Feature-based classification of optical water types in the Northwest Atlantic based on satellite ocean color data

Linda V. Martin Traykovski and Heidi M. Sosik

Biology Department, Woods Hole Oceanographic Institution, Woods Hole, Massachusetts, USA

Received 12 October 2001; revised 15 May 2002; accepted 18 July 2002; published 21 May 2003.

[1] We have developed an optical water type classification approach based on remotely sensed water leaving radiance, for application to the study of spatial and temporal dynamics of ecologically and biogeochemically important properties of the upper ocean. For CZCS and SeaWiFS imagery of the Northwest Atlantic region, pixels from several different locations projected into distinct clusters in water-leaving radiance feature space, suggesting that these waters can be distinguished using a few spectral bands of ocean color data. Based on these clusters, we constructed a Northwest Atlantic Training Set and developed two different classification techniques. The Euclidean Distance Classifier minimizes the raw distance between each pixel and the centroid of the class to which it is assigned, whereas the Eigenvector Classifier is based on scaling the raw distances by the variance of each class, thereby accounting for the shape of each class in feature space. We conducted an initial evaluation of these two classification techniques by constructing water type classes based on only half of the pixels of each water type (randomly selected) in the Northwest Atlantic Training Set; classification was then carried out on the remaining half of the training set data. Applying the Euclidean Distance Classifier resulted in an average of 97.4% correctly classified pixels over 20 trials; even higher success rates were achieved with the Eigenvector Classifier, which gave an average of 99.1% correctly classified pixels. The Euclidean Distance Classifier performed well with spherical classes, but with more ellipsoidal classes, classification success improved considerably using the Eigenvector Classifier. We then applied these classifiers to ocean color images of the Northwest Atlantic to elucidate the geographical location and extent of each water type. We interpreted classifier results based on our Classification Goodness of Fit measure, which indicates how closely a given pixel is associated with its assigned class. This revealed that sharp boundaries exist between water masses of different optical types, with pixels on either side of the boundaries being strongly associated with their water type class. We anticipate that our classification techniques will facilitate long-term time series studies by tracking optical water types through seasonal and interannual changes. *INDEX TERMS:* 4847 Oceanography: Biological and Chemical: Optics; 4275 Oceanography: General: Remote sensing and electromagnetic processes (0689); 4894 Oceanography: Biological and Chemical: Instruments and techniques; 4283 Oceanography: General: Water masses; *KEYWORDS:* classification, ocean color, optical water type, SeaWiFS, Northwest Atlantic

Citation: Martin Traykovski, L. V., and H. M. Sosik, Feature-based classification of optical water types in the Northwest Atlantic based on satellite ocean color data, *J. Geophys. Res.*, *108*(C5), 3150, doi:10.1029/2001JC001172, 2003.

1. Introduction and Background

[2] Remotely sensed ocean color data acquired with satellite sensors such as SeaWiFS (Sea-viewing Wide Field-of-view Sensor) and CZCS (Coastal Zone Color Scanner) provide temporally resolved synoptic views of ocean regions over long periods of time. The availability of satellite ocean color data has drastically altered our perception of the global ocean over the last 2 decades, and has opened up new opportunities to study the spatial and temporal variability of phytoplankton distributions. Traditionally, understanding of regional- and global-scale variability in phytoplankton abundance has been limited by sampling confined to point measurements obtainable from ship-based surveys. The CZCS sensor produced the first views of the surface ocean not subject to these sampling limitations; in combination with sea surface temperature imagery, these views have provided great insights into biological and physical interactions in the upper ocean. CZCS data has proven useful for studies of pigment dynamics at a variety of scales and for a wide range of environments [see *McClain*, 1993] and has set the stage for new and improved applications with current and future satellite ocean color data.

0148-0227/03/2001JC001172$09.00

[3] As evidenced by the successful use of CZCS imagery to study pigment distributions, the abundance of phytoplankton pigments plays an important role in determining spectral sea surface reflectance. Particularly in coastal waters however, both inherent and apparent optical properties are influenced by a wide array of physical, biological and chemical processes. These processes can lead to large sources of optical variability that may be independent of the abundance of phytoplankton pigments. In addition to these pigments, constituents such as colored dissolved organic matter (CDOM) of both marine and terrigenous origin, heterotrophic organisms, biological detritus, and inorganic particulate material can affect both the magnitude and spectral quality of reflected light. This complexity may interfere with accurate estimation of phytoplankton distributions based on optical signatures; however, it also presents the potential for deriving information about other water properties from space.

1.1. Pigment Retrieval

[4] Because ocean color signals vary in response to many factors, successful identification of optically different types of water is necessary for accurate retrieval of constituent concentrations. To the extent that the abundance and properties of absorbing and scattering materials are not constant and do not covary in space and time, application of generic or standard algorithms for pigment retrieval may result in errors. Algorithms used to estimate pigment concentrations from remotely sensed ocean color data typically rely on ratios of water-leaving radiance (L_w) or remote-sensing reflectance (R_{rs}) in the visible spectral bands, usually $L_w(443)/L_w(550)$ and/or $L_w(443)/L_w(520)$ for CZCS [e.g., *Clark*, 1981; *Gordon et al.*, 1983], and $R_{rs}(443)/R_{rs}(555)$, $R_{rs}(490)/R_{rs}(555)$, and/or $R_{rs}(510)/R_{rs}(555)$ for SeaWiFS [e.g., *O'Reilly et al.*, 1998]. While this approach has been very fruitful, these ratios can vary in response to factors besides chlorophyll concentration.

[5] Systematic errors in pigment retrieval from remotely sensed ocean color have been documented for a variety of specific conditions. In coastal environments with high levels of CDOM, pigment concentration can be overestimated [*Carder et al.*, 1989; *Hochman et al.*, 1994]. Even in waters optically dominated by phytoplankton, application of generic pigment algorithms can lead to overestimates or underestimates of pigment concentration due to differences in the spectrum of water-leaving radiance per unit of pigment biomass. This can result from pigmentation differences such as the presence of high surface concentrations of phycobilipigments in a bloom of cyanobacteria [*Sathyendranath*, 1986] or from physiological and ecological effects such as chronic low light acclimation in high latitude environments [*Mitchell and Holm-Hansen*, 1991; *Sosik et al.*, 1992]. Species-specific properties such as the production of highly scattering loose coccoliths by some types of coccolithophorids can also affect pigment estimates [*Holligan et al.*, 1983; *Balch et al.*, 1989, 1991]. All of these potential errors in pigment retrieval can be avoided if specific water types can be identified and appropriate algorithms or corrections applied. The application of optical water type classification techniques to ocean color imagery in problem regions will contribute to this effort.

1.2. Resolving Mesoscale Features From Satellite Ocean Color Observations

[6] Satellite ocean color images of large geographic areas often reveal mesoscale reflectance features that are associated with physical, biogeochemical, and biological processes in the upper ocean. Ocean color data has been exploited to help identify the scales associated with these features, and attempts have been made to correlate this data with in situ observations to identify the processes contributing to spatial and temporal variability. Examples where CZCS observations have been used for these purposes include studies of boundary current systems [*Peláez and McGowan*, 1986; *Smith et al.*, 1988; *Denman and Abbott*, 1988; *Thomas et al.*, 1994], coastal upwelling regions [*McClain et al.*, 1984; *Abbott and Zion*, 1985], and other dynamic coastal environments [*Eslinger and Iverson*, 1986; *Yoder et al.*, 1987; *Abbott and Zion*, 1987; *McClain et al.*, 1990], including areas dominated by river plumes [*Müller-Karger et al.*, 1989; *Hochman et al.*, 1994]. These types of studies provide fundamental information about upper ocean processes and can also be useful for improving the accuracy with which satellite data is interpreted, as is the case in high CDOM shelf waters [e.g., *Hochman et al.*, 1994].

[7] Identification of mesoscale features and local- to regional-scale water masses is also important for many biological and ecological questions. The study of pigment dynamics is an obvious example. In addition, the concept of biogeographic regions whose boundaries may vary in space and time is important for assessment of rate processes such as primary production using remotely sensed data [e.g., *Platt and Sathyendranath*, 1988; *Platt et al.*, 1991]. Accurate delineation of these regions is required, and methods to accomplish this which take advantage of the ideal spatial and temporal resolution of satellite-based observations should be far superior to those based on climatological observations. This idea had been explored, for example, in the use of CZCS data to identify biogeographic regions in the Gulf of California [*Santamarià-del-Angel et al.*, 1994], as well as by using Advanced Very High Resolution Radiometer (AVHRR) imagery in combination with local bathymetry to define water types in a study of productivity on Georges Bank [*Sathyendranath et al.*, 1991]. To date, efforts to identify mesoscale features or water type boundaries from remotely sensed ocean color data have generally relied only on pigment distributions or have involved relatively dramatic water type differences, such as those that occur near river plumes. The potential for using more information than is contained in pigment images and to discern more subtle differences in optical water types has not been fully explored.

1.3. Detecting Phytoplankton Blooms

[8] There have been efforts to use CZCS data for water type identification using specialized algorithms designed to recognize the unique optical properties of a particular type of phytoplankton. A successful method was developed to detect coccolithophore blooms using CZCS remotely sensed radiances based on a nonparametric parallelepiped super-

vised algorithm [*Brown and Yoder*, 1994a, 1994b], which was able to distinguish pixels within coccolithophore blooms from nonbloom pixels. *Subramaniam and Carpenter* [1994] developed a protocol to identify *Trichodesmium* blooms from CZCS imagery based on high reflectivity from gas vacuoles and a phycoerythrin absorption feature at 550 nm, and were able to distinguish two *Trichodesmium* blooms from sediment whitings and from some portions of coccolithophore blooms. Attempts have been made to detect cyanobacterial blooms using a supervised classification technique [*Zabicki*, 1995] based on the observed ratio of (total radiance at 750 nm) to (total radiance at 670 nm); although this ratio was always lower for suspected cyanobacterial blooms than for sediment conditions, it was not possible to distinguish coccolithophore blooms from *Trichodesmium* blooms with this method.

[9] These taxon-specific algorithms can indicate the presence of near mono-specific blooms in the analysis of particular ocean regions at times when blooms of that type are thought to occur. The utility of these approaches may be limited, however, in the identification and classification of a broad range of water types that may span many scales of spatial and temporal variability. To fully exploit ocean color data for the study of phytoplankton dynamics, it is necessary to develop a more universal scheme to optically classify many different types of phytoplankton blooms simultaneously by automatically distinguishing them from each other and from other nonphytoplankton dominated optical water types.

1.4. Northwest Atlantic Ocean

[10] The Northwest Atlantic Ocean has been the subject of intensive oceanographic study over the past several decades, and continues to be a target location for conducting multidisciplinary optical and ecological research. It encompasses a wide variety of oceanographic and optical regimes (Figure 1), including coastal and coastally influenced regions (the Gulf of Maine, Georges Bank, and the nearshore Central Mid-Atlantic Bight), open ocean waters (the Sargasso Sea), and a swiftly flowing western boundary current (the Gulf Stream). The coastally influenced regions exhibit the greatest seasonal optical variability, since both colored matter of terrigenous origin and phytoplankton biomass affect the inherent and apparent optical properties in these areas. Georges Bank, a shallow submarine bank lying along the outer continental shelf east of Cape Cod, Massachusetts, has been well studied both biologically and physically and is influenced by a wide range of important coastal processes which can contribute to optical variability [see *Backus*, 1987]. The Gulf of Maine is a semi-enclosed region bounded by shallow embankments including Georges Bank, and is tidally flushed principally through two deeper channels (Northeast Channel and Great South Channel). The Gulf encompasses several deep canyons, which remain stratified year-round, as compared to the seasonally well-mixed shallower zones [*Yentsch and Garfield*, 1981]. Optically, these shallow coastal waters are complex and the absolute accuracy of ocean color-based pigment estimates for the Georges Bank/Gulf of Maine region is limited by inadequate knowledge of the magnitude and sources of optical variability [e.g., see *Yentsch et al.*, 1994].

[11] To the south lies the oligotrophic, open ocean region of the Sargasso Sea, characterized by optically clear "blue" waters. This region has typically been categorized as Case I waters [after *Morel and Prieur*, 1977], and optical variability is more limited here than in the more coastally influenced Case II waters to the north. The Sargasso Sea is bounded to the west and north by the Gulf Stream, a powerful current whose structure, transport, and water mass properties vary both temporally and spatially, displaying seasonal, interannual, and mesoscale fluctuations. The path of the Gulf Stream meanders over a broad range, seasonally shifting considerable distances north/south, as well as exhibiting rapidly translating and evolving meanders on timescales as short as several days [*Watts*, 1983]. Frontal mesoscale eddies are formed when these meanders are cut off from the Stream, resulting in a ring of Gulf Stream water enclosing a body of water with very different physical, biological, and optical properties; the evolution of these eddies influences optical variability in the Sargasso Sea, the slope waters of the Mid-Atlantic Bight, and as far north as Georges Bank [*Ryan et al.*, 2001].

1.5. Classification Techniques

[12] A promising approach to identifying optical water types based on remotely sensed data is to develop a comprehensive framework within which different water types may be automatically and simultaneously distinguished from each other. Subsequently, additional information such as in situ observations can be used to categorize the water types in an ecologically relevant manner. The development of an automatic classification scheme essentially involves the inversion of observed data to retrieve a property of interest. One of us (Martin Traykovski) has previously explored the utility of both model-based and feature-based inversion techniques for the purposes of classifying the acoustic echoes from different types of zooplankton based on their unique spectral signatures. Classification schemes based on theoretical models proved reliable in high signal-to-noise conditions, whereas a feature-based classification approach was more robust in the presence of contaminating noise [*Martin et al.*, 1996]. Although these techniques have been useful in zooplankton studies [*Holliday et al.*, 1989; *Martin et al.*, 1996; *Martin Traykovski et al.*, 1998a, 1998b] as well as other oceanographic applications [e.g., *Munk and Wunsch*, 1979; *Fox et al.*, 1994], their utility for classifying optical water types from satellite spectral data has not been fully explored.

[13] Feature-based classification operates independently of a forward model, and relies only on the inherent characteristics of the observed data. It involves identifying and extracting the relevant features in the data (direct measurements and/or properties derived from these measurements) that will allow retrieval of the properties of interest. Once the relevant feature set is identified, it can then be employed to delineate classes or categories which share common properties or attributes. Data may subsequently be classified into categories based on these discriminating features. Statistical decision theory can be used to derive an optimum classification rule if multivariate probability density functions are known for each class or if an empirical probability model can be obtained by statistical estimation. Alternatively, a decision rule may be derived directly from the distribution of samples in feature space. Feature-based classification is also possible where no a

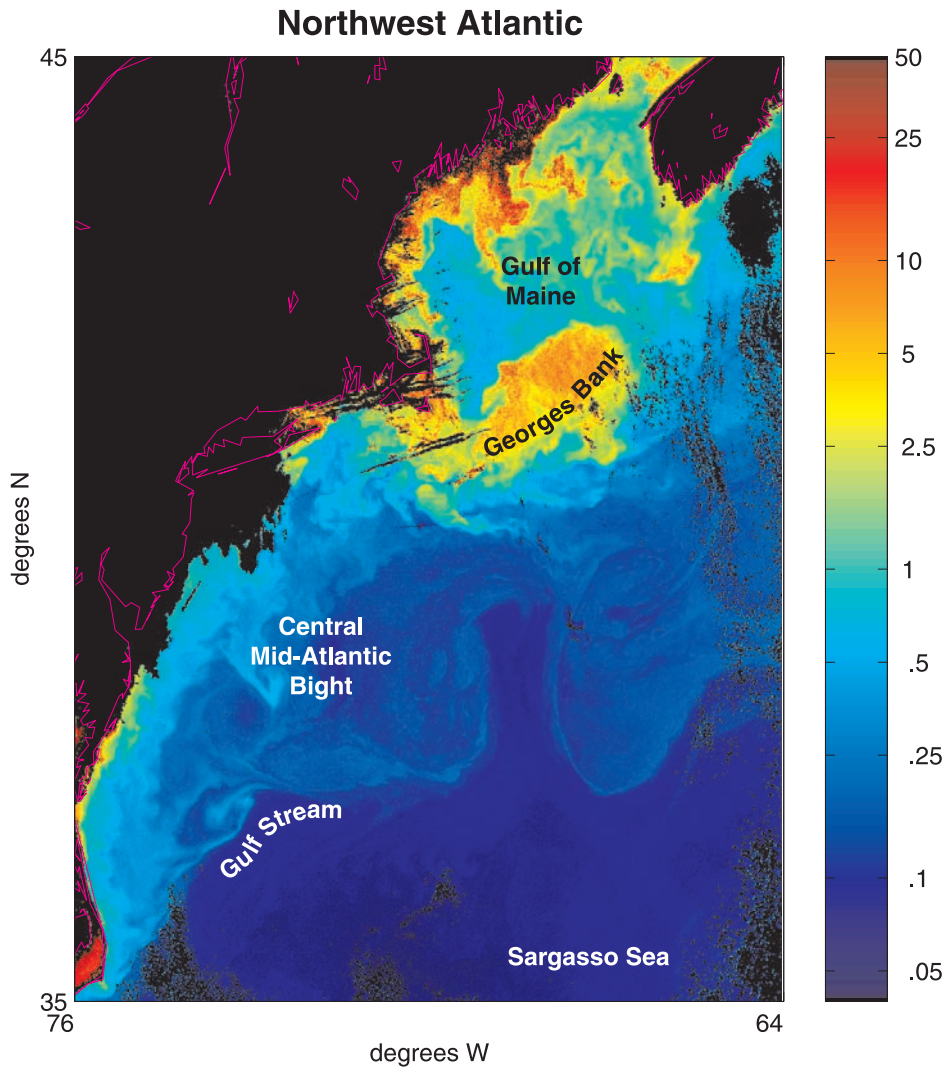

Figure 1. Optical variability in the Northwest Atlantic Ocean as revealed by satellite-derived chlorophyll *a* concentration (mg/m^3) from SeaWiFS on 8 October 1997, with oceanographic/optical regimes indicated. Land and clouds appear black; coastline is shown in pink. Low ratios of blue to green water leaving radiance (indicating high pigment concentrations) are often observed over Georges Bank. Episodic blooms of coccolithophorid phytoplankton [*Balch et al.*, 1991; *Brown and Yoder*, 1994a] as well as toxic dinoflagellates [*Anderson*, 1997] have been documented in the Gulf of Maine, contributing to optical variability in that region. Although phytoplankton biomass is believed to be the dominant variable constituent controlling the optical properties of the Sargasso Sea, recent studies have revealed the importance of CDOM and particulate matter in contributing to light attenuation in this region [*Siegel and Michaels*, 1996; *Nelson et al.*, 1998]. As a result of their unique properties, the formation and evolution of cold- and warm-core rings introduce a significant source of optical variability into the Sargasso Sea and the slope waters of the Mid-Atlantic Bight, with influences as far north as Georges Bank [*Ryan et al.*, 2001].

priori information is available, through the exploitation of naturally occurring groupings or clusters in the data. Features can then be empirically related to water type. Previous work on phytoplankton bloom identification [e.g., *Brown and Yoder*, 1994a, 1994b; *Subramaniam and Carpenter*, 1994; *Zabicki*, 1995] is a limited form of feature-based classification where a decision rule is applied to determine whether data fall inside or outside a single class boundary.

[14] In this paper we present the foundations for a more comprehensive feature-based approach to the optical classification of water type, and outline the development of specific classification techniques for application to remotely sensed ocean color data. The work presented here involves a regional study of waters in the Northwest Atlantic, including the Gulf of Maine and Georges Bank to the north, the Central Mid-Atlantic Bight, and the waters of the Gulf Stream and the northern Sargasso Sea to the south. We begin by presenting an analysis of selected CZCS and SeaWiFS data to demonstrate the feasibility of applying feature-based classification techniques to identify and delineate optical water types. The ultimate goal of our regional study is to establish feature-based classification approaches for discrim-

inating various optical water types present in the Northwest Atlantic based on satellite ocean color data. To this end, we have developed two different classification techniques and applied them to CZCS and SeaWiFS imagery of the region.

2. Methods

[15] We have automated the water type classification process by applying statistical decision criteria to define class boundaries and assign data to a particular class. In this section, we first describe the theoretical development of two feature-based classification techniques new to the analysis of ocean color data, the Euclidean Distance Classifier and the Eigenvector Classifier. Next we outline the approach we used to construct a classifier training set for the Northwest Atlantic region. Following this, we describe how the classifiers were applied to CZCS and SeaWiFS imagery of the Northwest Atlantic. Finally, we detail the development of a Classifier Goodness of Fit measure for application to the results of the Euclidean Distance Classifier.

2.1. Development of Classification Techniques

[16] Feature-based classification approaches separate observed data into classes based on the variability inherent in the data set. The data are generally multivariate, so that each data point has *nk* dimensions or features. In some cases, it may be desirable to chose the number of classes based on a priori knowledge, although this is not necessary for successful classification. For the classifiers described herein, the number of classes *ni* was predetermined based on a limited set of satellite observations, and a training set of data points $\mathbf{A}^{(i)}$ was constructed for each class. For each water type class *i* with *nj* training data points (each with *nk* features), the class centroid $\mathbf{m}^{(i)}$ is computed as the arithmetic *nk*-dimensional mean of the training data points $\mathbf{a}_j^{(i)}$ for that class, so that

$$\mathbf{m}_k^{(i)} = \left[\sum_{j=1}^{nj} \mathbf{a}_{jk}^{(i)}\right] / nj; \tag{1}$$

that is, the *k*th dimension of the centroid for class *i* is the mean of the *k*th dimension $\mathbf{a}_{jk}^{(i)}$ of all *nj* training data points in that class. The classification process involves projecting the multivariate data into an *nk*-dimensional feature space. Each *nk*-dimensional satellite data point $\mathbf{s}_v$ (for $v = 1$ to *nv*, the number of pixels in the image) may then be classified by evaluating its proximity to each class centroid according to a distance-based decision rule.

[17] The Euclidean Distance Classifier assigns each $\mathbf{s}_v$ to a water type class *i* based on the Euclidean distance between that data point and the centroid or mean of each class in feature space. The Euclidean distance D between data point $\mathbf{s}_v$ and the centroid for class *i* $\mathbf{m}^{(i)}$ is defined as

$$\mathrm{D}\left(\mathbf{s}_v, \mathbf{m}^{(i)}\right) = \left[\left(\mathbf{s}_v - \mathbf{m}^{(i)}\right)\left(\mathbf{s}_v - \mathbf{m}^{(i)}\right)^{\mathrm{T}}\right]^{1/2}. \tag{2}$$

A data point is assigned to the water type whose class centroid is the minimum Euclidean distance away, so that the decision rule for the Euclidean Distance Classifier may be written as

$$\mathbf{s}_v \in \text{class } a \text{ iff } \mathrm{D}\left(\mathbf{s}_v, \mathbf{m}^{(a)}\right) < \mathrm{D}\left(\mathbf{s}_v, \mathbf{m}^{(b)}\right) \text{for all } b \neq a. \tag{3}$$

The class decision boundaries for the Euclidean Distance Classifier are equidistant from the class centroid in all directions. An implicit assumption of this classifier is that all classes are spherical in shape; as a result, this formulation is not able to account for any variability in class shape.

[18] The Eigenvector Classifier was developed to include consideration of the *nk*-dimensional shape of each class. With this formulation, each water type class is defined in terms of a hyper-ellipsoid in feature space, the principal axes of which are given by the *nk* dominant eigenvectors ($\varphi_1, \varphi_2, \varphi_3, \ldots \varphi_{nk}$) of the covariance matrix $\mathbf{C}$ [see *Papoulis*, 1991] of the training data in that class,

$$\mathbf{C}^{(i)} = \left[\tilde{\mathbf{A}}^{(i)}\right]^{\mathrm{T}}\left[\tilde{\mathbf{A}}^{(i)}\right], \tag{4}$$

where each row of $\tilde{\mathbf{A}}^{(i)}$ contains a mean-subtracted, energy normalized *nk*-dimensional data point $\tilde{\mathbf{a}}_j^{(i)}$ from the training set for class *i*, and $0 \leq \mathbf{C}_{kk}^{(i)} \leq 1$. The extent of each class in the *nk* eigenvector directions is represented by the corresponding eigenvalues ($\lambda_1, \lambda_2, \lambda_3, \ldots \lambda_{nk}$). The result is that the classes occupy ellipsoidal regions in feature space, each oriented along their own eigenvector directions. With the Eigenvector Classifier, the distance E of a data point $\mathbf{s}_v$ from the centroid $\mathbf{m}^{(i)}$ of class *i* is computed in terms of components along the eigenvector directions for that class, and each component is scaled by the square root of the corresponding eigenvalue,

$$\mathrm{E}\left(\mathbf{s}_v, \mathbf{m}^{(i)}\right) = \left[\sum_{k=1}^{nk}\left(\mathbf{s}_{vk}^{*(i)} / \sqrt{\lambda_k}\right)^2\right]^{1/2}, \tag{5}$$

where $\mathbf{s}_v^{*(i)}$ is the distance between $\mathbf{s}_v$ and the centroid in the eigenvector reference frame,

$$\mathbf{s}_v^{*(i)} = \mathbf{R}\left[\mathbf{s}_v - \mathbf{m}^{(i)}\right]. \tag{6}$$

The data are projected into the eigenvector reference frame via the $nk \times nk$ transformation matrix $\mathbf{R}$, which is simply a matrix of dot products between unit row vectors (representing each of the *nk* dimensions in the original feature space) and the eigenvectors ($\varphi_1, \varphi_2, \varphi_3, \ldots \varphi_{nk}$). For example, with $nk = 3$ (three-dimensional case), $\mathbf{R}$ is the 3×3 matrix

$$\mathbf{R} = \begin{bmatrix} (\hat{\mathbf{x}} \cdot \varphi_1) & (\hat{\mathbf{y}} \cdot \varphi_1) & (\hat{\mathbf{z}} \cdot \varphi_1) \\ (\hat{\mathbf{x}} \cdot \varphi_2) & (\hat{\mathbf{y}} \cdot \varphi_2) & (\hat{\mathbf{z}} \cdot \varphi_2) \\ (\hat{\mathbf{x}} \cdot \varphi_3) & (\hat{\mathbf{y}} \cdot \varphi_3) & (\hat{\mathbf{z}} \cdot \varphi_3) \end{bmatrix}, \tag{7}$$

with $\hat{\mathbf{x}} = [1\ 0\ 0]$, $\hat{\mathbf{y}} = [0\ 1\ 0]$, and $\hat{\mathbf{z}} = [0\ 0\ 1]$. The scaled distance, measured as the number of standard deviations ($\sqrt{\lambda_k}$) between the data point and each class centroid, is computed in the transformed eigenvector space. A data point is assigned to the water type class for which this scaled distance is minimum, so that the decision rule for the Eigenvector Classifier may be written as

$$\mathbf{s}_v \in \text{class } a \text{ iff } \mathrm{E}\left(\mathbf{s}_v, \mathbf{m}^{(a)}\right) < \mathrm{E}\left(\mathbf{s}_v, \mathbf{m}^{(b)}\right) \text{for all } b \neq a. \tag{8}$$

2.2. Construction of a Training Set

[19] In order to apply the classifiers to satellite ocean color data of a specific region, a training set for that region

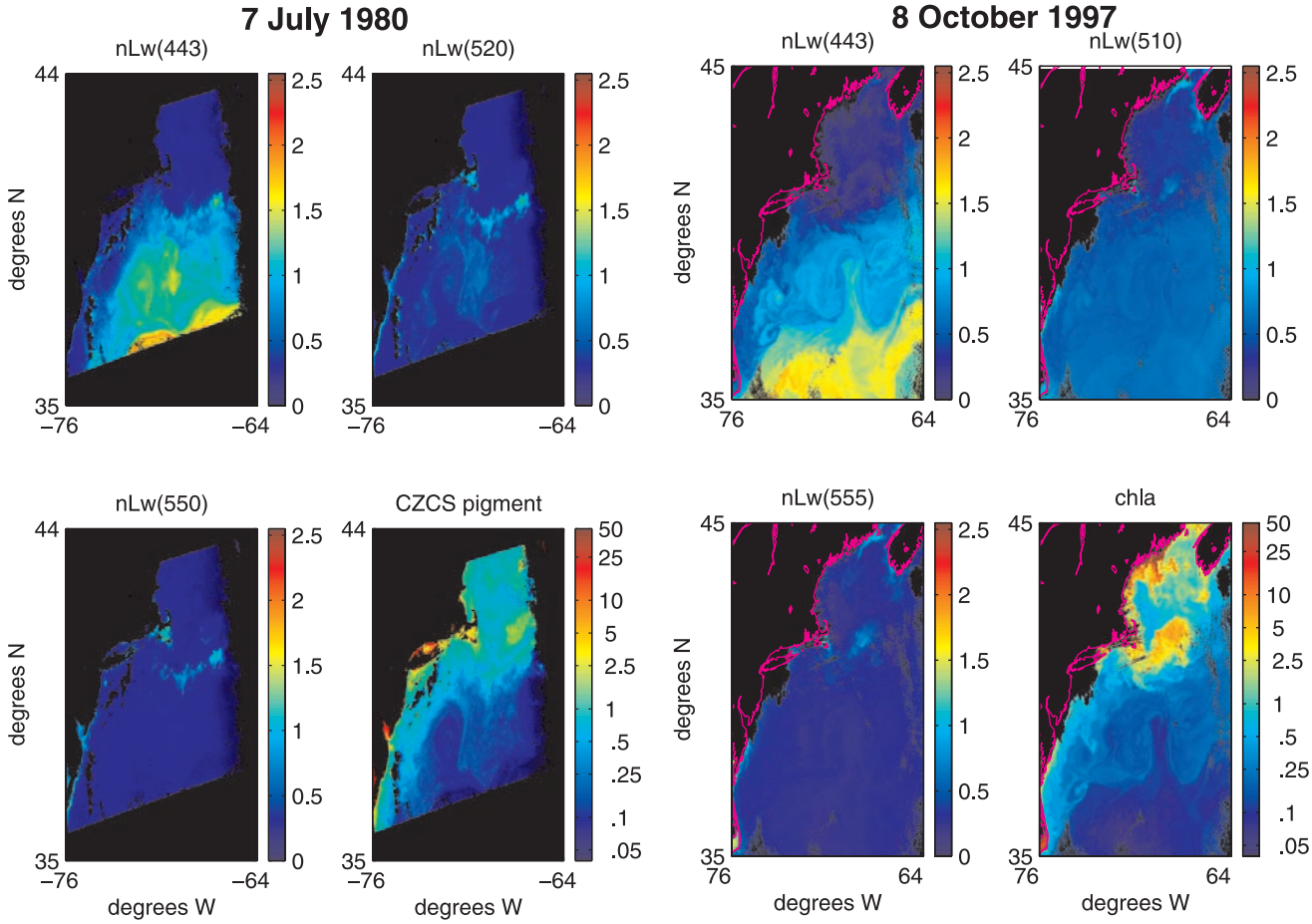

Figure 2. Satellite ocean color imagery of the Northwest Atlantic. (left) CZCS-derived normalized water-leaving radiances (mW/(sr·cm^2·μm)) in three spectral bands nLw(443), nLw(520), nLw(550), and pigment (mg/m^3) for 7 July 1980, and (right) SeaWiFS-derived normalized water-leaving radiances (mW/(sr·cm^2·μm)) in three spectral bands nLw(443), nLw(510), nLw(555), and chlorophyll a (mg/m^3) for 8 October 1997. Land and clouds appear black. Coastline is shown in pink (right-hand plots only) to delineate land under cloud cover.

must be constructed. The training set consists of a group of prototypical data points that represent the characteristics of each class, spanning the observed variability in feature values while maintaining class separability in feature space. Our analysis of optical water types in the Northwest Atlantic was focused on the region bounded by the Gulf of Maine to the north and the northern Sargasso Sea to the south, extending from 35°N to 45°N latitude, and 64°W to 76°W longitude. Based on general knowledge of the hydrography and bathymetry of the region, combined with an examination of CZCS and SeaWiFS imagery of the Northwest Atlantic (for examples, see Figure 2), we subjectively selected six geographic locations thought to have characteristic optical water types: the Northern Sargasso Sea (NSS), the waters in and around the northern boundary of the Gulf Stream (GS), the slope waters of the Central Mid-Atlantic Bight (CMAB), mineral-dominated waters, represented by a suspected coccolithophore bloom just south of Georges Bank (coc/min), the waters over Georges Bank (GB), and central Gulf of Maine waters (GM). To construct the Northwest Atlantic Training Set, we randomly selected 100 training data pixels from each of the six locations, and extracted normalized water-leaving radiances (nLw [*Gordon et al.*, 1988]) for each pixel from the 7 July 1980 CZCS image (Figure 2, left). For each training data point, three features were extracted from the ocean color imagery: nLw in each of three bands, 443 nm, 520 nm, and 550 nm. Thus, for this analysis of the Northwest Atlantic region, the training set $\mathbf{A}^{(i)}$(for i = 1 to ni) for each of the ni = 6 optical water type classes consisted of nj = 100 prototype data points, each represented by nk = 3 features: nLw(443), nLw(520), and nLw(550). We then projected the training set in a three-dimensional feature space (Figure 3). Class decision boundaries were computed for the Euclidean Distance Classifier (Figure 4, left), and class ellipsoids (spanning 1 standard deviation in all directions) were computed for the Eigenvector Classifier (Figure 4, right).

2.3. Classifier Application to Ocean Color Imagery

[20] To evaluate the performance of the two classification techniques, it is necessary to apply them to a test data set for which class membership is known. To this end, we constructed a test training set by randomly selecting half the

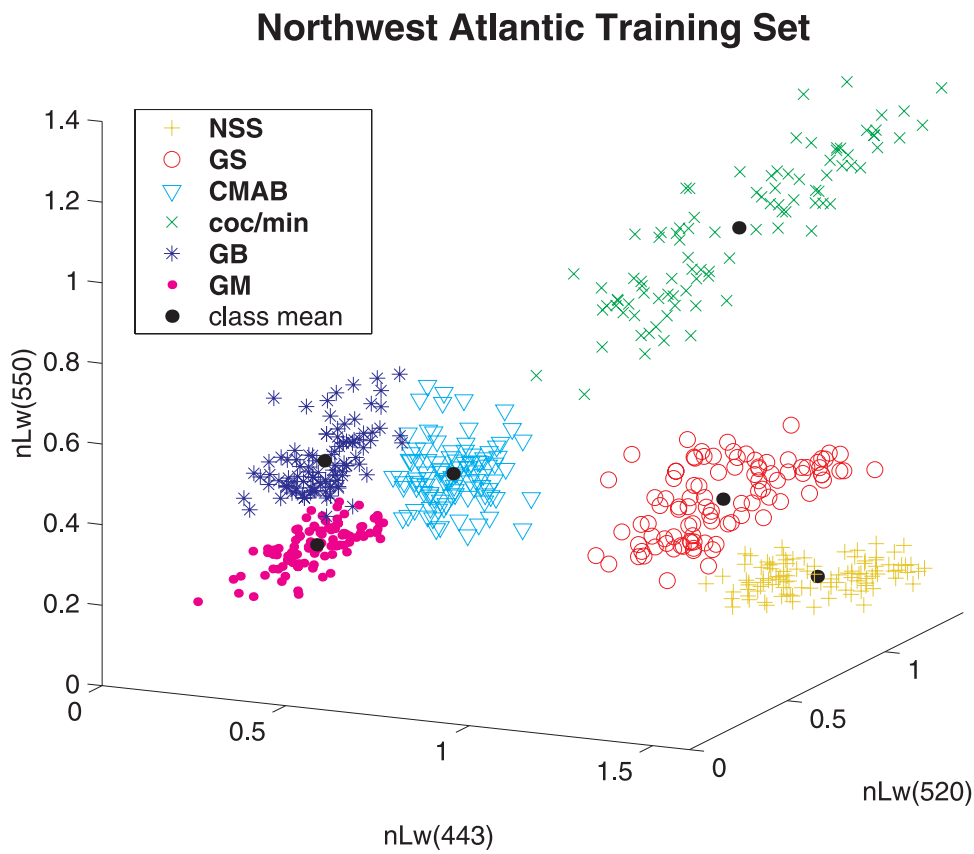

Figure 3. Northwest Atlantic Training Set. Water type classes shown in a three-dimensional single band feature space; nLw in mW/(sr·cm^2·μm). The training set consists of nj = 100 randomly chosen pixels from each of six different geographical regions. NSS = Northern Sargasso Sea waters; GS = Gulf Stream waters; CMAB = Central Mid-Atlantic Bight waters; coc/min = coccolithophore/mineral dominated waters; GB = Georges Bank waters; GM = Gulf of Maine waters; class mean $\mathbf{m}^{(i)}$ indicated as black dot at centroid of each cluster.

pixels (nj = 50) of each water type from the Northwest Atlantic Training Set. The test data set then consisted of the remaining pixels from each water type class, whose identity was known. We conducted 20 trials in which a different test training set was randomly selected, and the remainder of the pixels were assigned to the test data set. For each trial, each of the two classifiers was applied to the resulting test data set. We measured classifier performance as the percentage of the test data set correctly classified in each trial.

[21] We then applied the classifiers to satellite ocean color imagery of the Northwest Atlantic region. We extracted nLw at 443 nm, 520 (or 510) nm and 550 (or 555) nm (nk = 3) from the CZCS or SeaWiFS images for all pixels in the area from 35°N to 45°N latitude and 64°W to 76°W longitude. The classifier decision rules were applied to each pixel in the image, assigning it to one of the six water type classes. A map was produced of the classification results which indicates the class membership of each pixel from the original image.

2.4. Classification Goodness of Fit Measure

[22] A measure of Classification Goodness of Fit was developed to aid in the interpretation of the classification results. To explore the utility of this approach, we focused our efforts on the Euclidean Distance Classifier, since this formulation was less complex mathematically than the Eigenvector Classifier. A goodness of fit measure reveals the certainty with which each pixel is assigned to a given class. It provides information that is helpful in assessing geographical boundaries between water types, as well as in resolving questions that arise when groups or patches of pixels are assigned to a different class than the majority of surrounding pixels. A goodness of fit measure will also signal the presence of pixels not closely identified with any of the defined water type classes, possibly indicating the need to incorporate an additional class into the training set. To measure goodness of fit, a probability landscape is constructed, and the location of each pixel in the probability landscape determines its Classification Goodness of Fit.

[23] For each class, the class centroid $\mathbf{m}^{(i)}$ represents the center of several concentric probability regions; in three-dimensional feature space, these regions may be visualized as probability shells surrounding the centroid. The p% probability shell for class i, $S_p^{(i)}$, is the region surrounding the centroid $\mathbf{m}^{(i)}$ within which the closest p% of all the pixels in the classified image fall. The metric for "closeness" is unique to a particular classifier. For the Euclidean Distance Classifier, the p% probability shell includes pixels

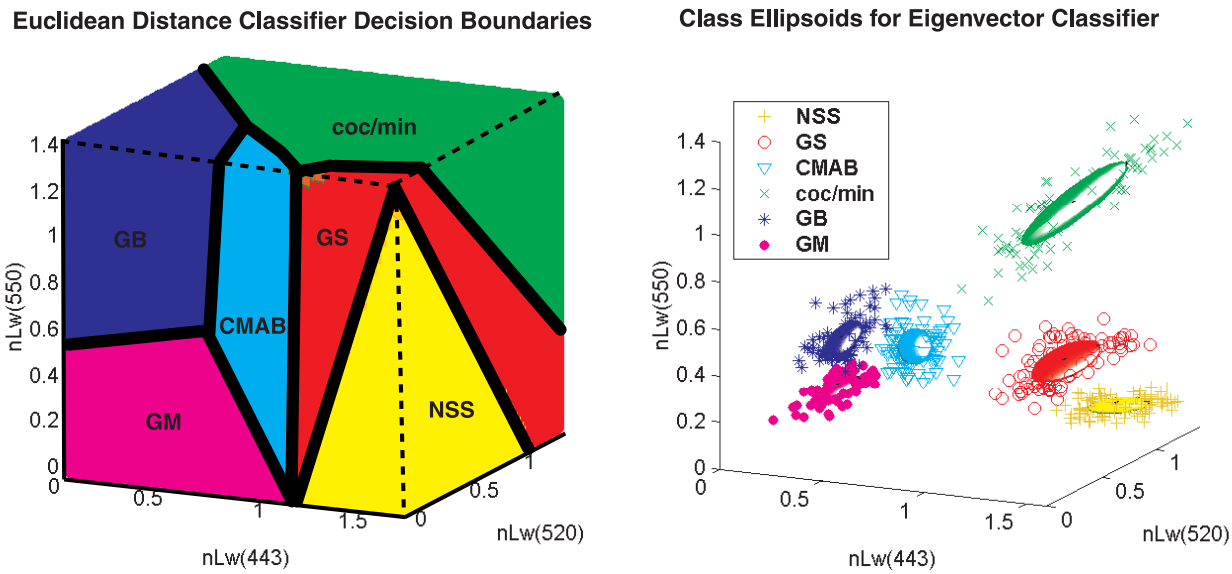

Figure 4. (left) Class decision boundaries for the Euclidean Distance Classifier; decision boundaries are equidistant from the class centroid in all directions. (right) Ellipsoids defining the classes for the Eigenvector Classifier shown with the Northwest Atlantic Training Set (nj = 100 training data points per class); ellipsoid orientation determined by the eigenvector directions for each class. The ellipsoid axis length in each eigenvector direction is equal to the square root of the corresponding eigenvalue, so that the ellipsoids shown encompass 1 standard deviation of the training data for each class. NSS = Northern Sargasso Sea waters; GS = Gulf Stream waters; CMAB = Central Mid-Atlantic Bight waters; coc/min = coccolithophore/mineral dominated waters; GB = Georges Bank waters; GM = Gulf of Maine waters; class mean $\mathbf{m}^{(i)}$ indicated as black dot at centroid of each cluster where visible.

whose Euclidean distance from the class centroid are in the smallest p% of all Euclidean distances for all pixels,

$$\mathbf{s}_v \in \mathrm{S}_p^{(i)} \text{ iff } \mathrm{D}\left(\mathbf{s}_v, \mathbf{m}^{(i)}\right) < \mathbf{d}_{n_p}^{(i)}. \tag{9}$$

Thus, a pixel $\mathbf{s}_v$ is enclosed within the p% probability shell for a given class if the distance between it and the class centroid is less than or equal to the n_pth-smallest distance $\mathbf{d}_{n_p}^{(i)}$. The nv-element vector $\mathbf{d}^{(i)}$ contains the distances, sorted in ascending order, between each of the nv pixels in the image and $\mathbf{m}^{(i)}$; n_p is the number of pixels in the p% probability shell, computed as

$$n_p = (p \times nv)/100 \text{ with } 0 \le p \le 100. \tag{10}$$

[24] The entire probability landscape consists of the combined probability shells for all the classes. The location of each pixel in the probability landscape is computed relative to the position of twenty probability shells (5%, 10%, 15%, . . ., 100%) for each class. A pixel is ultimately assigned a Goodness of Fit value (G) to each class corresponding to the lowest order probability shell p_o for that class which encloses it. The G value represents the $(100 - p_o)$th percentile of fit for that class. For example, for an image consisting of nv = 100,000 pixels, the p = 5% probability shell for a given class is the region of feature space surrounding the class centroid within which the closest n_p = 5000 pixels in the classified image fall, whereas the p = 10% probability shell is the region within which the closest n_p = 10,000 pixels fall. The pixel whose distance to the centroid ranks 23rd-closest is enclosed by both the 5% and 10% probability shells; it is assigned a Goodness of Fit value of G = 95 (corresponding to the lower order shell p_o = 5), indicating that it is in the 95th percentile for fit to that class, whereas a pixel whose distance ranks 6,015 is assigned G = 90. In this fashion, the classifier not only decides the class membership of a pixel, it also assigns a G value which reflects how strongly that pixel belongs to that class. Pixels which are assigned higher G values (e.g., 95 or 90) for a given class are identified very strongly with that class, whereas pixels which are assigned lower G values (e.g., 5 or 10) are enclosed only by the outer probability shells and are very weakly associated with that class.

3. Results

[25] The water types characteristic of these Northwest Atlantic locations were easily distinguishable as distinct, well-delineated clusters in feature space. In addition, these water types projected onto the same regions in feature space for different scenes over time. For example, our examination of CZCS and SeaWiFS images spanning 17 years revealed that randomly selected pixels from the waters over Georges Bank in the October 1997 SeaWiFS image (Figure 2, right) occupied the same region in feature space (not shown) as they did in the July 1980 CZCS image (Figure 2, left). In fact, this was true of randomly selected pixels for most of the other water types. The stationarity of the ocean color data over almost two decades contributes to the robustness of the training set.

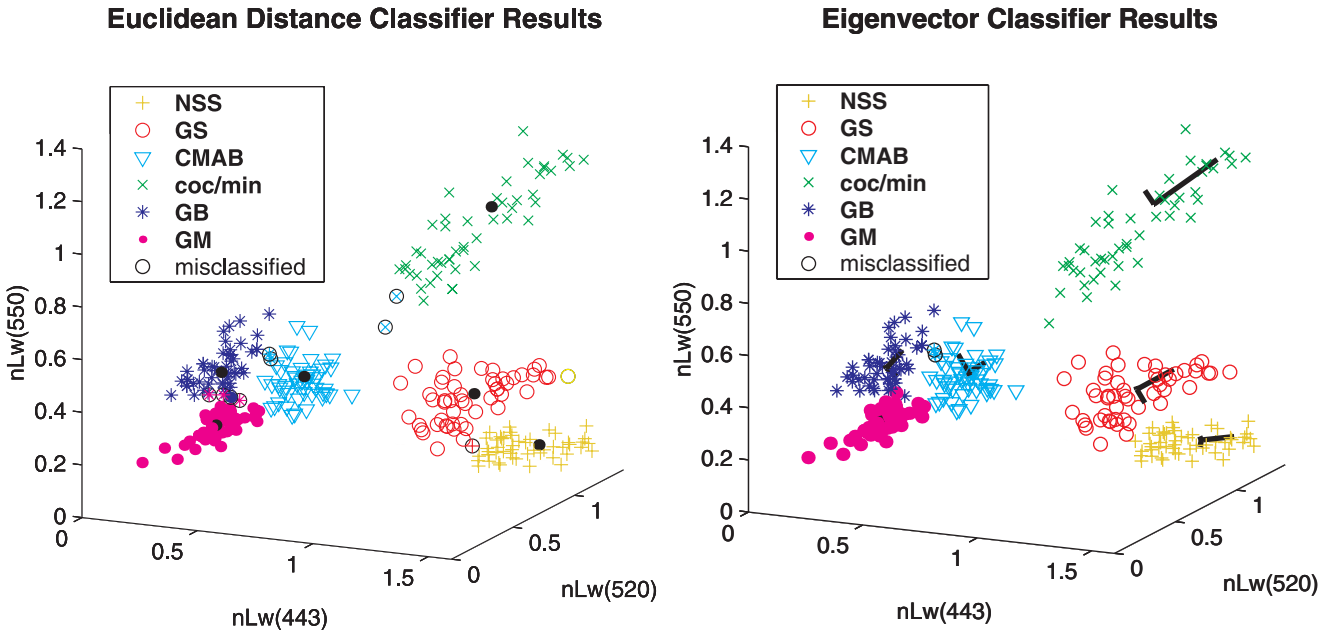

Figure 5. Example results of randomized classification runs of half the Northwest Atlantic Training Set. A test training set was constructed from half the pixels (nj = 50, randomly chosen, 20 trials) of each water type in the Northwest Atlantic Training Set; the remaining pixels (pictured above for one run) were classified with both classifiers. Misclassified pixels shown circled in black; misclassified pixels retain original symbol but take on color of the class to which they were assigned. (left) Euclidean Distance Classification results (average 97.4% correct); class mean $\mathbf{m}^{(i)}$ indicated as black dot at centroid of each cluster. (right) Eigenvector Classification results (average 99.1% correct); eigenvector directions indicated by black lines with line length designating 1 standard deviation of the test training data set in each eigenvector direction. NSS = Northern Sargasso Sea waters; GS = Gulf Stream waters; CMAB = Central Mid-Atlantic Bight waters; coc/min = coccolithophore/mineral dominated waters; GB = Georges Bank waters; GM = Gulf of Maine waters.

[26] Initial evaluation of the performance of each classification technique was accomplished by applying the classifiers to a test data set for which class membership is known. For each classifier, twenty trials were conducted in which a test training set (nj = 50 for each class) was randomly selected from the Northwest Atlantic Training Set, and the remaining 300 pixels were classified. For example, a single trial resulted in 10 misclassified pixels (96.7% correct) with the Euclidean Distance Classifier and only 2 misclassified pixels (99.3% correct) with the Eigenvector Classifier (Figure 5). Over the 20 trials, applying the Euclidean Distance Classifier resulted in an average of 97.4% correctly classified pixels, with a mean of 7.8 misclassified pixels (standard deviation (s.d.) 2.3). Even higher success rates were achieved with the Eigenvector Classifier; its consideration of the three-dimensional shape of each class reduced misclassification rates for the more elongated classes. Application of the Eigenvector Classifier gave an average of 99.1% correctly classified pixels, with a mean of 2.9 misclassified (s.d. 1.6) over 20 trials.

[27] Both the Euclidean Distance and Eigenvector Classifiers were applied to several cloud-free ocean color images of the Northwest Atlantic. Classification results revealed striking patterns of water type distribution throughout the region, as shown in the Euclidean Distance Classification results (Figure 6), and the Eigenvector Classification results (Figure 7) for both 7 July 1980 and 8 October 1997. The water types in each scene are clearly distinguishable, and classifier application reveals that waters of the same optical type form well-defined water masses that remain in the same general geographical regions over time. For example, the GM waters (Figures 6 and 7) occurred largely in the Gulf of Maine, whereas the CMAB waters occurred in the Central Mid-Atlantic Bight to the south of Georges Bank. The GB waters occurring over Georges Bank were quite distinct from the surrounding waters, indicating that these waters were characterized by a unique set of optical properties, and may have been more similar optically to the waters along the New England coast. In the southern half of the region, GS waters occurred to the north of NSS waters on both days, but both were found farther north on 7 July 1980 as compared to 8 October 1997. Mesoscale physical oceanographic features also become apparent after classification, since they often result from interactions between water masses characterized by different optical water types.

[28] The classifiers revealed an array of several patches of coc/min waters on 7 July 1980 between 40°N and 41°N (just south of Georges Bank), extending from approximately −70°(W) eastward to −66°(W). It is suspected that the presence of these optically distinct waters indicate a coccolithophore bloom. The training set for the coc/min class occupied a large region in feature space quite separate from the other optical water types (see Figure 3), and when subjected to the criteria of *Brown and Yoder* [1994b] for

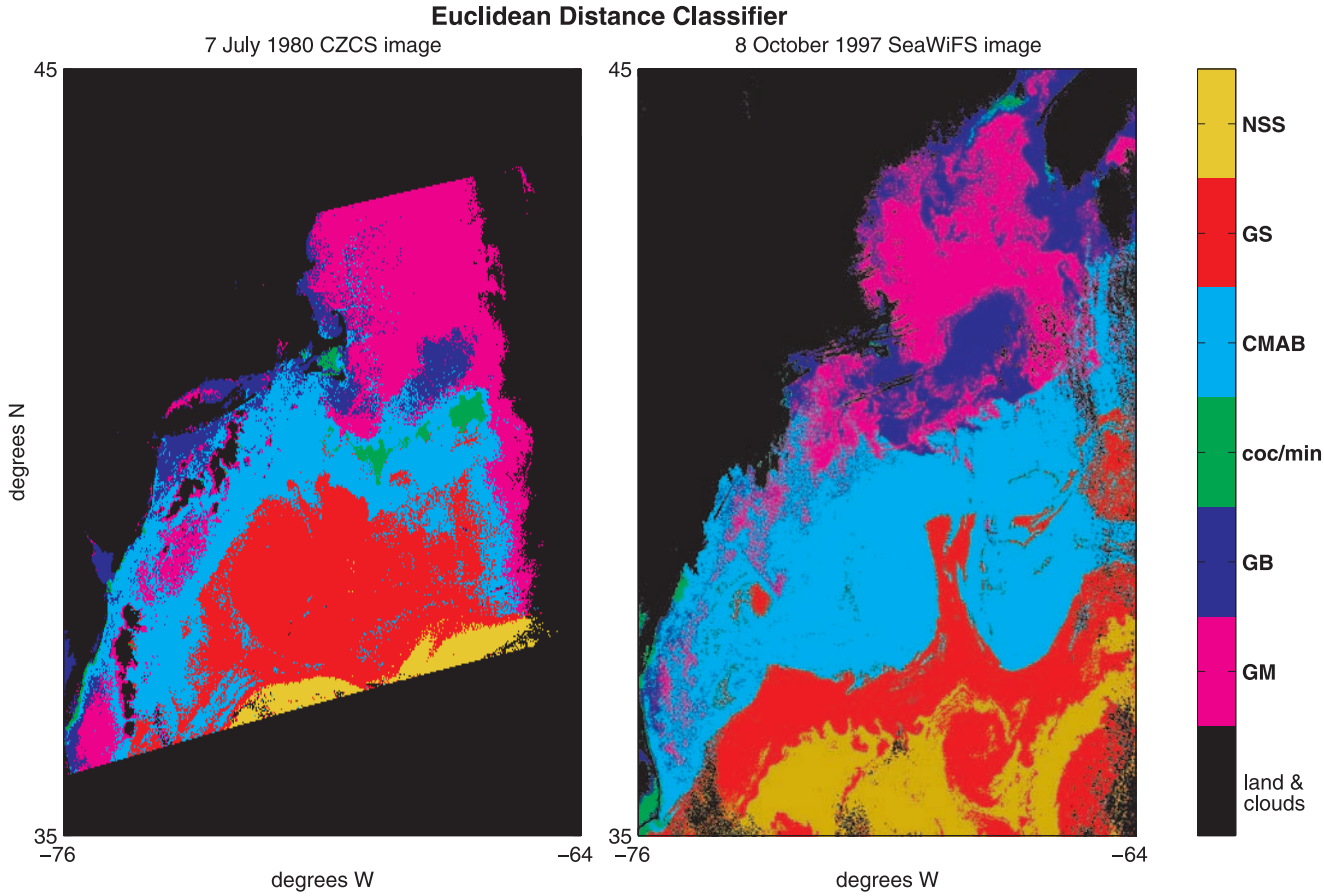

Figure 6. Classification results for the Euclidean Distance Classifier applied to the Northwest Atlantic on two different days: (left) 7 July 1980 (CZCS image) and (right) 8 October 1997 (SeaWiFS image). The water types are clearly distinguishable, and application of the classifier reveals that they form well-defined water masses. There are striking broad-scale similarities in the distribution of these water types between the two scenes, even though they are 17 years apart. Mesoscale physical oceanographic features are apparent; differences may represent seasonal and/or inter-annual variability.

CZCS data, 87% of the pixels were identified as coccolithophore pixels (Table 1). All the pixels (940) within the suspected coccolithophore bloom projected onto the same region in feature space as the coc/min training set, and over 50% met the Brown and Yoder coccolithophore criteria.

[29] The classifiers also assigned several other patches to the coc/min optical water type: the Nantucket Sound/Vineyard Sound/Buzzard's Bay region off Cape Cod, Massachusetts (7 July 1980), the coastal waters off the Eastern Shore (Maryland)/Delmarva Peninsula (both 7 July 1980 and 8 October 1997), a portion of the Maine Coastal Current between Maine and Nova Scotia, Canada (8 October 1997), and the Pamlico Sound/Outer Banks coastal waters off Hatteras Island, North Carolina (8 October 1997). Based on their location in feature space as well as the results of subjecting them to the appropriate coccolithophore criteria (see Table 1), it is evident that the pixels in these other patches do not represent coccolithophore blooms. In fact, the coc/min-classified pixels from the coastal waters of Maine, Massachusetts, Delaware, Maryland, and Virginia (from both 7 July 1980 and 8 October 1997) formed their own well-defined cluster centered above and to the left of the coc/min training set (see Figure 3 for reference). The mineral material in these coastal waters is most likely to be sediment of terrigenous origin, whose optical signature differs from that of the coccolithophore bloom-associated mineral material (e.g., detached liths) represented by the coc/min training set.

[30] In addition to the non-coccolithophore bloom patches classed as coc/min, close examination of the classification results revealed other examples of pixels assigned to a particular optical water type that do not appear to belong to that class. For example, the waters of Long Island Sound, Cape Cod Bay, and coastal Massachusetts were assigned to the GB class, as were the waters of Delaware Bay and the Chesapeake, and patches of GM waters appeared in the Central Mid-Atlantic Bight, far to the south of the Gulf of Maine. In these cases, it is not clear from the classifier results (Figures 6 and 7) whether the patches in question have similar optical properties to the prototype of the class to which they were assigned, or whether they were simply assigned to that class because a more appropriate prototype was not included in the training set. The Classification Goodness of Fit measure provides a means by which to interpret these classification results by assigning a G value to the classification of each pixel, measuring the certainty with which the

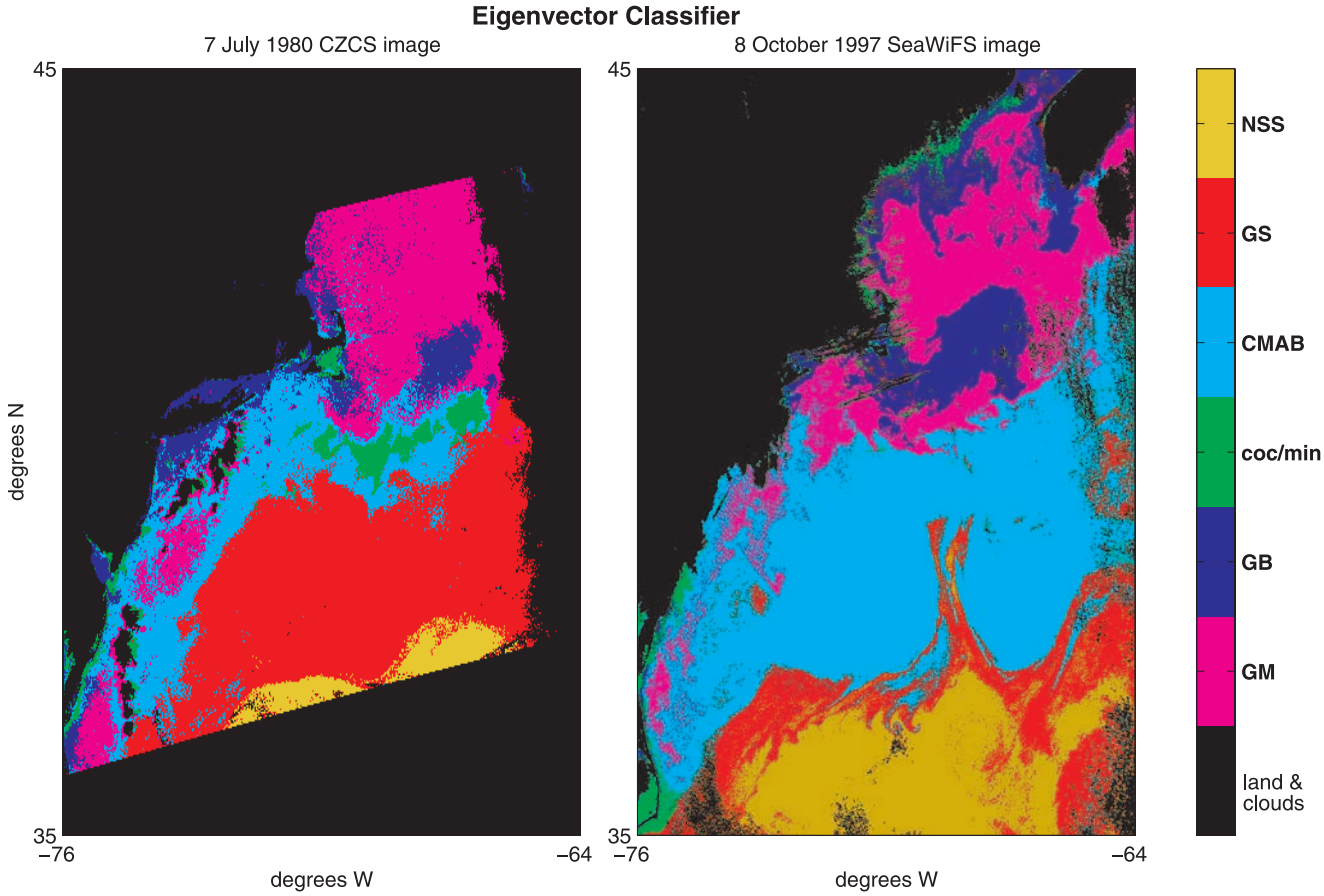

Figure 7. Classification results for the Eigenvector Classifier applied to the Northwest Atlantic on two different days: (left) 7 July 1980 (CZCS image) and (right) 8 October 1997 (SeaWiFS image). Results with this classifier are quite similar to the Euclidean Distance Classifier results shown in Figure 6. The Eigenvector Classifier seems to more correctly identify the pixels near the right cloud edge of the CZCS image as compared to the Euclidean Distance Classifier, but does not delineate the mesoscale eddy revealed by the Euclidean Distance Classifier in the 8 October 1997 image at ~36°N, 68°W (see Figure 6, right).

class assignment is made; pixels with high G values are strongly identified with their class, whereas lower G values reflect lower confidence in the classification result. This measure was particularly useful in evaluating pixels or groups of pixels for which the class assignments appeared to be incorrect, as well as in understanding the nature of the boundaries between water types.

[31] Applying Classification Goodness of Fit measures to the Euclidean Distance Classifier results (Figure 8) revealed that the boundaries between water masses of different optical types were quite distinct. Pixels on either side of these sharp boundaries were generally characterized by very high G values, indicating that they are very strongly associated with their respective water type class. The edges

Table 1. Application of Appropriate Coccolithophore Criteria to Pixels Classified as coc/min by Euclidean Distance Classifier

Region	Date	No. of coc/min Pixels	No. That Meet Criteria	Percent That Meet Criteria
South of Georges Bank MA	7 July 1980[a]	940	476	50.6%[b]
Nantucket Sound/Buzzards Bay MA	7 July 1980[a]	205	1	0.5%
Delmarva coast DE/MD/VA	7 July 1980[a]	187	2	1.1%
Maine Coastal Current ME	8 October 1997[c]	1060	0	0% {0%}
Delmarva coast DE/MD/VA	8 October 1997[c]	2579	174	6.7% {8.0%}
Pamlico Sound/Outer Banks NC	8 October 1997[c]	3134	57	1.8% {0.2%}
Northwest Atlantic Training Set	n/a	100	87	87%

[a]For CZCS data (7 July 1980) the criteria of *Brown and Yoder* [1994b] were applied.

[b]Only the patches comprising the suspected coccolithophore bloom south of Georges Bank on 7 July 1980 contained a significant percentage of pixels which met the criteria.

[c]For SeaWiFS data, pixels were subjected to the following criteria: $B2 \geq F(1)$; $B5 \geq F(2)$; $F(3) \leq B2/B5 \leq F(4)$; $F(5) \leq B4/B5 \leq F(6)$; $F(7) \leq B2/B4 \leq F(8)$; where B2 = nLw(443), B4 = nLw(510), and B5 = nLw(555). Results of applying two different limit vectors (F) are shown: for Chris Brown's SeaWiFS coccolithophore criteria (personal communication, 1999) F = [1.1 0.9 0.85 1.4 1.0 1.4 0.7 1.1]; for the SeaDAS criteria (results given in {}) F = [1.1 0.81 0.6 1.1 0.9 1.32 0.6 0.92].

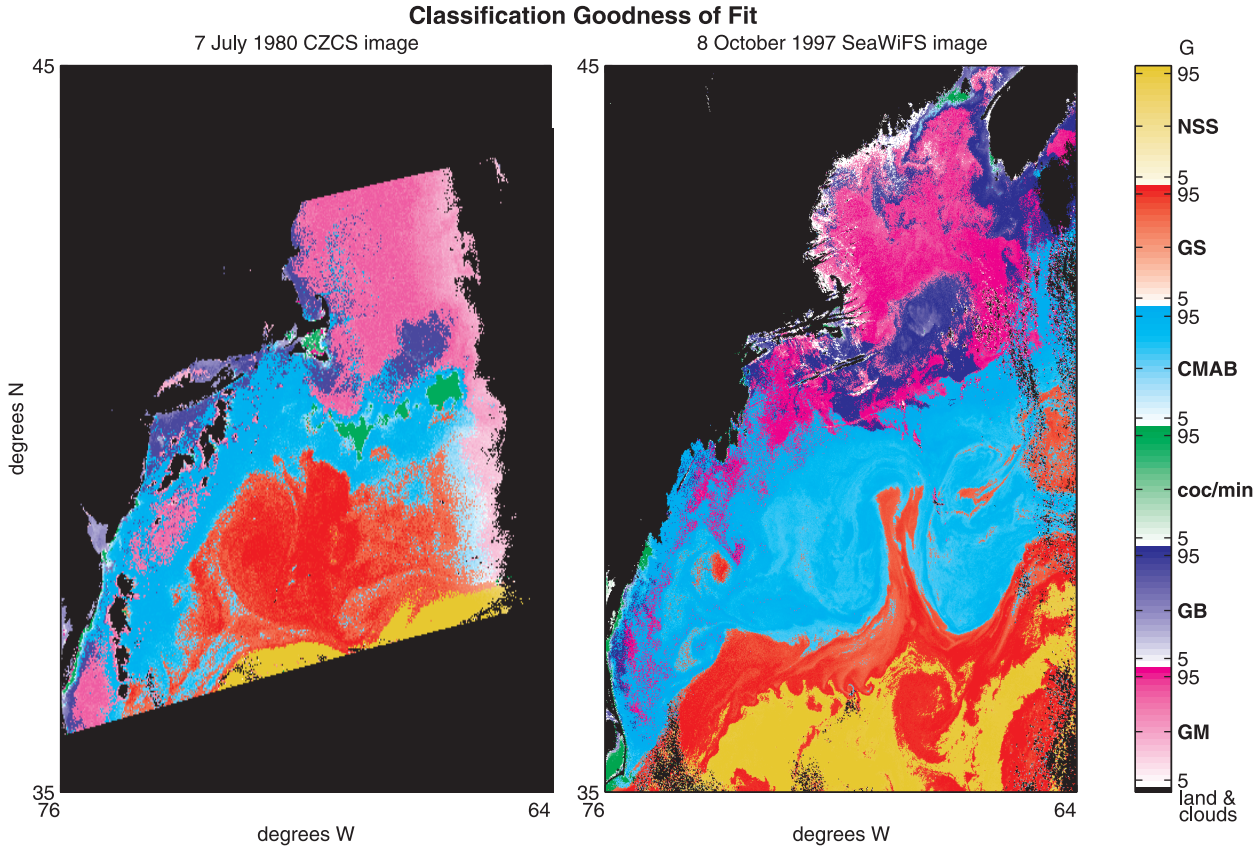

Figure 8. Classification Goodness of Fit for the Euclidean Distance Classifier (see Figure 6 for corresponding classifier results). Intensity of color indicates goodness of fit to class (white corresponds to very poor fit), as shown in the legend at right. A Goodness of Fit value (G) for each pixel is computed by determining where that pixel falls in relation to the concentric probability regions surrounding each class center; a higher G value indicates better fit. For pixels classified as GM, for example, the deepest pink colors (G = 95) indicate that those pixels are in the 95th percentile for fit to that class; that is, they are among the closest 5% of all pixels in the entire scene to the GM class centroid.

of the suspected coccolithophore bloom in the 7 July 1980 image (Figure 8, left) were an exception; it is likely that these low G value pixels surrounding the coccolithophore patches represent the mixing of the coc/min and CMAB water types. The Goodness of Fit measure also shed light on the identity of some groups of pixels for which the classification results appeared spurious. For example, for 7 July 1980, pixels in the longitudinally confined narrow band classed as GM to the extreme right of the CZCS image well south of the Gulf of Maine (see Figure 6, left) are shown to be only very weakly associated (G values ranging between ~10 and 20) with the GM optical type. Similarly, the CMAB-classified pixels in the narrow band extending southward from the suspected coccolithophore bloom, paralleling the eastern edge of this band of GM-classified waters, had G values of 10 or lower. It is likely that cloud-edge effects were confounding the optical signature of these pixels, and that their class assignment was an artifact of this noise in the image. Notably, the Eigenvector Classifier (see Figure 7, left) was better able to identify these pixels in spite of the noise in the data. The Delaware and Chesapeake Bay waters also have G values of 10 or lower, indicating that their optical properties differed from the GB waters found over Georges Bank. On the other hand, the G values for the New England coastal waters (GB), the waters off Nova Scotia (GB), and the patches of water in the Mid-Atlantic Bight (GM) were generally fairly high, indicating that their optical properties were quite similar to the class prototype.

4. Discussion

4.1. Classifiers Reveal Oceanographic Features

[32] The application of these distance-based optical water type classification techniques to satellite ocean color imagery of the Northwest Atlantic has revealed physically, ecologically, and biogeochemically relevant spatial patterns of water type distribution. Features ranging from regions of high phytoplankton biomass to mesoscale eddies to phytoplankton blooms were apparent in the classification results. For example, the waters over Georges Bank appeared as a unique water type. The Georges Bank area is of great interest

ecologically due to a seasonal cycle with periods of high productivity throughout the food chain, including important commercial fishes. Pigment distributions derived from these ocean color data are generally consistent with the spatial patterns and relatively high levels of phytoplankton pigment observed by traditional oceanographic survey methods in the shallow, tidally mixed waters over the Bank [*O'Reilly et al.*, 1987; *O'Reilly and Zetlin*, 1998]. These factors contribute to the unique optical signature of the GB water type.

[33] To the south, comparison with AVHRR sea-surface temperature data reveals that the GS optical water type was associated with the warm (~30°C) Gulf Stream waters as they flowed eastward. These Gulf Stream waters were also quite distinct optically from the slightly higher pigment Mid-Atlantic Bight waters to the north and the Sargasso Sea waters with similar pigment concentration to the south. The classification results for the summertime (7 July 1980) showed the GS waters extending much farther north than they did in the autumn (compare to 8 October 1997 results), which is consistent with the fact that the influence of the Gulf Stream moves northward in the summer months, retreating southward as winter approaches. The GS waters formed a fairly narrow band extending from the southwest to the northeast between 35° and 38°N latitude, and on 8 October 1997, a large meander of GS water extended northward from the path of the current into the colder waters of the Mid-Atlantic Bight. The NSS waters occurred in the slightly cooler Sargasso Sea, and the boundary between the NSS and the GS waters was punctuated with several small meanders and swirls (evident in both Figures 6 and 7, as well as in the AVHRR data), possibly indicating a region of turbulent mixing between optical water types.

[34] Two large, approximately circular patches of GS water were visible in the classification results on 8 October 1997, one far to the north (at ~40°N latitude and −65°(W) longitude) surrounded by CMAB waters, and the other just to the south of the Gulf Stream meander (at ~36°N latitude and −68°(W) longitude) surrounded by NSS waters. These were mesoscale eddies, which appear to be remnants of warm- and cold-core rings, respectively. In fact, cross-comparison of the classifier results with AVHRR imagery for that day demonstrated that the optical signature revealed by the classifiers was coincident with sea surface temperature anomalies. Examination of the sea-surface temperature data revealed that the northern eddy is a warm anomaly, with temperatures inside the eddy elevated ~2–4°C above the surrounding waters. The eddy south of the meander did not exhibit as strong a sea-surface temperature signature, with internal temperatures only ~1°C cooler than the Gulf Stream waters immediately to the north, and at most about 1°C warmer than the surrounding Sargasso waters, possibly indicating that it was in a more advanced stage of decay than the northern ring. This approach of validating the classifier results based on independent remotely sensed data would be further enhanced by the availability of sea-truth measurements; in situ sampling of these optical water masses will allow more complete characterization of their unique properties.

[35] Cold-core rings, those containing a mass of less saline continental slope water, are frequently observed drifting in the Sargasso Sea for many months [*Angel and Fasham*, 1983]. Comparison of coincident AVHRR and CZCS imagery of the Sargasso Sea revealed an inverse relationship between sea-surface temperature and pigment concentration, with cold temperature anomalies (ring centers) associated with higher pigment values [*McGillicuddy et al.*, 2001]. The decay of these rings brings about the deepening of the nutricline and pigment concentrations within the ring become intermediate between those of the slope water and the Sargasso water [*Wiebe*, 1976], until their unique biological and optical properties become indistinguishable from the surrounding Sargasso Sea.

[36] In a similar fashion, warm-core rings are formed when a Gulf Stream meander pinches off to enclose more saline oligotrophic Sargasso Sea water, thereby moving tropical and subtropical species into the temperate slope waters of the Mid-Atlantic Bight [*Craddock et al.*, 1992]. These eddies are generally shorter-lived and smaller in size than are cold-core rings, shrinking in diameter over time, and eventually being reabsorbed into the Gulf Stream [*Richardson*, 1983]. Following the formation of a warm-core ring, a rapid increase in phytoplankton biomass has been observed at the ring center [*Nelson et al.*, 1985], possibly as a consequence of enhanced nitrate flux there, whereas the highest bacterial biomass (reaching estuarine levels) has been observed in the frontal zone near the ring edge [*Ducklow*, 1986]. As these warm-core rings approach the Mid-Atlantic Bight shelf-break front in springtime, shelf and slope water is entrained seaward, and pigment concentrations along the front (which extends for several hundred km) are enhanced twofold over surrounding waters [*Ryan et al.*, 1999]. The influence of these rings has also been noted as far north as Georges Bank, where waters from the southern flank of the Bank were entrained into a warm-core ring in springtime, and exhibited much higher chlorophyll concentrations than the surrounding waters [*Ryan et al.*, 2001].

[37] The classifiers also revealed what appears to be the occurrence of a coccolithophore bloom (coc/min waters south of Georges Bank on 7 July 1980, Figures 6 and 7); blooms of *Emiliania huxleyi* are known to occur episodically in the region [*Balch et al.*, 1991; *Brown and Yoder*, 1994a], and represent a unique optical water type widely separated in feature space from all other water types in the Northwest Atlantic Training Set. Although no samples were available for sea-truthing, in situ measurements are the only means by which to gain further insight into this suspected coccolithophore bloom. For example, in-water sampling would have confirmed the identity of the bloom species, as well as the cell and detached lith concentrations responsible for the optical signature shown by the classifiers.

[38] In other work, this feature-based classification approach has proved valuable for the detection and identification of phytoplankton blooms around the globe. *Martin Traykovski and Sosik* [1998] (also, Optical classification of phytoplankton-dominated water types based on remotely sensed ocean colour, manuscript in preparation, 2003) have performed a cluster analysis of globally occurring phytoplankton-dominated waters differing not only in the phytoplankton taxa present, but also in the abundance of phytoplankton cells present. Data sets were compiled based on phytoplankton observations documented in the literature. Some represented bloom conditions of an identified phytoplankton taxon, e.g., *Trichodesmium* blooms in the Gulf of Thailand and off Northwest Australia [*Subramaniam and Carpenter*, 1994], and a bloom of the coccoli-

Table 2. Results of Randomized Trials in Which the Classifiers Were Applied to a Test Data Set With Known Class Membership

	Percent Correctly Classified (Mean of 20 Trials)[a]						
Classifier	NSS	GS	CMAB	coc/min	GB	GM	Overall
Euclidean Distance[b]	98.8%	99.3%	99.9%	97.7%	90.4%	98.5%	97.4%
Eigenvector[c]	99.8%	100%	96.8%	100%	98.8%	99.1%	99.1%

[a]Percent correctly classified (averaged over 20 trials) is shown; $nj = 1000$ (20 trials of 50 pixels each) for each water type class.

[b]The Euclidean Distance Classifier performed better with pixels belonging to the more spherical CMAB class.

[c]The Eigenvector Classifier performed better with pixels which either belonged to the more elongated classes (e.g., coc/min), or were adjacent to them (e.g., GB pixels adjacent to GM).

thophore *Emiliania huxleyi* over the northwest European Shelf [*Holligan et al.*, 1983]. Conditions with sub-bloom concentrations and mixed-taxa phytoplankton communities were also included in the analysis, e.g., a 25-km-long patch in the Southern California Bight containing the coccolithophore *Umbilicosphaera sibogae* (representing at most 60–75% of the phytoplankton biomass [*Balch et al.*, 1989]), present with other species including diatoms [*Eppley et al.*, 1984]. For each of the phytoplankton data sets, normalized water-leaving radiances in each spectral band were extracted from the corresponding ocean color data for selected pixels in the regions of interest. Using a simple clustering approach in a three-dimensional feature space, it was possible to distinguish between oceanic regions containing distinct blooms of different major taxa (i.e., *Trichodesmium*, coccolithophores, dinoflagellates) and also to separate coccolithophore waters into two optically different types (those with and without detached liths).

4.2. Classifier Performance

[39] When applied to a test data set with known class membership, both the Euclidean Distance Classifier and the Eigenvector Classifier had very high success rates, correctly classifying over 97% and 99% of pixels respectively. Examining the results of twenty randomized trials with the test data set (Table 2) revealed the strengths and weaknesses of each classification approach. The Euclidean Distance Classifier is well suited to spherical classes, i.e., classes whose boundaries in all directions are equidistant from the class centroid. In fact, this classifier performed best with pixels from the spherical CMAB water type class (see Figure 4, right, for class shape), misclassifying only 1 out of 1000 (20 trials of 50 pixels each). Performance was poorer with the coc/min and GB classes, with pixels being assigned to adjacent classes more often.

[40] For more ellipsoidal classes (e.g., coc/min, for which the region of feature space encompassed by the class was much more elongated in one dimension relative to the others) or for pixels adjacent to these more elongated classes (e.g., GB pixels adjacent to the elongated GM cluster), classification success can be improved significantly when the shape of the classes is taken into account. As a result, the Eigenvector Classifier was able to correctly identify the GB pixels much more often. This classifier consistently demonstrated a higher percentage of correct classifications than did the Euclidean Distance Classifier for the more ellipsoidal water type classes (e.g., NSS, GS, coc/min); only 2 pixels out of 3000 were misclassified for these three classes combined. The Eigenvector Classifier did not perform as well as the Euclidean Distance Classifier with pixels from more spherical classes in the proximity of ellipsoidal classes; for example, the highest misclassification rate for the Eigenvector Classifier occurred with CMAB pixels (see Table 2, column 3), which it assigned to the adjacent (elongated) GS, coc/min, and GM water type classes. Overall, each classifier appeared to perform best where the other was weakest, so that the two approaches complement each other. Notwithstanding these differences in performance, the resulting patterns of water type distribution were very similar when the two approaches were applied to ocean color imagery of the Northwest Atlantic (compare Figures 6 and 7). As a result, the two classifiers provided a consistent picture of the geographic extent of each optical water type in this region.

[41] In constructing a training set, it is important that the chosen class prototypes encompass the variability in each class while remaining separable in feature space. To ensure that the training set for the Northwest Atlantic region was robust, several different options were considered, including a training set developed strictly from SeaWiFS data, as well as a combined training set that included points from both CZCS and SeaWiFS data. Evaluation of these alternate training sets revealed that the Northwest Atlantic Training Set resulted in the best classifier performance for both classifiers, indicating that class separability was best maintained with this training set. In addition, the classification results for the 8 October 1997 SeaWiFS image (see Figures 6 and 7) using the chosen training set were very similar to those using the training set constructed strictly from SeaWiFS data, further establishing that the class prototypes in the Northwest Atlantic Training Set encompassed the variability in both the CZCS and SeaWiFS data.

[42] Once a relatively representative training set for a particular region is established, applying different classification approaches using this training set can help determine whether the training set is truly comprehensive or whether additional water type classes are necessary to fully capture the range of optical variability in the region of interest. Indications that the training set may not be complete include the appearance of scattered patches of a particular water type which are geographically removed from the region(s) dominated by that water type, as well as pixels or groups of pixels with very low G values, indicating very poor goodness of fit to class. For example, pixels from the 7 July 1980 and the 8 October 1997 data assigned to the coc/min optical water type included waters from a suspected coccolithophore bloom, as well as various patches of non-coccolithophore, likely sediment-dominated, coastal water. The pixels in these particular coastal areas projected into their own distinct cluster in feature space, adjacent to the coc/min class but far away from the other classes, and likely represent a unique water type. As a result, it would be possible to distinguish between at least two types of mineral-dominated optical water types; dividing the coc/min class into two subclasses and developing a training set for the sediment-dominated coastal waters would allow the classifiers to differentiate those waters from coccolithophore-dominated waters. Similarly, the waters of the Delaware and Chesapeake Bays were assigned to the GB class with very low G values; their optical properties clearly

differ from the GB class prototype, and would be better represented by a more appropriate optical water type. In this manner, the classification results provide specific information to guide improvements by identifying additional water types in the region of interest that were not well represented in the training set. Inclusion of these water types is likely to improve the classification results.

4.3. Pigment Content and Optical Water Type

[43] Previously, studies aimed at characterizing the spatial and temporal optical variability of mesoscale features based on ocean color imagery have relied to a large extent on pigment distributions [e.g., *Peláez and McGowan*, 1986; *Abbott and Zion*, 1987; *Yoder et al.*, 1987; *McClain et al.*, 1990; *Denman and Abbott*, 1988; *Thomas et al.*, 1994]. One of the advantages of the Euclidean Distance and Eigenvector Classifiers is that these feature-based approaches are able to exploit much more of the information inherent in ocean color data than is contained in the satellite-derived pigment image; the optical water type distribution patterns revealed by the classifiers were not necessarily apparent in the pigment images (compare Figures 6 and 7 with lower right-hand quadrant in each panel of Figure 2). Comparison of mean satellite-derived pigment concentrations (Figure 9) for the six water types in the Northwest Atlantic Training Set showed that the classifiers not only distinguish between waters with different pigment concentrations, they were also able to discriminate between different optical water types which did not differ in their pigment content. The oligotrophic NSS waters were characterized by very low pigment concentrations; this water type class was quite distinct from the slightly higher pigment GS waters, as well as the productive, high pigment GB waters, once projected in feature space (see Figure 3). An interpretation based on pigment concentration alone, however, will fail to distinguish the different optical water types that do not exhibit differences in pigment content. For example, although the coc/min and GM water types were indistinguishable based on their mean pigment concentrations (Figure 9), these optically distinct water types formed widely separated, well-delineated clusters in feature space (see Figure 3). In general, the distribution of clusters in feature space was not determined solely by their relative pigment content, and cluster proximity was not necessarily indicative of similar pigment concentrations (e.g., GB and GM). In fact, pixels of a particular water type tended to project in the same general region of feature space over time, even if pigment concentration changed significantly. As a result, temporal changes in water type distribution may be followed robustly with these classifiers, since water types are distinguished by their optical signatures which include much more than pigment information.

4.4. Studying Optical Variability

[44] This water type classification approach is ideally suited to the study of optical variability over large oceanic regions. The classifiers were designed to exploit the wealth of information in satellite ocean color imagery and synthesize it into a single map which reveals the distribution of optically distinct water types over a region. In this manner, the classification results provide a valuable composite of remotely sensed ocean color data which greatly facilitates the analysis of spatial variability in optical properties.

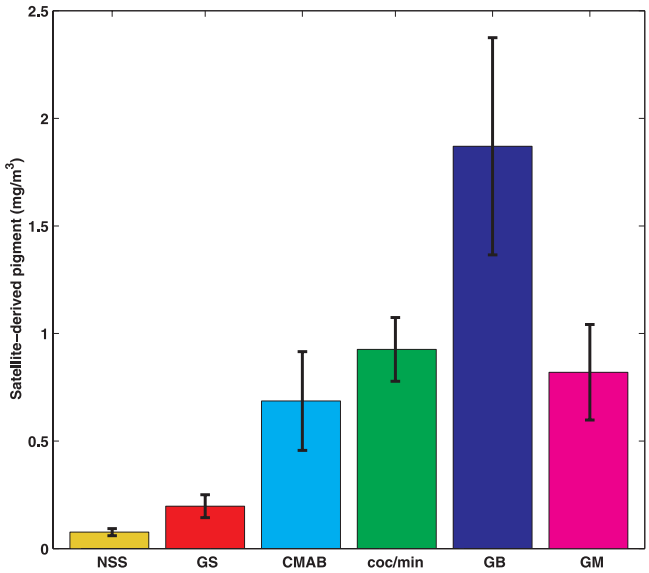

Figure 9. Mean satellite-derived chlorophyll-like pigment concentration (mg/m^3) for each water type in the Northwest Atlantic Training Set (see Figure 3, *nj* = 100 for each class). The classifiers are able to distinguish between different optical water types (e.g., CMAB, coc/min, GM) despite the fact that they have the same pigment concentration. NSS = Northern Sargasso Sea waters; GS = Gulf Stream waters; CMAB = Central Mid-Atlantic Bight waters; coc/min = coccolithophore/mineral dominated waters; GB = Georges Bank waters; GM = Gulf of Maine waters; black bar delineates 1 standard deviation.

Classifier-produced maps of the spatial distribution of optical water types in a region can be compared to bathymetry data, in situ measurements of physical and biological properties, and other remotely sensed data (e.g., sea-surface temperature, sea-surface height), yielding important insights into the sources of the observed optical variability.

[45] Assessment of temporal optical variability may be accomplished by tracking the occurrence of optical water types through time; applying the classifiers to a time series of images will allow changes in water type distributions to be followed easily. The classification results for 7 July 1980 and 8 October 1997 (17 years later) reveal that waters of the same optical type form well-defined water masses that remain in the same general geographical regions over time, demonstrating the utility of employing the classifiers to characterize temporal optical variability in the region. To date, efforts to apply the classifiers broadly to many images of the Northwest Atlantic in order to identify and track water types over time have been set back by SeaWiFS data processing difficulties. In particular, problems with the atmospheric correction algorithms implemented in SeaWiFS Data Analysis System (SeaDAS), have resulted in incorrect retrievals of nLw over the Northwest Atlantic region. Factors such as absorbing aerosols, turbid water (which leads to non-zero water-leaving radiance in the near-infrared (NIR)), and high solar zenith angle (particularly in the winter months), make this region particularly vulnerable to atmospheric correction algorithm failure. For example, ocean color data processing algorithms included in SeaDAS

3.3 resulted in widespread underestimates of nLw in the shorter wavelength bands (with a high frequency of negative nLw retrievals at 412 nm, 443 nm and in some cases, also at 490 nm), causing distortion of the spectral shape throughout the region. The atmospheric correction techniques were improved in May 2000 (Reprocessing 3; SeaDAS 4.0), with the inclusion of the Siegel NIR algorithm, which was designed to reduce the incidence of negative nLw in the violet and blue [*Siegel et al.*, 2000]. This approach resulted in substantial improvements in some areas (e.g., Chesapeake Bay). Unfortunately, SeaDAS 4.0 processing results still show markedly distorted spectra compared to in situ nLw measurements in the Gulf of Maine, with underestimates of nLw at 412 nm, 443 nm, and 490 nm, and a fairly high incidence of negative nLw at 412 nm.

[46] Since the Euclidean Distance and Eigenvector Classifiers rely on the unique spectral signature of each optical water type, these techniques are sensitive to atmospheric correction problems which distort spectral shape. The 8 October 1997 SeaWiFS image was one for which the effect of known artifacts introduced by the atmospheric correction problems was minimal (as determined by in situ match-ups of the nLw spectra). Further improvements to the SeaWiFS atmospheric correction algorithms are planned for the upcoming Reprocessing 4, and preliminary testing indicates that some of these modifications have the potential to improve the situation in the Northwest Atlantic. Once the atmospheric correction issues are resolved and the accuracy of the nLw retrievals for this region improves, these feature-based classification techniques will prove to be a powerful tool, facilitating time series studies of optical variability in the Northwest Atlantic based on a sequence of classifier-produced maps of optical water type distributions.

4.5. Future Directions

[47] Further development of these feature-based optical water type classification techniques will include exploring the utility of an automated method for the definition of water type classes. One approach would be to implement an adaptive scheme which searches through feature space for local minima in pixel density. In this scheme, isolines of low pixel density would represent class boundaries, and the location of class centroids could be mathematically defined based on these boundaries, facilitating the development and application of statistical decision rules for classification. With this method, the shape of each class would not be constrained to being spherical (as with the Euclidean Distance Classifier) or ellipsoidal (as with the Eigenvector Classifier); class shape could be defined uniquely for each class based on the location of the class boundaries. Alternatively, class centroids could be defined by implementing a clustering algorithm, which would provide an objective measure for optimum clustering; with this approach, results are quite sensitive to the number of classes chosen. In general, clustering techniques perform best when some a priori information regarding the number of classes is available.

[48] One of our top priorities as we continue this work is the collection of in situ optical data that can strengthen the classifier results and increase our understanding of the sources of optical variability revealed by the classifiers. The acquisition of sea-truth measurements across the region will contribute to validation of the optical water types identified by the classifiers. We have collected much in-water optical data over the last few years on several cruises to the Gulf of Maine. Although preliminary analysis of these data has provided some insights into the particular optical properties of the GM water type, we have not yet been able to collect comparable data for any of the other water types. We do have the opportunity to participate in two research cruises in the Northwest Atlantic region in the near future, during which we plan to sample a wide range of optical water types including any episodic features we encounter, making detailed measurements of the in-water optical properties across the region. We anticipate that this in situ data set will provide valuable insights into the sources of optical variability across the region, as well as contribute to further advancement of our classifiers.

[49] A logical extension of our feature-based classification work includes the development of a hybrid classification approach, which will involve integrating model-based inversion techniques with the feature-based classifiers, and incorporating sea-truth data into the classification scheme. To implement this hybrid approach, an inverse model (e.g., the semi-analytic radiance model of *Roesler and Perry* [1995] or the nearly backscattering independent model of *Zaneveld et al.* [1998]/*Barnard et al.* [1999]) could be employed to extract inherent optical properties (IOPs) for each pixel based on remotely sensed water-leaving radiances for the region. Development of a hybrid classification technique for Northwest Atlantic optical water types would be carried out in the context of in situ data collected during the ecological and hydrographic work as part of our research examining optical variability in this region. A priori information obtained from sampling the in-water optical properties at particular stations throughout the region could be used to constrain the inverse. The advantage of such a hybrid technique is that it can capitalize on the predictive power of existing semi-analytic models, while taking advantage of the intrinsic features in the data, which are independent of assumptions inherent in the models.

5. Conclusions

[50] Remote sensing of ocean color has significantly expanded our ability to study spatial and temporal variability in phytoplankton abundance and distribution; however, full exploitation of ocean color imagery requires both developments in modeling of upper ocean optical properties (and their relationships with biological, physical, and chemical properties) and more sophisticated data analysis techniques. Optical water type classification approaches based on remotely sensed water leaving radiance have great potential to contribute to the study of spatial and temporal dynamics of ecologically and biogeochemically important properties in the upper ocean. With the development of the feature-based Euclidean Distance and Eigenvector Classifiers, we have automated the water type classification process by applying statistical decision criteria to define class boundaries and assign pixels to a particular class. We have demonstrated that the application of these feature-based classification techniques to ocean color data facilitates discrimination between Northwest Atlantic optical water types, including those waters occurring within a spatially restricted region, where the interaction of tidal flow with

complex bottom topography can result in the formation of fronts between different water types. Application of these classification techniques will contribute to the interpretation of the underlying properties that define optical water types, facilitating region-wide examination of spatial variability in water types using satellite ocean color imagery. The classifiers also show promise as a valuable tool for analyzing patterns of temporal variability in water type distributions if applied to long-term time series studies tracking optical water types through seasonal and interannual changes.

[51] **Acknowledgments.** We are indebted to several generous people for making this work possible. John Ryan processed the 1980 CZCS image for the Northwest Atlantic. Thanks also to Jim Acker and the Goddard DAAC, Karen Baith and the SeaDAS development group, as well as the HNSG HRPT ground station. Many thanks to the members of our lab, Anne Canaday, Rebecca Green, and Ru Morrison, and summer student fellow Ashwini Deshpande. We appreciate the insight and perspective of Trevor Platt and two other anonymous reviewers, whose helpful suggestions resulted in significant improvements to the manuscript. This work was supported in part by grants from the NASA Ocean Biology/Biogeochemistry Program, grant NAG5-7445 (LVMT) and the ONR Environmental Optics Program, grant N00014-97-1-0646 (HMS). This is Woods Hole Oceanographic Institution contribution 10530.

References

Abbott, M. R., and P. M. Zion, Satellite observations of phytoplankton variability during an upwelling event, *Cont. Shelf Res.*, *4*, 661–680, 1985.

Abbott, M. R., and P. M. Zion, Spatial and temporal variability of phytoplankton pigment off northern California during the Coastal Ocean Dynamics Experiment, 1, *J. Geophys. Res.*, *92*, 1745–1755, 1987.

Anderson, D. M., Bloom dynamics of toxic *Alexandrium* species in the northeastern U.S., *Limnol. Oceanogr.*, *42*, 1009–1022, 1997.

Angel, M. V., and M. J. R. Fasham, Eddies and Biological Processes, in *Eddies in Marine Science*, edited by A. R. Robinson, pp. 492–524, Springer-Verlag, New York, 1983.

Backus, R. H., (Ed.), *Georges Bank*, 593 pp., MIT Press, Cambridge, Mass., 1987.

Balch, W. M., R. W. Eppley, M. R. Abbot, and F. M. H. Reid, Bias in satellite-derived pigment measurements due to coccolithophores and dinoflagellates, *J. Plankton Res.*, *11*, 575–581, 1989.

Balch, W. M., P. M. Holligan, S. G. Ackleson, and K. J. Voss, Biological and optical properties of mesoscale coccolithophore blooms in the Gulf of Maine, *Limnol. Oceanogr.*, *36*, 629–643, 1991.

Barnard, A. H., J. R. Zaneveld, and W. S. Pegau, In situ determination of the remotely sensed reflectance and the absorption coefficient: Closure and inversion, *Appl. Opt.*, *38*, 5108–5117, 1999.

Brown, C. W., and J. A. Yoder, Distribution pattern of coccolithophorid blooms in the western North Atlantic Ocean, *Cont. Shelf Res.*, *14*, 175–197, 1994a.

Brown, C. W., and J. A. Yoder, Coccolithophorid blooms in the global ocean, *J. Geophys. Res.*, *99*, 7467–7482, 1994b.

Carder, K. L., R. G. Steward, G. R. Harvey, and P. B. Ortner, Marine humic and fulvic acids: Their effects on remote sensing of ocean chlorophyll, *Limnol. Oceanogr.*, *34*, 68–81, 1989.

Clark, D. K., Phytoplankton pigment algorithms for the Nimbus-7 CZCS, in *Oceanography from Space*, edited by J. F. R. Gower, pp. 227–237, Plenum, New York, 1981.

Craddock, J. E., R. H. Backus, and M. A. Daher, Vertical distribution and species composition of midwater fishes in warm-core Gulf Stream meander/ring 82-H, *Deep Sea Res.*, *39*(Suppl. 1), S203–S218, 1992.

Denman, K. L., and M. R. Abbott, Time evolution of surface chlorophyll patterns from cross-spectrum analysis of satellite color images, *J. Geophys. Res.*, *93*, 6789–6798, 1988.

Ducklow, H. W., Bacterial biomass in warm-core Gulf Stream ring 82-B: Mesoscale distributions, temporal changes, and production, *Deep Sea Res.*, *33*, 1789–1812, 1986.

Eppley, R. W., F. M. H. Reid, and E. Stewart, Length of phytoplankton species patches on the Southern California Shelf, *Cont. Shelf Res.*, *3*, 259–266, 1984.

Eslinger, D. L., and R. L. Iverson, Wind effects on Coastal Zone Color Scanner chlorophyll patterns in the U.S. Mid-Atlantic Bight during spring 1979, *J. Geophys. Res.*, *91*, 12,985–12,992, 1986.

Fox, C. G., R. P. Dziak, H. Matsumoto, and A. E. Schreiner, Potential for monitoring low-level seismicity on the Juan de Fuca Ridge using military hydrophone arrays, *Mar. Tech. Soc. J.*, *27*, 22–30, 1994.

Gordon, H. R., D. K. Clark, J. W. Brown, O. R. Brown, R. H. Evans, and W. W. Broenkow, Phytoplankton pigment concentrations in the Middle Atlantic Bight: Comparison of ship determinations and CZCS estimates, *Appl. Opt.*, *22*, 20–36, 1983.

Gordon, H. R., O. R. Brown, R. H. Evans, J. W. Brown, R. C. Smith, K. S. Baker, and D. K. Clark, A semi-analytic radiance model of ocean color, *J. Geophys. Res.*, *93*, 10,909–10,924, 1988.

Hochman, H. T., F. E. Müller-Karger, and J. J. Walsh, Interpretation of the coastal zone color scanner signature of the Orinoco River plume, *J. Geophys. Res.*, *99*, 7443–7455, 1994.

Holliday, D. V., R. E. Pieper, and G. S. Kleppel, Determination of zooplankton size and distribution with multifrequency acoustic technology, *J. Cons. Cons. Int. Explor. Mer*, *46*, 52–61, 1989.

Holligan, P. M., M. Voillier, D. S. Harbour, P. Camus, and M. Champagne-Philippe, Satellite and ship studies of coccolithophore production along a continental shelf edge, *Nature*, *304*, 339–342, 1983.

Martin, L. V., T. K. Stanton, P. H. Wiebe, and J. F. Lynch, Acoustic classification of zooplankton, *ICES J. Mar. Sci.*, *53*, 217–224, 1996.

Martin Traykovski, L. V., and H. M. Sosik, Optical classification of water types based on remotely-sensed ocean colour, in *Ocean Optics XIV Conference Papers: Kailua-Kona*, *HI*, *10–13 November 1998*, vol. 2, Off. of Naval Res./NASA, Greenbelt, Md., 1998.

Martin Traykovski, L. V., T. K. Stanton, P. H. Wiebe, and J. F. Lynch, Model based Covariance Mean Variance Classification (CMVC) techniques: Algorithm development and application to the acoustic classification of zooplankton, *IEEE J. Ocean. Eng.*, *23*, 344–364, 1998a.

Martin Traykovski, L. V., R. L. O'Driscoll, and D. E. McGehee, Effect of broadband acoustic scattering of Antarctic krill Euphausia superba: Implications for inverting zooplankton spectral signatures for angle of orientation, *J. Acoust. Soc. Am.*, *104*, 2121–2135, 1998b.

McClain, C. R., Review of major CZCS applications: U.S. case studies, in *Ocean Colour: Theory and Applications in a Decade of CZCS Experience*, edited by V. Barale and P. M. Schlittenhardt, pp. 167–188, Kluwer Acad., Norwell, Mass., 1993.

McClain, C. R., L. J. Pietrafesa, and J. A. Yoder, Observations of Gulf Stream-induced and wind-driven upwelling in the Georgia Bight using ocean color and infrared imagery, *J. Geophys. Res.*, *89*, 3705–3723, 1984.

McClain, C. R., J. Ishizaka, and E. E. Hofmann, Estimation of the processes controlling variability in phytoplankton pigment distributions on the southeastern U.S. continental shelf, *J. Geophys. Res.*, *95*, 20,213–20,235, 1990.

McGillicuddy, D. J., V. K. Kosnyrev, J. P. Ryan, and J. A. Yoder, Covariation of mesoscale ocean color and sea-surface temperature patterns in the Sargasso Sea, *Deep Sea Res., Part II*, *48*, 1823–1836, 2001.

Mitchell, B. G., and O. Holm-Hansen, Bio-optical properties of Antarctic Peninsula waters: Differentiation from temperate ocean models, *Deep Sea Res.*, *38*, 1009–1028, 1991.

Morel, A., and L. Prieur, Analysis of variations in ocean colour, *Limnol. Oceanogr.*, *22*, 709–722, 1977.

Müller-Karger, F. E., C. R. McClain, T. R. Fisher, W. E. Esias, and R. Varela, Pigment distribution in the Caribbean Sea: Observations from space, *Prog. Oceanogr.*, *23*, 23–64, 1989.

Munk, W., and C. Wunsch, Ocean acoustic tomography: A scheme for large scale monitoring, *Deep Sea Res.*, *26*, 123–161, 1979.

Nelson, D. M., et al., Distribution and composition of biogenic particulate matter in a Gulf Stream Warm Core ring, *Deep Sea Res.*, *32*, 1347–1369, 1985.

Nelson, N. B., D. A. Siegel, and A. F. Michaels, Seasonal dynamics of colored dissolved material in the Sargasso Sea, *Deep Sea Res., Part I*, *45*, 931–957, 1998.

O'Reilly, J. E., and C. Zetlin, Seasonal, horizontal, and vertical distribution of phytoplankton chlorophyll a in the Northeast U.S. continental shelf ecosystem, *NOAA Tech. Rep., NMFS, 139*, 120 pp., 1998.

O'Reilly, J. E., C. Evans-Zetlin, and D. A. Busch, Primary production, in *Georges Bank*, edited by R. H. Backus, pp. 220–233, MIT Press, Cambridge, Mass., 1987.

O'Reilly, J. E., S. Maritorena, B. G. Mitchell, D. A. Siegel, K. L. Carder, S. A. Garver, M. Kahru, and C. McClain, Ocean color chlorophyll algorithms for SeaWiFS, *J. Geophys. Res.*, *103*, 24,937–24,953, 1998.

Papoulis, A., *Probability, Random Variables, and Stochastic Processes*, third ed., 666 pp., McGraw-Hill, New York, 1991.

Peláez, J., and J. A. McGowan, Phytoplankton pigment patterns in the California Current as determined by satellite, *Limnol. Oceanogr.*, *31*, 927–950, 1986.

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

Platt, T., C. M. Caverhill, and S. Sathyendranath, Basin-scale estimates of oceanic primary production by remote sensing: The North Atlantic, *J. Geophys. Res.*, *96*, 15,147–15,159, 1991.

Richardson, P. L., Gulf Stream rings, in Eddies in Marine Science, edited by A. R. Robinson, pp. 19–45, Springer-Verlag, New York, 1983.

Roesler, C. S., and M. J. Perry, In situ phytoplankton absorption, fluorescence emission, and particulate backscattering spectra determined from reflectance, *J. Geophys. Res.*, *100*, 13,279–13,294, 1995.

Ryan, J. P., J. F. Yoder, and P. C. Cornillon, Enhanced chlorophyll at the shelfbreak of the Mid-Atlantic Bight and Georges Bank during the spring transition, *Limnol. Oceanogr.*, *44*, 1–11, 1999.

Ryan, J. P., J. F. Yoder, and D. W. Townsend, Influence of a Gulf Stream warm-core ring on water mass and chlorophyll distributions along the southern flank of Georges Bank, *Deep Sea Res.*, *48*, 159–178, 2001.

Santamarià-del-Angel, E., S. Alvarez-Borrego, and F. E. Müller-Karger, Gulf of California biogeographic regions based on Coastal Zone Color Scanner imagery, *J. Geophys. Res.*, *99*, 7411–7421, 1994.

Sathyendranath, S., Remote sensing of phytoplankton: A review with special reference to picoplankton, in Photosynthetic Picoplankton, edited by T. Platt and W. K. W. Li, *Can. Bull. Fish. Aquat. Sci.*, *214*, 561–583, 1986.

Sathyendranath, S., T. Platt, E. P. W. Horne, W. G. Harrison, O. Ulloa, R. Outerbridge, and N. Hoepfner, Estimation of new production in the ocean by compound remote sensing, *Nature*, *353*, 129–133, 1991.

Siegel, D. A., and A. F. Michaels, Non-chlorophyll light attenuation on the open ocean: Implications for biogeochemistry and remote sensing, *Deep Sea Res., Part II*, *43*, 321–345, 1996.

Siegel, D. A., M. Wang, S. Maritorena, and W. Robinson, Atmospheric correction of satellite ocean color imagery: The black pixel assumption, *Appl. Opt.*, *39*, 3582–3591, 2000.

Smith, R. C., X. Zhang, and J. Michaelsen, Variability of pigment biomass in the California Current System as determined by satellite imagery, 1, Spatial variability, *J. Geophys. Res.*, *93*, 10,863–10,882, 1988.

Sosik, H. M., M. Vernet, and B. G. Mitchell, A comparison of particulate absorption properties between high- and mid-latitude surface waters, *Antarct. J. U. S.*, *27*, 162–164, 1992.

Subramaniam, A., and E. J. Carpenter, An empirically derived protocol for the detection of blooms of the marine cyanobacterium Trichodesmium using CZCS imagery, *Int. J. Remote Sens.*, *15*, 1559–1569, 1994.

Thomas, A. C., F. Huang, P. T. Strub, and C. James, Comparison of the seasonal and interannual variability of phytoplankton pigment concentrations in the Peru and California Current Systems, *J. Geophys. Res.*, *99*, 7355–7370, 1994.

Watts, D. R., Gulf Stream Variability, in *Eddies in Marine Science*, edited by A. R. Robinson, pp. 114–144, Springer-Verlag, New York, 1983.

Wiebe, P. H., The Biology of cold-core rings, *Oceanus*, *19*, 69–76, 1976.

Yentsch, C. S., and N. Garfield, Principal areas of vertical mixing in the waters of the Gulf of Maine, with reference to the total productivity of the area, in *Oceanography from Space*, edited by J. F. R. Gower, pp. 303–312, Plenum, New York, 1981.

Yentsch, C. S., D. A. Phinney, and J. W. Campbell, Color banding on Georges Bank as viewed by coastal zone color scanner, *J. Geophys. Res.*, *99*, 7401–7410, 1994.

Yoder, J. A., C. R. McClain, J. O. Blanton, and L.-Y. Oey, Spatial scales in CZCS-chlorophyll imagery of the southeastern U.S. continental shelf, *Limnol. Oceanogr.*, *32*, 929–941, 1987.

Zabicki, K. E., Determining the existence of cyanobacterial blooms using CZCS imagery, student report, NASA/Univ. of Md. Summer Fellowship Program in Remote Sens. of the Oceans, Greenbelt, Md., 1995.

Zaneveld, J. R. V., A. H. Barnard, and W. S. Pegau, A nearly backscattering independent algorithm for the retrieval of spectral absorption from remote sensing reflectance, paper presented at Ocean Sciences Meeting, AGU, San Diego, Calif., 9–13 February 1998.

L. V. Martin Traykovski and H. M. Sosik, Biology Department, MS 32, Woods Hole Oceanographic Institution, Woods Hole, MA 02543, USA. (lmartin@whoi.edu; hsosik@whoi.edu)