

## SeaWiFS discrimination of harmful algal bloom evolution

PETER I. MILLER\*, JAMIE D. SHUTLER, GERALD F. MOORE and  
STEVE B. GROOM

Remote Sensing Group, Plymouth Marine Laboratory, Prospect Place, Plymouth PL1  
3DH, UK

The discrimination of harmful algal blooms (HABs) from space would benefit both the capability of early warning systems and the study of environmental factors affecting the initiation of blooms. Unfortunately, there are no published techniques using global monitoring satellite sensors to distinguish the resulting subtle changes in ocean colour, so *in situ* sampling is needed to identify the species in any observed bloom. This paper investigates multivariate classification as an objective means to discriminate harmful and harmless algae and monitor their dynamics using ocean colour data derived from satellite sensors. The classifier is trained and tested using Sea-viewing Wide Field-of-view Sensor (SeaWiFS) data, though the method is designed to be generic for other sensors. Time-series results are presented using the new HAB likelihood index and suggest that SeaWiFS has some capability for observing the dynamic evolution of harmful blooms of *Karenia mikimotoi*, *Chattonella verruculosa* and cyanobacteria. Further, a multi-band spatial subtraction algorithm is described to automate the identification of bloom regions and improve the accuracy in discriminating HABs.

### 1. Introduction

Harmful algal blooms (HABs) are believed to be increasing in occurrence around the world (Anderson *et al.* 2002). Algal toxins can be concentrated by filter-feeding shellfish and cause amnesia or paralysis when ingested (Van Dolah 2000). Blooms also result in economic losses through closure of shellfish grounds, direct damage to fish farms, or contamination of tourist beaches. Before remedial action can be considered, research is needed to investigate initiation mechanisms. One such project, funded through the European Commission Framework 5 programme, is Harmful Algal Bloom Initiation and Prediction in Large European Marine Ecosystems (HABILE): this utilizes *in situ* measurements, coupled biological and physical models and Earth observation techniques to investigate environmental factors affecting HAB initiation and development.

HABs may comprise a small fraction of the phytoplankton ensemble or may dominate the flora, developing very high biomass or chlorophyll concentration that can be observed from satellite sensors, such as the Sea-viewing Wide Field-of-view Sensor (SeaWiFS). In European waters high biomass HAB species observed from satellite sensors include *Karenia mikimotoi* (figure 1(b), Holligan *et al.* 1983; Raine

---

\*Corresponding author. Email: pim@pml.ac.uk

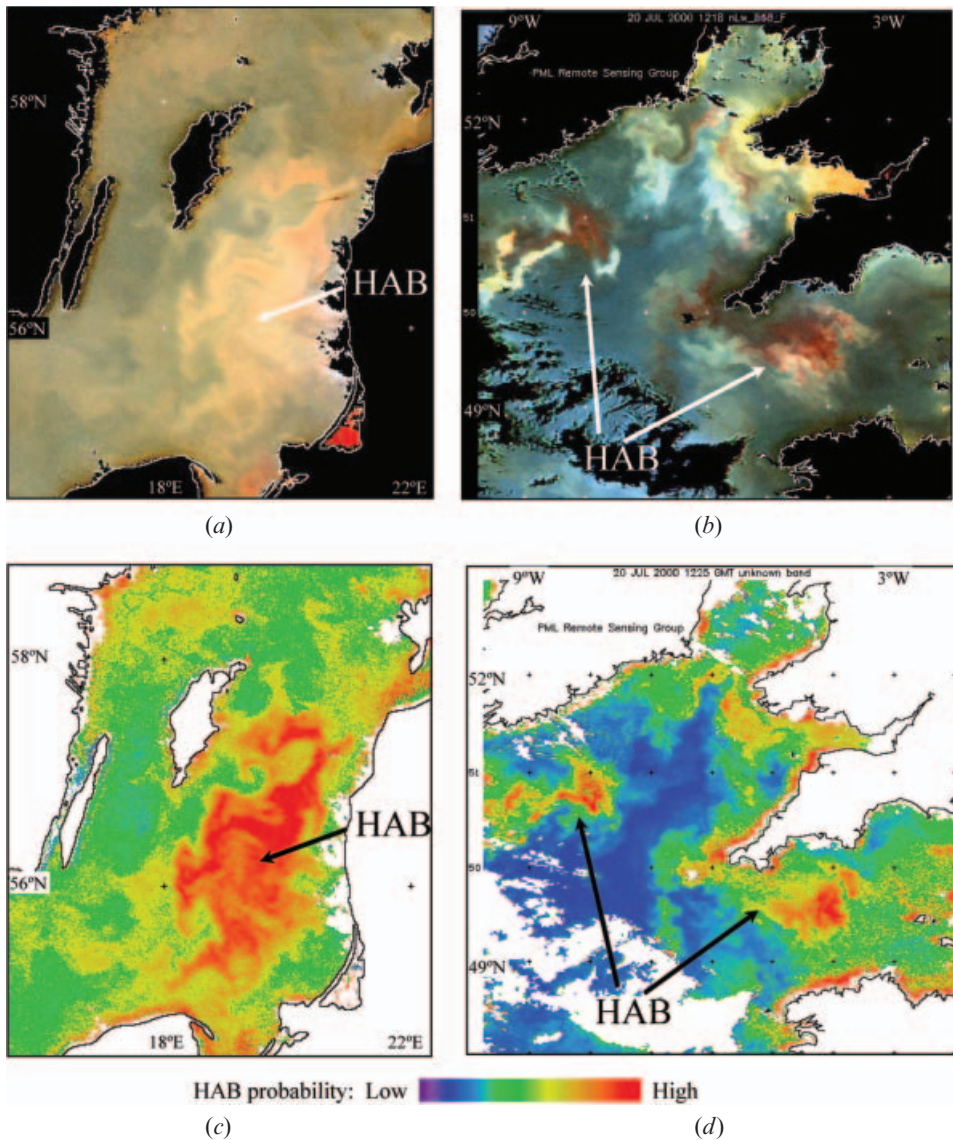


Figure 1. Example SeaWiFS contrast-enhanced false colour composites (555, 510 and 443 nm) of HABs: (a) Baltic Sea 5 June 2002 1128 UTC, with cyanobacteria bloom; (b) UK SW approaches 20 July 2000 1225 UTC, showing two *Karenia* blooms; (c)–(d) corresponding HAB likelihood maps.

*et al.* 2001), *Chattonella* spp. (Pettersson *et al.* 2001) and cyanobacteria (figure 1(a), Kahru *et al.* 2000). SeaWiFS data have already been integrated into an existing HAB monitoring system called ALGEINFO operating along the Norwegian coast (Pettersson *et al.* 2001) and, through visual interpretation, have enabled observation of the large-scale development of *Chattonella* blooms in 1998 and 2000. Furthermore, in 2001 SeaWiFS provided the first observation of the bloom (prior to *in situ* sampling). However, a weakness with this exercise was that it was not possible to differentiate harmful from non-harmful species: this required *in situ*

identification. The aim of this paper is to investigate automatic methods of objectively discriminating harmful and harmless algae using satellite ocean colour.

Multivariate analysis offers an objective approach for deriving ocean colour algorithms, based upon the observed characteristics of training samples. This approach was adopted to identify effective classifiers of the target HAB species using SeaWiFS ocean colour parameters. Promising results are described for manually training the classifier and its application to monitor dynamic ecosystem changes in time-series data. Further, the results of a multi-band spatial approach and how this could be used to produce a hybrid spectral/spatial method are presented.

## 2. Methods

Optical approaches to HAB monitoring and identification based on *in situ* hyperspectral measurements have been reviewed by Schofield *et al.* (1999) and these authors described one method of recognizing *Karenia brevis*. Liew *et al.* (2000) discriminated HAB species by the shape of their reflectance spectrum measured from a ship, and demonstrated the potential application of SeaWiFS using simulated radiance values. Application of similar methods to enable HAB discrimination from space requires high spectral resolution sensors. The SeaWiFS sensor provides ocean colour data in six spectral bands, and the National Aeronautics & Space Administration (NASA) MODerate resolution Imaging Spectroradiometer (MODIS) carried on-board the Aqua and Terra platforms measures ocean colour in nine of its 36 spectral wavebands. Further ocean colour sensors include the European Space Agency (ESA) MEdium Resolution Imaging Spectrometer (MERIS) and the Compact High Resolution Imaging Spectrometer (CHRIS), both of which sample data in more than 12 spectral bands.

In order to determine water constituents such as HABs, the absorption and backscatter of light in each band must be retrieved first. Smith *et al.* (submitted) developed an inherent optical properties (IOP) model, based on the slope of spectral signatures, producing estimates of absorption  $a(\lambda)$  and backscatter  $b_b(\lambda)$  from SeaWiFS water-leaving radiances  $L_{wn}(\lambda)$ . The method is applicable to turbid Case 2 water essential to the study of HABs in coastal regions; it is also less computationally expensive than other models requiring matrix inversion solutions (e.g. Hoge *et al.* 1999). This model has been integrated into the Plymouth Marine Laboratory SeaWiFS ocean colour processing system (Lavender and Groom 1999), to enable generation of absorption and backscatter maps. The system is based on SeaWiFS Data Analysis System (NASA) SeaDAS modules with the addition of the 'bright pixel' atmospheric correction designed for Case 2 water (Moore *et al.* 1999). Final products are mapped to the Mercator projection with a spatial resolution of approximately 1.1 km.

### 2.1 Multivariate discrimination of HABs

The Fisher (F) statistic can be used to rate the potential of a variable or combination of variables to discriminate between classes. This statistic is the ratio of the between-class variance to the within-class variance, so a high value ( $F \gg 1$ ) signifies potential differences between classes. Linear discriminate analysis (Flury and Riedwyl 1988) is an established statistical technique used to separate classes within a dataset, using matrix inversion to determine the optimal linear combination of variables that maximize the F-value. This linear combination is then used as a function to

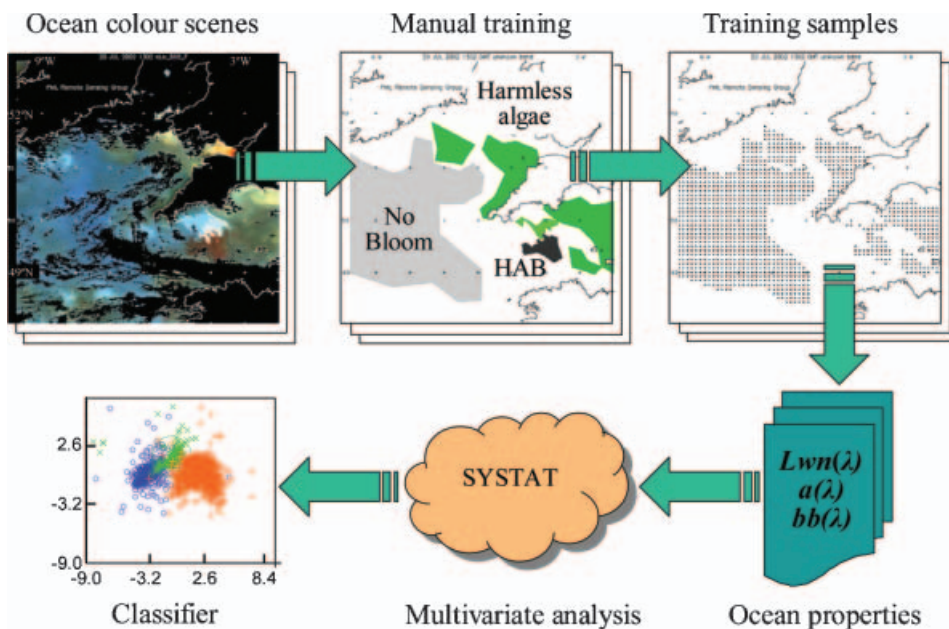


Figure 2. Schematic diagram of the training procedure for the multivariate HAB classifier.

discriminate further cases. This paper converts F-values into standard distance, which is the distance between class means divided by the pooled standard deviation (Flury and Riedwyl 1988). This provides an effective method to compare the discriminating power both of individual variables and of multivariate classifiers.

To develop the linear discriminant classifier (explained schematically in figure 2), a small training set of five SeaWiFS scenes were used, containing blooms confirmed by contemporaneous *in situ* sampling (top 5 rows of table 1) and also representative of mid-bloom conditions, as these are usually straightforward to delineate. Each

Table 1. Training (top five) and independent testing scenes used to develop multivariate classifier.

Region	Date	Time UTC	Bloom types
UK Southwest Approaches	20 July 2000	1225	<i>Karenia mikimotoi</i> <sup>1</sup>
UK Southwest Approaches	20 July 2002	1302	<i>Karenia mikimotoi</i> <sup>2</sup> , coccolithophores
North Sea	11 May 2000	1229	<i>Chattonella verruculosa</i> <sup>3</sup>
Baltic Sea	5 June 2002	1128	Cyanobacteria <sup>4</sup>
Baltic Sea	16 July 2002	1155	Cyanobacteria <sup>4</sup>
Celtic Sea, south of Ireland	4 August 1998	1346	<i>Karenia mikimotoi</i> <sup>5</sup>
Celtic Sea, south of Ireland	9 August 1998	1234	<i>Karenia mikimotoi</i> <sup>5</sup>
Celtic Sea, south of Ireland	15 August 1998	1344	<i>Karenia mikimotoi</i> <sup>5</sup>
North Sea	15 May 1998	1239	<i>Chattonella verruculosa</i> <sup>6</sup>
North Sea	2 June 1998	1254	<i>Chattonella verruculosa</i> <sup>6</sup>
North Sea	21 June 1998	1215	<i>Chattonella verruculosa</i> <sup>6</sup>
Baltic Sea	6 July 2001	1113	Cyanobacteria <sup>7</sup>

References to *in situ* validation of bloom types: 1, Groom *et al.* (2000); 2, Kelly-Gerrey *et al.* (2004); 3, Pettersson *et al.* (2003); 4, Rantajarvi (2003); 5, Raine *et al.* (2001); 6, Aure *et al.* (2000); 7, Hassink (2004).

training image was labelled manually with HAB, harmless and no-bloom regions, according to the published surveys and related ocean colour characteristics. To reduce interdependencies in the training data, the labelled areas were subsampled to separate training points by  $8 \times 8$  pixels, giving a total of 4090 training samples spread across the three categories. Each of the different radiance, absorption and backscatter parameters were then entered as variables to train the classifier, using stepwise discriminant analysis in the SYSTAT statistical package (supplied by SPSS Inc.). To reduce redundancy in these variables, SeaWiFS ocean products such as chlorophyll-*a* (chl-*a*) and  $K_d(490)$  were excluded as these are derived from radiance. To gain an unbiased estimate of classification performance, the classifiers were then applied to seven independent testing scenes, again with bloom species confirmed by *in situ* sampling (bottom 7 rows of table 1). Although the idea of a single HAB classifier is attractive, it was soon identified that the different characteristics of each species prevented reliable discrimination between HABs and non-harmful algae. The method was revised to generate a separate classifier for each target species, each based on sample blooms of that type, together with all training samples for harmless and non-bloom cases. This approach supports the hypothesis that the different algal species will have spectrally different responses. Example coefficients for these HAB classification functions can be found at [http://www.npm.ac.uk/papers/miller\\_ijrs2006/](http://www.npm.ac.uk/papers/miller_ijrs2006/).

The classifiers can be used to segment an image into regions most similar to the training categories. However, this will not respond to the early and subtle signs of developing HABs, so instead, the HAB classification function was used to derive a continuous scale of HAB likelihood. The following section explores how the detection of HABs can be automated and improved further through the use of spatial and spectral image analysis techniques.

## 2.2 Bloom detection using median subtraction techniques

Subtraction of the time-averaged background is a simple but effective way of determining areas within a scene that have changed, i.e. to detect foreground information. One system using this principle operates in the Gulf of Mexico, USA, for detecting *Karenia brevis* blooms with a combination of satellite, field and meteorological data (Stumpf *et al.* 2003). Rolling estimates of the mean background chl-*a* levels are determined and assumed to be HAB-free. These background estimates are then used to detect localized increases in chl-*a*, depicting areas of further interest, e.g. for guiding *in situ* sampling.

This approach was developed further to follow seasonal effects more closely and reduce the effects of high sediment. The use of the median, instead of the mean used by Stumpf *et al.* (2003), relaxes the need for a gap between the end of the time series and the image to be subtracted. This is because the median will be less affected by anomalies occurring within the time series and allows for seasonal changes to be followed more closely. The temporal median background image is calculated using data from the two months prior to the image of interest. Subtracting the image of interest from the background produces a difference map. Only those pixels that have increased in value by  $\geq 10\%$  of their background value are then retained. This process is repeated for chl-*a* and water-leaving radiances at 490 nm and 555 nm, producing three anomaly maps per satellite overpass:  $\Delta\text{chl-}a$ ,  $\Delta L_{\text{wn}}(490)$  and  $\Delta L_{\text{wn}}(555)$ . Figure 3 shows an example run of this subtraction technique on a HAB image for the UK Southwest Approaches.

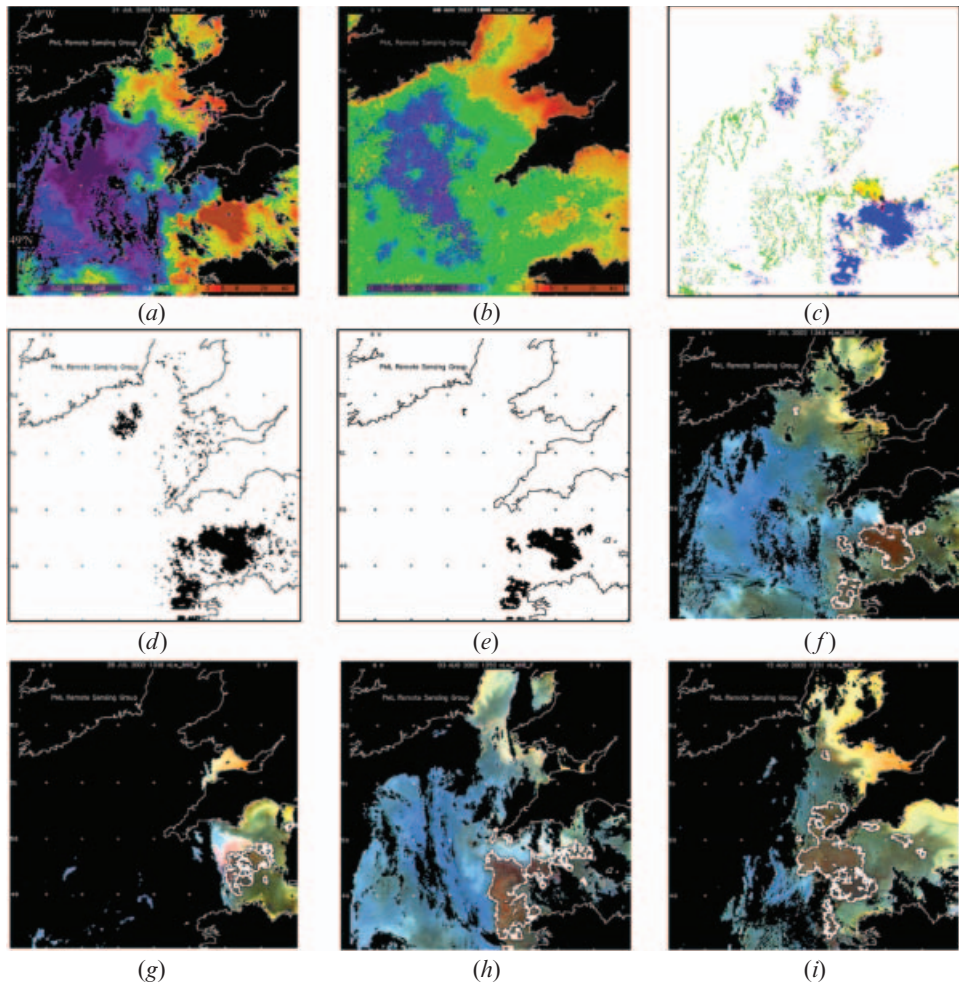


Figure 3. Example results illustrating the different stages of the spatial subtraction method: (a) SeaWiFS chl-*a* 21 July 2002; (b) median chl-*a* for the two previous months; (c) anomalous areas coloured red –  $L_{wn}(490)$ , green –  $L_{wn}(555)$ , blue – chl-*a*, or colour combinations; (d) areas with chl-*a* anomaly but no sediment anomaly; (e) following analysis of connected regions; (f) anomalies outlined in white on SeaWiFS enhanced true-colour image. Further example results from (g) 28 July; (h) 3 August and (i) 12 August.

As this study is specifically interested in chl-*a* anomalies, areas of high sediment that cause the SeaWiFS algorithm to give erroneous chl-*a* concentrations must be removed. Sediment is known to scatter light at 490 and 555 nm (Moore *et al.* 1999). Therefore, a mask image is created based on the chl-*a* anomaly map, using  $\Delta L_{wn}(490)$  and  $\Delta L_{wn}(555)$  to remove the effects of sediment anomalies. Each mask pixel  $P(x)$  is defined as:

$$P(x) = \begin{cases} 1 & \text{if } \Delta \text{chl-}a > 0 \\ 0 & \text{if } \Delta \text{chl-}a > 0, \Delta L_{wn}(490) > 0 \\ 0 & \text{if } \Delta \text{chl-}a > 0, \Delta L_{wn}(555) > 0 \end{cases} \quad (1)$$

### 3. Results

#### 3.1 Multivariate discrimination

The 2D scatterplot of the discriminant space for each species classifier shows promising separation between the three classes (figure 4). These plots depict the first two orthogonal linear combinations of input variables that best discriminate among the classes. In each case, the discriminant analysis identified a highly significant (Wilks' Lambda test,  $p < 0.001$ ) separation between the three classes. Experiments were used to test the effect of the bright pixel atmospheric correction and IOP model products on discrimination performance. The standard distance between class means for *Karenia*, *Chattonella* and cyanobacteria classifiers using  $L_{\text{wn}}(\lambda)$  derived using the SeaDAS v4.3 atmospheric correction (Gordon and Wang 1994) were 3.2, 2.4 and 2.6, respectively. Using the bright-pixel atmospheric correction to derive  $L_{\text{wn}}(\lambda)$  these distances became 4.1, 3.0 and 2.4; only cyanobacteria discrimination was not improved due to the unpredictable effect of surface accumulations on the colour signal. Adding the IOP model products resulted in further significant improvements, to 5.3, 5.5 and 3.1. These results justify the use of the bright pixel atmospheric correction and IOP model products in the final classifier, which is tested in all subsequent results.

The robustness of the classifier is best visualized by applying the HAB classification function to every pixel in a scene, to view the varying confidence in the presence of HABs across the region. Examples of HAB likelihood maps are presented in figures 1(c)–(d) for two of the training scenes. These show clear detection of cyanobacteria in the Baltic Sea, and *Karenia* in the English Channel and Celtic Sea, though false alarms are generated in areas of known high sediment, such as the Bristol Channel. The bloom off Cornwall was sampled from a research vessel on 25 July 2000 and found to contain *Karenia* in concentrations over 10 000 cells/ml, mixed with the bioluminescent *Noctiluca scintillans* (Groom *et al.* 2000).

Classification accuracy on training data is presented quantitatively using a classification matrix for each species (table 2). The top section of this table refers to

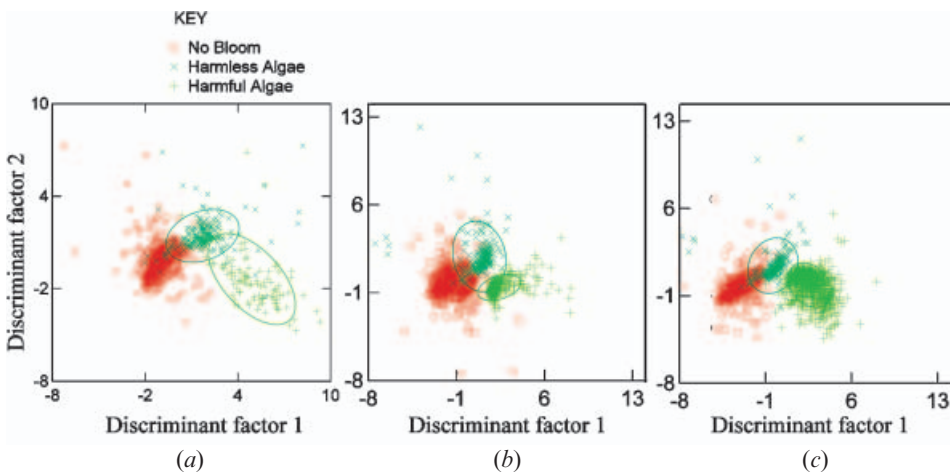


Figure 4. Scatterplots of discriminant space (arbitrary units), showing promising separation between HAB, harmless algae and no bloom pixels: (a) *Karenia mikimotoi*; (b) *Chattonella* spp.; (c) cyanobacteria.

Table 2. Classification accuracy on training set of 2400 samples from five SeaWiFS scenes.

	No bloom	Harmless	HAB	Total	% correct
No bloom	874	4	5	883	99
Harmless	33	153	21	207	74
<i>Karenia</i>	0	9	118	127	93
Total				1217	94
No bloom	942	33	54	1029	92
Harmless	23	138	6	167	83
<i>Chattonella</i>	17	3	391	411	95
Total				1607	92
No bloom	823	27	1	851	97
Harmless	20	183	3	206	89
Cyanobacteria	1	6	725	732	99
Total				1789	97

The actual class is indicated at the start of each row, and the classes predicted by the classifier are represented by columns. Shaded squares indicate the number of samples correctly classified. Results are shown separately for the three species classifiers.

samples used to train the *Karenia* classifier: the true classes are indicated at the start of each row, and the classes predicted by the classifier represented by columns. An ideal classifier will show all samples along the (shaded) diagonal, with no misclassifications in the remaining cells. Of the 127 *Karenia* samples, 118 were correctly identified, 9 classed as harmless algae and 0 as no-bloom; in addition, 26 of the 1090 harmless or no-bloom samples were classed as harmful. This represents a hit rate of 93% and false alarm rate of 2%. For *Chattonella*, a hit rate of 95% and false alarm rate of 5% was achieved. For cyanobacteria, a hit rate of 99% and false alarm rate of <1%. Note that these totals exclude a substantial number of samples for which one or more variables contained missing values; this problem is discussed below. These hit rates resulted from classifying each species individually, but an attempt was also made to identify the type of bloom by testing which of the three classifiers gave the highest HAB likelihood on each sample. *Karenia* was identified correctly on 82% of the samples, *Chattonella* 71%, and cyanobacteria 37%; overall the correct species was identified for 47% of the HAB samples. These are promising results but require further work on normalizing the classification functions for different species.

Table 3 presents the performance on independent test data, giving an unbiased validation of the HAB classifiers. For *Karenia*, the hit rate was 83%, with <1% false alarms. The *Chattonella* and cyanobacterial hit rates remained high (93% and 98%), with few false alarms (6% and <1%).

Comparing the discriminant power of parameters in each classifier reveals those that are most important for detecting particular species. This is shown using the change in multivariate standard distance that results from removing each parameter in turn from the classifier (figure 5). The methodology of training the classifier separately on each HAB species is justified by the varying importance of each ocean colour parameter. For *Karenia*, the most important parameters are  $L_{wn}(670)$  – due to its reddish appearance, and  $a(443)$  – illustrating the strong absorption of chl-*a* at 443 nm. For *Chattonella*,  $a(443, 490, 510)$  appear most significant, as would be expected for this strongly absorbing species. Cyanobacteria are discriminated by the increased backscatter in  $L_{wn}(490, 510, 670)$  due to the surface accumulation of this type of HAB (Kahru *et al.* 2000).

Table 3. Classification accuracy on independent testing set of 1600 samples from seven SeaWiFS scenes.

	No bloom	Harmless	HAB	Total	% correct
No bloom	554	1	1	556	100
Harmless	0	18	1	19	95
<i>Karenia</i>	12	15	136	163	83
Total				738	96
No bloom	538	11	36	585	92
Harmless	21	63	4	88	72
<i>Chattonella</i>	10	3	184	197	93
Total				870	90
No bloom	550	1	0	551	100
Harmless	0	18	1	19	95
Cyanobacteria	0	13	615	628	98
Total				1198	99

The actual class is indicated at the start of each row, and the classes predicted by the classifier are represented by columns. Shaded squares indicate the number of samples correctly classified. Results are shown separately for the three species classifiers.

It is interesting to note that  $L_{wn}(670)$  has some significance in all three classifiers, supporting previous work showing the importance of the 640–670 nm region of the spectrum for characterization of certain algal species from ship-borne optical measurements (Millie *et al.* 1997). The IOP  $a(443)$  is related to chl-*a* concentration and so contributes to the *Karenia* and *Chattonella* classifiers, and several other IOP products show greater discriminating power than their  $L_{wn}$  counterparts. However, it appears that the atmospheric correction and IOP model do not always cope with the full range of absorption and backscatter experienced in these dense blooms, and the resulting saturated and missing values reduce the training and testing sets.

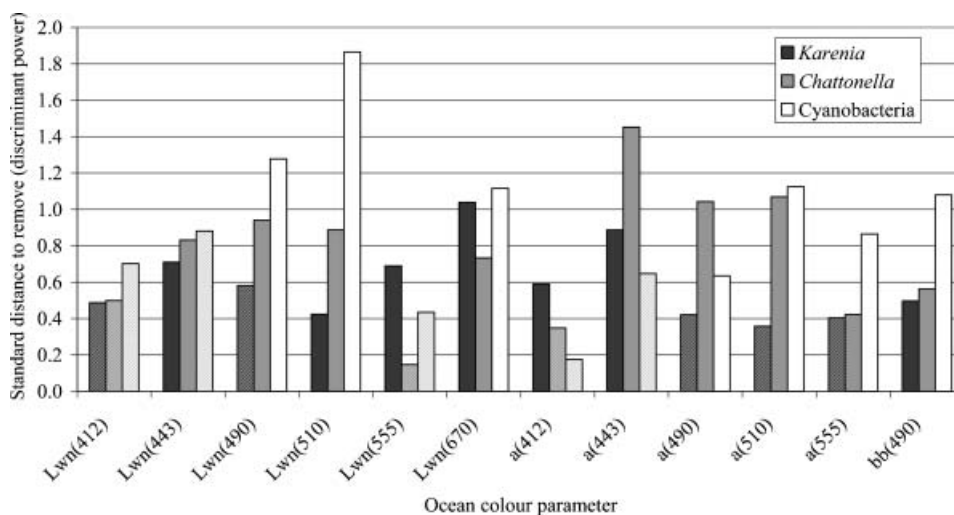


Figure 5. Relative importance of each SeaWiFS ocean colour parameter in the HAB classifiers. The bars hatched with diagonal lines represent variables identified as redundant in the stepwise discriminant analysis and excluded from the final classifier.

### 3.2 Time-series analysis of HAB evolution

Following development, the classifier can now be applied to the focus of this special issue: to study the dynamic nature of the initiation and evolution of HABs. Four time-series of SeaWiFS data were processed to generate HAB likelihood maps spanning the growth and decline of the blooms. As the training data are representative of mid-bloom conditions, the HAB likelihood index should depict the life cycle of the bloom, peaking during the period of highest abundance. These results also serve to test the robustness of the classifier on many sample images and blooms independent of the training set. Figure 6(a) presents selected HAB likelihood maps from the summer 2002 time-series generated using the *Karenia* classifier. This visualization matches well with qualitative analysis of the corresponding ocean colour scenes and with sampling conducted to validate FerryBox monitoring (Kelly-Gerreyn *et al.* 2004). The first scene on 24 June is prior to the bloom formation, though by 3 July there is a widespread increase in HAB likelihood indicating possible early bloom conditions. The gaps in cloud cover on 14 July reveal scattered HAB 'hot spots'. Comparing the 20 July image with the corresponding enhanced ocean colour scene (shown top-left in figure 2) it can be seen that both the true *Karenia* bloom and a false alarm for a coccolithophore bloom are assigned high HAB likelihood values. The false alarm is due to the considerable scattering by coccolithophores observed in  $L_{wn}(555)$  (Gordon *et al.* 2001). However, the other discriminant functions for this scene (not shown) indicate an even higher likelihood of 'harmless algae' for the coccolithophores. So this method offers the potential for detailed analysis to distinguish pixels on the boundary of two categories.

Between 28 July and 3 August the first bloom diminishes but a further bloom (suspected but not proven to be *Karenia*) begins to develop further west at the Channel entrance, peaking on 20 August before fading rapidly by 24 August following two cloudy days. Around this time, *Karenia* further east caused the temporary closure of shellfisheries in the Le Havre region of northern Brittany (Francis Gohin, 2002, personal communication).

This evolution of the bloom may also be studied by plotting changes over time in the range of HAB likelihood values contained within a specific region (figure 6(a) graph), in this case the small area labelled as harmful on the 20 July mid-bloom training image (centred near 49.5°N 4°W, shown in figure 2). The graph shows strong peaks associated with the middle of the bloom (Julian day 201–209) and renewed growth (day 232), as previously identified on the maps.

Figure 6(b) presents the evolution of cyanobacteria blooms in the Baltic Sea during summer 2002. The graph was generated for the large area covered by the bloom on 16 July. The white streaks on this scene are due to SeaDAS algorithm failures because of extremely high algal concentrations; this image is actually 100% cloud-free. It is interesting to note how quickly the situation changes from early to mid-HAB between 10 and 16 July (Julian day 191–197). The 30 July scene (day 211) is representative of late bloom conditions.

### 3.3 Background subtraction

The background subtraction algorithm was applied to a two-month sequence covering the development of an intense *Karenia mikimotoi* bloom in the English Channel during summer 2002. Figures 3(g)–(i) show three further subtraction results in the sequence following that in figure 3(f), and demonstrate that the bloom is

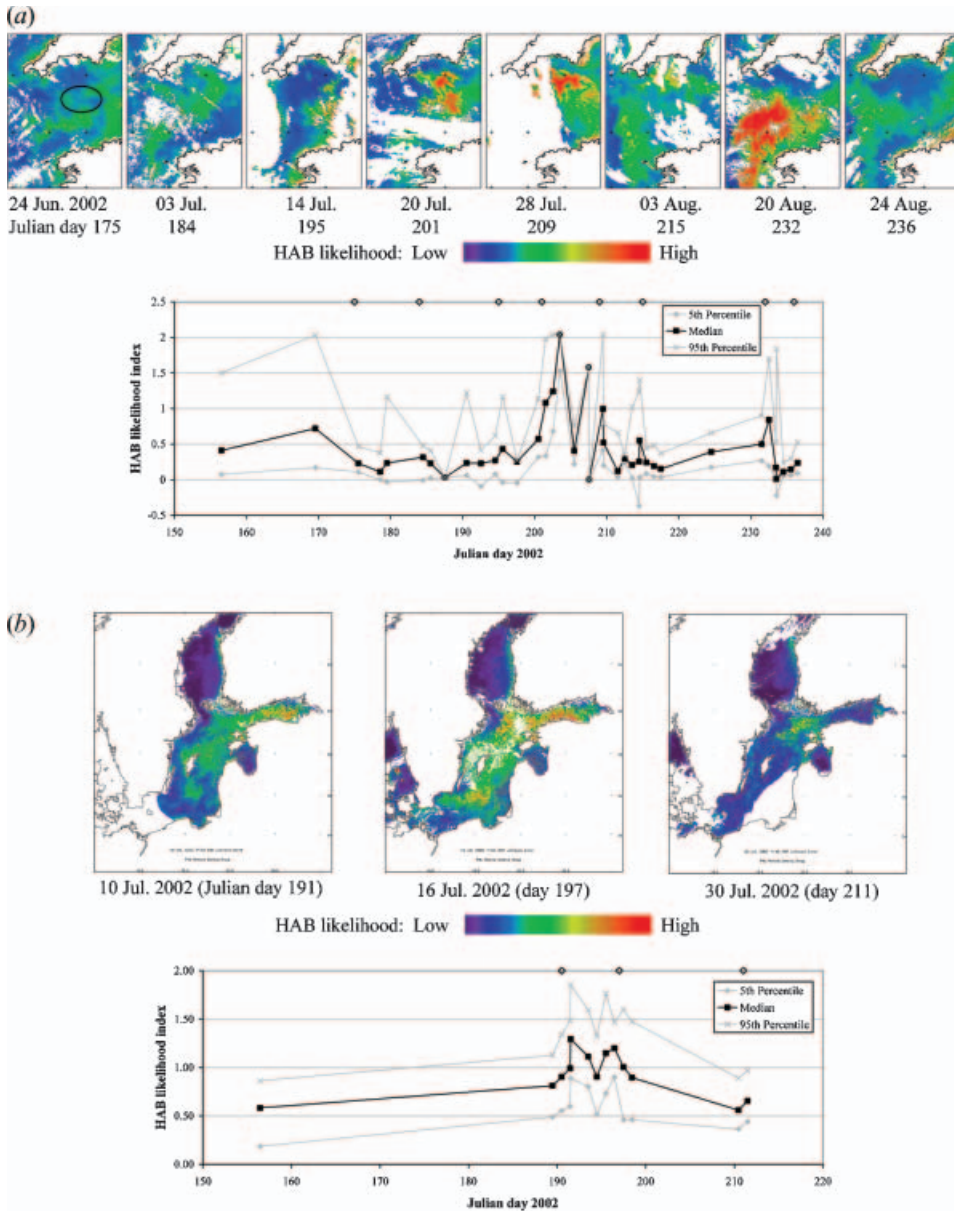


Figure 6. HAB likelihood time series for (a) a *Karenia* bloom in the English Channel and (b) a cyanobacteria bloom in the Baltic Sea, both during the summer of 2002. Selected HAB likelihood maps are shown, where white regions indicate cloud, land or invalid data and their dates are indicated by diamonds along the top of the plots.

effectively detected using this technique. It is of particular note that there are few false alarms compared with visual interpretation of the imagery, for instance there is no response for the coccolithophore bloom near the coast of Cornwall in the 28 July and 12 August scenes or for the significant increase in suspended sediment in the English Channel, Bristol Channel and Irish Sea seen on the 3 August scene.

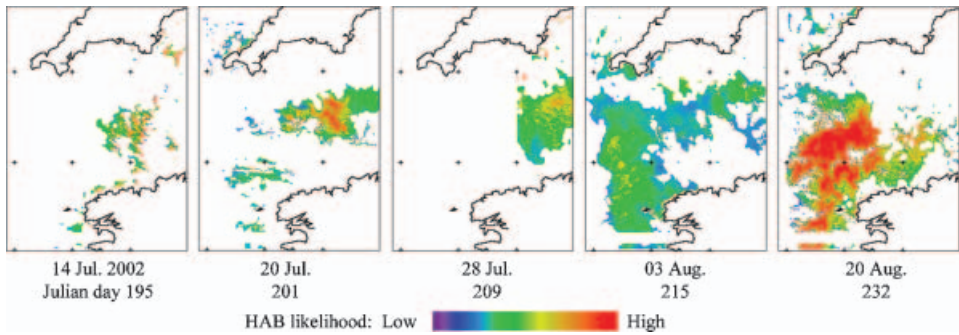


Figure 7. Regions automatically detected as chl-*a* anomalies using median subtraction, applied as masks to HAB likelihood maps selected from figure 6(a). This combined approach reduces false alarms, e.g. for coccolithophore blooms.

For these reasons subtraction may be a useful tool for an automated early warning system, so the integration of this technique with the multivariate HAB classifier was investigated, to enable the automated discrimination of HABs from other chl-*a* anomalies, and to reduce occurrences where HABs are indicated falsely in high-sediment regions. Figure 7 shows how the subtraction-detected anomalies were used to filter the HAB likelihood maps for the summer 2002 *Karenia* sequence shown in figure 6(a), demonstrating significant improvement. On 14 July both techniques agree on the early bloom region. On the subsequent three scenes the high HAB likelihood ratings for the coccolithophore bloom were removed cleanly using the subtraction mask, leaving a majority of the other algae. The 28 July and 3 August scenes demonstrated a successful integration of the techniques as the subtraction algorithm delineated the regions of increased algal concentration, within which the classifier indicates potentially harmful patches.

A final experiment was conducted to determine whether subtraction would assist in the detection of *Chattonella* in the North Sea, where intense absorption caused atmospheric correction and related classification problems. An example SeaWiFS scene from 13 May 2000 is shown in figure 8 together with the HAB likelihood map before and after masking using chl-*a* anomalies. The classifier produces a weak assessment, assigning a high rating to a coccolithophore bloom and a medium rating to the harmful bloom. The filtering successfully excludes the coccolithophores and retains some of the HAB, but the densest region is omitted, again due to the saturation problems.

#### 4. Discussion

The multivariate discrimination results were generated using the most appropriate classifier for each area: *Karenia* for south-west UK, cyanobacteria for the Baltic Sea and *Chattonella* for the North Sea. If multiple classifiers are to be used in a warning system, they could be combined by using the maximum HAB likelihood of the three species to rate each pixel, as illustrated by the preliminary results. Regional information could also be incorporated by weighting the classifier outputs according to the prior probability of observing those species at each location. The problem of missing values in the IOP products could be solved pragmatically by substituting a reduced classifier based only on  $L_{\text{wn}}(\lambda)$  for those pixels.

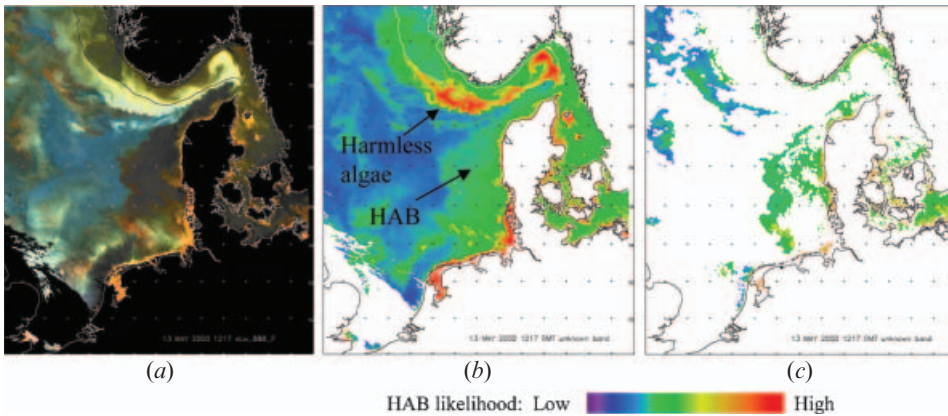


Figure 8. Detection of a *Chattonella* bloom in the North Sea on 13 May 2000 1217 UTC: (a) SeaWiFS enhanced true-colour; (b) HAB likelihood map showing weak detection of bloom compared to false alarm for coccolithophore bloom; (c) chlorophyll anomalies automatically detected and applied to mask HAB likelihood map.

The time-series analysis requires further development to objectively label specific periods of bloom development: no bloom; potentially pre-bloom; early non-harmful bloom; HAB; and late or post-bloom. Other temporal information, such as the rate of increase of HAB likelihood, may generate additional cues as to the type of bloom.

An ultimate aim would be to predict HABs on a European-wide basis, and this may be more practical to achieve at lower spatial resolution and with fewer parameters. Training and testing of the HAB classifier was performed on regional subsets of nominally 1.1 km resolution. Further work is needed to examine impacts on the efficacy of HAB discrimination due to averaging of ocean colour parameters at different spatial scales: global area coverage (GAC) 4.4 km resolution and global product 9 km resolution, and temporal scales: daily, eight-day and monthly composites. A multi-scale approach may be advantageous, to achieve fine resolution in the coastal areas most sensitive to HAB impacts, while using coarser resolution to monitor for large blooms in the open ocean.

Apart from reducing false alarms, there is a possibility that median subtraction could omit potentially harmful regions. This might arise if there are insufficient cloud-free data to model the background distribution, or if the anomaly persists long enough to be included in the background estimate. Also the seasonal progression of the ecosystem, e.g. from diatoms to dinoflagellates (Holligan *et al.* 1983), may change the likelihood of a HAB without an associated effect on chl-*a* concentration. For these reasons it is intended to conduct further research to find optimal parameters for the subtraction algorithm, and the best approach to highlight blooms detected by either subtraction or classification.

## 5. Conclusions

Three linear discriminate analysis classifiers were trained to recognize three HAB species using SeaWiFS data as their input. The spatial result of this classification was masked using the result from a background subtraction technique, designed to remove false-positives from the classifiers caused by high levels of sediment and the effects of non-harmful algal blooms. The need for a different classifier for each HAB species supports the hypothesis that the algal types are spectrally distinct and that

this difference is measurable from space using SeaWiFS. The resultant classifiers highlighted those spectral wavelengths that are particularly suitable for identifying the different algal types, supporting the results from previous *in situ* sampling studies (Millie *et al.* 1997, Liew *et al.* 2000). These approaches enabled preliminary time-series analysis of the dynamic evolution of multiple HABs, allowing the peaks of the blooms to be determined, and estimates of their strength and lifetime. It is hoped that further research on time-series analysis will identify automated methods to warn of HABs in their early stages of development, as this will increase the chance of success of mitigating actions to protect fisheries and aquaculture sites.

It is believed this is the first attempt to use objective HAB classification from SeaWiFS data. Its spectral channels are far from optimal for this discrimination, but this methodology provides a consistent framework for researching HAB detection using more advanced sensors such as MODIS and MERIS, and data fusion of complementary properties from multiple sensors and environmental factors such as sea surface temperature (SST), stratification, meteorology (including photosynthetically active radiation and precipitation) and terrestrial inputs from rivers. MODIS products, including fluorescence and sensor-specific absorption and backscatter algorithms, are currently being integrated into the classifier. The use of MERIS data should enable more exact spectral responses of the algal blooms to be determined.

Through further development of these techniques, the aim is to enable the automatic collation of data for long time-series studies of HAB frequency and extent, contributing to the understanding of environmental factors that impact upon their initiation.

### Acknowledgement

The authors acknowledge NERC Dundee Satellite Receiving Station for acquisition, and NASA SeaWiFS project and OrbImage for access to SeaWiFS data. Dr Paul Aplin and two anonymous reviewers are thanked for their helpful comments. This research was funded by Plymouth Marine Laboratory Core Strategic Programme and EC Framework 5 HABILE project (Contract EVK3-CT2001-00063).

### References

- ANDERSON, D.M., GLIBERT, P.M. and BURKHOLDER, J.M., 2002, Harmful algal blooms and eutrophication: Nutrient sources, composition, and consequences. *Estuaries*, **25**, pp. 704–726.
- AURE, J., DAHL, E., DANIELSSEN, D.S. and SØILAND, H., 2000, *Chattonella*— a new harmful algae in Norwegian coastal waters. *Fisken og Havet*, **2** (Special issue on environmental status of Norwegian oceans [in Norwegian]). Institute of Marine Research, Bergen.
- FLURY, B. and RIEDWYL, H., 1988, *Multivariate statistics: a practical approach* (London: Chapman and Hall).
- GORDON, H.R., BOYNTON, G.C., BALCH, W.M., GROOM, S.B., HARBOUR, D.S. and SMYTH, T.J., 2001, Retrieval of coccolithophore calcite concentration from SeaWiFS imagery. *Geophysical Research Letters*, **28**, pp. 1587–1590.
- GORDON, H.R. and WANG, M.H., 1994, Retrieval of water-leaving radiance and aerosol optical-thickness over the oceans with SeaWiFS – a preliminary algorithm. *Applied Optics*, **33**, pp. 443–452.
- GROOM, S.B., TARRAN, G.A. and SMYTH, T.J., 2000, Red-tide outbreak in the English Channel. *Backscatter*, Fall, pp. 8–11.
- HASSINK, U., 2004, *HELCOM: 30 years of protecting the Baltic Sea* (Helsinki Commission).

- HOGUE, F.E., WRIGHT, C.W., LYON, P.E., SWIFT, R.N. and YUNGEL, J.K., 1999, Satellite retrieval of inherent optical properties by inversion of an oceanic radiance model: a preliminary algorithm. *Applied Optics*, **38**, pp. 495–504.
- HOLLIGAN, P.M., VIOLLIER, M., DUPOUY, C. and AIKEN, J., 1983, Satellite studies on the distributions of chlorophyll and dinoflagellate blooms in the western English Channel. *Continental Shelf Research*, **2**, pp. 81–96.
- KAHRU, M., LEPPANEN, J.M., RUD, O. and SAVCHUK, O.P., 2000, Cyanobacteria blooms in the Gulf of Finland triggered by saltwater inflow into the Baltic Sea. *Marine Ecology-Progress Series*, **207**, pp. 13–18.
- KELLY-GERREYN, B.A., QURBAN, M.A., HYDES, D.J., MILLER, P. and FERNAND, L., 2004, Coupled “FerryBox” Ship of Opportunity and satellite data observations of plankton succession across the European Shelf Sea and Atlantic Ocean. In *International Council for the Exploration of the Sea (ICES) Annual Science Conference*, 22-25 September, Vigo, Spain.
- LAVENDER, S.J. and GROOM, S.B., 1999, The SeaWiFS automatic data processing system (SeaAPS). *International Journal of Remote Sensing*, **20**, pp. 1051–1056.
- LIEW, S.C., KWONG, L.K. and LIM, H., 2000, Classification of algal bloom types from remote sensing reflectance. In *21st Asian Conference on Remote Sensing*, GISdevelopment.net, 4-8 December, Taipei, Taiwan.
- MILLIE, D.F., SCHOFIELD, O.M., KIRKPATRICK, G.J., JOHNSEN, G., TESTER, P.A. and VINYARD, B.T., 1997, Detection of harmful algal blooms using photopigments and absorption signatures: A case study of the Florida red tide dinoflagellate, *Gymnodinium breve*. *Limnology and Oceanography*, **42**, pp. 1240–1251.
- MOORE, G.F., AIKEN, J. and LAVENDER, S.J., 1999, The atmospheric correction of water colour and the quantitative retrieval of suspended particulate matter in Case II waters: application to MERIS. *International Journal of Remote Sensing*, **20**, pp. 1713–1733.
- PETTERSSON, L.H., DURAND, D., JOHANNESSEN, A.M., SVENDSEN, E., NOJI, T., SOILAND, H., GROOM, S. and REGNER, P., 2001, Monitoring and model predictions of harmful algae blooms in Norwegian waters. In *Proceedings of the International Geoscience and Remote Sensing Symposium (IGARSS'01)*, 9-13 July, Sydney. *IEEE*, **3**, pp. 1146–1148. doi: 10.1109/IGARSS2001.976773.
- PETTERSSON, L.H., FUREVIK, B.R., DURAND, D. and JOHANNESSEN, J.A., 2003, Synergies in marine applications of ocean colour, infrared and SAR satellite EO data. Technical report no. 242, Nansen Environmental and Remote Sensing Center.
- RAINE, R., O'BOYLE, S., O'HIGGINS, T., WHITE, M., PATCHING, J., CAHILL, B. and MCMAHON, T., 2001, A satellite and field portrait of a *Karenia mikimotoi* bloom off the south coast of Ireland, August 1998. *Hydrobiologia*, **465**, pp. 187–193.
- RANTAJÄRVI, E., 2003, Alg@line in 2003: 10 years of innovative plankton monitoring and research and operational information service in the Baltic Sea. MERI Report Series of the Finnish Institute of Marine Research, **48**.
- SCHOFIELD, O., GRZYMSKI, J., BISSETT, W.P., KIRKPATRICK, G.J., MILLIE, D.F., MOLINE, M. and ROESLER, C.S., 1999, Optical monitoring and forecasting systems for harmful algal blooms: Possibility or pipe dream? *Journal of Phycology*, **35**, pp. 1477–1496.
- SMITH, T.J., MOORE, G.F., HIRATA, T. and AIKEN, J., A semi-analytical model for the derivation of ocean colour inherent optical properties: description, implementation and performance assessment. *Applied Optics*, submitted.
- STUMPF, R.P., CULVER, M.E., TESTER, P.A., TOMLINSON, M., KIRKPATRICK, G.J., PEDERSON, B.A., TRUBY, E., RANSIBRAHMANAKUL, V. and SORACCO, M., 2003, Monitoring *Karenia brevis* blooms in the Gulf of Mexico using satellite ocean color imagery and other data. *Harmful Algae*, **2**, pp. 147–160.
- VAN DOLAH, F.M., 2000, Marine algal toxins: Origins, health effects, and their increased occurrence. *Environmental Health Perspectives*, **108**, pp. 133–141.