Revised manuscript for: *Geophysical Research Letters*

Assessing the presence of discontinuities in the ocean color satellite record and their effects on chlorophyll trends and their uncertainties

Matthew L. Hammond[1,2], Claudie Beaulieu[1,3], Stephanie A. Henson[2] & Sujit K. Sahu[4]

1. Ocean and Earth Science, University of Southampton, SO14 3ZH, UK
2. National Oceanography Centre, Southampton, SO14 3ZH, UK
3. Ocean Sciences Department, University of California, Santa Cruz, CA, 95064, USA
4. Mathematical Sciences, University of Southampton, SO17 1BJ, UK

Key points

1) Discontinuities in multi-sensor ocean color chlorophyll records are detected in ~70 % of regions using a Bayesian space-time model

2) Discontinuities affect trend estimates in ~60 % of regions and can even bias the trends' sign (opposite sign in ~13% of regions)

3) The uncertainty of trend estimates increases by an average of 0.20 %yr^{-1} for a single discontinuity and 0.59 %yr^{-1} for two discontinuities

Abstract

Ocean color sensors are crucial for understanding global phytoplankton dynamics. However, the limited lifespans of sensors make multi-sensor datasets necessary for estimating long-term trends. Discontinuities may be introduced when merging data between sensors, potentially affecting trend estimates and their uncertainties. We use a Bayesian spatio-temporal model to investigate the presence of discontinuities and their impacts on estimated chlorophyll trends. The discontinuities considered are the introduction of MERIS, MODIS-Aqua, and VIIRS, and the termination of SeaWiFS. Discontinuities are detected in ~70 % of regions, affecting trend estimates (~60 % of regions have statistically different trends), and potentially even biasing trend estimates (opposite sign in ~13 % of regions). Considering a single discontinuity increases trend uncertainty by an average of 0.20% yr^{-1} (0.59% yr^{-1} for two discontinuities). This difference in trend magnitude and uncertainty highlights the importance of minimizing discontinuities in multi-sensor records and taking into account discontinuities when analyzing trends.

Index terms: 1635 1640 1986 1990 4855

Keywords: Chlorophyll, Bayesian Statistics, Spatio-Temporal Modeling, Discontinuities, Trend Estimation.

1 Introduction

Ocean color satellite records can be used to assess how global phytoplankton biomass may be affected by climate change. These records are especially suited to this task because of their high spatial coverage and temporal resolution (e.g. McClain, 2009). However, there are major challenges inherent to trend detection in chlorophyll-a (chl) derived from ocean color sensors. These include the low signal-to-noise ratio, the large degree of natural variability, and the shortness of the record (e.g. Beaulieu et al., 2013; Henson et al., 2010; Mélin et al., 2016; Saulquin et al., 2013). A comparison of observational, i.e. in situ and satellite, chl observations found that shorter datasets have conflicting, and larger magnitude, trend estimates when compared to longer records (Boyce & Worm, 2015). The large magnitude of natural variability can obscure a smaller magnitude long-term trend, thus challenging trend estimation.

To compensate for the shortness of any single ocean color record, multi-sensor datasets can be used. These combine the available ocean color sensors using various approaches (e.g. Lavender et al., 2015; Maritorena & Siegel, 2005). The four main ocean color sensors providing the longest overlapping period of coverage to date are: Medium Resolution Imaging Spectrometer (MERIS) (April 2002 to April 2012), Moderate Resolution Imaging Spectroradiometer aboard the Aqua satellite (MODIS-Aqua) (July 2002 to present), Sea-Viewing Wide Field-of-View Sensor (SeaWiFS) (September 1997 to December 2010), and Visible Infrared Imaging Radiometer Suite (VIIRS) (Jan 2012 to present). The approach used to combine satellite records must fully compensate for the differences between the individual datasets, which can vary temporally and spatially (Djavidnia et al., 2010). If the differences between datasets are not accounted for discontinuities may be introduced, trends estimated from the combined record may thus be biased and/or have increased uncertainty (Gregg & Casey, 2010). Such discontinuities may include a permanent mean-shift in the observed value, i.e. a mean-shift discontinuity (Weatherhead, 1998), which are considered here. Even with the use of multi-sensor records, the maximum available length of chl record is still only approximately 20 years, from the launch of SeaWiFS to present, shorter than the suggested ~30 years required to distinguish a climate change driven chl trend from natural variability (Henson et al., 2016; 2010).

To assess the effects of potential discontinuities on trend estimation, we model the discontinuities alongside the long-term trend as suggested in Weatherhead (1998). More specifically, we use a Bayesian spatio-temporal model, which has been shown to provide an accurate fit and complete assessment of uncertainty when estimating chl trends (Hammond et al., 2017). We consider three major discontinuities in the satellite record: the launch of both the MERIS and MODIS-Aqua sensors in the spring/summer of 2002, the termination of the

SeaWiFS sensor at the end of 2010, and the launch of the VIIRS satellite, providing data from the start of 2012.

2 Methods

2.1 Data

The chl data come from version 3.1 of ESA's OC-CCI project (Lavender et al., 2015; available at: http://www.esa-oceancolour-cci.org/). This product combines data from the SeaWiFS, MERIS, MODIS-Aqua (NASA R2014.0.1 reprocessing), and VIIRS sensors to create a continuous, bias-corrected monthly mean time-series running from September 1997 to December 2016 inclusive. Band-shifting and bias-correction techniques are used to combine the data from individual sensors. The band-shifting is performed using a bio-optical model inversion (Mélin & Sclep, 2012; 2015). The bias-correction is performed by adjusting pixel-level radiances to reduce the difference between SeaWiFS and the other sensors; a time window with increased central weight is used to correct seasonal biases (Chuprin et al., 2017; Djavidnia et al., 2010; Grant et al., 2017). We process this dataset by downscaling to a 1° grid (by averaging within 1° boxes) and by log-transforming chl values, after Campbell (1995).

As a comparison, we also perform the analysis on 1° gridded monthly mean data from the GlobColour dataset (available at: http://globcolour.info) in which SeaWiFS, MERIS, MODIS-Aqua (R2014.0.1), and VIIRS sensors are merged using the Garver, Siegel, Maritorena Model (GSM) process (Maritorena & Siegel, 2005; Maritorena et al., 2010). The GSM process combines sensor observations of water-leaving radiance to form a multi-source spectrum for each pixel. The multi-source spectrum is then inverted with a semi-analytical ocean color model, which describes the relationship between water-leaving radiance and the inherent optical properties of seawater, including backscattering and absorption coefficients (Maritorena et al., 2002; 2010). We use the Case 1 (open ocean) data only, as we do not consider coastal regions (see above). A log-transformation is also used on the GlobColour chl data. To help explain natural variability in the chl data, SST is used as a covariate. SST data are sourced from the NOAA optimum interpolation v2 monthly mean data product (Reynolds et al., 2002; available at: http://www.esrl.noaa.gov/psd/data/gridded/data.noaa.oisst.v2.html).

Trends are analyzed in 23 regions, based on those defined by Longhurst (1995, 1998). Coastal and polar waters are excluded due to issues with the availability and quality of data. Longhurst provinces are defined by biogeochemical and physical factors, and thus should have consistent trend amplitude and direction (Hammond et al., 2017).

2.2 Model Formulation

A hierarchical Bayesian spatio-temporal model is fitted separately in each of the 23 Longhurst regions retained for analysis (i.e. we use an un-pooled model with region-based independent fitting). This model uses all the data points inside each province and uses their spatial and temporal relationship to produce a province-wide set of parameter estimates (e.g. of trend and discontinuity). This approach provides a more accurate fit to observations and a more realistic assessment of uncertainty, when compared to averaging gridded trend estimates across the region (Hammond et al., 2017). The latter approach may also increase the risk of false positives (e.g. Wilks, 2016).

The key equations are presented below. First, the relationship between observed chl $Z_{n,t}$ and its true underlying value $O_{n,t}$ at location n and at month t is represented as:

$$Z_{n,t} = O_{n,t} + \varepsilon_{n,t} \quad (1)$$

where $\varepsilon_{n,t}$ is an independently normally distributed white noise process with zero mean and an unknown pure error variance, which primarily represents random measurement error (as well as environmental variability on scales finer than the grid spacing). A regression model is used to represent the true chl value (at grid point n and time t):

$$O_{n,t} = \boldsymbol{x}'_{n,t}\boldsymbol{\beta} + \boldsymbol{a}'_n\boldsymbol{w}_{m,t} \quad (2)$$

This regression model is composed of the covariates (including intercept) $\boldsymbol{x}_{n,t}$, the regression coefficients (constant for each region) $\boldsymbol{\beta} = (\beta_0, \beta_{Trend}, \beta_{SST}, \beta_{Disc}, \beta_{M1}, \ldots, \beta_{M12})$, and the term $\boldsymbol{a}'_n\boldsymbol{w}_{m,t}$ representing spatial and temporal correlation.

The spatial correlation is represented by an exponential decay away from site n, and the temporal correlation by a first order autoregressive process (i.e. a function of the preceding month). The term $\boldsymbol{a}'_n$ refers to the kriging coefficients at the grid $(n_1, n_2, \ldots, n_N)$ and the knot $(m_1, m_2, \ldots, m_M)$ locations. The knot locations are a reduced set of the grid locations, used to decrease the size of the spatial covariance matrix, allowing the large volumes of data used to be more efficiently computed. The term $\boldsymbol{w}_{m,t}$ represents the reduced spatio-temporal random effects at the knot locations.

The covariates include the date of the observation, the month (represented as factor levels where each month has an additional term, constant for all years), and SST. Time is used to estimate the temporal trend, the monthly factor is used to represent the seasonal cycle, and SST is used to isolate environmental variability. Including the SST term was shown to improve model fit as well as prevent issues with convergence (supporting information). As SST may capture a portion of the long-term chl trend, the trend estimated here represents the remaining long-term change not explained by SST variability. The regression coefficients correspond to the covariates as follows: β_0 to the intercept, β_{Trend} to the trend, β_{SST} to SST, β_{Disc} to the mean-shift discontinuity, and $\beta_{M1}, \ldots, \beta_{M12}$ to the monthly factor levels. Note that the monthly factor is not included in the Pacific Subarctic Gyres Province (East) (Region 18), because of the difficulty in identifying a stable phenology (supporting information Text S2).

The discontinuity covariate $\boldsymbol{x}^{Disc}$ indicates the presence of a mean-shift (we do not consider gradual drift between sensors) and is represented as a factor that is different either side of the known time of discontinuity t_{Disc} (Weatherhead, 1998):

$$\boldsymbol{x}_t^{Disc} = \begin{cases} 0, & t < t_{Disc} \\ 1, & t \geq t_{Disc} \end{cases} \quad (3)$$

We consider five scenarios based on major satellite inclusions and failures. The first is a scenario with no discontinuities (N-scenario). The second scenario has one discontinuity between the launches of the MERIS and MODIS-Aqua sensors in June 2002 (M-scenario). June 2002 is the time equidistant between their operational dates of April and July 2002, respectively. The third scenario has one discontinuity at the failure of the SeaWiFS satellite in December 2010 (S-scenario). The fourth scenario is when we consider both these discontinuities in the same model (MS-scenario). The final scenario is when all discontinuities mentioned above are considered, plus the launch of the VIIRS sensor in January 2012 (MSV-scenario). An additional scenario combining both the MERIS/MODIS discontinuity and the VIIRS discontinuity is considered in the supporting information (Text S3). For the multi-discontinuity scenarios (MS and MSV), the regression coefficient β_{Disc} includes additional t_{Disc} and $\boldsymbol{x}^{Disc}$ terms to estimate all discontinuities (i.e. two t_{Disc} and $\boldsymbol{x}^{Disc}$ terms for the MS-scenario and three for the MSV-scenario).

The modeling approach fits a full posterior distribution for each parameter. This study focuses on the trend and discontinuity parameters with their posterior mode representing the best estimate. The uncertainty of the trend and discontinuity estimates are represented by the 95 % credible interval of the posterior, defined as the 95 % highest density interval (Kruschke, 2015). We consider that a discontinuity is present if its magnitude is different from zero (i.e. its 95 % credible interval excludes zero). When comparing the trends in each region, we consider them likely to be statistically different from the baseline N-scenario if their 95 % credible intervals do not overlap with those of the N-scenario.

The spTimer package in R is used to estimate the model fit (Bakar & Sahu, 2015). See the supporting information and Hammond et al. (2017) for additional details on the model setup.

3 Results

3.1 Discontinuity magnitudes and their effect on trend estimates

The main text focuses on the ESA OC-CCI dataset; the scenarios using GlobColour data are analyzed in the supporting information (Text S4). In the majority of the regions in this study, we find that discontinuities are likely present and their magnitudes are large enough to affect trend estimates. The degree and direction of the effect is dependent on both the discontinuity scenario and region. We detect the presence of discontinuities in the majority of regions in all the discontinuity scenarios considered, although fewer are detected in the multi-discontinuity scenarios (Figure 1a). The majority of these regions also show that discontinuities affect trend estimates (Figure 1b).

The difference in trend estimates between the single discontinuity scenarios and the N-scenario is found to be inversely proportional to the discontinuity magnitude. The global average differences compared to the N-scenario, computed using weighting for the area and mean chl in each province, are as follows. We find that a discontinuity magnitude of 0.1 log(mg m^{-3}) leads to a trend that is -0.65% yr^{-1} different, based on global averages (Figures 1a and 1b). The discontinuity for the M-scenario is positive in most regions, leading to an overall negative trend difference (average of -0.54% yr^{-1}) (Figure 2). The opposite is found for the S-scenario (average of 0.59% yr^{-1}). For the MS-scenario, the sign of the difference is evenly distributed between positive and negative (average difference -0.028% yr^{-1}). In about half (12) of regions the trend difference for the MS-scenario lies between the trend differences for the two single discontinuity scenarios, suggesting they are partially cancelling out (Figure 1b). The MSV-scenario shows similar results to the MS-scenario with an average difference of 0.047% yr^{-1} (Figures 1b & 3a).

The average magnitude of trend differences (i.e. when the direction/sign of trend difference is omitted) is larger in the multi-discontinuity scenarios (1.1% yr^{-1} for the MSV-scenario and 0.85% yr^{-1} for the MS-scenario) than the single-discontinuity scenarios (0.65% yr^{-1} for the M-scenario and 0.81% yr^{-1} for the S-scenario). This can lead to a change of trend sign, i.e. from increasing to decreasing or vice versa, for example this occurs in 5 regions in the MSV-scenario. Despite differences between individual regions, there is no clear global pattern in either the trend difference or the discontinuity magnitude. The full results are presented in the supporting information, including an analysis using the GlobColour dataset that is found to show similar results, albeit with a slightly higher average trend difference in most scenarios (Table S2 & S3).

3.2 Effect of discontinuities on trend estimate uncertainties

Taking into account discontinuities increases uncertainty in all scenarios and regions. A single discontinuity increases trend uncertainty by an average of 0.21% yr^{-1} (Figure 1c). For the MS-scenario and MSV-scenario the increase in uncertainty is 0.64% yr^{-1} (Figure 1c & 3b). Individual regions show a disparity in the degree of uncertainty increase. The regions with the highest proportional increase in uncertainty, for the MSV-scenario relative to the N-scenario, are in the tropical to subtropical North Atlantic (average of 210 %). The regions with the smallest proportional uncertainty increase are typically found in the mid-latitude Pacific Ocean (average of 140 %). See supporting information for full results, including analysis using the GlobColour dataset, which is found to show similar results, albeit with a greater uncertainty difference in all scenarios.

4 Discussion

4.1 Ability to distinguish discontinuities and trends

Our results depend on our ability to distinguish trends and discontinuities accurately. We conduct a series of simulation studies to assess the model skill in accurately estimating trends and discontinuities (supporting information Text S5). We generate 100 synthetic datasets, of the same length as the present study, based on realistic values of chl, and its variability (with independent randomly generated noise in each dataset), and then superpose a range of realistic trends and discontinuities. We find that for these simulation studies the trend term is accurately estimated to within <1 %, and the discontinuity term is accurately estimated to within approximately 5 %. This suggests that our approach is highly capable of identifying trends and discontinuities, without confusing them with each other or with other components of chl variability.

4.2 How do discontinuities affect trend estimates?

The trend difference between the MSV-scenario and the N-scenario has an average magnitude of 1.1% yr^{-1}, and varies in the range ± 2.8% yr^{-1}, resulting in statistically different trends in 14 of the 23 regions. In a study that analyzed the effect of inter-sensor bias on trend detection, Mélin

(2016) showed that a 5 - 6 % bias between two sensors can lead to significantly different trends. This result was obtained by introducing artificial biases in the range 1 - 50 % when merging the SeaWiFS and MODIS-Aqua sensors. This illustrates the strong effect that discontinuities in the record can have, in agreement with the present study. However, Mélin (2016) also found that trends estimated for oligotrophic subtropical gyres are particularly sensitive to discontinuities in the record, which was attributed to the gyres' low natural variability. In our analysis, oligotrophic gyres do not seem to show such a pattern, except for the Pacific oligotrophic gyres, which show a larger than average trend difference (2.1% yr^{-1}) in the MSV-scenario relative to the N-scenario. The differences compared to Mélin (2016) are likely due to the substantial differences in the datasets and methodologies. Here we take into account discontinuities in a bias-corrected multi-sensor dataset using a spatio-temporal model with environmental variability isolated using SST whereas Mélin (2016) analyzed synthetic records with discontinuities induced prior to merging.

The discontinuity model in the present study represents a mean-shift, but biases between sensors can also increase over time and change over seasonal cycles (Djavidnia et al., 2010). A gradual drift in sensors' detected values may, like mean-shift discontinuities, directly affect trend estimates. Mélin (2016) determined that any drift greater than 2 % per decade can alter the conclusions of a trend analysis, which suggests this effect may be similarly important to mean-shift discontinuities. We do not consider drift here, as over the short-term period of drift (several years) it is likely to be confused with the trend estimate and lead to further increases in uncertainty and changes to the trend estimates.

The MODIS-Aqua sensor is known to be affected by sensor ageing, particularly towards the end of the study period, thus caution is advised for temporal analysis including the post-2012 period

(Mélin et al., 2017). To assess whether a drift in the MODIS-Aqua sensor may affect our results, we compare the trends detected over the period 1997-2016 in the present study to the trends detected over 1997-2013 in Hammond et al. (2017), which uses the ESA OC-CCI v2.0 dataset with the R2013.0.1 reprocessing MODIS-Aqua data. In Hammond et al. (2017), trends were detected in 17 of the 23 regions, as opposed to 19 such regions in the present study. The large-scale latitudinal pattern (whereby higher latitudes tend to have more positive trends) is also similar in both studies, 16 of the 23 regions in Hammond et al. (2017) have the same trend directions as the N-scenario. Although there are differences between the two studies which may be partly attributable to MODIS ageing effects, these are nevertheless minor and do not affect our conclusions.

4.3 How do discontinuities affect uncertainty in trend estimates?

Our results show that discontinuities in a record will increase the uncertainty of long-term trends, such that two discontinuities can double the uncertainty in trend estimates. Detection of trends in the current multi-sensor record may be particularly sensitive to the timing of discontinuities relative to decadal variability. The 1997/1998 El Niño event (Wolter & Timlin, 1998) lies before the MERIS/MODIS discontinuity, and the 2015/2016 El Niño event (Levine & McPhaden, 2016) follows the SeaWiFS and VIIRS discontinuities.

Trend detection may also be affected by the relative timing of discontinuities in the record. A discontinuity in the middle of a time-series is expected to have the greatest effect, which will decrease towards the beginning or end of the record (Beaulieu et al., 2013). The SeaWiFS discontinuity is further from either end of the record than the MERIS/MODIS discontinuity which may explain the larger uncertainty and trend differences seen in the S-scenario. Conversely, the VIIRS and SeaWiFS discontinuities are only separated by 1 year, potentially explaining the comparable results in the MSV-scenario and the MS-scenario. The increase in trend uncertainty when taking into account discontinuities is likely to make trend detection more challenging when using multi-sensor records, and will only increase as more sensors are introduced in to the record. However, the timing of these discontinuities is important; the effect on uncertainty of two temporally close discontinuities may be similar to one discontinuity.

The increase in trend estimate uncertainty when taking into account discontinuities occurs because the statistical model is estimating the magnitude of specific discontinuities. This leads to a greater degree of freedom as the model has extra terms to fit, which will increase with the number of discontinuity terms. These discontinuities still exist even if not specified in the model so studies neglecting to consider these terms will have a perceived, but inaccurate, smaller uncertainty.

4.4 Implications for multi-sensor ocean color records

Work by Brewin et al. (2014) suggests that trends in monthly log-transformed chl, estimated using least squares linear regression, show a similar regional pattern in the MERIS, MODIS-Aqua, and SeaWiFS sensors. Additionally, Mélin et al. (2017) found that these individual records, and VIIRS, show similar trends to the ESA OC-CCI dataset. However, the differences we find here imply that using a space-time model that specifically includes discontinuities and environmental variability (through the SST term) reveals additional information that would otherwise be missed.

We find similar results using both the ESA OC-CCI dataset and GlobColour dataset (full details in supporting information Text S4), i.e. that discontinuities are present in most regions and impact trend estimates. More specifically, the discontinuity magnitudes, trend differences, and trend uncertainty differences show a near 1:1 relationship between the two datasets. However, discontinuity magnitudes are on average slightly larger in the GlobColour dataset, and although this has a subtle effect on trend differences, the uncertainty differences in the GlobColour dataset are also larger on average. This result may suggest a slightly larger bias in the GlobColour dataset due to the different approaches used for merging satellite records. The ESA OC-CCI dataset has been corrected for bias (Lavender et al., 2015), whilst the GlobColour data are not explicitly bias-corrected but are instead merged by inversion with a bio-optical model (Maritorena et al., 2010). The larger discontinuities in GlobColour could also be attributed to the higher variance in this dataset (supporting information Table S4), which may impact quantities estimated within the model. Nevertheless, our results are consistent with both datasets used indicating the effect unaccounted discontinuities can have on trend detection.

5 Conclusion

We assess the presence of discontinuities in multi-sensor satellite records and their effect on estimation of chl trends using a Bayesian spatio-temporal method. We estimate discontinuities in our statistical model using a discrete factor, at the times dictated by three major discontinuities in the ocean color record corresponding to the introduction of the MERIS and MODIS-Aqua sensors in 2002, the loss of the SeaWiFS sensor at the end of 2010, and the introduction of the VIIRS sensor in 2012.

When modeling all three discontinuities, we find their effect in 16 of 23 regions. These discontinuities lead to a corresponding difference in trend estimates in 14 regions with a maximum difference of 2.9% yr^{-1}, which can even change the direction of trend. The effect on trend estimate uncertainty is dependent on the number of discontinuities taken in to account. If we model just one of the above discontinuities, there is a ~0.20%yr^{-1} increase in uncertainty. If we model two discontinuities, i.e. MERIS/MODIS & SeaWiFS or MERIS/MODIS & VIIRS, the uncertainty rises by at least 0.064% yr^{-1} and by up to 1.5% yr^{-1}, dependent on the region. Modeling all three discontinuities produces similar results to modeling the two discontinuities as listed above.

The bias in trend estimates and increase in their uncertainty when taking into account discontinuities challenges the detection of long-term trends in multi-sensor records and stresses the importance of using the best techniques to remove inter-sensor biases when creating these records. Such techniques may include advanced statistical methods, potentially including the use of spatio-temporal models, as well as launching missions with sufficient overlap in order to most effectively cross-calibrate and merge records.

Acknowledgements

The authors are grateful to the ESA for providing the OC-CCI dataset, ACRI-ST for providing the GlobColour dataset and NOAA for providing the Optimum Interpolation SST dataset used here. The data can be found at the following respective URLs: http://www.esa-oceancolour-cci.org/, http://globcolour.info, and https://www.esrl.noaa.gov. The code is made publicly available at: https://github.com/oceanstats/Discontinuities. M.L.H. was partially funded by a University of Southampton Vice Chancellor’s Studentship Award. C.B. was supported by a Marie Curie FP7-Reintegration-Grants within the 7th European Community Framework (project 631466 – *TROPHYZ*).

References

Bakar, K. S., and S. K. Sahu (2015), sp Timer: Spatio-Temporal Bayesian Modeling Using R, J Stat Softw, 63(15), 1-32.

Beaulieu, C., S. A. Henson, J. L. Sarmiento, J. P. Dunne, S. C. Doney, R. R. Rykaczewski, and L. Bopp (2013), Factors challenging our ability to detect long-term trends in ocean chlorophyll, Biogeosciences, 10(4), 2711-2724, doi: 10.5194/bg-10-2711-2013.

Boyce, D. G., and B. Worm (2015), Patterns and ecological implications of historical marine phytoplankton change, Mar Ecol Prog Ser, 534, 251-272, doi: 10.3354/meps11411.

Brewin, R. J. W., F. Mélin, S. Sathyendranath, F. Steinmetz, A. Chuprin, and M. Grant (2014), On the temporal consistency of chlorophyll products derived from three ocean-colour sensors, Isprs J Photogramm, 97, 171-184, doi: 10.1016/j.isprsjprs.2014.08.013.

Campbell, J. W. (1995), The lognormal distribution as a model for bio-optical variability in the sea, Journal of Geophysical Research: Oceans, 100(C7), 13237-13254, doi: 10.1029/95JC00458.

Chuprin, A., T. Jackson, M. Grant, and M. Zühlke (2017), System Prototype Specification (3.1.0). Ocean Colour Climate Change Initiative (OC_CCI) – Phase Two, Plymouth Marine Laboratory. Retrieved from http://www.esa-oceancolour-cci.org/?q=webfm_send/704

Djavidnia, S., F. Mélin, and N. Hoepffner (2010), Comparison of global ocean colour data records, Ocean Sci, 6(1), 61-76.

Gelfand, A. E., and A. F. M. Smith (1990), Sampling-Based Approaches to Calculating Marginal Densities, Journal of the American Statistical Association, 85(410), 398-409, doi: 10.1080/01621459.1990.10476213.

Geweke, J. (1992), Evaluating the accuracy of sampling-based approaches to calculating posterior moments. In *Bayesian Statistics 4* (ed JM Bernado, JO Berger, AP Dawid and AFM Smith). Clarendon Press, Oxford, UK.

Grant, M., T. Jackson, A. Chuprin, S. Sathyendranath, M. Zühlke, J. Dingle, T. Storm, M. Boettcher, and N. Fomferra (2017), Product User Guide (3.1.0). Ocean Colour Climate Change Initiative (OC_CCI) – Phase Two, Plymouth Marine Laboratory. Retrieved from http://www.esa-oceancolour-cci.org/?q=webfm_send/684

Gregg, W. W., and N. W. Casey (2010), Improving the consistency of ocean color data: A step toward climate data records, Geophys Res Lett, 37, doi: Artn L04605 10.1029/2009gl041893.

Hammond, M. L., C. Beaulieu, S. K. Sahu, and S. A. Henson (2017), Assessing trends and uncertainties in satellite-era ocean chlorophyll using space-time modeling, Global Biogeochem Cy, 31(7), 1103-1117, doi: 10.1002/2016gb005600.

Handcock, M. S., and M. L. Stein (1993), A Bayesian-Analysis of Kriging, Technometrics, 35(4), 403-410, doi: Doi 10.2307/1270273.

Handcock, M. S., and J. R. Wallis (1994), An Approach to Statistical Spatial-Temporal Modeling of Meteorological Fields, Journal of the American Statistical Association, 89(426), 368-378, doi: Doi 10.2307/2290832.

Henson, S. A., C. Beaulieu, and R. Lampitt (2016), Observing climate change trends in ocean biogeochemistry: when and where, Global Change Biol, 22(4), 1561-1571, doi: 10.1111/gcb.13152.

Henson, S. A., J. L. Sarmiento, J. P. Dunne, L. Bopp, I. Lima, S. C. Doney, J. John, and C. Beaulieu (2010), Detection of anthropogenic climate change in satellite records of ocean chlorophyll and productivity, Biogeosciences, 7(2), 621-640.

Kruschke, J. R. (2015), Doing Bayesian Data Analysis, 2nd ed., pp. 15–32, A Tutorial with R, JAGS, and Stan. Academic Press/Elsevier, Boston, Mass, isbn:9780124058880.

Lavender, S., T. Jackson, and S. Sathyendranath (2015), The Ocean Colour Climate Change Initiative, Ocean Challenge, 21(1), 3.

Levine, A. F. Z., and M. J. McPhaden (2016), How the July 2014 easterly wind burst gave the 2015-2016 El Nino a head start, Geophys Res Lett, 43(12), 6503-6510, doi: 10.1002/2016gl069204.Longhurst, A. (1995), Seasonal cycles of pelagic production and consumption, Prog Oceanogr, 36(2), 77-167, doi: Doi 10.1016/0079-6611(95)00015-1.

Longhurst, A. (1998), Ecological Geography of the Sea, 398 pp., Academic Press, San Diego.

Maritorena, S., and D. A. Siegel (2005), Consistent merging of satellite ocean color data sets using a bio-optical model, Remote Sens Environ, 94(4), 429-440, doi: 10.1016/j.rse.2004.08.014.

Maritorena, S., D. A. Siegel, and A. R. Peterson (2002), Optimization of a semianalytical ocean color model for global-scale applications, Appl Optics, 41(15), 2705-2714, doi: Doi 10.1364/Ao.41.002705.

Maritorena, S., O. H. F. d'Andon, A. Mangin, and D. A. Siegel (2010), Merged satellite ocean color data products using a bio-optical model: Characteristics, benefits and issues, Remote Sens Environ, 114(8), 1791-1804, doi: 10.1016/j.rse.2010.04.002.

Mélin, F. (2016), Impact of inter-mission differences and drifts on chlorophyll-a trend estimates, Int J Remote Sens, 37(10), 2233-2251, doi: 10.1080/01431161.2016.1168949.

Mélin, F., and G. Sclep (2012), Band Shift Correction (1.0). Ocean Colour Climate Change Initiative (OC_CCI) – Phase One, Plymouth Marine Laboratory. Retrieved From http://www.esa-oceancolour-cci.org/?q=webfm_send/226

Mélin, F., and G. Sclep (2015), Band shifting for ocean color multi-spectral reflectance data, Opt Express, 23(3), 2262-2279, doi: 10.1364/Oe.23.002262.

Mélin, F., Vantrepotte, V., Chuprin, A., Grant, M., Jackson, T., & Sathyendranath, S. (2017). Assessing the fitness-for-purpose of satellite multi-mission ocean color climate data records: A protocol applied to OC-CCI chlorophyll-a data. Remote Sensing of Environment, 203, 139–151. https://doi.org/10.1016/j.rse.2017.03.039

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

Reynolds, R. W., N. A. Rayner, T. M. Smith, D. C. Stokes, and W. Q. Wang (2002), An improved in situ and satellite SST analysis for climate, J Climate, 15(13), 1609-1625, doi: Doi 10.1175/1520-0442(2002)015<1609:Aiisas>2.0.Co;2.

Saulquin, B., R. Fablet, A. Mangin, G. Mercier, D. Antoine, and O. Fanton d'Andon (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, Journal of Geophysical Research: Oceans, 118(8), 3752-3763, doi: 10.1002/jgrc.20264.

Vantrepotte, V., and F. Mélin (2011), Inter-annual variations in the SeaWiFS global chlorophyll a concentration (1997–2007), Deep Sea Research Part I: Oceanographic Research Papers, 58(4), 429-441.

Weatherhead, E. C., et al. (1998), Factors affecting the detection of trends: Statistical considerations and applications to environmental data, J Geophys Res-Atmos, 103(D14), 17149-17161, doi: Doi 10.1029/98jd00995.

Wilks, D. (2016). "The stippling shows statistically significant grid points" How Research Results are Routinely Overstated and Overinterpreted, and What to Do about It. Bulletin of the American Meteorological Society, 97(12), 2263

Wolter, K., and M. S. Timlin (1998), Measuring the strength of ENSO events: How does 1997/98 rank?, Weather, 53(9), 315-324, doi: 10.1002/j.1477-8696.1998.tb06408.x.

Figures

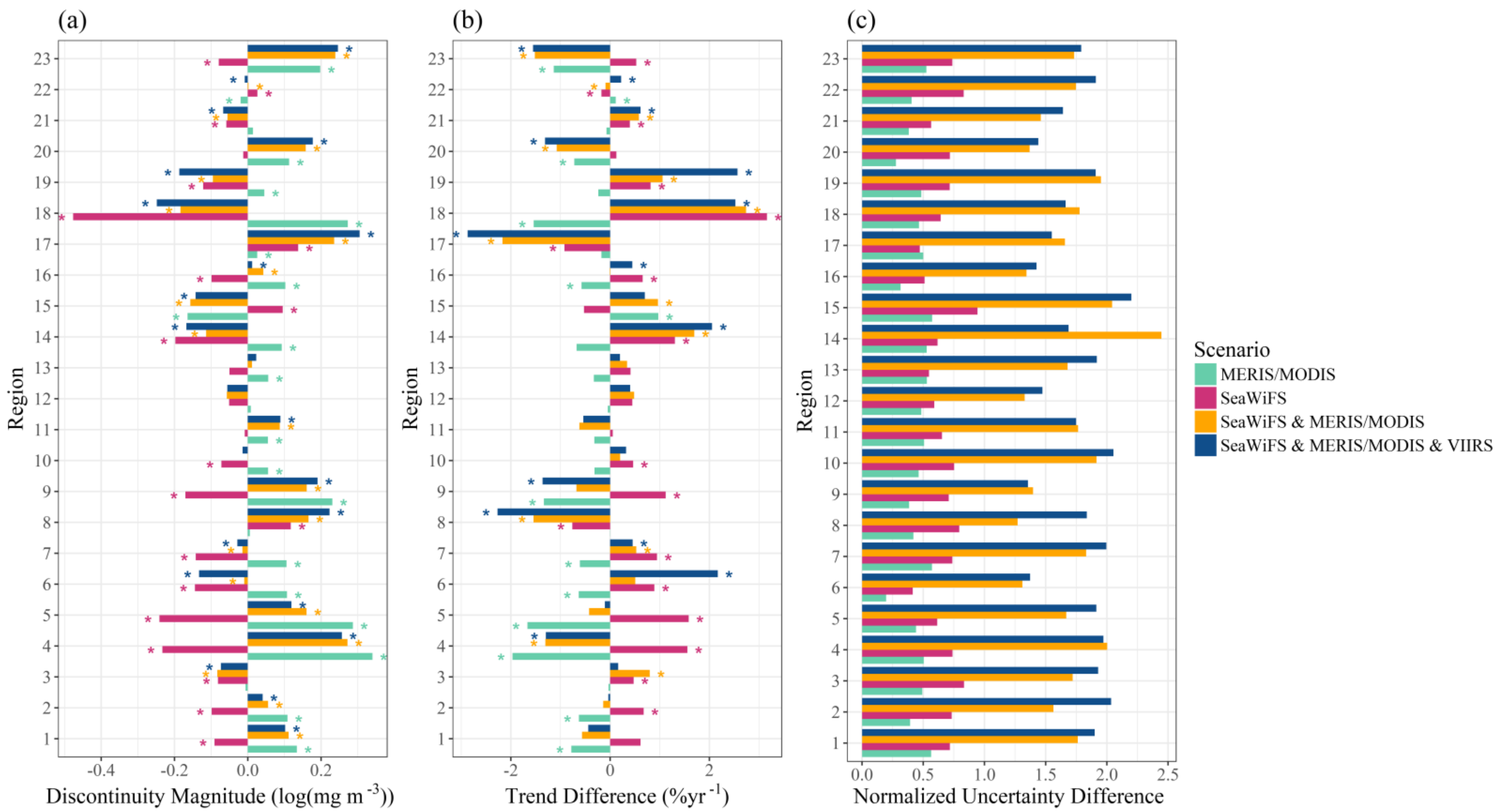

Figure 1. (a) Discontinuity magnitude for each region (averaged for the multiple discontinuity scenarios) as well as the differences in the (b) trend modal posterior density and (c) trend uncertainty (normalized to each region's trend uncertainty) between the models considering a discontinuity and the model with no discontinuity. For (a), * indicates that at least one discontinuity is different from zero, i.e. their 95 % credible intervals do not contain zero. For (b), * indicates regions where trends are different from the model with no discontinuity, i.e. their 95 % credible intervals do not overlap. The uncertainty is defined as the width of the 95 % credible intervals. The scenarios are abbreviated in the main text as follows: N-scenario, (no discontinuity scenario), M-scenario (MERIS/MODIS scenario), S-scenario (SeaWiFS scenario), MS-scenario

(SeaWiFS & MERIS/MODIS scenario), and MSV-scenario (SeaWiFS & MERIS/MODIS & VIIRS scenario). Region names are as follows: (1) Eastern Tropical Atlantic Province, (2) Indian Monsoon Gyres Province, (3) Indian South Subtropical Gyre Province, (4) North Atlantic Tropical Gyral Province, (5) North Pacific Equatorial Countercurrent Province, (6) North Pacific Tropical Gyre Province, (7) Pacific Equatorial Divergence Province, (8) South Atlantic Gyral Province, (9) West Pacific Warm Pool Province, (10) Western Tropical Atlantic Province, (11) Gulf Stream Province, (12) Kuroshio Current Province, (13) North Atlantic Drift Province, (14) North Atlantic Subtropical Gyral Province (East), (15) North Atlantic Subtropical Gyral Province (West), (16) North Pacific Polar Front Province, (17) North Pacific Subtropical Gyre Province (West), (18) Pacific Subarctic Gyres Province (East), (19) Pacific Subarctic Gyres Province (West), (20) South Pacific Subtropical Gyre Province, (21) South Subtropical Convergence Province, (22) Subantarctic Province, and (23) Tasman Sea Province. See Figure 3 for a map of the regions.

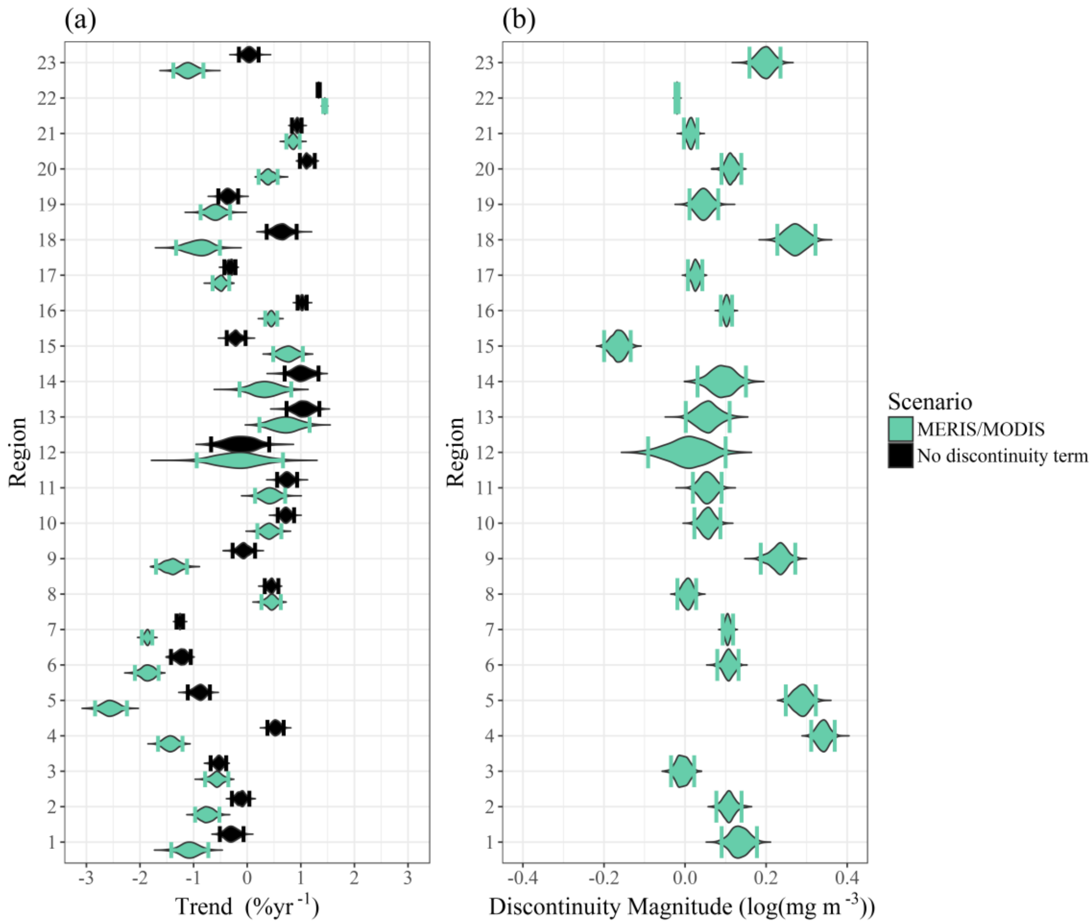

Figure 2. Posterior probability density of (a) the trend in the MERIS/MODIS discontinuity scenario and the no discontinuity scenario, and (b) the discontinuity magnitude in the MERIS/MODIS discontinuity scenario, for each region. We consider the trends, estimated for the two scenarios to be statistically different if their 95 % credible intervals do not overlap. Note the increase in uncertainty when considering discontinuities and the inverse relationship between the discontinuity magnitude and the trend difference. Corresponding figures for the other scenarios can be found in the supporting information (Figures S6 – S9). Regions are plotted in Figure 3 and their names are listed in the caption for Figure 1.

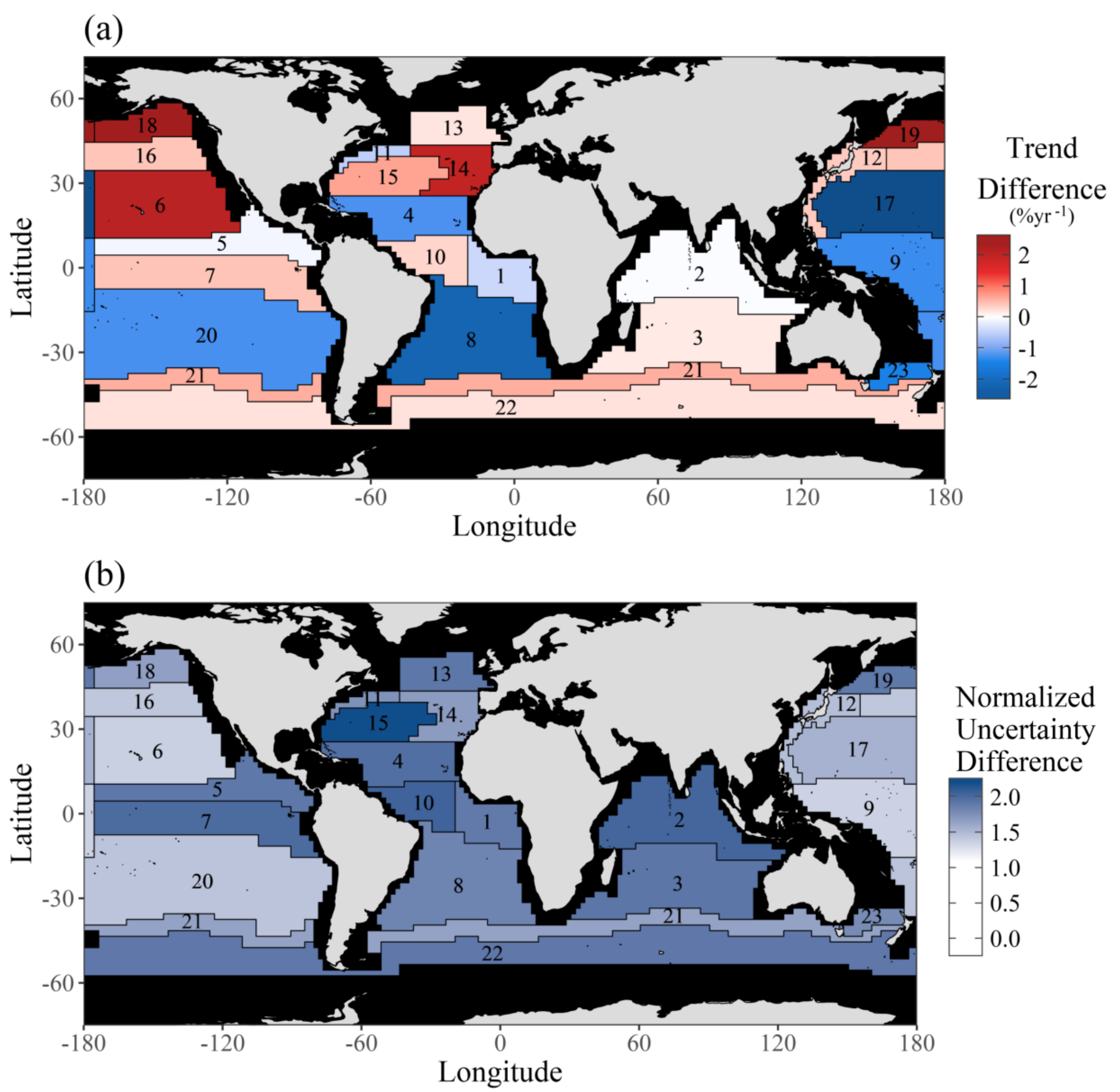

Figure 3. Regional differences in (a) estimated trend and (b) associated uncertainty (normalized to each region's uncertainty), comparing the scenario with all discontinuities and the scenario with no discontinuities. Region names are listed in the caption for Figure 1. See Figure S5 for the trend estimates from the scenario with no discontinuity.