AGU100 ADVANCING EARTH AND SPACE SCIENCE

Geophysical Research Letters

RESEARCH LETTER
10.1029/2017GL076928

Key Points:
- Discontinuities in multisensor ocean color chlorophyll records are detected in ~70% of regions using a Bayesian space-time model
- Discontinuities affect trend estimates in ~60% of regions and can even bias the trends' sign (opposite sign in ~13% of regions)
- The uncertainty of trend estimates increases by an average of 0.20%/year for a single discontinuity and 0.59%/year for two discontinuities

Supporting Information:
- Supporting Information S1

Correspondence to:
M. L. Hammond,
matthew.hammond@noc.ac.uk

Citation:
Hammond, M. L., Beaulieu, C., Henson, S. A., & Sahu, S. K. (2018). Assessing the presence of discontinuities in the ocean color satellite record and their effects on chlorophyll trends and their uncertainties. *Geophysical Research Letters*, *45*, 7654–7662. https://doi.org/10.1029/2017GL076928

Received 22 DEC 2017
Accepted 4 JUL 2018
Accepted article online 11 JUL 2018
Published online 3 AUG 2018

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], and Sujit K. Sahu[4]

[1]Ocean and Earth Science, University of Southampton, Southampton, UK, [2]National Oceanography Centre, Southampton, UK, [3]Ocean Sciences Department, University of California, Santa Cruz, CA, USA, [4]Mathematical Sciences, University of Southampton, Southampton, UK

Abstract Ocean color sensors are crucial for understanding global phytoplankton dynamics. However, the limited life spans of sensors make multisensor data sets 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 spatiotemporal model to investigate the presence of discontinuities and their impacts on estimated chlorophyll trends. The discontinuities considered are the introduction of Medium Resolution Imaging Spectrometer, Moderate Resolution Imaging Spectroradiometer-Aqua, and Visible Infrared Imaging Radiometer Suite and the termination of Sea-Viewing Wide Field-of-View Sensor. 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%/year (0.59%/year for two discontinuities). This difference in trend magnitude and uncertainty highlights the importance of minimizing discontinuities in multisensor records and taking into account discontinuities when analyzing trends.

Plain Language Summary Ocean color sensors are crucial for understanding global phytoplankton dynamics as they detect chlorophyll, a proxy for phytoplankton abundance. However, the limited life spans of sensors make multisensor data sets necessary for estimating long-term trends. Differences between sensors may be seen when combining them to form multisensor data sets, creating discontinuities in the record (a point in the record representing a sudden change) which could potentially affect estimates of trend and their uncertainties. We use a statistical model that explicitly considers the spatial relationship between observation locations to investigate the presence of discontinuities and their impacts on estimated chlorophyll trends. We consider three discontinuities relating to the introduction and termination of important ocean color sensors. Discontinuities are detected in the majority of regions, affecting trend estimates and potentially even changing the direction of trend. Considering a single discontinuity increases trend uncertainty, and considering multiple discontinuities increases this further. This difference in trend magnitude and uncertainty highlights the importance of minimizing discontinuities in multisensor data sets and taking into account discontinuities when analyzing trends.

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, 2016; Saulquin et al., 2013). A comparison of observational, that is, in situ and satellite, chl observations found that shorter data sets 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, multisensor data sets 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; January 2012 to present). The approach used to combine satellite records must fully compensate for the differences between the individual data sets, which can vary temporally and spatially (Djavidnia et al., 2010). If the differences between data sets are not accounted for, discontinuities may be introduced and 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, that is, a mean-shift discontinuity (Weatherhead et al., 1998), which is considered here. Even with the use of multisensor 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 et al. (1998). More specifically, we use a Bayesian spatiotemporal 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 the European Space Agency's Ocean Colour - Climate Change Initiative Project (ESA OC-CCI) (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 data set 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 data set (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 process (Maritorena et al., 2010; Maritorena & Siegel, 2005). The Garver, Siegel, Maritorena Model process combines sensor observations of water-leaving radiance to form a multisource spectrum for each pixel. The multisource spectrum is then inverted with a semianalytical 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, sea surface temperature (SST) is used as a covariate. SST data are sourced from the National Oceanic and Atmospheric Administration 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 spatiotemporal model is fitted separately in each of the 23 Longhurst regions retained for analysis (i.e., we use an unpooled 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

$$Z_{n,t} = O_{n,t} + \varepsilon_{n,t} \tag{1}$$

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):

$$O_{n,t} = \boldsymbol{x}_{n,t}^{'}\boldsymbol{\beta} + \boldsymbol{a}_{n}^{'}\boldsymbol{w}_{m,t} \tag{2}$$

This regression model is composed of the covariates (including intercept) $\boldsymbol{x}_{n,t}$, the regression coefficients (constant for each region) $\boldsymbol{\beta} = (\beta_0, \beta_{\text{Trend}}, \beta_{\text{SST}}, \beta_{\text{Disc}}, \beta_{M1}, \ldots, \beta_{M12})$, 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, …, n_N) and the knot (m_1, m_2, …, 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 spatiotemporal 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 β_{M1}, …, β_{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 $\boldsymbol{x}^{\text{Disc}}$ 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 et al., 1998):

$$\boldsymbol{x}_{t}^{\text{Disc}} = \begin{cases} 0, & t < t_{\text{Disc}} \\ 1, & t \geq t_{\text{Disc}} \end{cases} \tag{3}$$

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 multidiscontinuity scenarios (MS and MSV), the regression coefficient β_{Disc}

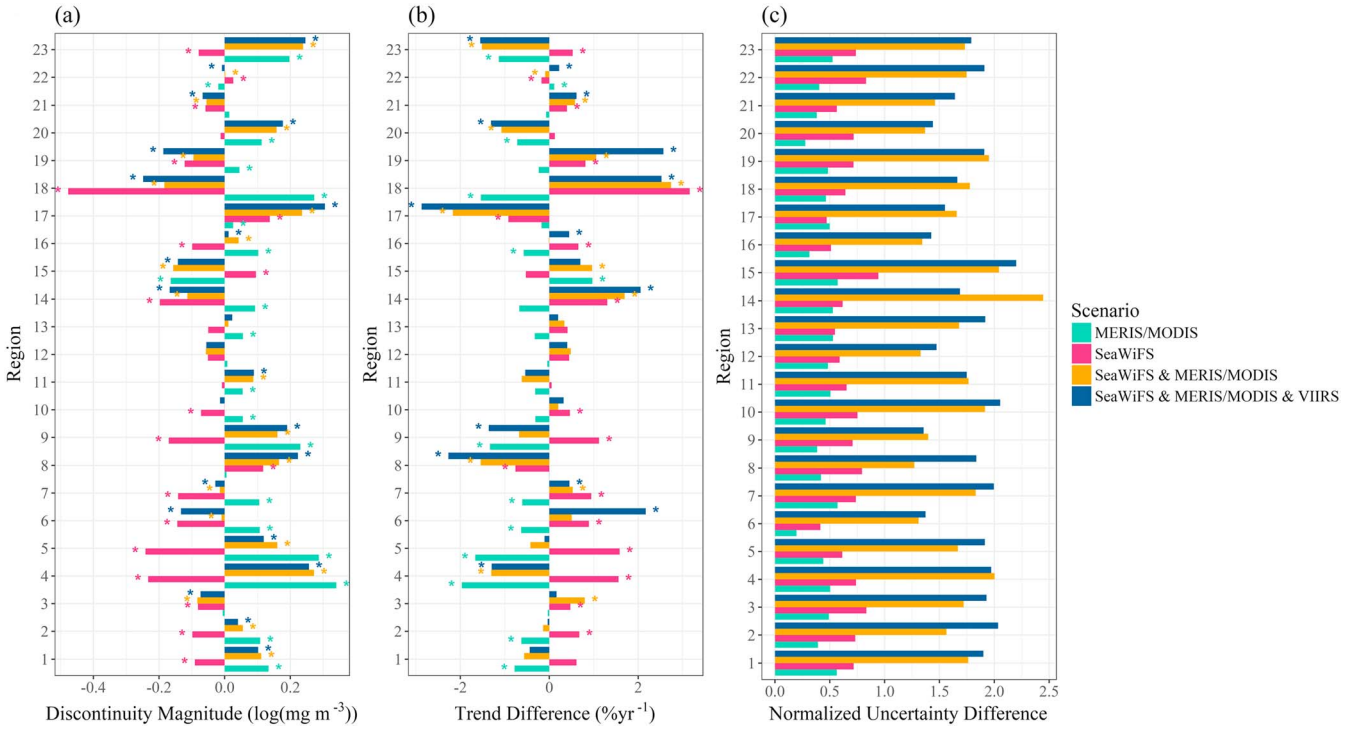

Figure 1. (a) Discontinuity magnitude for each region (averaged for the multiple discontinuity scenarios) as well as the differences in the (b) trend modal posterior density and (c) trend uncertainty (normalized to each region's trend uncertainty) between the models considering a discontinuity and the model with no discontinuity. For (a), * indicates that at least one discontinuity is different from 0; that is, their 95% credible intervals do not contain 0. For (b), * indicates regions where trends are different from the model with no discontinuity; that is, 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 and MERIS/MODIS scenario; and MSV scenario = SeaWiFS and MERIS/MODIS and 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. MERIS = Medium Resolution Imaging Spectrometer; MODIS = Moderate Resolution Imaging Spectroradiometer; SeaWiFS = Sea-Viewing Wide Field-of-View Sensor; VIIRS = Visible Infrared Imaging Radiometer Suite.

includes additional t_{Disc} and $\boldsymbol{x}^{\text{Disc}}$ terms to estimate all discontinuities (i.e., two t_{Disc} and $\boldsymbol{x}^{\text{Disc}}$ 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 0 (i.e., its 95% credible interval excludes 0). 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 data set; 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

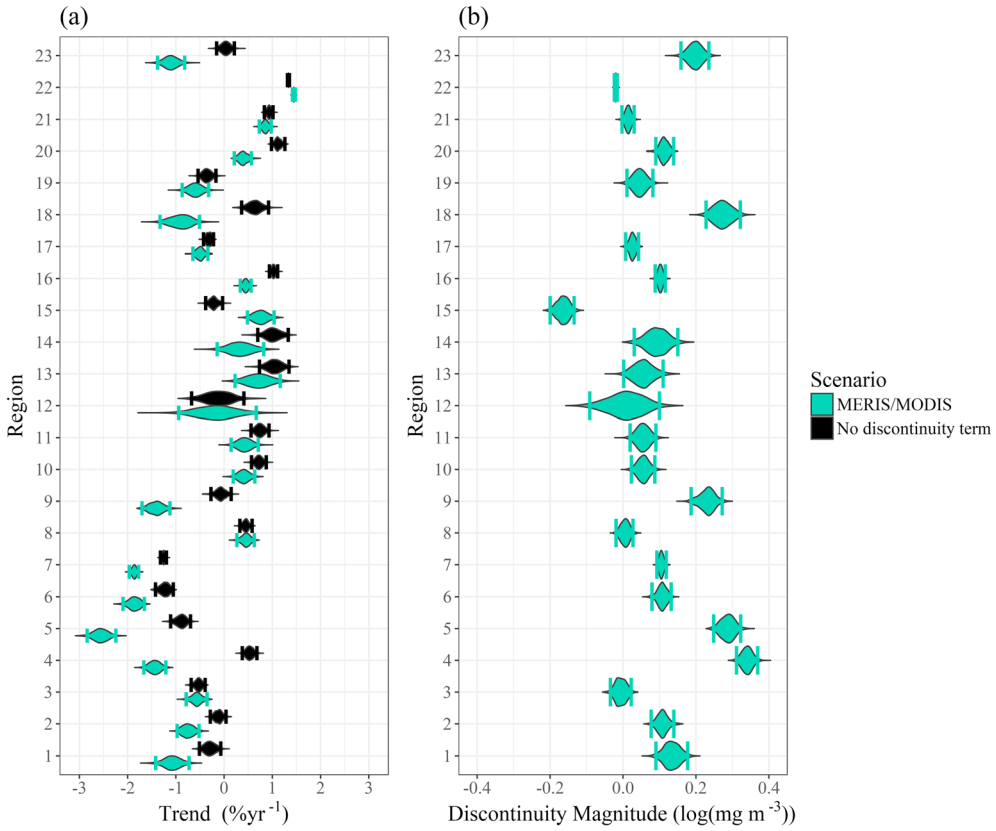

Figure 2. Posterior probability density of (a) the trend in the MERIS/MODIS discontinuity scenario and the no discontinuity scenario, and (b) the discontinuity magnitude in the MERIS/MODIS discontinuity scenario, for each region. We consider the trends, estimated for the two scenarios to be statistically 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 of Figure 1. MERIS = Medium Resolution Imaging Spectrometer; MODIS = Moderate Resolution Imaging Spectroradiometer.

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 multidiscontinuity 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%/year 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%/year; Figure 2). The opposite is found for the S scenario (average of 0.59%/year). For the MS scenario, the sign of the difference is evenly distributed between positive and negative (average difference −0.028%/year). 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 that they are partially canceling out (Figure 1b). The MSV scenario shows similar results to the MS scenario with an average difference of 0.047%/year (Figures 1b and 3a).

The average magnitude of trend differences (i.e., when the direction/sign of trend difference is omitted) is larger in the multidiscontinuity scenarios (1.1%/year for the MSV scenario and 0.85%/year for the MS scenario) than the single-discontinuity scenarios (0.65%/year for the M scenario and 0.81%/year for the S

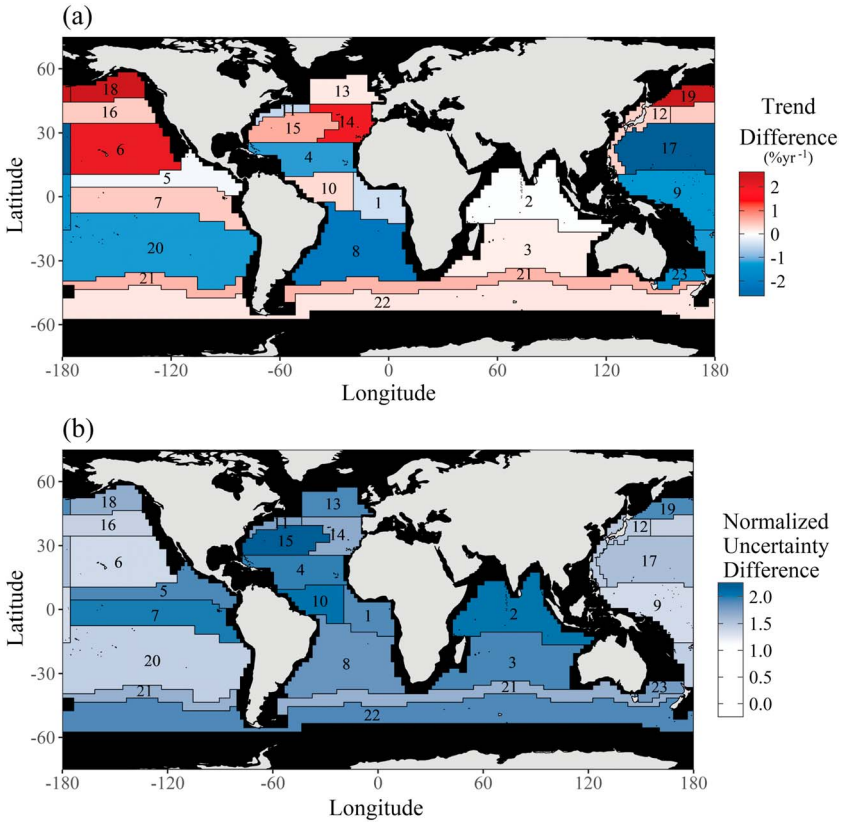

Figure 3. Regional differences in (a) estimated trend and (b) associated uncertainty (normalized to each region's uncertainty), comparing the scenario with all discontinuities and the scenario with no discontinuities. Region names are listed in the caption of Figure 1. See Figure S5 for the trend estimates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

scenario). This can lead to a change of trend sign, that is, from increasing to decreasing or vice versa; for example, this occurs in five 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 data set that is found to show similar results, albeit with a slightly higher average trend difference in most scenarios (Tables S2 and 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%/year (Figure 1c). For the MS scenario and MSV scenario the increase in uncertainty is 0.64%/year (Figures 1c and 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 midlatitude Pacific Ocean (average of 140%). See supporting information for full results, including analysis using the GlobColour data set, 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 data sets, of the same length as the present study, based on realistic values of chl and its variability (with independent randomly generated noise in each data set),

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%/year and varies in the range ± 2.8%/year, resulting in statistically different trends in 14 of the 23 regions. In a study that analyzed the effect of intersensor 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%/year) in the MSV scenario relative to the N scenario. The differences compared to Mélin (2016) are likely due to the substantial differences in the data sets and methodologies. Here we take into account discontinuities in a bias-corrected multisensor data set using a spatiotemporal 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 that 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 aging, particularly toward 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 data set 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 aging 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 multisensor 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 toward 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 multisensor records and will only increase as more sensors are introduced into 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 Multisensor 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 data set. 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 data set and GlobColour data set (full details in supporting information Text S4), that is, 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 data sets. However, discontinuity magnitudes are on average slightly larger in the GlobColour data set, and although this has a subtle effect on trend differences, the uncertainty differences in the GlobColour data set are also larger on average. This result may suggest a slightly larger bias in the GlobColour data set due to the different approaches used for merging satellite records. The ESA OC-CCI data set has been corrected for bias (Lavender et al., 2015), while 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 data set (supporting information Table S4), which may impact quantities estimated within the model. Nevertheless, our results are consistent with both data sets used indicating the effect unaccounted discontinuities can have on trend detection.

5. Conclusion

We assess the presence of discontinuities in multisensor satellite records and their effect on estimation of chl trends using a Bayesian spatiotemporal 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%/year, which can even change the direction of trend. The effect on trend estimate uncertainty is dependent on the number of discontinuities taken into account. If we model just one of the above discontinuities, there is a ~0.20%/year increase in uncertainty. If we model two discontinuities, that is, MERIS/MODIS and SeaWiFS or MERIS/MODIS and VIIRS, the uncertainty rises by at least 0.064%/year and by up to 1.5%/year, 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 multisensor records and stresses the importance of using the best techniques to remove intersensor biases when creating these records. Such techniques may include advanced statistical methods, potentially including the use of spatiotemporal models, as well as launching missions with sufficient overlap in order to most effectively cross calibrate and merge records.

Acknowledgments
The authors are grateful to the ESA for providing the OC-CCI data set, ACRI-ST for providing the GlobColour data set, and NOAA for providing the Optimum Interpolation SST data set 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., & Sahu, S. K. (2015). sp Timer: Spatio-temporal Bayesian modeling using R. *Journal of Statistical Software*, *63*(15), 1–32.

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

Boyce, D. G., & Worm, B. (2015). Patterns and ecological implications of historical marine phytoplankton change. *Marine Ecology Progress Series*, *534*, 251–272. https://doi.org/10.3354/meps11411

Brewin, R. J. W., Mélin, F., Sathyendranath, S., Steinmetz, F., Chuprin, A., & Grant, M. (2014). On the temporal consistency of chlorophyll products derived from three ocean-colour sensors. *ISPRS Journal of Photogrammetry and Remote Sensing*, *97*, 171–184. https://doi.org/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*, *100*(C7), 13,237–13,254. https://doi.org/10.1029/95JC00458

Chuprin, A., Jackson, T., Grant, M., & Zühlke, M. (2017). System prototype specification (3.1.0). Ocean colour climate change initiative (OC_CCI)—Phase two. *Plymouth Marine Laboratory* (Vol. 1, 37 pp.). Retrieved from http://www.esa-oceancolour-cci.org/?q=webfm_send/704

Djavidnia, S., Mélin, F., & Hoepffner, N. (2010). Comparison of global ocean colour data records. *Ocean Science*, *6*(1), 61–76. https://doi.org/10.5194/os-6-61-2010

Gelfand, A. E., & Smith, A. F. M. (1990). Sampling-based approaches to calculating marginal densities. *Journal of the American Statistical Association*, *85*(410), 398–409. https://doi.org/10.1080/01621459.1990.10476213

Geweke, J. (1992). Evaluating the accuracy of sampling-based approaches to calculating posterior moments. In J. M. Bernado, J. O. Berger, A. P. Dawid, & A. F. M. Smith (Eds.), *Bayesian statistics* (Vol. 4, pp. 169–193). Oxford, UK: Clarendon Press.

Grant, M., Jackson, T., Chuprin, A., Sathyendranath, S., Zühlke, M., Dingle, J., et al. (2017). Product user guide (3.1.0). Ocean colour climate change initiative (OC_CCI)—Phase two (Vol. 1, 47 pp.). *Plymouth Marine Laboratory*. Retrieved from http://www.esa-oceancolour-cci.org/?q=webfm_send/684

Gregg, W. W., & Casey, N. W. (2010). Improving the consistency of ocean color data: A step toward climate data records. *Geophysical Research Letters*, *37*, L04605. https://doi.org/10.1029/2009gl041893

Hammond, M. L., Beaulieu, C., Sahu, S. K., & Henson, S. A. (2017). Assessing trends and uncertainties in satellite-era ocean chlorophyll using space-time modeling. *Global Biogeochemical Cycles*, *31*, 1103–1117. https://doi.org/10.1002/2016gb005600

Handcock, M. S., & Stein, M. L. (1993). A Bayesian-analysis of kriging. *Technometrics*, *35*(4), 403–410. https://doi.org/10.2307/1270273

Handcock, M. S., & Wallis, J. R. (1994). An approach to statistical spatial-temporal modeling of meteorological fields. *Journal of the American Statistical Association*, *89*(426), 368–378. https://doi.org/10.2307/2290832

Henson, S. A., Beaulieu, C., & Lampitt, R. (2016). Observing climate change trends in ocean biogeochemistry: When and where. *Global Change Biology*, *22*(4), 1561–1571. https://doi.org/10.1111/gcb.13152

Henson, S. A., Sarmiento, J. L., Dunne, J. P., Bopp, L., Lima, I., Doney, S. C., et al. (2010). Detection of anthropogenic climate change in satellite records of ocean chlorophyll and productivity. *Biogeosciences*, *7*(2), 621–640. https://doi.org/10.5194/bg-7-621-2010

Kruschke, J. R. (2015). Doing Bayesian data analysis. In *A tutorial with R, JAGS, and Stan*, (2nd ed.pp. 15–32). Boston, mass: Academic press/Elsevier. isbn:9780124058880

Lavender, S., Jackson, T., & Sathyendranath, S. (2015). The ocean colour climate change initiative. *Ocean Challenge*, *21*(1), 3.

Levine, A. F. Z., & McPhaden, M. J. (2016). How the July 2014 easterly wind burst gave the 2015–2016 El Nino a head start. *Geophysical Research Letters*, *43*, 6503–6510. https://doi.org/10.1002/2016gl069204

Longhurst, A. (1995). Seasonal cycles of pelagic production and consumption. *Progress in Oceanography*, *36*(2), 77–167. https://doi.org/10.1016/0079-6611(95)00015-1

Longhurst, A. (1998). *Ecological geography of the sea*, (p. 398). San Diego: Academic Press.

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

Maritorena, S., & Siegel, D. A. (2005). Consistent merging of satellite ocean color data sets using a bio-optical model. *Remote Sensing of Environment*, *94*(4), 429–440. https://doi.org/10.1016/j.rse.2004.08.014

Maritorena, S., Siegel, D. A., & Peterson, A. R. (2002). Optimization of a semianalytical ocean color model for global-scale applications. *Applied Optics*, *41*(15), 2705–2714. https://doi.org/10.1364/Ao.41.002705

McClain, C. (2009). A decade of satellite ocean color observations. *Annual Review of Marine Science*, *1*(1), 19–42. https://doi.org/10.1146/annurev.marine.010908.163650

Mélin, F. (2016). Impact of inter-mission differences and drifts on chlorophyll-*a* trend estimates. *International Journal of Remote Sensing*, *37*(10), 2233–2251. https://doi.org/10.1080/01431161.2016.1168949

Mélin, F., & Sclep, G. (2012). Band shift correction (1.0). Ocean colour climate change initiative (OC_CCI)—Phase one (Vol 1, pp. 8–22). *Plymouth Marine Laboratory*. Retrieved From http://www.esa-oceancolour-cci.org/?q=webfm_send/226

Mélin, F., & Sclep, G. (2015). Band shifting for ocean color multi-spectral reflectance data. *Optics Express*, *23*(3), 2262–2279. https://doi.org/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

Reynolds, R. W., Rayner, N. A., Smith, T. M., Stokes, D. C., & Wang, W. Q. (2002). An improved in situ and satellite SST analysis for climate. *Journal of Climate*, *15*(13), 1609–1625. https://doi.org/10.1175/1520-0442(2002)015%3C1609:Aiisas%3E2.0.Co;2

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. *Journal of Geophysical Research: Oceans*, *118*, 3752–3763. https://doi.org/10.1002/jgrc.20264

Vantrepotte, V., & Mélin, F. (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., Reinsel, G. C., Tiao, G. C., Meng, X. L., Choi, D., Cheang, W. K., et al. (1998). Factors affecting the detection of trends: Statistical considerations and applications to environmental data. *Journal of Geophysical Research*, *103*(D14), 17,149–17,161. https://doi.org/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), 2263–2273. https://doi.org/10.1175/BAMS-D-15-00267.1

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