS record; the mission was a proof-of-concept, and thus the sensor was not continuously validated and suspected to drift after the first year of operation (Hooker and McClain, 2000; NRC, 2004). Ten years later the OCTS was launched and operated from July 1996 to June 1997. SeaWiFS became operational in September 1997 and remained remarkably stable for a decade, offering new opportunities for ocean biogeochemistry and climate research. Nevertheless, SeaWiFS began having telemetry problems in January 2008 and failed completely in December 2010. Three other ocean color sensors, MODIS-Terra, MERIS and MODIS-Aqua have been operating since December 1999, March and July 2002, respectively, but MERIS failed in May 2012. Several years of overlap between SeaWIFS and MODIS-Aqua allowed successful cross-calibrations to merge data from the two sensors (e.g. Fargion and McClain, 2003; Maritorena and Siegel, 2005; Pottier et al., 2006; Meister et al., 2012), increasing our potential for the detection of secular trends in ocean chlorophyll. However, MODIS Aqua and Terra are now beyond their operational lifetimes. Another ocean color instrument, the Visible Infrared Imager Radiometer Suite (VIIRS), is currently operational, but the data quality is, as of yet, undetermined.

Several authors have studied these satellite records in order to investigate trends in global ocean chlorophyll concentration and primary productivity (Gregg et al., 2005; Antoine et al., 2005; Behrenfeld et al., 2006; Vantrepotte and Mélin, 2009; Siegel et al., 2013) and to examine natural variability at the interannual and decadal time scales (Yoder and Kennelly, 2003; Martinez et al., 2009). However, none of these studies explicitly considers how the presence of autocorrelation may bias the ability to detect significant trends in the data. In climate time series, the autocorrelation is often represented by a first-order autoregressive process (red noise). The red noise arises from temporal persistence and it roughly approximates internal variability in the climate system in which slower response components such as the ocean and large ice sheets provide memory by responding slowly to a white noise forcing coming from weather systems (Hasselmann, 1976). Internal variability is also produced by coupled interactions between components of the climate system, such as climate oscillations (Hegerl et al., 2007). There is a great risk of misinterpreting changes in a relatively short time series when red noise is present, as it creates patterns that may be interpreted as trends or shifts with underlying mechanistic causality, but that are generated from a random process (e.g. Wunsch, 1999; Rudnick and Davis, 2003).

Long-term trends are detectable if the signal-to-noise ratio is large enough and a sufficient number of observations are available. Recent studies suggest that climate-change driven trends in satellite ocean color and inferred productivity are not yet distinguishable from red noise (Henson et al., 2010; Yoder et al., 2010). Due to the degree of internal variability in ocean productivity time series, approximately 39 yr of continuous data could be necessary to detect global climate-change-driven trends in ocean chlorophyll concentration and primary production (with a probability of detection of 0.9 and a significance level of 5 %) (Henson et al., 2010). This time frame assumes no interruption in satellite data – an unlikely scenario given the age of the current ocean color satellites (MODIS) and the unproven potential for VIIRS to provide the necessary data quality. The European Space Agency and the European Organisation for the Exploitation of Meteorological Satellites are planning the launch of the Ocean and Land Colour Instrument (OLCI) in 2013. If this is unsuccessful or MODIS fails before OLCI is operational, a discontinuity due to the change of instrument in the time series could seriously inhibit our ability to detect trends in ocean chlorophyll and productivity. Similarly, when OLCI exceeds its lifetime, a gap before launching a future satellite would again affect our ability to detect trends. Additional satellites have been launched or planned, but if the measurements are not made available for cross-calibration, the issue of potential discontinuity remains. Discontinuity can be introduced in the satellite records when a change of instrument occurs without an overlapping period during which the sensors in orbit may be cross-calibrated. While not ideal, the discontinuity due to a change of sensor might be estimated with some degree of uncertainty even without a period of overlap through careful calibration in orbit. However, with a period of overlap in orbit, discontinuities between sensors could be more accurately characterized through cross-calibration. Then, the magnitude and uncertainty of the discontinuity can be incorporated in the regression model used to detect long-term trends.

Discontinuities challenge the detection of trends in climate data since the discontinuity effect represents an additional parameter that must be estimated along with the magnitude of the trend, and therefore, longer time series of observations are required to achieve the same level of statistical confidence (Box and Tiao, 1975; Tiao et al., 1990; Weatherhead et al., 1998). For example, Weatherhead et al. (1998) estimated that in the worst-case scenario, a discontinuity could increase the number of years of data necessary to detect a linear trend (with a probability of detection of 90 %) by as

much as 50 %. Any potential discontinuity in satellite ocean color data must be taken into account in assessing the number of years of observations necessary to distinguish trends from internal variability in ocean chlorophyll and productivity.

The general objective of this study is to investigate the statistical factors that challenge the detection of trends in ocean color data and show why globally, and in most ocean basins, linear trends are not yet distinguishable from red noise (i.e., the internal variability) on a statistical basis. We use generalized least squares regression to detect trends in ocean chlorophyll satellite data and test the hypothesis that those trends detected are not an artifact of red noise. We quantify how a discontinuity in the time series would affect our ability to detect trends in ocean chlorophyll concentration given the observed variability and the expected trends estimated from a range of ocean models. More specifically, we assess how many additional years of satellite data would be needed to detect a trend if the current satellite fails before new satellite data are available. We also quantify how red noise affects the number of years of observations needed to detect trends in ocean chlorophyll concentration.

2 Data and methods

2.1 Data and models

We use monthly mean chlorophyll concentration data covering the January 1998–December 2007 period collected by SeaWiFS (version R2010.0; available at http://oceancolor.gsfc.nasa.gov/) averaged globally and in 14 biomes (shown in Fig. 1). The biome definition separates the regions where phytoplankton growth is seasonally light limited (for mid to high latitudes), regions where the ocean is gaining heat (equatorial regions) and oligotrophic regions (Henson et al., 2010). Observations taken after 2007 were not used as they are not continuous due to intermittent problems with the SeaWiFS instrument. The seasonal cycle was removed from the monthly means by subtracting from each month the mean of all observations taken during the same month for all years. To estimate long-term trends in surface ocean chlorophyll, we use the same three coupled physical-biogeochemical models as presented in Henson et al. (2010): GFDL-TOPAZ (Dunne et al., 2005, 2007), IPSL-PISCES (Aumont and Bopp, 2006) and NCAR-CCSM3 (Doney et al., 2009; Thornton et al., 2009). We estimate the trends in future climate change simulations forced with the IPCC A2 global warming scenario from 2001–2100 ($28.9\,\mathrm{Gt\,C\,yr^{-1}}$ from fossil fuel emissions by 2100). This scenario represents high cumulative carbon emissions due to human population growth and an increasing gap between the industrialized and developing nations (Nakicenovic et al., 2000). More details of the three model projections and biome definition are presented in Henson et al. (2010).

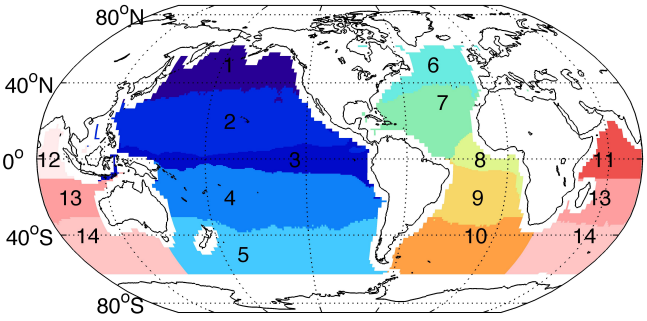

Fig. 1. Map of the 14 ocean biomes used in the analysis (1) High-latitude North Pacific, (2) Oligotrophic North Pacific, (3) Equatorial Pacific, (4) Oligotrophic South Pacific, (5) Southern Ocean Pacific, (6) High-latitude North Atlantic, (7) Oligotrophic North Atlantic, (8) Equatorial Atlantic, (9) Oligotrophic South Atlantic, (10) Southern Ocean Atlantic, 11) Arabian Sea, (12) Bay of Bengal, (13) Oligotrophic Indian, (14) Southern Ocean Indian.

2.2 Trend detection in presence of autocorrelation and discontinuity

A linear temporal trend can be expressed as

$$y_t = \mu + \omega t + N_t, \tag{1}$$

where y_t is the data (chlorophyll concentration) at time t, μ is the intercept, ω is the trend and N_t represents the residual noise at time t. This regression model was used in Henson et al. (2010) to represent trends in monthly mean chlorophyll and productivity data. When this model is fitted using ordinary least squares regression (OLS), it is assumed that the residuals are independent (white noise). However, it is often not reasonable to assume that successive observations of monthly chlorophyll concentration are independent from each other since there is memory being carried from month to month (red noise). In the presence of red noise, OLS tends to underestimate the variance and therefore inflates the test statistics on the regression coefficients, so that a trend can appear statistically significant when it is not (e.g. Wunsch, 1999).

Technically, we assume that the errors follow a first-order autoregressive process (AR(1)):

$$N_t = \phi N_{t-1} + \varepsilon_t, \tag{2}$$

where ϕ is the first-order autocorrelation and ε_t are normally distributed random errors (white noise) with a mean of zero and a common variance of σ_ε^2. It should be noted that the variance of the white noise process (ε_t) is directly related to the variance of the noise (N_t) by

$$\sigma_N^2 = \sigma_\varepsilon^2 / \left(1 - \phi^2\right), \tag{3}$$

where σ_N^2 is the noise variance. In this case, the first-order autoregressive process expresses the strength of the memory

being carried from one month to the other. The decorrelation time (or the time it will take in months to forget the current state of the process) is a more tangible measure of the strength of the memory and is expressed as a function of the first-order autocorrelation value ($(1+\phi)/(1-\phi)$). The regression parameters and their associated variance can be estimated using generalized least squares regression (GLS) to account for the presence of autocorrelation in the errors. To simplify these expressions we use matrix notation here. In matrix notation, the regression model presented in Eq. (1) is

$$\boldsymbol{y} = \mathbf{X}\boldsymbol{b} + \boldsymbol{N}, \tag{4}$$

where $\boldsymbol{y}$ is the $n \times 1$ data vector, $\mathbf{X}$ is a n $\times$ 2 design matrix with ones in the first column and the time in the second column, $\boldsymbol{b}$ is a 2×1 vector representing the intercept and trend and $\boldsymbol{N}$ is a $n \times 1$ vector representing the noise. The generalized least squares parameters ($\boldsymbol{b}$) and their variance ($\boldsymbol{V}(\boldsymbol{b})$) are given by

$$\boldsymbol{b} = \left(\mathbf{X}'\mathbf{S}^{-1}\mathbf{X}\right)^{-1}\mathbf{X}'\mathbf{S}^{-1}\boldsymbol{y} \tag{5}$$

$$\boldsymbol{V}(\boldsymbol{b}) = \left(\mathbf{X}'\mathbf{S}^{-1}\mathbf{X}\right)^{-1}, \tag{6}$$

where $\mathbf{S}$ is the $n \times n$ error-covariance matrix. The entries in the error covariance matrix represent the covariance of two errors depending on their separation in time s. For a first-order autoregressive process, the covariance of two residuals separated by s time units (months in this case) can be expressed as

$$C(N_t, N_{t+s}) = C(N_t, N_{t-s}) = \sigma_N^2 \phi^s. \tag{7}$$

More details on GLS estimation can be obtained in Brockwell and Davis (2002). Statistical analyses of chlorophyll concentrations are sometimes performed on log-transformed data (e.g. Campbell, 1995). However, a log-transformation did not help in stabilizing the variance of the model errors or making them more normally distributed, so we used the untransformed chlorophyll concentration by principle of parsimony.

Tiao et al. (1990) developed a simple equation allowing the estimation of the number of observations required to distinguish a trend (with a specified magnitude of ω_0) from red noise. The number of observations required depends on the signal-to-noise ratio and on the desired confidence level and power of detection for the test. It can be expressed as

$$n^* \approx \left[\frac{3.3\sigma_N}{|\omega_0|}\sqrt{\frac{1+\phi}{1-\phi}}\right]^{\frac{2}{3}}, \tag{8}$$

where n^* is the number of observations required to detect the trend. The magnitude of the trend can be expressed in absolute terms (e.g. changes in chlorophyll concentration in mg m^{-3} per year) or in relative terms (e.g. changes in percent per year or per decade). The factor 3.3 accounts for a power of detection of at least 0.90 and a significance level of 5 % (or a confidence level of 95 %) (Tiao et al., 1990; Weatherhead et al., 1998). The null hypothesis for the regression is that there is no trend. The significance level is the probability of incorrectly rejecting the null hypothesis when, in fact, it is true (false positive rate). The power of detection is the probability that the regression analysis will reject the null hypothesis when there is a trend (true positive rate). This means that if GLS regression is applied using a 5 % significance level to thousands of series having an approximate length of n^* and with the same parameters as presented in Eq. (8) (i.e. a lag-1 autocorrelation of ϕ, a variance of σ_N^2 and a trend with a magnitude of ω_0), in at least 90 % of the cases the trend would be detected if it exists. For a smaller significance level and/or a larger probability of detection, n^* would increase.

If a change of instrumentation or measurement procedures occurs, a discontinuity might be introduced in the observations and should be taken into account in the trend detection. An indicator function representing the effect of the change can be added to the model:

$$y_t = \mu + \omega t + \delta I_t + N_t \tag{9}$$

$$I_t = \begin{cases} 0, & t < T_0 \\ 1, & t \geq T_0 \end{cases}, \tag{10}$$

where I_t is a binary variable having a value of zero before the discontinuity and one after, T_0 represents the number of observations before the discontinuity and δ represents the magnitude (or the effect) of the discontinuity. The fraction of data before the discontinuity can be represented by

$$\tau = (T_0 - 1)/n \tag{11}$$

where n is the total number of observations.

Weatherhead et al. (1998) provided an estimate of n^* (with a specified trend magnitude of ω_0) in the presence of a discontinuity. They showed that n^* is larger in comparison to continuous data due to the additional parameter (δ) that needs to be estimated and depends on the time of the discontinuity:

$$n^* \approx \left[\frac{3.3\sigma_N}{|\omega_0|}\sqrt{\frac{1+\phi}{1-\phi}}\right]^{\frac{2}{3}} \frac{1}{[1-3\tau(1-\tau)]^{\frac{1}{3}}}. \tag{12}$$

It is important to note the distinction between a discontinuity and a gap. In the context of satellite data, a discontinuity would occur if no cross-calibration between one instrument and another was possible. In this case, n^* can be estimated using Eq. (12). In the presence of a gap only (i.e. the measurements are taken with the same instrument, but there is a gap in the time series), n^* can be estimated using Eq. (8) as there is no discontinuity effect to estimate.

Furthermore, it must be noted that the timing of the discontinuity (T_0) is known since we assume the discontinuity is caused by a change of instrument. However, if the timing of the discontinuity is unknown and needs to be estimated, Eq. (12) is not applicable and techniques developed to detect undocumented discontinuities would be required. Such techniques have been developed for in situ climate data for example (e.g. Peterson et al., 1998; Beaulieu et al., 2008).

2.3 Estimation of trends and noise in ocean color satellite data and biogeochemical models

To test for the presence of global and regional trends in ocean chlorophyll in SeaWiFS data, we fit the regression model presented in Eq. (1) to deseasonalized monthly mean SeaWiFS observations for 1998–2007, globally and in the 14 biomes presented in Fig. 1. We also estimate the first-order autocorrelation and standard deviation (Eqs. 2–3) in the ocean chlorophyll concentration from the deseasonalized monthly mean SeaWiFS observations. We estimate the trends predicted by the three ocean biogeochemical models also using the regression model presented in Eq. (1). For all the regression analyses, we verify the underlying assumptions of independently distributed normal errors with a constant variance using the Anderson–Darling normality test, the Breusch–Pagan homoscedasticity test and the Durbin–Watson independence test. All the statistical tests are performed using R (R Development Core team, 2008) and using a 5 % significance level.

We compute n^* for discontinuities at different points in the time series. The range of expected magnitudes in trends is estimated from the annual global and regional means of ocean chlorophyll for the 2001–2100 period in the three model runs described in the data and models section. Furthermore, we compute the number of years necessary to detect a trend for no discontinuity and the multi-model mean trend for different values of autocorrelation and standard deviation to show their effects on trend detection.

3 Results

3.1 Trends in satellite data and in ocean models

Table 1 presents the GLS trends for 1998–2007 in SeaWiFS satellite chlorophyll concentration for the global mean and average in 14 biomes. Globally and in most biomes, trends are not significant with the exception of the high-latitude North Atlantic (−1.3 % per year) and the Southern Ocean Pacific (0.65 % per year). In all cases, the Durbin–Watson test for the independence of the residual noise is significant, showing the necessity to take into account the autocorrelation in the analysis through GLS. Additional evidence for the presence of red noise is also presented in the Appendix A. Figure 2 presents the standardized (the mean is subtracted from the time series and then the time series is divided by its standard deviation) deseasonalized SeaWiFS chlorophyll concentration monthly means for all the regions. We present the standardized chlorophyll concentration to display all regions on the same scale since they have different ranges of concentrations. For the two biomes with significant trends (high-latitude North Atlantic and Southern Ocean Pacific), the fits are also presented. The deseasonalized SeaWiFS chlorophyll concentration monthly anomalies (not standardized) for all the regions are presented in the Appendix A.

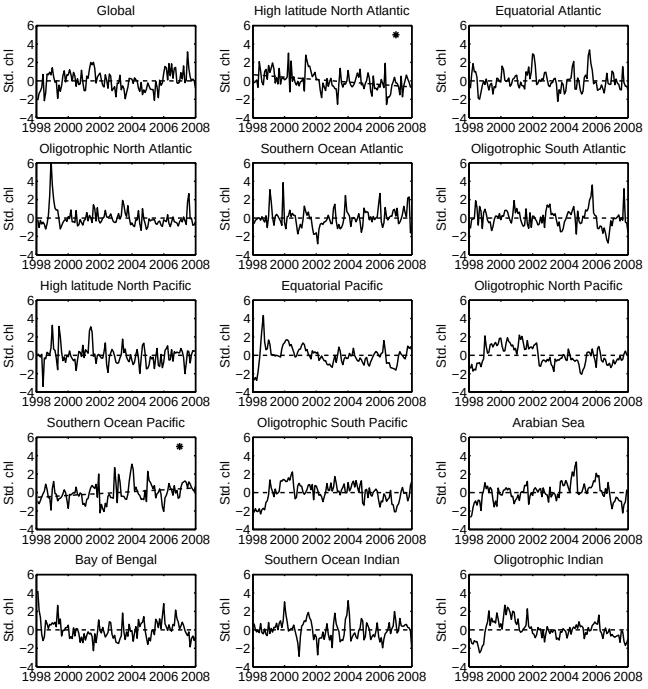

Fig. 2. Deseasonalized anomalies in SeaWiFS ocean chlorophyll concentrations from 1998–2007 averaged globally and in 14 biomes. Since the magnitude and variability varies among the different regions, the chlorophyll concentrations are standardized (the mean is subtracted from the time series and then the time series is divided by its standard deviation) to display on the same scale. The dotted lines represent the trends. The two panels with a star represent the trends that are significantly different from zero.

Table 2 presents the GLS trend estimates of global and biome-specific annual chlorophyll concentration in the three model projections for the period 2001–2100 (Fig. 3). We use output from three different models to give an estimate of the mean and range of possible future trends in surface chlorophyll concentration. Globally, IPSL-PISCES shows a strong and significant decreasing global trend. GFDL-TOPAZ also has a significantly decreasing global trend, but its magnitude is weaker than IPSL-PISCES. The global trend in NCAR-CCSM3 is not significant. In many biogeochemical models, the global trend reflects a balance between decreasing trends in some regions and increases in other regions. Thus, we also analyze the biomes trends that may give a clearer signal than the global mean. In the high-latitude North Atlantic, the

Table 1. Generalized least square trends, variability and autocorrelation of the SeaWiFS monthly ocean chlorophyll data. The trends (ω), noise standard deviation (σ_N) and autocorrelation (ϕ) are computed for the global mean and for the mean of each biome from January 1998–December 2007. For each region, we test whether the trends are significantly different from zero at the 5 % significance level.

Region	Chlorophyll concentration linear trends		
	ω[a] $\mathrm{mg\,m^{-3}\,yr^{-1}} \times 10^{-4}$ (% per year)	ϕ[b] $\mathrm{mg\,m^{-3}\,yr^{-1}}$	σ_N^2[c] $\times 10^{-2}$
Global mean	4.9 (0.11) [f]	0.52	1.26
High-latitude North Atlantic	−55.5 (−1.31)*,[f]	0.16	3.86
Equatorial Atlantic	−5.2 (−0.18)[d, f]	0.61	5.20
Oligotrophic North Atlantic	−12.7 (−0.95)[d, f]	0.59	2.02
Southern Ocean Atlantic	7.8 (0.26)[d, f]	0.39	2.31
Oligotrophic South Atlantic	−2.8 (−0.29)[e, f]	0.52	0.66
High-latitude North Pacific	−13.8 (−0.34)[d, e, f]	0.37	3.83
Equatorial Pacific	5.4 (0.30)[e, f]	0.87	2.51
Oligotrophic North Pacific	−3.0 (−0.34)[e, f]	0.84	0.60
Southern Ocean Pacific	10.8 (0.65)*, [d, f]	0.47	1.03
Oligotrophic South Pacific	2.0 (0.23)[e, f]	0.80	0.45
Arabian Sea	25.7 (1.45)[f]	0.73	2.69
Bay of Bengal	−7.6 (−0.50)[f]	0.59	1.51
Southern Ocean Indian	−4.4 (−0.21)[f]	0.59	1.03
Oligotrophic Indian	−2.9 (−0.32)[d, e, f]	0.80	0.62

[a] Linear trend as expressed in Eq. (1). The trends are computed using monthly data, but expressed yearly to be consistent with the trends computed using the models that are presented in Table 2.
* The trend is significantly different from zero, 5 % significance level.
[b] First-order autocorrelation of the noise estimated as presented in Eq. (2).
[c] Standard deviation of the noise as presented in Eq. (3).
[d] Normality hypothesis rejected, Anderson–Darling normality test, 5 % significance level.
[e] Homoskedasticity hypothesis rejected, Breusch–Pagan test, 5 % significance level.
[f] Independence hypothesis rejected, Durbin–Watson test, 5 % significance level.

IPSL-PISCES and NCAR-CCSM3 models project decreasing trends, while the GFDL-TOPAZ model projects an increasing trend. In the Southern Ocean Pacific, the trends projected by the GFDL-TOPAZ and NCAR-CCSM3 models are increasing. The IPSL-PISCES model projects a decreasing trend for the Southern Ocean Pacific. Whether the trends detected in the high-latitude North Atlantic and in the Southern Ocean Pacific in the SeaWiFS data might represent climate change or decadal variability cannot be answered without a detection and attribution study. Answering this question is even more challenging since the models often do not agree on the sign of the projected trends. Other regions where the three biogeochemical models do not all agree on the sign of the trend are Oligotrophic South Atlantic, Equatorial Pacific, Oligotrophic South Pacific and the Southern Ocean Indian.

3.2 Discontinuity and red noise effects on trend detection

Figure 4a presents the number of years of observation necessary to detect a global trend in satellite data for different trend magnitudes and different timing of a discontinuity. The range of the trend magnitude was set according to the global trend estimated from output of the three biogeochemical models. It can be seen that n^* increases with the fraction of data before the discontinuity. A discontinuity that occurs halfway in the time series (same number of observations before and after) has the most negative impact on trend detection (the results are symmetric above and below $\tau = 0.5$). Furthermore, trends of a small magnitude also need more observations to be detectable.

Approximately 27 yr of continuous observations would be needed to identify a trend in globally averaged chlorophyll concentration given the trend magnitude projected by the multi-model mean ($-1.53 \times 10^{-4}\,\mathrm{mg\,m^{-3}\,yr^{-1}}$) of the three biogeochemical models. This estimate is lower than the 39 yr estimate of Henson et al. (2010), because the latest SeaWiFS reprocessed version (R2010.0) used here has less variability than the version 5.2 used in the Henson et al. (2010) study. If a discontinuity occurs halfway through the time series, we estimate this would increase n^* from 27 to 43 yr. If the "real" global trend in the chlorophyll concentration were best represented by the IPSL-PISCES trend, then it would take approximately 13 yr of observations without discontinuity. On the other hand, if the "real" trend in the chlorophyll

Table 2. Ocean chlorophyll generalized least square trends for 2001–2100 estimated from the model outputs.

Region	Chlorophyll concentration linear trends $\mathrm{mg\,m^{-3}\,y^{-1}} \times 10^{-4}$ (% per year)		
	IPSL-PISCES	GFDL-TOPAZ	NCAR-CCSM3
Global mean	−4.4 (−0.21)*,f	−0.4 (−0.02)*	0. 2 (0.01)
High-latitude North Atlantic	−23.8 (−0.71)*,f	0.1 (0.00)*,d, f	−3.9 (−0.15)*,f
Equatorial Atlantic	−0.5 (−0.16)*,e, f	−1.4 (−0.11)*,e, f	−2.6 (−0.15)*,f
Oligotrophic North Atlantic	−1.3 (−0.24)*,d	−1.6 (−0.13)*	−1.5 (−0.10)*,e, f
Southern Ocean Atlantic	2.6 (0.09)*,f	1.6 (0.07)*,f	2.9 (0.09)*
Oligotrophic South Atlantic	−1.2 (−0.24)*,e	0.9 (0.07)*	−0.4 (−0.03)*,f
High-latitude North Pacific	−4.3 (−0.10)*,d, f	−3.1 (−0.10)*,f	−1.4 (−0.06)*
Equatorial Pacific	−4.2(−0.39)*,d	1.4 (0.06)*	−0.6 (−0.04)*,d
Oligotrophic North Pacific	−1.2 (−0.20)*	−2.2 (−0.23)*,e	−0.8 (−0.05)*,f
Southern Ocean Pacific	−3.9 (−0.16)*,f	1.1 (0.06)*,e	1.2 (0.05)*
Oligotrophic South Pacific	−1.3 (−0.23)*,e, f	0.2 (0.02)*,f	−0.3 (−0.02)*,d, f
Arabian Sea	−0.2 (−0.07)*,d	−2.9 (−0.29)*,e, f	−2.4 (−0.16)*,f
Bay of Bengal	−0.5 (−0.13)*,d, e	−1.9 (−0.15)*,e	−1.0 (−0.06)*,f
Southern Ocean Indian	−0.5 (0.02)f	0.9 (0.04)*,e	1.5 (0.07)*
Oligotrophic Indian	−1.3 (−0.22)*,f	−1.1 (−0.14)*,e	−0.7 (0.05)*,f

* The trend is significantly different from zero, 5 % significance level.
d Normality hypothesis rejected, Anderson–Darling normality test, 5 % significance level.
e Homoskedasticity hypothesis rejected, Breusch–Pagan test, 5 % significance level.
f Independence hypothesis rejected, Durbin–Watson test, 5 % significance level.

concentration is closer to the GFDL-TOPAZ trend, then it would take much longer to detect (from 66 yr of continuous observations to 105 observations with a discontinuity halfway through the time series). The NCAR-CCSM3 trend is not significantly different from zero, thus the estimation of the number of years of observations is not applicable here.

Figure 4b presents the effect of the first-order autocorrelation and standard deviation on the ability to detect trends of the same magnitude as the multi-model mean trend. One can see that strong, positive first-order autocorrelation and high standard deviation seriously inhibit the ability to detect trends. For example, in monthly observations with a standard deviation of $0.02\,\mathrm{mg\,m^{-3}}$ and a first-order autocorrelation of 0 (each month independent from the others), it would take approximately 25 yr of observations to detect a trend (without discontinuity) with the same magnitude as the multi-model mean trend ($-1.53 \times 10^{-4}\,\mathrm{mg\,m^{-3}\,yr^{-1}}$). In the presence of large monthly first-order autocorrelation (0.9), it would take approximately 65 yr of observations to detect the same trend. Figure 4b also presents the values observed in global monthly chlorophyll by SeaWiFS: a first-order autocorrelation of 0.52 and standard deviation of $0.01\,\mathrm{mg\,m^{-3}}$. The observed satellite autocorrelation corresponds to a decorrelation time of approximately 3 months, which is high enough that it must be taken into account in trend analyses (Fig. A2).

Figure 5 presents n^* in satellite data in six different biomes for different trend magnitudes and different times of discontinuity. The biomes were chosen to represent larger and smaller variability (quantified by the autocorrelation and variance of the red noise) and weaker and stronger trend projections. In the Equatorial Atlantic, which has relatively large variability, detecting a trend with the same magnitude as the biome multi-model mean trend would require approximately 54 yr of continuous observations and up to more than 90 yr of observations in presence of a discontinuity. In the Equatorial Pacific, even though the projected multi-model mean trend is smaller than in the Equatorial Atlantic, we would still require fewer years of observations to detect the trend (between 45 continuous years and up to 75 yr in presence of a discontinuity) since the variability is smaller in this biome. In the high-latitude North Pacific, which has large variability, we would require approximately between 37 yr (continuous) to 58 yr (discontinuity halfway) of observations to detect the multi-model mean trend. In the Oligotrophic North Pacific, which has very small variability and a relatively small projected decline, the multi-model mean trend should be detectable with approximately 29 yr of observations (or as much as 45 yr in presence of a discontinuity). The number of observations necessary to detect a trend in the eight remaining biomes is presented in the Appendix A (Fig. A4).

4 Discussion and conclusion

We have shown that ten-year trends in SeaWiFS ocean chlorophyll are distinguishable from the observed red noise in only two of our fourteen biomes: the high-latitude North Atlantic and the Southern Ocean Pacific (Fig. 1; Table 1).

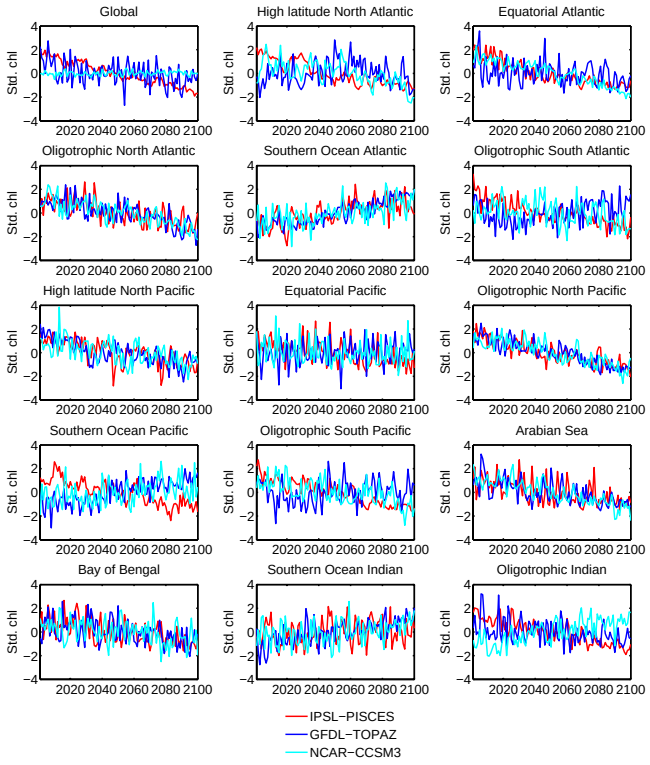

Fig. 3. Annual ocean chlorophyll concentrations projected for 2001–2100 averaged globally and in 14 biomes from three-ocean biogeochemical models forced with the A2 scenario from the Special Report on Emission Scenarios (Nakicenovic et al., 2000). The projections are standardized (the mean is subtracted from the time series and then the time series is divided by its standard deviation) to display concentrations on the same scale. The trends magnitudes are presented in Table 2.

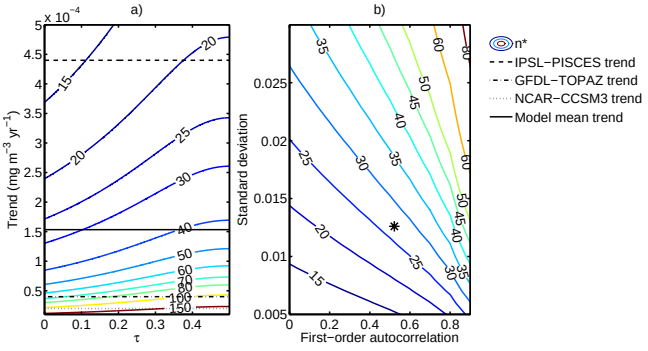

Fig. 4. (a) Number of years of observations (n^*) necessary to detect a statistically significant trend in satellite monthly ocean chlorophyll according to the magnitude of the trend and of the fraction of data before the discontinuity (τ). The standard deviation and autocorrelation used in the calculations were estimated from global SeaWiFS data from 1998–2007. The range for the trend magnitude was obtained from three biogeochemical models, and these magnitudes (in absolute value) are shown on the figure as well as the model mean trend (Table 2). We present the fraction of data before the discontinuity between 0 and 0.5 only since the results are symmetric. For example, the number of years necessary to detect a trend of the same magnitude will be the same if the discontinuity occurs after 25 % or 75 % of the data were collected. **(b)** Number of years of continuous observations necessary to detect a trend (with the magnitude of the multi-model mean trend) in satellite ocean chlorophyll according to the autocorrelation and standard deviation in the data. The star in the figure shows the values observed in global SeaWiFS data from 1998–2007.

The magnitude and sign of linear trends estimated in ten years of monthly SeaWiFS observations are in agreement with the trends detected by Vantrepotte and Mélin (2009) and Siegel et al. (2012), even though the analysis was performed on biome means in the present study. The high-latitude North Atlantic and the Southern Ocean Pacific biomes exhibit strong linear trend signals, but a formal detection and attribution study to distinguish between trends expected from natural variability (e.g. decadal oscillations, natural forcings) and climate change would be necessary to assign causality. Alternatively, a simple methodology that would facilitate the distinction of a long-term trend from a suspected dominant source of variability, such as the El Niño Southern Oscillation (ENSO) for example, could be applied. Such a method might include an ENSO index term in the regression model (Henson et al., 2010) to account for the variability related to ENSO and thus highlight any residual trend that may be attributable to other factors such as climate change.

We discussed the importance of accounting for the red noise in trend analyses of ocean productivity using ocean color satellite data and quantified its effect on the time necessary to detect a trend signal. In addition, we showed that a discontinuity in the satellite data measurements could have a large negative impact on our ability to understand ocean productivity's response to climate change. We estimate approximately 27 yr of continuous observations are required to detect global trends in surface chlorophyll concentration.

The above presents an idealized scenario. In reality, the scenario could be that the remaining ocean color sensors, MODIS Aqua and Terra, fail during the present year (2012) or next year (2013) before a new satellite is launched, and VIIRS fails to achieve the necessary data quality. During this period, no measurements would be taken, cross-calibration would not be possible and calibration would rely exclusively on in situ observations. Assuming that the real trend in chlorophyll is best represented by the mean trend of the three models, the trend would be distinguishable from the noise only after at least 40 yr (in 2037) of observations. This is because the discontinuity would occur at a time when approximately 43 % (17 yr) of the data were collected (as opposed to 27 yr of continuous observations, if there is no interruption). In this case, the discontinuity would amount to 13 additional years of observations necessary before a trend could be detected. This estimation does not include the duration of the gap (e.g. if the gap lasts two years, a trend would not be detectable until 2039). If the discontinuity occurs during

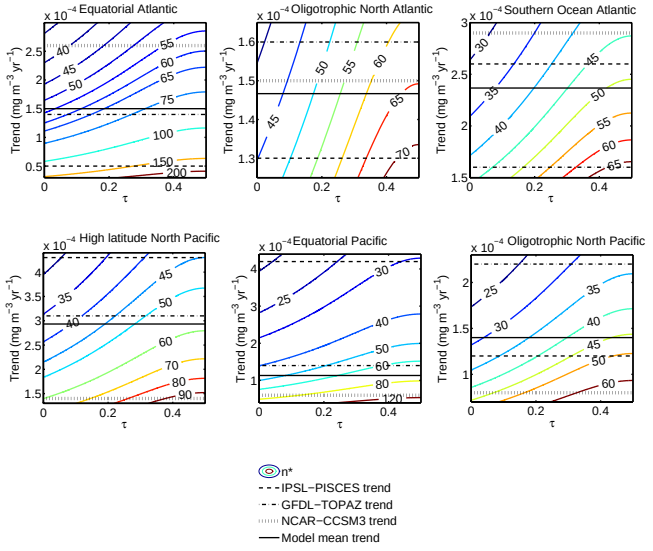

Fig. 5. Number of years of observations (n^*) necessary to detect a trend in satellite monthly ocean chlorophyll in six different biomes according to the magnitude of the trend and of the fraction of data before the discontinuity (τ). The standard deviation and autocorrelation used in the calculations were estimated from SeaWiFS data from 1998–2007. The range for the trend magnitude was obtained from three biogeochemical models, and these magnitudes (in absolute value) are shown on the figure as well as the model mean trend (Table 2). We present the fraction of data before the discontinuity between 0 and 0.5 only since the results are symmetric. For example, the number of years necessary to detect a trend of the same magnitude will be the same if the discontinuity occurs after 25 % or 75 % of the data were collected.

2018 instead, we would need 43 yr of observations, and the trend would not be detectable before 2040. In this case, the discontinuity occurs at a time when approximately 50 % of the data were collected. Overall, it could take an additional 13–16 yr of observations to detect a trend in satellite ocean chlorophyll under the idealized scenario if the two MODIS sensors fail and VIIRS remains of questionable reliability before OLCI is launched. Of course, this could be reduced if a cross-calibration was possible with another instrument that made overlapping measurements for a few years, so that a consistent, continuous time series was available. Other satellites have been launched or are planned, but if the data is not available for cross-calibration, the same problem will occur.

Several assumptions were made in this study and our results are valid only if they are reasonable. These assumptions include:

1. Following Henson et al. (2010), we made the assumption that the trends in ocean chlorophyll concentration are linear since the biogeochemical models used projected trends that are approximately linear over time. More development would be necessary in order to detect spatio-temporal trends, nonlinear trends or step-like behavior changes or to assess the number of years of observations required to do so.

2. To assess n^*, we used the trends projected by three ocean biogeochemical models coupled with climate models. We used several models to represent the uncertainty associated with the trend magnitude. The different trends in chlorophyll concentration estimated by each of the three models are due in part to their different representations of the ecosystem. In particular, changes in the relative proportions of large (diatoms) versus small phytoplankton can contribute substantially to the magnitude of the estimated climate-change driven trend (Steinacher et al., 2010). Diatoms and other large phytoplankton are expected to decline more rapidly in response to increasing nutrient limitation than small phytoplankton (e.g. Bopp et al., 2005), and so models that exhibit a more substantial contribution by large phytoplankton under current conditions may project larger climate-driven declines in chlorophyll than those that exhibit larger contributions by small phytoplankton as nutrient limitation increases in response to increased vertical stratification under climate change (Sarmiento et al., 2004). If real observed trends are greater than or less than the range predicted by the ocean biogeochemical models, n^* may be fewer or more than our estimates, respectively. For example, the trends could be different if other coupled climate-ocean biogeochemistry models were used. Coupled climate-ocean biogeochemistry models projections from the Coupled Model Intercomparison Project Phase 5 database that are or will be made available should be considered in future work.

Table 3. List of notations.

ε_t	errors following a white noise process
μ	intercept in the linear regression model
σ_ε^2	white noise variance
σ_N^2	red noise variance
τ	fraction of data before the discontinuity
ϕ	first-order autocorrelation
ϖ	magnitude of the trend
ϖ_0	expected magnitude of the trend
$\boldsymbol{b}$	2×1 vector of regression parameters
n	number of observations
n^*	number of observations necessary to detect a trend
$\boldsymbol{N}$	$n \times 1$ vector of errors following a red noise process
N_t	errors following a red noise process
$\mathbf{S}$	$n \times n$ error-covariance matrix
t	time
T_0	number of observations before the discontinuity
$\mathbf{X}$	$n \times 2$ matrix of explanatory variables
$\boldsymbol{y}$	$n \times 1$ vector of observations
y_t	time series of observations

3. Similarly, n^*could vary if a different biome definition was used and if we assumed the biomes were also expanding as suggested in Polovina et al. (2008, 2011). For example, if we were using a different biome definition exhibiting smaller internal variability and/or larger signal, n^* would decrease. For chlorophyll at greater depths, n^* may also vary, as it cannot be estimated using satellite data. Gliders and floats provide complementary information about the vertical structure of ocean chlorophyll as well as surface ocean chlorophyll in cloudy conditions (Boss et al., 2008; Perry et al., 2008). A global network of bio floats could provide additional opportunity to detect long-term trends in ocean chlorophyll concentration.

4. The results are based on the assumption that the average seasonal cycle remains the same year after year. Violation of this hypothesis might confuse trend detection. Some studies have suggested that the seasonal cycle in chlorophyll concentration may be changing with time in some regions (Vantrepotte and Mélin, 2009; Henson et al., 2013). However, since we do trend detection on biomes means, the changing seasonal cycle seems to be cancelling out when averaging and it seems reasonable to assume that the cycle approximately repeats itself year after year.

5. The results are also based on the assumption that the red noise estimated by the standard deviation and first-order autocorrelation from ten years of satellite data observations is representative of long-term internal variability. We assumed that these statistical properties are stationary, but the results could vary slightly if these properties change in time or if it takes longer to estimate internal variability.

6. We have used SeaWiFS satellite data only to make sure that we analyze a self-consistent time series. However, the results could vary if longer merged time series from different satellites (e.g. SeaWiFS and MODIS-Aqua) were used. Furthermore, we ignore the possible utility of other platforms such as ships, moorings, floats, gliders, and aircraft for estimating long-term trends in ocean chlorophyll and the use of these platforms to eliminate or reduce potential discontinuities when merging satellite data records.

7. Finally, we fixed the desired significance level at 5 % and a probability of detection of 0.9 to estimate the number of years of observations necessary to detect a trend, but if a larger significance level or a smaller probability of detection were targeted, it would require fewer observations to detect trends in ocean chlorophyll.

This work demonstrates the necessity of continuous monitoring of global ocean chlorophyll. This requires ensuring overlap in operation between satellites and collecting consistent in situ observations, so that validation, monitoring sensor degradation and cross-calibration of instruments is possible (NRC, 2004). In situ measurements have been successfully used to validate and reduce uncertainty in satellite ocean color data (e.g. McClain et al., 2009). Data from careful cross-calibrations that fully eliminate discontinuities across satellites should be capable of detecting trends with the same confidence as data from a single satellite. Cross-calibration methods allowing generation of unbiased time series from SeaWiFS-MODIS-VIIRS-OLCI would be crucially useful to increase our potential for the detection of climate change effects on ocean productivity.

Appendix A

Additional details on the data and models and additional results

In this appendix, we provide more details about the SeaWiFS data and the three models projections that were used. Figure A1 presents the SeaWiFS ocean chlorophyll concentrations anomalies. The variability is very different between the biomes. High-latitude North Atlantic, Equatorial Atlantic, the Southern Ocean Atlantic and the high-latitude North Pacific regions exhibit the largest variability, while the oligotrophic regions have very small variability in chlorophyll concentrations.

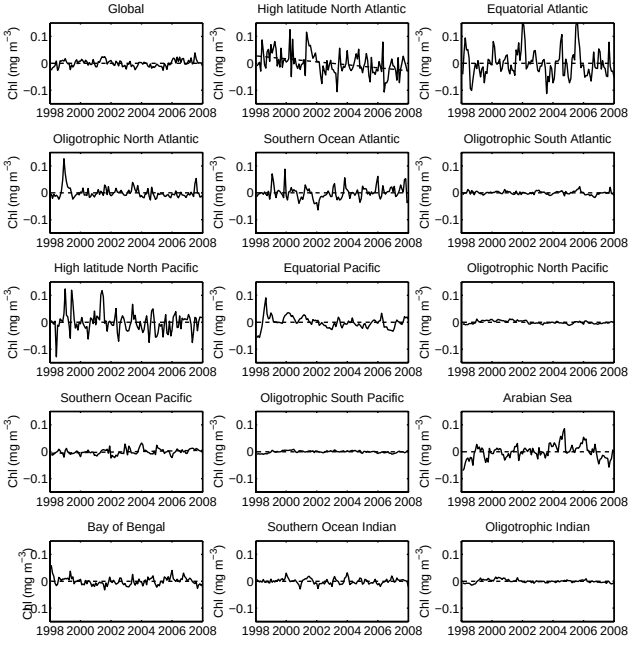

Fig. A1. SeaWiFS ocean chlorophyll concentrations anomalies from 1998–2007 averaged globally and in 14 biomes. The dotted lines represent the trends.

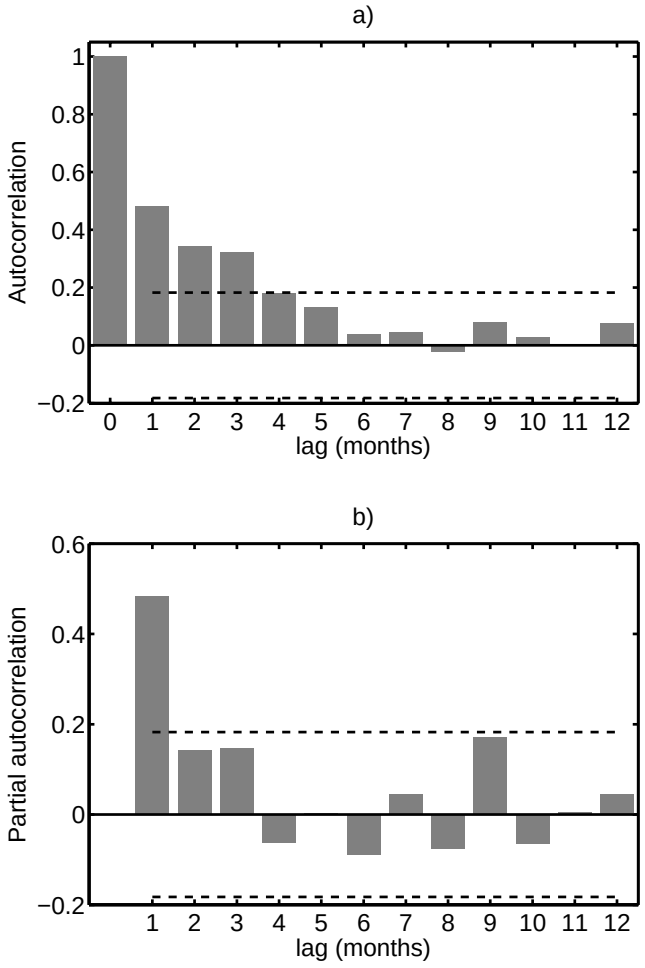

Fig. A2. (a) Sample autocorrelation function and **(b)** sample partial autocorrelation function of the SeaWiFS ocean global chlorophyll monthly anomalies from 1998–2007. The dashed lines present the 95 % confidence interval for the partial autocorrelation. The exponential decay in the autocorrelation function and significance of only the lag 1 in the partial autocorrelation function indicates a first-order autoregressive process (red noise).

Figure A2 presents the sample autocorrelation function and partial autocorrelation function of the SeaWiFS globally averaged anomalies in chlorophyll concentration. The autocorrelation function and partial autocorrelation functions are commonly used in autoregressive moving average model selection. For an autoregressive model of order p (AR(p)), the theoretical autocorrelation function tails off as an exponential decay or damped sine wave, and the theoretical partial autocorrelation function is equal to zero past lag-p (Wei, 1990).

The exponential decay shape of the autocorrelation function (Fig. A2a) and the partial autocorrelation function drop after lag-1 (Fig. A2b), indicating that a first-order autocorrelation model appropriately fits the noise and justifies the choice of GLS regression to study trends in ocean color chlorophyll concentration.

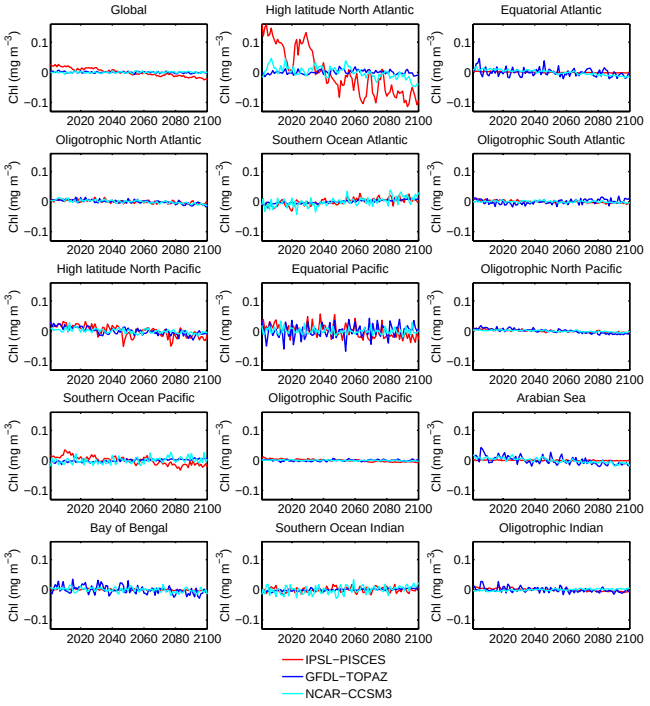

Fig. A3. Ocean chlorophyll concentrations anomalies projected for 2001–2100 averaged globally and in 14 biomes from three-ocean biogeochemical models forced with the A2 scenario from the Special Report on Emission Scenarios (Nakicenovic et al., 2000). The trends magnitudes and significance are presented in Table 2.

In order to present differences in variability of chlorophyll concentrations anomalies from the three models, we also present the model projections in all biomes (Fig. A3). The variability depends on the biome and model. In general, the oligotrophic regions also show the smallest variability, in agreement with the satellite data.

Figure A4 presents the number of observations necessary to detect a trend in the biomes that were not presented in Fig. 5: high-latitude North Atlantic, oligotrophic South Atlantic, Southern Ocean Pacific, oligotrophic South Pacific, Arabian Sea, Bay of Bengal, Southern Ocean Indian and oligotrophic Indian.

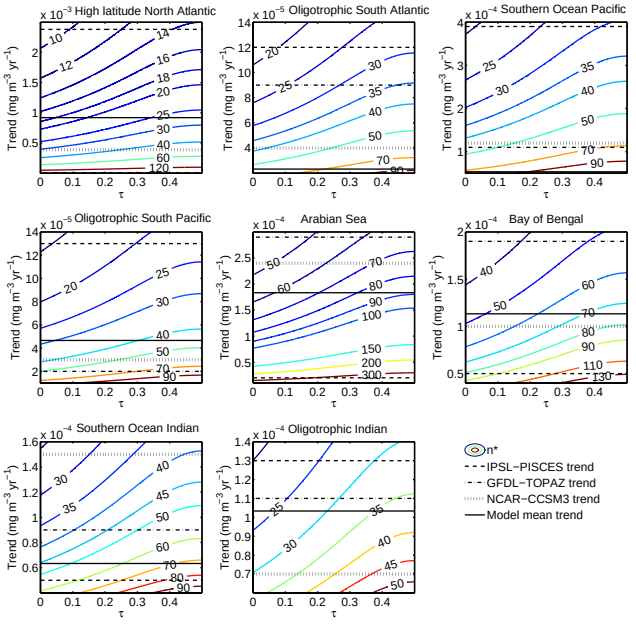

Fig. A4. Number of years of observations (n^*) necessary to detect a trend in satellite monthly ocean chlorophyll in eight different biomes according to the magnitude of the trend and of the fraction of data before the discontinuity (τ). The standard deviation and autocorrelation used in the calculations were estimated from SeaWiFS data from 1998–2007. The range for the trend magnitude was obtained from three biogeochemical models and these magnitudes (in absolute value) are shown on the figure as well as the model mean trend. We present the fraction of data before the discontinuity between 0 and 0.5 only since the results are symmetric. For example, the number of years necessary to detect a trend of the same magnitude will be the same if the discontinuity occurs after 25 % or 75 % of the data was collected, assuming a linear trend.

Acknowledgements. The authors thank GSFC/NASA for providing the SeaWiFS data that was used in accordance with the SeaWiFS Research Data Use Terms and Conditions Agreement. The authors also thank V. Saba and two anonymous reviewers for useful comments on the manuscript. CB and JLS acknowledge financial support from the Carbon Mitigation Initiative with support from BP. JLS and RRR were partly supported by the NF-UBC Nereus Program. SAH was supported by NERC grant NE/G013055/1. SCD acknowledges support from NSF grant EF-0424599.

Edited by: G. Herndl

References

Aksnes, D. L. and Ohman, M. D.: Multi-decadal shoaling of the euphotic zone in the southern sector of the California Current System, Limnol. Oceanogr., 54, 1272–1281, doi:10.4319/lo.2009.54.4.1272, 2009.

Antoine, D., Morel, A., Gordon, H. R., Banzon, V. F., and Evans, R. H.: Bridging ocean color observations of the 1980s and 2000s in search of long-term trends, J. Geophys. Res.-Oceans, 110, C06009, doi:10.1029/2004jc002620, 2005.

Aumont, O. and Bopp, L.: Globalizing results from ocean in situ iron fertilization studies, Global Biogeochem. Cy., 20, GB2017, doi:10.1029/2005gb002591, 2006.

Beaulieu, C., Seidou, O., Ouarda, T., Zhang, X., Boulet, G., and Yagouti, A.: Intercomparison of homogenization techniques for precipitation data, Water Resour. Res., 44, W02425, doi:10.1029/2006wr005615, 2008.

Behrenfeld, M. J., O'Malley, R. T., Siegel, D. A., McClain, C. R., Sarmiento, J. L., Feldman, G. C., Milligan, A. J., Falkowski, P. G., Letelier, R. M., and Boss, E. S.: Climate-driven trends in contemporary ocean productivity, Nature, 444, 752–755, doi:10.1038/nature05317, 2006.

Bopp, L., Aumont, O., Cadule, P., Alvain, S., and Gehlen, M.: Response of diatoms distribution to global warming and potential implications: A global model study, Geophys. Res. Lett., 32, L19606, doi:10.1029/2005GL023653, 2005.

Boss, E., Swift, D., Taylor, L., Brickley, P., Zaneveld, R., Riser, S., Perry, M. J., and Strutton, P. G.: Observations of pigment and particle distributions in the western North Atlantic from an autonomous float and ocean color satellite, Limnol. Oceanogr., 53, 2112–2122, doi:10.4319/lo.2008.53.5_part_2.2112, 2008.

Box, G. E. P. and Tiao, G. C.: Intervention analysis with applications to economic and environmental problems, J. Am. Stat. Assoc., 70, 70–79, doi:10.2307/2285379, 1975.

Boyce, D. G., Lewis, M. R., and Worm, B.: Global phytoplankton decline over the past century, Nature, 466, 591–596, doi:10.1038/nature09268, 2010.

Brockwell, P. J. and Davis, R. A.: Introduction to time series and forecasting, 2nd Edn., Springer, New York, NY, 2002.

Campbell, J. W.: The Lognormal distribution as a model for bio-optical variability in the sea, J. Geophys. Res.-Oceans, 100, 13237–13254, doi:10.1029/95jc00458, 1995.

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

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

Dunne, J. P., Armstrong, R. A., Gnanadesikan, A., and Sarmiento, J. L.: Empirical and mechanistic models for the particle export ratio, Global Biogeochem. Cy., 19, GB4026, doi:10.1029/2004gb002390, 2005.

Dunne, J. P., Sarmiento, J. L., and Gnanadesikan, A.: A synthesis of global particle export from the surface ocean and cycling through the ocean interior and on the seafloor, Global Biogeochem. Cy., 21, GB4006, doi:10.1029/2006gb002907, 2007.

Fargion, G. S. and McClain, C. R.: MODIS validation, data merger and other activities accomplished by the SIMBIOS project: 2002-2003, NASA/TM-2003-212249, National Aeronautics and Space Administration, Goddard Space Flight center, Greenbelt, Maryland, 2003.

Gregg, W. W., Casey, N. W., and McClain, C. R.: Recent trends in global ocean chlorophyll, Geophys. Res. Lett., 32, L03606, doi:10.1029/2004gl021808, 2005.

Hasselmann, K.: Stochastic climate models. 1. Theory, Tellus, 28, 473–485, 1976.

Hegerl, G. C., Zwiers, F. W., Braconnot, P., Gillett, N. P., Luo, Y., Marengo Orsini, J. A., Nicholls, N., Penner J. E., and Stott, P. A.: Understanding and Attributing Climate Change, in: Climate Change 2007: The Physical Science Basis, Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change, Cambridge University Press, Cambridge, UK and New York, NY, USA, 2007.

Henson, S., Cole, H., Beaulieu, C., and Yool, A.: The impact of global warming on seasonality of ocean primary production, Biogeosciences Discuss., 10, 1421–1450, doi:10.5194/bgd-10-1421-2013, 2013.

Henson, S. A., Sarmiento, J. L., Dunne, J. P., Bopp, L., Lima, I., Doney, S. C., John, J., and Beaulieu, C.: Detection of anthropogenic climate change in satellite records of ocean chlorophyll and productivity, Biogeosciences, 7, 621–640, doi:10.5194/bg-7-621-2010, 2010.

Hooker, S. B. and McClain, C. R.: The calibration and validation of SeaWiFS data, Prog. Oceanogr., 45, 427–465, doi:10.1016/s0079-6611(00)00012-4, 2000.

Karl, D. M., Bidigare, R. R., and Letelier, R. M.: Long-term changes in plankton community structure and productivity in the North Pacific subtropical gyre: The domain shift hypothesis, Deep-Sea Res. II, 48, 1449–1470, doi:10.1016/S0967-0645(00)00149-1, 2001.

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

Mackas, D. L.: Does blending of chlorophyll data bias temporal trend?, Nature, 472, E4–E5, doi:10.1038/nature09951, 2011.

Maritorena, S. and Siegel, D. A.: Consistent merging of satellite ocean color data sets using a bio-optical model, Remote Sens. Environ., 94, 429-440, doi:10.1016/j.rse.2004.08.014, 2005.

Martinez, E., Antoine, D., D'Ortenzio, F., and Gentili, B.: Climate-Driven Basin-Scale Decadal Oscillations of Oceanic Phytoplankton, Science, 326, 1253–1256, doi:10.1126/science.1177012, 2009.

McClain, C. R.: A Decade of Satellite Ocean Color Observations, Annual Review of Marine Science, 1, 19–42, doi:10.1146/annurev.marine.010908.163650, 2009.

McQuatters-Gollop, A., Reid, P. C., Edwards, M., Burkill, P. H., Castellani, C., Batten, S., Gieskes, W., Beare, D., Bidigare, R. R., Head, E., Johnson, R., Kahru, M., Koslow, J. A., and Pena, A.: Is there a decline in marine phytoplankton?, Nature, 472, E6–E7, doi:10.1038/nature09950, 2011.

Meister, G., Franz, B. A., Kwiatkowska, E. J., and McClain, C. R.: Corrections to the Calibration of MODIS Aqua Ocean Color Bands Derived From SeaWiFS Data, IEEE T. Geosci. Remote, 50, 310–319, doi:10.1109/tgrs.2011.2160552, 2012.

Nakicenovic, N., Alcamo, J., Davis, G., de vries, B., Fenhann, J., Gaffin, S., Gregory, K., Grübler, A., Tae Yong, J., Kram, T., La Rovere, E. L., Michaelis, L., Mori, S., Morita, T., Pepper, W., Pitcher, H., Price, L., Riahi, K., Roehrl, A., Rogner, H.-H., Sankovski, A., Schlesinger, M., Shukla, P., Smith, S., Swart, R., van Rooijen, S., Victor, N., and Dadi, Z.: Special report on emissions scenarios., Working Group III of the Intergovernmental Panel on Climate Change, Cambridge University Press, Cambridge, 2000.

National Research Council – Committee on Climate Data Records from NOAA Operational Satellites: Climate data records from environmental satellites, National Academies Press, Washington, DC, 2004.

Perry, M. J., Sackmann, B. S., Eriksen, C. C., and Lee, C. M.: Seaglider observations of blooms and subsurface chlorophyll maxima off the Washington coast, Limnol. Oceanogr., 53, 2169–2179, doi:10.4319/lo.2008.53.5_part_2.2169, 2008.

Peterson, T. C., Easterling, D. R., Karl, T. R., Groisman, P., Nicholls, N., Plummer, N., Torok, S., Auer, I., Boehm, R., Gullett, D., Vincent, L., Heino, R., Tuomenvirta, H., Mestre, O., Szentimrey, T., Salinger, J., Forland, E. J., Hanssen-Bauer, I., Alexandersson, H., Jones, P., and Parker, D.: Homogeneity adjustments of in situ atmospheric climate data: A review, Int. J. Climatol., 18, 1493–1517, doi:10.1002/(sici)1097-0088(19981115)18:13< 1493::aid-joc329>3.0.co;2-t, 1998.

Polovina, J. J., Howell, E. A., and Abecassis, M.: Ocean's least productive waters are expanding, Geophys. Res. Lett., 35, L03618, doi:10.1029/2007gl031745, 2008.

Polovina, J. J., Dunne, J. P., Woodworth, P. A., and Howell, E. A.: Projected expansion of the subtropical biome and contraction of the temperate and equatorial upwelling biomes in the North Pacific under global warming, ICES J. Mar. Sci., 68, 986–995, doi:10.1093/icesjms/fsq198, 2011.

Pottier, C., Garcon, V., Larnicol, G., Sudre, J., Schaeffer, P., and Le Traon, P. Y.: Merging SeaWiFS and MODIS/Aqua ocean color data in North and Equatorial Atlantic using weighted averaging and objective analysis, IEEE T. Geosci. Remote, 44, 3436–3451, doi:10.1109/tgrs.2006.878441, 2006.

R Development Core Team, R: A Language and Environment for Statistical Computing, R Foundation for Statistical Computing, 409 pp., available at: http://www.R-project.org., 2008

Rudnick, D. L. and Davis, R. E.: Red noise and regime shifts, Deep-Sea Res. I, 50, 691–699, doi:10.1016/s0967-0637(03)00053-0, 2003.

Rykaczewski, R. R. and Dunne, J. P.: A measured look at ocean chlorophyll trends, Nature, 472, E5–E6, doi:10.1038/nature09952, 2011.

Saba, V. S., Friedrichs, M. A. M., Carr, M. E., Antoine, D., Armstrong, R. A., Asanuma, I., Aumont, O., Bates, N. R., Behrenfeld, M. J., Bennington, V., Bopp, L., Bruggeman, J., Buitenhuis, E. T., Church, M. J., Ciotti, A. M., Doney, S. C., Dowell, M., Dunne, J., Dutkiewicz, S., Gregg, W., Hoepffner, N., Hyde, K. J. W., Ishizaka, J., Kameda, T., Karl, D. M., Lima, I., Lomas, M. W., Marra, J., McKinley, G. A., Melin, F., Moore, J. K., Morel, A., O'Reilly, J., Salihoglu, B., Scardi, M., Smyth, T. J., Tang, S. L., Tjiputra, J., Uitz, J., Vichi, M., Waters, K., Westberry, T. K., and Yool, A.: Challenges of modeling depth-integrated marine primary productivity over multiple decades: A case study at BATS and HOT, Global Biogeochem. Cy., 24, GB3020, doi:10.1029/2009gb003655, 2010.

Sarmiento, J. L., Slater, R., Barber, R., Bopp, L., Doney, S. C., Hirst, A. C., Kleypas, J., Matear, R., Mikolajewicz, U., Monfray, P., Soldatov, V., Spall, S. A., and Stouffer, R.: Response of ocean ecosystems to climate warming, Global Biogeochem. Cy., 18, GB3003, doi:10.1029/2003gb002134, 2004.

Siegel, D. A., Behrenfeld, M. J., Maritorena, S., McClain, C. R., Antoine, D., Bailey, S. W., Bontempi, P. S., Boss, E. S., Dierssen, H. M., Doney, S. C., Eplee, R. E. J., Evans, R. H., Feldman, G. C., Fields, E., Franz, B. A., Kuring, N. A., Mengalt, C., Nelson, N. B., Patt, F. S., Robinson, W. S., Sarmiento, J. S., Swan, C. M., Werdell, P. J., Westberry, T. K., Wilding, J. G., and Yoder, J. A.: Regional to global assessments of decadal scale phytoplankton dynamics from the SeaWiFS mission, Remote Sens. Environ., in press, 2013.

Steinacher, M., Joos, F., Frölicher, T. L., Bopp, L., Cadule, P., Cocco, V., Doney, S. C., Gehlen, M., Lindsay, K., Moore, J. K., Schneider, B., and Segschneider, J.: Projected 21st century decrease in marine productivity: a multi-model analysis, Biogeosciences, 7, 979–1005, doi:10.5194/bg-7-979-2010, 2010.

Thornton, P. E., Doney, S. C., Lindsay, K., Moore, J. K., Mahowald, N., Randerson, J. T., Fung, I., Lamarque, J.-F., Feddema, J. J., and Lee, Y.-H.: Carbon-nitrogen interactions regulate climate-carbon cycle feedbacks: results from an atmosphere-ocean general circulation model, Biogeosciences, 6, 2099–2120, doi:10.5194/bg-6-2099-2009, 2009.

Tiao, G. C., Reinsel, G. C., Xu, D. M., Pedrick, J. H., Zhu, X. D., Miller, A. J., Deluisi, J. J., Mateer, C. L., and Wuebbles, D. J.: Effects of autocorrelation and temporal sampling schemes on estimates of trend and spatial correlation, J. Geophys. Res.-Atmos., 95, 20507–20517, doi:10.1029/JD095iD12p20507, 1990.

Vantrepotte, V. and Mélin, F.: Temporal variability of 10-year global SeaWiFS time-series of phytoplankton chlorophyll a concentration, ICES J. Mar. Sci., 66, 1547–1556, doi:10.1093/icesjms/fsp107, 2009.

Weatherhead, E. C., Reinsel, G. C., Tiao, G. C., Meng, X. L., Choi, D. S., Cheang, W. K., Keller, T., DeLuisi, J., Wuebbles, D. J., Kerr, J. B., Miller, A. J., Oltmans, S. J., and Frederick, J. E.: Factors affecting the detection of trends: Statistical considerations and applications to environmental data, J. Geophys. Res.-Atmos., 103, 17149–17161, doi:10.1029/98jd00995, 1998.

Wei, W. W. S.: Time series analysis – univariate and multivariate methods, Addison-Wesley, New York, NY, 1990.

Wunsch, C.: The interpretation of short climate records, with comments on the North Atlantic and Southern Oscillations, Bulletin of the American Meteorological Society, 80, 245–255, doi:10.1175/1520-0477(1999)080¡0245:tioscr¿2.0.co;2, 1999.

Yoder, J. A. and Kennelly, M. A.: Seasonal and ENSO variability in global ocean phytoplankton chlorophyll derived from 4 years of SeaWiFS measurements, Global Biogeochem. Cy., 17, 1112, doi:10.1029/2002gb001942, 2003.

Yoder, J. A., Kennelly, M. A., Doney, S. C., and Lima, I. D.: Are trends in SeaWiFS chlorophyll time-series unusual relative to historic variability, Acta Oceanol. Sin., 29, 1–4, doi:10.1007/s13131-010-0016-o, 2010.