# Image processing for smarting browsing of ocean color data products: investigating algal blooms

Jer Hayes [a] *, Edel O'Connor [a], King-Tong Lau [a], Noel E. O'Connor [a], Alan F. Smeaton [a], Dermot Diamond [a]
[a]CLARITY – centre for sensor web technologies, Dublin City University (Ireland)
*IBM, Mulhuddart, Ireland
hayesjer@ie.ibm.com

## ABSTRACT

Remote sensing technology continues to play a significant role in the understanding of our environment and the investigation of the Earth. Ocean color is the water hue due to the presence of tiny plants containing the pigment chlorophyll, sediments, and colored dissolved organic material and so can provide valuable information on coastal ecosystems. We propose to make the browsing of Ocean Color data more efficient for users by using image processing techniques to extract useful information which can be accessible through browser searching. Image processing is applied to chlorophyll and sea surface temperature images. The automatic image processing of the visual level 1 and level 2 data allow us to investigate the occurrence of algal blooms. Images with colors in a certain range (red, orange etc.) are used to address possible algal blooms and allow us to examine the seasonal variation of algal blooms in Europe (around Ireland and in the Baltic Sea). Yearly seasonal variation of algal blooms in Europe based on image processing for smarting browsing of Ocean Color are presented.

**Keywords:** Ocean Color, image processing, algal blooms.

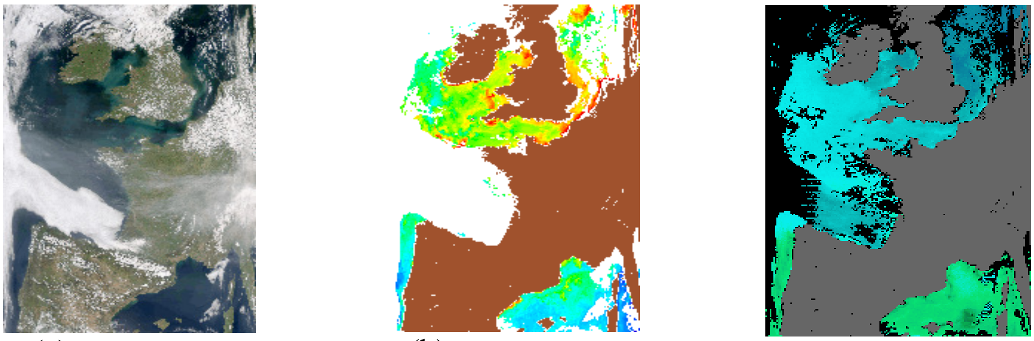
## 1. INTRODUCTION

Many environmental phenomena occur over a spatial range that in-situ sensors (or sensor networks) have difficulty in monitoring, e.g., harmful algal blooms (HABS). Harmful algal blooms are of interest for two primary reasons: human health, and the large range of effects they have on the marine ecosystems. The frequency of occurrence of HABS has increased across the globe in recent decades in coastal waters [1,2]. Remote sensing is ideal for detecting and monitoring algal blooms over larger spatial scales than sensor networks and long time scales as it offers unique large scale synoptic data to capture the range and variability of many complex processes. Although, a complete picture of algal blooms with respect to their causes and occurrence also require the use of in-situ sensors as one of the contributing factors to HAB is increased nutrient pollution [3] which enters watersheds and ultimately the sea. Better management of nutrient inputs, e.g., phosphate, to the watershed can lead to significant reductions in HABs. The development of in-situ sensors which are capable of monitoring nutrient levels, e.g., [4,5], suggest that a more holistic approach where multi-modal data is analysed will lead to better management than currently exists.

Algal blooms consist of phytoplankton or cyanobacteria and the bloom is due to a rapid increase in the population of the species involved. Not all algal blooms are harmful but harmful blooms can affect human health and can have large adverse effects on the marine ecosystems, including the death of plants and animals due the reduction of dissolved oxygen caused by the HAB. Much research has been carried out into satellite image analysis techniques to distinguish between toxic species and other species (e.g. [6,7,8,9,10]). Algal blooms are a natural phenomenon; however the frequency, duration and distribution appear to have increased in recent years [11] most probably due to increased nutrient pollution [3]. Many studies are investigating the use of satellite imagery and in-situ field measurements to predict the occurrence of algal blooms in order to mitigate their effects in a timely fashion (e.g. [3,12,13,14] ).

Despite the huge benefits that satellite remote sensing has brought to marine monitoring applications, there are also many issues and drawbacks with using this type of data. Generally ocean colour satellite sensors only operate in the solar reflective spectral range; hence they only gather useful data on cloud-free days during periods of daylight (i.e. when illumination conditions are suitable). In regions where cloud cover is a predominant weather condition such as in coastal regions around Ireland and the UK, this can prove hugely problematic for obtaining satellite ocean colour information at high temporal scales. There are also issues in the coastal zone for the reliable estimation of parameters especially that of chlorophyll concentration. Thus a marine monitoring/event detection network that incorporates a number of sensing modalities is desirable. Work carried out in [13] demonstrates the necessity for integrating both satellite and in-situ information sources in an environmental sensor network monitoring SST related events.

Ocean color is defined as the water hue due to the presence of tiny plants containing the pigment chlorophyll, sediments and colored dissolved organic material and so can provide valuable information on coastal ecosystems The 'Ocean Color project' collects data from various satellites (e.g. MERIS, MODIS) and makes this data freely available online (see: http://oceancolor.gsfc.nasa.gov/). The project is an exemplar of data sharing for scientific use (and public use). One method of searching the Ocean Color project data is to visually browse level 1 and level 2 data. Users can search via location (based on predefined regions or user selected areas), time and data type. They are then presented with images which cover chlorophyll, quasi-true color and sea surface temperature (11 μ) and links to the source data for particular days (see Figure 1). However it is often preferable for users to search by event and analyze the distribution of color in an image before examination of the source data.



(a) (b) c

Figure 1: Summary images of raw data representing (a) quasi-true colour, (b) chlorophyll, and (c) Sea Surface Temperature (11 μ)

We propose to make the browsing of Ocean Color data more efficient for users by using image processing techniques to extract useful information which can be accessible through browser searching. The automatic image processing of the visual level 1 and level 2 data allow us to investigate the occurrence of algal blooms. Images with specific color signatures (red, orange etc.) can be flagged as possible algal blooms and allow us to examine the seasonal variation of algal blooms in Europe (around Ireland and in the Baltic sea).

The advantage of using color signature can be seen in Figure 1?, where area with similar color signature can be grouped together to represent if an algal boom event has/has not occurred in these area. This color-coding technique simplifies the search process through making available the color and /or area of interest relating to the algae concentration.

Hence, an image analysis system that allows for smarter browsing has been developed [16]. This system, amongst other things, allows a user, given an image, to select the region they want to analyze and the histogram they want to view for this image (i.e. All Band, 5-Band). The bars of the histogram are colored according to the color value that particular bin represents. The user can also select a start date, an end date, a region and a color and to examine its distribution over time. The system is designed to be extended to allow the distribution of data from other modalities to be examined. An actual determination of whether an algal bloom is present is best supported by a variety of evidential data (ship surveys, buoy data etc.). Thus the system allows for incorporation into an event detection framework where signals from various modalities are compared, source data is incorporated where events are disputed or validated.

A search system was also developed which allows the user to browse data by choosing a time period, a region, a color, and a percentage of the region. For example the user may want to search for images where there has been a high chlorophyll event (e.g. red pixels) or where at least 75% of the region is not covered by cloud. This allows the user to search for images which are not contaminated by cloud cover and which contain a particular color representing a certain percentage of the water pixels in that region. The user can choose to sort the results by percentage or by date. The user can also browses the image data and if they see an event of interest in a particular region, the system will retrieve other images with similar events in this region. Similar regions are retrieved through calculating the Euclidean distance between the histograms (the user can choose this comparison to be a low-level or a high-level comparison through selecting All-Band or 5-Band). This is a standard technique in the image processing literature, however the system will be extended to include other histogram comparison techniques and investigate those which are more suitable. Candidate events (e.g. algal blooms) can be identified and the conditions surrounding them which can help predict these events in the future and mitigate the effects in a timely fashion.

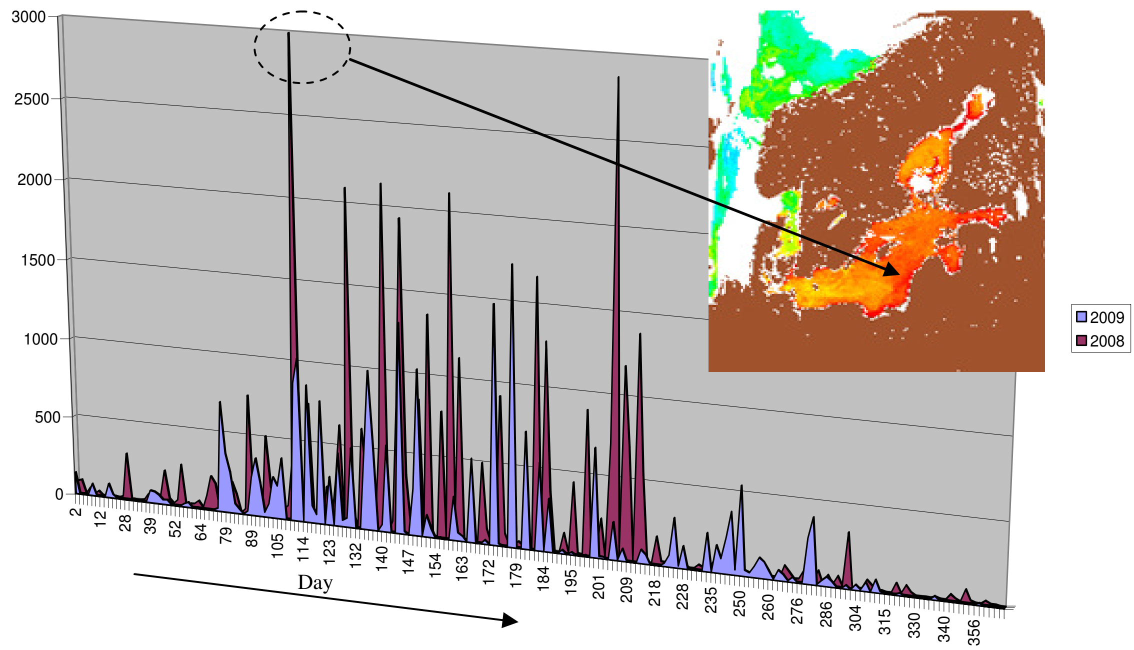



Figure 2: Image browsing based on a simple analysis of colour allows a user to quickly identify possible events of interest. Data from a two year period are presented as images that had 75% Baltic coverage with color intensity in a specific range from Images being mapped against days of the year.

The proposed search system made use of simple color information to allow quick and easy access to data which meets certain criteria for further analysis. It can also easily be extended into a framework whereby images which meet certain criteria can be compared against other modalities, e.g., sea surface temperatute, in-situ sensors, meteorological conditions etc. If an onset of an algal boom event is suspected, source data can subsequently be requested and a more in depth study can be carried out, resulting in early warning of marine events. It also allows for the quick analysis of trends in a region over time. Other modalities can be investigated for dates or periods where specific types of events have been detected in a certain region and an in depth analysis can be carried out.

## 2. INVESTIGATING ALGAL BLOOMS: METHODS AND DATA

To undertake the initial development of the system images from The ‘Ocean Color project’ (http://oceancolor.gsfc.nasa.gov/) were automatically downloaded. The images were similar to the quasi-true color, chlorophyll, and Sea Surface Temperature (11 μ) as shown in Figure 1. The images were grouped into those that covered 75% of the Baltic Sea and those that contained the English Channel from 2003 to 2010. These images are generated from MODIS swaths. Only MODIS images were considered at this stage for this investigation but the incorporation of data from other satellite sensors such as MERIS could also be  included very simply if required. For a given day tdownloaded images generally included quasi-true color, chlorophyll and SST images. Although occasionally an additional fourth SST related image was sometimes available and was downloaded but it was not used in the context of this work. The images grouped into those that covered 75% of the Baltic Sea contained 14,350 images (group 1) while the English channel group contained 2265 images (group 2) – see Table 1 for more details.

Table 1 – Breakdown image types and number for the data sets

| | Quasi-true color | Chlorophyll | Sea Surface Temperature (11 μ) |
|---|---|---|---|
| Group 1 | 4308 | 4308 | 5734 |
| Group 2 | 725 | 725 | 714 |

An analysis of the color shows that brown represents land, white represents no data available. Other colors range from purple to red, with purple representing the lowest amount of chlorophyll and red representing the highest. To search for possible algal blooms we searched for pixels that had a red channel > 250, a green channel > 250, and a blue channel < 10. Colors in this range are considered as candidate algal bloom events and we dub these 'pixels of interest' and represent an orange to red color range.

## 3. RESULTS

We now present yearly results which show the seasonal variation of algal blooms in Europe based on the image processing for smarting browsing of Ocean Color data products outlined in the previous sections. We first look at the Baltic Sea region and then examine the English Channel region (as these images often include Ireland an area of particular interest to the authors).

In Figure 2 is shown the average number of pixels of interest in images per month. The pixels of interest are those that would be associated with an algal bloom event. This gives a coarse view of the seasonal variation. The plot for the 2008 data suggests that this was a year with more candidate bloom events than other years. The frequency of algal boom occurrence is clearly seasonal between April and August and there is not a year on year relationship. A caveat is that cloud cover effects the analysis in that algal blooms may exist in the ocean but do not appear in the image data as they are blocked by the cloud cover. The lack of data for given days would reduce the pixel of interest frequency. However, even with this in mind we can quickly determine times of the year when algal blooms occur. Some of the images which gave rise to the highest average pixels of interest are shown Figure 4 (which cover May 2008). In Figure 5, a summary view of the seasonal variation for the English Channel group images is shown. As with Figure 3 the plot for the 2008 data suggests that this was a year with more candidate bloom events than other years.

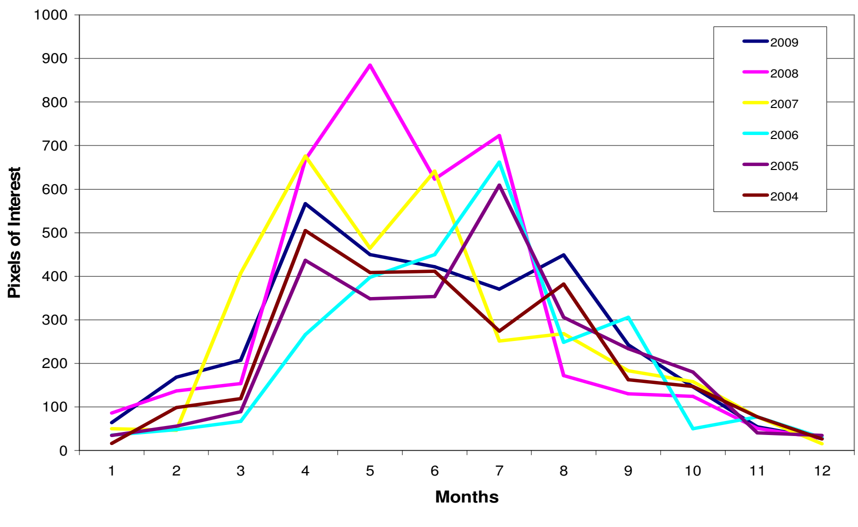
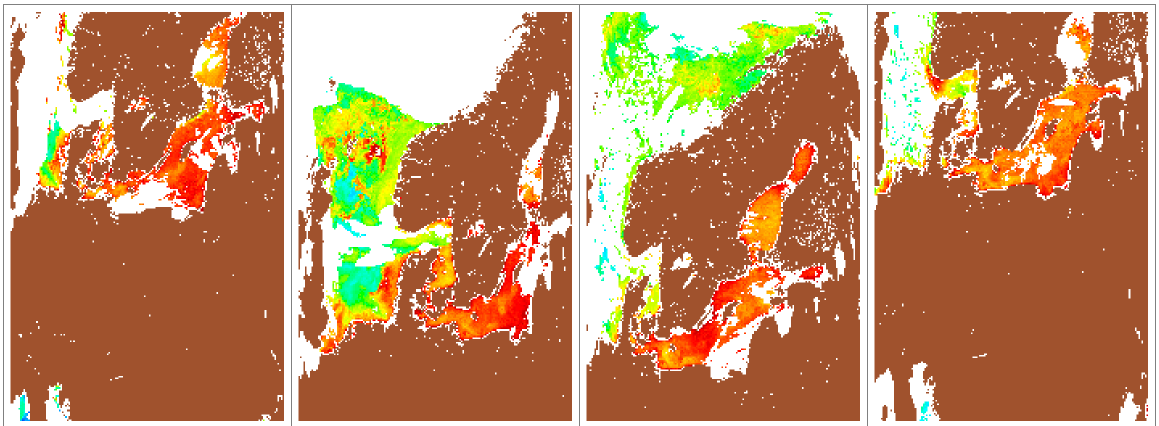



Figure 3: Seasonal variation of candidate algal bloom events in the data set where images covered at least 75% of the Baltic region from the Ocean color browser.



| Day 125, 2008 | Day 130, 2008 | Day 143, 2008 | Day 147, 2008 |
|---|---|---|---|

*Figure 4: Examples of images which have a large number of pixels of interest for the Baltic Region.*

The analysis of chlorophyll images to detect possible candidate events is best used as a means to quickly browse the data. Yet even a simple analysis of the RGB values of data can highlight when event hotspots are likely to occur and also to highlight the relative differences between different years. However, each image file is associated with the name of the file that represents the source data (in this case MODIS) so that further investigation can be undertaken with the raw data.

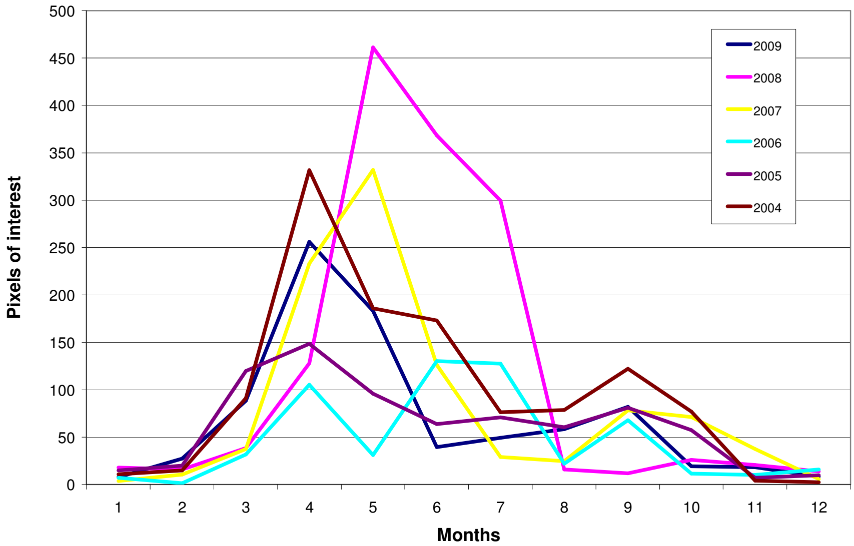



Figure 5: Seasonal variation of candidate algal bloom events in the data set where images covered the English Channel from the Ocean color browser.

In conjunction with the detection of possible candidate events the image analysis can also be used as a means to determine when we are likely not to find useful information, e.g., when there is heavy cloud cover. In Figure 6, we analysis the average count of white pixels in the chlorophyll images which gives an indication of the levels of cloud cover for the years 2007 to 2009. As would be expected winter months have the highest cloud cover.

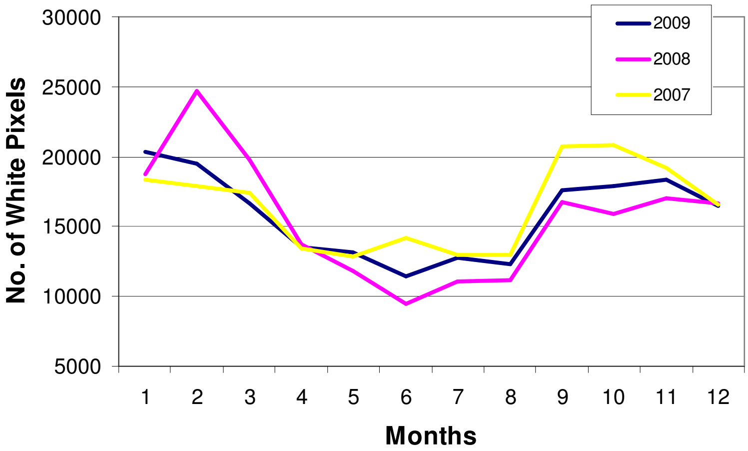



Figure 6: Frequency of white pixels (which indicate cloud cover) for that Baltic Sea data set.

## 4. CURRENT DEVELOPMENTS – IMAGE ANALYSIS SYSTEM

In order to acquire information for specific regions of interest, e.g. Ireland. We need to analyse the satellite images to identify regions of interest. This regional identification is currently being worked on [16]. In order to detect events in specific regions around the coast of Ireland, six main regions were delineated – North West, West, South West, North East, East, South East. This is quite a difficult task due to the fact that the shape, size and orientation of Ireland changes from image to image. An algorithm which uses the centroid, bounding box and orientation of Ireland in the image under consideration is used in order to regionalise Ireland. This algorithm uses the chlorophyll images and works well in most cases (except where Ireland is located in the far left of the image). In this algorithm, the orientation of the connected component representing Ireland is used in order to orient the image whereby Ireland is in an upright position.

## 5. CONCLUSIONS

We proposed to make the browsing of Ocean Color data more efficiently for users by using information extracted by image processing techniques to simplify the search process. The automatic image processing of the visual level 1 and level 2 data also allow users to investigate the occurrence of algal blooms. Two data sets which contained images centered on the Baltic Sea and English Channel were downloaded and processed. The yearly results show that although the frequency, duration and distribution appear to have increased in recent years [11] no apparent trend can be observed. This is not unexpected as the weather conditions conducive to algal blooms varies every year and better water management in the future should decrease the nutrient load entering coastal areas.

Current work on an image analysis system has focused on the use of machine learning to detect regions and the ability to integrate more data sources. So rather than analyzing whole images (as in Figure 4) these regions within this images are now processed to offer a more fine-grained analysis and browsing experience.

## 6. ACKNOWLEDGEMENTS

Based on research funded by the Dept. of Communications, Marine & Natural Resources under the Strategy for Science, Technology and Innovation (2006-2013) and by Science Foundation Ireland under grant 07/CE/I1147.

## 7. REFERENCES

1. Hallegraeff, 1993 G.M. Hallegraeff, (1993). A review of harmful previous termalgal bloomsnext term and their apparent global increase, Phycologia 32, pp. 79–99.
2. Anderson, D.M. (1997). Turning back the harmful red tide, Nature 388 (1997), pp. 513–514.

3. J. Heisler, P.M. Glibert, J.M. Burkholder, D.M. Anderson, W. Cochlan, W.C. Dennison, Q. Dortch, C.J. Gobler, C.A. Heil, E. Humphries, A. Lewitus, R. Magnien, H.G. Marshall, K. Sellner, D.A. Stockwell, D.K. Stoecker, M. Suddleson (2008). Eutrophication and harmful algal blooms: A scientific consensus, Harmful Algae, Volume 8, Issue 1, HABs and Eutrophication, pp. 3-13.

4. Cleary, John and Slater, Conor and Diamond, Dermot (2009) Analysis of phosphate in wastewater using an autonomous microfluidics-based analyser. In: ICEST 2009 - International Conference on Environmental Science and Technology, 28-30.

5. C. Slater, J. Cleary, C. M. McGraw, W.S. Yerazunis, K.T. Lau, and D. Diamond, "Autonomous field-deployable device for the measurement of phosphate in natural water" Proc. of SPIE, Vol. 6755, Sept. 2007, 67550L1-67550L8.

6. Amin, R., Zhou, J., Gilerson, A., Gross, B., Moshary, F., and Ahmed, S., (2009). Novel optical techniques for detecting and classifying toxic dinoflagellate Karenia brevis blooms using satellite imagery, Opt. Express, 17, pp. 9126-9144.

7. Astoreca R, Rousseau V, and Lancelot, C. (2009). Coloured dissolved organic matter (CDOM) in Southern North Sea waters: Optical characterization and possible origin.Estuarine, Coastal and Shelf Science, Volume 85, Issue 4, pp. 633-640.

8. Kahru, M. and Mitchell, B. G. (1998), Spectral Reflectance and absorption of a massive red tide off Southern California. Journal of Geophysical Research, 103, pp. 601-621.

9. David F. Millie, Oscar M. Schofield, Gary J. Kirkpatrick, Geir Johnsen, Patricia A. Tester and Bryan T. Vinyard (1997). Limnology and Oceanography, Vol. 42, No. 5, Part 2: The Ecology and Oceanography of Harmful Algal Blooms (Jul., 1997), pp. 1240-1251

10. Staehr, P.A. and Cullen, J.J. (2003). Detection of Karenia mikimotoi by spectral absorption signatures, Journal of Plankton Research, 25, pp. 1237-1249.

11. Muller-Karger, F.E., Andrefouet, S., Varela, R., and Thunell, R., (2007). The colour of the coastal ocean and applications in the solution of research and management problems, Remote Sensing of Coastal Environments, chapter 5, pp. 101-127.

12. Pitcher, G.C., Figueiras, F.G., Hickey, B.M. and Moita, M.T., (2010). The physical oceanography of upwelling systems and the development of harmful algal blooms, Progress in Oceanography, 65, 1:2, pp. 5-32.

13. Pitcher, G.C. and Weeks, S.J., (2007). The variability and potential for prediction of harmful algal blooms in the southern Benguela ecosystem, Large Marine Ecosystems, 14, pp. 125-146.

14. Zingone, A., and Enevoldsen, H.,O., (2000). The diversity of harmful algal blooms: a challenge for science and management. Ocean and Coastal Management, 43, 8:9, pp. 725-748.

15. O'Connor, E. and Smeaton, A.F. and O'Connor, N.E. (2010) A reputation and trust based multi-modal sensor network for environmental monitoring. In: Environ 2010, 17-19 February 2010, Limerick, Ireland.

16. O'Connor, E., Hayes, J., Lau, K.T., Smeaton, A.F., O'Connor, N.E., and Diamond, D. (2010). Image processing for smart browsing of oceancolor data products and subsequent incorporation into a multi-modal sensing framework. To be presented at the Joint conference of the UK RSPSoc and IEOS, Cork, Ireland.