Fall 12-3-2018

Multispectral Classification of Gulf of Maine Surface Waters: Seasonal and Interannual Variability

Brianna M. King
University of Maine, brianna.m.king@maine.edu

MULTISPECTRAL CLASSIFICATION OF GULF OF MAINE SURFACE WATERS: SEASONAL AND INTERANNUAL VARIABILITY

By

Brianna M. King

B.S. Rensselaer Polytechnic Institute, 2012

A THESIS

Submitted in Partial Fulfillment of the

Requirements for the Degree of

Master of Science

(in Oceanography)

The Graduate School

The University of Maine

December 2018

Advisory Committee:

Andrew C. Thomas, Professor of Oceanography, Advisor

David Townsend, Professor of Oceanography

Huijie Xue, Professor of Oceanography

MULTISPECTRAL CLASSIFICATION OF GULF OF MAINE SURFACE WATERS: SEASONAL AND INTERANNUAL VARIABILITY

By Brianna M. King

Thesis Advisor: Dr. Andrew C. Thomas

An Abstract of the Thesis Presented
in Partial Fulfillment of the Requirements for the
Degree of Master of Science
(in Oceanography)
December 2018

Bio-optically, surface waters of the northeastern US continental shelf are strongly heterogeneous, exhibiting highly variable distributions in both time and space of suspended sediment, colored dissolved organic matter (CDOM), and phytoplankton concentration and community structure. These render the standard global NASA satellite chlorophyll algorithm suspect. However, spectral signatures of the water are well quantified by the actual satellite-based multispectral reflectance (Rrs) measurements. Here, we use 6 bands of Rrs measurements from 19 years (1998-2016) of monthly composite SeaWiFS and MODIS data to identify the spectral signatures of dominant water types present over the northeast shelf. A merged multivariate clustering approach that incorporates neural network-based Self-Organizing Maps and agglomerative hierarchical clustering is used to group similar spectral signatures across the entire time and space domain. These sort the surface water into 8 dominant groups. Remapped results provide a climatological view of seasonal cycles of these bio-optical groups and their spatial geography and then quantify their interannual variability over 19 years. Results indicate the same 2-3 spectral groups dominate the study area each year. Several less spatially predominant groups are present mainly along the coast and over George's Bank. Spectral signatures of the water groups suggest that three groups represent the clearest waters and two groups the most particle and

absorption dominated waters, with other groups representing varying concentrations of CDOM and phytoplankton. These descriptions are consistent with both the geographic position of the groups and calculated inherent optical properties of the groups. Interannual variability of the groups at each location effectively captures boundaries between ecological regions, reflecting frontal zones reported in the literature. Quantified trends in the spatial coverage of the groups over the 19 years shows a significant decline in the time/space coverage of the clearest water groups. These trends are strongest and significant in summer months, consistent with strong summer trends in sea surface temperature. There are also systematic trends in water groups in March and May on either side of the April spring bloom, suggesting phenology shifts in bloom characteristics. Over interannual time scales, the more turbid coastal water groups have correlations to wind strength and river discharge. These results provide new satellite-based ocean color measurements of shelf variability over seasonal and interannual time scales and present managers with new tools to monitor biologically important change over this economically critical and rapidly changing ecosystem.

ACKNOWLEDGMENTS

I would like to thank my advisor, Dr. Andrew Thomas, for his guidance and support along each step of the way towards the completion of my degree. I would like to thank my committee members, Dr. Huijie Xue and Dr. David Townsend for sharing their expertise and perspectives. I would also like to thank my friends and family, especially Andre Bucci, Stephen Moore, Gabby Hillyer, my fiancé Mike Lemieux and my parents David and Denise King, who have been an unwavering source of support and encouragement along this journey.

This work was funded by a NASA EPSCoR award to Dr. Andrew Thomas through the Maine Space Grant Consortium, the University of Maine Office of the Vice President for Research, and the University of Maine School of Marine Sciences.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

1.1. Overview

The surface ocean exhibits variations in color as a result of the absorption and reflectance of the water molecules themselves combined with the absorption and reflectance of any particles or dissolved material in the water (Mobley, 1994). In coastal zones, varying concentrations and types of suspended sediment, colored dissolved organic matter (CDOM) and phytoplankton alter the color of the water from blue to green to shades of yellow and brown (Arnone and Gould, 1998; Yoder and Kennelley, 2006; McClain, 2009; Mélin and Vantrepotte, 2015). Quantifying this color provides insights into the bio-optical and biogeochemical characteristics of the water.

Satellite ocean color sensors such as the Sea-viewing Wide Field-of-view Sensor (SeaWiFS) and the Moderate Resolution Imaging Spectroradiometer (MODIS) measure reflectance at multiple bands of the visible spectrum. Repeat orbits over many years offer large-scale, consistent and repetitive data over decades (SeaWiFS 1997-2010; MODIS 2003-present). Measurements of ocean color have revolutionized our understanding of many aspects of ocean biogeochemistry and its variability (McClain, 2009). Specific examples include chlorophyll variability from small (e.g. Kahru et al., 2004) to large (Polovina et al., 2008) spatial scales and from short (e.g. Fiedler, 1984; Thomas et al., 2003) to longer time scales (e.g. Signorini et al., 2015), primary production assessments (e.g. Moore and Abbott, 2000; Goes et al., 2004; Behrenfeld et al., 2006), river plume variability (e.g. Thomas and Weatherbee, 2006; Molleri et al., 2010; Hopkins et al., 2013), biological effects of storms (Babin et al., 2004), phytoplankton community structure (e.g. Alvain et al., 2005) and phytoplankton phenology (Ji et al., 2007; Platt et al., 2009; Song et al., 2010; Ardyna et al., 2017). One important application of these satellite-measured ocean color data sets is to isolate and map time/space boundaries of ecological provinces that have similar

biogeochemical characteristics and/or variability (e.g. Longhurst, 1998; Oliver and Irwin, 2008; Irwin and Oliver, 2009; Vantrepotte and Méin, 2011; Reygondeau et al., 2013), something that is not possible with in situ data. Such maps provide an important tool to monitor change.

The overarching goals of this study are to isolate and describe the dominant bio-optical surface water types on the northeast US shelf (Fig. 1) and then quantify their time and space variability. Nineteen years (1998-2016) of SeaWiFS and MODIS multispectral measurements are used to develop a climatological view of the seasonality of dominant water color types and then quantify their interannual variability over the 19-year period. The work builds on previous efforts to identify and map regions of similar bio-optical variability reported in the literature and circumvents known problems with estimating oceanographic variables such as chlorophyll from space-borne ocean color measurements in optically complex coastal water.

1.2. Bio-optical geography from satellite data

Past studies subdivide the ocean into geographic regions based on various measures of similarity between properties of interest. Longhurst (1998, 2006) provided a global ocean identification of static ecological provinces based on satellite ocean color data, in-situ measured nutrient data and various physical parameters. Subsequent studies inspired by Longhurst's provinces worked to improve classification of surface waters globally and regionally. Oliver and Irwin (2008) identified 81 global biogeographic provinces using a more objective approach based on clustering techniques applied to 4 years (2003-2006) of MODIS water leaving radiance (Rrs) data and sea surface temperature (SST) measurements. Mélin and Vantrepotte (2015) used 7 years of global SeaWiFS Rrs data and an unsupervised clustering method to distinguish water masses in coastal and shelf regions worldwide. Their results led to the identification of 16 optically distinct water classes in coastal waters globally, of which 5-6 described the US northeast shelf region. These 5-6 classes represented the most turbid waters identified in their study, demonstrating the significant tidal mixing, river influence and high

phytoplankton concentrations that are characteristic to this region. Kavanaugh et al. (2014) combined concurrent SeaWiFS chl-a and PAR measurements with satellite SST data over the North Pacific from 1998-2010. They used a two-step data reduction clustering technique involving probabilistic self-organizing maps and hierarchical agglomerative clustering to identify unique seascapes and document the seasonal migration of these seascapes. Cluster analyses on MODIS SST and chl-a data from 3 time periods in 2003 (Devred et al., 2007) subdivided the Northwest Atlantic into ecological provinces, documented their seasonal variation and compared these provinces to those of Longhurst's for the same region. Traykovski and Sosik (2003) avoided chlorophyll measurements and compared 2 multivariate feature-based classification methods based on 3 bands of Rrs values from 2 days (July 7, 1980 and October 8, 1997) of Coastal Zone Color Scanner (CZCS) and SeaWiFS data. Their results over their Northwest Atlantic study area isolated 6 different water types, with the northwest shelf region dominated by 2 optical water types they label as Gulf of Maine and Georges Bank.

1.3. Coastal optical complexity

Coastal regions and semi-enclosed shelf seas such as the northeast shelf are characterized by strong spatial and temporal hydrographic and biological variability due to impacts from runoff and river discharge, tidal mixing and variations in tidal magnitude, wind-driven mixing, advection of different water masses of varying nutrient concentrations (Townsend et al., 2015) and varying bathymetry. Mid-latitude shelves are also subject to strong seasonal variability (Townsend et al., 2006). These influences interact to produce strong variability in the concentration and species composition of phytoplankton, suspended matter size and type, and CDOM, leading to complex and varying optical properties of coastal waters (IOCCG, 2000; Balch et al., 2012; He et al., 2013). This complexity poses challenges to the quantification of satellite-based oceanographic quantities such as chlorophyll (IOCCG, 2000; Vantrepotte et al., 2012; Ye et al., 2016).

After atmospheric correction, multispectral satellite instruments measure remote sensing reflectance (Rrs) across bands of select wavelengths within the visible spectrum. Standard NASA algorithms developed for the calculation of chlorophyll concentration are based on ratios of the blue to green Rrs values and have been widely and successfully utilized for open ocean waters (Uitz et al., 2006; Brewin et al., 2010; Vantrepotte and Mélin, 2011) where color is primarily a function of chlorophyll concentration. In coastal waters, however, strong reflectance over the visible spectrum by non-algal particles suspended in surface waters and strong absorbance in the blue wavelengths by CDOM affect the Rrs values and lead to an overestimation of chlorophyll (Darecki and Stramski, 2004; Dall'Olmo et al., 2005; Garcia et al., 2005; Werdell et al., 2009). This bias is especially problematic in the Gulf of Maine (GOM) where significant quantities of CDOM delivered by local rivers remain in the surface water (Balch et al., 2004). Other approaches to the use of Rrs data to estimate oceanographic variables take a semi-analytic approach, inverting spectra to derive estimates of phytoplankton and dissolved substance absorption, and particulate backscatter (e.g. Garver and Siegel, 1997; Lee et al., 2002). In their standard application, all these approaches are optimized to a global average and most of their validation is carried out on open ocean measurements making their direct application in localized coastal regions such as the Gulf of Maine challenging.

Here, we attempt to avoid biases in the estimation of oceanographic quantities such as chlorophyll by directly examining the actual Rrs measurements. These offer a multivariate view of the color spectrum of each parcel of water, in a systematic time series, over the duration of our study period.

1.4. Study area background

The study area (Fig. 1) consists of the North American shelf extending from immediately south of Massachusetts including Nantucket Shoals, northeast to the western portion of the Scotian shelf. The area is dominated by the Gulf of Maine, a semi-enclosed shelf sea with bathymetry restricted interaction with the Atlantic Ocean. Low salinity surface water flows southwest down the Scotian Shelf and enters the eastern GOM around Nova Scotia (Smith 1989). During the spring, Scotian Shelf Water flows into the Bay of Fundy along the eastern side, where it is tidally mixed and then carried out of the Bay of Fundy along the western side to join the Eastern Maine Coastal Current (EMCC). During the remainder of the year, Scotian Shelf Water flows more directly into the EMCC, entering the current at the mouth of the Bay of Fundy (Pettigrew et al., 1998; Xue et al., 2000). Residual circulation in the GOM is cyclonic (Pettigrew et al., 1998). The EMCC flows southwest along the Maine coast as a relatively cold, well-mixed and nutrient-rich body of water (Townsend et al., 1987). The EMCC seasonally bifurcates in the

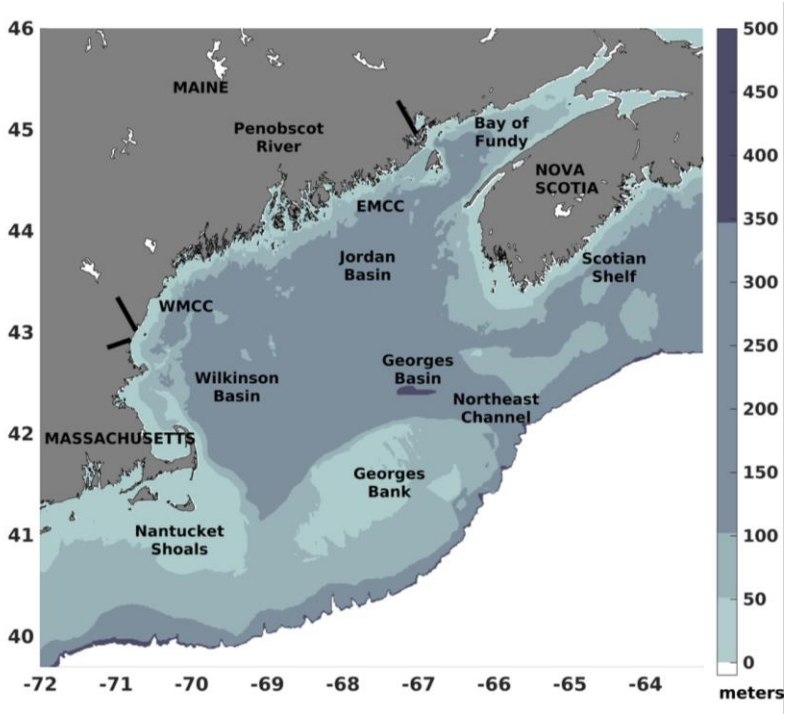

Fig. 1. The study area showing geographic names of specific locations, political boundaries and bathymetry. Regions of two coastal currents along the Maine coast are shown as EMCC and WMCC. Offshore areas (> 500 m) have been masked out and are shown in white.

vicinity of Penobscot Bay. One branch continues west along the coast as the Western Maine Coastal Current (WMCC), while the other branch turns offshore, contributing to the cyclonic gyre over Jordan Basin (Pettigrew et al., 2005; Townsend et al., 2006). The WMCC flows westward along the coast towards Massachusetts, where it contributes to the cyclonic gyre around Wilkinson Basin. The coastal flow is influenced by seasonally varying river discharge, the largest of which are the St. John and Penobscot Rivers (Townsend et al., 2006). The contribution of freshwater to the coastal current, especially during the spring, impacts the circulation patterns along the coast in addition to nearshore stratification (Townsend, 1991). Some water from western GOM is entrained in the clockwise residual circulation around Georges Bank, while other water exits the GOM directly through the Northeast Channel (Pettigrew et al., 1998, 2005). South of Cape Cod, shelf flow is to the west.

Regional bathymetry (Fig. 1) plays a major role in circulation, vertical mixing and resulting biological and turbidity patterns. The deepest areas in the GOM are the three basins, Jordan, Wilkinson and Georges Basin, and the Northeast Channel. These areas reach depths of >200 m, with the basins all reaching depths greater than 250 m. These basins stratify earliest in the season and most strongly in the summer months, leaving a nutrient-poor, oligotrophic upper water column. In contrast, much of the upper Bay of Fundy, coastal regions off the southern tip of Nova Scotia, Georges Bank, and Nantucket Shoals are shallow (Brooks, 1985). Strong tidal currents in the region induce vertical mixing that keeps these shallow regions better mixed throughout the year, injecting nutrients into the upper water column and creating regions of elevated phytoplankton biomass (Thomas et al., 2003; Townsend et al., 2006; Hasegawa et al., 2011) and increased suspended sediment and turbidity (Hargrave et al., 1983; Brooks, 1985). Hydrographic frontal zones separate the well-mixed waters over shallow regions from the stratified deeper waters. These features are prominent, especially in summer, in the Bay of Fundy, the southwestern coast of Nova Scotia, around Nantucket Shoals, around Georges Bank and on the southern boundary of the cold EMCC (Loder and Greenberg, 1986; Townsend et al., 2006). A shelf-break front is

often present along the seaward side of the study region. In the southern portion of the study area, this front separates the fresher and cooler waters of the shelf from the warmer, more saline waters of the continental slope that have a stronger influence of the Gulf Stream (e.g. Gawarkiewicz et al., 1996). In the northern portion of the study area, this front separates the fresher Scotian Shelf water from slope water of northern origin (e.g. Fournier et al., 1979; Townsend et al., 2006). A systematic mapping of frontal features in satellite sea surface temperature data (Ullman and Cornillon, 1999) tracks the seasonal position of these fronts. They also document a winter frontal feature nearshore in the GOM that separates colder, shallow coastal water from warmer water over the deeper basins resulting from differing impacts of strong winter atmospheric cooling.

The study region is highly biologically productive (O'Reilly et al., 1987), with clear seasonal and spatial patterns in phytoplankton concentration (Thomas et al., 2003). In winter (December-February), chlorophyll concentrations throughout the region are at their lowest. Concentrations peak during the spring bloom in March-May, decline in summer from June-July and then increase again between August-November during the fall bloom. Seasonal cycles are weaker over shallow bathymetry, where vertical mixing brings new nutrients into the upper water column throughout the year (e.g. Townsend and Thomas, 2001) and strongest over the deeper basins (Thomas et al., 2003). Yoder et al. (2002) show a clear relationship between SST patterns and surface chlorophyll patterns over the study area. Shallow regions also create shifts in the timing of the spring bloom (Townsend et al., 1992, 1994) as winter light limitation is relived earlier over shallow water columns. There is evidence that the relatively fresh surface water from the Scotian Shelf impacts the timing of the spring bloom in the GOM (Ji et al., 2007), driving earlier blooms upstream in the east and delayed blooms in the western GOM. Seasonality in phytoplankton community structure contributes to the bio-optical complexity of the region. Species succession over Georges Bank (Townsend and Thomas 2001, 2002; Gettings et al., 2014) begins as early as January or February, when diatoms dominate. Dinoflagellates become dominant after this and by

June diatom concentrations begin to increase again in response to silicate regeneration. Over deeper basins, the spring bloom, dominated by diatoms, gives way to summer conditions dominated by smaller cells. Together, the interaction of bathymetry and vertical mixing create a strongly heterogeneous phytoplankton regime in both time and space.

CHAPTER 2

METHODS

2.1. Data

Daily 1 km resolution, atmospherically-corrected remote sensing reflectance (Rrs) data from NASA's MODIS and SeaWiFS missions were acquired from NASA's Ocean Color server (oceancolor.gsfc.nasa.gov) for the study area. The SeaWiFS data used here cover full years from 1998-2007 when the incidence of missing data due to mission problems was minimal. MODIS data cover the period 2003-2016. Together, these sensors provide a 19-year study period, 1998-2016. Daily swath data were remapped to a standard map projection. To reduce data volume, data gaps due to cloud cover and strong daily variability in Rrs retrievals likely caused by poor cloud/fog masking or variability in atmospheric correction success, monthly average data were deemed an effective temporal resolution to view the seasonal and interannual variability we were interested in. Monthly composite images were formed using a multi-step smoothing approach. Daily scenes were first median filtered in space using a 5x5 operator. The monthly mean at each location was then formed using the 50% of valid retrievals centered on the median Rrs value within a 5x5 pixel region over that month. The resulting monthly fields were then further smoothed with a 5x5 median operator. This approach was effective in dramatically reducing the impact of Rrs outliers from the composite averaging and producing final monthly composite images that were relatively smooth. All further analyses were performed on the 19-year monthly composite data time series.

SeaWiFS and MODIS sensors each have slightly different spectral bands and wavelength centers. The monthly compositing was applied to the 6 visible spectral bands of Rrs measurements from the bands of each sensor that were most complementary (Table 1).

The 5-year period of mission overlap (2003-2007) was used to test whether, within the context of our goals and analysis approach, the spectral data from the two sensors could be treated as equal to form a single time series. Four separate analyses were conducted, subjected to our grouping analysis (described in detail below) and compared. In the first two analyses, SeaWiFS and MODIS monthly data from 2003-2007 were analyzed separately. The third and fourth analyses utilized a combination of both SeaWiFS and MODIS data. The third analyzed a time series with SeaWiFS data from 2003-2004 and MODIS data from 2005-2007 and the fourth performed the grouping analysis using SeaWiFS data from 2003-2004 and MODIS data from 2003-2007. Select months from all four analyses were compared by both mapping the final group membership from each combination and by quantifying the number of pixels assigned to each group. Between all four analyses, very few pixels changed groups and monthly maps and the spectral characteristics of the final groups from each test were essentially identical. We concluded that for the level of grouping we were conducting, the bands in the 6 wavelengths we used from the two missions could be considered identical. A single 19-year monthly time series was created using SeaWiFS data for the period of 1998-2002 and MODIS data for 2003-2016.

A bathymetry mask was fitted to each monthly composite image excluding waters deeper than 500 m to focus analysis on shelf waters.

Concurrent environmental metrics representing SST anomalies (SSTa), river discharge and wind strength were acquired. Daily, optimally interpolated NOAA OISST 0.25 x 0.25 degree resolution data (Reynolds et al., 2007; Banzon and Reynolds, 2017) from January 1982 to December 2016 were downloaded from the NOAA Earth System Research Laboratory (http://www.esrl.noaa.gove/psd/).

Table 1. Spectral Rrs bands unique to SeaWiFS and MODIS sensors used in this analysis.

Sensor	Rrs(λ) nm band 1	Rrs(λ) nm band 2	Rrs(λ) nm band 3	Rrs(λ) nm band 4	Rrs(λ) nm band 5	Rrs(λ) nm band 6
SeaWiFS	412	443	490	510	555	670
MODIS	412	443	488	531	555	667

Monthly averages and climatologies were calculated for the study area and used to calculate monthly SST anomalies. A series of grid cells from the center of the GOM were subsampled and averaged to characterize interannual SST variability over the study period for the entire study region. Daily mean Penobscot River discharge data were downloaded from the USGS Water Resources (https://waterdata.usgs.gov/usa/nwis). Monthly average discharge was calculated for 1998-2016 and on interannual time scales is considered representative of relative discharge from rivers in the study region. Monthly wind velocities from the National Centers for Environmental Prediction (NCEP) from grid locations (42-44 N and 290-294 E) in the central GOM were downloaded as u-wind and v-wind reanalysis vectors from the NOAA Earth System Research Lab (https://www.esrl.noaa.gov). These were used to calculate monthly average wind speed for the study region over the study period January 1998 to December 2016.

2.2. Two-step multivariate cluster analysis

We group optical water types through time and space based on their similarity over the 6 spectral bands of Rrs data. Each data pixel (approximately 243 k) in each of the 228 monthly composite images was assigned a group following a two-step multivariate cluster analysis. We first use Self-Organizing Maps (Kohonen, 2001) to affect a large data reduction and partition data pixels into a relatively large number of initial groups. We then further simplify the output of the Self-Organizing Map to fewer groups through hierarchical clustering.

Prior to clustering, Rrs values in each of the 6 spectral bands were normalized by their mean and standard deviation. The overall mean of each band was subtracted from the original Rrs values, and then divided by the band's standard deviation. The resulting normalized Rrs data values weigh each band equally, allowing each to contribute equally to the distance-based multivariate clustering analysis below.

Self-Organizing Maps (SOMs) are an effective data reduction technique (Liu and Weisberg, 2011) for large, multi-dimensional data sets and have been utilized extensively in social and geographic sciences

(e.g. Hara et al., 1994; Kaski and Kohonen, 1996; Chen and Shrestha, 2000; Nerbonne and Heeringa, 2001). Previous examples of their use within the field of oceanography include (Ainsworth, 1999; Richardson et al., 2003; Risien et al., 2004; Iskandar et al., 2008). SOMs are a form of artificial neural network, utilizing competitive and cooperative learning to train a group of nodes, modeling the input data onto a reduced dimension node field based on measurements of attribute similarity. In this study, these attributes are the Rrs values in the 6 bands. After iterating through the entire data set many times, and continuously updating the SOM node field, the final node field is a 2-dimensional time and space model of the original data where each node is defined by 6 attributes representing Rrs values from the 6 satellite bands. Upon completion, each original data point can be assigned an output node that most closely represents it, effectively grouping the data (Kohonen, 2001).

We used the SOM package made available by the Department of Information and Computer Science at the Helsinki University of Technology (http://www.cis.hut.fi). The selected distance metric for this SOM analysis was the commonly used Euclidean distance. The number of nodes and their 2D MxN arrangement is user defined, usually determined through trial and error (Kohonen, 2013), with consideration of the 2D shape of the input data (Kohonen, 2001). We tested 40 different sized and shaped node arrangements and tracked the quantization error, the measure of the final average Euclidean distance between all data points and their best matching node, to track the effectiveness of the 2D node arrangement (Kohonen, 2013). Based on these values, we determined that 36 output nodes in a 6x6 arrangement provided an effective compromise between the level of detail/separation we were seeking and an effective representation of the original data. Default values for the training rate and cooperative learning distance metrics in the SOM were used. Each pixel in all 228 monthly composite images were analyzed with the SOM as a 6-value vector in Rrs space and were modeled onto the 36 nodes, that have a 6-value vector representing the Rrs values. Upon completion, grouping is

accomplished by assigning each original data point a number from 1-36, indicative of the node that most closely represented it, measured as Euclidean distance.

The SOM algorithm is effective at retaining maximum diversity of the optical characteristics of the data pixels while assigning more nodes to areas of the data attribute space that have the most data. Examination of the spectral shape of each of these 36 nodes indicated that the characteristics of many of the nodes were reasonably similar. We wanted to further simplify the Rrs characteristics with the goal of distinguishing dominant oceanographic features throughout the study area time/space window. The 36 SOM nodes were further grouped into more optically distinct groups using a simple agglomerative hierarchical clustering with Euclidean distance and average linkage. Combinations of other separation and grouping metrics were examined, but the resulting dendrogram (Fig. 2) of the chosen method provided the simplest, most effective partitioning of the SOM nodes. A stopping rule was applied to determine the optimal separation distance to cut the dendrogram tree. While somewhat subjective, the stopping rule represents a balance based on vertical distance between merged groups and the number of final groups desired (Vesanto and Alhoniemi, 2000; Thomas et al., 2010). Large distances indicate an increased dissimilarity between groups. Here we identified the lowest region of the tree that exhibited the largest distances and resulted in merging the SOM nodes into the 8 most similar groups. Two other

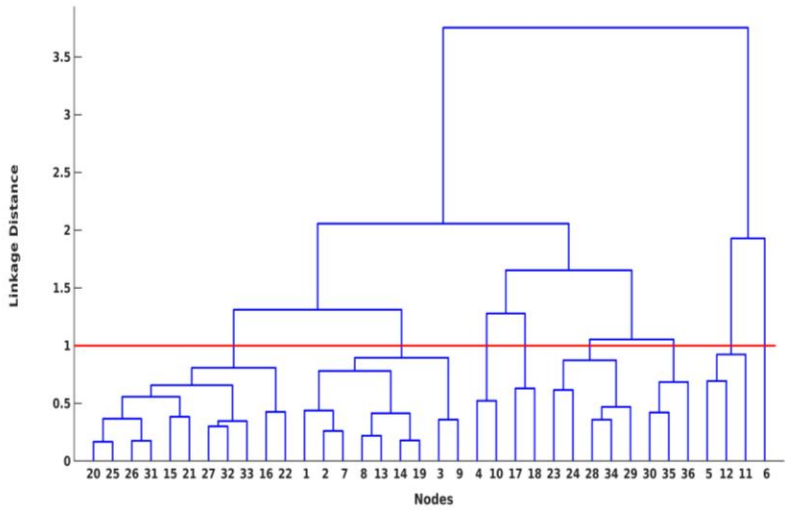

Fig. 2. Dendrogram of hierarchical clustering of 36 SOM nodes. The red line indicates where the stopping rule was applied, resulting in 8 final groups.

stopping distances were examined, resulting in 11 and 13 groups respectively. The spectral characteristics of these groupings showed that several of the groups exhibited strong similarities. We concluded that clustering the 36 SOM nodes into 8 groups was an effective grouping into sufficiently optically distinct water types. Each pixel in each month was then assigned a label from 1-8, indicative of which group it belonged to.

2.3. Monthly climatologies and interannual variability

Following the two-step cluster analysis, a monthly climatology time series was generated. At each location, in each calendar month, the climatological group number was defined as the group that most often occurred (the mode) over the 19 years of data. For the rare data points (< 0.0005% of the total data points) that had more than one group identified as the mode, the group with the lowest value was assigned to represent that data point in the climatology. Due to the small number of data points impacted by this arbitrary decision, the ultimate effect on the final results was deemed to be negligible.

The stability of these climatological values over the 19 years for each of the locations/months provides one estimate of interannual variability and was estimated by quantifying the number of years the assigned climatological value (mode) at each data location occurred. Values approaching 19 suggest locations that were highly stable over the 19 years and were consistently placed into the same group. Locations with much lower values were frequently placed into different groups, suggesting stronger interannual variability. Climatological seasonal variability was quantified as the percentage of the study area occupied by each group in each month. Interannual variability was quantified as the percentage of the study area occupied by each group in each of the 228 monthly composite images of the 19-year time series.

The above climatology and interannual variability calculations were also conducted on the original 36 SOM nodes to provide an increased level of detail in the analysis. These results are presented in the appendix.

CHAPTER 3

RESULTS

3.1. Group spectral characteristics

Mean spectra over all months and locations determined by the 6 Rrs values characterize the optical signature of each of the 8 groups separated by the clustering (Fig. 3). Based on the shape of the spectra, these 8 groups can be broadly described as three general bio-optical types: the clearest water in the study area/time series (groups 1, 2 and 5), water with an increased influence of varying concentrations of phytoplankton and CDOM (groups 3, 4 and 6), and more turbid/particle dominated water (groups 7 and 8).

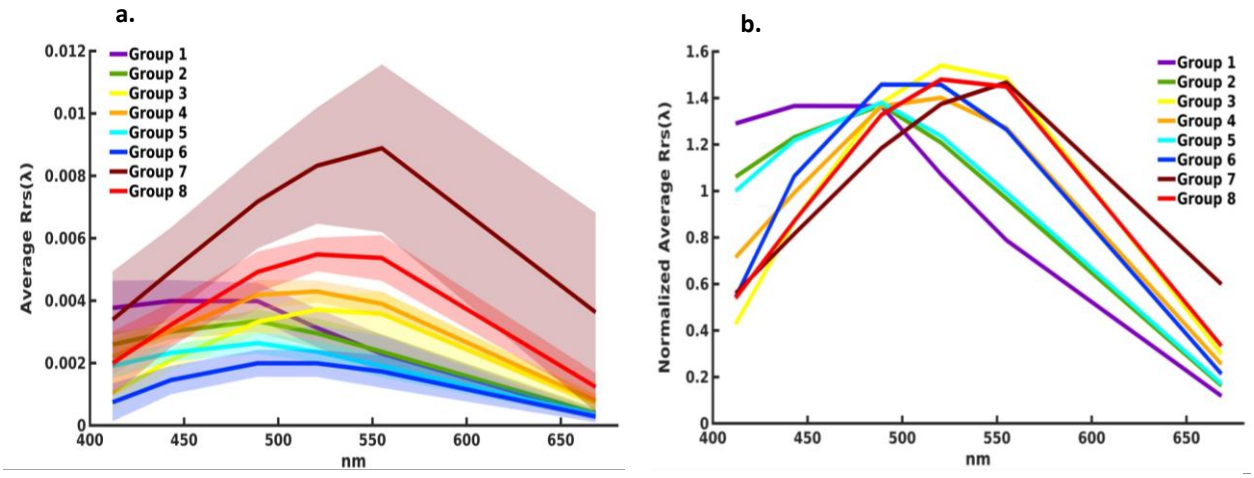

Fig. 3: Spectral signatures in the 6 bands for each of the 8 groups identified by the two-step multivariate clustering technique. a.) Mean spectra for each group. The shaded areas represent ± 1 standard deviation. b.) The same mean spectra normalized to the mean Rrs value of each band to provide a different view of differences in spectral shape.

Groups 1, 2 and 5 have spectra with relatively low Rrs values in the spectral region 550-670 nm (Fig. 3a) and relatively high and flat Rrs values in the blue-green region in the normalized spectra (Fig. 3b), indicative of the clearest waters in our study area. Group 1 represents the clearest waters. In the shorter (blue) portion of the spectrum, this group has the highest and flattest normalized reflectance

values (Fig. 3b), characteristic of weak phytoplankton and CDOM absorption, and water most closely approaching oligotrophic conditions. This group also exhibits the lowest normalized Rrs values from 510-670 nm (Fig. 3b), indicative of decreased absorption due to lower concentrations of particulate and dissolved matter (Yoder and Kennelly, 2006; Lohrenz et al., 2009; Mélin and Vantrepotte, 2015). In groups 2 and 5, the lower normalized values in the blue region of the spectrum (McClain, 2009; Vantrepotte et al., 2012), the spectral peak located around 490 nm (Ye et al., 2016), and the higher normalized values from 510-670 nm compared to group 1 suggests an increased influence of CDOM and suspended particle concentration. Groups 2 and 5, however, still exhibit significantly less influence by phytoplankton and non-algal particles than other groups. Together these three groups represent varying degrees of the clearest waters present in the GOM.

The groups with increased phytoplankton and CDOM influence (groups 3, 4 and 6) exhibit a plateau in their spectra maximum values (Fig. 3a) from approximately 500-555 nm and very low reflectance values at 412 nm. Of the 8 groups, group 3 is likely most influenced by CDOM, indicated by the lowest normalized Rrs value at 412 nm (Fig. 3b). Group 3 has high values between 510-670 nm, suggesting more suspended particle influence than groups 4 and 6. With the lowest spectral values across the shorter wavelengths (Fig. 3a), group 6 likely represents the water type with the strongest phytoplankton influence, as chlorophyll concentration is roughly inversely proportional to reflectance in the blue region of the spectrum due to absorbance (Yoder and Kennelly, 2006).

Groups 7 and 8 exhibit stronger Rrs values across most of the longer wavelengths of the spectra (Fig. 3a), indicative of stronger backscatter and increased particulate influence compared to the other groups. Of these 2 groups, group 7 has the strongest of these characteristics. The low normalized values in the shortest wavelength suggest increased influence of CDOM within these waters. The peak at ~555 nm suggests an increased amount of suspended sediment and other particulate matter (Lohrenz et al., 2009; Mélin and Vantrepotte, 2015; Ye et al., 2016).

3.2. Climatological seasonal variability

Mapping the climatological group at each location in each calendar month provides a spatial view of the climatological seasonal progression of bio-optical water types over the year (Fig. 4). The maps show a distinct seasonal progression in the dominant water groups present within the study area. Beginning in January, groups 4, 7 and 8 are present along much the entire Maine coast and in the Bay of Fundy. These three groups also dominate over Nantucket Shoals, the southwest coast of Nova Scotia and the center of Georges Bank (see Fig. 1). Group 5 dominates the western and central GOM whereas group 6 dominates the Scotian Shelf and eastern GOM. In February, group 6 occupies more of the central GOM and group 5 less. Groups 4, 7 and 8 have receded shoreward around Nantucket Shoals. The center of Georges Bank now includes contributions from group 3. Other patterns are similar to those of January. In March, group 6 covers most of the study area, including most of Nantucket Shoals and Georges Bank. The Bay of Fundy, the coast of Maine and the southern coast of Nova Scotia remain dominated by groups 7 and 8 with an increased contribution from group 3. In April, the study area is almost completely occupied by group 6, with the exception of the Bay of Fundy and the eastern Maine coast. Group 5 is almost entirely absent. By May, group 6 still dominates, but groups 7 and 8 are concentrated further up into the Bay of Fundy, groups 3 and 4 increase over Nantucket Shoals and group 5 reappears in patches in the central GOM and along much of the shelf break south of Nantucket Shoals and Georges Bank. In June, the dominant group covering most of the study area has shifted from group 6 to group 5, with group 6 now occupying coastal regions in the western GOM. The extent of groups 3, 7 and 8 has further contracted into the Bay of Fundy. Groups 4 and 8 appear over Georges Bank and group 1 appears along the shelf break, with group 2 separating it from group 5. By July, the dominant water type in the central GOM is group 1, extending along the entire shelf break, over the Northeast Channel and the deeper basins of the GOM. Again, group 2 can be found at the boundary between groups 1 and 5. Groups 3, 4, 7 and 8 remain low in areal coverage and continue to contract even further into the Bay of

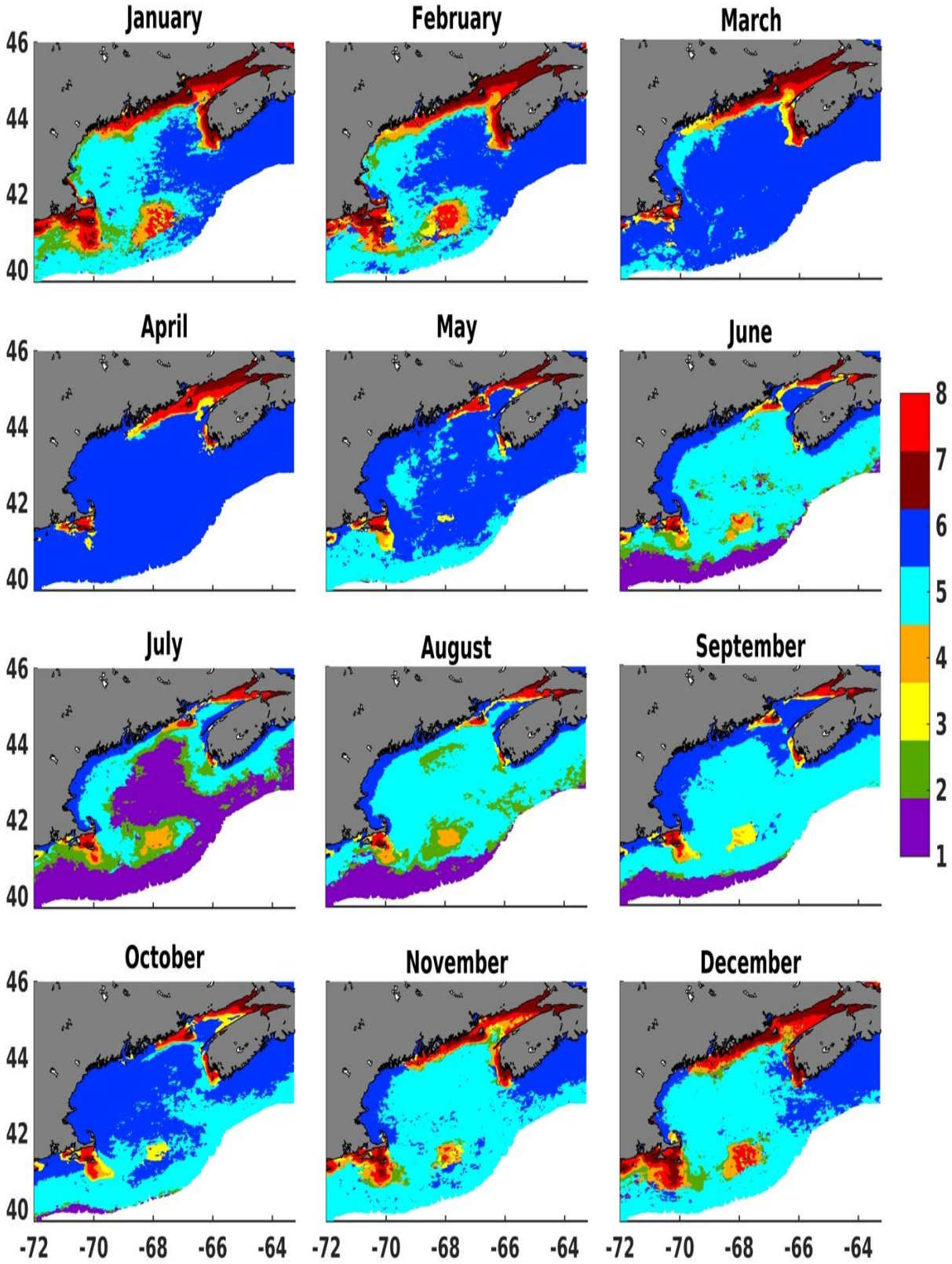

Fig. 4: Monthly climatology maps of group distribution over the 19-year study period. Climatological value at each location is the mode over the 19 years for each month.

Fundy. In August, group 5 resumes as the dominant group over much of the study area, with group 2 evident around Georges Bank and in patches in the central GOM and on the Scotian Shelf. Group 1 recedes to the southwestern shelf break region. By September, group 5 remains dominant in the central GOM, but group 6 has expanded outward from the coast and group 2 has largely disappeared. The area covered by groups 3, 4, 7 and 8 remains low and is primarily concentrated in the Bay of Fundy region. Group 3 replaces group 4 over the central portion of Georges Bank. Along the southwestern portion of the shelf break, the extent of groups 1 and 2 has decreased. In October, the central GOM is dominated by group 6 again, with group 5 along the eastern offshore edge of the study area. Groups 3, 4, 7 and 8 increase in area within the Bay of Fundy and over Nantucket Shoals. Group 1 and 2 are present, but strongly reduced at the southern shelf break. By November, group 6 has receded to the Scotian Shelf and the western coastal region of the GOM and group 5 once again dominates the central GOM. Groups 3, 4, 7 and 8 have expanded west down the coast from the Bay of Fundy and expanded around Nantucket Shoals and Georges Bank. December strongly resembles January, with the GOM dominated by group 5, group 6 on the Scotian Shelf and western GOM coast, and groups 4, 7 and 8 having a large areal coverage extending down the Maine coast from the Bay of Fundy and spreading out more over Nantucket Shoals and Georges Bank.

A quantitative summary of this climatological seasonality (Fig. 5) shows the spatial coverage of each group as a percentage of the total study area in each month. Comparing the y-axis of these plots shows the extent to which the study area is dominated by groups 5 and 6 in most months. Groups 1 and 2 both have maxima in the summer months (June, July, August), but group 1 is almost entirely absent in other months (fall, winter and spring), whereas group 2 has a secondary maximum in late fall/winter (November-February). Groups 3 and 6 have maxima in spring, peaking in March and April respectively, a second maxima in the fall peaking in October, and summer and winter minima. Group 4 has a seasonal cycle that resembles group 2 but has strong maxima in late fall/winter (November-February), and

smaller, secondary maxima in summer (June-August). Group 5 exhibits no obvious seasonal distribution pattern; the most defining feature of this group is the very low areal coverage in April. Groups 7 and 8 have similar seasonal cycles, with winter maxima and summer minima and group 7 having the strongest seasonality.

3.3. Interannual variability

Although there are strong and obvious seasonal cycles in the climatology, the location and extent of each of the 8 optical groups varies between years and months. One measure of interannual variability is the frequency of occurrence of the group at each location identified as the climatology. Stability maps (Fig. 6) identify locations that were more consistently placed into the climatological group each month

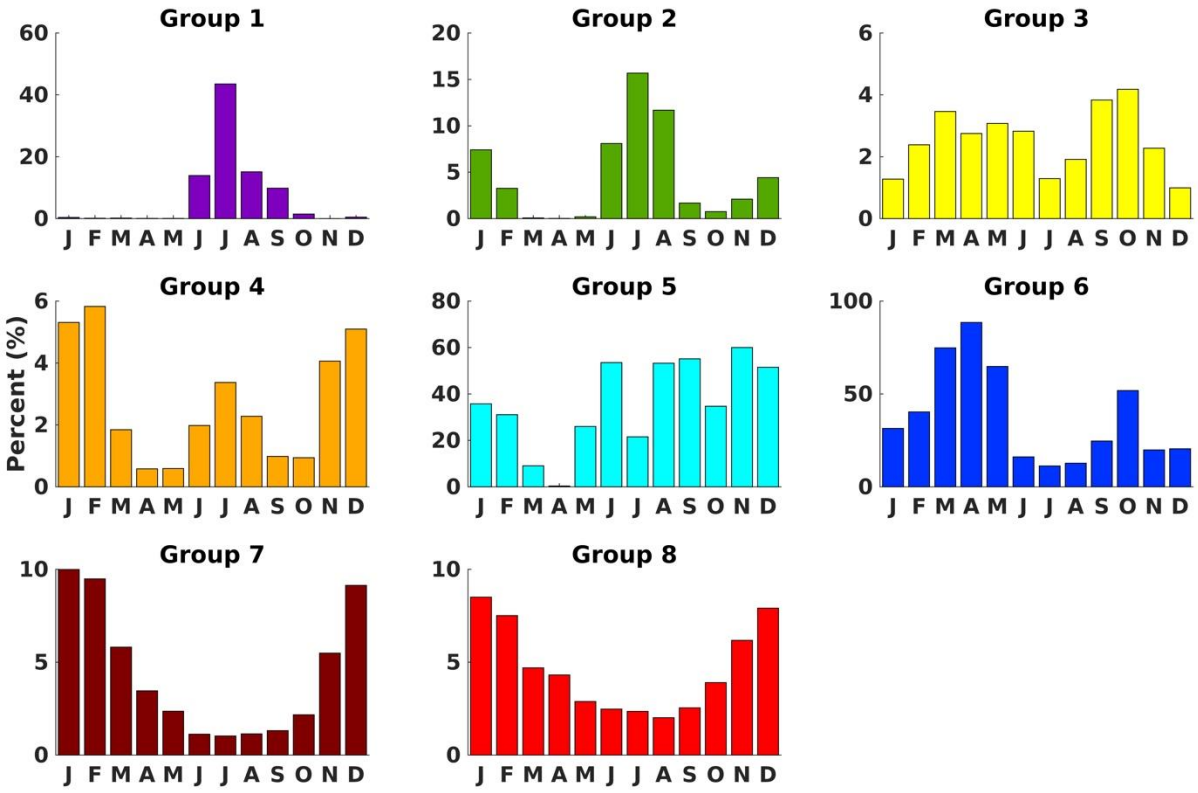

Fig. 5: Quantification of group climatological distribution. Percentage of the total study area occupied by each group in each monthly climatology map from Fig. 4.

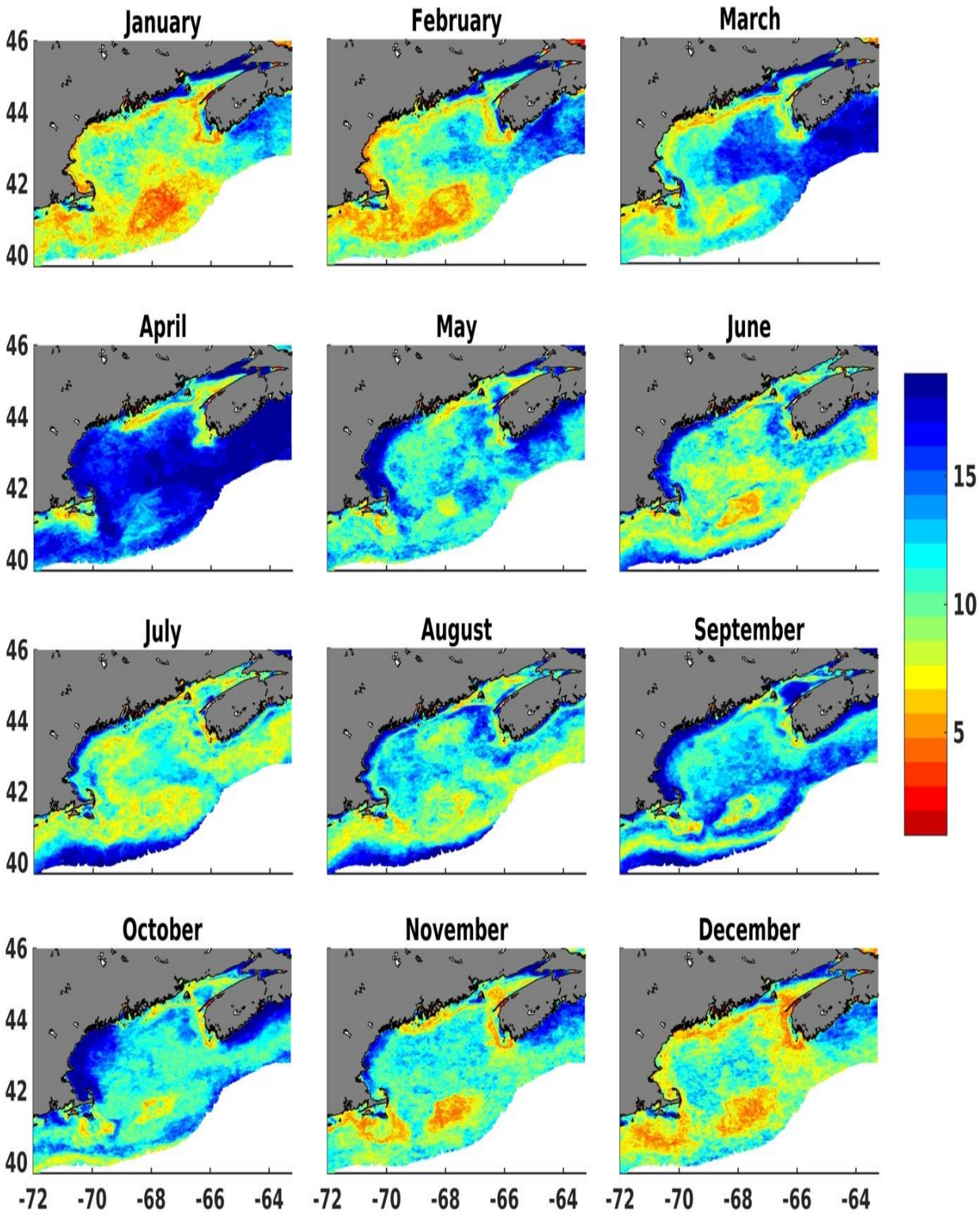

Fig. 6: Monthly stability maps of group climatological values. Regions of high stability (blue) are frequently classified as the single climatological group across the 19 years. Regions of low stability (red) are more frequently classified as a different group across the 19 years.

over the 19 years and those locations where the climatology (Fig. 4) hides stronger interannual variability. There are clear seasonal and spatial patterns to this interannual stability. In late fall and winter months (November-February), the stability of the climatological group in much of the study area is lowest. In contrast, spring (March and April) is the most stable, with April as the most stable month. Georges Bank and Nantucket Shoals are consistently the least stable regions in the study area. The southwestern coast of Nova Scotia and the Maine coastal region are also regions of strong interannual variability in most months. The Scotian Shelf, the western GOM coastal regions and the upper portions of the Bay of Fundy are the most stable regions in many months, especially in spring and summer. Similarly, the shelf-break region in the southern-most portion of our study area also exhibits strong interannual stability, especially in summer months.

Overall interannual variability is represented in Fig. 7 as the percent of the total study area that each group occupies in each month over the study period. The seasonality of many groups shown in the climatology (Fig. 4 and 5) is clearly evident, but differences from year to year are also evident. We supplement this view of interannual variability by plotting the monthly percent coverage of each group separately (Fig. 8) and simplify this further as the annual total percent coverage in each year (Fig. 9). Among the clearer patterns in this interannual variability are decreasing coverage of the study area by groups 1 and 2, strongest in the last 5-6 years of the time series. Both have a significant ($p < 0.05$) negative trend over the study period of 0.40 and 0.21% $year^{-1}$ respectively. We used the non-parametric Sen's slope to calculate these trends to avoid issues associated with the unknown underlying distribution of the data and to minimize the impact of outliers on each end of the times series. Conversely, groups 3 and 6 suggest an increasing trend, again, most strongly in the latter part of the time series. Group 3 has weak interannual variability over the initial part of the time series. However, after 2012 when it reached its lowest value, a steady increase in its areal coverage is evident. The last three years have the highest values of the entire time series. Although individual months of groups 6

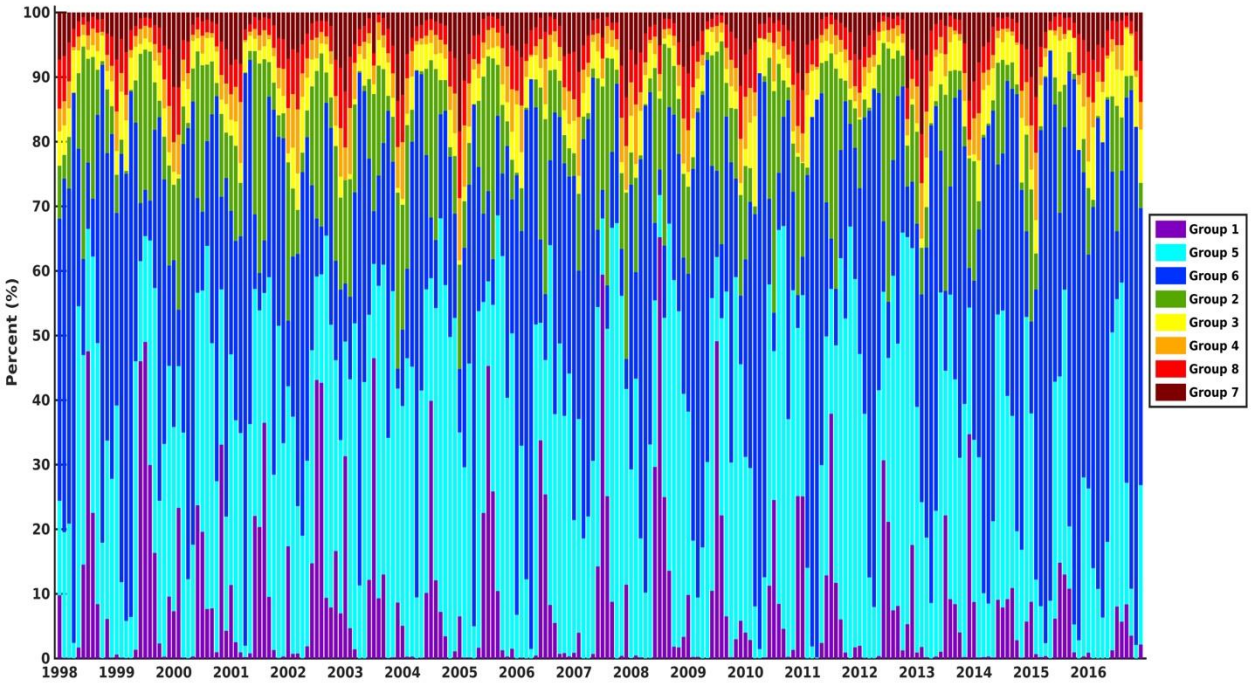

Fig. 7: Summary of group interannual variability over 19 years. The time series shows the percent of the total study area that each group covers for each monthly composite image.

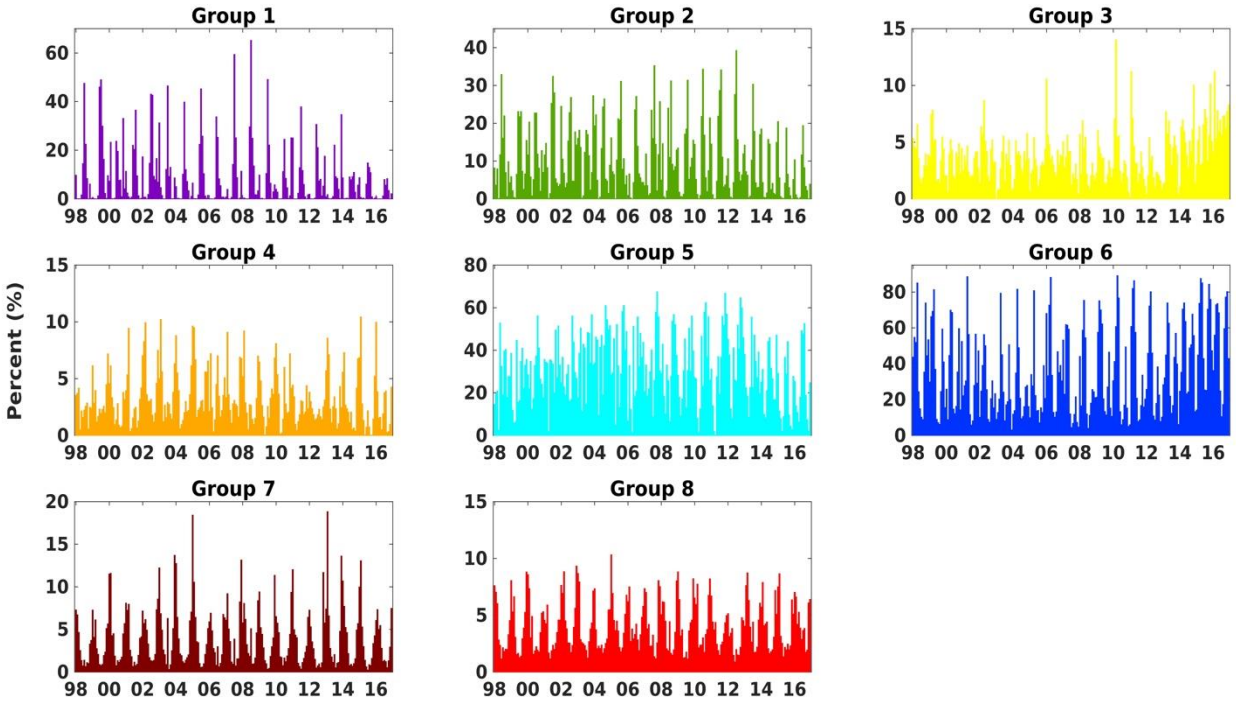

Fig. 8: Interannual variability of each group shown as the percent of the total study area that the group covers in each month over the study period. Separated by group to better observe interannual variability. Note the changing y-axis scale.

(Fig. 8) do not show a clear increasing trend, over annual periods (Fig. 9) the interannual variability shows a decreasing trend from 1998-2005, a large value in 2006, and then an increasing trend from 2007-2016, with highest values in the time series from 2014-2016. Groups 4, 5, 7 and 8 do not have obvious trends or patterns in Figures 8 or 9.

Beyond multi-year trends, some individual years show interannual events consistent across many of the 8 groups (Fig. 9). For groups 3, 4 and 8, the year 2012 has the smallest coverage of all years, and for group 5, this year is among the strongest. The year 2002 has the largest coverage for groups 1, 2 and 8 and is among the largest for group 4.

We separated the 19-year trends for each group into individual months to examine whether the trends evident in Fig. 9 occurred preferentially within certain seasons and whether certain groups had trends in specific seasons that the annual values did not show. A number of groups have significant trends in spring and summer months (Fig. 10 and 11) (see appendix for all plots). In spring, both March and May (Fig. 10) had significant decreasing trends in groups 2 and 5 over the time series and group 6 had a significant increasing trend. April trends were not significant. In summer (Fig. 11), group 1 has a significant decreasing trend in all three months (June-August), group 3 had significant increasing trends in June and August (and was also increasing in July, but not significantly), and groups 5 and 6 had significant increasing trends in June and August respectively. Both groups 5 and 6 were also increasing (but not statistically significant) in the other 2 summer months (July-August and June-July respectively). In winter months, although strong interannual variability was evident, trends over the study period (December and January, not shown) were not significant for any group, and even in adjacent months (November and February), only 1 group had a significant trend in each month, and neither was one of the two dominant groups (5 and 6). (See appendix for tables with all slope and significance values.)

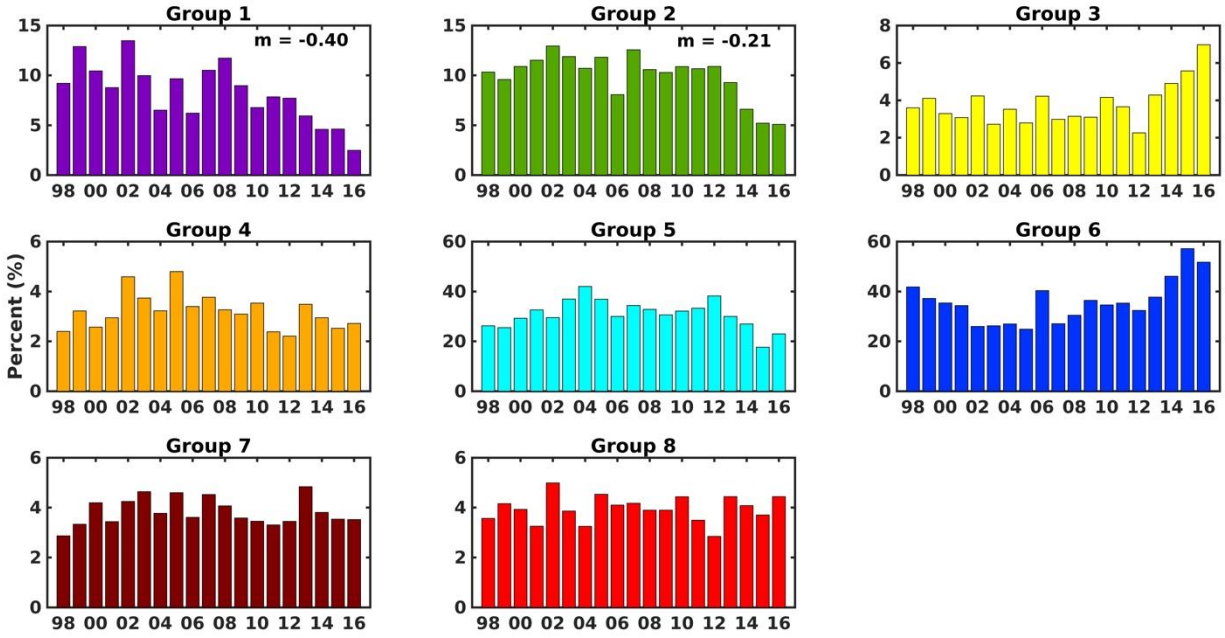

Fig. 9. Annual total of percent coverage in each year. Annual value was calculated using the total number of data points from each year and the total number of points of each group.

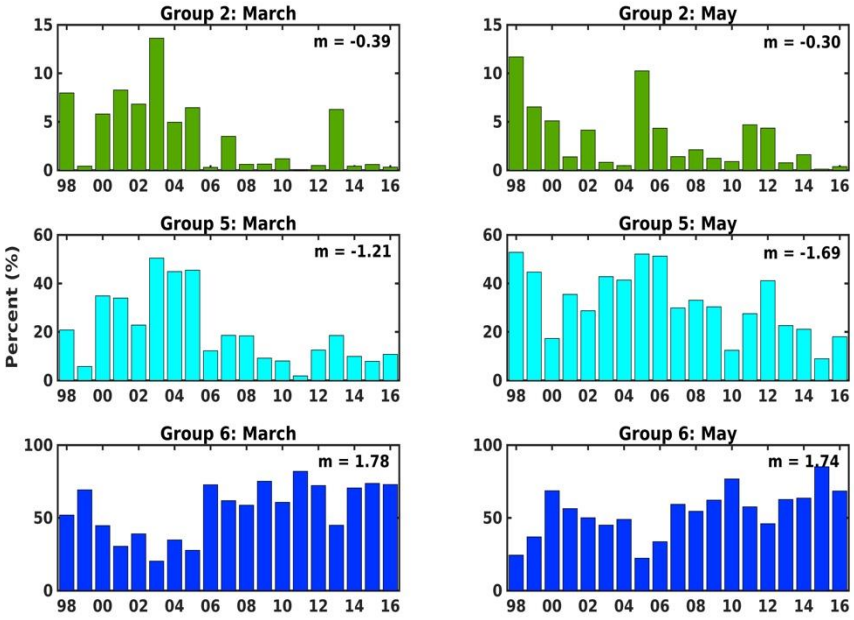

Fig. 10: Groups with significant trends over the time series in individual months in spring (March and May). April did not have significant changes in group distribution patterns. Slopes (m, as % year^{-1}) were calculated using the non-parametric Sen's slope calculation.

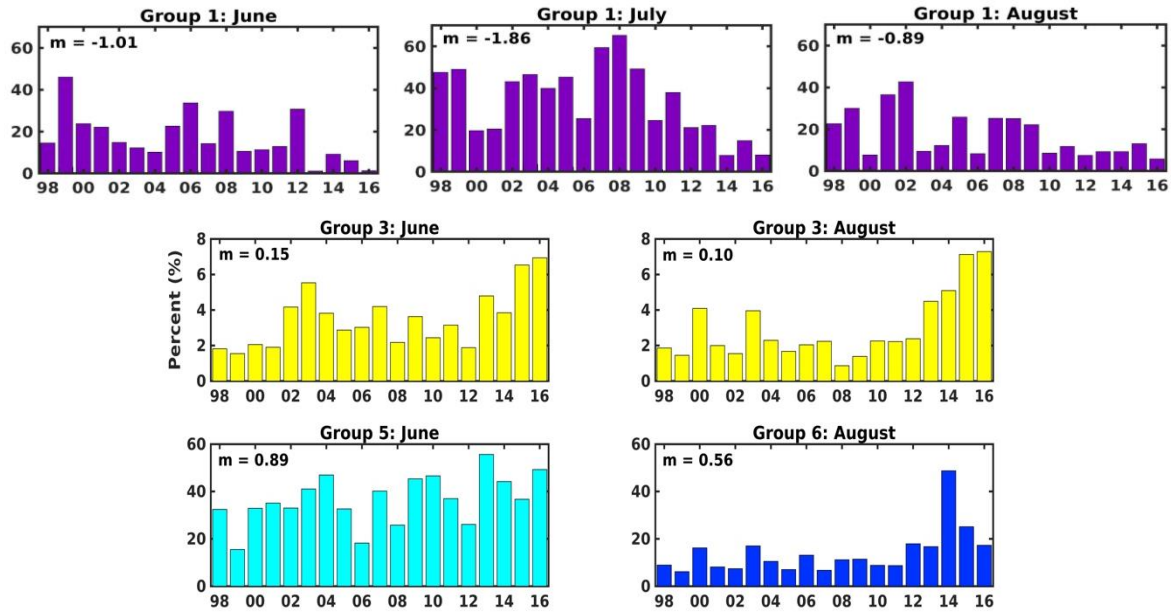

Fig. 11: Groups with significant trends over the time series in individual months in summer (June, July, August). Slopes were calculated using the non-parametric Sen's slope calculation.

CHAPTER 4

DISCUSSION

4.1. Climatological patterns and group spectral and bio-optical characteristics

Inherent optical properties (IOPs) calculated over our 8 groups provide a check on the bio-optical interpretation of water characteristics based on the Rrs spectral shapes evident in Fig. 3. Marine IOPs quantify contributions to the absorption and scattering properties of the light field in the water column and so provide a link between the satellite-measured remote sensing reflectance (Rrs) and the biogeochemical properties of the water being measured (Werdell et al., 2018). Three IOP variables were calculated using data from the generalized IOP (GIOP) formulation (Werdell et al., 2013) provided by the NASA ocean color database (oceancolor.gsfc.nasa.gov): phytoplankton spectral absorption coefficient (a_{ph}), absorption coefficient from dissolved organic matter and non-algal particles (a_{dg}) and the particulate backscattering coefficient (b_{bp}). MODIS daily GIOP data from 443 nm from daily scenes for the period 2003-2015 were converted to monthly composites with the same methodology used for the Rrs data (see Methods). GIOP values corresponding to the location of each of the 8 groups were

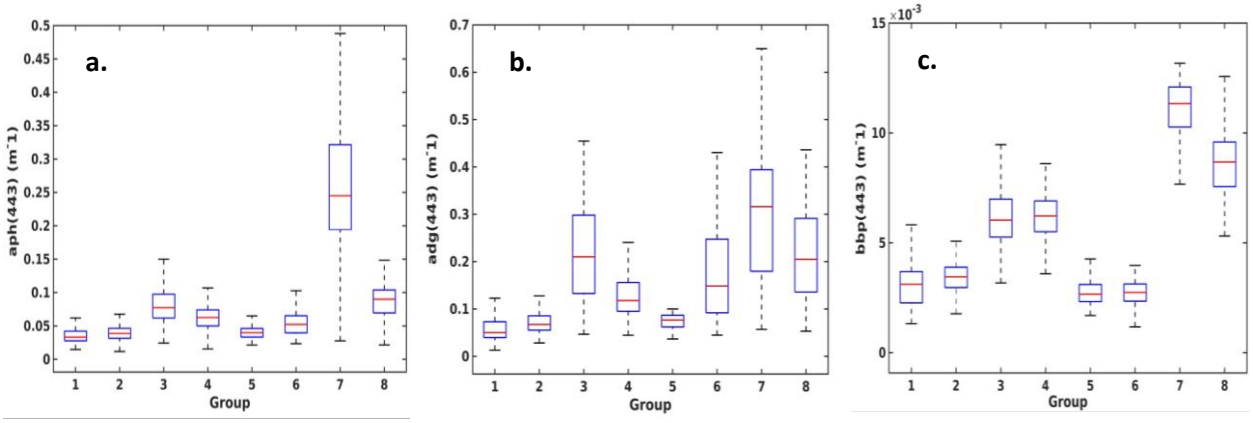

Fig. 12: MODIS GIOP data from 2003-2015 for each group time/location. Median values are indicated by red line with upper and lower bounds of boxes indicating the 75th and 25th percentiles respectively. Box whiskers include extreme data points not considered to be outliers. a.) Phytoplankton absorption for each group. b.) Dissolved and particulate matter absorption for each group. c.) Particulate backscattering for each group.

extracted from each monthly composite and their median values and associated statistics over the study period were calculated for each of the 3 IOP variables (Fig. 12).

IOP results for each group are consistent with the general description given to each group based on their Rrs spectral characteristics. IOPs indicate that groups 1, 2 and 5 have the lowest values for a_{ph}, a_{dg} and b_{bp}, indicative of the clearest and least biologically productive waters in the study area. Groups 7 and 8 have the highest values for each IOP, indicative of the largest relative load of backscattering particulate matter, and highest dissolved absorption and phytoplankton concentrations. Of these, group 7 has significantly elevated phytoplankton influence and particulate scattering (a_{ph} and b_{bp}). Groups 3, 4 and 6 are intermediate between these two extremes, with increased particulate load and a stronger influence of phytoplankton and non-particulate absorption than groups 1, 2 and 5, but less particulate load and weaker absorption than groups 7 and 8.

These descriptions are also consistent with expected surface water bio-optical characteristics based on their geographic position, the overall circulation patterns within the study area and seasonal cycles in stratification and bloom occurrence. Groups 5 and 6 have relatively similar spectral and IOP characteristics and dominate the study area both spatially and temporally. Of these, group 6 has slightly higher phytoplankton and dissolved absorption, dominates during the spring and fall bloom periods (Fig. 5), is associated with the western Maine coastal region throughout the summer and maintains dominance over the Scotian Shelf throughout winter (Fig. 4). Group 5 dominates during winter and summer. Group 1 is present primarily along the shelf break on the southern edge of the study area and is present inside the GOM only in mid-summer (Fig. 4), disappearing entirely during the winter (Fig. 5). These patterns are consistent with relatively oligotrophic surface water influenced by the Gulf Stream intruding onto the study area's shelf (e.g. Gawarkiewicz et al., 2012) and strongly stratified, nutrient-poor conditions over the deeper central basins of the GOM in mid-summer (e.g. Townsend, 1991; Thomas et al., 2003). Group 5 appears to represent an optical intermediate between groups 1 and 6.

Groups 3 and 6 both have seasonal distribution patterns (Fig. 5) that match the timing of spring and fall phytoplankton blooms (e.g. Thomas et al., 2003) and as such are likely strongly driven by the surface nutrient regime, light and stratification. The timing of their respective climatological seasonal peaks (Fig. 5) suggests that group 3 isolates bio-optical characteristics most prevalent during early bloom conditions and group 6 best represents peak bloom conditions. Group 2 also represents relatively clear water and occurs at the boundaries of other groups, especially between group 1 and 5, and groups 4 and 5, over the Scotian Shelf, Nantucket Shoals and Georges Bank regions (Fig. 4), primarily in summer and secondarily in winter (Fig. 5).

In contrast, groups 7 and 8 are present in times and locations consistent with water having the highest spectral Rrs values and strongest absorption and backscatter. These are located along shallow coastal zones, especially along the Maine coast, Bay of Fundy and southwestern Nova Scotia coast, and over Nantucket Shoals and Georges Bank: regions of strongest tidal mixing, strongest river influence, highest potential for suspended sediment load and elevated nutrient concentrations supporting higher chlorophyll concentrations (e.g. Townsend et al., 1987). These groups are strongest in winter and weakest in summer, consistent with increased winter wind mixing and reduced vertical stratification over shallow coastal waters. With its exceptionally strong tidal currents, it is no surprise that group 7, the most turbid water in the GOM, characterizes a significant proportion of the waters within the upper reaches of the Bay of Fundy year-round. Groups 3 and 4 have intermediate bio-optical properties and occur primarily at the boundary between these more turbid waters and groups 5 and 6. Along the Maine coast, these groups, together with Groups 7 and 8, follow the path of the well-mixed, nutrient-rich EMCC (Townsend et al., 2006). Groups 3, 4 and 8 dominate waters on top of Georges Bank, a shallow region of strong tidal mixing, increased vertical nutrient flux, elevated phytoplankton concentrations and the potential for sediment resuspension (O'Reilly et al., 1987; Townsend et al., 2006; Hu et al., 2008).

4.2. Climatological seasonal and regional stability

The study area is a highly variable shelf region that experiences strong changes in bio-optical properties in both time and space (Fig. 6). Interannual variability is strongest during late fall/winter months (November-February). This is especially true over shallow regions such as Georges Bank and Nantucket Shoals, where variability in winter conditions from year to year likely create strongly differing conditions of suspended sediment, phytoplankton concentrations and phytoplankton community structure. Another possible contributor to increased winter group interannual variability is the increased cloud cover that reduced valid satellite retrievals in winter months. With fewer days of data available, the monthly averaged composite images are more strongly influenced by daily variability leading to different group classification in any one year and stronger interannual variability in data point classification. Conversely, April experiences the highest stability in group classification over the study period. April is the climatological month of the annual spring bloom over the study region (Thomas et al., 2003) and the data show that the bio-optical conditions during this bloom are highly stable from year to year. There is a second region-wide increase in interannual stability during the period of the fall bloom, centered on September, again suggesting recurring similar bio-optical conditions over most years.

Interannually, the most consistently unstable region over the seasons in the study area is Georges Bank (Fig. 6). This region is known to be highly dynamic, influenced by tidal mixing, wind-driven mixing, strong gradients in nutrient concentration, phytoplankton concentration, community structure and frontal zones (Gettings et al., 2014; Hu et al. 2008; Townsend and Pettigrew, 1996; Townsend et al., 2006). These combine to keep the optical properties of this shallow bank highly dynamic. The only month when group classification of data points on Georges Bank is reasonably consistent is in April, when the spring bloom completely covers this area. In contrast, the upper Bay of Fundy is perhaps the most stable region in the study area. Due to intense tidal mixing, sediments are continuously mixed into

the water column. With one group best representing waters dominated by suspended sediment, this leaves little variability in data point classification for this region. In summer months when this tidal mixing is no longer augmented as strongly by wind-driven mixing, the variability of group classification increases slightly as the suspended sediment load likely decreases and other bio-optical factors become more important.

An oceanographic feature that the interannual group stability (Fig. 6) effectively isolates is the position of the dominant frontal zones of the region. These appear as localized regions of elevated group variability from year to year due to slight variations in their exact position and/or bio-optical characteristics. Prominent in Fig. 6 are the shelf break front along the southern portion of the study area south of Cape Cod and Georges Bank that is evident in all months except April and May and fronts around Georges Bank. Also prominent are winter frontal zones separating shallow, well mixed colder water from warmer deeper water along the coast of Maine, frontal zones at the seaward edge and western terminus of the EMCC in spring, summer and fall, and the frontal structure over the southern Nova Scotia shelf. The position and seasonality of these fronts is strongly similar to those shown by Ullmann and Cornillon (1999) who isolated these fronts in 12 years of satellite SST data. These authors show that because their position is strongly controlled by bathymetry, they reoccur in the same position from year to year. Not surprisingly, these SST frontal zones are also localized regions of increased variability in bio-optical characteristics.

4.3. Interannual variability and comparisons to environmental variables

The data show that 2 of the clearest water groups (1 and 2) have a significant decline in coverage over the study period (Fig. 9). These trends are dominated by decreases during the summer months (Fig. 11) and are coincident with significant increases in groups 3 and 6. Both groups 3 and 6 represent waters with increased phytoplankton and CDOM influence whose seasonal timing suggests they peak in coverage during the spring and fall blooms. These trend patterns suggest an increase in the time and/or

space window of phytoplankton and/or absorption dominated conditions over the study period. They are consistent with the increasing trend in summer CDOM concentration along a transect across the GOM over the past decade reported by Balch et al. (2012). These authors suggest this CDOM increase is due to climate-related increases in precipitation and river runoff into the GOM, resulting in increased surface dissolved absorption. Trends were not evident in April group coverage but were evident in the months on either side of the April spring bloom (Fig. 10). In both March and May, groups 2 and 5 decline over the time series, while group 6 increases. These trends suggest shifts in the timing of bloom conditions and likely reflect changes in the bio-optical properties associated with the strong phenology shifts in SST shown by Thomas et al. (2017) that were strongest in spring, summer and fall. The combination of group distribution patterns from May-July suggest that group 5 represents a transition bio-optical type between the bloom conditions represented by group 6 and the most oligotrophic conditions represented by group 1. The increase in group 6 in May (Fig. 10) appears to delay the insurgence of group 5, indicated by the increasing June coverage of group 5 over the time series. This in turn affects the timing and total areal coverage of group 1 in the summer, with a decrease in group 1 presence in the central GOM. Summer group distribution patterns suggest that the clearest waters represented by group 1 are covering gradually less area over time and are being replaced by other groups, mainly groups 3, 5 or 6.

A comparison of group interannual variability to concurrent metrics representing sea surface temperature anomalies (SSTa), GOM river discharge and wind mixing, provides insight into the extent to which interannual variability in the coverage of the bio-optical groups is related to easily-tracked environmental forcing. Correlations were calculated between the interannual variability in group percent coverage in individual months and the concurrent environmental metric monthly mean. A Spearman rank correlation was used to avoid potential biases caused by the unknown underlying probability distributions of group sizes and to decrease the importance to strong outliers.

Groups 7 and 8 represent the most turbid and strongest absorbing waters in the study area and are closely associated with the coast especially along the Maine coast and Bay of Fundy. Regional river discharge is strongly seasonal, with minima in winter and summer, and a strong maximum in April during the spring freshet when discharge is 5-10 times that of mid-summer (Fig.12). An examination of the extent to which interannual variability in these groups was related to river discharge showed that March variability in the percent coverage of group 8 was positively correlated with March discharge, and April variability of group 7 was positively correlated with April discharge. Group 3, which is also often closely associated with these coastal groups (Fig. 4) was also positively correlated to river discharge, in both March and April, as well as in January and February. There was no correlation of these groups to summer or fall river discharge. However, group 6, which in July is usually restricted to a narrow coastal band along the western GOM shore (Fig. 4), had a significant positive correlation of July percent coverage to July river discharge. Together, these correlations show that the satellite data and multivariate grouping methodology is effective in capturing variability in these coastal bio-optical water groups and the area coverage of these groups in the study region is linked to river discharge, especially in spring, expanding in years of stronger river discharge.

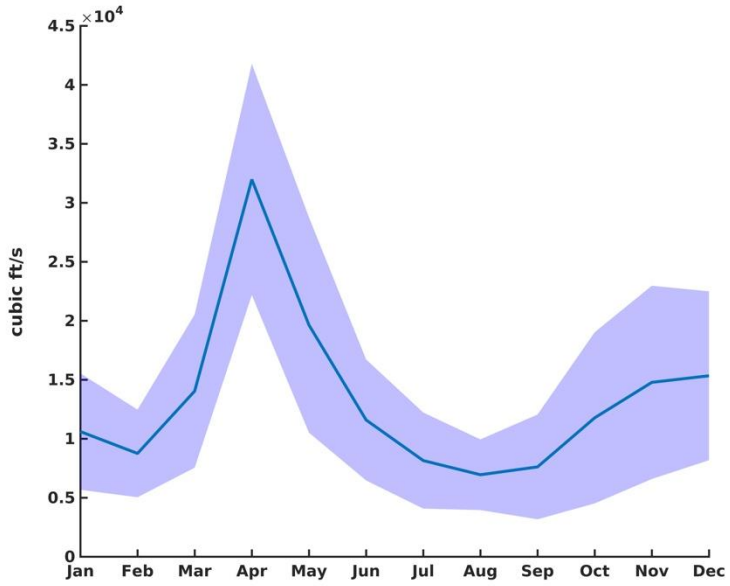

Fig. 13: Climatological monthly Penobscot River discharge 1998-2016 (±1 s.d.). Calculated from daily mean USGS Water Resources data collected in West Enfield. ME.

Wind-driven mixing impacts vertical stability. In shallow regions, wind mixing can play a role in increased turbidity and over deeper basins can play a role in the vertical flux of nutrients and light availability. Climatological wind speeds over the study area are maximum in late fall/winter (November-February) and minimum in summer (Fig. 14). Comparisons of the interannual variability in group percent coverage to monthly mean wind speed over the study period showed that both of the most turbid coastal groups (7 and 8) had correlations to wind speed. The percent coverage of group 7 was positively correlated to wind speed in January, March, July and October. Group 8 was positively correlated to wind speed in October and December. A second pattern of interaction was evident over deeper basins in the spring. In the study region, the April spring bloom is climatologically dominated by group 6 (Fig. 4). However, significant negative correlations between April wind speeds and April coverage of groups 4 and 5 suggest that years of decreased wind mixing in spring are associated with increases in the percent coverage of these two other groups.

We examined correlations between monthly SST anomalies calculated for the GOM region and monthly group size over the 19-year study period. The strong climate-induced increasing trend in SST

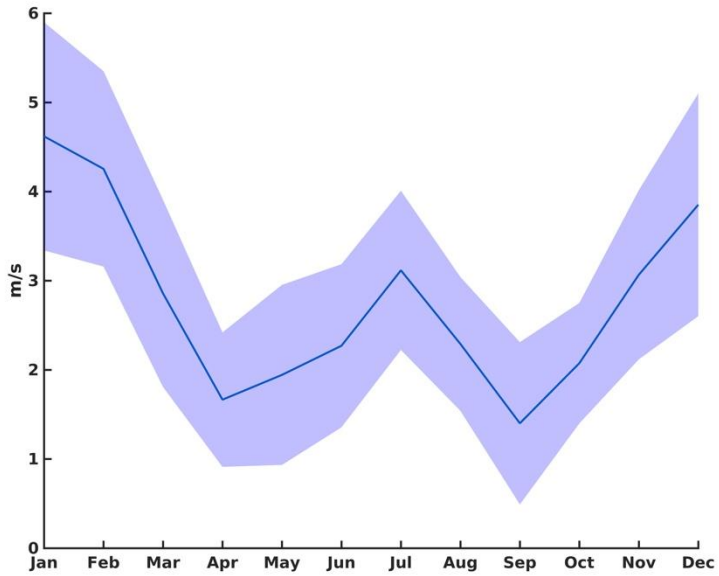

Fig. 14: Average wind speed from grid locations in the central GOM (±1 s.d.). Monthly NCEP reanalysis data from the NOAA Earth System Research Lab.

(Pershing et al., 2015; Thomas et al., 2017) was removed from the SST anomalies prior to correlation calculations. No systematic correlations were evident between the interannual variability of group percent coverage and the monthly SST anomalies over the study period. It is interesting to note, however, that the large marine heat wave that impacted the region in 2012 (Pershing et al., 2015; Scannell et al., 2016) was coincident with annual changes in the coverage of many bio-optical groups (Fig. 9), with groups 3, 4 and 8 all having a minima in their coverage and group 5 a maxima.

CHAPTER 5

SUMMARY AND CONCLUSIONS

Nineteen years (1998-2016) of SeaWiFS and MODIS multispectral Rrs data were used to quantify seasonal and interannual variability in the surface optical properties of the US northeastern shelf. Data were grouped over time and space based on similarity of their Rrs spectra using a two-step multivariate clustering technique resulting in the 8 most dominant bio-optical water types. These 8 groups were broadly identified as 3 groups representing the clearest waters in the study area, 2 groups with the highest Rrs values across their spectra indicative of increased turbidity and particle load, and a series of groups with intermediate spectral shapes representing varying concentrations of phytoplankton and CDOM. These descriptions were consistent with the geographic and bathymetric group locations and also with IOP values calculated for the same groups.

Seven of the 8 dominant groups had strong seasonal cycles in their distribution patterns indicating the groups isolated here effectively track the seasonal evolution of bio-optical conditions in the study region. Two of the groups, representing the clearest and least productive waters of the study window, exhibited a significant decline in annual coverage over the time series that parallel the strong warming trend of the study region and increases in CDOM concentrations reported in the literature. Examination of trends in individual months over the 19-year time series sheds light on seasonally specific shifts. Trends were not evident in winter months (December-February), likely due to strong mixing, low temperatures and low light conditions that create similar bio-optical conditions each winter. Multi-year trends were also not evident in April, the peak of the spring bloom, again likely due to recurrent bio-optical conditions each year. However, months on either side of the climatological spring bloom (March and May) exhibited a decrease in groups characterizing clearer waters and an increase in the groups associated with the spring bloom. These patterns are indicative of bloom bio-optical phenology shifts that parallel spring SST phenology shifts reported in the literature. The monthly data also show that the

multi-year decreasing trend in clearest waters was primarily due to trends in summer months, consistent with when the strongest trends in SST are occurring.

The analysis approach and results offer a highly objective view of satellite-measured ocean color variability within the study area. Results based on the actual Rrs spectral signatures of each data point avoid assumptions and biases made by in-water algorithms that use band ratios or invert poorly constrained spectral shapes in attempts to estimate specific oceanographic characteristics such as chlorophyll concentration or absorption. In such optically complex waters, such approaches are best applied using locally-tuned coefficients assisted by in situ data as ground truth. Biases in such approaches can also lead to discontinuities in derived time series over multiple missions. Differences in ocean color presented here are based on the most direct satellite measurements available, removing these additional sources of uncertainty. The study area is known to be a region where these biases can be strong. Allowing the clustering algorithms to identify the most similar optical characteristics and grouping data points based on their quantifiable statistical similarity avoids biases in more subjective group identification and pattern interpretation. The actual number of groups to retain and describe the data remains subjective and as with any modeling approach, needs to balance the number of variables used in the model with ease of interpretation. The results reported here focus on simplifying the multi-spectral data to the dominant 8 groups evident over the time/space window of the study and an analysis of their seasonal and interannual variability. However, the initial SOM analysis resulted in 36 groups, offering a much richer view of ocean color variability over the study window. Variability within these 36 groups has the potential to be useful in other studies more focused on specific environmental processes or more detailed analyses of sub-regions than the broad overview of variability presented here. Such applications await further analysis.

Results presented here offer insight into some of the ocean color trends and changes that are occurring in the study area concurrent with the well documented rapid warming of the region (e.g.

Pershing et al., 2015; Saba et al., 2015; Thomas et al., 2017). Only a longer time series can determine if the trends documented here are the new norm for the region or if they are unresolved low frequency interannual variability. However, the trends in group coverage and their seasonal timing imply changes in the bio-optical characteristics of the surface water and changes in the amount, timing and/or composition of phytoplankton. Each of these changes has implications for the trophic functioning of this highly productive region and likely will have socio-economic implications in a region strongly dependent on marine resources.

REFERENCES

Ainsworth, E.J., 1999. Visualization of ocean colour and temperature from multi-spectral imagery captured by the Japanese ADEOS satellite. *Journal of Visualization*, 2(2), 195-204.

Alvain, S., C. Moulina, Y. Dandonneau, and F.M., Breon, 2005. Remote sensing of phytoplankton groups in case 1 waters from global SeaWiFS imagery. *Deep-Sea Research*, 52, 1989-2004.

Ardyna, M., H. Claustre, J.B. Sallée, F. D'Ovidio, B. Gentili et al., 2017. Delineating environmental control of phytoplankton biomass and phenology in the Southern Ocean. *Geophysical Research Letters,* 44. dio:10.1002/2016GL072428.

Arnone, R.A. and R.W. Gould, 1998. Coastal monitoring using ocean color. *Sea Technology*, 39(9), 18-27.

Babin, S.M., J.A. Carton, T.D. Dickey, and J.D. Wiggert, 2004. Satellite evidence of hurricane-induced phytoplankton blooms in an oceanic desert. *Journal of Geophysical Research*, 109, C03043. dio:10.1029/2003JC001938.

Balch, W.M., D.T. Drapeau, B.C. Bowler, and T.G. Huntington, 2012. Step-changes in the physical, chemical and biological characteristics of the Gulf of Maine, as documented by the GNATS time series. *Marine Ecology Progress Series*, 450, 11-35. doi: 10.3354/meps09555.

Balch, W.M., D.T. Drapeau, B.C. Bowler, E.S. Booth, J.I. Goes et al., 2004. A multi-year record of hydrographic and bio-optical properties of the Gulf of Maine: I. Spatial and temporal variability. *Progress in Oceanography*, 63, 57-98. doi: 10.1016/j.pocean.2004.09.003.

Banzon, V.R., R.W., and National Center for Atmospheric Research Staff (eds.), Modified 21 Apr 2017. *The Climate Data Guide: SST data: NOAA High-resolution (0.25 x 0.25) Blended Analysis of Daily SST and Ice, OISSTv2*. Retrieved from: http://climatedataguide.ucar.edu/climate-data/sst-data-noaa-high-resolution-0.25x0.25-blended-analysis-daily-sst-and-ice-oisstv2.

Behrenfeld, M.J., R.T. O'Malley, D.A. Siegel, C.R. McClain, J.L. Sarmiento et al., 2006. Climate-driven trends in contemporary ocean productivity. *Nature Letters,* 444, 752-755. doi:10.1038/nature05317.

Brewin, R.J.W., S.J. Lavender, N.J. Hardman-Mountford, T. Hirata, 2010. A spectral response approach for detecting dominant phytoplankton size class from satellite remote sensing. *Acta Oceanologica Sinica*, 29(2), 14-32. doi: 10.1007/s13131-010-0018-y.

Brooks, D.A., 1985. Vernal circulation in the Gulf of Maine. *Journal of Geophysical Research*, 90(C3), 4687-4705.

Chen, C.H. and B. Shrestha, 2000. Classification of multi-sensor remote sensing images using Self-Organizing Maps and Radial Basis Function Networks*. International Geoscience and Remote Sensing Symposium*, 2, 711-713.

Dall'Olmo, G., A.A. Gitelson, D.C. Rundquist, B. Leavitt, T. Barrow et al., 2005. Assessing the potential of SeaWiFS and MODIS for estimating chlorophyll concentration in turbid productive waters using red and near-infrared bands. *Remote Sensing of Environment*, 96, 176-187. doi: 10.1016/j.rse.2005.02.007.

Darecki, M. and D. Stramski, 2004. An evaluation of MODIS and SeaWiFS bio-optical algorithms in the Baltic Sea. *Remote Sensing of Environment*, 89, 326-350. doi: 10.1016/j.rse.2003.10.012.

Devred, E., S. Sathyendranath, and T. Platt, 2007. Delineation of ecological provinces using ocean colour radiometry. *Marine Ecology Progress Series*, 346, 1-13. doi: 10.3354/meps07149.

Fiedler, P.C., 1984. Satellite observations of the 1982-83 El Nino along the U.S. Pacific coast. *Science*, 224, 1251-1254.

Fournier, R.O., M. Van Det, J.S. Wilson, and N.B. Hargreaves, 1979. The influence of the shelfbreak front off Nova Scotia on phytoplankton standing stock in late winter. *Journal of the Fisheries Research Board of Canada*, 36, 1228-1237.

Garcia, C.A.E., V.M.T Garcia, and C.R. McClain, 2005. Evaluation of SeaWiFS chlorophyll algorithms in the Southwestern Atlantic and Southern Oceans. *Remote Sensing of Environment*, 95, 125-137. doi: 10.1016/j.rse.2004.12.006.

Garver, S.A. and D.A. Siegel, 1997. Inherent optical property inversion of ocean color spectra and its biogeochemical interpretation: 1. Time series from the Sargasso Sea. *Journal of Geophysical Research*, 102(C8), 18,607-18,625.

Gawarkiewicz, G.G., R.E. Todd, A.J. Plueddemann, M. Andres, and J.P. Manning, 2012. Direct interaction between the Gulf Steam and the shelfbreak south of New England. *Scientific Reports*, 2(553). doi: 10.1038/srep00553.

Gawarkiewicz, G.G., T.G. Ferdelman, T.M. Church, and G.W. Luther III, 1996. Shelfbreak frontal structure on the continental shelf north of Cape Hatteras. *Continental Shelf Research*, 16(14), 1751-1773.

Gettings, R.M., D.W. Townsend, M.A. Thomas, and L. Karp-Boss, 2014. Dynamics of late spring and summer phytoplankton communities on Georges Bank, with emphasis on diatoms, *Alexandrium* spp., and other dinoflagellates. *Deep-Sea Research II*, 103, 120-138. doi: 10.1016/j.dsr2.2013.05.012.

Goes, J.I., H.R. Gomes, A. Limsakul, and T. Saino, 2004. The influence of large-scale environmental changes on carbon export in the North Pacific Ocean using satellite and shipboard data. *Deep-Sea Research II*, 51, 247-279. doi: 10.1016/j.dsr2.2003.06.004.

Hara, Y., R.G Atkins, S.H Yueh, R.T. Shin, and J.A. Kong, 1994. Application of neural networks to radar image classification. *IEEE Transactions on Geophysics and Remote Sensing*, 32, 100-111.

Hargrave B.T., N.J. Prouse, G.A. Phillips, and P.A. Neame, 1983. Primary production and respiration in pelagic and benthic communities at tow intertidal sites in the upper Bay of Fundy. *Can. J. Fish. Aq. Sci*., 40, 229-243.

Hasegawa, D., J. Sheng, D.A. Greenberg, and K.R. Thompson, 2011. Far-field effects of tidal energy extraction in the Minas Passage on tidal circulation in the Bay of Fundy and Gulf of Maine using a nested-grid coastal circulation model. *Ocean Dynamics*, 61(11), 1845-1868. doi: 10.1007/s10236-011-0481-9.

He, X., Y. Bai, D. Pan, N. Huang, X. Dong et al., 2013. Using geostationary satellite ocean color data to map the diurnal dynamics of suspended particulate matter in coastal waters. *Remote Sensing of Environment*, 133, 225-239. doi: 10.1016/j.rse.2013.01.023.

Hopkins, J., M. Lucas, C. Dufau, M. Sutton, J. Stum et al., 2013. Detection and variability of the Congo River plume from satellite derived sea surface temperature, salinity, ocean colour and sea level. *Remote Sensing of Environment,* 139, 365-385. doi: 10.1016/j.rse.2013.08.015.

Hu, S., D.W. Townsend, C. Chen, G. Cowles, R.C. Beardsley et al., 2008. Tidal pumping and nutrient fluxes on Georges Bank: A process-oriented modeling study. *Journal of Marine Systems*, 74, 528-544. doi: 10.1016/j.marsys.2008.04.007.

IOCCG, 2000. Remote Sensing of Ocean Colour in Coastal, and Other Optically-Complex Waters. Sathyendrandath, S. (ed.), *Reports of the International Ocean-Colour Coordinating Group*, No. 3, IOCCG, Dartmouth, Canada.

Irwin, A.J. and M.J. Oliver, 2009. Are ocean deserts getting larger? *Geophysical Research Letters,* 36, L18609. doi: 10.1029/2009GL039883.

Iskandar, I.T., T. Tozuka, Y. Masumoto, and T. Yamagata, 2008. Impact of Indian Ocean dipole on intraseasonal zonal currents at 90E on the equator as revealed by self-organizing map. *Geophysical Research Letters*, 35, L14S03. doi: 10.1029/2008GL033468.

Ji, R., C.S. Davis, C. Chen, D.W. Townsend, D.G. Mountain et al., 2007. Influence of ocean freshening on shelf phytoplankton dynamics. *Geophysical Research Letters,* 34, L24607. dio:10.2029/2007GL032010.

Kahru, M., E. DiLorenzo, M. Manzano-Sarabia, and B.G. Mitchell, 2004. Ocean-color variability in the Gulf of California: scales from days to ENSO. *Deep Sea Research*, 51, 139-146.

Kaski, S. and T. Kohonen, 1996. Exploratory data analysis by the Self-organizing Map: Structures of welfare and poverty in the world. In *Neural Networks in Financial Engineering*, A.P.N. Refenes, Y.Abu-Mostafa, J. Moody, A. Weigend (eds.). World Singapore: Scientific, 498-507.

Kavanaugh, M.T., B. Hales, M. Saraceno, Y.H. Spitz, A.E. White et al., 2014. Hierarchical and dynamic seascapes: A quantitative framework for scaling pelagic biogeochemistry and ecology. *Progress in Oceanography*, 120, 291-304. doi: 10.1016/j.pocean.2013.10.013.

Kohonen, T., 2001. *Self-Organizing Maps (3 ed.)*. Springer-Verlag, ISBN 978-3-540-67921-9, New York, Berlin, Heidelberg.

Kohonen, T., 2013. Essentials of the self-organizing map. *Neural Networks*, 37, 52-65. doi: 10.1016/j.neunet.2012.09.018.

Lee, Z.P., K.L. Carder, and R.A. Arnone, 2002. Deriving inherent optical properties from ocean color: a multiband quasi-analytical algorithm for optically deep waters. *Applied Optics*, 41, 5755-5772.

Liu, Y. and R.H. Weisberg, 2005. Patterns of ocean current variability on the West Florida Shelf using the self-organizing map. *Journal of Geophysical Research*, 110, C06003. doi: 10.1029/2004JC002786.

Loder, J.W. and D.A. Greenberg, 1986. Predicted positions of tidal fronts in the Gulf of Maine region. *Continental Shelf Research*, 6(3), 397-414.

Lohrenz, S.E., X. Chen, M. Tuel, and W.-J. Cai, 2009. Hyperspectral remote sensing of water mass properties in a river-influenced coastal region. *IEEE, 2009*. doi: 10.23919/OCEANS.2009.5422387.

Longhurst, A.R., 1998. *Ecological Geography of the Sea*, first ed. Academic Press, San Diego, 398 pages.

Longhurst, A.R., 2006. Ecological geography of the sea. Academic Press, 560 pages.

McClain, C.R., 2009. A decade of satellite ocean color observations. *Annual Review of Marine Science,* 1, 19-42. doi: 10.1146/annurev.marine.010908.163650.

Mélin, F. and V. Vantrepotte, 2015. How optically diverse is the coastal ocean? *Remote Sensing of Environment,* 160, 235-251. doi: 10.1016/j.rse.2015.01.023.

Mobley, C.D., 1994. *Light and Water: radiative transfer in natural water*. Academic Press. 592 pages.

Molleri, G.S.F., E.N.L.M. Novo, and M. Kampel, 2010. Space-time variability of the Amazon River plume based on satellite ocean color. *Continental Shelf Research*, 30, 342-352. doi: 10.1016/j.csr.2009.11.015.

Moore, J.K. and M.R. Abbott, 2000. Phytoplankton chlorophyll distributions and primary production in the Southern Ocean. *Journal of Geophysical Research*, 105, C12, 28,709-28,722.

Nerbonne, J. and W. Heeringa, 2001. Computational comparison and classification of dialects. *Dialectologia et Geolinguistica*, 9, 69-83.

Oliver, M.J. and A.J. Irwin, 2008. Objective global ocean biogeographic provinces. *Geophysical Research Letters,* 35, L15601. doi: 10.1029/2008GL034238.

O'Reilly, J.E., C. Evans-Zeltin, and D.A. Busch, 1987. Primary production. Pp. 220-233. In: Backus, R.H. (ed.). 1987 Georges Bank. MIT Press, Cambridge, Mass. 593 pp.

Pershing, A.J., M.A. Alexander, C.M. Hernandes, L.A. Kerr, A. LeBris et al., 2015. Slow adaptation in the face of rapid warming leads to collapse of Gulf of Maine cod fishery. *Science*. doi: 10.1126/science.aac9819.

Pettigrew, N.R., 2005. The kinematic and hydrographic structure of the Gulf of Maine Coastal Current. *Deep-Sea Research II*, 52, 2369-2391. doi: 10.1016/j.dsr2.2005.06.033.

Pettigrew, N.R., D.W. Townsend, H. Xue, J.P. Wallinga, P.J. Brickley et al., 1998. Observations of the Eastern Maine Coastal Current and its offshore extensions in 1994. *Journal of Geophysical Research*, 103(C13), 30,623-30,639.

Platt, T., G.N. White III, L. Zhai, S. Sathyendranath, and S. Roy, 2009. The phenology of phytoplankton blooms: Ecosystem indicators form remote sensing. *Ecological Modelling,* 220, 3057-3069. doi:10.1016/j.ecolmodel.2008.11.022.

Polovina, J.J., E.A. Howell, and M. Abecassis, 2008. Ocean's least productive waters are expanding. *Geophysical Research Letters*, 35, L03618.

Reygondeau, G., A. Longhurst, E. Martinez, G. Beaugrand, D. Antoine et al., 2013. Dynamic biogeochemical provinces in the global ocean. *Global Biogeochemical Cycles*, 27, 1046-1058. doi: 10.1002/gbc.20089.

Reynolds, R.W., T. Smith, C. Liu, D.B. Chelton, K.S. Casey et al., 2007. Daily high-resolution-blended analyses for sea surface temperature. *Journal of Climate*, 20, 5473-5496. doi: 10.1175/2007JCLI1824.1.

Richardson, A., C. Risien, and F.A. Shillington, 2003. Using self-organizing maps to identify patterns in satellite imagery. *Progress in Oceanography*, 59(2-3), 223-239.

Risien, C.M., C.J.C. Reason, F.A. Shillington, and D.B. Chelton, 2004. Variability in satellite winds over the Benguela upwelling system during 1999-2000. *Journal of Geophysical Research*, 109(C3, C03010). doi: 10.1029/2003JC001880.

Saba, V.S., S.M. Griffies, W.G. Anderson, M. Winton, M.A. Alexander et al., 2016. Enhanced warming of the Northwest Atlantic Ocean under climate change. *Journal of Geophysical Research: Oceans*, 121, 118-132. doi: 10.1002/2015JC011346.

Scannell, H.A., A.J. Pershing, M.A. Alexander, A.C. Thomas, and K.E. Mills, 2016. Frequency of marine heatwaves in the North Atlantic and North Pacific since 1950. *Geophysical Research Letters*. doi: 10.1002/2015GL067308.

Signorini, S.R., B.A. Franz, and C.R. McClain, 2015. Chlorophyll variability in the oligotrophic gyres: mechanisms, seasonality and trends. *Frontiers in Marine Science*. doi: 10.3389/fmars.2015.00001.

Smith, P.C., 1989. Seasonal and interannual variability of current, temperature and salinity off southwest Nova Scotia. *Can. J. Fish. Aquat. Sci.*, 46.

Song, H., R. Ji, C. Stock, and Z. Wang, 2010. Phenology of phytoplankton blooms in the Nova Scotian Shelf – Gulf of Maine region: remote sensing and modeling analysis. *Journal of Plankton Research*, 32(11), 1485-1499. doi:10.1093/plankt/fbq086.

Thomas, A.C., A.J. Pershing, K.D. Friedland, J.A. Nye, K.E. Mills et al., 2017. Seasonal trends and phenology shifts in sea surface temperature on the North American northeastern continental shelf. *Elementa Science of the Anthropocene*, 5(48). doi: 10.1525/elementa.240.

Thomas, A.C. and R.A. Weatherbee, 2006. Satellite-measured temporal variability of the Columbia River plume. *Remote Sensing of Environment,* 100, 167-178. doi: 10.1016/j.rse.2005.10.018.

Thomas, A.C., D.W. Townsend, and R. Weatherbee, 2003. Satellite-measured phytoplankton variability in the Gulf of Maine. *Continental Shelf Research*, 23, 971-989.

Thomas, A.C., P.T. Strub, and P. Brickley, 2003. Anomalous satellite-measured chlorophyll concentrations in the northern California Current in 2001-2002. *Geophysical Research Letters*, 30(15). doi: 10.1029/2003GL017409.

Thomas, A.C., R. Weatherbee, H. Xue, and G. Liu, 2010. Interannual variability of shellfish toxicity in the Gulf of Maine: Time and space patterns and links to environmental variability. *Harmful Algae*, 9, 458-480. doi: 10.1016/j.hal.2010.03.002.

Townsend, D.W., 1991. Influences of oceanographic processes on the biological productivity of the Gulf of Maine. *Reviews in Aquatic Sciences,* 5(3-4), 211-230.

Townsend, D.W. and A.C. Thomas, 2001. Winter-spring transition of phytoplankton chlorophyll and inorganic nutrients on Georges Bank. *Deep-Sea Research II*, 48, 199-214.

Townsend, D.W., A.C. Thomas, L.M. Mayer, and M.A. Thomas, 2006. Oceanography of the northwest Atlantic continental shelf. *The Sea: The Global Coastal Ocean: Interdisciplinary Regional Studies and Syntheses*. A.R. Robinson and K.H. Brink (eds). Cambridge, MA: Harvard University Press, 119-168.

Townsend, D.W. and M. Thomas, 2002. Springtime nutrient and phytoplankton dynamics on Georges Bank. *Marine Ecology Progress Series*, 228, 57-74.

Townsend, D.W. and N.R. Pettigrew, 1996. The role of frontal currents in larval fish transport on Georges Bank. *Deep-Sea Research II*, 43(7-8), 1773-1792.

Townsend, D.W., J.P. Christensen, D.K. Stevenson, J.J. Graham, and S.B. Chenoweth, 1987. The importance of a plume of tidally-mixed water to the biological oceanography of the Gulf of Maine. *Journal of Marine Research*, 45, 699-728.

Townsend, D.W., L.M. Cammen, P.M. Holligan, D.E. Cambell, and N.R. Pettigrew, 1994. Causes and consequences of variability in the timing of spring phytoplankton blooms. *Deep-Sea Research*, 41, 747-765.

Townsend, D.W., M.D. Keller, M.E. Sieracki, and S.G. Ackleson, 1992. Spring phytoplankton blooms in the absence of vertical water column stability. *Nature*, 360, 59-62.

Townsend, D.W., N.R. Pettigrew, M.A. Thomas, M.G. Neary, D.J. McGillicuddy et al., 2015. Water masses and nutrient sources to the Gulf of Maine. *Journal of Marine Research*, 73, 93-122.

Traykovski, L.V.M and H.M. Sosik, 2003. Feature-based classification of optical water types in the Northwest Atlantic based on satellite ocean color data. *Journal of Geophysical Research*, 108, C5. doi: 10.1029/2001JC001172.

Ullman, D.S. and P.C. Cornillon, 1999. Satellite-derived sea surface temperature fronts on the continental shelf off the northeaster U.S. coast. *Journal of Geophysical Research*, 104, 23,459-23,478.

Uitz, J., H. Claustre, A. Morel, and S.B. Hooker, 2006. Vertical distribution of phytoplankton communities in open ocean: An assessment based on surface chlorophyll. *Journal of Geophysical Research*, 111, C08005. doi: 10.1029/2005JC003207.

Vantrepotte, V. and F. Mélin, 2011. Inter-annual variations in SeaWiFS global chlorophyll *a* concentration (1997-2007). *Deep-Sea Research I*, 58(4), 429-441. doi: 10.1016/j.dsr.2011.02.003.

Vantrepotte, V., H. Loisel, D. Dessailly, and X. Mériaux, 2012. Optical classification of contrasted coastal waters. *Remote Sensing of Environment*, 123, 306-323. doi: 10.1016/j.rse.2012.03.004.

Vesanto, J. and E. Alhoniemi, 2000. Clustering of the Self-Organizing Map. *IEEE Transactions on Neural Networks*, 11(3).

Werdell, P.J., B.A. Franz, S.W. Bailey, G.C. Feldman, E. Boss, 2013. Generalized ocean color inversion model for retrieving marine inherent optical properties. *Applied Optics*, 52(10).

Werdell, P.J., L.I.W. McKinna, E. Boss, S.G. Ackleson, S.E. Craig et al., 2018. An overview of approaches and challenges for retrieving marine inherent optical properties from ocean color remote sensing. Progress in Oceaography, 160, 186-212. doi: 10.1016/j.pocean.2018.01.001.

Werdell, P.J., S.W. Bailey, B.A. Franz, L.W. Harding Jr., G.C. Feldman et al., 2009. Regional and seasonal variability of chlorophyll-a in Chesapeake Bay as observed by SeaWiFS and MODIS-Aqua. *Remote Sensing of Environment*, 113, 1319-1330. doi: 10.1016/j.rse.2009.02.012.

Xue, H., F. Chai, and N.R. Pettigrew, 2000. A model study of the seasonal circulation in the Gulf of Maine. *Journal of Physical Oceanography*, 30.

Ye, H., J. Li, T. Li, Q. Shen, J. Zhu et al., 2016. Spectral classification of the Yellow Sea and implications for coastal ocean color remote sensing. *Remote Sensing*, 8(321). doi: 10.3390/rs8040321.

Yoder, J.A. and M.A. Kennelly, 2006. What have we learned about ocean variability from satellite ocean color imagers? *Oceanography*, 19(1).

Yoder, J.A., S.E. Schollaert, and J.E. O'Reilly, 2002. Climatological phytoplankton chlorophyll and sea surface temperature patterns in continental shelf and slope waters off the northeast U.S. coast. *Limnology and Oceanography*, 47, 672-682.

Appendix A

SELF-ORGANIZING MAP RESULTS

Figures below show results from the Self-Organizing Map analysis, classifying data points into 36 nodes. 2-Dimensional node arrangement, Rrs spectra, climatological distribution and stability maps, and interannual variability graphs of the 36 nodes are included.

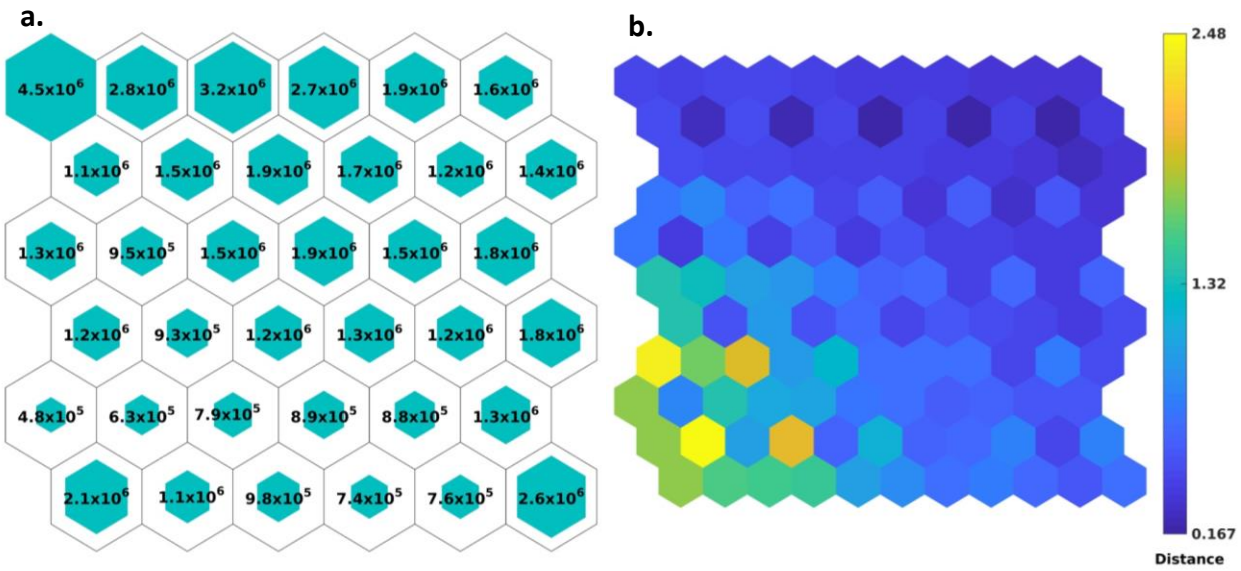

Fig. 15: SOM hits and u-matrix maps. a.) Hits map shows the relative (filled area) and actual number of data points classified into each of the 36 SOM nodes. b.) U-matrix is a measure of the Euclidean distance between neighboring nodes.

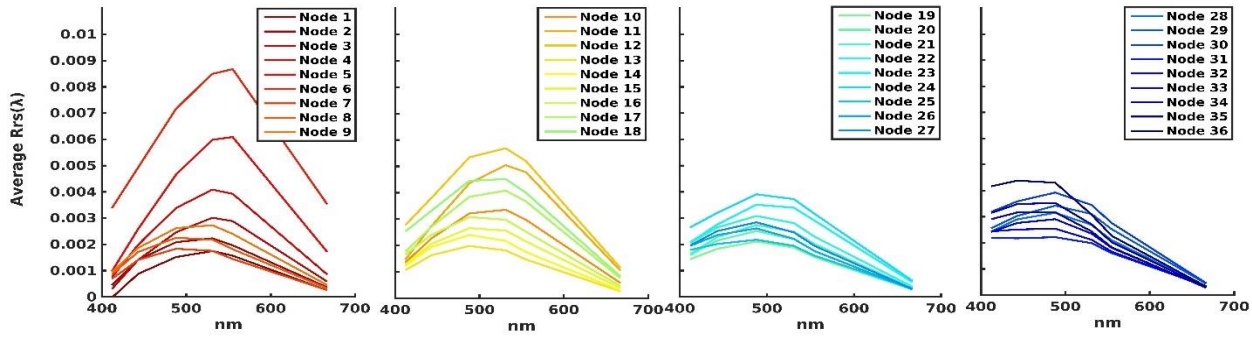

Fig. 16: Average spectral characteristics for data grouped by each of the 36 SOM nodes. Data points over the 19-year time series associated with each node were identified and their 6 spectral bands averaged. In general, spectra are arranged from more turbid waters in the left plot to clearest waters on the right.

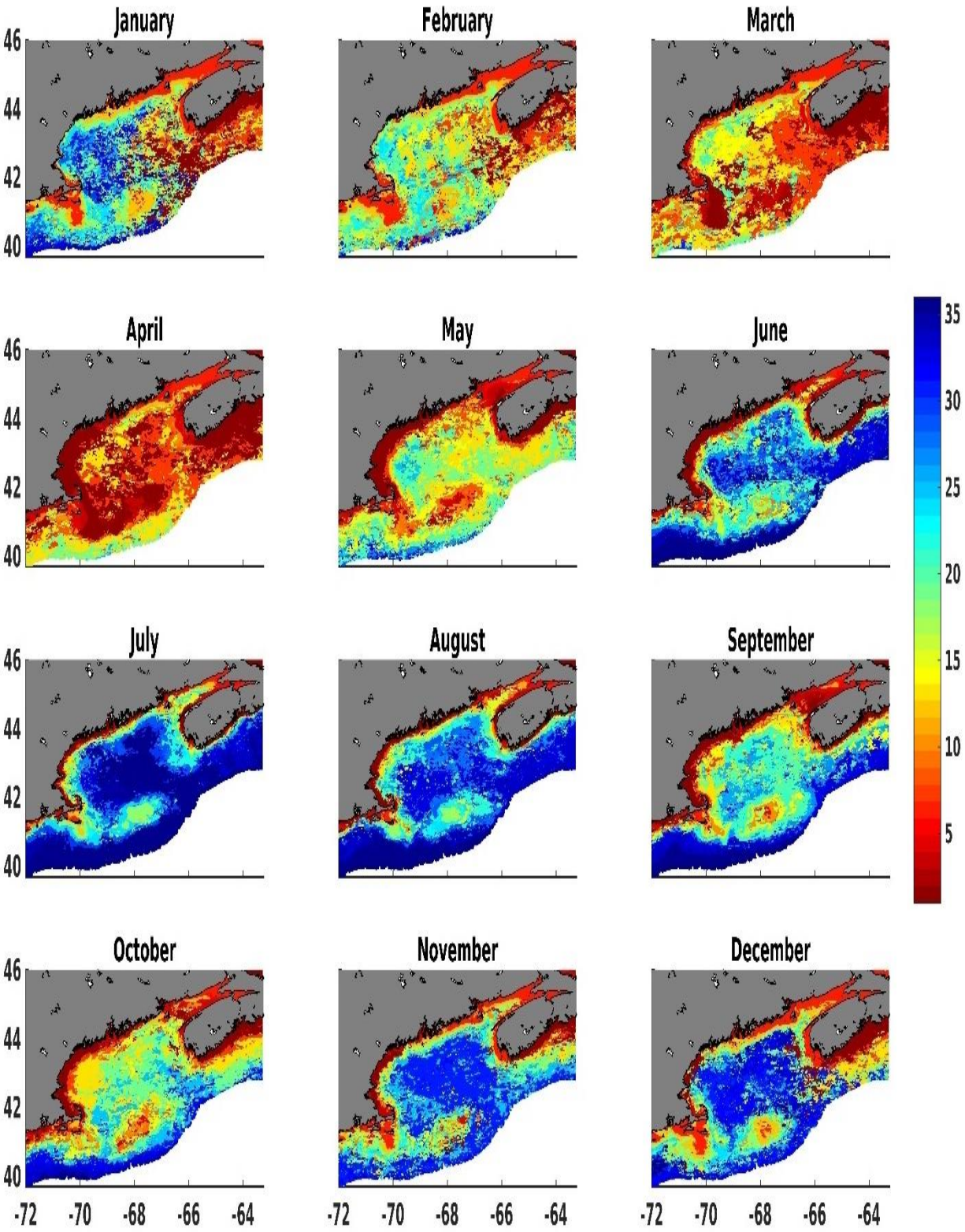

Fig. 17: Monthly climatology maps of the 36 nodes. At each data location, the climatological value is the most often occurring node (the mode) over the 19 years of data.

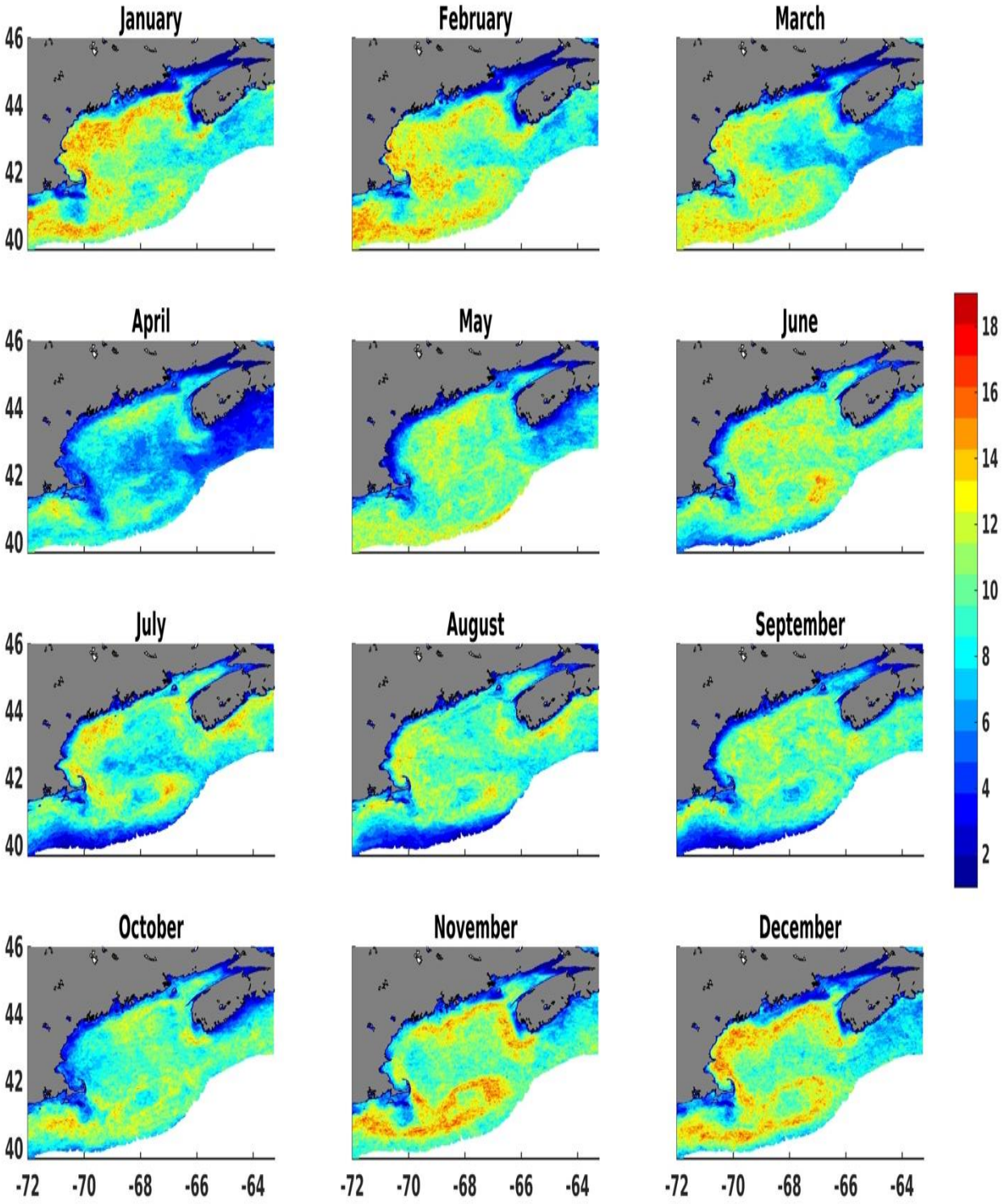

Fig. 18: Interannual stability at each location of the monthly climatological values over the 19 years for the 36 SOM nodes. Quantifies how frequently each location was classified as a different node over the 19 years. Regions of low interannual variability (blue) are most often classified as the same node across the 19 years. Regions of high interannual variability (red) are more frequently classified as a different node across the 19 years.

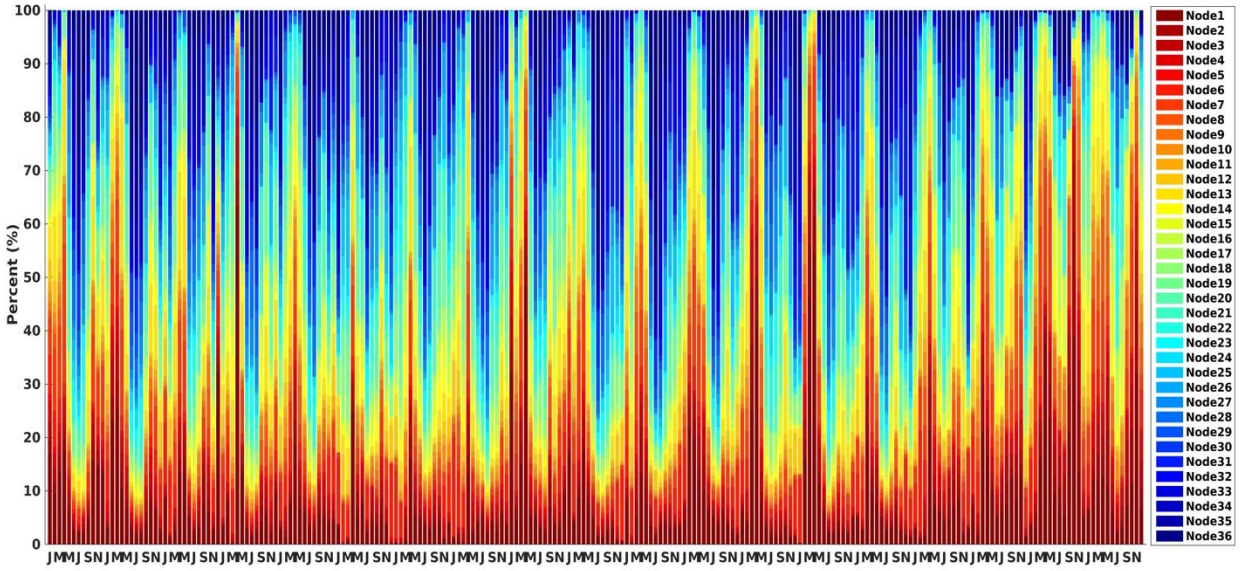

Fig. 19: Summary of interannual variability of the 36 SOM node coverage of the study area. Time series of the percent of the total study area that each node covers for all monthly composite images. Every other month indicated on the x-axis. Seasonal cycles in node distribution are evident in addition to interannual variability and study area-wide distribution shifts in the last 5-6 years.

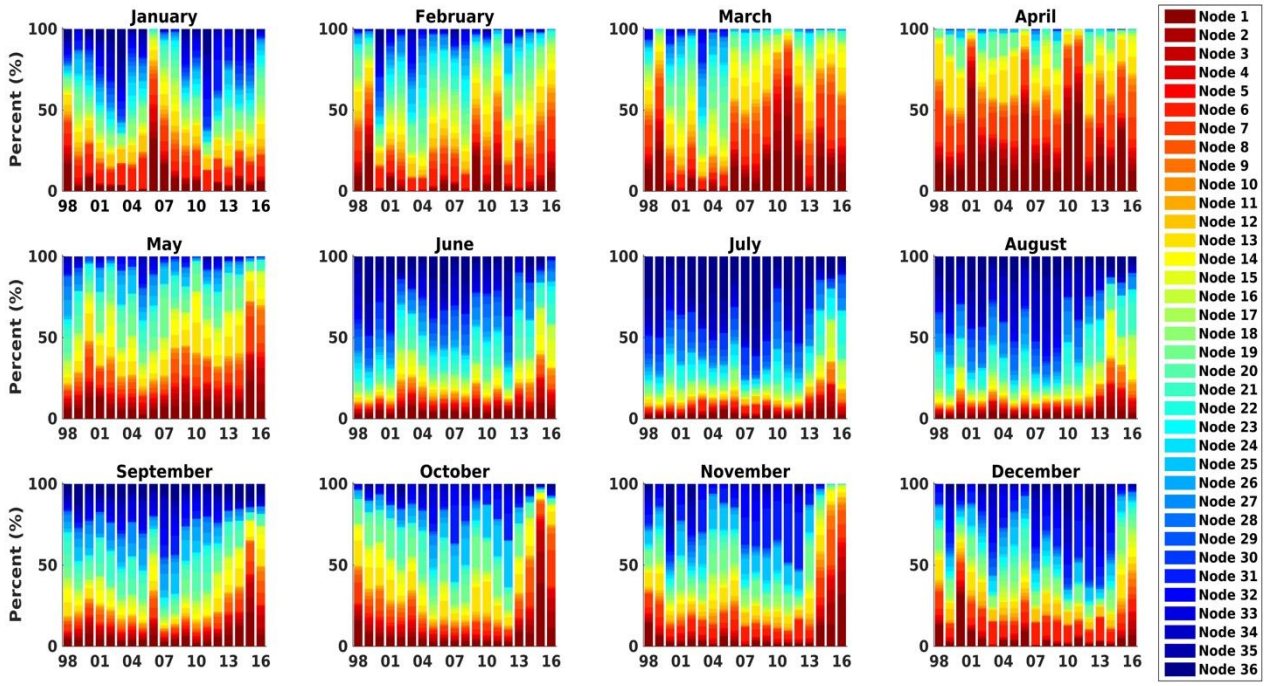

Fig. 20: Percentage of the total study area that each of the 36 SOM nodes occupies separated by month to show interannual variability over the 19-year time series. Anomalous bio-optical characteristics in specific months and years (e.g. March 2003, September 2006) become evident, in addition to shifts in overall distribution patterns in the last several years of the time series (e.g. July-November).

Appendix B

MONTHLY COMPOSITE GROUP DISTRIBUTION MAPS

Figures below show the distribution of the 8 final bio-optical groups for each month in each of the 19 years.

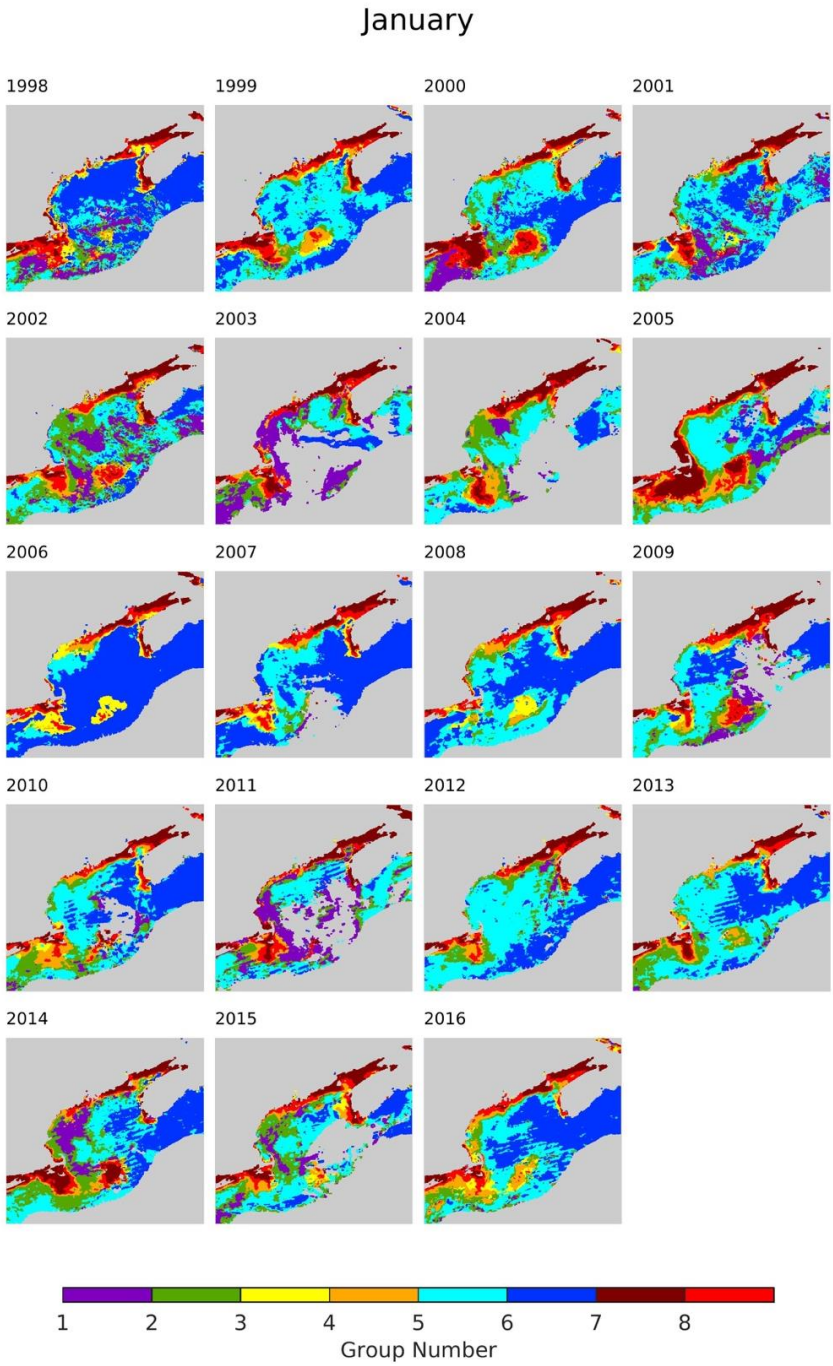

Fig. 21: January interannual distribution of 8 bio-optical groups over the 19-year time series.

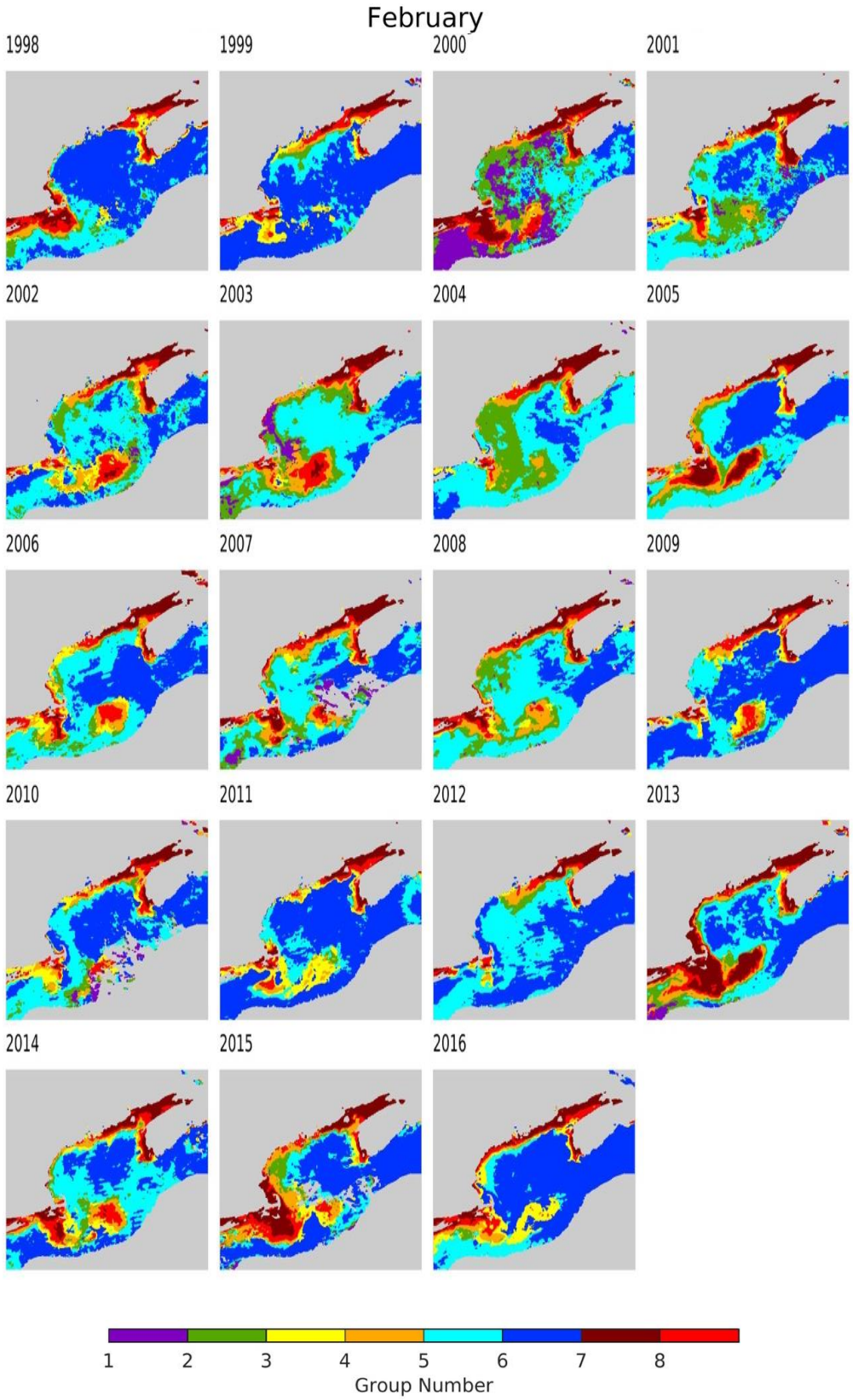

Fig. 22: February interannual distribution of 8 bio-optical groups over the 19-year time series.

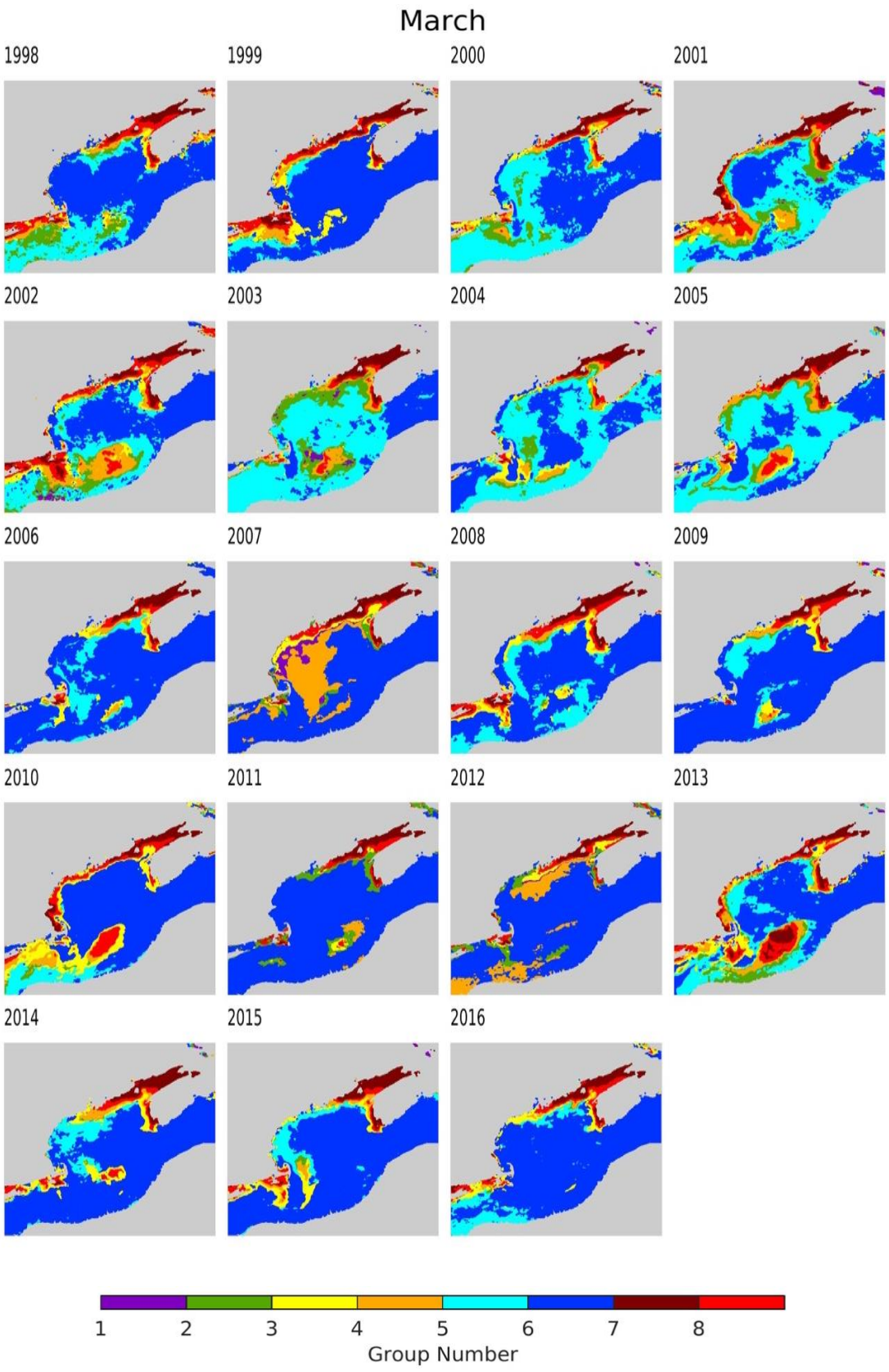

Fig. 23: March interannual distribution of 8 bio-optical groups over the 19-year time series.

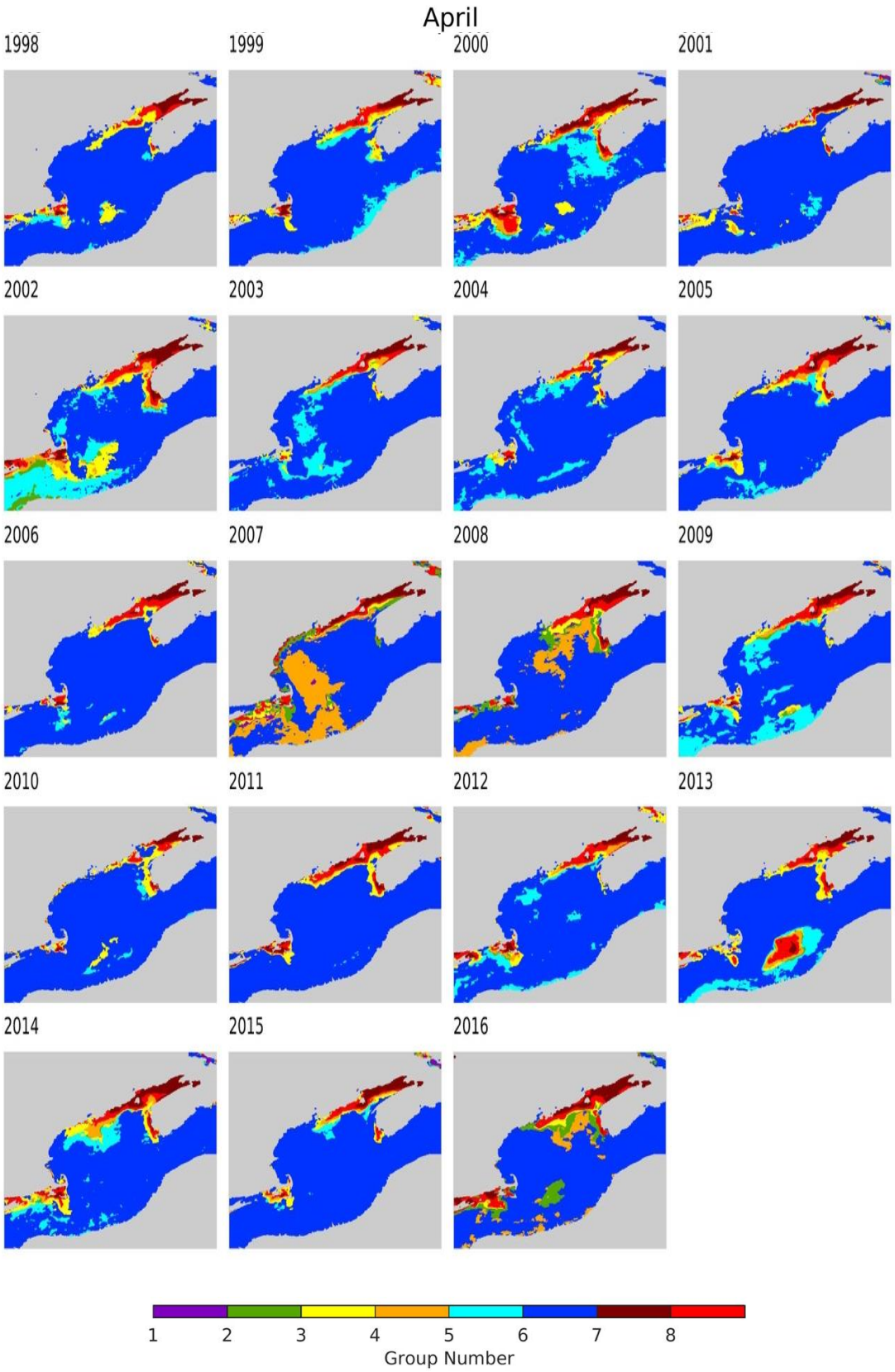

Fig. 24: April interannual distribution of 8 bio-optical groups over the 19-year time series.

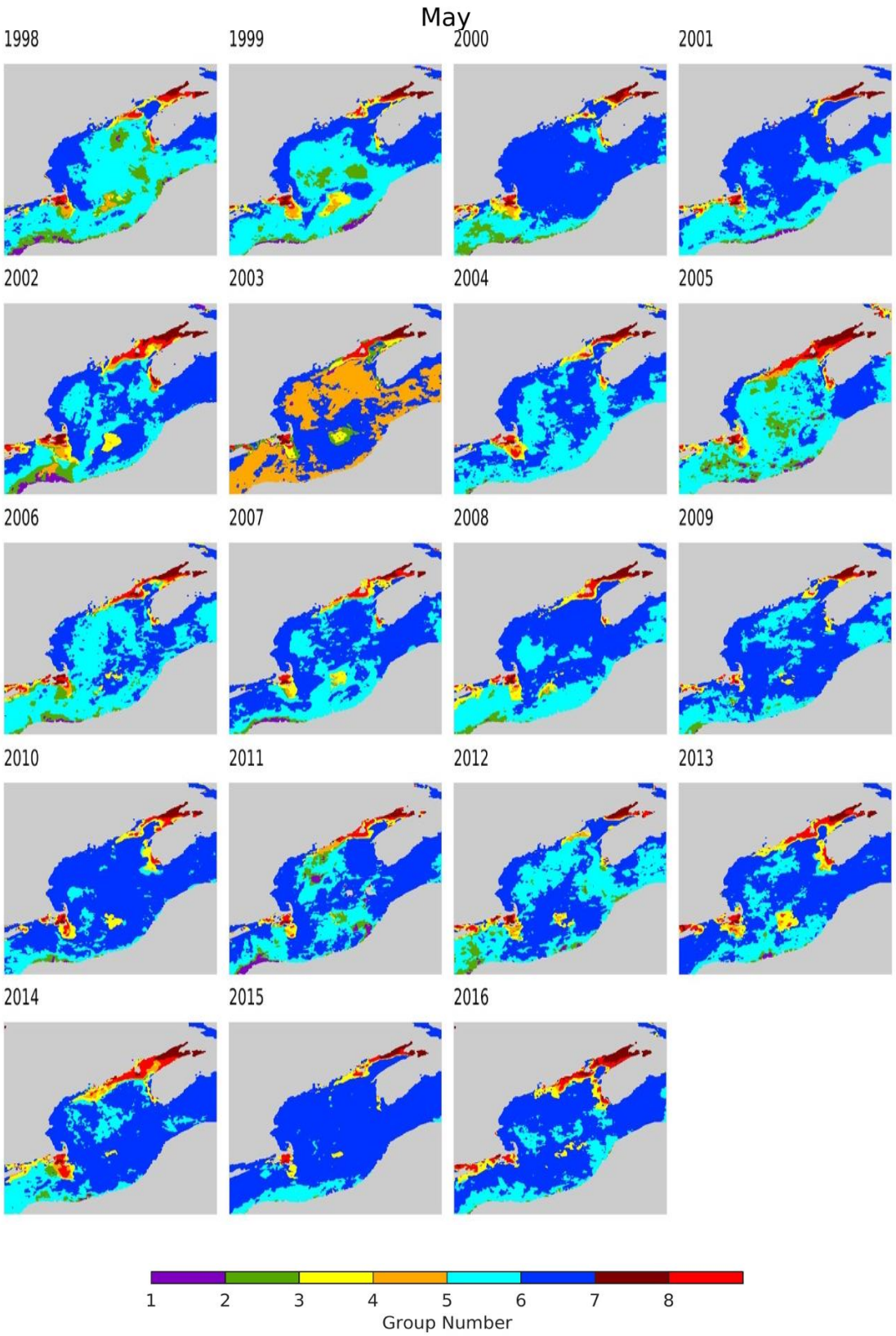

Fig. 25: May interannual distribution of 8 bio-optical groups over the 19-year time series.

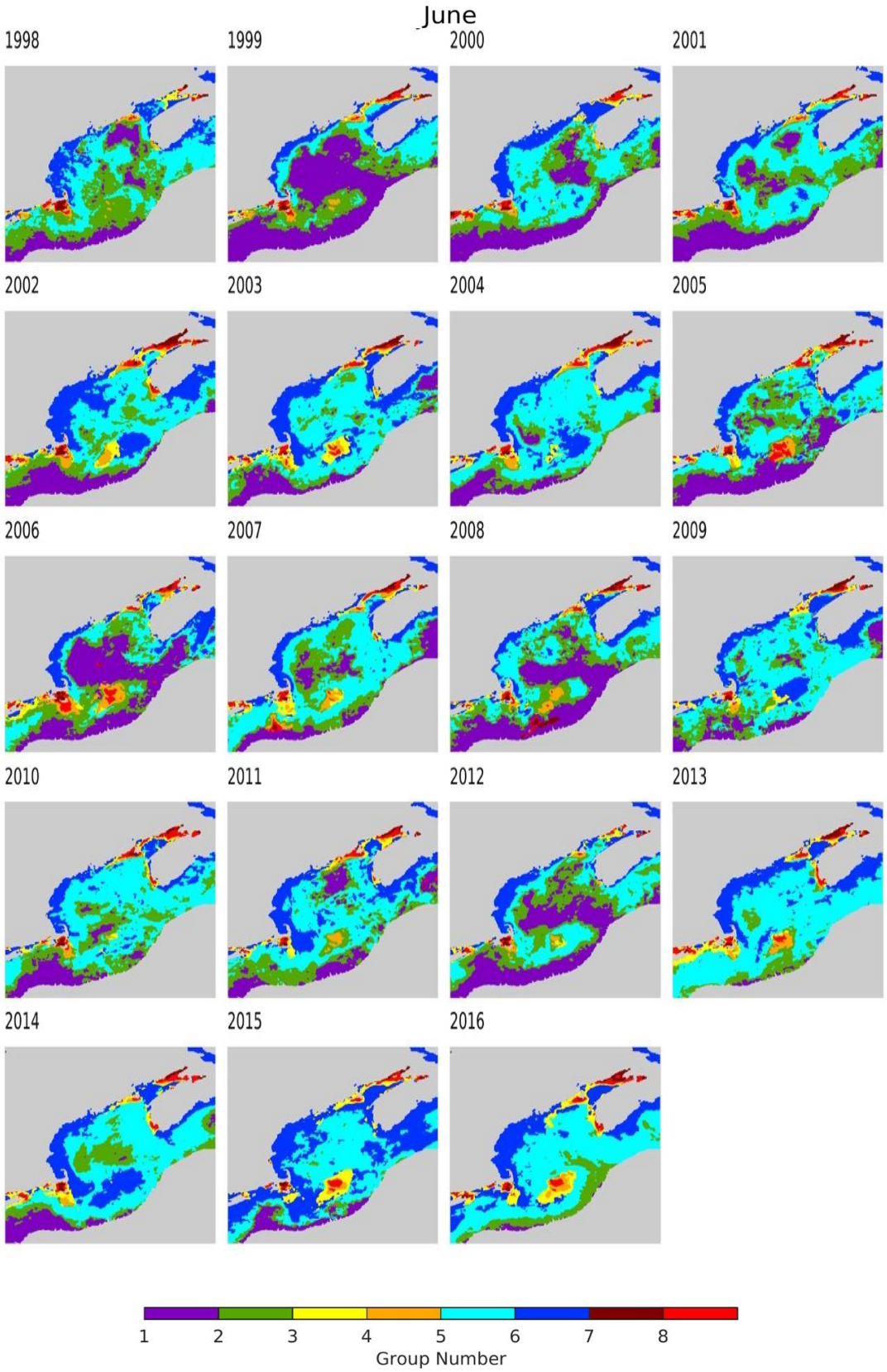

Fig. 26: June interannual distribution of 8 bio-optical groups over the 19-year time series.

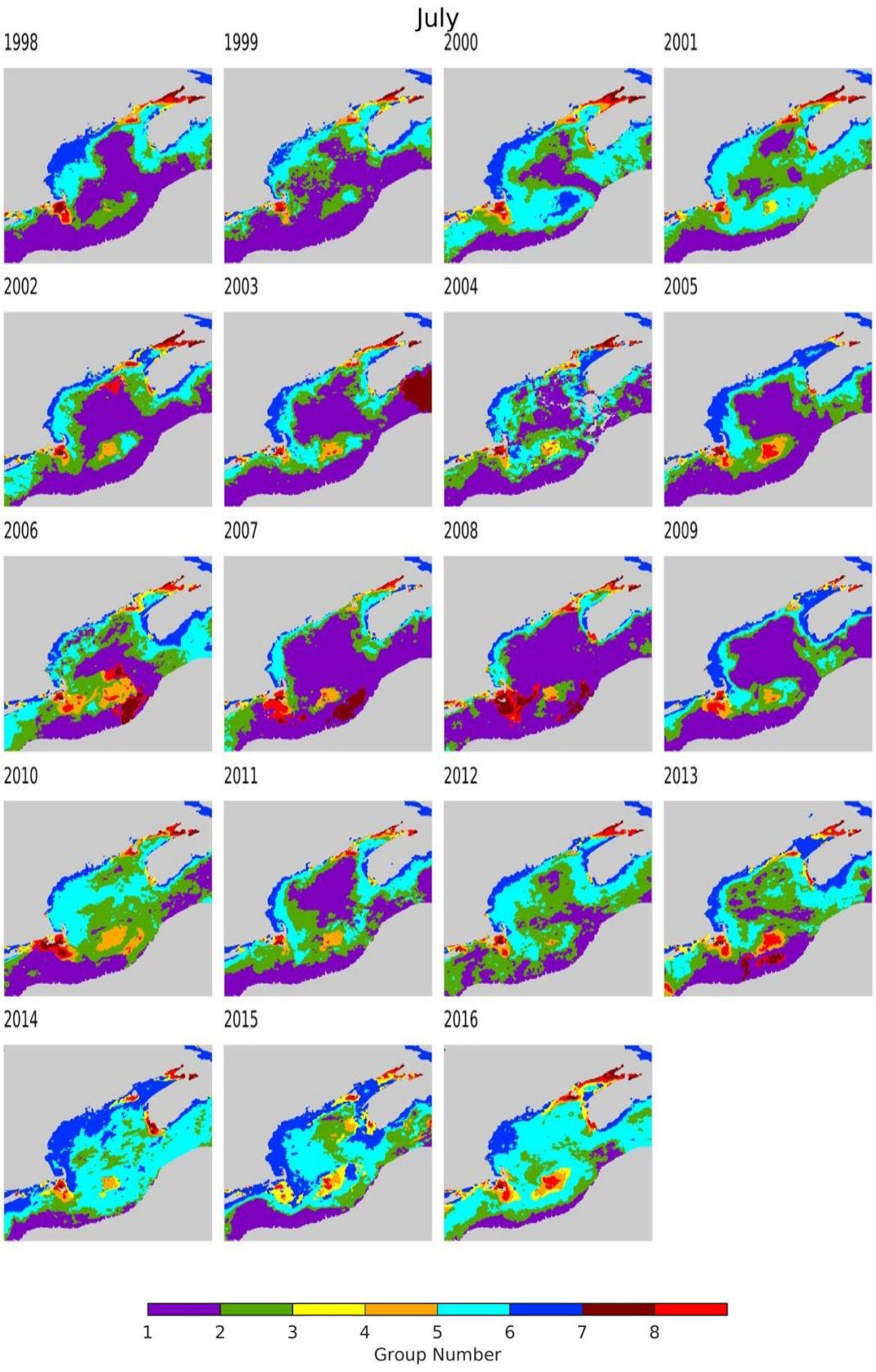

Fig. 27: July interannual distribution of 8 bio-optical groups over the 19-year time series.

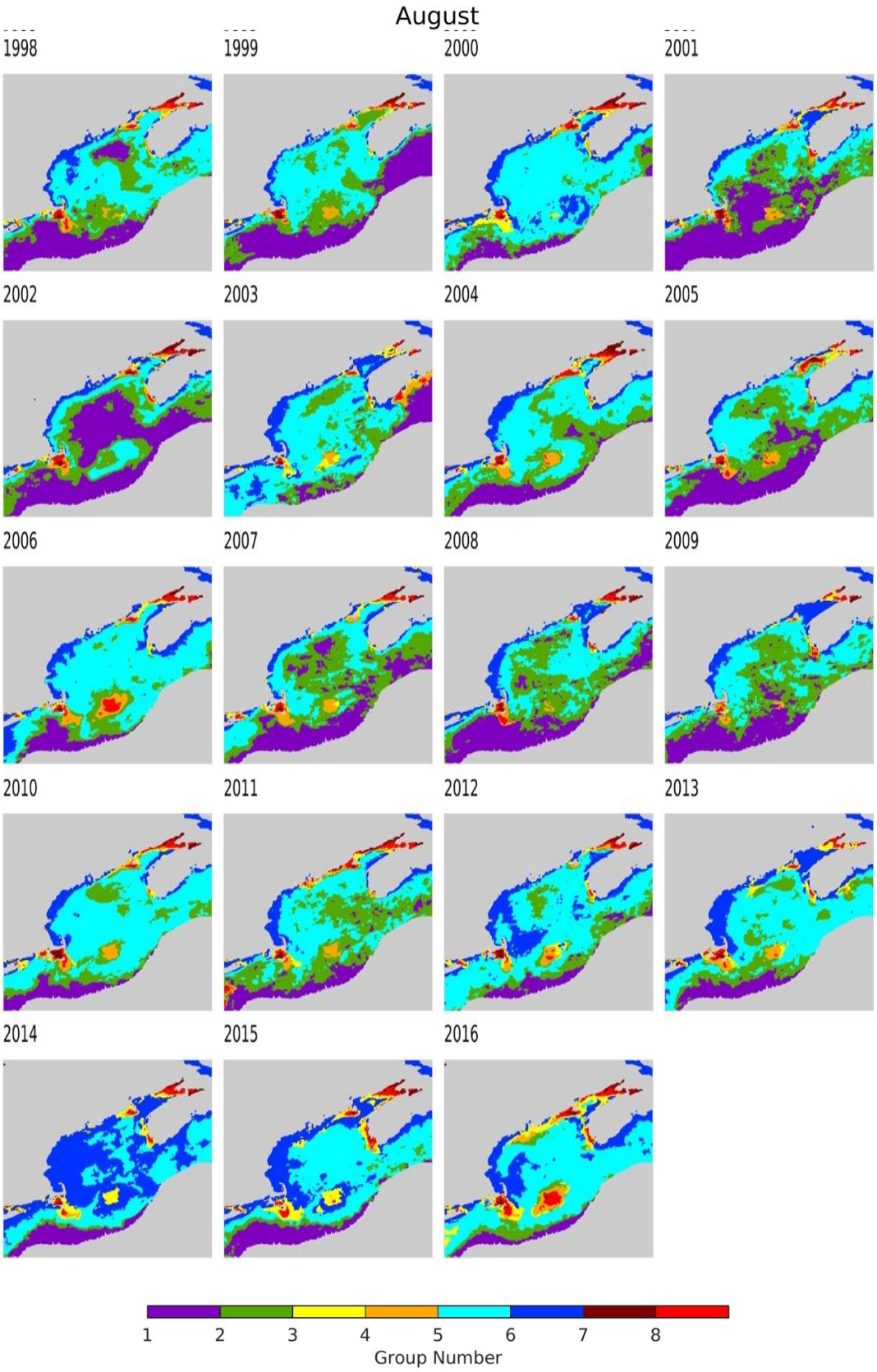

Fig. 28: August interannual distribution of 8 bio-optical groups over the 19-year time series.

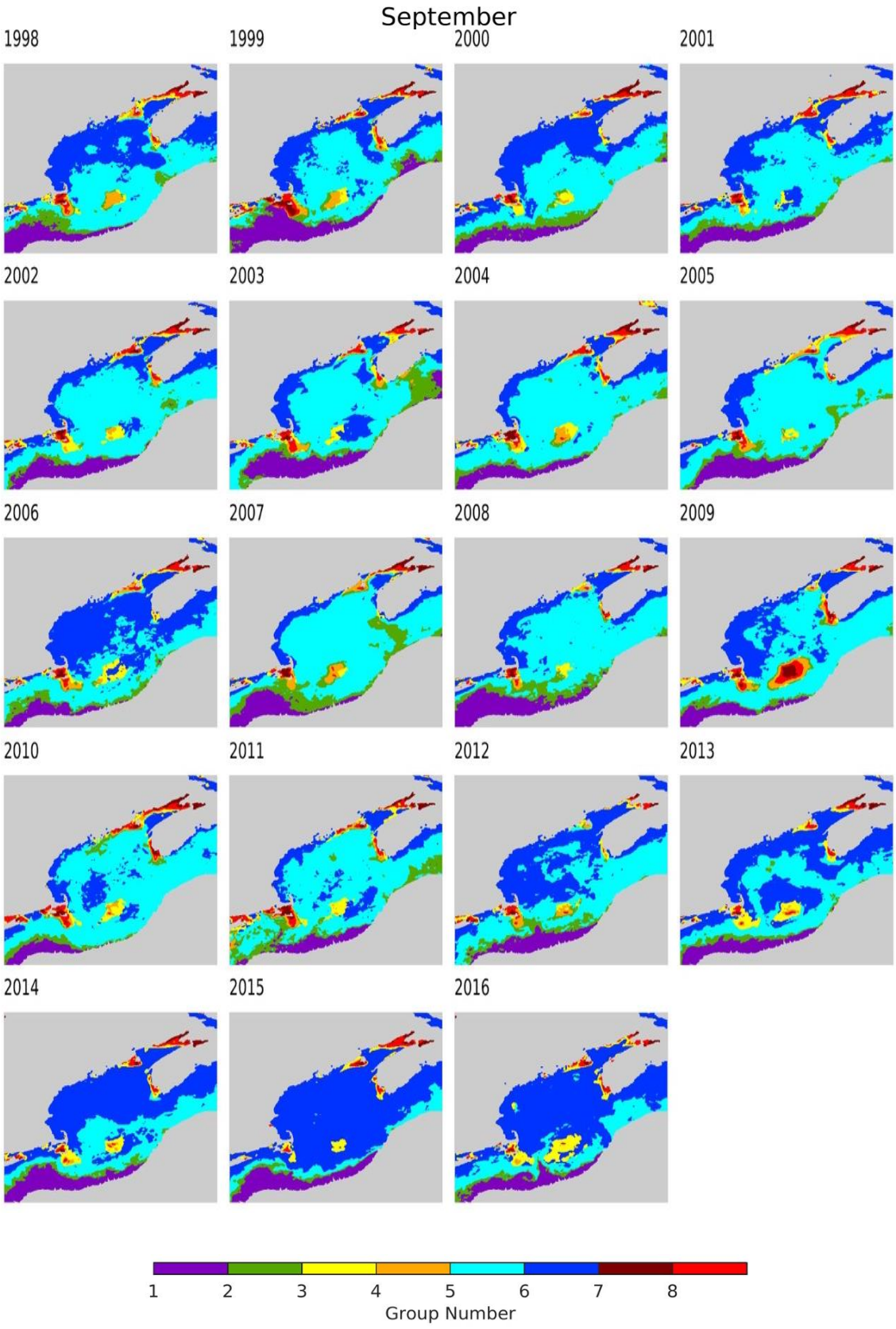

Fig. 29: September interannual distribution of 8 bio-optical groups over the 19-year time series.

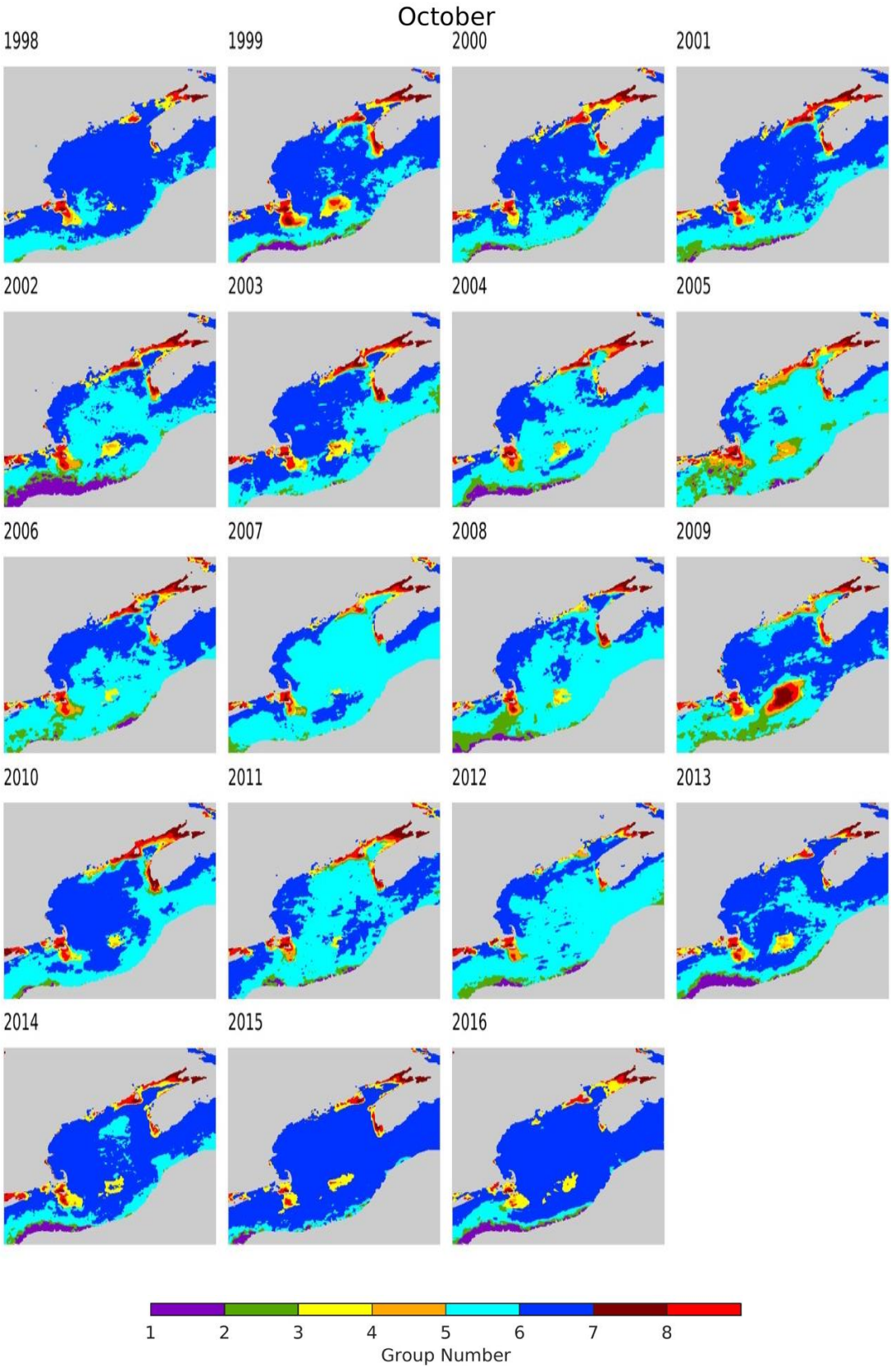

Fig. 30: October interannual distribution of 8 bio-optical groups over the 19-year time series.

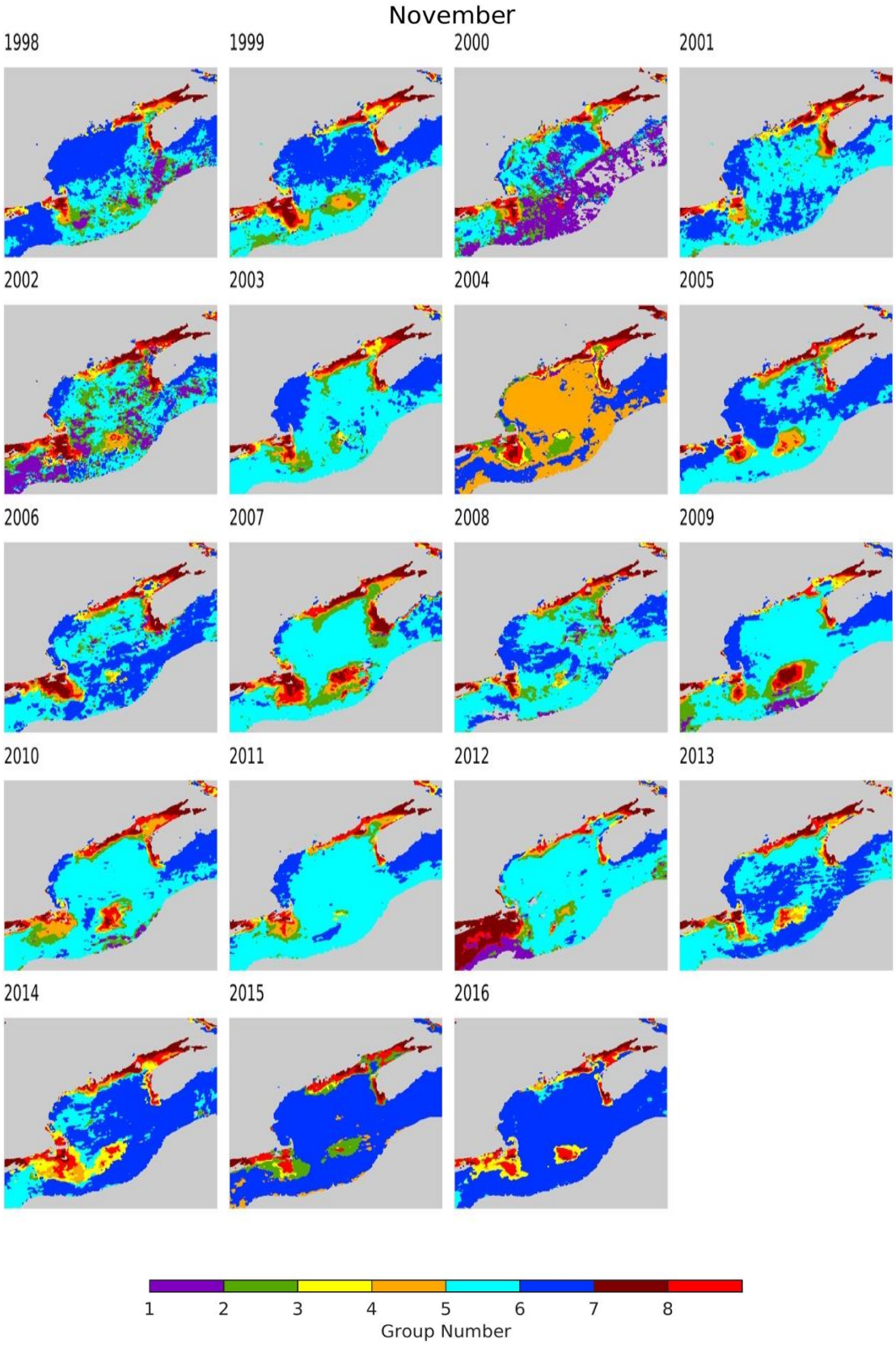

Fig. 31: November interannual distribution of 8 bio-optical groups over the 19-year time series.

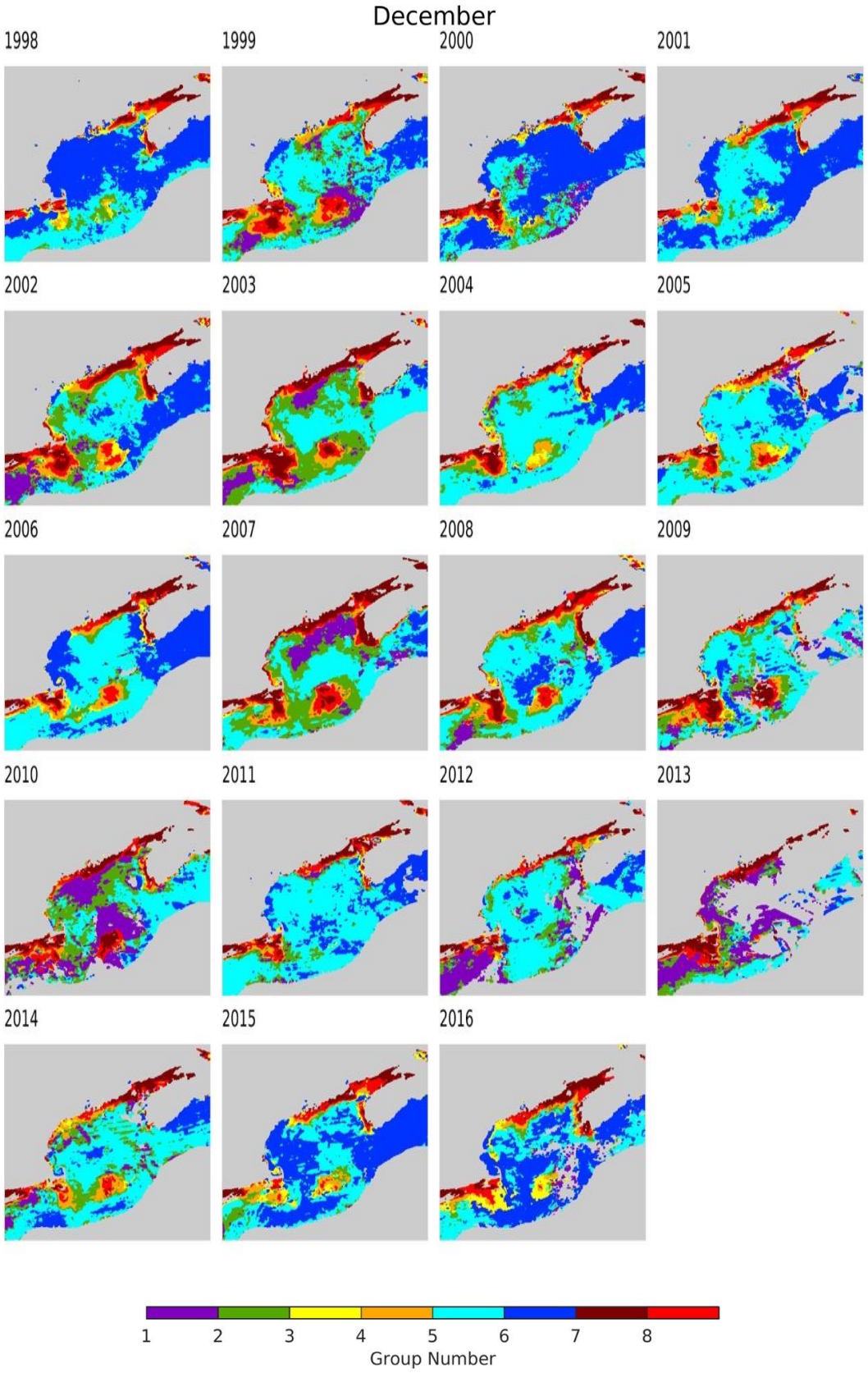

Fig. 32: December interannual distribution of 8 bio-optical groups over the 19-year time series.

Appendix C

MONTHLY PERCENT GROUP COVERAGE BAR PLOTS

Figures below show the percent coverage of the study region of the 8 final bio-optical groups for each of the 19 years, separated by month.

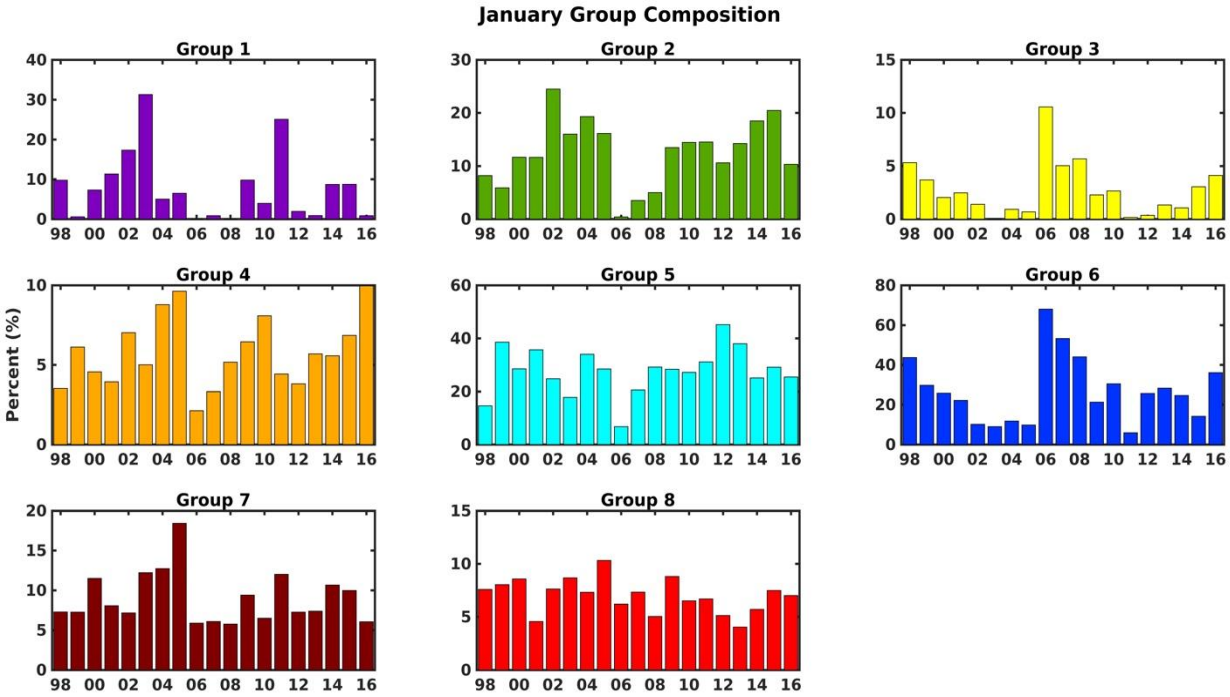

Fig. 33: January interannual variability in total percent study area coverage of each of the 8 bio-optical groups over the 19-year time series. Note the varying y-axis scaling.

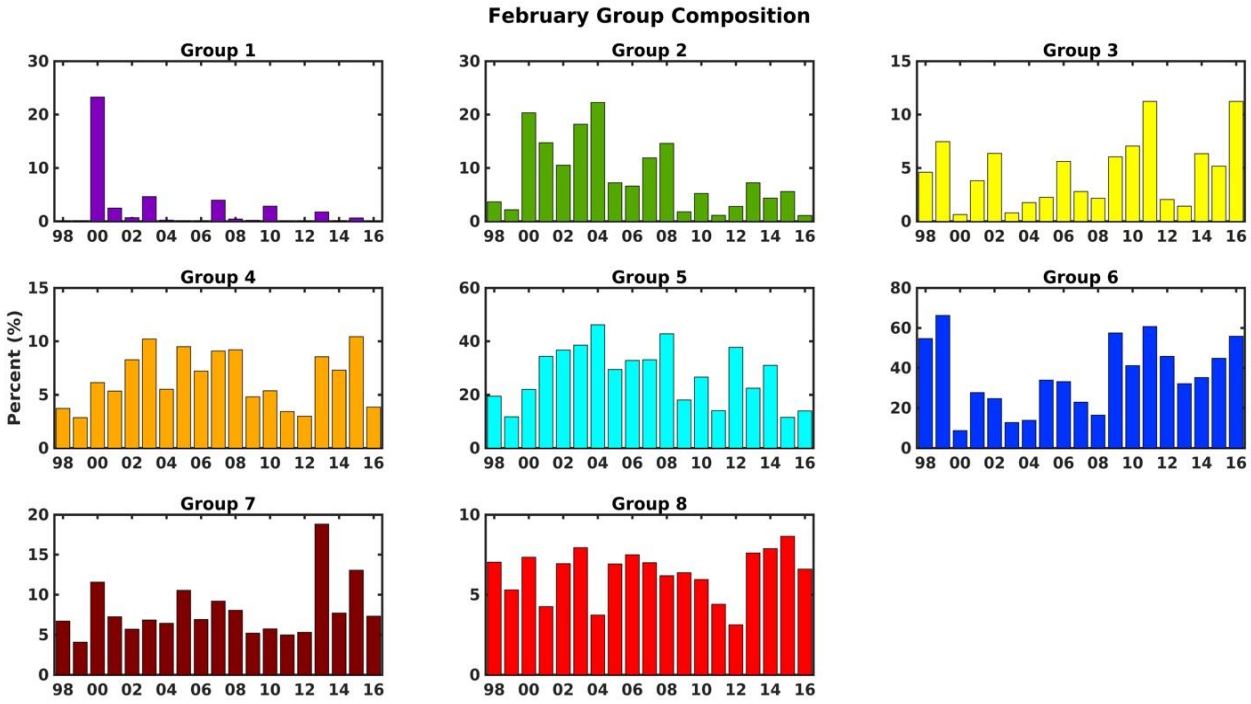

Fig. 34: February interannual variability in total percent study area coverage of each of the 8 bio-optical groups over the 19-year time series. Note the varying y-axis scaling.

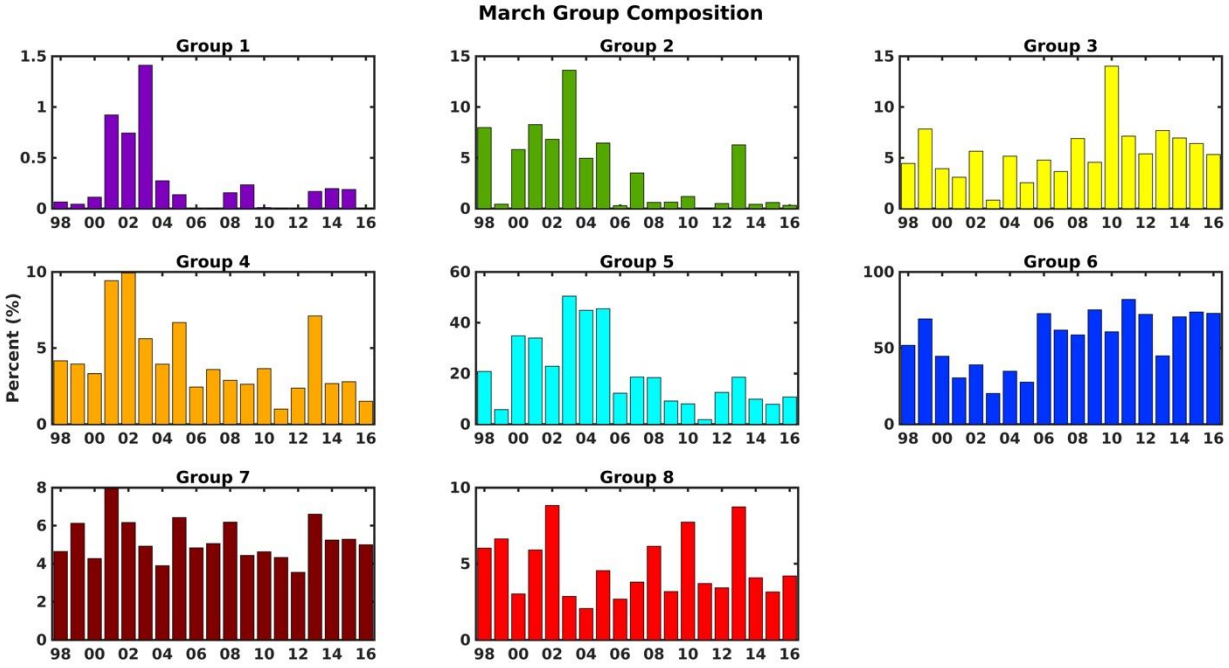

Fig. 35: March interannual variability in total percent study area coverage of each of the 8 bio-optical groups over the 19-year time series. Note the varying y-axis scaling.

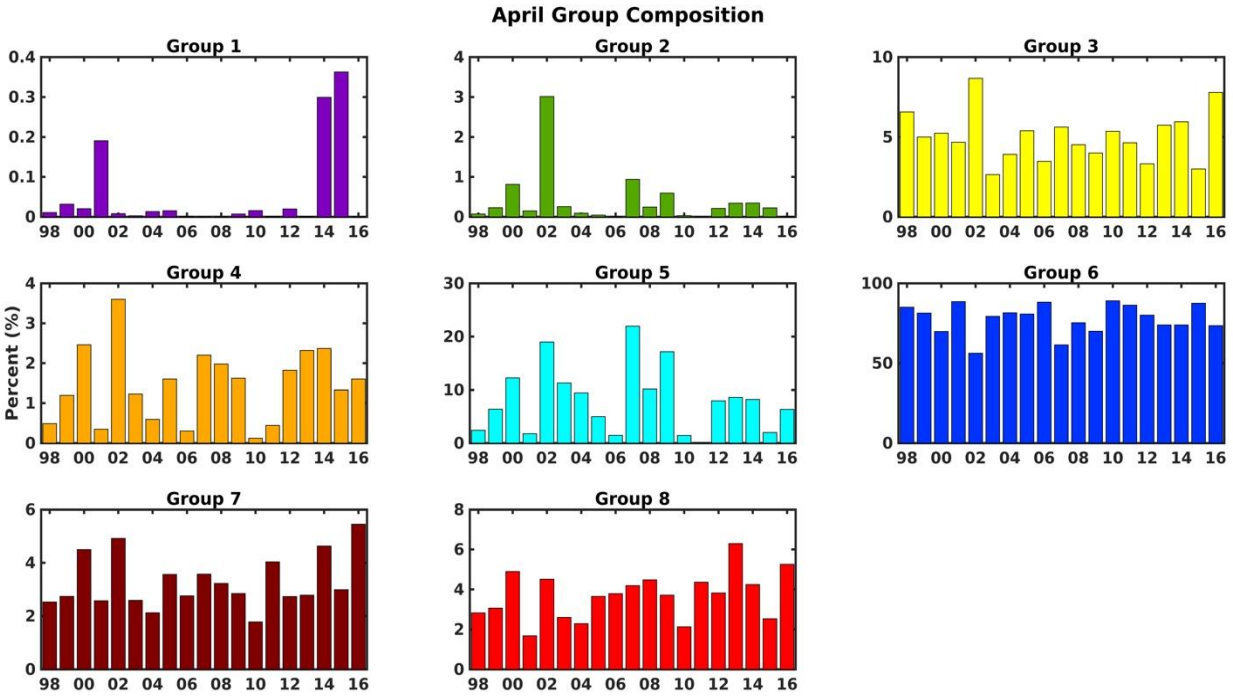

Fig. 36: April interannual variability in total percent study area coverage of each of the 8 bio-optical groups over the 19-year time series. Note the varying y-axis scaling.

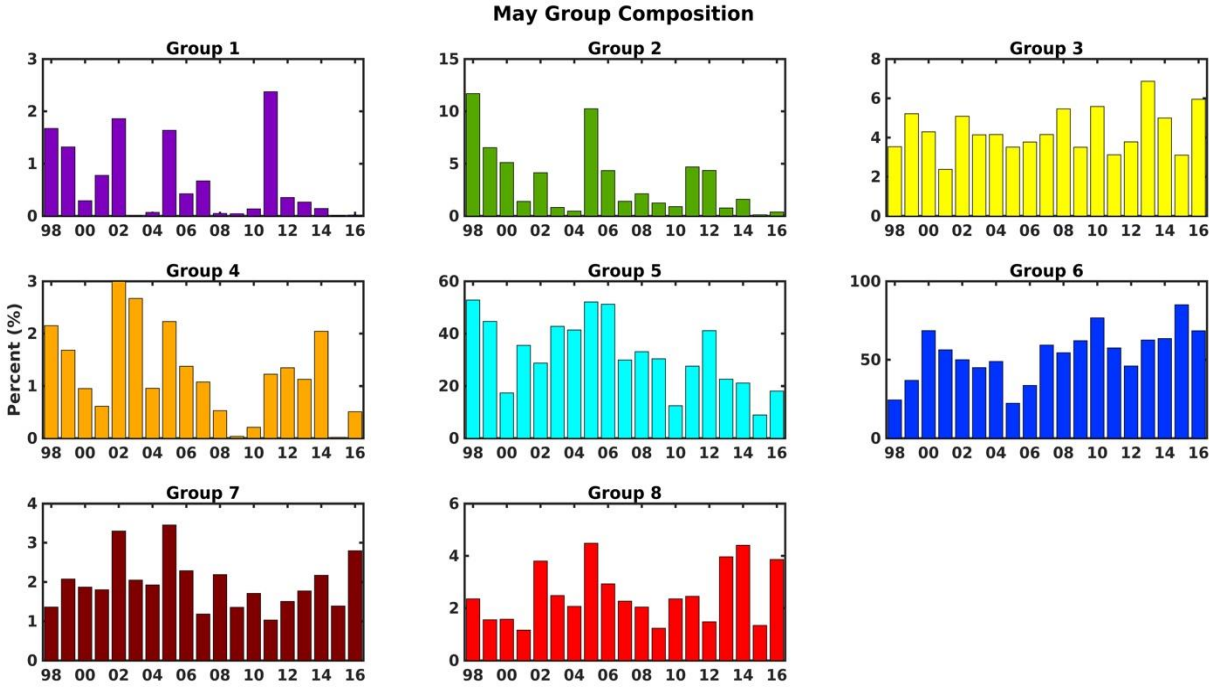

Fig. 37: May interannual variability in total percent study area coverage of each of the 8 bio-optical groups over the 19-year time series. Note the varying y-axis scaling.

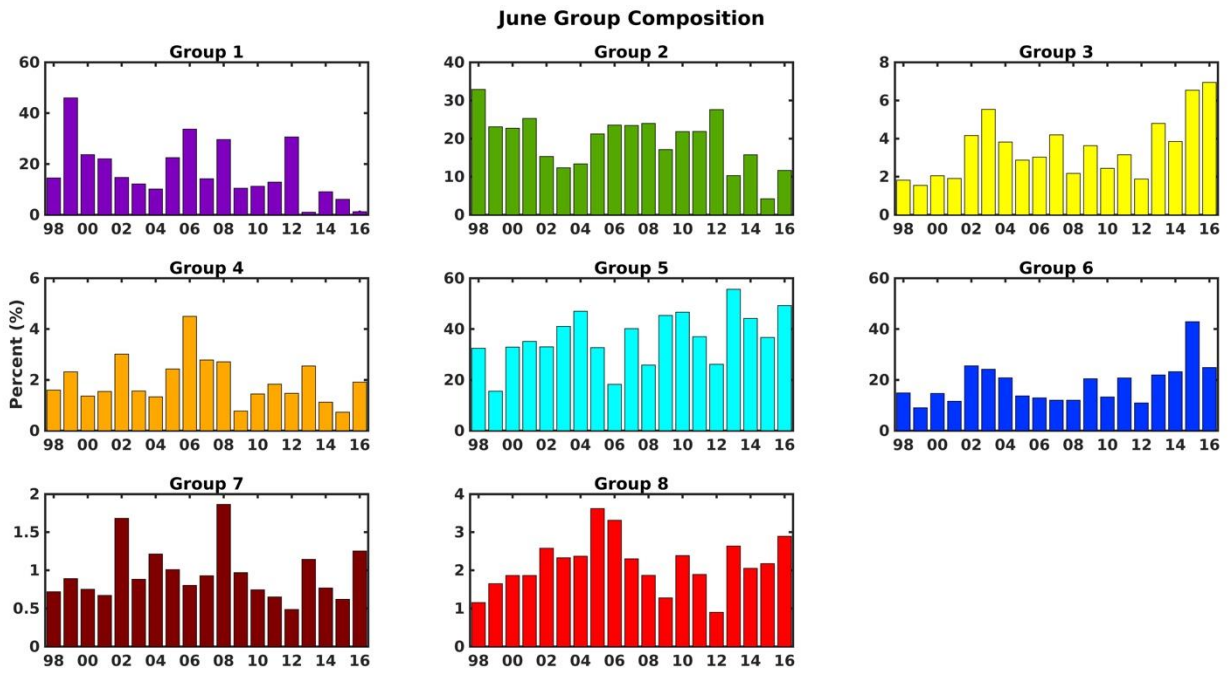

Fig. 38: June interannual variability in total percent study area coverage of each of the 8 bio-optical groups over the 19-year time series. Note the varying y-axis scaling.

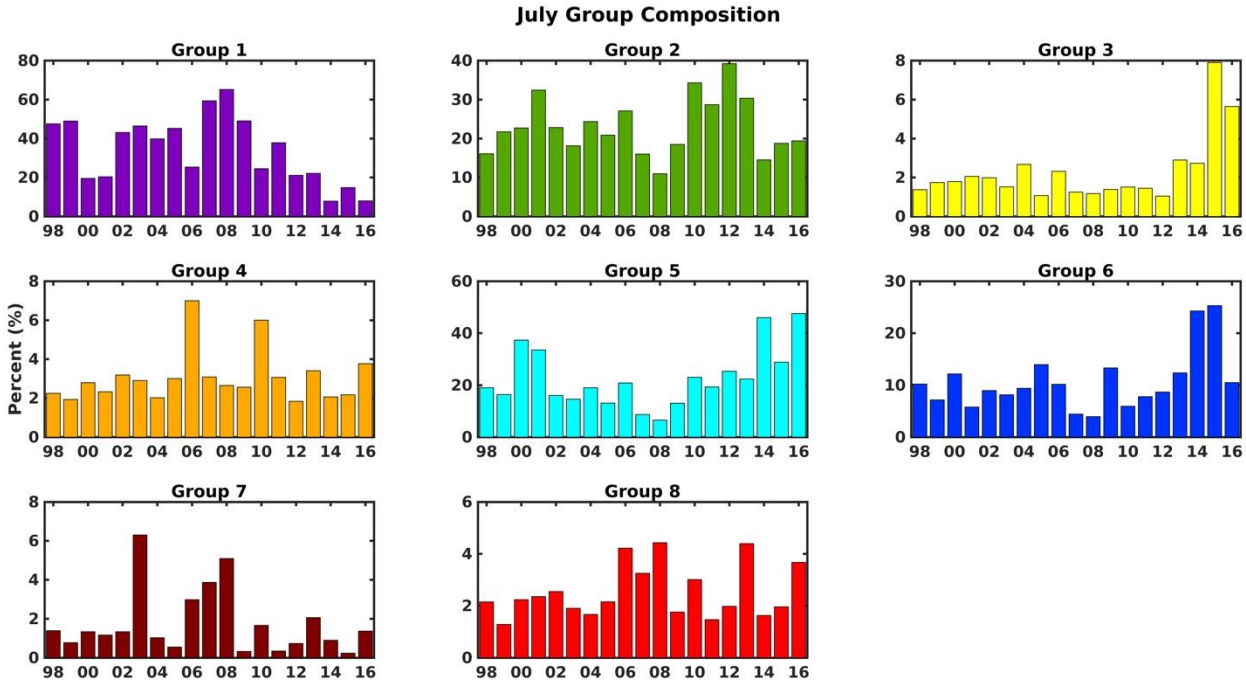

Fig. 39: July interannual variability in total percent study area coverage of each of the 8 bio-optical groups over the 19-year time series. Note the varying y-axis scaling.

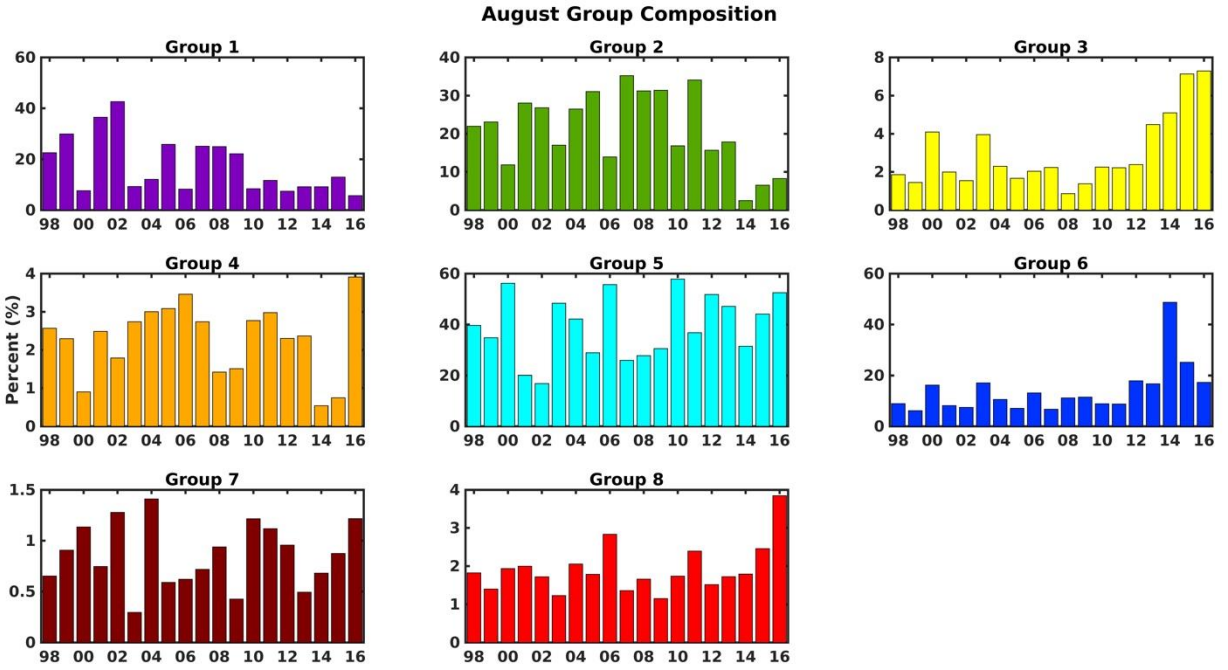

Fig. 40: August interannual variability in total percent study area coverage of each of the 8 bio-optical groups over the 19-year time series. Note the varying y-axis scaling.

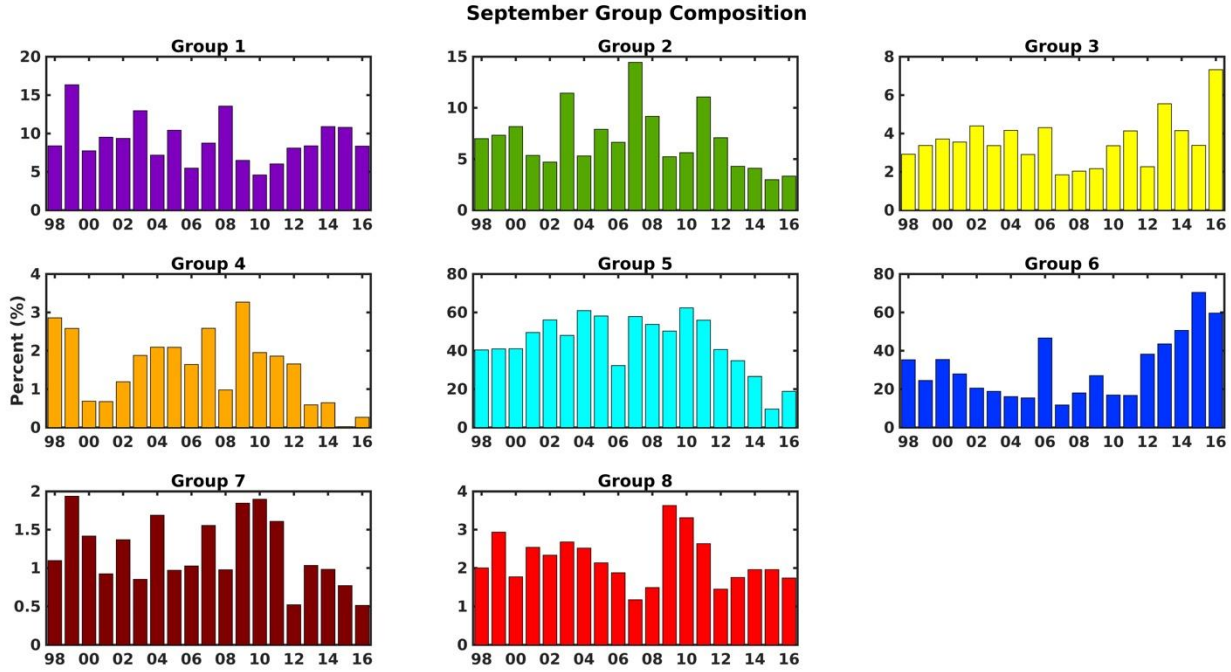

Fig. 41: September interannual variability in total percent study area coverage of each of the 8 bio-optical groups over the 19-year time series. Note the varying y-axis scaling.

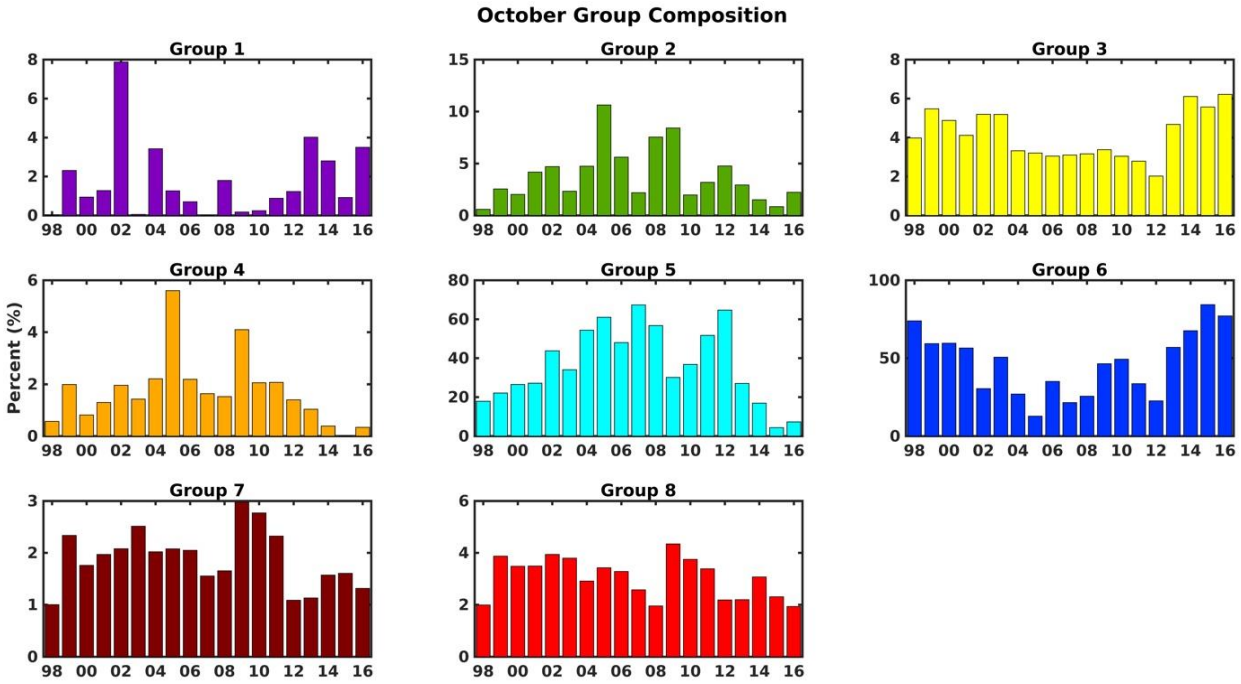

Fig. 42: October interannual variability in total percent study area coverage of each of the 8 bio-optical groups over the 19-year time series. Note the varying y-axis scaling.

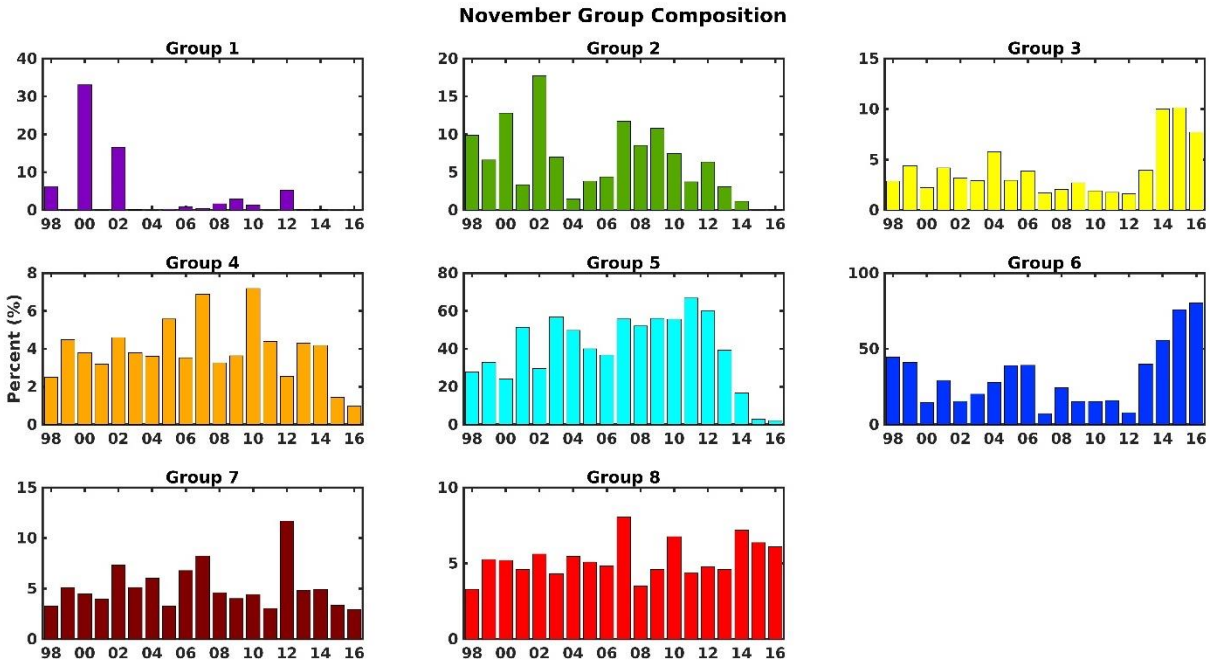

Fig. 43: November interannual variability in total percent study area coverage of each of the 8 bio-optical groups over the 19-year time series. Note the varying y-axis scaling.

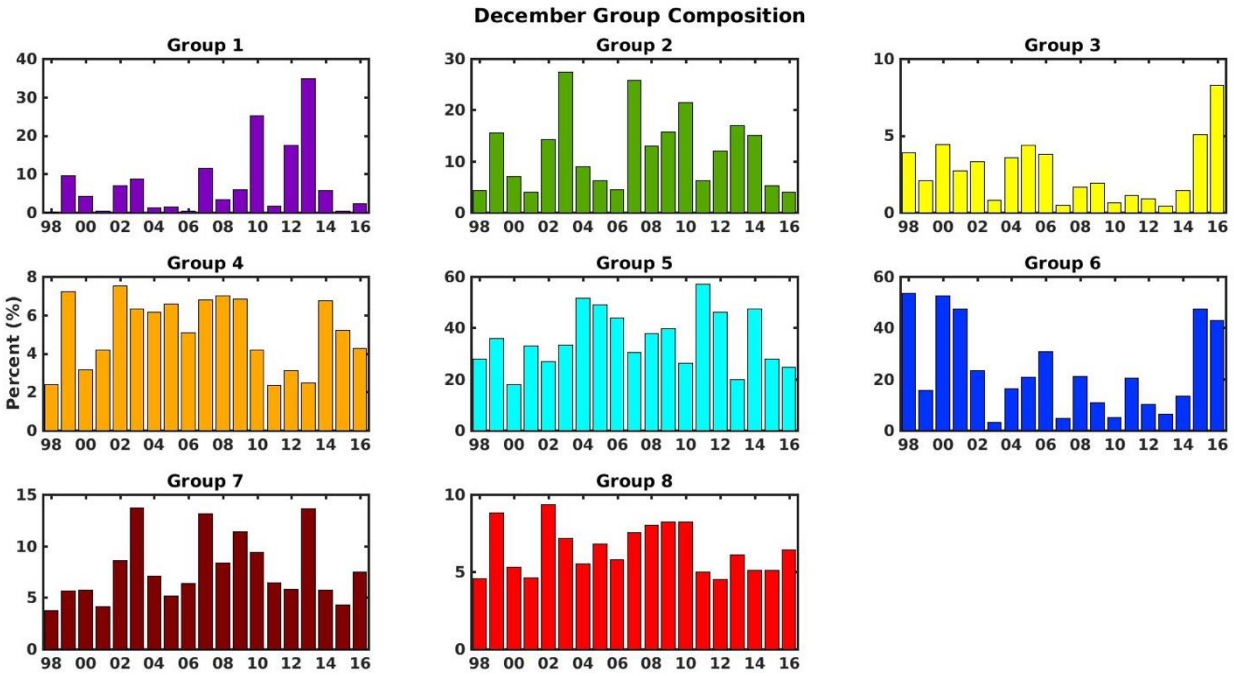

Fig. 44: December interannual variability in total percent study area coverage of each of the 8 bio-optical groups over the 19-year time series. Note the varying y-axis scaling.

Table 2: Monthly group slopes, p values and significance. Tends (% of study area year^{-1}) in each of the 8 groups, in each month, over the 19-year study period, their statistical significance (as p values) and a summary of the trend significance > 95%. Trends were calculated using the non-parametric Sen's slope calculations, p values test the null hypothesis that the actual trend is =0 and significance summary is binary, assigning a value of 1 to 95% significant trends.

Slope

Group	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1	-0.192	-0.016	-0.005	0.000	-0.055	-1.014	-1.885	-0.891	-0.095	0.052	-0.040	0.140
2	0.260	-0.637	-0.386	-0.004	-0.295	-0.536	0.132	-0.570	-0.203	-0.019	-0.487	0.004
3	-0.067	0.146	0.157	-0.006	0.037	0.153	0.070	0.103	0.051	-0.047	0.012	-0.079
4	0.128	0.047	-0.148	0.028	-0.071	-0.024	0.037	0.001	-0.078	-0.039	-0.043	-0.041
5	0.130	-0.468	-1.206	-0.206	-1.685	0.891	0.640	0.618	-0.957	0.042	0.360	0.212
6	-0.110	1.495	1.775	-0.095	1.742	0.527	0.310	0.563	1.248	0.238	0.625	-0.673
7	-0.031	0.062	-0.024	0.029	-0.015	-0.004	-0.034	0.004	-0.028	-0.029	-0.029	0.098
8	-0.113	0.025	-0.024	0.097	0.048	0.027	0.039	0.030	-0.031	-0.074	0.066	-0.015

p value

Group	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1	0.53	0.26	0.34	0.51	0.04	0.01	0.04	0.03	0.53	0.44	0.13	0.29
2	0.40	0.04	0.01	0.48	0.01	0.07	0.73	0.36	0.06	0.89	0.01	0.97
3	0.62	0.26	0.11	0.94	0.62	0.01	0.23	0.01	0.44	0.62	0.89	0.40
4	0.21	0.67	0.01	0.48	0.07	0.40	0.40	0.97	0.04	0.29	0.53	0.58
5	0.73	0.40	0.03	0.40	0.01	0.04	0.16	0.29	0.23	0.97	0.83	0.78
6	0.78	0.14	0.03	0.78	0.01	0.14	0.18	0.02	0.21	0.83	0.44	0.29
7	0.78	0.48	0.83	0.16	0.67	0.83	0.44	0.78	0.18	0.36	0.58	0.36
8	0.14	0.78	0.89	0.16	0.44	0.23	0.53	0.33	0.18	0.04	0.33	0.83

Significance

Group	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1	0	0	0	0	1	1	1	1	0	0	0	0
2	0	1	1	0	1	0	0	0	0	0	1	0
3	0	0	0	0	0	1	0	1	0	0	0	0
4	0	0	1	0	0	0	0	0	1	0	0	0
5	0	0	1	0	1	1	0	0	0	0	0	0
6	0	0	1	0	1	0	0	1	0	0	0	0
7	0	0	0	0	0	0	0	0	0	0	0	0
8	0	0	0	0	0	0	0	0	0	1	0	0

BIOGRAPHY OF THE AUTHOR

Brianna King was born in Pittsfield, Massachusetts where she graduated salutatorian of Taconic High School in 2009. She went on to attend Rensselaer Polytechnic Institute in Troy, New York and received a Bachelor of Science in Environmental Science in Fall 2012. She worked in the Department of Biology and Paleo-Environment at Columbia University's Lamont-Doherty Earth Observatory prior to beginning her graduate studies at the University of Maine in Fall 2016. She is a candidate for the Master of Science degree in Oceanography from the University of Maine in December 2018.