Assessing the fitness-for-purpose of satellite multi-mission ocean color climate data records: A protocol applied to OC-CCI chlorophyll- a data

F. Mélin, Vincent Vantrepotte, A. Chuprin, M. Grant, T. Jackson, S. Sathyendranath

Remote Sensing of Environment 203 (2017) 139–151

Contents lists available at ScienceDirect

Remote Sensing of Environment

journal homepage: www.elsevier.com/locate/rse

Assessing the fitness-for-purpose of satellite multi-mission ocean color climate data records: A protocol applied to OC-CCI chlorophyll-*a* data

F. Mélin[a,*], V. Vantrepotte[b], A. Chuprin[c], M. Grant[c], T. Jackson[c], S. Sathyendranath[d]

[a] European Commission, Joint Research Centre (JRC), TP270, via Fermi 2749, Ispra 21027, Italy
[b] INSU-CNRS, UMR 8187, Laboratoire d'Océanologie et des Géosciences, Université Lille Nord de France, ULCO, France
[c] Plymouth Marine Laboratory (PML), Prospect Place, The Hoe, Plymouth PL1 3DH, UK
[d] National Centre for Earth Observation, Plymouth Marine Laboratory (PML), Prospect Place, The Hoe, Plymouth, PL1 3DH, UK

ARTICLE INFO

Article history:
Received 3 October 2016
Received in revised form 17 March 2017
Accepted 29 March 2017
Available online 12 April 2017

Keywords:
Climate
Chlorophyll-a
Ocean color
CCI

ABSTRACT

In this work, trend estimates are used as indicators to compare the multi-annual variability of different satellite chlorophyll-*a* (Chl*a*) data and to assess the fitness-for-purpose of multi-mission Chl*a* products as climate data records (CDR). Under the assumption that single-mission products are free from spurious temporal artifacts and can be used as benchmark time series, multi-mission CDRs should reproduce the main trend patterns observed by single-mission series when computed over their respective periods. This study introduces and applies quantitative metrics to compare trend distributions from different data records. First, contingency matrices compare the trend diagnostics associated with two satellite products when expressed in binary categories such as existence, significance and signs of trends. Contingency matrices can be further summarized by metrics such as Cohen's κ index that rates the overall agreement between the two distributions of diagnostics. A more quantitative measure of the discrepancies between trends is provided by the distributions of differences between trend slopes. Thirdly, maps of the level of significance P of a t-test quantifying the degree to which two trend estimates differ provide a statistical, spatially-resolved, evaluation. The proposed methodology is applied to the multi-mission Ocean Colour-Climate Change Initiative (OC-CCI) Chl*a* data. The agreement between trend distributions associated with OC-CCI data and single-mission products usually appears as good as when single-mission products are compared. As the period of analysis is extended beyond 2012 to 2015, the level of agreement tends to be degraded, which might be at least partly due to the aging of the MODIS sensor on-board Aqua. On the other hand, the trends displayed by the OC-CCI series over the short period 2012–2015 are very consistent with those observed with VIIRS. These results overall suggest that the OC-CCI Chl*a* data can be used for multi-annual time series analysis (including trend detection), but with some caution required if recent years are included, particularly in the central tropical Pacific. The study also recalls the challenges associated with creating a multi-mission ocean color data record suitable for climate research.

1. Introduction

In a context of climate change and anthropogenic pressures of various types bearing on ocean ecosystems (Roberts, 2003; Lotze et al., 2006; Galloway et al., 2008; Fabry et al., 2008; Stewart et al., 2010), marine phytoplankton are likely to be affected (e.g. Sarmiento et al., 2004; Steinacher et al., 2010; Boyce et al., 2014) and their evolution needs to be monitored and understood. Phytoplankton being a keystone of marine ecosystems, chlorophyll-a concentration (Chl*a*) is listed as an Essential Climate Variable (ECV, Bojinski et al., 2014) by the Global Climate Observing System (GCOS, 2011). Ocean color remote sensing allows us to follow the variations of the marine near-surface phytoplankton at the scale of global oceans. But long time series (i.e., of the order of multiple decades) are required to separate the signature of climate change from the background of natural seasonal and inter-annual variability (Henson et al., 2010; Yoder et al., 2010). Considering the typical lifetime of satellite platforms, this implies the combination of successive missions into one data record suitable to address climate time scales, in effect operating a shift from a mission-centric approach to a variable-centric one.

Without underestimating the relevance of precursor missions (Hovis et al., 1980; Zimmermann and Neumann, 1997; Kawamura et al., 1998), it seems fair to say that the Sea-viewing Wide Field-of-view Sensor (SeaWiFS, McClain et al., 1998) heralded the beginning of a systematic global monitoring of the ocean (McClain et al., 2004a):

* Corresponding author.

http://dx.doi.org/10.1016/j.rse.2017.03.039

it was followed by other global-scale missions, and currently new platforms are being launched or designed to ensure continuity, such as the Joint Polar Satellite System (JPSS) from the National Oceanic and Atmospheric Administration (NOAA), which started in 2011, the Sentinel-3 program from the European Space Agency (ESA, Donlon et al., 2012) with its Ocean and Land Colour Imager (OLCI), the Global Change Observation Mission-Climate (GCOM-C) from the Japan Aerospace Exploration Agency (JAXA), and the Plankton, Aerosol, Cloud, ocean Ecosystem (PACE) mission from the National Aeronautics and Space Administration (NASA). Following the availability of a multi-mission data record, several investigations relied on multiple missions to study time series of ocean color-derived variables going beyond the SeaWiFS period (Mélin et al., 2009, 2011; Kahru et al., 2012; Bélanger et al., 2013; Coppini et al., 2013; Saulquin et al., 2013; Gregg and Rousseaux, 2014; Park et al., 2015; Signorini et al., 2015; Soppa et al., 2016). These studies showed an uneven level of attention to the issue of inter-mission differences that appear to be variable in time and space as well as being significant when creating climate data records (Djavidnia et al., 2010; Mélin, 2010, 2011; Mélin et al., 2016a). These differences may propagate artifacts into a combined multi-mission data set, potentially affecting the conclusions that can be drawn from a time series analysis, for instance regarding trends (Gregg and Casey, 2010; Beaulieu et al., 2013). A specific sensitivity study quantified how close mission-specific Chl*a* data from two satellite missions need to be to allow a trend analysis with a combined data set (Mélin, 2016). The study concluded that inter-mission biases of the order of 5%, if not corrected, generally lead to a trend (represented by the slope of linear regression) significantly different from the trend obtained from a reference series unaffected by these biases (this reference series was built by operating a simple climatological bias correction before averaging the data from the two missions). Moreover, the requirement becomes more stringent with low Chl*a* values, which implies that analyses conducted in oligotrophic waters are extremely sensitive to the effects of inter-mission differences.

Various initiatives have been carried out to create ocean color time series spanning successive missions for different applications (e.g. Kwiatkowska and Fargion, 2003; Maritorena et al., 2010; Pottier et al., 2006; Mélin and Zibordi, 2007; Mélin et al., 2011; IOCCG, 2007). Considering the background introduced above, constructing multi-mission data records of Chl*a* suitable for climate research is very challenging. The availability of multi-mission data sets thus begs the question of their fitness-for-purpose as climate data records (CDR), that is to say, their capacity to fulfill the purpose of a time series analysis suitable to reach robust climate-relevant conclusions.

This paper introduces a methodology to assess the fitness-for-purpose of multi-mission ocean color CDRs by comparing trend estimates obtained with a multi-mission time series and with single-mission products, the rationale being that trend distributions associated with a multi-mission data set should be consistent with those of all successive single-mission series when computed over a common time period. If this be the case, it might be considered safe to conduct climate research on the time scales allowed by the multi-mission time series. To the contrary, the occurrence of trends from a multi-mission series that significantly differ from those of a single mission indicates that undesirable artifacts have been introduced in the construction of the combined data set and that its status as a CDR is questionable. This reasoning relies on the assumption that single-mission products are themselves CDRs and can be treated as time series of reference free from temporal artifacts such as drifts from insufficiently characterized calibration. It also implies that the major challenge to achieving the status of CDR for a multi-mission data set is represented by how single-mission products are combined to create the CDR (Mélin, 2016). There is a solid basis for using single-mission products as reliable series of reference, supported by the efforts conducted by space agencies on the calibration of instruments and their stability in time (Xiong et al., 2010; Eplee et al., 2012; Cao et al., 2013; Eplee et al., 2015; IOCCG, 2013), even though some caution should still be exercised. This is a key point because the construction of multi-mission CDRs would be strongly jeopardized if the temporal variations of single-mission products could not be trusted. It is also worth mentioning that there is a lack of alternatives, since time series of in situ data collected with the same protocol and at the same location and suitable for trend analysis are likely to remain a rarity.

Even if the methodology can be applied on any data and region, it is illustrated here using the example of the Chl*a* global data set developed by the ESA Ocean Colour (OC) component of the Climate Change Initiative (CCI, Hollmann et al., 2013) that specifically aimed at producing climate-quality data records from ocean color data. The data are first described and briefly compared with single-mission products. The core of the paper is then a description of the methodology and its application to the OC-CCI data set, whose distribution of trend estimates is assessed using three complementary analyses.

2. Satellite data and methods

2.1. Single-mission products

Five single-mission Chl*a* products were obtained from the NASA distribution system in association with the missions SeaWiFS, the ESA MEdium Resolution Imaging Spectrometer (MERIS) onboard the Envisat platform (Rast et al., 1999), the Moderate Resolution Imaging Spectroradiometer (Esaias et al., 1998) onboard the Aqua and Terra platforms (MODIS-A and MODIS-T, respectively), and the Visible/Infrared Imager Radiometer Suite (VIIRS, Schueler et al., 2002) onboard the Suomi National Polar-orbiting Partnership spacecraft. All data sets are Level-3 gridded monthly Chl*a* products with a spatial resolution of a 12th-degree. The five missions were treated with a common strategy for calibration and processed with the same atmospheric correction (Franz et al., 2007). The product version corresponds to NASA reprocessing R2014 (R2014.0 for SeaWiFS and MODIS-T, R2014.0.1 for MODIS-A, and R2014.0.2 for VIIRS; NASA, 2015) except for MERIS (R2012.1). One change from R2012 to R2014 was the adoption of the Ocean Color Index (OCI) algorithm for Chl*a* computation in oligotrophic waters (Hu et al., 2012), while the OC4 algorithm (O'Reilly et al., 2000) was retained for other conditions.

2.2. Multi-mission product

The multi-mission Chl*a* monthly time series was obtained from the OC-CCI version 3 data set (CCI, 2016a) that was the result of merging data from SeaWiFS, MERIS, MODIS-A, and VIIRS. The data from SeaWiFS and VIIRS were obtained from NASA as Level-2 files (associated with reprocessing R20104.0 and R2014.0.1, respectively). The OC-CCI processing included both Global and Local Area Coverage (LAC and GAC) SeaWiFS imagery in its data stream. While restricted by the operation of limited on-board recording capacity and of download by ground stations, LAC data have a full 1-km resolution, therefore bringing a much richer information content (particularly around continental masses) than the sub-sampled GAC imagery (Mélin et al., 2002). MODIS-Aqua data were obtained as Level-1A (un-calibrated) data and then processed to attain the same calibration as for NASA's R2014.0, while MERIS data were from the 3rd ESA reprocessing (also input to NASA reprocessing R2012). Then the atmospheric correction for MODIS-A and MERIS was performed through the POLYMER algorithm (Steinmetz et al., 2011). Additionally, cloud flagging for MERIS and SeaWiFS adopted an advanced scheme (CCI, 2016b).

In the framework of the OC-CCI data stream, merging was carried out at the level of the remote sensing reflectance R_{RS} that were first expressed on the wavebands of SeaWiFS through a band-shifting

scheme (Mélin and Sclep, 2015). A bias correction was then operated using SeaWiFS as a data set of reference: a spatially-resolved bias climatology was computed for each waveband using the period of overlap between SeaWiFS on one hand, and MERIS and MODIS-A on the other hand (2003–2007) (CCI, 2016c), so that the bias could be removed on any grid point and day. The same operation was done between VIIRS R_{RS} and bias-corrected MODIS-A data (using the period of overlap 2012–2013). This simple bias correction accounts for the average spatial variability of the bias distribution as well as its average annual cycle, and places the various R_{RS} data on the SeaWiFS baseline. Merging was then operated by simply averaging band-shifted bias-corrected R_{RS} data.

From the R_{RS} distributions, Chl*a* was computed using the same approach as NASA for the reprocessing R2014, that is, combining the OC4 and OCI algorithms. Additionally, the OC5 algorithm (Gohin et al., 2002) was adopted in coastal regions, with the blending of the open ocean and coastal conditions performed using a framework of optical classes, whereby the algorithm outputs were weighted by the relevant class memberships (Mélin et al., 2011; Moore et al., 2014).

Overall, the OC-CCI data stream shows commonalities with single-mission products, but there are a few major differences, such as the use of a different atmospheric correction for MODIS-A and MERIS, a different cloud masking for MERIS and SeaWiFS, the adoption of another Chl*a* algorithm for coastal conditions and the inclusion of LAC SeaWiFS data. These elements were then combined in a complex way as the data went through the band-shifting and bias correction procedures.

2.3. Trend estimates and comparison of trends

The methods employed for trend analysis were used in previous works (Vantrepotte and Mélin, 2009, 2011) and are briefly summarized here. For each grid point, the Chl*a* monthly series first went through a pre-processing step. If a month was associated with a missing value in more than 50% of the cases (i.e., the number of years), then all values for that month were excluded, in practice creating an annual cycle of varying length ($\leq$12 months). If more than 30% of the remaining data were missing, then the whole series was excluded; otherwise missing values were filled in by an eigenvectors filtering method (Ibañez and Conversi, 2002).

The calculation of the linear trend (expressed in % year^{-1}) associated with the pre-processed series was performed after the removal of its annual cycle (de-seasonalized series) (Vantrepotte and Mélin, 2009). The standard error of the associated slope β is a function of the unexplained sum of squares of the linear regression and is written as (Sokal and Rohlf, 1994):

$$s_\beta = \sqrt{\frac{1}{(N-2)} \frac{\sum_{i=1}^{N} (x_i - (\alpha + \beta t_i))^2}{\sum_{i=1}^{N} (t_i - \bar{t})^2}} = \sqrt{\frac{1}{(N-2)} \left(\frac{\sigma(x)^2}{\sigma(t)^2} - \beta^2 \right)} \quad (1)$$

where t and x represent the N-element series associated with time (of average $\bar{t}$) and geophysical data, respectively, and α is the intercept of the regression, while σ indicates the standard deviation operator. The level of significance p of the trend was computed with a t-test where $t = \beta / s_\beta$ and with N-2 degrees of freedom.

Trend estimates obtained by various time series can be compared with a statistical test (Mélin, 2016). Considering the slopes of linear regression β_1 and β_2 at a given location (and their standard errors $s_{\beta,1}$ and $s_{\beta,2}$) associated with two satellite products, a t-test was used to establish if β_1 and β_2 could be considered equal with a level of significance P: when P was small, this hypothesis could be rejected and the slopes were considered different (for clarity, the notation P was preferred when comparing slopes, while p was used to quantify the significance of a single trend). The t-test relied on the following value:

$$t = \frac{\beta_1 - \beta_2}{\sqrt{s_{\beta,1}^2 + s_{\beta,2}^2}} \quad (2)$$

with a degree of freedom df equal to:

$$df = \frac{\left(s_{\beta,1}^2 + s_{\beta,2}^2 \right)^2}{\frac{1}{N_1} s_{\beta,1}^4 + \frac{1}{N_2} s_{\beta,2}^4} \quad (3)$$

following Andrade and Estévez-pérez (2014, their equation 8). The numbers of months N_1 and N_2 varied according to the coverage given by each satellite product.

Other metrics were used to compare trend estimates; for the sake of clarity they are introduced as results are presented in Section 4.

3. Comparison of products

Before addressing the trend estimates, the various products introduced above were compared. First, it is worth stressing that, notwithstanding significant differences, the various Chl*a* products are generally consistent in terms of global distribution (e.g. Djavidnia et al., 2010) and this is also true for the OC-CCI products (Couto et al., 2016). Even though the scope of this study was not to document differences between products, providing general results in this respect helped illustrating different periods in the data sets and some of their characteristics. The unbiased relative difference ψ served to quantify the average bias (relative difference) between two series $(x_{1,i})_{i=1,N}$ and $(x_{2,i})_{i=1,N}$:

$$\psi = \frac{2}{N} \sum_{i=1}^{N} \frac{(x_{2,i} - x_{1,i})}{(x_{2,i} + x_{1,i})} \quad (4)$$

In the definition of ψ, the reference for the relative difference (denominator) is the average of the two products, which avoids arbitrarily selecting one product as reference, while numerically preventing cases where only the denominator is close to zero.

Fig. 1 shows the globally-averaged relative difference ψ for a certain number of sensor pairs, while Fig. 2 illustrates differences of OC-CCI Chl*a* with respect to single-mission products. In general the differences between single-mission products ψ are lower (in modulus) than 10%. The value of ψ between SeaWiFS and MODIS-A is the smallest, oscillating mostly between $\pm$5%; this is in line with their recent common reprocessing and with the fact that these missions have gone through extensive characterizations. Differences are higher when MERIS Chl*a* are compared with SeaWiFS or MODIS-A values (with MERIS Chl*a* lower), which can be partly explained by a different reprocessing version. The manifestation of deactivating the MERIS Offset Control Loop (OCL, operating a correction for the instrument dark current) from December 2004 to October 2006 is also seen (Fig. 1). This underlines the sensitivity of ocean color records to small changes in space-sensor calibration functions. During the first part of their overlap period (2003–2010), ψ comparing MODIS-A and MODIS-T is mostly in the interval $\pm$5%, but displays larger oscillations in the following years. It is important to mention that the MODIS-T mission has met difficulties in characterizing the instrument and its calibration history (Franz et al., 2008) . In fact, the calibration of MODIS-T is constrained using other ocean color missions (Kwiatkowska et al., 2008), so that its data record does not constitute a fully independent time series. Over the period 2012–2015, the difference between MODIS-A and VIIRS averages 4.4% (with VIIRS Chl*a* lower). A factor potentially explaining larger oscillations

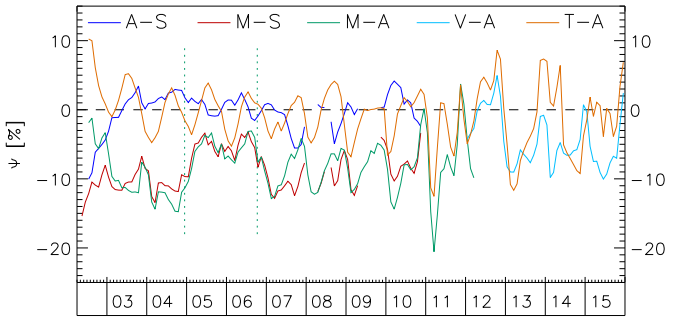

Fig. 1. Time series of globally-averaged relative differences between Chl*a* single-mission products. The letters S, M, A, T, and V are associated with SeaWiFS, MERIS, MODIS-A, MODIS-T, and VIIRS, respectively. The dotted vertical green lines show the period when the MERIS OCL was deactivated (see text).

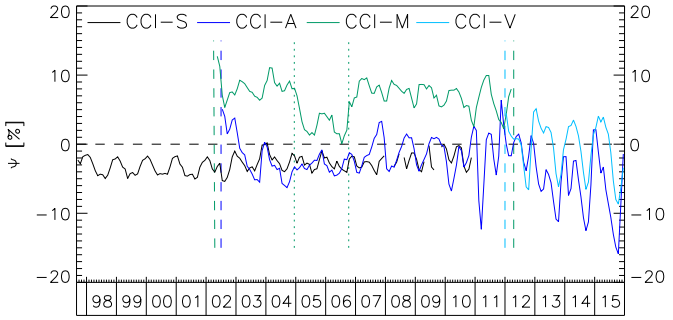

Fig. 2. Time series of globally-averaged relative differences between OC-CCI Chl*a* and single-mission products (with letters as in Fig. 1). The dashed vertical green lines show the start and end date of the MERIS data stream, the dashed blue and light-blue lines show the start of the MODIS-A and VIIRS data, respectively. The dotted vertical green lines show the period when the MERIS OCL was deactivated (see text).

in ψ in the last years is the radiometric degradation of MODIS-A as the sensor is aging (Meister et al., 2012; Meister and Franz, 2014).

When conducting the same exercise with OC-CCI data (Fig. 2), several observations can be made. Differences between SeaWiFS and OC-CCI Chl*a* are the lowest (in modulus) with small oscillations. This agreement is favored by two factors (see Section 2.2): first, by construction the OC-CCI series should mostly reproduce the SeaWiFS record over the years when only SeaWiFS data are available (up to 2002); then, for subsequent years, the bias correction scheme should bring the products from the other missions largely in line with the SeaWiFS record. In that respect, the bias correction appears to fulfill its task since the temporal stability of the ψ series is remarkable even when MERIS and MODIS-A data are added in the data stream (see also Couto et al., 2016) . This being said, a residual effect can still be noticed, the average ψ changing from −3.4 % over the period 1998–2001 (when only SeaWiFS was available) to −2.6 % over the period 2003–2007.

Between 2002 and 2009, ψ between OC-CCI and MODIS-A Chl*a* (Fig. 2) is generally below 5% (in modulus) and negative (with exceptions in 2002 and from 2007), but larger oscillations (with ψ increasingly negative) are noticed in the later part of the record, which again might be partly explained by calibration issues. On average OC-CCI Chl*a* appears higher than MERIS values, which is coherent with the comparison between SeaWiFS and MERIS data (Fig. 1); the signature of the OCL deactivation period is also clearly seen. The difference between OC-CCI and VIIRS Chl*a* shows seasonal variations between +4% and −6 %.

Of course these results would show significant regional variations (e.g. Djavidnia et al., 2010; Mélin, 2010), the analysis of which is beyond the scope of this study and would deserve additional dedicated studies. The existence of regional variations implies that the methods developed to assess the trends associated with the multi-mission data set need to accommodate a spatial dimension. However, from the global average of ψ, some general characteristics of the time series can be observed that are relevant for the rest of the study. First, if considering SeaWiFS, MERIS and MODIS-A data, the ψ series display a relative temporal stability, notwithstanding some oscillations (and the OCL issue for MERIS). This further supports the use of single-mission products as series of reference. But this might not be fully applicable to recent years considering the issue of aging associated with MODIS-A. Such a stability of the differences with respect to OC-CCI data seems encouraging (in line with results obtained with OC-CCI version 1 data, Couto et al., 2016) but the comparison with MODIS-A suggests that, again, including data of the last years (say, after 2012) could lead to spurious results.

4. Assessment of trends in multi-mission data records

After having provided background information on the various time series and their relative agreement (at least at global scale), the study is now focusing on the assessment of trend distributions. Trend results are mostly based on the period up to 2012, generally using SeaWiFS, MERIS and MODIS-A before manifestation of its aging (the case of more recent years is addressed subsequently). After a description of observed trends, the methodology of trend assessment follows three lines of analysis. First, the use of contingency matrices is introduced to confront the main diagnostics of trend estimates. Then trend slopes from different products are compared on a point-by-point basis. Finally, trend spatial distributions are compared by statistical tests.

4.1. Observed trends

The trends obtained for SeaWiFS over the period 1998–2007 are presented in Fig. 3a. The analysis is limited to the ten years when SeaWiFS functioned in an optimal way without data gaps, and only statistically significant trends are shown ($p < 0.05$). Significant negative trends (that can reach $-5\%\,\text{year}^{-1}$) are observed generally in oligotrophic subtropical gyres (except in the south Atlantic), as well as in the northeast Atlantic, while significant positive trends are noticed in various regions such as the southeast Pacific and the Tasman Sea (Vantrepotte and Mélin, 2009, 2011; Vantrepotte et al., 2011). The negative trends in Chl*a* in the subtropical gyres have attracted substantial attention (McClain et al., 2004b; Polovina et al., 2008; Polovina and Woodsworth, 2011; Signorini and McClain, 2012), as they appeared consistent with hypotheses of a more stratified and warming ocean (Doney, 2006).

The trend distribution obtained over 10 years (August 2002–July 2012) using MODIS-A shows that several patterns have been reversed (Fig. 3b). This is particularly the case in the subtropical Pacific, with opposite patterns in the South Pacific, along the Equator in the western Warm Pool region, or in the northwest quarter. The negative trends observed in the northeast Atlantic with SeaWiFS are also replaced by a positive signal (with an extension northward to Iceland), while significant positive trends have extended in the northern Atlantic subtropics with respect to SeaWiFS. In the Indian Ocean, the trends observed with SeaWiFS in the western part of the basin and in the southern subtropical gyre (positive and negative, respectively) are also reversed. The clear pattern of positive trend found in the Tasman Sea is no longer significant over the following period.

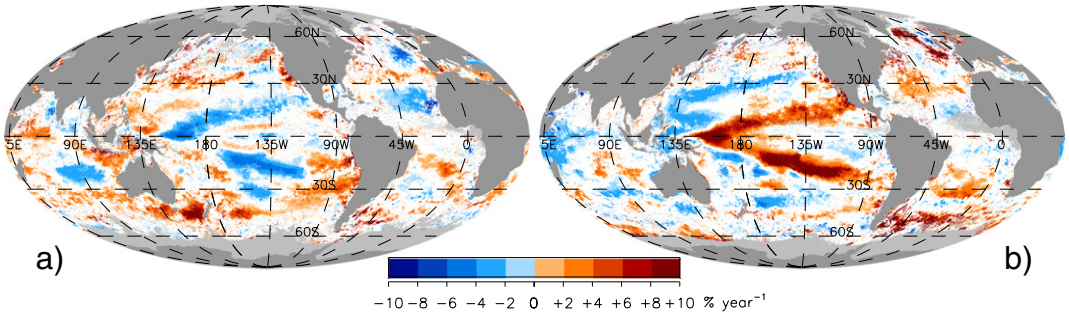

Fig. 3. Trends for a) SeaWiFS Chl*a* over the period Jan. 1998–Dec. 2007, and b) MODIS-A Chl*a* over the period Aug. 2002–Jul. 2012. Only significant trends ($p < 0.05$) are represented. Grey shows land, while light-grey is associated with areas where the series are insufficient for analysis.

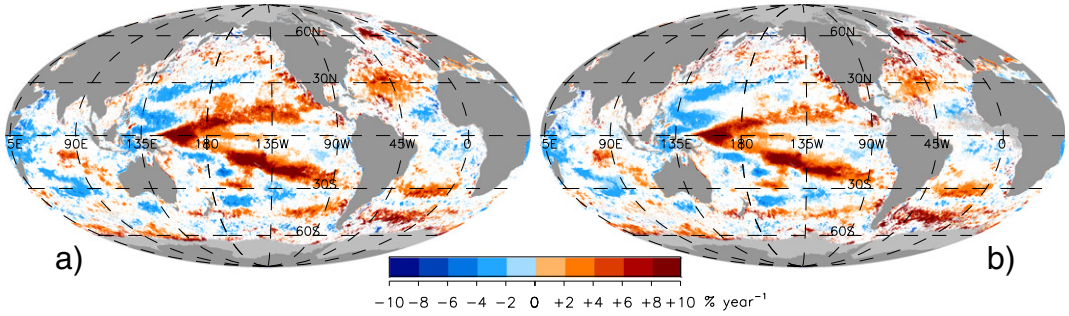

Fig. 4. Trends for a) MERIS Chl*a*, and b) MODIS-A Chl*a*, over the period Aug. 2002–Jul. 2011. Only significant trends ($p < 0.05$) are represented. Grey shows land, while light-grey is associated with areas where the series are insufficient for analysis.

In Fig. 4a, the trends associated with MERIS over 9 years (August 2002–July 2011) show a distribution very close to the results obtained with MODIS-A over the same period (Fig. 4b). The distributions associated with MODIS-A over the two slightly different periods (a difference of one year, Figs. 3b and 4b) generally agree but some differences can be noticed. For instance, the feature of negative trends in the southern Indian Ocean is clearer over the period 2002–2011. For completeness, it is mentioned that the trend distribution obtained with MODIS-T is also very similar to those shown by MODIS-A or MERIS over 2002–2011 or 2002–2012 (not shown).

A first conclusion based on the comparison between MERIS and MODIS is that trend distributions associated with the single-mission products are remarkably consistent when computed over a common period. A similar conclusion is reached when comparing the single-mission trends with those derived from the OC-CCI data sets, computed over the same time intervals: January 1998–December 2007 in Fig. 7a, to be compared with the SeaWiFS results (Fig. 3a), August 2002–July 2011 in Fig. 8a, to be compared with MERIS or MODIS-A (Fig. 4), and August 2002–July 2012 in Fig. 9a, to be compared with MODIS-A results (Fig. 3b) (the OC-CCI maps will be further commented below, Section 4.4). However, these trend maps only illustrates a qualitative agreement (when determined over a common period). The comparison of two trend distributions requires more quantitative criteria that are presented in the next sections. In that context, the availability of almost a decade of coincident measurements by MODIS and MERIS is a great asset as it provides a comparison of reference, with which results obtained with multi-mission products can be confronted. So the criteria of comparison are first tested on the pair MODIS/MERIS, providing benchmark results, before being applied to the OC-CCI data.

4.2. Contingency matrix

Beyond the visual agreement illustrated in the previous section, trend distributions computed from two products over a common period should provide the same kind of diagnostic over general questions such as: “is Chl*a* increasing or decreasing?”. Contingency matrices (also termed multi-way tables, Sokal and Rohlf, 1994) are used in statistics to summarize the relationship between several diagnostic variables. Here, they can be used for the purpose of comparing the output of a trend analysis expressed in simple terms of increase/decrease. First, each grid point of the global map is put into one of four categories: negative or positive trend slopes for both products, negative slope for one product and positive for the other, and vice-versa. Then the contingency matrix summarizes such a comparison of outcomes by summing the surface of the domain

Table 1
Contingency matrices comparing trend analysis outcomes. The comparison MODIS-A/MERIS is made over the period Aug. 2002–Jul. 2011, MODIS-A/MODIS-T over Aug. 2002–Jul. 2012. Percentage values quantify the amount of the ocean where the diagnostics on the sign of trend slopes β apply.

%	MERIS		MODIS-T	
MODIS-A	$\beta \geq 0$	$\beta < 0$	$\beta \geq 0$	$\beta < 0$
$\beta \geq 0$	48.7	8.4	44.6	11.2
$\beta < 0$	9.4	33.5	4.6	39.5

Table 2
Contingency matrices as in Table 1, but where diagnostics consider the significance of the trends. 'n.s.' stands for non-significant; '*' indicates significant trends ($p < 0.05$).

%	MERIS			MODIS-T		
MODIS-A	n.s.	$\beta^* \geq 0$	$\beta^* < 0$	n.s.	$\beta^* \geq 0$	$\beta^* < 0$
n.s.	58.0	6.1	3.4	57.5	2.4	6.2
$\beta^* \geq 0$	5.4	15.0	0.0	5.9	14.7	0.0
$\beta^* < 0$	4.3	0.0	7.7	2.0	0.0	11.2

(expressed in %) associated with these four categories. In the case of the comparison between MODIS-A and MERIS trends (Aug. 2002–Jul. 2011), the slope of linear regression β is of the same sign over 82.1% of the domain (sum of diagonal elements in Table 1) whereas it differs over 17.8% of the domain. Comparing MODIS-A and MODIS-T (Aug. 2002–Jul. 2012), the agreement in sign is fulfilled over 84.1% of the ocean.

A more meaningful comparison takes into account the statistical significance of the trends, as presented in Table 2 with the three following categories: non-significant trends, significant positive or negative slopes. For the comparison MODIS-A/MERIS, the trend diagnostics agree for 80.7% of the domain, with 58.0% associated with non-significant trends, and 22.7% with significant trends of the same sign. To the contrary, 19.3% of the domain is characterized by a contrasted diagnostic but it is worth underlining that the worst case, that is, significant trends for both products with opposite signs, almost never occurs (0.017%). Similar results are observed when comparing both MODIS data sets (Table 2).

For the sake of comparison, a similar analysis was conducted to summarize the differences in trends observed between SeaWiFS and MODIS-A over different periods (1998–2007 and 2002–2012, respectively). In that case, 56.9% of the domain shows trends of opposite signs. Considering the level of significance, significant trends of the same sign are observed over 2.8% of the ocean only, whereas 10.4% of the ocean is characterized by significant trends of opposite signs (a very rare occurrence when series are compared over the same period). These results confirm that Chl*a* trends have indeed substantially changed from the period 1998–2007 to 2002–2012 as was anticipated in Section 4.1 from a qualitative analysis of the trend maps (Fig. 3).

A common way to synthesize a square contingency table is to introduce the Cohen's κ index that quantifies the magnitude of the agreement between two raters (Cohen, 1960; Viera and Garrett, 2005; Warrens, 2011). Considering a contingency matrix $(\rho_{i,j})_{i,j=1,n}$ made of the proportions of n diagnostics given by two missions:

$$\kappa = \frac{\rho - \rho_c}{1 - \rho_c} \tag{5}$$

where ρ is the proportion of observed agreement (i.e., common diagnostic from two raters) written as the sum of diagonal terms $\sum_{i=1}^{n} \rho_{i,i}$, and ρ_c is the proportion of agreement expected from chance alone, expressed by $\sum_{i=1}^{n} \left(\sum_{j=1}^{n} \rho_{i,j} \sum_{k=1}^{n} \rho_{k,i} \right)$ (Warrens, 2011). For a perfect agreement, κ would reach 1 whereas values close to 0 indicate an agreement not better than chance. Here the number of diagnostics is $n = 3$ (non-significant trend, positive and negative significant trend) and the matrix terms $\rho_{i,j}$ are the percentage values (as in Table 2) divided by 100. The κ index associated with the comparison between MODIS-A and MERIS trends is 0.60, which indicates a moderate-to-substantial agreement (Viera and Garrett, 2005). Comparing MODIS-A and -T (Aug. 2002–Jul. 2012), κ is 0.68 (substantial agreement). On the other hand, the comparison between the SeaWiFS and MODIS-A trends computed over two partly overlapping periods is equal to zero (−0.009), again in line with the clear differences in trends observed over the different periods (Fig. 3).

A similar analysis was conducted with the OC-CCI record, comparing it with the series from SeaWiFS (Jan. 1998–Dec. 2007), MERIS (Aug. 2002–Jul. 2011) and MODIS-A (Aug. 2002–Jul. 2012) over their respective periods. Without considering the level of significance, the slopes obtained with the OC-CCI product agree with the single-mission products for approximately 85% of the domain (85.8%, 83.7% and 84.1%, respectively). When considering the significance of the signals, this agreement is at least 80% (Table 3), the highest being 82.3% when comparing OC-CCI with SeaWiFS (77% for MODIS-T). These results are associated with κ of approximately 0.64. As when comparing trends associated with single-mission products, the discrepancies in analysis outcomes are made of cases with trends significant in one case and not the other, and virtually never happen with significant trends of opposite signs.

As anecdotal result, it is worth noticing that trends are more often positive than negative, regardless of the level of significance, the period considered and the product (Tables 1 to 3).

4.3. Comparison of trend slopes

A more quantitative approach to compare trend estimates is simply to analyze differences in trend slopes. For various pairs of products, the differences in slopes computed for all grid points at global scale are summarized in Fig. 5. Focusing first on the differences between single-mission products, the comparison between MERIS and MODIS-A trend slopes (period Aug. 2002–Jul. 2011) shows a median difference close to null, an inter-quartile distance of 1.6% year^{-1} and the 10th and 90th percentiles of the distribution of approximately 2% year^{-1} in modulus (Fig. 5). Despite the limitations noticed for MODIS-T data, the results are also given for the pairs MODIS-A/MODIS-T (period Aug. 2002–Jul. 2012), with slopes associated with MODIS-T on average lower than those of MODIS-A (median of −0.27% year^{-1}); on the other hand, the inter-quartile distance is lower (1.2% year^{-1}).

When comparing the OC-CCI slopes with those of single-mission products, results are very similar or better, with a median difference from 0.05% year^{-1} (comparison with SeaWiFS over 1998–2007) to 0.27% year^{-1} (comparison with MODIS-A over Aug. 2002–Jul.

Table 3
Contingency matrices as in Table 2 but where OC-CCI trends are compared with SeaWiFS (period Jan. 1998–Dec. 2007), MERIS (Aug. 2002–Jul. 2011) and MODIS-A (Aug. 2002–Jul. 2012).

%	SeaWiFS			MERIS			MODIS-A		
OC-CCI	n.s.	$\beta^* \geq 0$	$\beta^* < 0$	n.s.	$\beta^* \geq 0$	$\beta^* < 0$	n.s.	$\beta^* \geq 0$	$\beta^* < 0$
n.s.	56.1	9.6	2.6	54.4	8.7	4.9	51.7	10.2	4.2
$\beta^* \geq 0$	3.1	16.0	0.0	3.5	17.5	0.0	2.0	18.6	0.0
$\beta^* < 0$	2.3	0.0	10.1	2.3	0.0	8.7	3.1	0.0	10.0

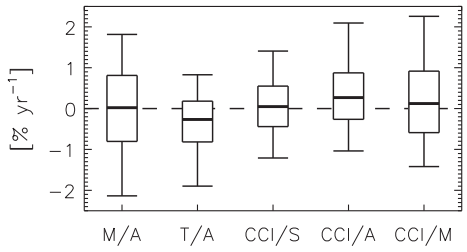

Fig. 5. Statistics of differences between trend slopes obtained over a common period, displayed as box-and-whiskers plots, with boxes indicating the 25th and 75th percentiles, the inner bar associated with the median, and the whiskers being the 10th and 90th percentiles of the distributions. The acronyms S, M, A, T, and CCI are associated with SeaWiFS, MERIS, MODIS-A, MODIS-T, and OC-CCI, respectively.

2012), and inter-quartile distance lower than 1.5% year^{-1}. These results indicate that the distributions of trend slopes obtained from theOC-CCI data are very consistent with those of the single-mission products over several periods. However, such an analysis does not tell us how significant the differences in slope actually are, so the addition of a statistical test in the comparison of trends is added in the next section.

4.4. Statistical comparison

For each pair of time series, the level of significance P of the t -test comparing trend slopes, introduced in Section 2.3 and related to Eqs. (2) and (3), quantifies the degree to which two trend estimates differ. In the description of results, the value of P is considered as a measure of the difference between two slopes of linear trends: the smaller P is, the more the slopes differ. Fig. 6 illustrates the P distribution for the comparison MERIS versus MODIS-A (2002-2011): P values appear low only in scattered and small areas (in the northern subtropical Pacific and Atlantic or in the Baltic Sea), confirming the high degree of coherence between the trend estimates associated with these two missions. P is lower than 0.5 for 38.0% of the ocean, and lower than 0.05 for only 1.9% of the domain of analysis, which means that a very small fraction of the ocean is characterized by significantly different trends when these two products are considered.

Again, the comparison between the two single-mission products is useful to assess the results obtained with OC-CCI Chl*a*, displayed in Figs. 7 to 9, when compared to SeaWiFS (1998–2007), MERIS (2002–2011) and MODIS-A (2002–2012). As already noticed, the trends derived from the OC-CCI series are remarkably consistent with the SeaWiFS record (compare Figs. 3a and 7a), an agreement exemplified by the distribution of P (Fig. 7b) that shows low values only for rather small patterns in the northeast Atlantic, close to the northwest

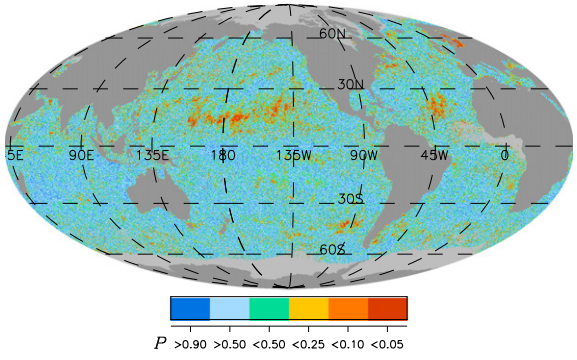

Fig. 6. Level of significance P of the t-test comparing the slopes of linear regression obtained for MERIS and MODIS-A over Aug. 2002 to Jul. 2011. Grey shows land, while light-grey is associated with areas where the series are insufficient for analysis.

African coast and in the tropical southern Atlantic. The latter regions are particularly challenging for atmospheric correction because of cloud cover and dust inputs (Kaufman et al., 2005). Statistically significant differences in trends ($P < 0.05$) are found only in 1.2% of the domain (28.2% are associated with P lower than 0.5). Obviously, the OC-CCI record is very close to SeaWiFS by construction for the initial part of the time series. But the introduction in 2002 of the data records from MERIS and MODIS-A could have resulted in artifacts in the OC-CCI trends because of inter-mission biases (Mélin, 2016; Mélin et al., 2016b); the consistency between the OC-CCI and SeaWiFS trends (actually superior to that obtained when comparing MERIS and MODIS-A) suggests that the OC-CCI processing stream (including its bias correction) performs well.

The trends derived from the OC-CCI series also show a good visual agreement with MERIS (2002–2011, Fig. 8) and MODIS-A (2002–2012, Fig. 9), confirmed by the P distributions: in the comparison with MERIS, P is lower than 0.05 for 1.6% of the domain ($P < 0.5$ for 35% of the ocean), while in the comparison with MODIS-A, this is the case for 2.2% of the domain ($P < 0.5$ for 38% of the ocean), results that are similar to those of the MERIS versus MODIS-A comparison. However some distinct patterns of P can be seen, particularly in the tropical central Pacific south of the Equator (around the meridian 135° W) and along the subtropical front in the Southern Ocean (Figs. 8b and 9b). In these cases, the patterns in trends are actually similar but their amplitudes appear higher in the OC-CCI product. Low P's are also observed in the Baltic Sea when OC-CCI and MODIS-A are compared (but this is also true for the comparison MERIS/MODIS-A, Fig. 6).

4.5. Application with recent years

As explained above, the assessment of the trends derived from the OC-CCI data essentially relied on the SeaWiFS, MERIS and MODIS-A missions, considering the limitations associated with MODIS-T and the short time interval available with VIIRS. Moreover, the use of MODIS-A was restricted to the period 2002–2012, considering the time series of the differences ψ seen in Section 3, and to avoid any artifact associated with uncertain calibration after 2012. The OC-CCI record extending to 2015, it appears worth discussing how the resulting trends are affected, and possibly degraded, as more years are included. This analysis has to rely on the last data years from MODIS (A and T) and the first years of VIIRS, so that extreme caution is due in the interpretation of the results.

Trends obtained from the OC-CCI data were compared with those from MODIS-A over the periods from August 2002 to July of 2011, 2012, 2013, 2014 and 2015. From the contingency matrices, the share of the ocean with differing diagnostics (i.e., trend significant for one product and not for the other, and trends significant for both products but with different signs, see Section 4.2) regularly increases with time, going from 18.4% for the period 2002–2011 to 23.3% for 2002–2015, while κ decreases from 0.65 to 0.55. The part of the ocean with low P values also increases: P is lower than 0.05 for 1.6% of the ocean for the period 2002–2011, increasing to 5.1% for 2002–2015, but the part of the ocean with P lower than 0.5 stays rather constant (in the interval 34.9%–38.2 %). This indicates that the patterns are rather consistent between products but the magnitude of the trends and their related level of significance tend to diverge as more years are added, which is confirmed by an examination of the trend maps (not shown).

The trends obtained by MODIS-A and MODIS-T were compared over the same periods. The contingency matrices indicate that the part of the ocean with differing diagnostics remains constant and fairly low, between 16.1% and 17.4% (κ between 0.62 and 0.68). From the statistical comparison of trends, only 0.7–0.9 % of the domain have P lower than 0.05 (from 29.8% to 32.7% for $P < 0.5$). These results are consistent with the fact that the calibration of MODIS-T is

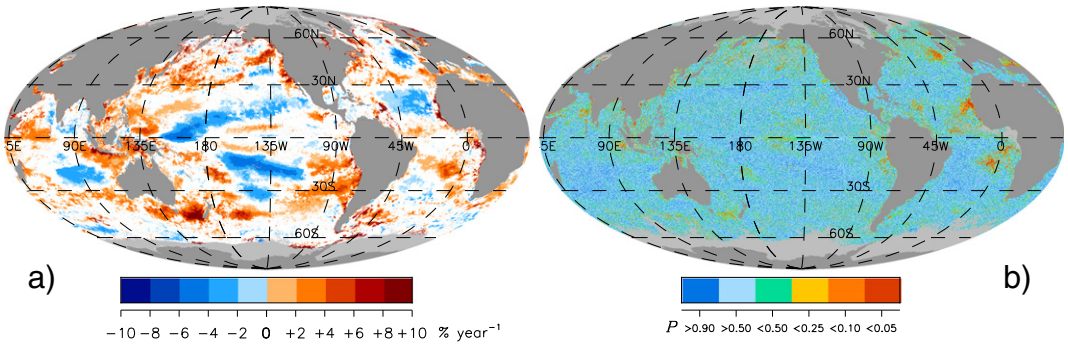

Fig. 7. a) Trends for OC-CCI Chl*a* over the period Jan. 1998–Dec. 2007 (only significant trends, $p < 0.05$, represented), and b) level of significance P of the t-test comparing the slopes of linear regression obtained for SeaWiFS and OC-CCI over that period. Grey shows land, while light-grey is associated with areas where the series are insufficient for analysis.

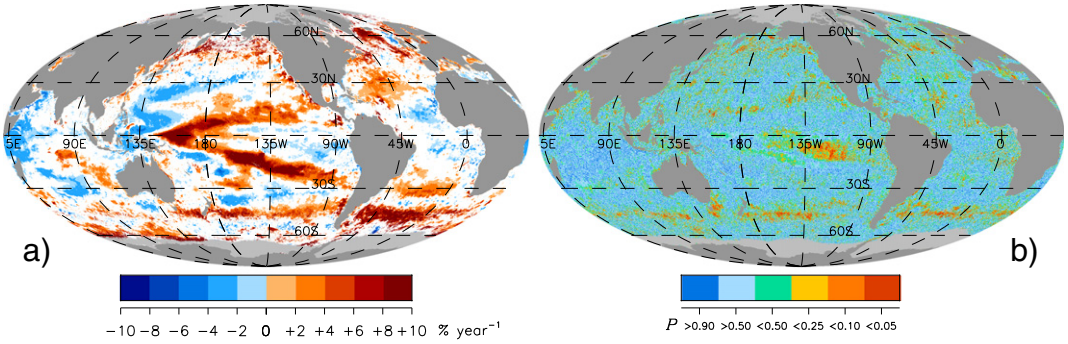

Fig. 8. a) Trends for OC-CCI Chl*a* over the period Aug. 2002–Jul. 2011 (only significant trends, $p < 0.05$, represented), and b) level of significance P of the t-test comparing the slopes of linear regression obtained for MERIS and OC-CCI over that period. Grey shows land, while light-grey is associated with areas where the series are insufficient for analysis.

tied to that of MODIS-A, but they suggest that MODIS-T can be used for the purpose of evaluating OC-CCI data over a longer period.

A similar comparison was then made with OC-CCI and MODIS-T data over the period from March 2000 to February of 2011, 2012, 2013, 2014 and 2015. The results appear degraded as soon as the analysis goes beyond 2011: κ is equal to 0.55 (which is already relatively low) for the series 2000–2011 but between 0.50 and 0.52 if later years are included. From the associated contingency matrix, differing diagnostics are observed over an approximately constant (and fairly high) percentage of the ocean (26–27 %) when the period includes years exceeding 2011, while it is 23% for the interval 2000–2011. The areas with significantly different trends ($P < 0.05$) increase from 2.4% to 4.1% when data from March 2011 to February 2012 are included, and further increases when other years are added, up to 6.5%. Again, these results are mostly due to differing magnitude and significance of the trends while the patterns are similar (see Fig. 10 as an example for the period 2000–2015). Low P values are prominently found in the central Pacific (in a similar area as when comparing OC-CCI and MODIS-A), as well as in the northern subarctic Pacific and Atlantic oceans.

The relative degradation of the results with the addition of the latest years suggests that the comparison with data from VIIRS might be of interest in spite of the short duration of this record. Statistics were therefore derived by comparing the VIIRS series with MODIS-A,

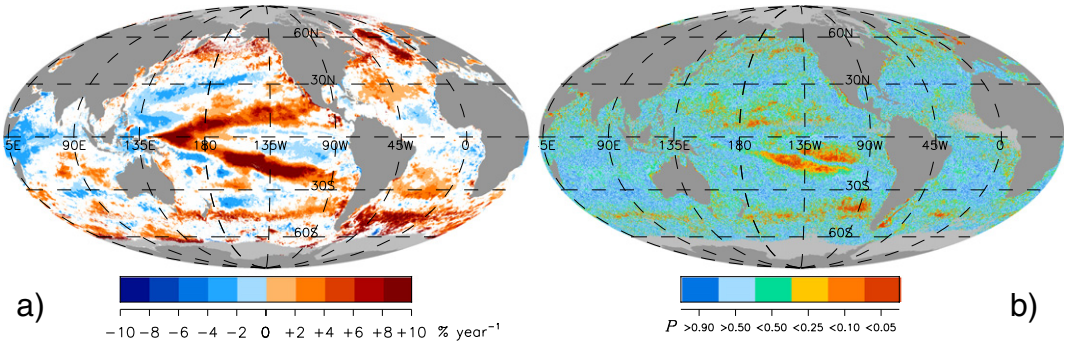

Fig. 9. a) Trends for OC-CCI Chl*a* over the period Aug. 2002–Jul. 2012 (only significant trends, $p < 0.05$, represented), and b) level of significance P of the t-test comparing the slopes of linear regression obtained for MODIS-A and OC-CCI over that period. Grey shows land, while light-grey is associated with areas where the series are insufficient for analysis.

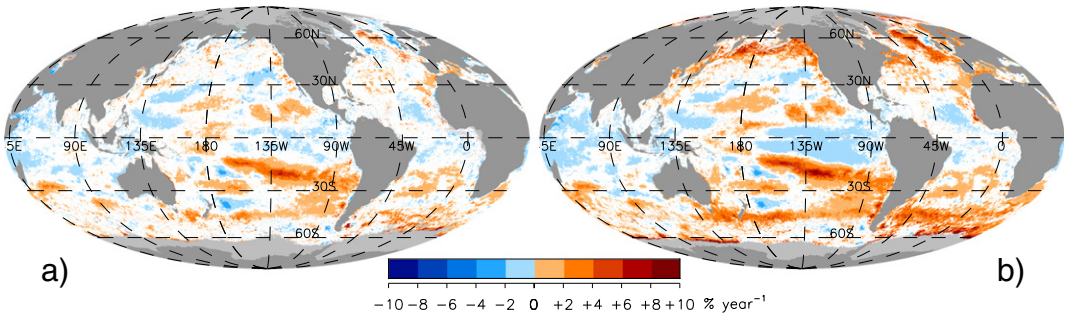

Fig. 10. Trends for a) MODIS-T Chl*a*, and b) OC-CCI Chl*a*, over the period Mar. 2002–Feb. 2015. Only significant trends ($p < 0.05$) are represented. Grey shows land, while light-grey is associated with areas where the series are insufficient for analysis.

MODIS-T and OC-CCI data over the period January 2012 to December 2015 (four years). The part of the ocean with differing trend diagnostics is approximately 18% when comparing MODIS (-A or -T) with VIIRS series (with κ of approximately 0.65), while P is found lower than 0.05 for 1.7–2.0 % of the ocean. When comparing OC-CCI and VIIRS data, these statistics are 14.9% (associated with κ as high as 0.72) and 1.3%, respectively, which means that the trends found from both series are actually very consistent. The median difference between trend slopes is only −0.3% year $^{-1}$, which is very low considering that trend slopes tend to be higher over a short period. It is worth recalling that over the considered time interval, the OC-CCI data are mostly the result of merging VIIRS and MODIS-A data (the MERIS mission ending in April 2012). The agreement is here illustrated with the trend maps associated with both products (Fig. 11). It is underlined that the scale is extended with respect to other maps to accommodate higher values of trend slopes associated with only four years of data. There is a strong signature of a decreasing Chl*a* in the central and eastern tropical Pacific that is reminiscent of El Niño events (e.g., Radenac et al., 2012) and at least partly explained by the event taking place in 2015 (Stramma et al., 2016; Bell et al., 2016). There is also a negative signal along the west coast of North America that could be the result of the warm anomaly affecting the northeast Pacific from 2013 to 2016 coupled with the impact of El Niño (Bond et al., 2015; Cavole et al., 2016; Jacox et al., 2016).

The OC-CCI series now extends over 18 years, which is nearing the period required in tropical regions to detect anthropogenic climate change over the signal of natural variability (Henson et al., 2010). For completeness, Fig. 12 shows the trend map obtained for OC-CCI Chl*a* over October 1997–September 2015. Noticeable are positive trends in the northern Atlantic and Pacific oceans, in the southern Atlantic as well as along a horseshoe pattern that goes from the western equatorial Pacific to mid-latitude South America and then westward across the South Pacific. Negative patterns are noted in the Indian Ocean, in some mid-latitude and equatorial Pacific regions and in the subtropical gyre of the North Atlantic. Considering the elements discussed above about the use of recent data, this trend map is to be taken with appropriate caution and should be confirmed by additional work.

5. Discussion

In this section, various aspects of the proposed methodology as well as its application to the OC-CCI data are discussed.

5.1. Methodological considerations

This study described a protocol to estimate the consistency of satellite Chl*a* series from the point of view of their multi-annual variations, with the objective of assessing the fitness-for-purpose of multi-mission data sets as climate data records (CDR). The trend of linear regression was adopted as a bulk indicator of multi-annual variations that could be used in comparisons between data sets. It is acknowledged that this indicator is a very crude descriptor of ocean Chl*a* multi-annual variability (Vantrepotte and Mélin, 2011) but such trends are indeed expected to appear in ocean Chl*a* series over climate temporal scales (Henson et al., 2010). Moreover this study showed that a comprehensive comparison between products was allowed by results from the trend analysis coupled with fairly simple methods applicable at global scale to satellite data. Particularly, trends of linear regression can easily be compared through statistical tests at the same resolution as ocean color products. It is stressed here that presenting, interpreting and understanding the observed trends were not among the objectives of the present study.

The proposed method was applied to the OC-CCI Chl*a* data set in relation to single-mission products. After having set the stage by

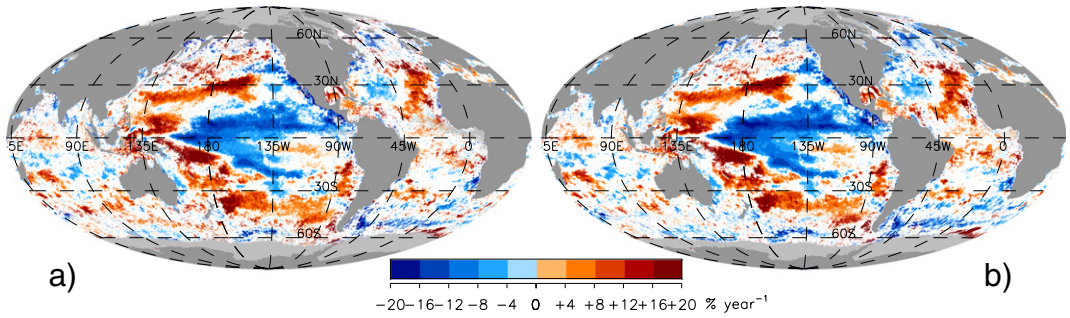

Fig. 11. Trends for a) VIIRS Chl*a*, and b) OC-CCI Chl*a*, over the period Jan. 2012–Dec. 2015. Only significant trends ($p < 0.05$) are represented. Grey shows land, while light-grey is associated with areas where the series are insufficient for analysis.

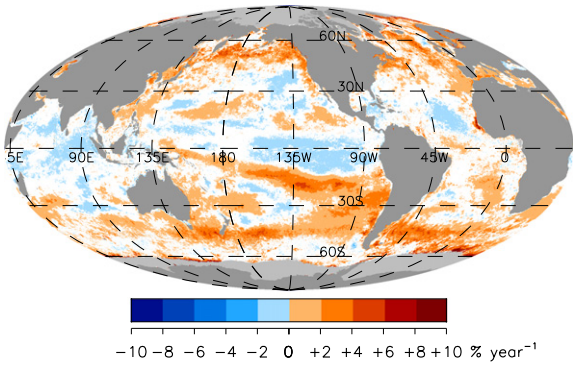

Fig. 12. Trends for OC-CCI Chl*a* over the period Oct. 1997–Sep. 2015. Only significant trends ($p < 0.05$) are represented. Grey shows land, while light-grey is associated with areas where the series are insufficient for analysis.

presenting the time series through their globally averaged differences (Figs. 1 and 2, Section 3), the general approach was based on comparisons between the trends obtained with OC-CCI data and those associated with single-mission series. Considering the activities carried out by space agencies on the calibration of instruments and their stability in time (Xiong et al., 2010; Eplee et al., 2012; Cao et al., 2013; Eplee et al., 2015; IOCCG, 2013), single-mission series were assumed free of spurious temporal artifacts so that multi-mission series should reproduce the trends observed by single-mission products over common periods. In other words, the trends observed from single-mission series were assumed to be sufficiently faithful to reality to serve as benchmarks for multi-mission data (this point is further discussed below). In that context, the use of an independent series (i.e., not included in the multi-mission data set) would be ideal. This was here attempted using the MODIS-T data (not included in the OC-CCI data stream) but as mentioned above (Section 3) these do not constitute a fully independent data record. More generally, the number of ocean color missions is fairly limited for any given period so that excluding one mission from merging is a serious limitation. Moreover, as far as trends are concerned, the multi-mission series should in any case agree with all available single-mission series over the full time interval it covers.

Conversely, if trends shown by the multi-mission data disagreed with results from single-mission products, it would be likely that the construction of the multi-mission data set introduced artifacts in the series that would preclude its use as CDR. It is however worth stressing that a perfect agreement was not to be expected even if the various data sets were perfectly consistent. First, undetected artifacts affecting the single-mission data might still contaminate merged products. Discrepancies could as well originate from differing temporal sampling: multi-mission data sets such as OC-CCI showed a significantly improved coverage with respect to standard single-mission products by relying on inputs from several missions (Mélin et al., 2009; Maritorena et al., 2010; Couto et al., 2016). This phenomenon is reinforced in the case of the OC-CCI data through the use of more robust algorithms allowing more retrievals in challenging atmospheric conditions found in regions such as northwest Africa, the Arabian Sea, the Red Sea (Racault et al., 2015) or the Gulf of Guinea (Nieto and Mélin, 2017), and in fact low P values were noticed offshore Africa when comparing SeaWiFS and OC-CCI trends (Section 4.4, Fig. 7). A richer information content in the Chl*a* series might make trend detection easier with respect to patchier data (Weatherhead et al., 1998). The question of properly including the influence of temporal sampling in the assessment of trends observed by multi-mission products would be worth additional work. In any case, when comparing multi-mission and single-mission data, it is useful to have a baseline for reference. In that context, the ocean color community has the chance to have almost a decade of coincident ocean color data from two well-performing missions (MODIS-A and MERIS). Processed with the same strategy and algorithms, they may represent the best possible agreement to be expected from two products. Additional overlapping data sets are available from SeaWiFS, MODIS-T, and VIIRS.

Different criteria of comparison were presented that covered several characteristics of trend distributions:

- contingency matrices, easily summarized by metrics such Cohen's κ, were useful to reveal how much of the global ocean shared the same diagnostic about the existence, significance and signs of trends;
- the distributions of differences in trend slopes gave a quantitative measure of the discrepancies between trends;
- maps of the level of significance P of a t-test quantifying the degree to which two trend estimates differed provided a statistical, spatially-resolved, evaluation.

With the calculation of a few bulk indicators and the P map, such a framework allows the evaluation and comparison of candidate CDRs (other examples are seen in Mélin et al., 2016b) and monitoring the effectiveness of reprocessing efforts. By the same token, the evolution of these indicators as more years are added in the data stream might reveal issues related to the most recent data in the merged series and/or in the single-mission data records.

The proposed methodology could be applied in a straightforward manner to other quantities, derived from ocean color (such as the remote sensing reflectance or inherent optical properties), or from other remote sensing sources. It could also be adapted to evaluate how outputs from biogeochemical models represent long-term variability with respect to satellite products. The analysis relied mostly on statistical indicators applied to trend distributions. The methodology could therefore support the effort required to demonstrate that a given data set fulfills the GCOS requirements in terms of temporal stability (GCOS, 2011). However it is highlighted that it is by no means the only way CDRs should be evaluated. First, more advanced statistical techniques could certainly be developed, particularly taking into account the uncertainties associated with the different products. Then, following a complementary route, additional, more qualitative, criteria can be listed (such as the main characteristics of algorithms, or their applicability to all missions) to evaluate if a data set meets the requirements for climate research (Sathyendranath et al., 2017, this issue). More generally, comprehensive sets of metrics have been proposed to assess the maturity of data sets for climate research (Bates and Privette, 2012) which are not considered here.

5.2. Interpretation of the results

The application of these criteria to the OC-CCI data showed that the agreement with single-mission products was similar to the consistency shown by these products. The agreement of OC-CCI with the SeaWiFS series (1998–2007) was remarkable with only 1.2% of the domain of analysis characterized by P lower than 0.05 and a κ index of 0.66, while the differences in trend slopes were nicely centered on zero (Fig. 5). When compared with MERIS or MODIS-A data over August 2002–July 2011 or 2012, P was lower than 0.05 for approximately 2% of the ocean, and κ was ~0.64. However, an area of lower P was observed in the south central equatorial Pacific, associated with different magnitudes of the trends but not with differences in patterns (Figs. 8 and 9). As baseline comparison, the agreement between MODIS-A and MERIS over the period 2002–2011 could be summarized by a proportion of the ocean with $P < 0.05$ of 1.9% and a κ of 0.60. These results suggest that the OC-CCI record can be considered

fit to conduct temporal analyses over the period 1997 to 2012 (with some caution in the south central Pacific).

Considering that the OC-CCI data set relies on the SeaWiFS, MERIS and MODIS-A data for its construction, this conclusion stating a good agreement might have been expected, but this reasoning is in large part misleading. In particular, trend detection based on multi-mission (merged or concatenated) data appears extremely sensitive to inter-mission biases (Mélin, 2016), so that the agreement observed between OC-CCI data and single-mission products by no means should be taken for granted. In fact the results have evolved with the OC-CCI data set versions: for instance the comparison between OC-CCI version 2 and MERIS or MODIS-A data did not show the pattern of low P values in the south central Pacific, whereas the agreement with SeaWiFS series was significantly degraded from version 1 to version 2 (not shown) and improved again with version 3. Moreover, the methodology applied to other multi-mission data sets that do not include bias corrections in their processing stream showed much degraded results with respect to OC-CCI (see examples in Mélin et al., 2016b; Mélin, 2016), in some cases to the point of being unsuitable for climate research. It might seem an easy task to create a multi-mission data record consistent, as far as trends are concerned, with single-mission data included in its construction, but this is actually not the case. This further underlines that properly addressing inter-mission differences remains a crucial issue in the construction of multi-mission CDRs.

5.3. Single-mission records and the issue of recent years

In line with the work performed by space agencies on sensor calibration, the assumption that single-mission series from SeaWiFS, MERIS and MODIS-A are free from spurious artifacts appears fairly robust over the period 1997–2012, even though the signature of the deactivated OCL can be seen in the time series related to MERIS (Figs. 1 and 2) and the MODIS-A band at 412 nm shows a reduced confidence after 2007 (B. Franz, personal communication). The assumption is further questioned over the recent period by calibration issues related to the aging of the MODIS-A sensor. MODIS-T now presents an impressively long series but issues associated with calibration over its life time (Franz et al., 2008) suggest great caution when analyzing its temporal variability while its dependence to MODIS-A calibration (Kwiatkowska et al., 2008) limits its use for independent assessments. In fact, the trend distributions associated with MODIS-A and MODIS-T appeared very consistent, for instance with a percentage of the ocean with $P < 0.05$ lower than 1% regardless of the considered period. In the case of OC-CCI, extending the series after 2012 meant a degraded agreement with respect to MODIS-A, the part of the ocean characterized by $P < 0.05$ increasing from 1.6% to 5.1% as four more years were included (2012 to 2015). It is recalled that the OC-CCI data relied exclusively on MODIS-A and VIIRS over the period from April 2012 (demise of MERIS) to 2015. Part of this degradation in trend consistency might be due to the aging of the MODIS-A sensor (Meister et al., 2012; Meister and Franz, 2014), even though it can not be excluded that residual inter-mission differences had an impact on the construction of the OC-CCI record. It is also worth noting that this degradation happened largely in relation with the pattern of low P values observed in Fig. 9 in the central Pacific south of the Equator: as more years were added, P further decreased and the pattern spread. The rest of the ocean appeared mostly unaffected. In conclusion it seems that the OC-CCI data are suitable for temporal analysis, with more caution due (particularly in the equatorial Pacific) as the time interval of interest exceeds 2012. As a positive note, and considering a history of regular reprocessing and progress, future improvements in the characterization of the MODIS-A calibration including the final part of the data record are possible.

Unfortunately, the period covered by VIIRS is still rather short and with a small overlap with MERIS. The overlap with MODIS-A is more substantial but it is a period when that instrument showed signs of radiometric degradation. Nevertheless the comparison of trends over four years of data gave an excellent agreement between products, particularly when OC-CCI and VIIRS data were considered (the proportion of the ocean with $P < 0.05$ being 1.3% and κ equal to 0.72). The only noticeable pattern with low P was again seen in the south central Pacific. So, with some caution, the OC-CCI data record appears appropriate to study inter-annual variability over the last few years. These considerations highlight the challenges inherent to building a multi-mission data record suitable for climate studies, challenges that are compounded when using aging satellites or successive missions with little overlap. In particular properly bridging the period from 2012 (arbitrarily taken as a date when MODIS data start losing quality) to 2016 (when the overlap between VIIRS and OLCI is initiated) seems critical. As a corollary, this study recalls the importance of operational missions to maintain ocean color time series over multiple decades.

6. Conclusions

The detection of climate change signals above the natural variability requires long time series of ocean color data that can only be obtained by combining data sets from successive missions. The existence of inter-mission differences or other artifacts can easily be transferred to a multi-mission time series and create spurious signals in the derived trends (Mélin, 2016) so that methodologies are required to detect these occurrences and ultimately to assess if multi-mission data sets are suitable for climate investigations and to inform interested users adequately. Such a methodology described here relies on the comparison between trends associated with the multi-mission data and those found for single-mission series. Under the assumption that those are free from spurious temporal artifacts, the trends observed with the multi-mission data should be consistent with the single-mission series. This agreement can be evaluated with a three-fold approach with contingency matrices, the direct comparison of the slopes and statistical tests. The following conclusions are drawn:

- the proposed methodology, the outputs of which can be summarized with a few statistics and a map (the P distribution), offers suitable tools to evaluate the climate quality for trend detection of multi-mission data records;
- when the approach is applied to the Chl*a* data series from OC-CCI, the results suggest that it can safely be used for temporal analysis when limited to the period before 2012;
- when more recent years are included, the amplitude of the trends might be affected by issues likely related to sensor insufficient characterization or degradation, but the trend patterns seem preserved;
- nonetheless, the trends displayed by the OC-CCI series over the short period 2012–2015 are very consistent with those of VIIRS.

Overall, the study shows encouraging results about our capacity to create Chl*a* climate data records. But creating climate-quality single-mission series and ensuring a cross-mission consistency allowing climate research is still a major task (Mélin and Franz, 2014) that requires sustained efforts if we want to make best use of the first twenty years of continuous ocean color data.

Acknowledgments

This work is a contribution to the Ocean Colour Climate Change Initiative (OC-CCI) of the European Space Agency. Additional support came from the National Centre for Earth Observation of the Natural

Environment Research Council of UK. Acknowledgments are due to ESA and NASA for the distribution of satellite data.

References

Andrade, J.M., Estévez-pérez, M.G., 2014. Statistical comparison of the slopes of two regression lines: a tutorial. Anal. Chim. Acta 838, 1–12.

Bates, J.J., Privette, J.L., 2012. A maturity model for assessing the completeness of climate data records. EOS Trans. Am. Geophys. Union 93, 441.

Beaulieu, C., Henson, S.A., Sarmiento, J.L., Dunne, J.P., Doney, S.C., Rykaczewski, R.R., Bopp, L., 2013. Comment on factors challenging our ability to detect long-term trends in ocean chlorophyll. Biogeosciences 9, C8200–C8209.

Bélanger, S., Babin, M., Tremblay, J.-E., 2013. Increasing cloudiness in Arctic damps the increase in phytoplankton primary production due to sea ice receding. Biogeosciences 10, 4087–4101.

Bell, G.D., Halpert, M., L'Heureux, M., 2016. ENSO and the tropical Pacific. State of the Climate 2015. 97. Bulletin of the American Meteorological Association., pp. S93–S98.

Bojinski, S., Verstraete, M., Peterson, T.C., Richter, C., Simmons, A., Zemp, M., 2014. The concept of Essential Climate Variables in support of climate research, applications, and policy. Bull. Am. Meteorol. Soc. 95, 1431–1443.

Bond, N.A., Cronin, M.F., Freeland, H., Mantua, N., 2015. Causes and impacts of the 2014 warm anomaly in the NE Pacific. Geophys. Res. Lett. 42, 3414–3420. http://dx.doi.org/10.1002/2015GL063306.

Boyce, D.G., Dowd, M., Lewis, M.R., Worm, B., 2014. Estimating global chlorophyll changes over the past century. Prog. Oceanogr. 122, 163–173.

Cao, C., Xiong, J., Blonski, S., Liu, Q., Uprety, S., Shao, X., Bai, Y., Weng, F., 2013. Suomi NPP VIIRS sensor data record verification, validation, and long-term performance monitoring. J. Geophys. Res. 118. http://dx.doi.org/10.1002/2013JD020418.

Cavole, L.M., Demko, A.M., Diner, R.E., Giddings, A., Koester, I., Pagniello, C.L.M.S., Paulsen, M.L., Ramirez-Valdez, A., Schwenck, S.M., Yen, N.K., Zill, M.E., Franks, P.J.S., 2016. Biological impacts of the 2013–2015 warm-water anomaly in the Northeast Pacific: winners, losers, and the future. Oceanography 29, 273–285.

CCI, 2016a. Ocean Colour Climate Change Initiative Product User Guide. , pp. 47. http://www.esa-oceancolour-cci.org/.

CCI, 2016b. OC-CCI ATBD Pixel Identification. Ocean Colour-Climate Change Initiative, Algorithm Theoretical Basis Document, Version 2.0. pp. 57.. http://www.esa-oceancolour-cci.org/.

CCI, 2016c. Ocean colour data bias correction and merging. Ocean colour-Climate Change Initiative, algorithm theoretical basis document, version 3.7. pp. 36. http://www.esa-oceancolour-cci.org/.

Cohen, J., 1960. A coefficient of agreement for nominal scales. Educ. Psychol. Meas. 20, 213–220.

Coppini, G., Lyubarstev, V., Pinardi, N., Colella, S., Santoleri, R., Christiansen, T., 2013. The use of ocean-colour data to estimate Chl-a trend in European seas. Int. J. Geosci. 4, 927–949.

Couto, A.B., Brotas, V., Mélin, F., Groom, S., Sathyendranath, S., 2016. Inter-comparison of OC-CCI chlorophyll-a estimates with precursor datasets. Int. J. Remote Sens. 37, 4337–4355.

Djavidnia, S., Mélin, F., Hoepffner, N., 2010. Comparison of global ocean colour data records. Ocean Sci. 6, 61–76.

Doney, S.C., 2006. Plankton in a warmer world. Nature 444, 695–696.

Donlon, C., Berruti, B., Buongiorno, A., Ferreira, M.H., Féménias, P., Frerick, J., Goryl, P., Klein, U., Laur, H., Mavrocordatos, C., Nieke, J., Rebhan, H., Seitz, B., Stroede, J., Sciarra, R., 2012. The global monitoring for environment and security (GMES) Sentinel-3 mission. Remote Sens. Environ. 120, 37–57.

Eplee, R.E., Meister, G., Patt, F.S., Barnes, R.A., Bailey, S.W., Franz, B.A., McClain, C.R., 2012. On-orbit calibration of SeaWiFS. Appl. Opt. 51, 8702–8730.

Eplee, R.E., Turpie, K.R., Meister, G., Patt, F.S., Franz, B.A., Bailey, s.W., 2015. On-orbit calibration of the Suomi National Polar-Orbiting Partnership Visible Infrared Imaging Suite for ocean color applications. Appl. Opt. 54, 1984–2006.

Esaias, W.E., Abbott, M.R., Barton, I., Brown, O.B., Campbell, J.W., Carder, K.L., Clark, D.K., Evans, R.H., Hoge, F.E., Gordon, H.R., Balch, W.M., Letelier, R., Minnett, P.J., 1998. An overview of MODIS capabilities for ocean science observations. IEEE Trans. Geosci. Remote Sens. 36, 1250–1265.

Fabry, V.J., Seidel, B.A., Feely, R.A., Orr, J.C., 2008. Impacts of ocean acidification on marine fauna and ecosystem processes. ICES J. Mar. Sci. 65, 414–432.

Franz, B.A., Bailey, S.W., Werdell, P.J., McClain, C.R., 2007. Sensor-independent approach to the vicarious calibration of satellite ocean color radiometry. Appl. Opt. 46, 5068–5082.

Franz, B.A., Kwiatkowska, E., Meister, G., McClain, C.R., 2008. Moderate. J. Appl. Remote. Sens. 2, 023525. http://dx.doi.org/10.1117/1.2957964.

Galloway, J.N., Townsend, A.R., Erisman, J.W., Bekunda, M., Cai, Z., Freney, J.R., Martinelli, A., Seitzinger, S.P., Sutton, M.A., 2008. Transformation of the nitrogen cycle: recent trends, questions, and potential solutions. Science 320, 889–892.

GCOS, 2011. Systematic observation requirements for satellite-based data products for climate, 2011 update: supplemental details to the satellite-based component of the implementation plan for the global observing system for climate in support of the UNFCCC. GCOS. World Meteorological Organization., pp. 154.

Gohin, F., Druon, J.N., Lampert, L., 2002. A five-channel chlorophyll concentration algorithm applied to SeaWiFS data processed by SeaDAS in coastal waters. Int. J. Remote Sens. 23, 1639–1661.

Gregg, W.W., Casey, N.W., 2010. Improving the consistency of ocean color data: a step toward climate data records. Geophys. Res. Lett. 37, L04605. http://dx.doi.org/10.1029/2009GL041983.

Gregg, W.W., Rousseaux, C., 2014. Decadal trends in global pelagic ocean chlorophyll: a new new assessment integrating multiple satellites, in situ data, and models. J. Geophys. Res. 119, 5921–5933. http://dx.doi.org/10.1002/2014JC010158.

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

Hollmann, R., Merchant, C.J., Saunders, R., Downy, C., Buchwitz, M., Cazenave, A., Chuvieco, E., Defourny, P., de Leeuw, G., Forsberg, R., Holzer-Popp, T., Paul, F., Sandven, S., Sathyendranath, S., van Roozendael, M., Wagner, W., 2013. The ESA Climate Change Initiative — satellite data records for essential climate variables. Bull. Am. Meteorol. Soc. 94, 1541–1552.

Hovis, W.A., Clark, D.K., Anderson, F., Austin, R.W., Wilson, W.H., Baker, E.T., Ball, D., Gordon, H.R., Mueller, J.L., El-Sayed, S.Z., Sturm, B., Wrigley, R.C., Yentsch, C.S., 1980. Nimbus-7 coastal zone color scanner: system description and initial imagery. Science 210, 60–63.

Hu, C., Lee, Z.P., Franz, B., 2012. Chlorophyll a algorithms for oligotrophic oceans: a novel approach based on three-band reflectance difference. J. Geophys. Res. 117, C01011. http://dx.doi.org/10.1029/2011JC007395.

Ibañez, F., Conversi, A., 2002. Prediction of missing values and detection of 'exceptional events' in a chronological planktonic series: a single algorithm. Ecol. Model. 154, 9–23.

IOCCG, Gregg, W.W., Aiken, J., Kwiatkowska, E., Maritorena, S., Mélin, F., Murakami, H., Pinnock, S., Pottier, C. (Eds.), 2007. Ocean colour data merging. Reports of the International Ocean-Colour Coordinating Group. 5. IOCCG, Dartmouth, Canada, pp. 65.

IOCCG, Frouin, R. (Ed.), 2013. In-flight calibration of satellite ocean-colour sensors. Reports of the International Ocean-Colour Coordinating Group. 14. IOCCG, Dartmouth, Canada, pp. 106.

Jacox, M.G., Hazen, e.L., Zaba, K.D., Rudnick, D.L., Edwards, C.A., Moore, A.M., Bograd, S.J., 2016. Impacts of the 2015–2016 El Niño on the California Current system: e assessment and comparison to past events. Geophys. Res. Lett. 43, 7072–7080. http://dx.doi.org/10.1002/2016GL069716.

Kahru, M., Kudela, R.M., Manzano-Sarabia, M., Mitchell, B.G., 2012. Trends in the surface chlorophyll of the California Current: merging data from multiple ocean color satellites. Deep-Sea Res. II 77-80, 89–98.

Kaufman, Y.J., Koren, I., Remer, L.A., Tanré, D., Ginoux, P., Fan, S., 2005. Dust transport and deposition observed from the terra-moderate resolution imaging spectroradiometer (MODIS) spacecraft over the Atlantic Ocean. J. Geophys. Res. 110, D10S12. http://dx.doi.org/10.1029/2003JD004436.

the OCTS Team, Kawamura, H., 1998. OCTS mission overview. J. Oceanogr. 54, 383–399.

Kwiatkowska, E.J., Fargion, G.S., 2003. Application of machine-learning techniques towards the creation of a consistent and calibrated chlorophyll concentration baseline dataset using remotely sensed ocean color data. IEEE Trans. Geosci. Remote Sens. 41, 2844–2860.

Kwiatkowska, E.J., Franz, B.A., Meister, G., McClain, C.R., Xiong, X., 2008. Cross-calibration of ocean-color bands from Moderate Resolution Imaging Spectroradiometer on Terra platform. Appl. Opt. 47, 6796–6810.

Lotze, H.K., Lenihan, H.S., Bourque, B.J., Bradbury, R.H., Cooke, R.G., May, M.C., Kidwell, S.M., Kirby, M.X., Peterson, C.H., Jackson, J.B.C., 2006. Depletion, degradation, and recovery potential of estuaries and coastal seas. Science 312, 1806–1809.

Maritorena, S., Hembise Fanton d'Andon, O., Mangin, A., Siegel, D.A., 2010. Merged satellite ocean color data products using a bio-optical model: characteristics, benefits and issues. Remote Sens. Environ. 114, 1791–1804.

McClain, C.R., Cleave, M.L., Feldman, G.C., Gregg, W.W., Hooker, S.B., Kuring, N., 1998. Science quality SeaWiFS data for global biosphere research. Sea Technol. 39, 10–16.

McClain, C.R., Feldman, G.C., Hooker, S.B., 2004a. An overview of the SeaWiFS project and strategies for producing a climate research quality global ocean bio-optical time series. Deep-Sea Res. II 51, 5–42.

McClain, C.R., Signorini, S.R., Christian, J.R., 2004b. Subtropical gyre variability observed by ocean-color satellites. Deep-Sea Res. II 51, 281–301.

Meister, G., Franz, B.A., Kwiatkowska, E.J., McClain, C.R., 2012. Corrections to the calibration of MODIS Aqua ocean color bands derived from SeaWiFS data. IEEE Trans. Geosci. Remote Sens. 50, 310–319.

Meister, G., Franz, B.A., 2014. Corrections to the MODIS Aqua calibration derived from MODIS Aqua calibration derived from MODIS Aqua ocean color products.. IEEE Trans. Geosci. Remote Sens. 52, 6534–6541.

Mélin, F., Steinich, C., Gobron, N., Pinty, P., Verstraete, M.M., 2002. Optimal merging of LAC and GAC data from SeaWiFS. Int. J. Remote Sens. 23, 801–807.

Mélin, F., Zibordi, G., 2007. An optically-based technique for producing merged spectra of water leaving radiances from ocean color. Appl. Opt. 46, 3856–3869.

Mélin, F., Zibordi, G., Djavidnia, S., 2009. Merged series of normalized water leaving radiances obtained from multiple satellite missions for the Mediterranean Sea. Adv. Space Res. 43, 423–437.

Mélin, F., 2010. Global distribution of the random uncertainty associated with satellite derived Chla. IEEE Geosci. Remote Sens. Lett. 7, 220–224.

Mélin, F., 2011. Comparison of SeaWiFS and MODIS time series of inherent optical properties for the Adriatic Sea. Ocean Sci. 7, 351–361.

Mélin, F., Vantrepotte, V., Clerici, M., D'Alimonte, D., Zibordi, G., Berthon, J.F., Canuti, E., 2011. Multi-sensor satellite time series of optical properties and chlorophyll a concentration in the Adriatic Sea. Prog. Oceanogr. 91, 229–244.

Mélin, F., Franz, B.A., 2014. Assessment of satellite ocean colour radiometry and derived geophysical products. In: Zibordi, G., Donlon, C., Parr, A. (Eds.), Optical Radiometry for Oceans Climate Measurements. vol. 47. Academic Press, Experimental Methods in the Physical Sciences., pp. 722. chapter 6.1.

Mélin, F., Sclep, G., 2015. Band shifting for ocean color multi-spectral reflectance data. Opt. Express 23, 2262–2279.

Mélin, F., 2016. Impact of inter-mission differences and drifts on chlorophyll-a trend estimates. Int. J. Remote Sens. 37, 2061–2079.

Mélin, F., Sclep, G., Jackson, T.J., Sathyendranath, S., 2016a. Uncertainty estimates of remote sensing reflectance derived from comparison of ocean color satellite data set. Remote Sens. Environ. 177, 107–124.

Mélin, F., Vantrepotte, V., Chuprin, A., Grant, M., Jackson, T., Sathyendranath, S., 2016b. Global trends in chlorophyll concentration observed with the satellite ocean colour data record. Proceedings of the Living Planet Symposium 2016, Prague, Czech Rep., SP-740. 7. pp. 9–13. May 2016.

Moore, T.S., Dowell, M.D., Bradt, S., Ruiz Verdu, R., 2014. An optical water type framework for selecting and blending retrievals from bio-optical algorithms in lakes and coastal waters. Remote Sens. Environ. 143, 97–111.

NASA, 2015. Reprocessing 2014.0. http://oceancolor.gsfc.nasa.gov/cms/reprocessing/OCReproc20140.html.

Nieto, K., Mélin, F., 2017. Variability of chlorophyll-a concentration in the Gulf of Guinea and its relation to physical oceanographic variables. Prog. Oceanogr. 151, 97–115.

O'Reilly, J.E., Maritorena, S., Siegel, D.A., O'Brien, M.C., Toole, D.A., Mitchell, B.G., Kahru, M., Chavez, F.P., Strutton, P., Cota, G.F., Hooker, S.B., McClain, C.R., Carder, K.L., Mueller-Karger, F., Harding, L., Magnusson, A., Phinney, D., Moore, G.F., Aiken, J., Arrigo, K.R., Letelier, R., Culver, M., Hooker, S.B., Firestone, E.R., 2000. Ocean color chlorophyll a algorithms for SeawiFS, OC2, and OC4:. NASA Technical Memorandum 2000-206892, 20. NASA-GSFC, Greenbelt, Maryland. , pp. 9–23. Version 4.

Park, J.-Y., Kug, J.-S., Bader, J., Rolph, R., Kwon, M., 2015. Amplified Arctic warming by phytoplankton under greenhouse warming. Proc. Natl. Acad. Sci. 112, 5921–5926.

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

Polovina, J.J., Woodsworth, P.A., 2011. Declines in phytoplankton cell size in the subtropical oceans estimated from satellite remotely-sensed temperature and chlorophyll, 1998-2007. Deep-Sea Res. II 77-80, 82–88.

Pottier, C., Garçon, V., Larnicol, G., Sudre, J., Schaeffer, P., Le Traon, P.-Y., 2006. Merging SeaWiFS and MODIS/Aqua ocean color data in North and Equatorial Atlantic using weighted averaging and objective analysis. IEEE Trans. Geosci. Remote Sens. 44, 3436–3451.

Racault, M.F., Raitsos, D.E., Berumen, M.L., Brewin, R.J.W., Platt, T., Sathyendranath, S., Hoteit, I., 2015. Phytoplankton phenology indices in coral reef ecosystems: application to ocean-color observations in the Red Sea. Remote Sens. Environ. 160, 222–234.

Radenac, M.H., Léger, F., Singh, A., Delcroix, T., 2012. Sea surface chlorophyll signature in the tropical Pacific during eastern and central Pacific ENSO events. J. Geophys. Res. 117, C04007. http://dx.doi.org/10.1029/2011JC007841.

Rast, M., Bezy, J.L., Bruzzi, S., 1999. The ESA Medium resolution imaging spectrometer MERIS — a review of the instrument and its mission. Int. J. Remote Sens. 20, 1681–1702.

Roberts, C.M., 2003. Our shifting perspectives on the oceans. Oryx 37, 166–177.

Sarmiento, J., Slater, R., Barber, R., Bopp, L., Doney, S., Hirst, A., Kleypas, J., Matear, R., Mikolajewicz, U., Monfray, P., Soldatov, V., Spall, S., Stouffer, R., 2004. Response of ocean ecosystems to climate warming. Global Biogeochem. Cycles 18, GB3003. http://dx.doi.org/10.1029/2003GB002134.

Sathyendranath, S., Brewin, R.J.W., Jackson, T., Mélin, F., Platt, T., 2017. Ocean–colour products for climate-change studies: what are their ideal characteristics?. Remote Sens. Environ. submitted.

Saulquin, B., Fablet, R., Mangin, A., Mercier, G., Antoine, D., Fanton d'Andon, O., 2013. Detection of linear trends in multisensor time series in the presence of autocorrelated noise: application to the chlorophyll-a SeaWiFS and MERIS data sets and extrapolation to the incoming Sentinel 3 — OLCI mission. J. Geophys. Res. 118, 3752–3763. http://dx.doi.org/10.1002/jgrc.20264.

Schueler, C.F., Clement, J.E., Ardanuy, P.E., Welsch, C., DeLuccia, F., Swenson, H., 2002. NPOESS VIIRS sensor design overview. Proceedings SPIE 4483, Earth Observing Systems. VI. http://dx.doi.org/10.1117/12453451.

Signorini, S.R., McClain, C.R., 2012. Subtropical gyre variability as seen from satellites. Remote Sens. Lett. 3, 471–479.

Signorini, S.R., Franz, B.A., McClain, C.R., 2015. Chlorophyll variability in the oligotrophic gyres: mechanisms, seasonality and trends. Front. Mar. Sci. 2, http://dx.doi.org/10.3389/fmars.2015.00001.

Sokal, R.R., Rohlf, F.J., 1994. Biometry: The Principles and Practice of Statistics in Biological Research. Third, W.H. Freeman and Company, pp. 887.

Soppa, M.A., Völker, C., Bracher, A., 2016. Diatom phenology in the Southern Ocean: mean patterns, trends and the role of climate oscillations. Remote Sens. 8, 420. http://dx.doi.org/10.3390/rs805420.

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., Segschneider, J., 2010. Projected 21st century decrease in marine productivity: a multi-model analysis. Biogeosciences 7, 979–1005.

Steinmetz, F., Deschamps, P.-Y., Ramon, D., 2011. Atmospheric correction in the presence of sun glint: application to MERIS. Opt. Express 19, 9783–9800.

Stewart, K.R., Lewison, R.L., Dunn, D.C., Bjorkland, R.H., Kelez, S., Halpin, P.N., Crowder, L.B., 2010. Characterizing fishing effort and spatial extent of coastal fisheries. PLoS ONE 5, e14451. http://dx.doi.org/10.1371/journal.pone.0014451.

Stramma, L., Fischer, T., Grundle, D.S., Krahmann, G., Bange, W., Marandino, C.A., 2015. Observed El Niño conditions in the eastern tropical Pacific in October 2016. Ocean Sci. 12, 861–873.

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

Vantrepotte, V., Mélin, F., 2011. Inter-annual variations in the SeaWiFS global chlorophyll a concentration (1997–2007). Deep-Sea Res. I 58, 429–441.

Vantrepotte, V., Loisel, H., Mélin, F., Desailly, D., Duforêt-Gaurier, L., 2011. Global particulate matter pool temporal variability over the SeaWiFS period (1997–2007). Geophys. Res. Lett. 38, L02605. 2010GL046167.

Viera, A.J., Garrett, J.M., 2005. Understanding interobserver agreement: the kappa statistics. Fam. Med. 37, 360–363.

Warrens, M.J., 2011. Cohen's kappa is a weighted average. Stat. Methodol. 8, 473–484.

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

Xiong, X., Sun, J., Xie, X., Barnes, W.L., Salomonson, V.V., 2010. On-orbit calibration and performance of Aqua MODIS reflective solar bands. IEEE Trans. Geosci. Remote Sens. 48, 535–546.

Yoder, J.A., Kennelly, M.A., Doney, S.C., Lima, I.D., 2010. Are trends in SeaWiFS chlorophyll time-series unusual relative to historic variability? Acta Oceanol. Sin. 29, 1–4.

Zimmermann, G., Neumann, A., 1997. The spaceborne imaging spectrometer MOS for ocean remote sensing. Proceedings of the 1st International Workshop on MOS-IRS and ocean color. pp. 1–9. Berlin, Apr. 28–30 1997.