Discrimination of diatoms from other phytoplankton using ocean-colour data

Shubha Sathyendranath[1,2], Louisa Watts[3], Emmanuel Devred[1,2,*],
Trevor Platt[2], Carla Caverhill[2], Heidi Maass[2]

[1]**Department of Oceanography, Dalhousie University, Halifax, Nova Scotia B3H 4J1, Canada**
[2]**Biological Oceanography Division, Bedford Institute of Oceanography, Box 1006, Dartmouth, Nova Scotia B2Y 4A2, Canada**
[3]**Atmospheric Science Team, Natural Environment Research Council, Polaris House, North Star Avenue, Swindon SN2 1EU, UK**

ABSTRACT: Recent papers have highlighted the differences between the absorption characteristics of phytoplankton populations dominated by diatoms and those of other types of phytoplankton populations from the North West Atlantic. It has been suggested that these differences could introduce a bias in satellite-derived concentrations of the phytoplankton pigment, chl *a*. In this paper, these differences in optical properties of diatoms are exploited to develop a bio-optical algorithm to distinguish diatom populations from other types of phytoplankton populations in the region. The algorithm is applied to SeaWiFS data on ocean colour, and the results are compared with *in situ* data on phytoplankton population types based on HPLC data. The comparison shows that the algorithm successfully distinguishes between diatoms and non-diatom populations in the majority of cases studied. A branching algorithm is then applied to the satellite data to estimate chl *a* concentration in the region: a diatom-specific algorithm is used when diatoms are identified in a pixel, and another algorithm for mixed populations when this is not the case. The estimated chl *a* concentrations are compared with *in situ* estimates when matching observations exist. The results show that the branching bio-optical algorithm often performs better than the OC4 algorithm used in standard processing of SeaWiFS data. However, the results may be poor when the initial identification of population types is wrong. Finally, the new algorithm is used to map the distribution of diatoms in the region in spring and summer: the patterns that emerge are consistent with the known features of diatom distributions in the region.

KEY WORDS: Phytoplankton community structure · Ocean colour · Diatoms · Remote sensing · SeaWiFS · North West Atlantic

INTRODUCTION

Over the last decade a new generation of satellite ocean-colour sensors has been launched into space. The German Modular Optoelectronic Scanner (MOS), the Japanese Ocean Colour and Temperature Scanner (OCTS) and the French Polarisation and Directionality of the Earth's Reflectances (POLDER) were all launched in 1996. The Sea Viewing Wide-Field-of-View (SeaWiFS) and the Moderate Resolution Imaging Spectroradiometer (MODIS-Terra) were launched in 1997 and 1999, respectively, as US initiatives. The European sensor called the Medium Resolution Imag-

ing Spectrometer (MERIS) was launched in 2002. These improved instruments acquire radiometric data in several spectral wavebands in the visible and infra-red domains, with the number of wavebands ranging from 8 (SeaWiFS and POLDER) to 36 (MODIS). The resulting wealth of information is proving to be a powerful resource to the scientific community for basin-scale biogeochemical research, since it provides a window into the synoptic state of the pelagic ecosystem.

The algorithm developed by the National Aeronautics and Space Administration (NASA) is applied to radiometric data measured over 4 of the spectral bands on the SeaWiFS sensor to retrieve the main phyto-

plankton pigment in the ocean, chl *a* (O'Reilly et al. 1998). This is the standard type of algorithm currently in use for producing maps of chl *a* distribution in the world ocean from SeaWiFS data. In turn, these maps can be used to study the photosynthetic activity of marine phytoplankton and its impact on bio-geochemical processes in the ocean at basin scales. Such studies are crucial to the aims of large, international oceanographic programs such as the Joint Global Ocean Flux Study (JGOFS), the Global Ocean Ecosystems Dynamics (GLOBEC) and Surface Ocean-Lower Atmosphere Study (SOLAS).

Remote-sensing algorithms, such as Version 4 of the NASA algorithm (known as OC4), that are designed for use at global scales (O'Reilly et al. 1998) may be less accurate at local and regional scales. This is due to the sensitivity of such algorithms to changes in the composition of phytoplankton species present in the water (Sathyendranath et al. 2001, Bouman et al. 2003). Phytoplankton and associated particulate and dissolved material absorb and scatter light, and their collective influence dominates the optical variability of the water column, at least in the open-ocean waters generally referred to as Case 1 waters (Morel 1980). In such waters, it is the phytoplankton and substances that covary with them that determine the magnitude and spectral quality of the radiance (flux per unit area and unit angle) that leaves the sea-surface, for a given incident irradiance at the sea's surface. This flux eventually reaches the satellite sensor after passing through the atmosphere. The spectral values of the water-leaving radiance (corrected for the influence of the atmosphere and for variations in incoming solar radiation) are then used in an algorithm to retrieve chl *a* concentrations. Errors arise in such retrievals because the corrected water-leaving radiance may vary independently of changes in the concentration of chl *a*. Several factors may be responsible for such effects, but this paper deals primarily with one of them: variability in the optical characteristics of phytoplankton populations.

The specific absorption coefficients of phytoplankton cells (the absorption coefficient of phytoplankton at a particular wavelength normalised to its chl *a* concentration) can vary because of differences in the pigment composition and size structure of phytoplankton populations (Sathyendranath et al. 1987, Mitchell & Kiefer 1988, Babin et al. 1993). Modifications in the pigment composition of the phytoplankton population (changes in the type and amount of auxiliary pigments present relative to the concentration of chl *a* in the sample) can arise from adaptations of existing populations to modifications in the light and nutrient regimes, or from alterations in the phytoplankton population itself (Hoepffner & Sathyendranath 1992, 1993, Johnsen & Sakshaug 1993, Bricaud et al. 1995, Culver & Perry

1999, Marra et al. 2000). When the pigment composition changes, there are corresponding changes in the spectral absorption characteristics of phytoplankton (Sathyendranath et al. 1987, Bidigare et al. 1990, Hoepffner & Sathyendranath 1991). Furthermore, the absorption efficiency of pigments decreases with increases in either the cell size or the intracellular pigment concentration (Duysens 1956, Kirk 1975, 1976, 1994, Platt & Jassby 1976, Morel & Bricaud 1981, Sathyendranath et al. 1987, Hoepffner & Sathyendranath 1991, Sosik & Mitchell 1994).

Indeed, a recent study (Stuart et al. 2000) of the Labrador Sea during the spring of 1996 has shown that diatom populations exhibit much lower specific absorption coefficients than other phytoplankton populations (predominantly smaller prymnesiophytes) from the same region. The effect is attributable to increased pigment packaging in the larger diatom cells as well as an increase in intra-cellular chl *a* concentrations resulting from adaptation to low light levels. This may lead to a significant under-estimation in retrieved chl *a* concentrations in this area if a universal algorithm is used (Sathyendranath et al. 2001).

The water-leaving radiances from the blue and green parts of the spectrum are used in ocean-colour algorithms for retrieval of chl *a* concentrations. Hoepffner & Sathyendranath (1992) note that variations in the proportion of the pigment fucoxanthin (a pigment common to diatoms) relative to chl *a* can change the blue-green ratio of phytoplankton absorption by a factor of up to 6, and hence could impact water-leaving radiances. Morel (1997) has shown that the blue-green ratio of waters dominated by a *Synechococcus* bloom may be different from that of waters with a 'normal' phytoplankton population having the same amount of chl *a*. He showed that this could result in an overestimate of chl *a* values by as much as a factor of 3. In situations such as those described above, the application of a universal algorithm for the retrieval of chl *a* biomass may not be ideal, and the use of a regionally parameterised algorithm becomes more appealing.

Here, we address the development of a regional algorithm for chlorophyll retrieval on the East Coast of Canada in the North Atlantic: an area that encompasses the Scotian Shelf, Georges Bank, the Laurentian Channel and part of the Grand Banks. Biogeochemical studies in these areas have shown the phytoplankton composition to be dominated by diatom cells during the spring bloom period (Bouman 2003). Application of a universal algorithm to such an area might yield inaccuracies in the retrieved chlorophyll concentrations. We have therefore developed a branching algorithm that first distinguishes populations dominated by diatoms from other phytoplankton

assemblages (referred to here as the 'mixed population'). This involved the use of a theoretical model of ocean colour, whose parameters have been tuned using local measurements of phytoplankton absorption and pigment composition. The model is used to derive an algorithm that distinguishes diatoms from mixed populations, which is then applied to ocean-colour data from SeaWiFS to retrieve maps of diatom distribution for the study area. Finally, the branching algorithm is used for the retrieval of chl a concentrations. The algorithm modifies the procedure for estimating chlorophyll concentrations, depending on whether it is being applied to a diatom bloom, or not. We examine whether this leads to improvements in the accuracy of the retrieved concentrations, over the standard algorithm (OC4). The results are critically evaluated, and the limitations and possible avenues for improvements are identified.

METHODS

The ocean-colour model. The model used here is a combination of the models developed by Sathyendranath & Platt (1997, 1998), as implemented by Sathyendranath et al. (2001), with some modifications. It relates absorption and backscattering coefficients to reflectance at the sea surface. Variations in reflectance at the sea surface determine the intrinsic colour of the ocean. This model is designed for use in Case 1 waters: waters in which the optical properties are determined by the phytoplankton and covarying substances present (Morel 1980).

Reflectance (dimensionless) is defined as the ratio of upwelling irradiance (flux per unit area) to downwelling irradiance at the same depth:

$$R(\lambda, z) = \frac{E_u(\lambda, z)}{E_d(\lambda, z)} \qquad (1)$$

where $R(\lambda, z)$ is the reflectance at wavelength λ and depth z, $E_u(\lambda, z)$ is the upwelling irradiance at the same wavelength and depth and $E_d(\lambda, z)$ is the corresponding downwelling irradiance. For given chl a concentrations, a model is used to compute reflectances at the 6 wavelengths that match the mid-values of the 6 SeaWiFS wavebands in the visible portion of the spectrum: 412, 443, 490, 510, 555 and 670 nm. Look-up tables are generated for matched pairs of chl a concentrations and reflectances.

The model assumes that reflectance at a given wavelength just below the sea-surface ($z = 0$) is the sum of reflectances associated with Raman and elastic scattering. The contribution of elastic scattering to reflectance at the sea-surface, $R^E(\lambda, 0)$, is represented in the model as:

$$R^E(\lambda, 0) = r \frac{b_b(\lambda)}{a(\lambda) + b_b(\lambda)} \qquad (2)$$

where $a(\lambda)$ is the absorption coefficient at λ, $b_b(\lambda)$ is the backscattering coefficient, and r is a proportionality factor. The inherent optical properties a and b_b are taken to be depth-independent in this model. The model used here is a Case 1 model: i.e. all the inherent properties are modelled as a function of chl a concentration.

Reflectance due to Raman scattering is computed according to Sathyendranath & Platt (1998) and includes a first-order Raman scattering term that accounts for single upward Raman scatter (R^R) and 2 second-order terms accounting for a combination of elastic and Raman scattering events (R^{RE} and R^{ER}). The total reflectance at the sea-surface is then computed as the sum of the contributions due to the elastic and Raman scattering.

The backscattering coefficient is expressed in the model as the sum of 2 components: backscattering by pure seawater, and particles in suspension (including phytoplankton). The backscattering by water is computed here according to Morel (1974). The particle backscattering is expressed as a function of chl a concentration, using the theoretical and empirical results of Ulloa et al. (1994) and the observations of Loisel & Morel (1998), as detailed in Sathyendranath et al. (2001).

The total absorption coefficient a is calculated as the sum of the absorption coefficients of pure water (a_w), yellow substances (a_y) and phytoplankton (a_p):

$$a(\lambda) = a_w(\lambda) + a_p(\lambda) + a_y(\lambda) \qquad (3)$$

Absorption by pure seawater was computed according to Pope & Fry (1997) for all the 6 SeaWiFS wavelengths and their corresponding Raman source wavelengths, except for the source wavelength for 412 nm (362 nm, corresponding Raman wavelength), for which we relied on data from Sogandares & Fry (1997). Since the model is designed for Case 1 conditions, absorption by yellow substances is computed as a function of chl a. Details of this implementation are given in Sathyendranath et al. (2001).

Phytoplankton absorption model. In modelling the absorption coefficient for particulate material (i.e. the absorption coefficient of phytoplankton, given the assumption of Case 1 conditions), we follow Sathyendranath et al. (2001), with the following main difference: the phytoplankton absorption is parameterised here using data from the North West Atlantic Zone (NWAZ) limited by the co-ordinates 39° N, 71° W; 62.5° N, 42° W. This region was chosen to encompass the areas of interest listed in the 'Introduction' (see Fig. 1). The total number of stations used for the para-

meterisation of phytoplankton absorption was 222. The ship-borne data used were collected over 6 oceanographic expeditions during the months of April, May, June and October, over the years 1996 to 1998. Data from 1999 and 2000 were kept aside as an independent data set for testing the performance of the algorithm developed here. Ship-based measurements of phytoplankton absorption spectra and the corresponding HPLC (High Performance Liquid Chromatography) data on pigments (all data collected at depths less than or equal to 10 m) are used in the study. Details on the methodologies used can be found in Stuart et al. (1998), Sathyendranath et al. (1999) and Head & Horne (1993).

The HPLC data were used to separate diatom-dominated samples from the mixed population based on the concentrations of chl c_3 and fucoxanthin relative to chl *a*. If the HPLC sample for a station was characterised by a ratio of chl c_3 to chl *a* less than 0.02 and a ratio of fucoxanthin to chl *a* greater than 0.4, then those stations were categorised as diatom-dominated stations, as in Stuart et al. (2000) and Sathyendranath et al. (2001). Out of the 222 stations, we found a total of 75 stations that fit these criteria. The remaining data were assigned to the mixed population. Stuart et al. (2000) and Sathyendranath et al. (2001) note that blooms of prymnesiophytes (*Phaeocystis* sp.) also occur in this region. But Sathyendranath et al. (2001) also point out that the prymnesiophyte blooms have optical characteristics that are very similar to the mixed populations from low and mid latitudes, at least for the SeaWiFS wavelengths. Therefore, in this study we have not attempted to distinguish between prymnesiophytes

and other phytoplankton. The effort is focussed here on identifying diatoms.

The absorption coefficients of the diatoms and the mixed populations a_p at each of the 6 SeaWiFS wavelengths and their corresponding Raman source wavelengths were then parameterised as non-linear functions (Sathyendranath et al. 2001) of chl *a* concentration *C* determined from HPLC measurements:

$$a_p(\lambda) = U(\lambda)\{1 - \exp(-S[\lambda]C)\} + Ca_2^\star(\lambda) \qquad (4)$$

where U (m^{-1}), S (m^3 (mg chl *a*)$^{-1}$) and $a_2^\star$ (m^2 (mg chl g)$^{-1}$) are parameters fitted as a function of wavelength λ and *C* is the chlorophyll concentration (mg chl *a* m^{-3}). In the fitting programme, *S* was constrained to take a value between 0 and 2 (Sathyendranath et al. 2001). Examples of the fits of Eq. (4) to the data for diatom and mixed populations are shown in Fig. 2. Eq. (4) was derived (Sathyendranath et al. 2001) as a simple representation of absorption by 2 populations of phytoplankton: one with high specific absorption coefficients that maintains a low concentration, and a second population with low specific absorption coefficients capable of attaining high concentrations. The non-linearity in the model used to describe diatom populations admits implicitly that the population itself could be constituted of cells with higher specific absorption coefficients at low concentrations, and of cells with lower specific absorption coefficients at higher pigment concentrations. Such changes may result from a change in either the size of the cells or the intra-cellular concentration of pigments.

Fitting the equation to all the wavelengths of interest for the 2 groups of data yields 24 sets of parameters
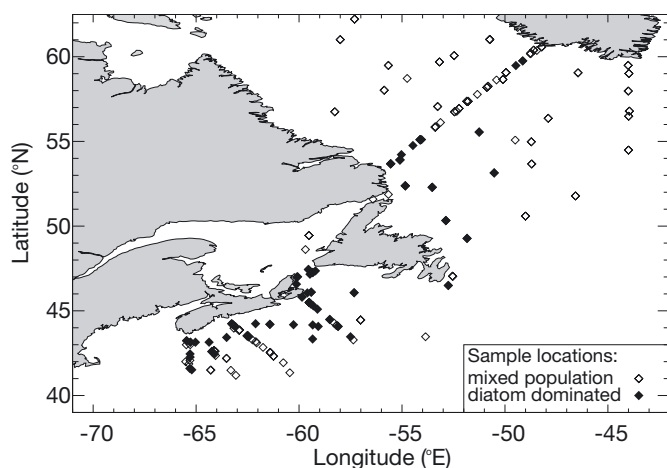

Fig. 1. North West Atlantic Zone (39° N, 71° W; 62.5° N, 42° W) study area. Locations of *in situ* sampling are also shown. Different symbols are used to indicate locations where phytoplankton populations dominated by diatoms were encountered, and those where other types of phytoplankton populations were found
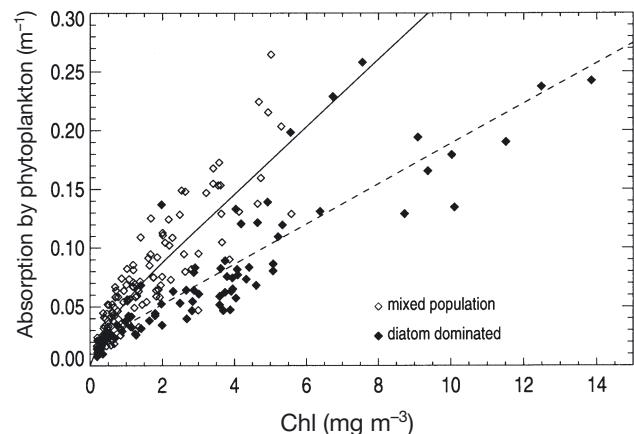

Fig. 2. Phytoplankton absorption at 443 nm, a_p(443), plotted as a function of chl *a* concentration, for mixed populations and for diatoms. The non-linear function fitted to the diatom data (dashed line) and to the mixed populations (continuous line) are also shown. Parameters corresponding to the functions fitted to all the wavelengths used in the study are given in Table 1

Table 1. The parameters U (m^{-1}), $a_2^\star$ (m^2 mg chl a^{-1}) and S (m^3 mg chl a^{-1}) for the mixed and diatom-dominated populations, as a function of wavelength λ (nm). Parameters are given for SeaWiFS wavelengths (lower member of each wavelength pair), and for their corresponding Raman wavelength (upper member)

| λ (nm) | Mixed | | | Diatom | | |
	U	$a_2^\star$	S	U	$a_2^\star$	S
362	0.0624	0.0273	0.541	0.0254	0.0234	8.72
412	0.0376	0.0269	1.01	0.103	0.00950	0.234
386	0.0424	0.0236	0.753	0.0277	0.0161	2.43
443	0.0281	0.0310	1.80	0.0747	0.0118	0.270
421	0.0340	0.0285	1.23	0.0939	0.0102	0.252
490	0.0248	0.0146	1.68	0.0122	0.00981	4.68
435	0.0306	0.0316	1.65	0.0897	0.0113	0.248
510	0.0178	0.00979	1.41	0.00786	0.00810	4.77
468	0.0236	0.0249	2.12	0.0177	0.0128	2.95
555	0.00386	0.00554	2.13	0.00302	0.00443	2.35
547	0.00542	0.00644	1.93	0.00349	0.00522	3.12
670	0.00586	0.0169	2.29	0.0110	0.0104	1.51

(Table 1) that are then used in the model to describe absorption of light by phytoplankton (a_p) as a function of the chl a concentration. The ocean-colour model was then run to create look-up tables of predicted reflectances, given a range of sea-surface chl a concentrations, from 0.01 to 40 mg chl a m^{-3}. Two sets of look-up tables were created: one for the diatom population and the other for the mixed population. A uniform chl a profile was assumed, and the model was run for the middle of April and for latitude 50.75° N and longitude 56.5° W, corresponding to the centre of the study region. April was selected for the initial study because it is known that diatoms dominate in the region during spring-bloom conditions (Bouman 2003). The computations are insensitive to the longitude assigned, but the date and the latitude determine the angular distribution of the light field underwater. In principle, therefore, they influence the proportionality factor r in Eq. (2). However, in the developed algorithms, we always used ratios of reflectances at 2 wavelengths. This helps reduce errors in the results due to any potential errors in r, to the extent that the variations in r are independent of the wavelengths used in the ratios.

Model approximations. We also investigated the effect of non-uniform chlorophyll profiles, and of using fixed date and location in the computations, by running the model for the following cases: (1) Non-uniform chl a profiles, whose shapes are described by 4-parameter Gaussian distribution functions (Platt & Sathyendranath 1988) fitted to observed shipboard data from the study region in April, were used in the model. The model was then run for the particular dates and locations corresponding to the ship-borne obser-

vations. (2) An average of the shapes of non-uniform profiles observed in April was computed, and the model was run for the specific dates and locations corresponding to the *in situ* observations. (3) Finally, the chl a profile was assumed to be uniform, and the model was run for the middle of April and for the central location of the study area defined above.

Comparison of the look-up tables created showed only small differences ($<5\%$) in the computed reflectance ratios. This is not unexpected, given that (1) the angular structure of the irradiance field reaching the sea-surface, as computed using the clear-sky model of Bird (1984), does not vary significantly over the space of 2 wk in April (the time over which the profiles were measured at sea); (2) the incoming irradiance does not vary significantly over the latitudinal differences observed between stations at which the profiles were measured; and (3) the depth to which the satellite sees is approximately 20% of the photic depth, which is approximately 50 m at this time of year (B. Irwin pers. comm.), so that the satellite records only (approximately) the top 10 m of the chl a profile. Since the results were seen to be fairly insensitive to the simplifying assumptions, the model runs for the uniform water column and the central pixel of the study area and for the mid-point of the period of study were used in our analyses.

Model implementation. An example of the look-up tables created for the 510:555 nm ratio for the diatom and mixed populations using the model is given in Fig. 3. In principle, all the visible wavelengths of SeaWiFS can be used to distinguish diatoms from other phytoplankton populations. In practice though, we recognise
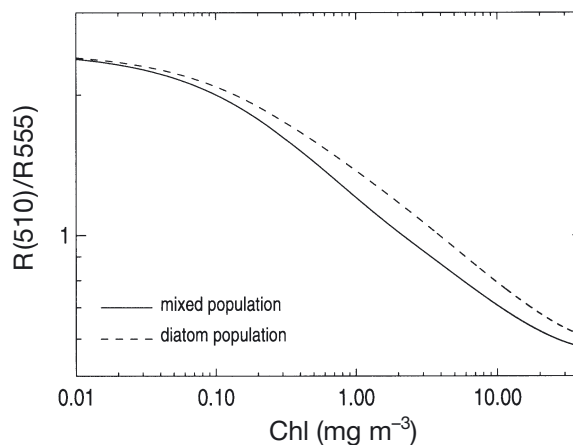

Fig. 3. Ratio of reflectance at 510 nm to that at 555 nm modelled as a function of concentration of chl a. The 2 curves indicate models for diatoms and mixed populations

that the water-leaving radiances at short wavelengths of 412 and 443 nm may be highly influenced by absorption by yellow substances (Bricaud et al. 1981, Bukata et al. 1981, Roesler et al. 1989, Gallegos et al. 1990). Since the model used here is designed for Case 1 waters, it does not account for any potential changes in the concentration of yellow substances relative to that of phytoplankton. Furthermore, in processing Sea-WiFS data for the study region, it was noted that a large number of pixels returned unrealistic negative values for water-leaving radiances at 412 and 443 nm, indicating problems with the atmospheric correction algorithm, especially for these wavelengths. Hence, the signals from these 2 wavelengths were excluded, and only the signals at 490, 510, 555 and 670 nm have been used in this analysis.

To determine whether a given pixel in a satellite image was dominated by diatoms or mixed populations, we proceeded as follows: For each pixel, the normalised remote-sensing reflectance R_{RS} was computed for the 4 wavelengths used in the algorithm. Here, $R_{RS}(\lambda)$ is defined as:

$$R_{RS}(\lambda) = L(0, \lambda)/E_d(0, \lambda) \qquad (5)$$

where $L(0, \lambda)$ is the upwelling radiance from the sea-surface in the zenith direction at wavelength λ, and $E_d(0, \lambda)$ is the downwelling irradiance at the surface at the same wavelength. Remote-sensing reflectance has dimensions of inverse steradian (sr^{-1}). The chlorophyll concentration was first estimated using both the 510:555 and 490:670 ratios, assuming a diatom population. The difference between the 2 computed chlorophyll values was then estimated and normalised to the mean of the 2 values. The exercise was repeated, assuming a mixed phytoplankton population. The correct model should yield smaller differences in the concentrations retrieved using 2 different waveband pairs than should the wrong model. Therefore, each pixel was assigned to diatom or mixed population, on the basis of the computed normalised differences for the 2 cases.

Each of the scenes available for the first fortnight of April 2000 was analysed in the same way. Then, for each pixel, the information from all the images was used to compute the number of times the pixel was identified as being diatom-dominated. This number was divided by the total number of images with valid data at the pixel, to determine the probability that the location was indeed dominated by diatoms during the period investigated. The entire analysis was also repeated for the first fortnight of August 2000, as an example of a different season when the diatoms are not expected to be dominant.

To test the performance of the algorithm, we used all the HPLC chl *a* data for April from depths ≤ 10 m for which cloud-free SeaWiFS data was available for the same day and for the same geographic location. A total of 44 stations with matching *in situ* and satellite data was found for years 1998, 1999 and 2000 combined. The HPLC data from these stations were used to identify locations with diatom populations and those with mixed populations. The algorithm performance was evaluated by comparing the computed satellite results against *in situ* observations.

RESULTS

When comparing the model-based classification of phytoplankton populations against the classification of corresponding *in situ* stations by HPLC data, we found that the identifications were correct for 7 out of 10 stations in 1998 (agreement between model and observation for 70% of all cases tested). For the April 1998 data set, the model identifications of diatoms were correct for 6 out of 8 cases (success rate of 75%). The success rate for the whole data set was 72%. For the 1999 data, there were no matching observations of diatoms, but the mixed populations were identified correctly in 8 out of 9 cases. The HPLC data from 1998 were included in the database that was used to establish the phytoplankton absorption model, whereas the data sets for 1999 and 2000 constituted our independent data set for testing the model predictions. No degradation of results is evident for the independent data set: if anything, there is an improvement in the performance of the algorithms for 1999 and 2000. These results are very encouraging, especially when one considers the inherent limitations of the type of comparison attempted here: the *in situ* HPLC results are based on a small sample volume (typically 0.5 l), whereas the satellite pixel has a footprint of ca. 1 km². Furthermore, the matching of *in situ* and satellite data required only that both data be collected in the same day: the samples could therefore be separated from each other by several hours. Both these spatial and temporal disparities in the matched pairs may introduce differences in the results that are related more to the variabilities in the field and the dynamic nature of the field, rather than to the quality of the algorithm. Such mismatches in the time and space scales of sampling have to be borne in mind when evaluating this type of validation.

A qualitative approach to evaluating the results is to examine the distribution patterns for diatoms obtained using the algorithm. To this end we processed all the images available for the study area for the first fortnight of April 2000. In each image, every pixel was classified as diatom- or mixed-population dominated, using the discriminatory algorithm. Next, all the
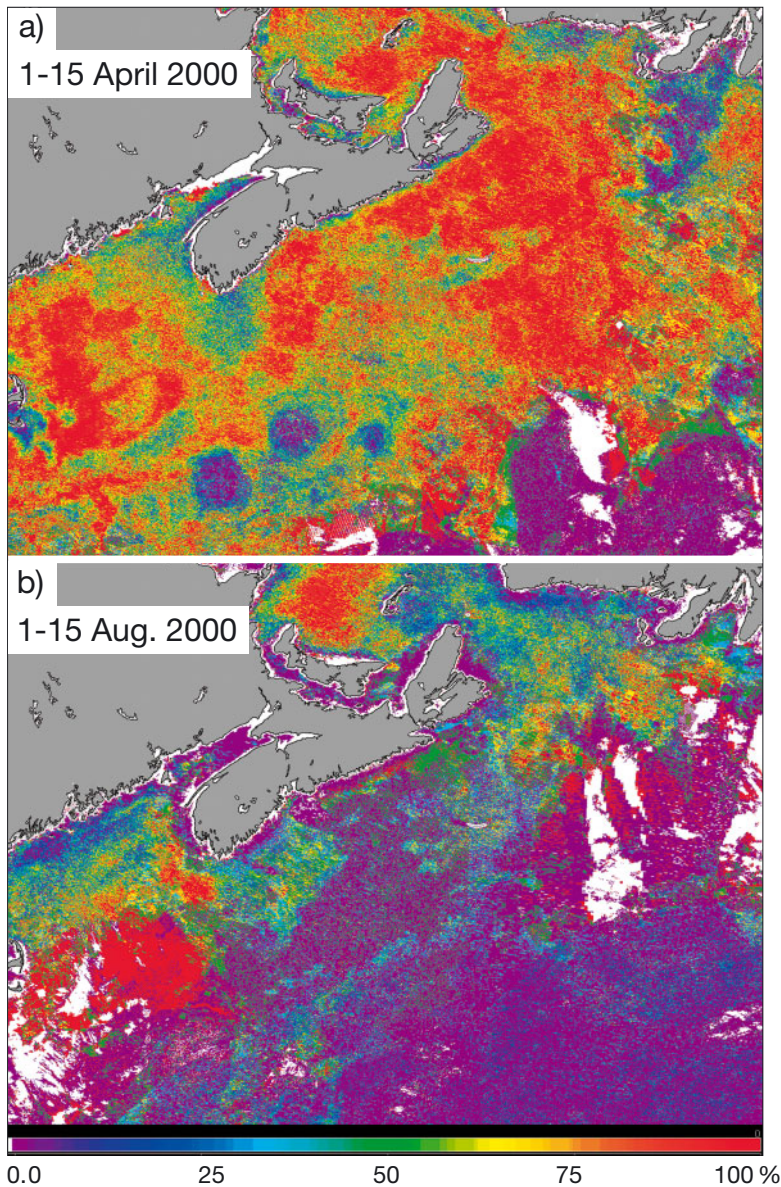
Fig. 4. Probability that the phytoplankton population at a particular location is dominated by diatoms. Results for the first fortnight of (a) April 2000 and (b) August 2000

probability maps show are consistent with what we know of the study area. In the April map, we see large areas with a high probability that the phytoplankton populations are dominated by diatoms. Notable exceptions are the Gulf Stream (bottom right-hand portion of the image), and the eddies that have spun away from the Gulf Stream. As noted earlier, diatoms are known to dominate in the region during the spring bloom, whereas the Gulf Stream is typically dominated by small phytoplankton. It is also encouraging that the diatom field has shrunk considerably in the August image, with a notable exception being the Georges Bank area. It has been reported (Hoepffner & Sathyendranath 1992, 1993) that diatoms dominate on Georges Bank in August, but not the slope or Gulf Stream waters.

Since the absorption characteristics of diatoms and other phytoplankton are sufficiently different that we are able to distinguish one from the other, it raises the question of whether the possibility exists for improving the retrieval accuracy of chl *a* concentrations by applying algorithms that are tuned for diatoms and mixed populations. This was done using a branching algorithm: if a pixel was identified as being dominated by diatoms, then the bio-optical model for diatoms was used to compute the chlorophyll concentration, and a mixed-population model was used for the remaining pixels. The satellite-derived chl *a* concentration is plotted against *in situ* chlorophyll in Fig. 5 for all the pixels for which we have overlapping data. The root-mean-square error for the OC4 algorithm is 51%, and that for the semi-analytic model tuned for diatoms and mixed populations is almost the same, at 52%. However, the performance of the semi-analytic algorithm improves considerably if we eliminate 3 points from the analysis. The root-mean-square error for the remaining 41 points then decreases to 31%. It is interesting to note that the 3 points with high errors (identified in Fig. 5) are all pixels that were mis-classified as diatom-dominated, when in fact the populations were mixed. This points to the potential that errors could increase for mis-classified pixels, whereas the errors would decrease typically for those pixels that were classified correctly. Note that Carder et al. (1999) had made a similar point about application of their regionally tuned semi-analytical algorithm to MODIS data.

images for the fortnight were used to obtain the number of times in which a given pixel was identified as being dominated by diatoms. This number was divided by the pixel depth for the fortnight (that is to say, the number of images with valid data at that pixel location for the fortnight). This yields a probability of occurrence of a diatom-dominated population at that pixel location over the particular time interval. A similar processing was done on all the images available for the first fortnight of August, 2000. The maps showing the probability that the populations observed were dominated by diatoms in the first fortnights of April and August 2000 are shown in Fig. 4. The patterns that the
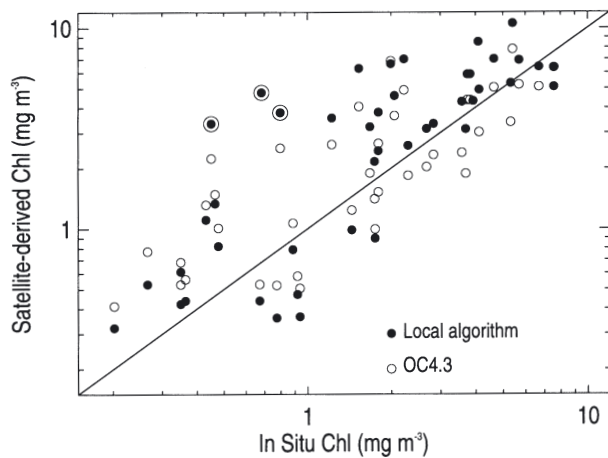
Fig. 5. Satellite-derived chlorophyll concentration as a function of *in situ* measured concentrations at the same pixel location and on the same day. Different symbols are used for results obtained using the NASA SeaWiFS OC4.3 algorithm (open circles), and for those obtained using the local branching algorithm based on the bio-optical properties of diatoms and mixed populations of phytoplankton (filled circles). The 3 misclassified pixels excluded from the computation of the mean root-square errors are encircled

Composite images of the chlorophyll concentrations for the first-fortnights of April and August 2000 computed using the branching bio-optical algorithm are shown in Fig. 6. Regional and seasonal algorithms to account for variations in the species composition of phytoplankton have been recommended in the past. In fact, for the North West Atlantic Zone, such algorithms have been shown to perform better than global algorithms (Devred et al. unpubl., Fuentes-Yaco et al. unpubl.). The discriminating, branching algorithm of the type presented in this paper has an advantage over a seasonal or regional approach, in that it yields a smoother transition between algorithms, and avoids sharp discontinuities in images such as those which occur in images made by shifting algorithms.

DISCUSSION AND CONCLUSION

Over the last decade or so, ocean-colour satellites have become increasingly sophisticated, with increased spectral resolution and signal-to-noise ratios. Such improvements open the door for the development of better algorithms for retrieval of chlorophyll concentrations. One of the avenues to achieving such improvements is to use branching algorithms tailored for specific groups of phytoplankton with their own characteristic optical properties. The approach we have used here is only the beginning of the development of such a school of algo-
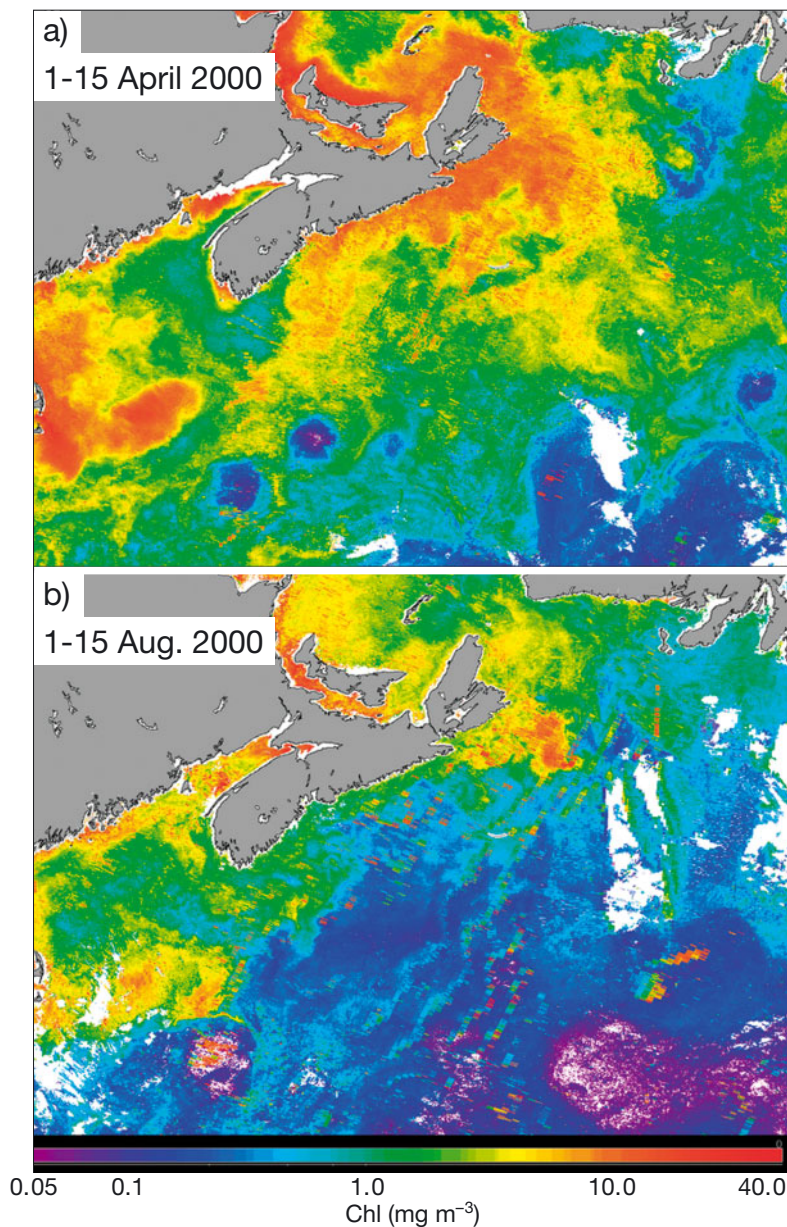

Fig. 6. Composite image of satellite-derived chlorophyll concentration for the North West Atlantic Zone. Computations based on the branching bio-optical algorithm for the region. Results for the first fortnight of (a) April 2000 and (b) August 2000

rithms. No doubt, as our understanding of the optics of ocean colour progresses, further refinements to the model can be incorporated. Algorithms have already been in effect for mapping the distribution of coccolithophore blooms (Brown & Podestá 1997), and algorithms for mapping *Trichodesmium* blooms have been proposed (Subramaniam et al. 1999, Subramaniam et al. 2002). The list will certainly grow as we understand the bio-optics of various types of phytoplankton better.

Not only do branching algorithms of the type used here have the potential to improve algorithms for chlorophyll retrieval, but they also yield additional valuable information on the distribution of major phytoplankton groups: for the first time ever, it has now become possible to produce synoptic maps of distribution of diatoms. We have used SeaWiFS data for the analysis presented here. Sensors with higher spectral resolution will certainly allow further refinements to the algorithm in the future. But for now, SeaWiFS, with its multi-year global database and easy accessibility, remains the best choice for implementation of such algorithms.

The algorithm has been implemented and tested for the North West Atlantic Zone. We are aware that diatoms from other geographic areas exhibit optical properties that differ from those of the NWAZ (Stuart et al. 2004). Caution is therefore essential before the algorithm can be safely transferred to other regions: fine-tuning of the model would no doubt be essential before the model can be adapted for implementation elsewhere.

LITERATURE CITED

Babin M, Therriault JC, Legendre L, Condal A (1993) Variations in the specific absorption coefficient for natural phytoplankton assemblages: impact on estimates of primary production. Limnol Oceanogr 38:154–177

Bidigare RR, Ondrusek ME, Morrow JH, Kiefer DA (1990) *In vivo* absorption properties of algal pigments. Proc SPIE Ocean Opt X 1302:290–302

Bird RE (1984) A simple, solar spectral model for direct-normal and diffuse horizontal irradiance. Solar Energy 32: 461–471

Bouman HA (2003) Relationship between community structure and the bio-optical properties of marine phytoplankton. PhD dissertation, Dalhousie University, Halifax

Bouman HA, Platt T, Sathyendranath S, Li WKW and 7 others (2003) Temperature as indicator of optical properties and community structure of marine phytoplankton: implica-

tions for remote sensing. Mar Ecol Prog Ser 258:19–30

Bricaud A, Morel A, Prieur L (1981) Absorption by dissolved organic matter of the sea (yellow substance) in the UV and visible domains. Limnol Oceanogr 26:43–53

Bricaud A, Babin M, Morel A, Claustre H (1995) Variability in the chlorophyll-specific absorption coefficients of natural phytoplankton: analysis and parameterization. J Geophys Res 100:13321–13332

Brown CW, Podestá GP (1997) Remote sensing of Coccolithophore blooms in the Western South Atlantic Ocean. Remote Sens Environ 60:83–91

Bukata RP, Jerome JH, Bruton JE, Jain SC, Zwick HH (1981) Optical water quality model of Lake Ontario. 1. Determination of the optical cross sections of organic and inorganic particulates in Lake Ontario. Appl Optics 20: 1696–1703

Carder KL, Chen FR, Lee ZP, Hawes SK, Kamykowski D (1999) Semianalytic moderate-resolution imaging spectrometer algorithms for chlorophyll *a* and absorption with bio-optical domains based on nitrate-depletion temperatures. J Geophys Res 104:5403–5421

Culver ME, Perry MJ (1999) The response of photosynthetic absorption coefficients to irradiance in culture and in tidally mixed estuarine waters. Limnol Oceanogr 44:24–36

Duysens LNM (1956) The flattening of the absorption spectrum of suspensions, as compared to that of solutions. Biochim Biophys Acta 19:1–12

Gallegos CL, Correll DL, Pierce JW (1990) Modeling spectral diffuse attenuation, absorption, and scattering coefficients in a turbid estuary. Limnol Oceanogr 35:1486–1502

Head EJH, Horne EPW (1993) Pigment transformation and vertical flux in an area of convergence in the North Atlantic. Deep-Sea Res II 40:329–346

Hoepffner N, Sathyendranath S (1991) Effect of pigment composition on absorption properties of phytoplankton. Mar Ecol Prog Ser 73:11–23

Hoepffner N, Sathyendranath S (1992) Bio-optical characteristics of coastal waters: absorption spectra of phytoplankton and pigment distribution in the western North Atlantic. Limnol Oceanogr 37:1660–1679

Hoepffner N, Sathyendranath S (1993) Determination of the major groups of phytoplankton pigments from the absorption spectra of total particulate matter. J Geophys Res 98: 22789–22803

Johnsen G, Sakshaug E (1993) Bio-optical characteristics and photoadaptive responses in the toxic and bloom-forming dinoflagellates *Gyrodinium aureolum*, *Gymnodinium galatheanum*, and two strains of *Prorocentrum minimum*. J Phycol 29:627–642

Kirk JTO (1975) A theoretical analysis of the contribution of algal cells to the attenuation of light within natural waters. I. General treatment of suspensions of pigmented cells. New Phytol 75:11–20

Kirk JTO (1976) A theoretical analysis of the contribution of algal cells to the attenuation of light within natural waters. III. Cylindrical and spheroidal cells. New Phytol 77: 341–358

Kirk JTO (1994) Light and photosynthesis in aquatic ecosystems. Cambridge University Press, Cambridge

Loisel H, Morel A (1998) Light scattering and chlorophyll concentration in case 1 waters: a reexamination. Limnol Oceanogr 43:847–858

Marra J, Trees CC, Bidigare RR, Barber RT (2000) Pigment absorption and quantum yields in the Arabian Sea. Deep-Sea Res II 47:1279–1299

Mitchell BG, Kiefer DA (1988) Variability in pigment specific particulate fluorescence and absorption spectra in the

northeastern Pacific Ocean. Deep-Sea Res I 35:665–689

Morel A (1974) Optical properties of pure seawater. In: Jerlov NG, Steemann Nielsen E (eds) Optical aspects of oceanography. Academic Press, New York, p 1–24

Morel A (1980) In-water and remote measurement of ocean color. Boundary-Layer Meteorol 18:177–201

Morel A (1997) Consequences of a *Synechococcus* bloom upon the optical properties of oceanic (case 1) waters. Limnol Oceanogr 42:1746–1754

Morel A, Bricaud A (1981) Theoretical results concerning light absorption in a discrete medium, and application to specific absorption of phytoplankton. Deep-Sea Res I 28A:1375–1393

O'Reilly JE, Maritorena S, Mitchell BG, Siegel DA, Carder KL, Garver SA, Kahru M, McClain C (1998) Ocean color chlorophyll algorithms for SeaWiFS. J Geophys Res 103: 24937–24953

Platt T, Jassby AD (1976) The relationship between photosynthesis and light for natural assemblages of coastal marine phytoplankton. J Phycol 12:421–430

Platt T, Sathyendranath S (1988) Oceanic primary production: estimation by remote sensing at local and regional scales. Science 241:1613–1620

Pope RM, Fry ES (1997) Absorption spectrum (380–700 nm) of pure water: II. Integrating cavity measurements. Appl Optics 36:8710–8723

Roesler CS, Perry MJ, Carder KL (1989) Modeling in situ phytoplankton absorption from total absorption spectra in productive inland marine waters. Limnol Oceanogr 34: 1510–1523

Sathyendranath S, Platt T (1997) Analytic model of ocean color. Appl Optics 36:2620–2629

Sathyendranath S, Platt T (1998) Ocean-colour model incorporating transspectral processes. Appl Optics 37: 2216–2227

Sathyendranath S, Lazzara L, Prieur L (1987) Variations in the spectral values of specific absorption of phytoplankton.

Limnol Oceanogr 32:403–415

Sathyendranath S, Stuart V, Irwin BD, Maass H, Savidge G, Gilpin L, Platt T (1999) Seasonal variations in bio-optical properties of phytoplankton in the Arabian Sea. Deep-Sea Res 46:633–654

Sathyendranath S, Stuart V, Cota G, Maass H, Platt T (2001) Remote sensing of phytoplankton pigments: a comparison of empirical and theoretical approaches. Int J Remote Sensing 22:249–273

Sogandares FM, Fry ES (1997) Absorption spectrum (340–640 nm) of pure water: I. Photothermal measurements. Appl Optics 36:8699–8709

Sosik HM, Mitchell BG (1994) Effects of temperature on growth, light absorption, and quantum yield in *Dunaliella tertiolecta* (Chlorophyceae). J Phycol 30:833–840

Stuart V, Sathyendranath S, Platt T, Maass H, Irwin BD (1998) Pigments and species composition of natural phytoplankton populations: effect on the absorption spectra. J Plankton Res 20:187–217

Stuart V, Sathyendranath S, Head EJH, Platt T, Irwin B, Maass H (2000) Bio-optical characteristics of diatom and prymnesiophyte populations in the Labrador Sea. Mar Ecol Prog Ser 201:91–106

Stuart V, Ulloa O, Sathyendranath S, Platt T, Alarcón G, Major H, Head EJH (2004) Characteristics of phytoplankton populations in the upwelling system off the coast of Chile. Rev Chilena Hist Nat (in press)

Subramaniam A, Carpenter EJ, Falkowski PG (1999) Bio-optical properties of the marine diazotrophic cyanobacteria *Trichodesmium* spp. II. A reflectance model for remote sensing. Limnol Oceanogr 44:618–627

Subramaniam A, Brown CW, Hood RR, Carpenter EJ, Capone DG (2002) Detecting *Trichodesmium* blooms in SeaWiFS imagery. Deep-Sea Res II 49:107–121

Ulloa O, Sathyendranath S, Platt T (1994) Effect of the particle-size distribution on the backscattering ratio in seawater. Appl Optics 33:7070–7077