

THE USE OF MODIS DATA TO DEFINE NATURAL BOUNDARIES AND REGIONS IN THE MARINE WATER COLOUMN.

J.H.J. Leach* and A. Kitchingman

Department of Infrastructure Engineering, University of Melbourne, Parkville, Victoria, Australia.

KEYWORDS: Marine, Mapping, Classification, Satellite, Multitemporal

ABSTRACT

This study looks at the elements that define a natural boundary and its marine variant. It then aims to quantify and delimit natural marine boundaries from satellite sea surface temperature and ocean colour imagery. Natural boundaries often have a fuzzy or zonal nature, can be dependent on and independent from the character of the region which they bound, and have temporal characteristics that occur over a variety of temporal scales. Natural marine boundaries have the added complexity of occurring in three dimensions and of being temporally dynamic throughout the water column. A monthly time-series of MODIS satellite imagery from July 2002 to June 2006 was used in this study. The standardised principle component analysis (sPCA) was used to identify the mean state of identifiable natural process and to quantifying a degree of spatio-temporal (seasonal) dichotomy in detected marine processes. Component images from the sPCA provided a spatial representation of marine phenomena, while sPCA loadings indicated the temporal fluxes exhibited in a given component. Classification of significant sPCA components was conducted to derive regional boundaries. Initially, this was performed in an unsupervised manner using an ISO process. Subsequent classified images were then simplified using a combination of dendrograms and knowledge-based examination. A comparison of these results with similarly derived regionalisation showed that the spatio-temporal representations have the potential to further inform marine spatial management regimes. It was also shown that the delimitation of natural phenomena is feasible and can have seasonal dynamics represented.

INTRODUCTION

This study investigated the potential for deriving boundaries from natural processes in the marine environment. This study looks at the elements that define a natural boundary and its marine variant. It was able to quantify and delimit natural marine boundaries from satellite sea surface temperature and ocean colour imagery.

NATURE OF NATURAL MARINE BOUNDARIES

The nature and definition of a natural boundary is complicated and often subjective.

Natural boundaries often have a fuzzy or zonal nature, can be dependent on, and independent from, the character of the region which they bound, and have temporal characteristics that occur over a variety of temporal scales. The nature of the marine environment is invariably three-dimensional, and sometimes four-dimensional (time), to account for the variety of marine resources utilised. Fisheries are exploited throughout the water column and along the sea-floor, while mineral resources in the same location can be mined. In terms of a cadastre, a terrestrial equivalent would be a spatial administration system simultaneously encompassing roads, flight paths, cattle herds and flocks of birds. Theoretically the sea-floor can be treated in a similar way to land and its mineral resources. The sea-floor's similarity to land means most traditional cadastral approaches can be used, as seen in Australian Government's Offshore Acreage Leases for oil and gas exploration.

Key differences to existing cadastres or spatial management systems occur in the water column. Although there is legislation to cover three-dimensional areas, little is available for dealing with the interaction of multi-dimensional entities as well as their representation in a cadastre. This leads to cases where activities and

subsequent regulations on the sea-floor overlap and potentially conflict with those in the water column. Current boundaries that regulate water column resources such as fisheries are delimited in land parcel-like regions that are overlaid on the ocean surface. They are often designed for statistical purposes and have no or little reflection of water column structure that may govern the distribution and quantity of a fisheries species. There is also little reflection of the temporal nature of a water column resource other than the opening and closing of fishing zones. Many pelagic fisheries species are highly migratory and sometimes follow thermoclines which can occur at various depths over different latitudes. Temperate southern bluefin tuna have been found in colder thermoclines below the hotter equatorial waters tropics during spawning aggregations (Davis and Farley 1999), while tropical and sub-tropical skipjack tuna are known to follow warm water bodies across the Pacific Ocean (Lehodey et al. 1997).

The pelagic realm has only just begun to be adequately sampled. Until the innovation and introduction of recent technologies (satellite imagery, GPS, etc.) many aspects, such the dynamics of the sea surface and water column, were inaccessible or poorly represented. Since dynamic elements are intrinsic to the marine environment (Mann and Lazier 2006) greater representational emphasis must be directed towards three dimensions and/or spatio-temporal dynamism. The variety and accessibility of maritime datasets at global scales is continuing to expand. The advent of near real-time processing of these datasets is providing the opportunity to re-assess and/or calibrate spatial entities based on natural marine phenomena. Challenges still lie in the acquisition of broad scale high resolution data, as well as combining pelagic and sea-floor regionalisation in a cohesive form for simplified spatial management planning.

MODIS DATA

Imagery from the Moderate Resolution Imaging Spectroradiometer (MODIS) was chosen as the most suitable data source for this project. MODIS data has become a key source in providing global environmental data for a vast range of scientific and monitoring applications. MODIS data is freely available in a wide variety of formats and temporal bins. Its time-series spanning several years and ongoing supply of data makes it suitable for temporal investigations.

The MODIS sensor encompasses atmospheric, land and ocean imaging in a single instrument. There are two processes such as surface water movement (i.e. currents and eddies) are indirectly detected through the effects they have on SST or Ocean Colour. Satellite imagery can also show evidence of oceanic events, such as upwellings, which occur through the water column from depths greater than ones where satellite sensors can directly detect. The MODIS SST and Ocean Colour data have proven to be reliable (EOS Aqua in particular) and widely used in oceanographic research. The combination of near global coverage and relatively high resolution has made MODIS a valuable source of environmental data.

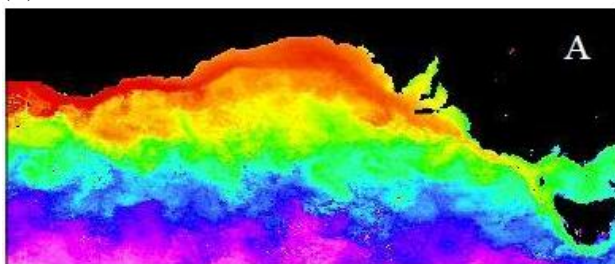
SST indicates the temperature of the very thin surface layer of water (less than 1mm thick). It is often slightly cooler than the 'bulk' sea surface temperature, which is taken from the top several meters of ocean surface. Satellite SST is essentially measuring the temperature at the transition point between the ocean and atmosphere, which makes it very

MODIS instruments on board the Earth Observing System (EOS) satellites, which are aimed at furthering knowledge of global dynamics and process. The EOS satellites are in circular, sun-synchronous, near-polar orbits approximately 705km from the Earth's surface. Almost all of the Earth is covered every 1-2 days. The MODIS instrument has a swath at 10km along track and 2330km across track. There are 36 spectral bands covering wavelengths of 0.4µm to 14.5µm.

The SST and Ocean Colour data, available through MODIS, are two of the key oceanic properties that enable global insight into oceanic surface processes. Many of the important in oceanic/atmospheric transfer research. SST has become one of the most important climate and oceanographic variables available (Kaplan 1998).

Initially, SST was processed in its raw form using the degree Celsius values. Data in this form are largely affected by the first order latitudinal trend of the Earth's temperature gradients. While this approach eventually provided useful results in identifying the latitudinal gradients in SST, it did not satisfy the effort to identify oceanographic features such as currents, eddies, etc. that may have proved more useful and relevant at a regional scope. A simple approach to removing the first order latitudinal trend was by obtaining the residuals from a Least Squares Linear Regression fit between raw temperature values and latitude. These SST residuals enhanced clarification of SST water bodies and allowed interpretation without the influence of latitude (Figure 1).

(A)



(B)

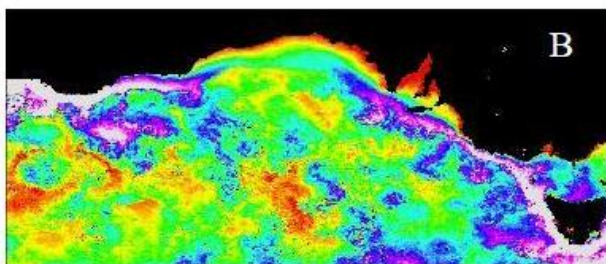


Figure 1, June 2004 data, (A) Original and (B) Linear Least Squares residuals

of the major challenges to CSAT accuracy is situations where non-phytoplankton contaminates in the ocean alter pixel values. These contaminates are often suspended particulates or dissolved organic matter (CDOM). CDOM is essentially made up of disturbed substrate and/or detritus resulting from a myriad of shelf/coastal processes. In order to minimise large regions of no data in the final results, efforts have been made to fill cloud cover pixels from the average value of nearby valid pixels. A neighbourhood

A key goal of ocean colour sensing is the determination of phytoplankton chlorophyll-a concentration in the near surface of ocean waters (Yoder and Kennelly 2006). Phytoplankton are microscopic plants and principle photosynthetic organisms in the oceans which form the basis for many food webs and primary production (Yoder and Kennelly 2003). The concentration of the active photosynthetic agent, chlorophyll-a, in phytoplankton is derived from ocean colour values giving a proxy for the level of primary biological activity (Ryther and Yentsch 1957).

Ocean colour data has been collected via a variety of satellite sensors since the late 1970s. Over this time the accuracy and limitations of ocean colour data has been determined and examined widely. Ocean colour is detected in the visible wavelength spectrum, with the concentration of chlorophyll-a derived from the green and blue spectrum bands (Robinson 2004, Yoder and Kennelly 2006). As with most remotely sensed properties, there are contaminates which decrease the accuracy of detected pixel values. Since ocean colour is detected in the visible spectrum, cloud cover plays a major contaminate role with sometime vast areas completely masked. Sun glint from waves is also another source of reflectance contamination. Probably one

interpolation from the averaged value of valid pixels in a nine pixel radius (~42km) was found suitable to fill small to medium sized regions of cloud cover, while maintaining an indicative pixel value reflecting an appropriate influence of its surrounds.

A monthly time-series of MODIS satellite imagery from July 2002 to June 2006 was used in this study. The standardised principle component analysis (sPCA) was used

to identify the mean state of identifiable natural process and to quantifying a degree of spatiotemporal (seasonal) dichotomy in detected marine processes. Component images from the sPCA provided a spatial representation of marine phenomena, while sPCA loadings indicated the temporal fluxes exhibited in a given component. Classification of significant sPCA components was conducted to derive regional boundaries.

DATA PROCESSING

Standardised principle component analysis (sPCA) has been used to spatially quantify the annular mode and inter-annular dynamics of ocean surface processes. This was conducted on the four-year monthly binned dataset of MODIS Aqua SST and CSAT. In order to validate sPCA results a knowledge-based examination was conducted and key oceanographic processes were identified. Results of this examination are important to assess the ability of the analysis to maintain important oceanographic information as well as aid the eventual classification of oceanic regions for further marine zonation research.

Principle component analysis is a well established statistical technique used to condense multivariate information into a set of typical values. Within the realm of remote sensing it has often been used for data compression, image enhancement, change detection and the examination of temporal dimensionality (Hirosawa et al. 1996, Cheng et al. 2006, Piwowar and Millward 2002, Vlahakis and Pollatou 1993). For multi-temporal studies, it was suggested by Singh and Harrison (1985) that an sPCA yields improved image enhancement over PCA (otherwise known as unstandardised PCA) especially within the scope of spatio-temporal analyses. Standardised principle component analysis differs from the normal PCA with the use of a correlation matrix instead of a covariance matrix. This essentially standardises the original dataset to a mean of 0 and deviation of 1. Individual standardisation of the original data and the use of a covariance PCA would lead to the same result as using unaltered data in a correlation PCA or sPCA.

Standardisation produces an equal weighting of all the input images and prevents certain features from dominating the analysis due to large numerical values (Eklundh and Singh 1993). The need to cater for the influence of variance difference in time-series satellite imagery is important in reducing error factors inherent in image acquisition such as detector calibration, sun angle and atmospheric transmission (Singh and Harrison 1985). Standardisation also reduces the influence of large scale singular anomalies on higher components that aim to represent an overall spatio-temporal. To maintain relevance to marine cadastral and boundary research it is necessary to produce vector regional bounds for natural features identified in the sPCA results. The sPCA components in this study were regionalised to produced simplified regionalised layers

CLASSIFICATION AND VERIFICATION

The results of the sPCA produced rasters of continuous coverage. In themselves they can provide a useful baseline layer to inform the spatial character and degree of seasonal dynamics. However, in terms of management and ultimately cadastral needs there is a preference for some

representing the key detected oceanic features. The initial classification was conducted with an unsupervised clustering algorithm, while the final simplification of regions was based on the use of dendrograms and knowledge of important features. Spatial components were processed separately, while relative spatiotemporal components were combined into single spatiotemporal layers. The spatial components produced interesting results, with a high degree of regional partitioning in coastal regions. Spatiotemporal components produced much more complex regionalisation, although many important seasonal features were easily identified and regionalised.

As CSAT images were used without alteration, subsequent principle components can be interpreted in conjunction with loading signatures to indicate relative level of pixel CSAT concentration. Although producing and using the SST residuals enhances clarification of possible regions of interest, the underlying values become non interpretive to physical SST values (degrees Celsius). Colour contrast in SST residual components (Figure 2) is only indicative of the degree of difference from the maxima. Interpretation of SST residual components is possible with the help of the component loadings. However, the bulk of relevant features are identified through knowledge-based examination. For an indication of the relevance to real SST, the original images were correlated against the components. Although subsequent coefficients are much lower than the loadings with input residual images, it is still possible to see how representative components are of particular seasons.

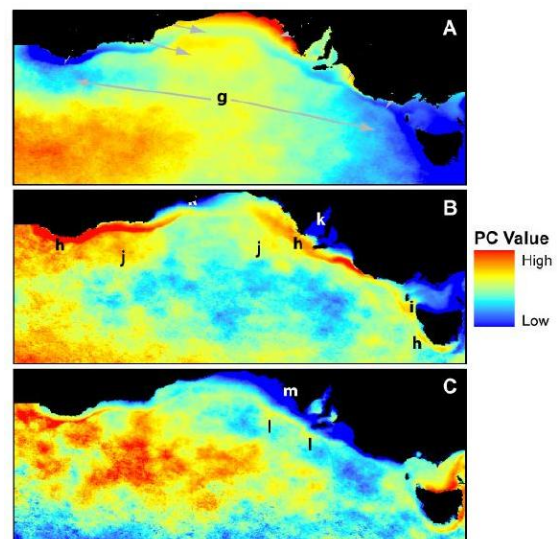


Figure 2: Southern Australian SST sPCA components: (A) Component 1 – Spatial Mean; (B) Component 2 – summer/winter dichotomy; (C) Component 3 – autumn/spring dichotomy

degree of delineation of features in the sPCA results. As the result shows regional similarities, a regional classification approach to defining boundaries was deemed the most appropriate for this study. Defining particular regional boundaries would then provide both a visual reference and an analysis tool. The potential vector dataset that can be produced from classification results could then be incorporated into vector dominated cadastral systems.

Two factors ultimately led to the final classification procedure used in this study. First, there was the identification of the sPCA components into spatial (component-1) and temporal (components 2 and 3). These representations led to a natural division for interpretable layers. Secondly, the efficiency of utilising classified layers for assessment or management purposes had to be taken into account. As a large number of separate layers can be unwieldy for a conclusive assessment, all possible effort was made to reduce the number of final layers. In this respect the combination of relevant SST and CSAT layers for a single classification was explored. This method would have resulted in a single layer for both the spatial and temporal dynamic representations. Although in the spatial realm this is a feasible option, the temporal dynamic realm proved more problematic. Important areas of high dynamic in one dataset were often cancelled by low dynamic values in the same area in the second dataset. For example, the Leeuwin Current, which is an important feature in the distribution of many commercial marine species, has a strong dynamic signature in SST components. However, due to its oligotrophic nature it is almost absent in CSAT dynamics. The combination of these two datasets tends to reduce the signature of the Leeuwin, which could lead to an assumption of reduced importance. On this factor it was decided to maintain SST and CSAT datasets separately.

The Iterative Self Organising Data Analysis (ISO) clustering technique was used in this study. This method iteratively classifies an image, recalculates statistics and reclassifies again. Cells are assigned to a specific class

Each existing regional polygon was given a unique identifier and the vector layer was converted back to a raster image with the same extent and resolution as the original. The final classified rasters were able to have statistics extracted per the individual regions, allowing distinct spatial regions to be assessed independently or in conjunction with surrounding regions.

In the final result of the classification process, relevant oceanographic regions were isolated. The Leeuwin Current in the west was separated into regions, reflecting its weakening signal as the current moves eastward. The northern Bass Strait showed the opposing influences from both western and eastern sources. Upwelling regions in the Great Australian Bight, and the Bonney Coast leading into Encounter Bay, were maintained as well as major eddy regions such as those South of the Leeuwin and in the central GAB. The classification results for southern Australia are shown in Figure 3 and the results for the west Australian coast are shown in Figure 4.

define it is a more important factor. In other words, a uniform degree of high variability within a region is not a problem, as long as the region's bounds delimit the region of variability relatively accurately.

using minimum spectral distance to class means. The initial first pass is based on arbitrary spectral means. The process is repeated until little change in classes occurs or, either a minimum spectral distance or a maximum number of iterations is reached. The raw classified images were imported into the ArcGIS environment and dendrograms were created. Using the grouping hierarchies in the relevant dendrograms and knowledge of key oceanographic systems, similar class groupings were assessed on repetitiveness and merged. One of the important knowledge-based assumptions in the classification process is the identification of spatially significant regions. Maintaining these often large areas while removing redundant regions can be as simple as eliminating regions under a certain area threshold. In this case these processes took several steps and were conducted in the vector environment after converting the classified raster images. Vector features were eliminated by being merged into a larger bordering feature which shared the longest border. Initially a hard threshold required all features with areas smaller than 0.5 decimal degree squared (dD2) to be eliminated. Features up to 0.1dD2 were assessed and despite a few exceptions all were eliminated. Notable were relatively small regions located in the open ocean regions. As the open ocean is subject to a high degree of spatio-temporal influences, some of the regions fail to spatially define a significantly common or periodic occurrence. Open ocean regions with depths greater than 500m were subjected to a larger area threshold of 1 dD2, although care was taken not to eliminate features attributed to islands and well known anomalies.

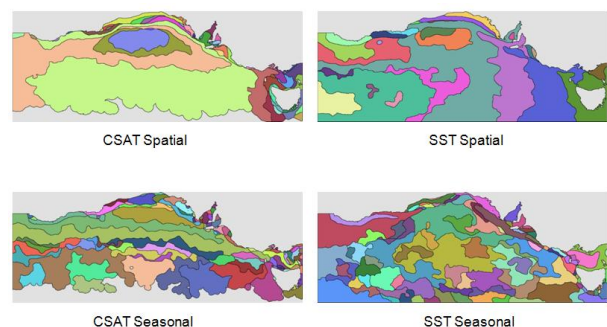


Figure 3: Southern Australian Boundaries

The verification of these results proved a problem since conventional accuracy assessment techniques did not apply. A pixel's relative fit to a region's character is not in itself an accurate measure of the suitability of regional boundaries. While interpreting error magnitude in the case of this project, it should be remembered that a region that evenly contains a high degree of internal heterogeneity is just as viable as a region that is heterogeneous. The degree to which the region bounds

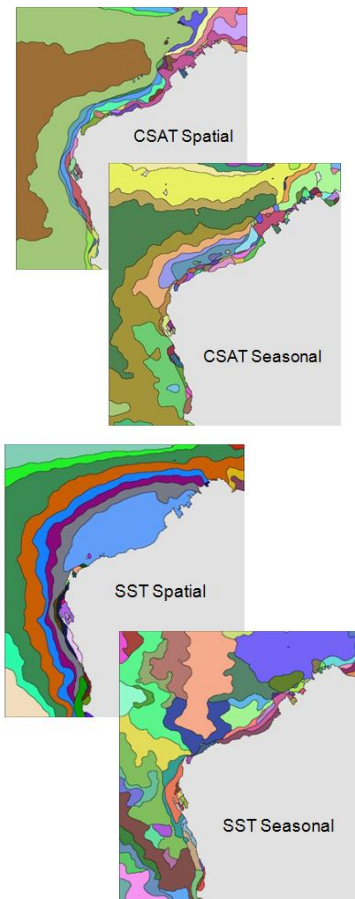


Figure Four: Australian West Coast Boundaries

The process used to assess the accuracy of the classification in this study is as follows: the basic regional mean (μ_x) and standard deviations (s_x) were calculated using the individual components used to derive regional boundaries. The standard deviation image gave a regional indication of areas that contained a large degree of variability. In order to gain a clearer pixel-based image of each pixel's association with its regional deviation, a binning operation was performed producing three bins classifying pixels into their relevant deviation (1, 2 and 3 or more) from the regional mean (e.g. Figure 5). This provided a discretely contrasted view of pixels that were further away from the regional mean character. Some pixels highlighted areas where the region boundary had been compromised through the ISO classification process using two components or had been amalgamated during the class reduction process using dendrograms. The pixel indicator of deviations from the regional mean in this form only indicates distribution characteristic within the classified region. Across the whole image all pixels that fall into the three or more deviations have the same value. However, should one want to reassess particular anomalies of the classification, then there is little indication of where best to start in order to find some degree of significance. To counter these problems two approaches were taken to give a pixel level indicator of its scale of regional variability over the whole regional scope. Essentially the desired result required a high score for a pixel in a high deviation bin in a region with a large deviation. A pixel close to the region's mean but in a region with a high deviation still scored high but ranked less. A region of low deviation and pixels close to the mean were ranked the lowest.

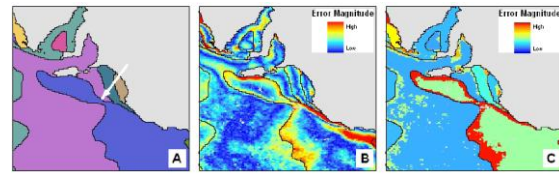


Figure Five: Local example showing A – Classes; B – Standardized region pixel deviation magnitude; and C – pixel deviation magnitude.

DISCUSSION

A comparison of these results with similarly derived regionalisation showed that the spatio-temporal representations have the potential to further inform marine spatial management regimes. It was also shown that the delimitation of natural phenomena is feasible and can have seasonal dynamics represented. This provides a potential baseline for near real-time boundary assessments and an addition to adaptive management techniques. Ultimately the results have shown that there is scope for the improved incorporation of marine dynamics into marine spatial management systems.

The methodology presented in this study has provided a simple framework for the identification and delimitation of both spatial averages and spatio-temporal trends in single oceanic properties, namely SST and CSAT. As well as creating a systematic approach, the componentised nature provides scope for added complexity and changes to individual components. This approach reduces major redundancy issues and need for potentially frequent large scale redesign, and is far more suitable for a process which is likely to make advances over the near future. Similar approaches are being developed in a number of other scientific fields such as bioregionalisation that may eventually feed into the overall marine cadastre. The methodology presented can also utilise other sources of satellite data (or any raster data) that can be at any number of spatial resolutions. The key requirements for suitable results would be a long enough time-series of data and input datasets to be of equal resolution and extents.

REFERENCES

- Cheng, Q., Jing, L. & Panahi, A. (2006) Principal component analysis with optimum order sample correlation coefficient for image enhancement. *International Journal of Remote Sensing*, 27, 3387 - 3401.
- Eklundh, L. & Singh, A. (1993) A comparative analysis of standardised and unstandardised principle components analysis in remote sensing. *International Journal of Remote Sensing*, 14, 1359 - 1370.
- Hirosawa, Y., Marsh, S. E. & Lkliman, D. H. (1996) Application of standardized principal component analysis to land-cover characterization using multitemporal AVHRR data. *Remote Sensing of the Environment*, 58, 267 - 281.

- Kaplan, A., Cane, M. A., Kushnir, Y., Clement, A. C., Blumenthal, M. B. & Rajagopalan, B. (1998) Analyses of global sea surface temperature 1856 - 1991. *Journal of Geophysical Research, Oceans*, 103, 18567 - 18589.
- Lehodey, P., Bertignac, M., Hampton, J., Lewis, A. & Picaut, J. (1997) El Niño Southern Oscillation and tuna in the western Pacific. *Nature*, 389.
- Mann, K. H. & Lazier, J. R. N. (2006) *Dynamics of Marine Ecosystems: Biological-Physical Interactions in the Oceans*, U.K., Blackwell.
- Piowar, J. M. & Millward, A. A. (2002) Multitemporal change analysis of multispectral imagery using principal components analysis. *International Geoscience and Remote Sensing Symposium 2002 (IGARSS '02)*. Toronto.
- Ryther, J. H. & Yentsch, C. S. (1957) The estimation of phytoplankton production in the ocean from chlorophyll and light data. *Limnology and Oceanography*, 2, 281–286.
- Singh, A. & Harrison, A. (1985) Standardized principle components. *International Journal of Remote Sensing*, 6, 883 - 896.
- Vlahakis, G. N. & Pollatou, R. S. (1993) Temporal variability and spatial distribution of sea surface temperatures in the Aegean Sea. *Theoretical and Applied Climatology*, 47, 15 - 23.
- Yoder, J. A. & Kennelly, M. A. (2003) Seasonal and ENSO variability in global ocean phytoplankton chlorophyll derived from 4 years of SeaWiFS measurements. *Global Biogeochemical Cycles*, 17, 23:1 - 23:14.
- Yoder, J. A. & Kennelly, M. A. (2006) What have we learned about ocean variability from satellite ocean color imagers. *Oceanography*, 19, 152-171.