Remote sensing reflectance anomalies in the ocean

Yannick Huot\textsuperscript{a,*}, David Antoine\textsuperscript{b,c}

\textsuperscript{a} Centre d'Applications et de Recherches en Télédétection, Département de géomatique appliquée, Université de Sherbrooke, Sherbrooke, Québec J1K 2R1, Canada
\textsuperscript{b} Remote Sensing and Satellite Research Group, Department of Physics, Astronomy and Medical Radiation Sciences, Curtin University, GPO Box U1987, Perth, Western Australia 6845, Australia
\textsuperscript{c} Sorbonne Universités, UPMC Univ Paris 06 and CNRS, Laboratoire d’Océanographie de Villefranche, Villefranche-sur mer 06238, France

\textbf{A R T I C L E  I N F O}

Article history:
Received 12 August 2015
Received in revised form 4 May 2016
Accepted 1 June 2016
Available online xxxx

Keywords:
Remote sensing reflectance, Anomalies, Ocean color, Remote sensing

\textbf{A B S T R A C T}

Small spectral differences from the mean remote sensing reflectance (Rrs) of the ocean – anomalies – can provide unique environmental information from ocean color satellite data. First, we describe the average relationship between three input spectral bands and an output band by developing a look-up table (LUT) based on the fully normalized Rrs from the MODIS AQUA sensor. By dividing the Rrs measured at the output wavelength by the prediction from the LUT, we obtain several anomalies depending on the combination of input and output bands. None of these anomalies are correlated with chlorophyll concentration on the global scale. Some anomalies are strongly correlated with previously described data products (e.g., CDOM index, backscattering coefficients from semi-analytical inversion models), but others are not correlated with any product currently distributed by NASA. In the latter case, new information about oceanic optical properties is extracted from the ocean color spectra, which allows identification of water masses that was otherwise impossible with standard ocean color products. It was not possible, in some cases, to identify the optical source of this information, which may be spatially and temporally variable. We also show that by removing the main source of variability, the anomalies show interesting potential to identify subtle shifts in sensor response in satellite time series.

© 2016 Elsevier Inc. All rights reserved.

1. Introduction

A multitude of numerical inversion approaches have been developed to quantify geophysical quantities from the light spectrum leaving the ocean, in particular as recorded by ocean color satellite instruments. A large array of variables is now available, including phytoplankton chlorophyll (cf. Hu & Campbell, 2014), the absorption of colored dissolved organic matter (CDOM), the particulate backscattering coefficient (Garver & Siegel, 1997; Roesler, Perry, & Carder, 1989), phytoplankton groups and species (Brewin et al., 2011), particulate inorganic (Balch, Gordon, Bowler, Drapeau, & Booth, 2005) and organic carbon (Pabi & Arrigo, 2006; Stramski, Reynolds, Kahru, & Mitchell, 1999), and the particle size distribution (Kostadinov, Siegel, & Maritorena, 2010).

The physics underlying the inversion of the light leaving the ocean is well understood (Gordon, 2002; e.g. Gordon et al., 1988). By normalizing the water-leaving radiance \( (L_w, \text{W m}^{-2} \text{nm}^{-1} \text{sr}^{-1}) \) by the downward irradiance \( (E_d, \text{W m}^{-2} \text{nm}^{-1}) \) just above the surface \( (0^+) \) at each wavelength \( (\lambda, \text{nm}) \), we define the remote sensing reflectance:

\[
R_{rs}(\lambda) = \frac{L_w(\lambda)}{E_d(\lambda,0^+)}
\]

(1)

* Corresponding author.

The \( R_{rs}(\lambda) \) can, in turn, be expressed in terms of inherent optical properties (IOPs); those that depend only on the nature and concentration of a substance and not on the geometrical distribution of the light field (Preisendorfer, 1976). This allows the concentration of substances to be retrieved. Among different approximations (e.g. Gordon et al., 1988; Lee et al., 2011; Morel & Prieur, 1977; Zaneveld, 1995) is

\[
R_{rs} = \frac{\pi f_b a}{Q^2 a^2}.
\]

(2)

where \( b_a \left( \text{m}^{-1} \right) \) and \( a \left( \text{m}^{-1} \right) \) are two IOPs, the total backscattering and absorption coefficients, respectively, \( f \) is a factor that accounts for the reflection and refraction effects at the air–water interface, and \( f/Q \left( \text{sr}^{-1} \right) \) describes the bidirectional nature of reflectance (cf. Morel, Antoine, & Gentili, 2002). The \( f/Q \) factor, which is itself dependent on several variables, accounts for the geometry and conditions of observation in the water. In Eq. (2), as well as throughout this paper, to simplify the notation, we dropped the variable dependence (solar zenith angle, observation angle, IOPs, etc.) for \( f/Q \) as well as the dependence on wavelength. We did the same for \( f \) which depends on the geometry of observation.

When multiple absorbing and scattering substances are present, the total IOP value can be represented as the sum of the contributions of each substance (Preisendorfer, 1976). We can also separate the
known contribution of water from the contributions of all the other variable constituents,

\[ R_{rs} = \mathcal{R}_b b_{bbw} + \sum_i \frac{b_i}{Q} a_{wi} + \sum_j a_j, \]  

where \( b_{bbw} \) and \( a_{wi} \) are respectively the backscattering (Morel, 1974; Werdell, Franz, Leffler, Robinson, & Boss, 2013; Zhang, Hu, & He, 2009) and absorption coefficients of seawater (Lee et al., 2015; Pope & Fry, 1997), and \( b_i \) and \( a_j \) are respectively the backscattering and absorption coefficients for water constituents \( i \) and \( j \). Eq. (3) highlights that the sources of variability in the remotely sensed reflectance originate from the amplitudes and spectral shapes of the inherent optical properties and the \( Q_i \) and \( f_j/Q \) factors. To retrieve IOPs accurately, variability in \( f_j/Q \) as a function of the observation geometry must be accounted for. In open ocean waters, the approach currently taken by NASA is to use pre-computed look-up tables that describe the amplitude and spectral change in this factor as a function of chlorophyll concentration and the geometry of observation (Morel et al., 2002). This approach provides so-called exactly normalized remote sensing reflectance \( R_{rs}^{ex}(\lambda) \), which, if the look-up table used is accurate, provides \( R_{rs}^{ex} \) corrected as if the observations were made at nadir with the Sun at zenith. Thereby, using \( R_{rs}^{ex} \) means that the dependency on the geometry of observation has been removed. This equation does not explicitly show that another source of variability in \( R_{rs}^{ex}(\lambda) \) is the vertical distribution of IOPs (Gordon & Clark, 1980; Piskozub, Neumann, & Woźniak, 2008; Sathyendranath & Platt, 1989; Stramska & Stramski, 2005; Zaneveld, Barnard, & Boss, 2005).

To estimate the value of individual IOPs in the presence of multiple absorption and scattering components, a given IOP must have a distinct spectral shape and a sufficiently large relative contribution to the total absorption or backscattering coefficients. Under these conditions, it can have an observable influence on the remotely measured signal (see discussion in IOCCG (1998)). This impact should, at the very least, lead to differences in \( L \) that are a few times higher than the noise equivalent radiance of the sensor. The satellite sensors of the current and recent generation have between 5 (e.g., VIIRS) and 7 (e.g., MODIS) bands in the visible. Since the number of different IOPs that can be measured from a \( R_{rs}^{ex}(\lambda) \) spectrum is equal at most to the number of bands, this means that the numerous variables currently retrieved from \( R_{rs}^{ex}(\lambda) \) either all have a distinct and strong impact on \( R_{rs}^{ex}(\lambda) \) or are highly correlated to one another. The latter is the case, for example, for certain empirical algorithms retrieving \( K_a^s \) and chlorophyll that use different functions applied to the same band ratios (cf. Morel et al., 2007). If the algorithm can explain a significant fraction of the in situ variance, correlations between the retrieved variables is the only plausible scenario when there are more variables retrieved than the number of bands.

Considering the limited number of bands available on past and current satellites, a reductionist approach to the inversion problem has been mostly followed. Three parameters, that have been chosen because they have the strongest impact on \( R_{rs}^{ex}(\lambda) \), are generally estimated from semi-analytical models: phytoplankton absorption, CDOM (together with non-algal particle absorption) and the amplitude (and sometimes shape) of backscattering (IOCCG, 2006 and references therein). Validation of these approaches (e.g. IOCCG, 2006) has shown that all three parameters can be retrieved well. These semi-empirical algorithms were designed to avoid biases (e.g. Dierssen, 2010; Loisel, Lubac, Desaillly, Dufret-Gaurier, & Vantrepotte, 2010) occurring with empirical band ratios algorithms. They generally cannot, however, account for the variable shape of the inherent optical properties, which appear to be very important in explaining some of the errors in the empirical algorithms (Sauer, Roesler, Werdell, & Barnard, 2012; Szeto, Werdell, Moore, & Campbell, 2011).

This research has generally been driven by the desire to measure a given biogeochemical variable from remote sensing. One such variable is the phytoplankton species composition. Modelers are particularly keen to obtain such data remotely to validate global multispecies models (e.g. Bopp, Aumont, Cadule, Alvein, & Gehlen, 2005). As such, a significant amount of research has attempted to provide this information (see reviews by Brewin et al. (2011); IOCCG (2014); Nair et al. (2008)). An approach developed by Alvein, Moulin, Dandonneau, and Breon (2005) for such retrievals looks for small variations (or anomalies) in \( R_{rs}^{ex}(\lambda) \) by dividing a measured ocean color spectrum at a given pixel by the mean spectrum that would be measured for the same blue to green ratio. This approach has been shown to be particularly sensitive to variations in backscattering and CDOM absorption, and also shows promise for developing empirical algorithms to estimate these variables (Brown, Huot, Werdell, Gentili, & Claustre, 2008). Different approaches looking at such anomalies have also led to methods to find regions that depart from the mean oceanic trend that chlorophyll alone would predict. They include algorithms for quantification of CDOM absorption (Morel & Gentili, 2009), backscattering (Morel & Bélanger, 2006) or quantum yield of Sun-induced fluorescence (Huot, Franz, & Fradette, 2013), and algorithms for identification of water types (Lee & Hu, 2006).

Building on such methods, but without a desired variable as an outcome, we develop an approach to simultaneously account for the most important effects of absorption and backscattering allowing us to extract finer spectral anomalies in \( R_{rs}^{ex}(\lambda) \) and examine if additional information is available. The anomalies are obtained using an estimate of \( R_{rs}^{ex}(\lambda) \) at one wavelength from a combination of three other wavelengths.

2. Methods

2.1. Datasets

We used all available 2007 daily MODIS level 3 mapped \( R_{rs}^{ex}(\lambda) \) data at 412, 443, 488, 531, 547 and 667 nm (processing version 2012.0). From this dataset, approximately 146 million pixels were randomly selected to create look-up tables (LUTs). In addition, selected level 2 scenes were also downloaded for closer examination of computed anomalies.

2.2. Computing the LUTs and anomalies

The LUTs were created by sectioning the three-dimensional space spanned by the \( R_{rs}^{ex}(\lambda) \) data for three independent wavelengths into 27,000 small cubes. This was achieved by using 30 \( R_{rs}^{ex}(\lambda) \) intervals for each wavelength. While the number of intervals was found to have a
very small influence on the results, the number of cubes was selected by trial and error to provide sufficiently large cubes that include enough points to calculate statistically-significant averages, and yet to provide a sufficient number of divisions to capture the variability. A minimum of 40 measurements within a cube was deemed necessary to compute anomalies (this leads to some regions without data in the figures presented in the result section). We then identified from all the randomly selected pixels those which fell within each of these small cubes and computed the average of a dependent (predicted) $R_{\text{hit}}(\lambda)$ at a fourth wavelength in each cube. The LUT hence returns the predicted $R_{\text{hit}}(\lambda)$, referred to as $R_{\text{hit}}(\lambda)$, at an output band for a given combination of three $R_{\text{hit}}(\lambda)$. The predicted, $R_{\text{hit}}(\lambda)_{\lambda_1\lambda_2\lambda_3}$, represents the average oceanic $R_{\text{hit}}(\lambda)$ for a given combination three different input bands ($\lambda_1\ldots\lambda_3$).

To compute our anomalies ($R_{\text{hit}}(\lambda)_{\lambda_1\lambda_2\lambda_3}$, unitless) – in fact, ratios – we divided the measured $R_{\text{hit}}(\lambda)$ at the wavelength we are predicting (the dependent wavelength) by the predicted value ($R_{\text{hit}}(\lambda)_{\lambda_1\lambda_2\lambda_3}$) based on the LUT interpolated for the three input $R_{\text{hit}}(\lambda)$ (i.e., $R_{\text{hit}}(\lambda_1), R_{\text{hit}}(\lambda_2), R_{\text{hit}}(\lambda_3)$).

\[
R_{\lambda_1\lambda_2\lambda_3} = \frac{R_{\text{hit}}(\lambda)}{R_{\text{hit}}(\lambda)_{\lambda_1\lambda_2\lambda_3}} \quad (4)
\]

Table 1

| Values of $R^\prime$ between the anomalies and several ocean color products. To save space, all values have been multiplied by 100 and rounded to the nearest unit; anomaly notation is $\lambda R_{\lambda_1\lambda_2\lambda_3}$ instead of $R_{\text{hit}}(\lambda)_{\lambda_1\lambda_2\lambda_3}$ in the text. To highlight the better relationships, light shading was applied to values of $R^\prime$ between 0.40 and 0.79, and $R^\prime$ between 0.40 and 0.79 while darker shading was applied to values of $R^\prime$ between 0.80 and 1.0. Data used to compute $R^\prime$ were the global data for the month of March 2007 (e.g., Figs. 2 and 3). |

Abbreviations and references are as follow GIOIP (Werdell, Franz, Bailey, et al. (2013a), QAA (Lee et al., 2002), GSM (Maritorena et al., 2002), OCM3 for the variables denoted with these acronyms. References for the other products are POC = particulate organic carbon (Stramski et al., 2008), CDO M index (Morel & Cencini, 2009), bbb anomalies (see text), Particulate inorganic carbon (Balch et al., 2005; Gordon et al., 2001), Zeu lee = euphotic zone depth (Lee et al., 2007), Zeu Morel = euphotic zone depth (Morel et al., 2007), IPAR = instantaneous Ångström exponent, Aer. opt. thick. 869 = aerosol optical thickness at 869 nm. |

where $\lambda$ is the dependent or predicted wavelength at which the anomaly is computed and $\lambda_1, \lambda_2, \lambda_3$ are the three wavelengths used to estimate $R_{\text{hit}}(\lambda)_{\lambda_1\lambda_2\lambda_3}$, Fig. 1 shows an example of a LUT obtained for $R_{\text{hit}}(\lambda)_{443}(412.488.547)$. We chose to work with 3 input bands as these should, in most cases, be able to capture the three main IOPs currently derived from ocean color algorithms: phytoplankton and CDOM absorption coefficients as well as the backscattering coefficient.

We computed monthly composites by summing daily level 3 mapped images and dividing by the number of days when unflagged data were present for each pixel. This method leads to slightly different map products from those distributed by NASA, which are based on level 2 data.

2.3. Computing signed determination coefficients

To examine the possible non-linear relationships between the anomalies and other MODIS AQUA products, we fitted a third-order polynomial to the monthly binned log-transformed data for the whole globe between any two MODIS products or between these products and the anomalies for the month of March 2007. We computed the coefficient of determination ($R^2$) between the fitted and measured data.
To provide information about the direction of the relationship, we also computed the correlation coefficient ($\rho$) between the two log-transformed datasets (not fitted). We constructed a signed coefficient of determination ($\bar{R}^2$), which uses the value of the coefficient of determination from the third-order polynomial fit, and the sign from the $\rho$ on the log transformed data.

### 3. Results and discussion

#### 3.1. General observations

A correlation analysis of $R_{rs}(\lambda)$ shows (see Table 1) that wavelengths below and including 488 nm are well correlated with each other, so are wavelengths above and including 531 nm. However, wavelengths in these two groups show very little correlation with wavelengths in the other group: the absolute value of $\bar{R}^2$ ($\bar{R}^2|$) between $R_{rs}(\lambda)$ at any wavelengths across these two groups is always below 0.2 (see Table 1). By design of the algorithms, the first group is well and negatively correlated with all “absorption products” which are largely derived from similar bands, while the second group is well and positively correlated with the products that are linked to backscattering as they mainly used information from the green bands. As long as the dataset covers most typical oceanic conditions, these results are only very weakly modified by different selection of data for the correlation analyses as they compare the output of algorithms applied on different samples from the same “population” ocean color data. Here, for consistency with data presented below, we used the month of March 2007 which covers most of the globe.

Fig. 2 presents global maps of six anomalies for the month of March 2007 (See supplementary material for larger version and other months), and Fig. 3 shows other ocean color products. Visual inspection shows that none of the anomalies reproduce the patterns observed in the chlorophyll maps. Corroborating this observation, the $\bar{R}^2|$ between the anomalies and all chlorophyll or phytoplankton absorption estimates are always below 0.35 (see Table 1). Because chlorophyll algorithms have been extensively validated, this suggests that the patterns observed are generally independent of chlorophyll at the global scale. Because the accepted origin of the main source of variability in so-called “case-1” waters is phytoplankton absorption and the concentration of covarying substances, the division by the LUT outputs during the creation of the anomalies removes, as it should, this first order effect. It should also, at the same time, remove the effect of the covarying CDOM concentration (Morel & Gentili, 2009; Siegel, Maritorena, Nelson, & Behrenfeld, 2005) and particulate backscattering (Antoine et al., 2011; Brewin, Dall’Olmo, Sathyendranath, & Hardman-Mountford, 2012; Huot, Morel, Twardowski, Stramski, & Reynolds, 2008; Westberry, Dall’Olmo, Boss, Behrenfeld, & Moutin, 2010) if appropriate bands are included in the LUT. This last condition is important as, given the observation above regarding the correlation of the $R_{rs}(\lambda)$ bands, a LUT that does not include input bands higher than 488 nm may not be able to capture the effect of backscattering. Similarly, we could expect a LUT with only wavebands below 531 nm to not capture variability associated with CDOM and phytoplankton absorption.

![Fig. 2. Anomalies obtained for the month of March 2007 as per the title in each panel. The histogram at the right of each panel indicates the relative frequency of data for the value of the anomaly. The percentage given at the bottom of each histogram represents the fraction of the total data within the red, white and blue parts of the color bar. Values above and below these limits are colored yellow and green respectively in the maps. Values that fall outside the LUT as well as values that are not observed by the sensors (masked) are colored in black; in Fig. 3a only masked pixels are colored black such that a comparison allow identifying the pixels that fall outside the LUT.](image-url)
3.2. The $R_{412,488,547}^2$ anomaly is caused by CDOM anomalies

The $R_{412,488,547}^2$ anomaly has an $R^2 = -0.59$ with $R_{443,488,547}^2$ and $R^2 = -0.67$ (see Table 1) with the CDOM index (Morel & Gentili, 2009), highlighting a common source of variability. Fig. 4 illustrates the origin of the inverse relationship with $R_{412,488,547}$. This figure was developed by varying the concentrations of CDOM and chlorophyll using standard case 1 relationships as described in the figure legend. In panel C, the reflectance arising from the spectrum with additional CDOM (Chl = 0.3, 1.25 at CDOM) has a lower $R_{412}^2$ than the standard Case 1 water spectrum for the same chlorophyll concentration but very similar values at 443, 488 and 547 nm, resulting in a $R_{443,488,547}^2$ that is lower than 1. The spectrum from the standard model that best fits the $R_{412}^2$ spectrum with higher CDOM at 412, 488 and 547 nm is for a higher chlorophyll concentration (Chl = 0.355 mg m$^{-3}$; the thin line on Fig. 4A). Such a spectrum should be underlying the LUT as it represents the average spectrum of the standard model. For that spectrum, the $R_{412,488,547}^2$ anomaly is greater than 1 (the standard spectrum is below the spectrum with more CDOM at 443 nm). The inverse is true for the $R_{412}^2(\lambda)$ resulting from the absorption spectrum with more CDOM absorption than the mean oceanic spectrum. The $R_{412,488,547}^2$ thus has, in large part, the same origin as $R_{412,488,547}^2$ though about 40% of the variability originates from other sources.

The CDOM index (Morel & Gentili, 2009) was developed to provide the ratio between the actual absorption by CDOM and the expected absorption by CDOM for a given chlorophyll concentration (actually band ratios). Based on case 1 waters statistics, the index uses the ratios of reflectance 412 to 443 nm and 488 to 555 nm to describe a function (in the form of a LUT or grid) that returns the CDOM index. Assuming the variability retrieved in the CDOM index has been attributed correctly by Morel and Gentili (there is no reason to believe otherwise), the high determination coefficient with this index, together with the low

Fig. 3. Remote sensing products distributed by the NASA ocean color web site for the month of March 2007. The name of the product is given in the title of each panel. The algorithms used are as follow: Chlorophyll OC3M (oceancolor.gsfc.nasa.gov/REPROCESSING/R2009/); sea surface temperature — standard algorithm 11 μm daytime; CDOM and NAP absorption — QAA (Lee et al., 2002); CDOM index — Morel and Gentili (2009); Particulate backscattering coefficient — QAA (Lee et al., 2002). Averages were computed as for the anomalies by using the level 3 daily images (see Section 2.2).
Here, the backscattering is accounted for by directly entering absolute values of $R_{412}^{\text{ex}}(\lambda)$ in the LUT and the remotely measured dataset is used to obtain the “mean” (LUT) value. The advantage of not using ratios is that waters with identical ratios but different values of $R_{412}^{\text{ex}}(\lambda)$ can be addressed as different classes. Avoiding the modeling step has the advantage of preventing any potential biases between the model and remotely sensed reflectance (e.g., measurement errors, spectral band definition, sensor calibration, erroneous water IOPs). The disadvantage is that interpretation of the anomalies in terms of CDOM and potential correction of chlorophyll algorithms are not as easily carried out (e.g. Brown et al., 2008; Morel & Gentili, 2009). By their definition, the anomaly and the CDOM index are inversely related; in our case, increasing CDOM leads to a decreasing anomaly.

The distribution of this anomaly highlights once again (see also Brown et al., 2008; Morel & Gentili, 2009) the very low CDOM concentration relative to chlorophyll in the Equatorial Pacific Ocean compared with the Atlantic Ocean, reflecting the different CDOM sources and sinks in these basins (Nelson & Siegel, 2013). Similarly, the time series at BOUSSOLE and BATS (Fig. 5) show an increase of the $R_{412,443,547}^\text{fl}$ during the spring bloom. As observed in previous studies (e.g. Brown et al., 2008; Morel & Gentili, 2009), the bloom is associated with an increased phytoplankton absorption while the CDOM remains more stable, leading to an increase in the ratio of phytoplankton to CDOM absorption hence to an increase of $R_{443,488,547}^{\text{fl}}$ and a decrease of $R_{412,443,547}^{\text{fl}}$. It is interesting to note also the strong correlation between the times series of the two anomalies at BOUSSOLE and BATS after ~2008 which suggests similar seasonal balances between production and degradation of CDOM at these two sites.

### 3.3. The unknown sources of variability for the $R_{412,443,547}^{\text{fl}}$ anomaly

In the month of March 2007, the $R_{412,443,547}^{\text{fl}}$ (Fig. 2) tended to be slightly above the yearly average in the global ocean (i.e., above 1), while 95% of the data were within 10% of the LUT table estimates (see histogram and color bar). The strongest positive anomalies occurred in the Southern Ocean where low values were observed in the North Atlantic and subtropical western North Pacific Ocean. Compared to other anomalies, stronger seasonal patterns are observed on the global scale (see supplementary material); large regions (e.g. the equatorial Pacific) switch from positive to negative with an annual cycle (see also Fig. 5 described below). This anomaly is the most intriguing, as it is not functionally related to any other NASA-distributed ocean color product. Atmospheric correction artifacts on global scales do not appear to be at the origin of this anomaly as it is not correlated with the aerosol optical depth at 869 nm or to the Angström exponent; the same is true of all other anomalies (see Table 1). It is most related, though only very weakly, to the particulate backscattering coefficients from the Garver, Siegel and Maritorena (GSM, Maritorena, Siegel, & Peterson, 2002) algorithm, the generalized IOP algorithm (GIOP, Werdell, Franz, Bailey, et al. (2013a), and the Quasi Analytical Algorithm (QAA, Lee, Carder, & Arnone, 2002)) with $R_{412}^\text{fl}$ between 0.27 and 0.31. It also shows a low $R_{412}^\text{fl}$ of 0.22 with $R_{412,443,547}^{\text{fl}}$, highlighting more their differences than similarities despite their similar definitions and prediction wavelengths (488 and 531 nm) that are next to each other in the measured MODIS spectra (though on opposite sides of the absorption-driven vs backscattering-driven parts of the spectrum, see above). Because the IOPs of the different in-water constituents are not independent we cannot use the sum of the $R_{412}^\text{fl}$ to explain a larger fraction of the variability in $R_{412,443,547}^{\text{fl}}$. Furthermore, the unknown response of algorithms to anomalies means that more advanced statistical analyses (e.g. stepwise multiple regression analyses) are inappropriate as they may reflect more the response of the algorithms than explain the origin of the anomaly through the algorithm outputs. As opposed to anomalies that show strong correlation with existing validated products, the use of

![](image)
remotely sensed products has, therefore, clear limitations in trying to interpret $R_{412,443,547}$. The three bands used in the LUT should, in principle, account together for the mean changes in the amplitudes of 1) the absorption by CDOM (strongest impact at 412 nm), 2) the absorption by phytoplankton (strongest impact at 443 nm) and 3) backscattering by particles (most discernable impact at 547 nm). Therefore, it is likely that the signal observed here originates from small changes in one or a combination of absorption or backscattering component shapes relative to the mean shape in the ocean for the same $R_{rs}$ ($\lambda$). While amplitude effect cannot be completely excluded, given the constraints above, changes in the spectral shapes of the IOPs, which are indeed well known to occur (most obvious in phytoplankton absorption), appear to be the most likely source. It does not, however, have to be the same components everywhere.

In addition to computing the $R^2$, other tests were carried out to identify the source of this anomaly. Clear oceanic patterns in this anomaly (compare Figs. 2 and 3) allow us to reject the variation observed as simply noise. We have examined several scenes to search for clues of its origin. Fig. 6 shows one of these scenes measured around New Zealand. While the $R_{412,443,547}$ is overall high in this scene, there is a continuous region of low anomalies along and offshore of the South Island. The feature is seen neither in the other anomalies (the offshore portion appears in the $R_{412,443,547}$ anomaly) nor in the chlorophyll concentration. It appears to very closely follow the subtropical front (~15 °C isotherm in SST, see also Pinkerton et al., 2005), highlighting a particular optical feature of the front. Lower values of $R_{412,443,547}$ are also observed in Fig. 2 near New Zealand and extend beyond Australia. This scene also highlights that all the anomalies are clearly identifying mesoscale processes.

At the Mediterranean BOUSSOLE site (43°22′ N 7°54′ E), high values of this anomaly are generally associated with the spring bloom (see Fig. 5), with sometimes very low values just preceding or in the first part of the spring bloom. One of the potential sources of variability in this signal could originate from different spectral shapes of phytoplankton absorption. Using the BOUSSOLE dataset we have tried to examine such potential links. In situ observations have shown that the spring blooms at BOUSSOLE tend to be dominated either by nanophytoplankton or microphytoplankton (Gernez, Antoine, & Huot, 2011). During the time series presented in Gernez et al., years 2002, 2003, 2004 and 2007 were dominated by nanophytoplankton while years 2005 and 2006 were dominated by microphytoplankton. A qualitative comparison of the blooms and anomalies for these years do not show any trends. We carried out a more quantitative analysis by using the in situ HPLC-derived size classes (microphytoplankton, nanophytoplankton, picophytoplankton, according to Uitz, Claustre, Morel, & Hooker, 2006), which we converted into a continuous size.
variable according to Bricaud et al. (2004). We then divided this size value by the expected value for the measured chlorophyll concentrations according to the Uitz et al. (2006) statistics (working from the assumption that anomalies should be linked to a departure from a mean state). For all samples, when waters were stratified (parameterization is provided in 2006) from 2002 until 2007 (N = 65) for which there was a satellite matchup within 2 days, we compared the $R_{412,443,547}^{488}$ with this departure from the mean phytoplankton size ratio. No relationships emerged. This appears to exclude phytoplankton species as a dominant control on this anomaly at the BOUSSOLE site.

Antoine et al. (2011) published a time series of the particulate backscattering coefficient and particulate backscattering spectral slopes for the BOUSSOLE site. Neither of the parameters covary with the $R_{412,443,547}^{488}$ anomaly. This leaves very few obvious candidates for the source of this anomaly at BOUSSOLE. Among them are: the shapes of the colored detrital components, which could include the ratio of detritus to dissolved matter and variability in the depth profiles of the IOPs.

We are not aware of statistics and datasets, which would allow us to test these hypotheses.

At the BATS and HOT sites (Fig. 5), the anomaly shows a regular annual cycle with a maximum in summer and a minimum in winter. At the BOUSSOLE site, in addition to a less regular annual cycle, the anomaly generally drifts to lower values than the mean between 2008 and 2013 while the other sites remain near the mean value through the time series.

### 3.4. The $R_{412,443,555}^{488}$ and $R_{412,443,488}$ anomalies

Fig. 6. Chlorophyll concentration ($\log_{10}$[mg m$^{-3}$]) and sea surface temperature (°C) for the waters surrounding New Zealand as well as six anomalies as indicated in the panel titles. Color scales for the anomalies are the same as those used in Fig. 2. The scene was taken on 23 March 2007.

Fig. 7. The $R_{412,443,555}^{488}$ anomaly for the year 2007 for different latitudes (see legend) in the Pacific Ocean. Daily level 2 images for the four points at different latitudes were downloaded, the average of $11 \times 11$ pixels taken, and a five-point running average was applied to the resulting time series.
Given the correlation with backscattering estimates, one possible source of variability in these anomalies would be “backscattering anomalies”, that is, departures from the mean relationship between total absorption or chlorophyll and backscattering. As a simple test of this hypothesis, we fitted a polynomial between the phytoplankton absorption obtained from the QAA algorithm and the backscattering coefficient from the QAA algorithm on the monthly composite dataset both at 443 nm. We then built an anomaly by dividing the remotely sensed QAA bbp by the output from the fit for each pixel. The $R^2$ between this backscattering anomalies and $R_{412,443,547}^{bbp}$ and $R_{412,443,488}$ for the March dataset (bbp anomaly in Table 1) were $0.22$ for both, which is much lower than for $R_{443}$ (443). We also found low $R^2$ with estimates of the ratio of $bbp(443)$ over total absorption obtained from semi-analytical algorithms (data not shown). As such, it does not appear that departures from the mean backscattering represent a stronger source of variability in this anomaly than $bbp(443)$ itself. It therefore seems that in the case of the $R_{412,443,488}$ anomaly, the effect of backscattering is not sufficiently strong on the input bands (412, 443 and 488 nm) relative to the effect of absorption such that this anomaly reflects strongly the amplitude of backscattering and does not represent “departures from the mean” backscattering.

The $R_{412,443,547}$ and $R_{412,443,488}$ anomalies show strong annual cycles at BOUSSOLE that are synchronized to each other (Fig. 5), with higher values during the spring bloom and lower values in summer and winter. This is consistent with the measurements of the particulate backscattering at 555 nm made at BOUSSOLE (2011). At HOT and BATS (Fig. 5) both anomalies show annual cycles, but the $R_{412,443,488}$ is noisy and has low amplitude. Nevertheless, in both cases, they show higher values in summer and lower values in winter.

### 3.5. The $R_{412,443,547}$ anomaly: coastal anomalies and potential solar zenith angle biases

Outside coastal waters, as mentioned above, our initial working hypothesis concerning this anomaly was that it would be related to the spectral slope of the backscattering coefficient; since the input bands in the LUT (and in particular the 547 nm bands) provide much of the information on the backscattering amplitude, we would expect that departures from the mean oceanic values at 667 nm would be caused by a change in the slope of the backscattering spectrum. However, it is at this time difficult to test this hypothesis because there appears to be a large effect of the solar zenith angle on the $R_{412,443,547}$ anomaly arising from a likely difference in the way the solar zenith angle affects $R_{412}^{bbp} (\lambda)$ in the input bands and at 667 nm. Strong latitudinal and temporal seasonal effects and the opposite phases in the Northern and Southern hemisphere (Fig. 7) of this anomaly are pointing strongly to such an artifact. This effect leads to deviations up to 60 to 70% from the mean. It thus far outweigh biogeochemical effects in this anomaly in most oceanic water (see also Fig. S6) and is probably the reason why the coastal waters anomalies are only strongly highlighted in the lower latitudes in Fig. 2. As opposed to this anomaly, the other anomalies did not show such latitudinal dependence in the annual cycles (compare Fig. S6 with other figures in the supplementary material). The alternative hypothesis that water temperature could cause this artifact is considered unlikely because the temperature dependence of water absorption is very small at 667 nm (Röttgers, McKee, & Utschig, 2014) neither is consistent with the observed distributions in Fig. S6. Similarly, the possibility that varying IOP depth profiles could be the reason for these changes appears remote since the optical depth (and thus the depth observed...
by the satellite sensor) is limited mostly to the mixed layer due to the high water absorption at this wavelength.

Apart from the BOUSSOLE site, which shows annual cycles (likely linked to the solar zenith angle discussed above), the anomalies at BATS and HOT (lower latitudes) are very noisy (data not shown).

3.6. Small sensor shifts are highlighted by the anomalies

One interesting side result of our study is that anomalies allow small temporal shifts in the response of the sensor to be identified. This is because the anomalies, by design, remove most of the first order variability in $R_{fi}^{412}(\lambda)$, so that they can provide a useful tool to examine a sensor’s spectral response. Fig. 8 shows the time series of $R_{fi}^{412,488,547}$ at the Hawaiian Ocean Time series, which provides a particularly stable environment with essentially no seasonal effects in $R_{fi}^{412,488,547}$ and a value near 1. However, sometime in 2008 or the beginning of 2009, there was a slight decrease observed in $R_{fi}^{412,488,547}$ which appears to last until the end of the time series. This shift is accompanied by a change in the behavior of the $R_{fi}^{443,488,547}$ anomaly towards stronger variability and higher values (See Fig. 5). Though at face value this would represent a shift in the gyre biogeochemistry, the hypothesis can be rejected because two other distant sites show a similar shift in $R_{fi}^{412,488,547}$ (Fig. 8 B, C and D). While the times series show more variability at these sites, the different distribution of points before and after 2009 also points to such a sensor shift. The most likely cause is thus a slight shift in the sensor response that has not been fully characterized and corrected. Future reprocessing will likely allow this potential artifact to be corrected (a new reprocessing has occurred since the data in this paper was processed).

4. Conclusion

We have defined several anomalies in ocean color whose variability originates from biogeochemical sources or algorithm biases. All of these anomalies cannot, obviously, be independent of each other and of other data products. Some, however, appear to meet this criterion and could represent new sources of information about oceanic processes or biases in algorithms (in particular, the $R_{fi}^{412,443,547}$ and $R_{fi}^{443,488,547}$ anomalies). In some cases, we were able to identify the source of the variability in these anomalies while, in others, these sources eluded our analyses and further work will be required to obtain a better understanding. Computing such anomalies appears to provide a powerful means of looking at shifts or biases in the calibration of ocean color sensors, especially in stable environments. Future work should also allow us to identify a biogeochemical source for more of these anomalies and algorithms will be developed to provide new insights into ocean processes. This will be difficult, though, as they represent only very small spectral shifts, less than a few percent, which will require well-calibrated in situ sensors to correctly observe. The stark differences in the $R_{fi}^{412,443,547}$ and $R_{fi}^{443,488,547}$ anomalies emphasize that adjacent spectral bands in current ocean color sensors can convey different information highlighting the potential usefulness of more spectral bands on ocean color sensors (cf. Lee, Shang, Hu, & Zibordi, 2014). At this point, beyond the magnitudes of CDOM (and non-algal particulate) and phytoplankton absorption and backscattering coefficient, the next most important source of variability in remotely sensed reflectance is largely unknown; however, it is likely that the variable shape of one of these constituents is important. Different authors, therefore, have made educated guesses when building inversion models as to which additional variable can be retrieved (e.g. the spectral slope of $b_{sp}$, the spectral slope of CDOM, or phytoplankton type). At this point, however, it appears that our limited knowledge of the “next largest source” of IOPs variability limits our ability for interpreting the anomalies developed herein, and to develop inverse models anchored in an understanding of fine optical variability in the ocean.

Acknowledgments

We thank Aymeric Ambert for the help with computations and ZhongPing Lee for the comments on an earlier draft. This research was funded by grants to Huot from NSERC Discovery program and the Canada research chair program.

Appendix A. Supplementary data

Supplementary data to this article can be found at [online link].

References
