DIAGNOSING CDOM DYNAMICS IN A RIVER-DOMINATED COASTAL MARGIN: APPLICATION TO THE TEXAS-LOUISIANA SHELF USING MACHINE LEARNING AND AUTONOMOUS VEHICLES

A Dissertation

by

ROBERT LEE ILES IV

Submitted to the Graduate and Professional School of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Chair of Committee,	Steven F. DiMarco
Co-Chair of Committee,	Gerardo Gold-Bouchot
Committee Members,	Henry Potter
	Antonietta Quigg
Head of Department,	Shari Yvon-Lewis

December 2021

Major Subject: Oceanography

ABSTRACT

Both marine and terrestrial dissolved organic matter can play a major role in regulating the biogeochemistry and oxygen demand of the coastal ocean and continental shelf region. Where organic matter originates is not the only factor controlling ocean biogeochemistry, but also how recently it was created (i.e. freshness). Samples from two northern Gulf of Mexico cruises in 2019 were analyzed to determine dissolved organic matter fluorescence and subsequently used to create a predictive algorithm for estimating organic matter freshness and source (i.e. terrestrial vs. marine) using a number of variables, including depth, salinity, temperature, and oxygen among others. Additionally, a novel approach for using an established machine learning method, k-means clustering, was used to partition towfish and glider datasets into constituent groups. The derived algorithms were applied to multiple northern Gulf of Mexico datasets to determine the source and freshness of modeled organic matter components. Results showed that, in the early summer near the Mississippi River delta, dissolved organic matter is primarily fresh and terrestrial in origin. By the late summer, the freshest organic matter was found on the western shelf. Riverine impacts to shelf organic matter are reduced when river discharge is lower. Oxygen demand was well-correlated ($R^2 > 0.9$) to freshness, regardless of organic matter source. Both terrestrial and marine organic matter were found to correlate to observations of hypoxia, implying that northern Gulf of Mexico hypoxia can be driven by organic matter from sources other than byproducts from algal blooms exclusively. Results suggest that mixing and photodegradation were more impactful to organic matter variability in the upper layers, whereas in the bottom layers,

microbial uptake, as deduced by measurements of oxygen demand, was more impactful. Dataset cluster decomposition reveals both the physical structure of the water column as well as allows for quantification of the biogeochemical processes therein. The clusters produced in this study are shown to represent water masses distinguished by river plumes, wind-induced upwelling effects, shifts in currents, density-induced stratification, and microbial activity. This research shows the impacts that organic matter source and freshness can have on the biogeochemistry of a continental shelf region, in particular with respect to oxygen demand. The methods and algorithms produced in this study could also be applied to other river-dominated regions to investigate further the role of dissolved organic matter in ocean biogeochemistry.

DEDICATION

This work is dedicated to my wife, Bryanna, and to my two children, Levi and Abigail.

You all are everything to me.

ACKNOWLEDGEMENTS

I extend sincere appreciation to my committee chair, Dr. Steven DiMarco, my committee co-chair, Gerardo Gold-Bouchot, and my committee members, Dr. Henry Potter and Dr. Antonietta Quigg, for their guidance, instruction, and support throughout the course of my time at Texas A&M. I thank Dr. DiMarco for securing the funding for me to conduct this research, for fostering my passion for oceanography, and for helping me to mature as a researcher. With Dr. DiMarco's help over these last three years, I now sincerely feel able, willing, and confident to pursue the challenges ahead in my career. I am grateful to Dr. Gold-Bouchot for his help with the CDOM data processing and analysis, and especially for his patience in answering many of my questions over the course of the algorithm development process. Also, Dr. Gold-Bouchot's feedback on numerous revisions of the various manuscripts that comprise this work was crucial to its successful completion. I thank Dr. Henry Potter for his help in shaping this research from its inception, for providing helpful feedback and instruction with statistical analysis, and for allowing me to participate in a number of his cruises. Thank you, Dr. Potter, for also being an encouraging friend to me. Finally, I thank Dr. Antonietta Quigg for helping me to consider the broader role of my research in the context of the global ocean and the overall body of science.

I was fortunate to have such a wonderful support structure of friends and colleagues in the Department of Oceanography. I thank Laura Caldwell for the role she played in getting me in touch with Dr. DiMarco initially and, ultimately, helping me to get to Texas A&M. To my labmates, Sakib Mahmud, Xiao Ge, Meng Lyu, Xin Yang,

Alex Sun, Brian Buckingham, Dr. Franke Hsu, and Sam Longridge, thank you all for your companionship and help throughout this process. We made many fun memories that I will cherish forever. In particular, thank you to Sakib, Xiao, Meng, Franke, Brian, and Laura for all your help in getting me settled in at A&M in my first year. Also, thank you to Xiaoqi Wang for being the first to suggest that I explore the machine learning clustering approach for my glider data; thank you, Xiaoqi, also for taking care of Meng for us!

Thank you to GCOOS for funding much of this research and to those at GCOOS who helped with various aspects of my research assistantship, in particular Dr. Barb Kirkpatrick, Felimon Gayanilo, Dr. Shin Kobara, Dr. Franke Hsu, and Marion Stoessel. To Dr. Kerri Whilden and the GERG glider team, thank you for all that you did in helping me with glider data access and processing. I also would like to thank Dr. Kristen Thyng and Dr. Chrissy Wiederwohl for their help with developing my proficiency with Python and Matlab, respectively. To Dr. Wiederwohl and Dr. Jess Fitzsimmons, I thank you for the work that you both do with the GLC, which was a significant help to me getting acclimated in my first year. I would also like to acknowledge my master's committee members from LSU, Dr. Nan Walker, Dr. Robert Rohli, and Dr. John White, for encouraging me to pursue a PhD and for preparing me to achieve success in my doctoral research.

I cannot express the fullness of my gratitude to my family. To my parents, Trey and Deanna, and my sisters, Julie and Jannah, I am so grateful that I grew up in a loving, warm, and encouraging household with you all. Thank you all for the support and help

that you have provided to me not only over these last three years, but throughout my life. To my children, Levi and Abigail, you both make me so happy and I am so blessed to have you in my life. Although having infants and toddlers while in graduate school is certainly a challenge, I would not have had it any other way. Finally, to my wife, Bryanna, I will never be able to convey the full depth of my appreciation for all that you have done for me. You encouraged me when I did not think I was capable, you were my greatest help in difficult times, and you were the most vocal in cheering my successes. I could not have done this without you and it is as much your achievement as it is mine.

CONTRIBUTORS AND FUNDING SOURCES

Contributors

This work was supervised by a dissertation committee consisting of Professor Steve DiMarco as advisor, Professor Gerardo Gold-Bouchot as co-advisor, and Professor Henry Potter of the Department of Oceanography and Professor Antonietta Quigg of the Department of Marine Biology at Texas A&M Galveston.

The glider data analyzed in this work were supplied by Dr. Steven DiMarco and the Texas A&M Geochemical and Environmental Research Group. Cruise data from June 2019 that were used in Chapter 3 were supplied by Dr. Gerardo Gold-Bouchot. All other work conducted for the dissertation was completed by the student independently, including data collection accomplished aboard cruises funded by Dr. Henry Potter.

Funding Sources

Graduate study was supported by an assistantship from GCOOS. This work was also made possible in part by Texas Sea Grant Grants-in-Aid of Graduate Research Institutional Grant NA18OAR4170088, to the Texas Sea Grant College Program from the National Sea Grant Office, National Oceanic and Atmospheric Administration, U.S. Department of Commerce. Additional funding from NSF, GOMRI, IOOS, Texas One Gulf, and NOAA helped to support data collection for various aspects of this project.

NOMENCLATURE

ANOVA	Analysis of Variance
AOU	Apparent Oxygen Utilization
BIX	Biological Index (i.e. “freshness”)
Chl-*a*	Chlorophyll-*a*
CDOM	Chromophoric (colored) Dissolved Organic Matter
CMR	Peak C : Peak M Ratio
DOM	Dissolved Organic Matter
EEM	Excitation-emission Matrix
FDOM	Fluorescent Dissolved Organic Matter
GERG	Geochemical and Environmental Research Group
MARS	Mississippi-Atchafalaya River System
MCH	Texas A&M University Mechanisms Controlling Hypoxia Project
NDBC	NOAA National Data Buoy Center
PCA	Principal Component Analysis
SSH	Sea Surface Height
TXLA	Texas-Louisiana (Shelf)
USGS	United States Geological Survey

TABLE OF CONTENTS

Page

LIST OF FIGURES

Page

LIST OF TABLES

Page

1. INTRODUCTION

1.1. Background

Dissolved organic matter (DOM) is an important control on marine biogeochemistry, in particular as the foundational part of the microbial loop as it provides the fuel necessary for microbial remineralization and respiration. A comprehensive understanding of oceanic DOM dynamics can help to further knowledge of the global carbon cycle (Osburn et al. 2016), ocean acidification (Cai et al., 2011; Kealoha et al., 2020), harmful algal blooms (Bargu et al., 2019; D'Sa, 2008; Hetland & Campbell, 2007), and hypoxia (Bianchi et al., 2010; Rowe & Chapman, 2002). DOM generally enters the marine environment from either marine (e.g. biomass decay or photodegradation of other material) or terrestrial sources (e.g. rivers). The varying characteristics of marine and terrestrial DOM have been extensively researched and qualified (Coble 1996, 2007; Carlson & Hansell 2015; Repeta 2015). In estuarine and coastal environments with significant river influences, the contrast between terrestrial and marine DOM can result in significant variations in biogeochemistry, as DOM from differing sources differs with respect to its chemical structure, bioavailability, and relative molecular weight (Hansen et al., 2016; Zhou et al., 2018). A key metric to consider when analyzing DOM is how recently the DOM was created, often known as its "freshness". DOM freshness can reveal how biologically active a study area is because fresher DOM is often more bioavailable.

A common method for studying DOM is to examine the chromophoric portion of the bulk DOM pool (CDOM) using spectral analysis techniques. CDOM is simply DOM that absorbs light; measuring the amount of absorption of light and at what wavelengths light is absorbed can provide information on the characteristics (e.g., source, freshness, chemistry, etc.) of the CDOM in the sample being analyzed. Fluorescent DOM (FDOM) is a subset of CDOM that both absorbs and emits light (i.e., "fluoresces"). As with CDOM absorption analysis, measuring the amount of light emitted at a particular wavelength of an FDOM sample, and furthermore comparing the amount of light emitted at a different wavelength, can reveal the nature of the FDOM, including its origin and how recently it was created (i.e., freshness). Fluorometers are commonly used aboard marine technology systems including CTDs, towfish, and submersible gliders to measure FDOM. However, such fluorometers generally only measure FDOM at one wavelength, thus providing only an estimate of bulk FDOM (sometimes called "bulk CDOM").

The Texas-Louisiana (TXLA) shelf is an area of the northwestern Gulf of Mexico that experiences extreme variability in both physical processes and biogeochemistry, in large part because of the influence of rivers. Nutrients from the Mississippi-Atchafalaya River System (MARS) help to fuel algal blooms along the TXLA shelf, the transformation of which by heterotrophs results in export of massive amounts of DOM to the lower depths. In the spring and summer, density gradients arise from the MARS freshwater plumes and warming of the surface layers such that the water column becomes stratified. The stratification of the water column inhibits

ventilation of the lower layer such that degradation of dead phytoplankton cells and consumption of DOM in the ocean bottom can lead to severe depletion of dissolved oxygen and, ultimately, hypoxic conditions. While hypoxia on the TXLA shelf has been extensively researched for decades, a gap exists in the knowledge of the actual nature of the DOM that helps fuel hypoxia in this region (Pauer et al., 2014; Bianchi et al., 2010; Bianchi et al., 2011). Hypoxia-causing DOM is commonly thought to originate almost exclusively from algal blooms fueled by MARS resources; however, DOM from terrestrial sources, from the MARS endmembers for example, has not been fully explored as an additional potential pathway leading to hypoxia formation on the TXLA shelf.

1.2. Research Questions

The primary objective of this study is to use the spectral characteristics and mixing dynamics of FDOM in order to develop novel methods for detailing the biogeochemistry of the TXLA shelf region of the northern Gulf of Mexico. While fluorometer data from submersible buoyancy gliders or towed systems, for example, are useful for estimating bulk CDOM concentrations, bulk CDOM alone is an insufficient metric for gaining a more complete understanding of how terrestrial and marine FDOM sources impact TXLA shelf biogeochemistry. Comparing bulk CDOM to other oceanographic properties, such as salinity, can help get a sense of the relative rate of nonconservative mixing is occurring in a system; novel analytical methods are required

to ascertain where and why nonconservative behavior (e.g. consumption or creation of DOM) occurs.

This dissertation utilized discrete water samples, data from multiple ocean technology platforms, including autonomous and towed systems, and ancillary data from buoys and satellites in order to answer the following key questions:

1. Can machine learning methods be used to explain the mixing behavior of CDOM as measured by subsurface ocean technology platforms?

2. Can machine learning methods and high-resolution spectroscopic analysis of discrete water samples be used in tandem to characterize CDOM dynamics, beyond bulk CDOM measurements, measured by 3-channel fluorometers commonly used aboard underwater vehicles and towed system?

3. How does the recent autochthonous contribution (e.g. as determined by the CDOM freshness) and DOM source (e.g. terrestrial vs. marine) of the TXLA shelf bulk CDOM pool change:

   (a) with respect to salinity gradients, both relative to the pycnocline as well as distance from major regional freshwater endmembers (e.g. MARS endmembers); and,

   (b) in relation to the time scales of summertime seasonal hypoxia?

1.3. Scientific Hypotheses

The project's research questions lead to the following scientific hypotheses to be tested:

- **H1**: Machine learning methods can classify and quantify water column biochemical variability into water mass structure that can evolve at spatiotemporal scales on the order of hours and kilometers in TXLA shelf waters.
- **H2**: The TXLA shelf sub-pycnocline bulk CDOM pool is comprised of a greater portion of recent autochthonous material on the western side of the shelf than on the eastern side.
- **H3**: The recent autochthonous constituent (i.e., freshness) of the TXLA shelf bulk CDOM pool is inversely correlated with oxygen concentration in the spring and early summer, but is not correlated with dissolved oxygen in the late summer and autumn as recalcitrant material becomes more prevalent.

1.4. Organization

The dissertation is comprised of five chapters. Chapter 1 presented a concise introduction to DOM, its importance on the TXLA shelf, and the limitations of methods of fluorescence analysis of DOM by subsurface ocean technology platforms. Chapters 2 thru 4 each contain a more comprehensive introduction section that expounds on the material that those chapters cover. Chapter 2 details the development of a novel technique involving established machine learning methods that can be applied to large datasets in order to identify water masses and associated biogeochemical processes as measured by subsurface ocean technology platforms. Chapter 3 describes the creation of two separate algorithms that can be used to estimate FDOM freshness and source (i.e. terrestrial or marine) in existing glider and towed system datasets. Chapter 4 uses the machine learning methods from Chapter 2 and the algorithms from Chapter 3 to estimate

FDOM freshness and source across a 3-year timeframe. Furthermore, Chapter 4 compares the FDOM dynamics to oxygen data to bridge the gap between the nature of FDOM and the occurrence of hypoxia. Chapter 5 summarizes the results presented in Chapters 2-4, provides conclusions, and offers future research suggestions.

2. DIAGNOSING COASTAL PROCESSES USING MACHINE LEARNING AND OCEAN BUOYANCY GLIDERS

2.1. Introduction

Ocean buoyancy gliders (Stommel, 1989) function as the "satellites" of the subsurface; i.e. collecting a multitude of oceanographic variables to an extent rivaling what earth-orbiting satellites offer for the marine surface layer. Modern Slocum gliders are increasingly being used to advance the understanding of coastal and blue ocean processes (Testor et al., 2019). Gliders function by adjusting their buoyancy in order to rise and fall in the water column, according to pre-programmed instructions or as directed remotely by a glider pilot, to collect vertical profiles of a variety of hydrographic and biogeochemical observations (DiMarco & Knap, 2018). Gliders surface at programmed regular intervals and transmit their collected data via satellite or cellular networks while simultaneously receiving from a glider pilot instructions of way point correction, whether to activate or deactivate individual sensors (e.g. to conserve battery life to help extend the mission duration), and maximum sampling depth (Glenn et al., 2011). Gliders are reusable, can be adapted to different research applications through customization of sensor payloads, and can be deployed for weeks or months, as battery power allows.

Numerous projects have utilized gliders, for example, to enhance understanding of tropical cyclone intensification (Goni et al., 2017), to explore oxygen dynamics and associated biogeochemical processes (Johnson et al., 2009), and vertical mixing impacts

to ocean heat content (Ruiz et al., 2012). In the Gulf of Mexico specifically, glider data have been used to investigate monitoring strategies for coastal hypoxia (Ramey et al., 2017), quantify water column turbulence for improving oil spill forecasts (Wang et al., 2018; Mahmud et al., 2020), and improve predictive capabilities of Loop Current dynamics (Mahmud et al., 2020; Rudnick et al., 2016). Glider data have been used to explore processes in various marine environments. However, the substantial spatiotemporal scales exhibited by quasi-Lagrangian devices, such as gliders, do not readily support precise description of biogeochemical processes within erratic coastal environments like the northern Gulf of Mexico.

The Texas-Louisiana Shelf is a shallow, river-dominated region, situated within the northwestern portion of the marginal Gulf of Mexico, often recognized for its annual intense summer hypoxia (Rabalais et al., 2007) and frequent hurricane landfalls. Hypoxia in the region is controlled by terrestrial nutrient supply and stratification (Bianchi et al., 2010), both of which arise chiefly from nutrient-laden freshwater discharge from the Mississippi-Atchafalaya River System (Bianchi et al., 2010; Rabalais et al., 2007). Nutrients from wind-driven upwelling or autochthonous nutrient supply from microbial remineralization can additionally amplify primary production, thereby further sustaining a hypoxic event (Feng et al., 2012; Feng et al., 2014; Joshi et al., 2015). Rapid intensification of tropical cyclones along the Texas-Louisiana Shelf has been observed to occur as a result of enhanced ocean heat content in the shallow nearshore region (Potter et al., 2019) as well as from tropical cyclone passage over warm-core eddies shed from the Loop Current (Hong et al., 2000; Scharroo et al., 2005).

Additionally, the Texas-Louisiana Shelf is a global leader in oil and gas production, with numerous offshore oil platforms providing a considerable contribution to the regional economy of the Gulf States. Ocean currents within the Texas-Louisiana Shelf vary seasonally with winds (Cochrane & Kelly, 1986; Nowlin et al., 2005) and are a vital consideration for the structural integrity of offshore mineral exploration activities. Tidal currents are generally small (<10 cm/s) owing to the physical geography of the marginal sea (DiMarco and Reid, 1998). Comprehensive understanding of subsurface currents, strengthened in part by the use of gliders among other technology, is critical for mitigating future oil spill disasters as well as for forecasting the movement of oil in the event of a spill.

Submersible gliders can collect millions of observations per 30-day mission, thereby providing unique insight into the subsurface processes that cannot be otherwise explored using satellite or moored observations. However, it can often be difficult to parse out the key data within such immense datasets. Averaging data over the considerable spatiotemporal scales present within glider datasets is not necessarily conducive to quantifying processes with high precision within the volatile environment of the Texas-Louisiana Shelf. State-of-the-art methods of analyzing AUV and towfish data will further clarify the fundamental mechanisms driving both biogeochemical and physical variability within the dynamic environment of the Texas-Louisiana Shelf. Ultimately, these analyses can lead to new approaches for abating anthropogenic impacts to other similar coastal and marine environments.

One technique often used for water mass delineation and endmember tracking in estuarine and coastal environments is to examine the relationship between dissolved organic matter (DOM) and salinity using a scatter plot. Bulk DOM typically exhibits a conservative relationship when plotted against salinity (i.e., an approximately constant linear slope); however, the relationship becomes tenuous in coastal regions where several marine and freshwater endmembers combine to create behavior that is more non-conservative. Numerous other research projects have demonstrated the significance of exploring DOM sources and dynamics in river-dominated continental shelf regions. In the Changjiang River basin and adjacent East China Sea, Gao et al. (2019) used comparisons of salinity and DOM to track water mass mixing and associated nutrient and carbon sources and sinks. Stepanauskas et al. (2002) found that riverine organic material was a significant control of Baltic Sea primary productivity, even suggesting that organic material from rivers influences coastal primary production on the same order as riverine inorganic nutrients. Kim et al. (2019) compared how nitrogen dynamics varied as a function of source (e.g., river vs. atmosphere) in the northern Gulf of Mexico with those of the coastal sea off Korea. Non-conservative mixing of DOM can occur when DOM is created or removed by microbial sources. Another common method of removal of the light-absorbing or "colored" portion of DOM (CDOM) in the photic zone of the water column is through photodegradation of CDOM by sunlight (Vodacek et al., 1997), including in the northern Gulf of Mexico (D'sa, 2008). In this study, we exploit the relationship between CDOM and salinity using an innovative machine learning approach.

Machine learning is an artificial intelligence method of data analysis that employs the use of training datasets or algorithms to enable researchers to recognize patterns and relationships within other datasets and improve their predictive performance over multiple iterations. In essence, machine learning enables a computer software component, for example, to accomplish a task often without being deliberately taught or programmed by a user. Machine learning techniques have been employed in a range of medical and scientific disciplines; for example, in genome sequencing (Libbrecht & Noble, 2015), in cancer prevention (Kourou et al., 2015), weather forecasting (Bauer et al., 2015; Reichstein et al., 2019), and pattern recognition in seismic tremor data (Rouet-Leduc et al., 2019), among many others. Within marine science, machine learning methods have been used to investigate the Gulf of Mexico Loop Current (Oey et al., 2005; Zeng et al., 2015), for modeling the temporal evolustion of the surface salinity field of the TXLA shelf (Nowlin et al. 2005), for modeling coastal hypoxia (Hetland & DiMarco, 2008), and in the development of the regional ocean modelling system (Moore et al., 2004; Shchepetkin & McWilliams, 2005), among other applications (see Ahmad (2019) for a comprehensive review).

One specific machine learning method, the k-means clustering algorithm (Forgy, 1965; Lloyd, 1982), is a type of *unsupervised* machine learning technique by which distinct portions of a dataset are quantifiably differentiated into a user-defined number of groups ("clusters") of points with shared commonality. The k-means algorithm is a non-hierarchical clustering method that determines the location of cluster centroids within a dataset and assigns each point of data to a cluster, depending on the point's respective

Euclidean distance to the cluster centroids (i.e. the closest centroid determines the cluster to which the point belongs). The optimal number of clusters for a given dataset can be determined using the silhouette score method (Rousseeuw, 1986). Silhouette scoring iterates through a user-defined range of cluster numbers (k) and determines what k allows for points to be located nearest (in Euclidean distance) to their respective cluster centroid. This study uses the k-means algorithm to capitalize upon the highly resolved spatiotemporal scales over which submersible gliders survey and is one of the first applications of k-means to a subsurface oceanographic dataset.

Our results demonstrate the value of partitioning a glider dataset into clusters, rather than analyzing an entire glider dataset altogether across time and space. By exploiting the typical linear mixing relationship of CDOM and salinity, we used machine learning clustering analysis of glider data to reveal where in the shelf and to what degree the effects of an intense river flood pulse, wind-driven upwelling, and shifts in wind-driven currents contributed to summertime oceanographic variability along the Texas-Louisiana Shelf. Additionally, we demonstrate how variability within the glider clusters is not uniformly distributed throughout the entire record. Variability is dominated principally by density in the clusters further from the coast, where stable stratification is established as a result of the warmer, fresher water overlying the dense, colder bottom water. Conversely, the variability of the glider clusters nearer to the coast, and thus nearer to riverine nutrient sources, is a function of diurnal periods of phytoplankton biomass and the associated respiration fluctuations. A clear delineation of distinct regions in the water column, both vertically as well as horizontally, further enhances our

knowledge of river impacts upon the biogeochemistry of river-dominated continental shelf regions.

This section is organized accordingly: section 2.2 identifies the data and methodology employed in this study. Section 2.3 outlines the results of the clustering analysis and associated statistical analysis, followed by section 2.4, a discussion and explanation of the results. Finally, section 2.5 provides a concise summary of the chapter and concluding comments.

2.2. Data and Methods

2.2.1. Underwater Autonomous Gliders

Texas A&M University Geochemical and Environmental Research Group (GERG) maintains a fleet of 4 Teledyne G2 Slocum Gliders, each of which is assigned a unique numeric serial number (S/N) identifier (e.g. "Glider 540"). Every glider deployment mission is also given a numeric identifier, beginning in sequential order from the first deployment ("Mission 1") in 2013. Two glider deployments are reviewed in this study: Missions 13 (M13; Glider 540) and 14 (M14; Glider 308), both of which were deployed simultaneously in the northern Gulf of Mexico from 1 – 20 July 2015 (NOAA NCCOS Award No. NA15NOS4780168). The gliders were deployed approximately 14 nautical miles southeast of Galveston Bay and recovered 65 nautical miles southeast of their deployment location. Both Glider 504 and Glider 308 were equipped with a Seabird SBE41CP CTD and a WetLABS ECO Puck fluorometer for measuring chlorophyll-a (chl-*a*), bulk CDOM, and turbidity. Oxygen measurements, however, were collected

using a different sensor aboard each unit: Aanderaa Optode for Glider 540 (M13) and a fast-response RINKO-II for Glider 308 (M14).

Both M13 and M14 each collected > 400,000 data points, followed a near-parallel course to one another (Figure 2.1), and remained in shallow (< 50m) waters throughout their deployment. Bulk CDOM concentration was < 7 mg L^{-1} within 3 standard deviations for both M13 and M14 datasets, thus all analyses were calculated on the subset of CDOM < 7 mg L^{-1} for both gliders. Notably, at the time of deployment, the M14 fluorometer had not been serviced in three years, whereas the M13 instrumentation was purchased and calibrated in the months immediately preceding deployment. The resulting disparity in instrument performance provided an opportunity to test whether the machine learning algorithm could overcome the resulting noise in the M14 dataset in order to identify patterns consistent with those observed by the comparatively cleaner, and more recently serviced, M13 record. Glider data were subject to standard GERG QA/QC protocols, including inspection for in/out range of historical observations from the LATEX Shelf (range checking), outlier removal of unphysical values, in and out of water values, time step, data velocity (i.e., first order differencing in time), calibration checking, etc

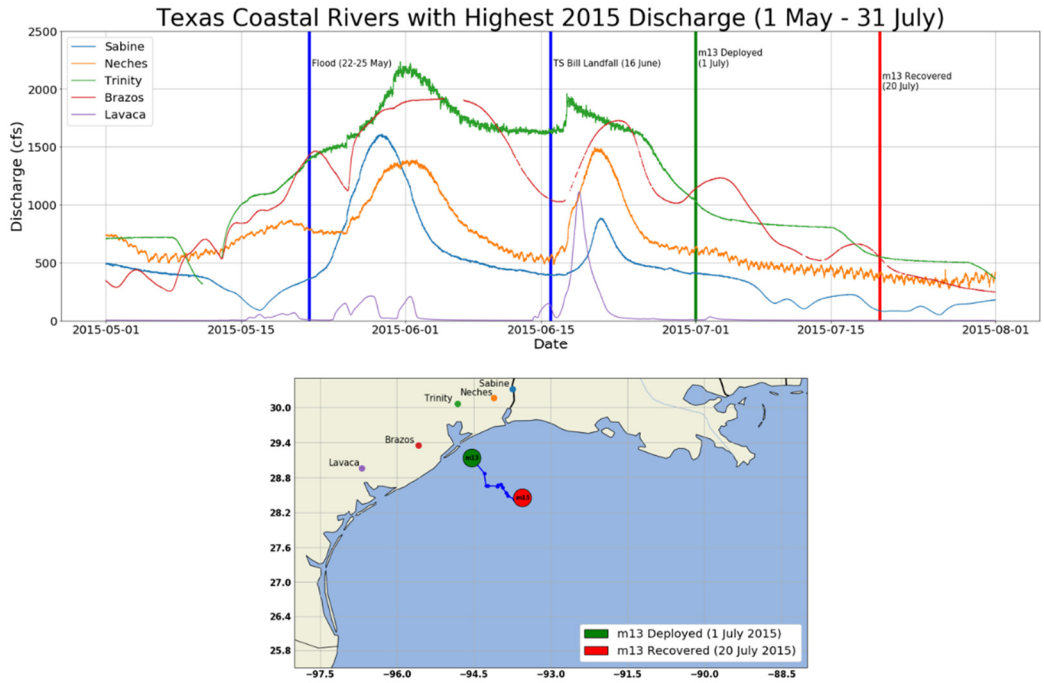

Figure 2.1 (**a**) 1 May - 1 August 2015 discharge rate (m^3 s^{-1}) of the five Texas rivers with highest discharge among all Texas rivers in 2015. Vertical blue lines indicate timing of the 2015 Texas-Oklahoma flood (22-25 May) and Tropical Storm Billy landfall (16 June); vertical green line indicates date of glider deployment (1 July) and red line indicates glider recovery date (20 July); (**b**) map of glider trajectory, green circle indicating location of glider deployment and red the point of glider recovery; location of USGS river discharge meters in Texas are labeled

2.2.2. Machine Learning Clustering

Clustering analysis was accomplished by applying the k-means clustering algorithm separately to CDOM/salinity data from each of the M13 and M14 glider datasets. The only input parameter that was provided to the k-means algorithm was the number of clusters, a value selected based on silhouette scoring of each dataset. A test of the efficacy of the k-means method for analysis of these glider data utilized M13 cluster centroids as training data in a separate iteration of the k-means clustering technique upon the M14 dataset (see Section 2.3.3).

2.2.3. USGS River Discharge

Measurements of river discharge rate of five Texas rivers (Figure 2.1), each of which exhibited the bulk of peak discharge rates amongst Texas rivers in 2015, was obtained from United States Geological Survey (USGS) stations at: the Sabine River (USGS station 8030500), the Neches River (USGS station 8041780), the Trinity River (USGS Station 8067000), the Brazos River (USGS station 08116650), and the Lavaca River (USGS station 08164000). USGS data are provided in $ft^3\ s^{-1}$, but are herein converted and reported as $m^3\ s^{-1}$.

2.2.4. Buoy Data

Measurements of wind speed and direction were obtained from National Data Buoy Center (NDBC) buoy 42035 to support the idea that wind-driven upwelling influenced water mass structure and to help explain the role of wind-driven currents on glider trajectory.

2.2.5. Satellite Sea Surface Height

Sea surface height (SSH) observations, as measured by the NOAA AVHRR satellite sensor, were provided by the Louisiana State University Earth Scan Lab (https://www.esl.lsu.edu/; Nan Walker, personal communication). SSH measurements helped to supplement the buoy data for investigating the potential impact of currents on the glider trajectories and associated observations of water mass structure.

2.3. Results

2.3.1. Dataset Partitioning with the k-means Clustering Algorithm

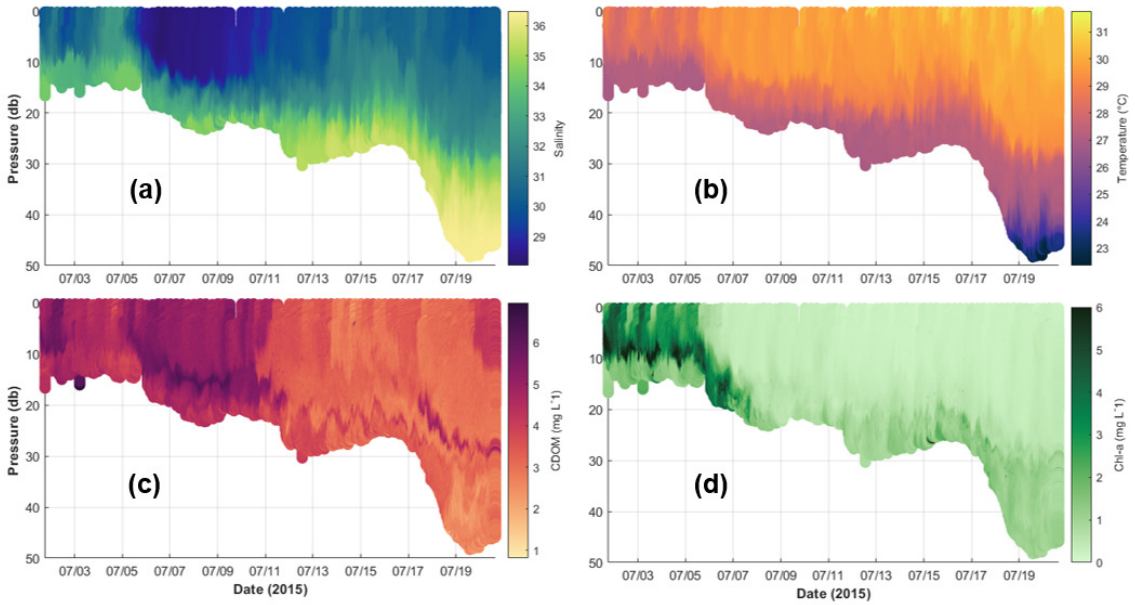

Figure 2.2: Hovmöller time-series of glider M13 (a) salinity, (b) temperature (°C), (c) CDOM (mg L^{-1}), and (d) chl-a (mg L^{-1}) observations.

The relationship between CDOM and salinity reveals linear mixing processes in coastal and estuarine waters. Typically, CDOM and salinity exhibit an inverse linear correlation; i.e. fresher waters generally have higher CDOM concentrations than saline waters. However, the July 2015 glider data (Figure 2.2) were marked by an anomalous departure from the conventional linear CDOM-salinity relationship (Figure 2.3). Pulses of Texas river discharge, brought about by the May 2015 Texas-Oklahoma flood and the June 2015 landfall of Tropical Storm Bill in Texas, presumably introduced a considerably higher concentration of CDOM and freshwater onto the Texas-Louisiana

Shelf than is typical for the summer (Nowlin et al., 1998). The July 2015 glider data were consequently marked by an anomalous departure from the conventional linear CDOM-salinity conservative mixing relationship. In order to explain this irregularity, the k-means algorithm was applied to a scatterplot comparison of CDOM and salinity in order to segment the dataset into distinct clusters; i.e. areas of the dataset wherein CDOM and salinity were inversely correlated were distinguished from those regions wherein a positive or no relationship existed between the two parameters. Each of the individual clusters were then analyzed separately, as well as compared to one another, in order to determine the underlying mechanisms contributing to the observed variability in CDOM-salinity relationship.

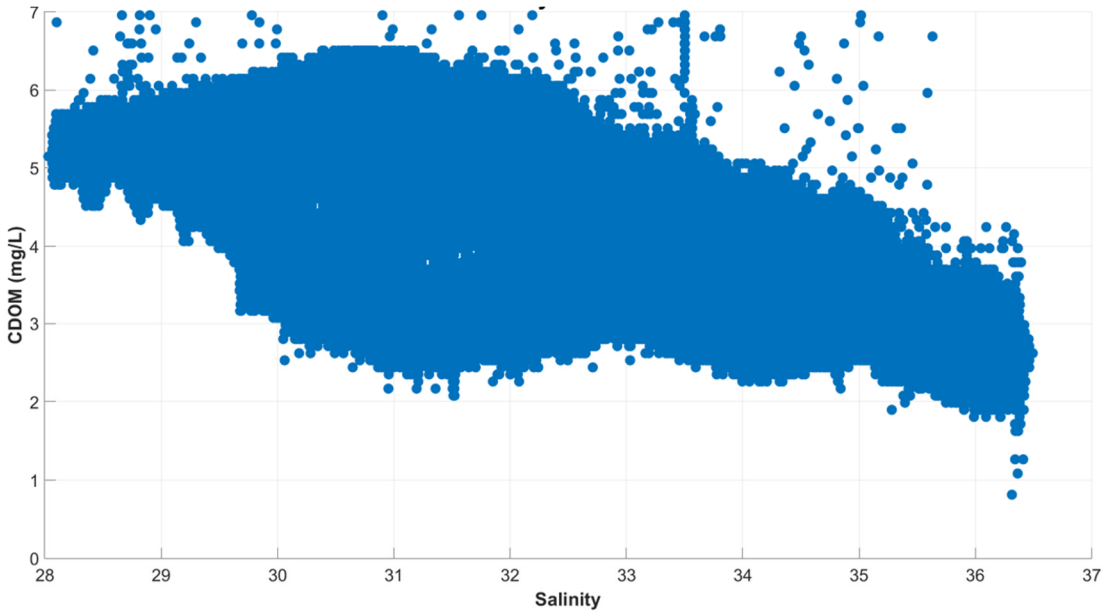

Figure 2.3 Observations of chromophoric dissolved organic matter (CDOM; mg L^{-1}) as a function of salinity for the entirety of the M13 glider record (n = 407,294)

Silhouette scoring recommended five clusters as the optimal cluster number for the M13 datasets (Figure B1), thus the CDOM-salinity scatterplots were partitioned accordingly (Figure 2.4a). Each of the five clusters corresponds to a unique region in the water column through which the glider traveled (Figure 2.4b): two nearshore clusters (nearshore surface cluster = NSC: green, nearshore bottom cluster = NBC: purple), two offshore clusters (offshore surface cluster = OSC: yellow, offshore bottom cluster = OBC: orange), and the pycnocline cluster (PYC: blue) which spans the length of the time-series and separates the OSC from the OBC.

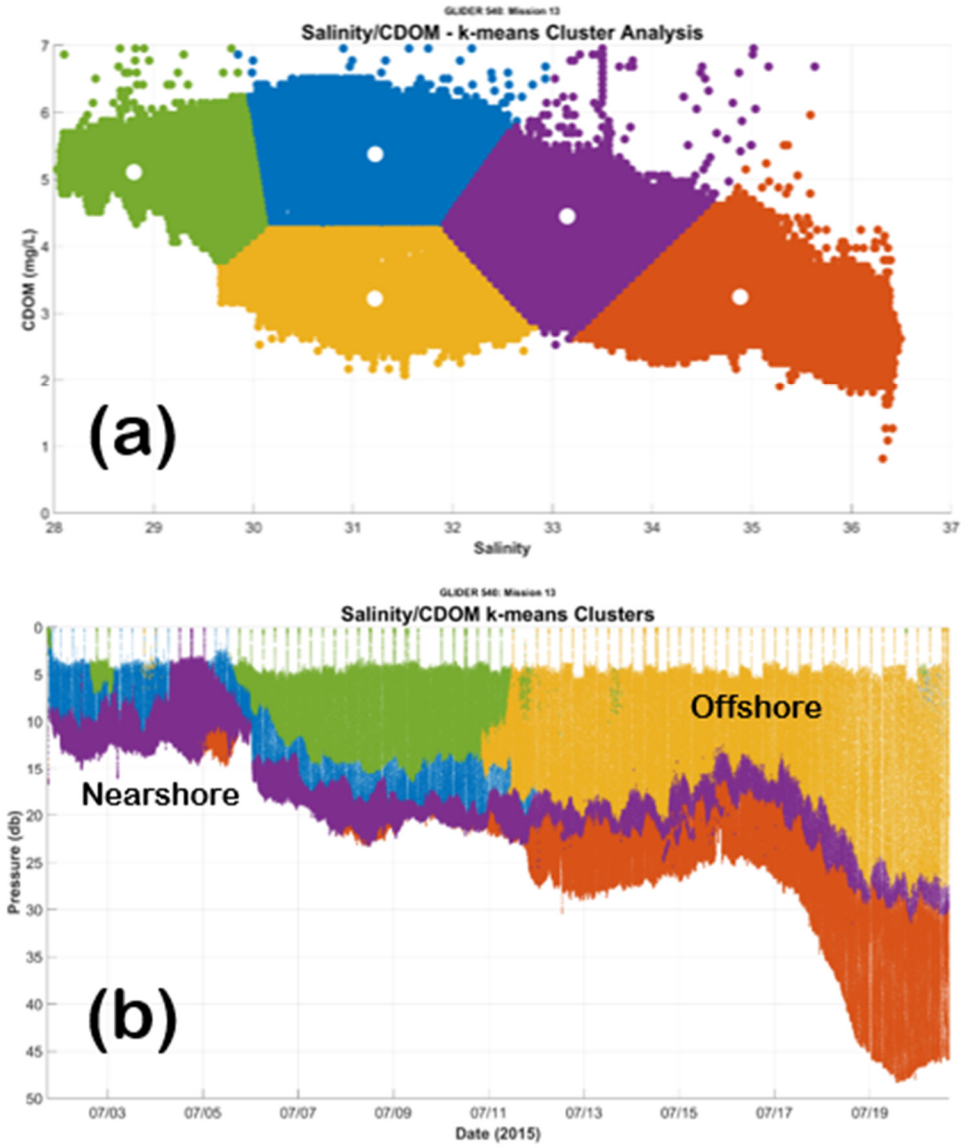

Figure 2.4 (**a**) Result of the k-means clustering process upon observations of chromophoric dissolved organic matter (mg L^{-1}) as a function of salinity for the entirety of the M13 glider record (n = 407,294), individual points are colored according to cluster classification, white circles indicate cluster centroid; (**b**) Time-series of M13 glider observation depth as a function of time (note that ocean bottom is ~1-2m deeper than the bottom observations on this figure); points are colored according to their cluster classification as depicted in Figure3a

Neither of the nearshore clusters (Fig 2.3, blue and green) penetrated greater than ~20m depth. The NSC in particular (corresponding to the area of lowest average salinity) did not penetrate below ~16m depth. The PYC depth varied as a function of distance from the coast and local bathymetry, averaging ~14m deep across the entirety of the glider record. The OBC extended to the deepest of M13 observations at ~48m. However, some OBC points were also identified in the nearshore at depths as shallow as ~9m (and will be analyzed separately). Both CDOM and dissolved oxygen concentration were observed to be relatively higher in the nearshore clusters in comparison to offshore clusters (Table 2.1). Similarly, average salinity was lower in the nearshore clusters.

Table 2.1: Mean oxygen and CDOM for each of the five M13 k-means clusters; green highlight indicates cluster with highest mean concentration.

Cluster	**Nearshore**		**Offshore**		**PYC**
	NSC	**NBC**	**OSC**	**OBC**	
Oxygen (mg/L)	5.81	6.08	5.63	5.26	5.93
CDOM (mg/L)	5.11	5.37	3.21	3.24	4.45

An increase in continuous wind speed, from ~3 to ~9 m/s, was observed at NDBC Buoy 42035 over an eleven-hour timespan from 3-4 July. The timing of the increase in wind speed corresponded to the only observation of the clustering algorithm designating waters shallower than ~8 meters as part of the PYC. The wind speed increase likewise corresponded to the shallowest observations of the OBC at ~9m depth.

A shift in continuous wind direction at NDBC Buoy 42035 from predominantly easterly (~88° average from 1-12 July) to predominantly southerly (~175° average from 12-20 July) corresponded to a northward shift in glider trajectory. Prior to the wind direction change around 12-13 July, the glider moved across-shelf in the nearshore (depth < 20m), but thereafter was diverted along-shelf upcoast (toward Louisiana) in the

offshore region (maximum depth from 20-48m) until the end of its deployment. The shift in wind direction was coupled with observations of decreasing SSH slope with increasing latitude (from ~10cm below MSL at 29°N to ~50cm below MSL at 29.5°N) in the region of the glider deployment. Both observations of wind direction and SSH slope decrease are consistent with previous observations of directional shift in the Texas Coastal Current in the summer (Cochrane & Kelly, 1986; Cho et al. 1998; Nowlin et al., 2005; Walker, 2005). The change in glider direction with the shift in the Texas Coastal Current was also coincident with changes in cluster classification; the offshore clusters, bisected by the PYC, dominate both the cluster-partitioned time-series following the Texas Coastal Current transformation.

2.3.2. Variability Among and Within Clusters

The five cluster groups derived from the k-means clustering algorithm were used as treatment groups within an analysis of variance (ANOVA) test wherein the glider parameters not used for cluster classification (dissolved oxygen, chl-*a*, temperature, turbidity, and sample depth) were used as dependent variables. Effect sizes (ω^2) were also calculated for each parameter (Table 2.2). The 95% confidence interval of ω^2 did not include 0 for any parameter (Nakagawa & Cuthill, 2007), thus demonstrating the effectiveness of the clustering technique in establishing clusters that are statistically distinct. Temperature and depth exhibited the largest ω^2 of the parameters (0.76 and 0.57, respectively), which, when coupled with the salinity observations, indicates that the primary distinguishing factor across all five clusters is density.

Table 2.2: Effect size (ω^2), 95% ω^2 confidence interval, and F-statistic and associated degrees of freedom for ANOVA test for each parameter using M13 k-means clusters as treatment groups; green highlight indicates parameters with highest ω^2.

ANOVA Parameter	F-statistic $_{df}$	Effect size			
		ω^2	Min	Max	95% CI
Oxygen	9,928 $_{4, 100,670}$	0.2828	0.2784	0.2872	0.0088
Chlorophyll-a	51,953 $_{4, 407,289}$	0.3378	0.3357	0.34	0.0043
Temperature	321,740 $_{4, 407,289}$	0.7596	0.7586	0.7606	0.002
Turbidity (10^{-3})	30,149 $_{4, 407,289}$	0.2284	0.2263	0.2306	0.0043
Depth	136,690 $_{4, 407,289}$	0.5731	0.5714	0.5748	0.0034

Following ANOVA, principal component analysis (PCA) of z-scored data was used in order to illuminate to what degree the variability within each cluster could be attributed to the different glider-measured oceanographic parameters. PCA was performed on both the M13 dataset in its entirety (i.e. irrespective of clustering, n = 407,294) as well as on each cluster individually (i.e. 5 clusters, 5 separate PCA). PCA of the entire dataset (i.e. no clustering) indicated that 90% of variance was attributed to Components 1 and 2. Component 1, accounting for ~70% of the overall variance, found density to be the dominant process as temperature and salinity had the highest principal component scores (inverse of one another; Figure 2.5). Chl-*a* was a secondary control to Component 1, but was the dominant process identified within Component 2, signifying the substantial variability in primary production observed within the dataset. Notably, Component 2 turbidity was inverted with respect to chl-*a*, not surprisingly confirming that light availability was a significant constraint upon primary productivity.

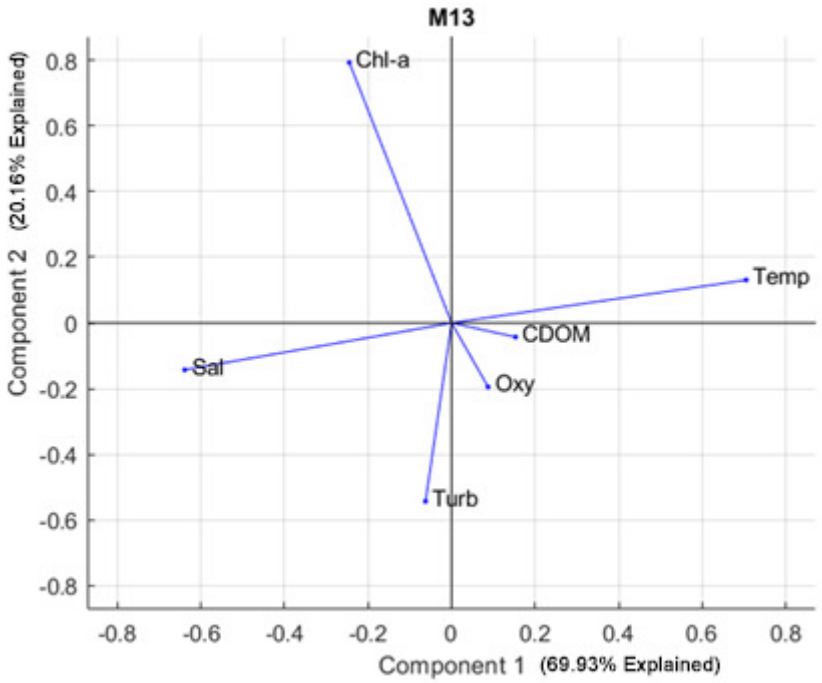

Figure 2.5 Biplot of Component 2 loadings (20.16% variance explained) as a function of Component 1 loadings (69.93% variance explained) derived from the principal component analysis of the composite M13 glider record for the six glider parameters: salinity (Sal), chlorophyll-*a* (Chl-a), turbidity (Turb), dissolved oxygen (Oxy), chromophoric dissolved organic matter (CDOM), and water temperature (Temp)

While density differences explained variance within the glider record as a whole, results from the cluster-specific PCAs found that dominant processes differed depending on glider depth and distance from the coast (Figure 2.6). Nearshore clusters, for example, were marked by the variable levels of microbial activity, while density dominance upon offshore cluster variability mirrored the PCA observations of the comprehensive dataset (i.e. not cluster-specific). While microbial activity in the nearshore was attributed to variance in the rate of primary production, CDOM and oxygen (notably inverse of one another, suggesting microbial production of CDOM) were a more significant control to offshore variance, with chl-*a* and turbidity effectively

absent in the offshore PCA. Additionally, no apparent correlation between microbial activity and density was indicated by PCA within any of the five clusters. Furthermore, the first two principal components combined to account for at least 85% of variance in all clusters and as much as 93% in the OSC and NBC. The Component 3 explained variance did not exceed 9% (NSC) for any of the five clusters; therefore, the effects of the first two principal components represent the most prominent processes within each respective cluster.

Within Component 1 of both the NBC (Component 1 = 83% variance explained) and the PYC (Component 1 = 75% variance explained), primary productivity, as represented by chl-*a,* was the overwhelming influence upon variability as indicated by the elevated chl-*a* score (Figure 2.6d, 2.6b). As with the comprehensive dataset PCA, turbidity was inverse of chl-*a*, further substantiating the well-established importance of light availability to primary productivity. While temperature and salinity loadings were effectively absent from Component 1, density scores dominated within Component 2 for both the NBC and PYC.

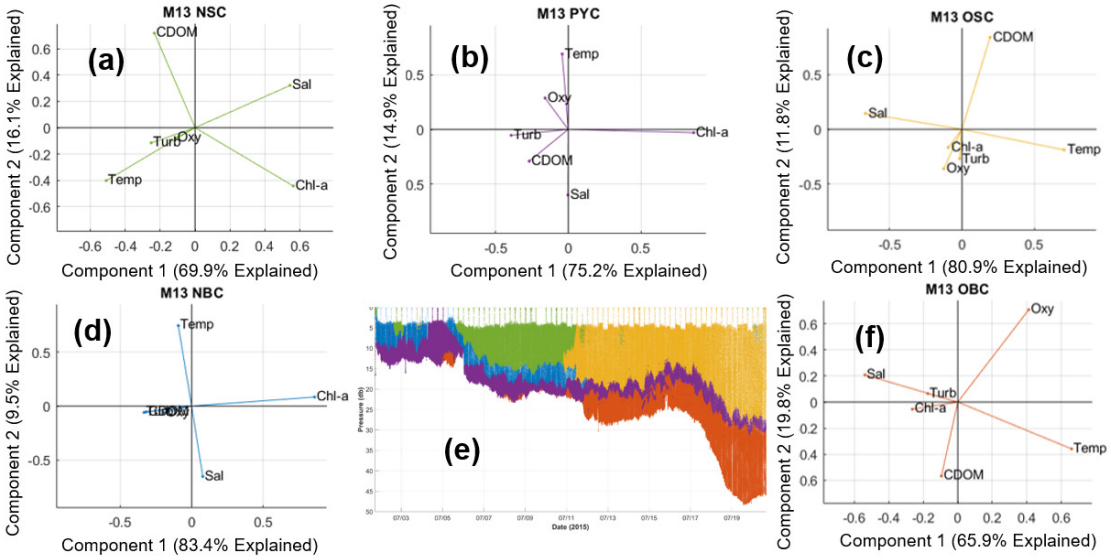

Figure 2.6 Biplot of Component 2 loadings as a function of Component 1 loadings derived from the principal component analysis of (**a**) the M13 nearshore surface cluster (NSC), (**b**) the M13 pycnocline cluster (PYC), (**c**) the M13 offshore surface cluster (OSC), (**d**), the M13 nearshore bottom cluster (NBC), and (**f**) the M13 offshore bottom cluster (OBC), for the six glider parameters: salinity (Sal), chlorophyll-*a* (Chl-a), turbidity (Turb), dissolved oxygen (Oxy), chromophoric dissolved organic matter (CDOM), and water temperature (Temp); biplot lines in Figure 5a-d, f are colored according to (**e**) cluster classification

The offshore cluster PCAs (Figure 2.6c, 2.6f) demonstrated a nearly direct contrast to that of the NBC and PYC. Density, rather than primary production, dominated across Component 1 for both the OBC (Component 1 = 66% variance explained) and OSC (Component 1 = 81% variance explained) while the effects of microbial activity, as estimated by CDOM and oxygen relationships, were instead evident within Component 2 (OBC = 20% variance explained; OSC = 12% variance explained). However, the effects of microbial activity were attributable to different processes in the offshore. Across Component 2 in both offshore clusters, CDOM scored prominently while chl-*a* was essentially absent, a notable contrast to the prominence of phytoplankton biomass

observed within the NBC and PYC. A distinguishing feature of the OBC PCA was the prominence of dissolved oxygen across both Components 1 and 2. Within OBC Component 1, oxygen was inverse with respect to salinity, implying ventilation or physical controls. Oxygen was inverted with respect to CDOM within Component 2, indicating that microbial respiration varied as a function of organic matter availability (i.e. increases in organic material concentration encouraged oxygen consumption, and vice versa).

The NSC represented a buoyant mass of relatively fresher water that was likely derived from coastal rivers (Zhang et al., 2012) as a result of heavy flooding in Texas in the weeks preceding glider deployment (Figure 2.1a). Neither density nor microbial activity was more prevalent within the NSC; instead, their respective contributing impacts to variance were comparable across both NSC Components 1 and 2 (Figure 2.6a). Within the NSC, chl-*a* was inversely correlated to CDOM, rather than turbidity (i.e., light availability), revealing the different ongoing dominant processes within the NSC as opposed to the NBC and PYC. Therefore, NSC phytoplankton biomass was modulated by microbial remineralization of river-borne organic material.

2.3.3. Assessment of Model Effectiveness

While increased data scatter (i.e., noise, as noted in Section 2.2.1) rendered the M14 glider data unsuitable for the same unsupervised clustering that was performed upon the M13 dataset, the similarity between and close proximity of the M13 and M14 glider trajectories (Figure 2.1b) provided a unique opportunity to test the overall efficacy of the

k-means clustering technique for glider data analysis. Many similarities were observed in a comparison of the resulting M14 clustering (Figure 2.7) to that of the M13 data. The parallels were found not only in terms of the structural composition of the water column through which the gliders translated (i.e. offshore clusters, nearshore clusters, and PYC), but also within the underlying processes contributing to variance within the resulting clusters.

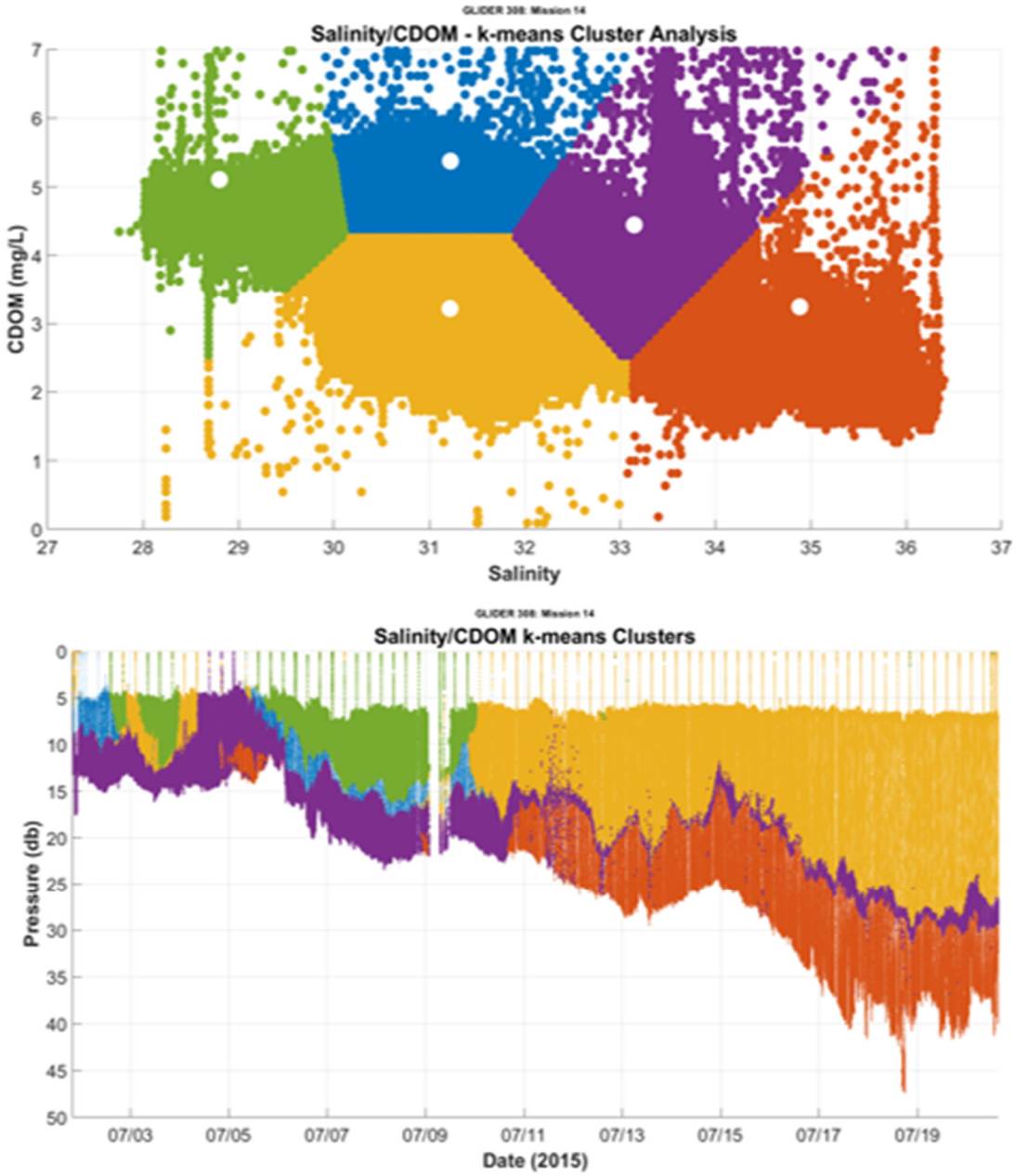

Figure 2.7 (**a**) Result of the k-means clustering process upon observations of chromophoric dissolved organic matter (mg L^{-1}) as a function of salinity for the entirety of the M14 glider record (n = 445,500), individual points are colored according to cluster classification, white circles indicate M13 cluster centroids; (**b**) Time-series of M14 glider observation depth as a function of time; points are colored according to their cluster classification as depicted in Figure 2.6a

The only significant difference in the physical makeup of the water column of the M14 dataset was a thinner PYC in the offshore as well as the classification of some of the nearshore points as belonging to the offshore clusters. This disparity reflects that the M13 centroids were each located at points ~0.5 mg L^{-1} higher on the M14 scatterplot than M13, thus the offshore clusters (i.e. lower CDOM concentration) encroached upward on the cluster scatterplot. However, despite these discrepancies and the increased noise within the M14 dataset, the clustering technique was again consistent in its performance, as confirmed by the overall structural composition of the water column (i.e. nearshore vs. offshore clusters, deep vs. shallow clusters) as well as accompanying statistical analysis. ANOVA effect size (ω^2) 95% confidence intervals again did not include 0 for any parameter. Temperature and depth again exhibited the largest ω^2 of the parameters (0.82 and 0.69, respectively), reflecting the importance of density in cluster classification. PCA results showed density as a stronger control of variance in the offshore than in the nearshore, but microbial activity, as evidenced by highly variable chl-*a* and oxygen, was the chief control on variability in all five M14 clusters.

2.4. Discussion

2.4.1. Physical and Biogeochemical Processes Elucidated through Dataset Clustering

Significant river influence in the Texas-Louisiana Shelf region introduces high CDOM, low salinity water into the marine environment (Nowlin et al., 1998; Nowlin et al., 2005). Over time, physical mixing of river water compounds with biogeochemical

processes, including photodegradation and microbial remineralization, to decrease CDOM concentration and increase salinity. The addition and removal of CDOM in this manner typically results in an inverse relationship between salinity and CDOM in estuarine and coastal environments.

In 2015, however, river flood pulses brought about by the May Texas-Oklahoma flood and the June landfall of Tropical Storm Bill in Texas introduced a considerably higher concentration of CDOM and freshwater onto the Texas-Louisiana Shelf than is typical for the summer. These anomalous flood events modified the CDOM-salinity relationship such that the typical inverse relationship was not well-defined within the agglomerative record of two gliders that were deployed in July 2015 on the Texas-Louisiana Shelf. While portions of the CDOM-salinity scatterplots exhibited an overall inverse relationship, other sections (e.g., portions of the nearshore clusters) reflected a potentially positive interaction (i.e., increasing CDOM with increasing salinity). In order to help explain this deviation from the traditional inverse CDOM-salinity relationship that characterizes the Texas-Louisiana Shelf coastal environment, the k-means clustering algorithm was employed. A scatterplot of M13 CDOM-salinity observations was subdivided into five clusters each of which provided a more precise insight into the physical mixing processes on the shelf. When the glider time-series data are colored according to their respective k-means-derived cluster classification, the physical composition of the shelf water column is clearly discernible.

Two offshore clusters, each respectively representing the shallow and deep portions of the water column, are divided by a distinctive pycnocline. The observed physical

dynamic of the offshore clusters reflects the stable stratification of the offshore water column that fits well with climatological expectations of density gradients in the shallow Gulf of Mexico in summertime. The nearshore clusters, while still distinguishable, do not exhibit the same defined vertical layering as those offshore. Instead, the nearshore clusters were marked by pronounced variations in microbial activity, owing to the substantial influence of material brought about by the elevated river discharge. The freshwater pulses also combined with the effects of winds to create a variable environment in the nearshore, which inhibited the establishment of vertical layering, unlike that which was observed further offshore (in > 20m waters). The upcoast shift of the Texas Coastal Current around 12-13 July, as substantiated by satellite-measured SSH and shifts in wind direction at NDBC buoy 42035, was coupled with an increase in horizontal current velocity which served to advect the CDOM-laden freshwater out of the gliders' path and helped to produce the density-driven stratification observed offshore. A composite review of these observations indicates that the clustering algorithm resolved not only the river influence, but also effects of wind-induced upwelling in the nearshore. On 3-4 July, for example, elevated wind speeds at nearby NDBC buoy 42035 corresponded to the shallowest observations of the PYC within the time-series.

Furthermore, while the clustering technique produces a clear image of the physical makeup of the shelf water column, the influence of the underlying biogeochemical processes are further elucidated by cluster-specific PCA. While PCA indicated that density differences (i.e. temperature and salinity) explained the bulk of the variance in

the offshore clusters (with the notable contribution of oxygen and CDOM upon bottom-water variability in the offshore), microbial processes dominated in the nearshore and the PYC. In particular, chl-*a* was the overwhelming influence upon PCA loading in the NBC and PYC, wherein the latter of which turbidity loading was inverted with respect to chl-*a*, indicating the importance of light availability upon production in the PYC. Within the NSC, however, PCA loadings of both density and microbial activity were nearly equivalent. Additionally, an inverse correlation of NSC loadings of CDOM and chl-*a* suggested microbial remineralization was the leading constraint on primary productivity therein. Therefore, the majority of riverine inorganic material uptake likely occurred in the NBC and PYC, whereas remineralization of CDOM promoted primary production in the NSC. These observations, combined with scarce primary production offshore, hint at the horizontally-layered structure of the northern Gulf of Mexico coastal zone, as suggested by Rowe and Chapman (2002) and Kim et al. (2020).

The simultaneous deployment of two gliders, whose respective trajectories were nearly parallel to one another, allowed for an aptitude assessment of the k-means clustering technique. The glider M13 cluster centroids generated by the k-means algorithm were selected as the starting input parameters around which M14 clusters were constrained. The M14 dataset, although fraught with noise as a result of instrumentation complications, still resolved the same physical structure of the total shelf water column: two offshore clusters, vertically divided by a pycnocline which was itself continuous throughout the time-series, and two nearshore clusters. The utility of employing the k-

means clustering technique for glider data is underscored by its consistency in resolving the same physical and biogeochemical features in both the M13 and M14 datasets.

One notable disparity between the M13 and M14 datasets was the observation of a spike in M14 CDOM concentration between 33 and 35 salinity, the bulk of which was concentrated at the deeper layers of the pycnocline cluster in the nearshore. The physical location of elevated CDOM within the water column reveals that the most vigorous microbial activity was found at depth. The M14 CDOM spike is evidence of the dynamic and abrupt changes that can occur on short spatiotemporal scales in a shallow coastal marine environment. Both M13 and M14 followed a near-identical path throughout their deployment and yet M14 encountered CDOM conditions that were substantially elevated in the nearshore. The CDOM pulses were coupled with a reduction in oxygen concentration (Figure B6), signifying that the CDOM spike was attributable to microbial activity and not simply river input of allochthonous CDOM material. Ultimately, these observations indicate that inferential spatial statistic approaches, which estimate biogeochemical conditions over large spatial scales from a select few discrete samples (as opposed to the >400,000 glider observations per glider in this study) do not capture the profound variability that occurs in coastal environments.

2.5. Summary and Conclusion

The k-means clustering algorithm was used to divide two large datasets of glider observations (M13 n = 407,294; M14 n = 445,500) into five clusters per dataset, each of which is representative of a distinct region of the shelf water column both in depth and

in relative distance from the coast. Cluster partitions correspond well to expectations of shelf dynamics based on known climatological patterns as well as observed atmospheric and hydrodynamic variables. Wind measurements from NDBC Buoy 42035, located near the location of glider deployment, were also used in concert with satellite SSH observations to explain the patterns in cluster distribution. Wind-induced upwelling, for example, was potentially observed from 3-4 July when a significant increase in wind speed at 42035, over a period less than 12 hours, corresponded to the only surface observations of the PYC, which was otherwise located in the mid-to-lower water column throughout the remainder of both M13 and M14 time-series. Shifts in predominant wind direction from easterly to southerly (meteorological convention) occurred from 12-13 July, at which time SSH slope was found to be decreasing northward (from ~10-50 cm above MSL), both of which observations fit with climatological expectations of summertime shifts in the Texas Coastal Current (Cochrane & Kelly, 1986). However, these observations may be circumstantial, as the time scales of shelf upwelling are generally longer than a few hours (Kämpf & Chapman, 2016). The shift in current was revealed both in the pattern of the gliders' forward translation as well as by the k-means clustering algorithm, which identified a change in cluster type coincident with the timescale of the wind direction change.

ANOVA indicated that the clustering approach was successful in dividing the glider data into unique regions that were distinguishable from one another in terms of the processes contributing to their physical and biogeochemical variability. PCA further elucidated the variance within clusters, finding that variance within the entire glider

datasets (i.e. irrespective of clustering) is predominantly a function of density. However, cluster-specific PCA found that clusters in the nearshore are more variable with respect to microbiology, being governed by disparity arising from diurnal periods of minima and maxima in chl-*a* and oxygen, whereas offshore cluster variability is density-controlled. Offshore, however, oxygen was also significantly impactful to bottom water variance, an observation that has profound implications for the understanding of Texas-Louisiana Shelf oxygen, which is often believed to remain static throughout the summer months.

The simplicity of the k-means clustering algorithm results in a lower resolution of cluster boundaries as compared to other more complex clustering algorithms, such as DBSCAN. However, k-means clustering does provide a useful approach for demarcation of a dataset into discrete regions in space and time. Each of the clusters identified by the k-means clustering help to explain how oceanographic processes vary as a function of time, depth, proximity to the coast, or with respect to anomalous atmospheric or hydrographic events (e.g. wind or river discharge extremes).

Future research applications could utilize *in-situ* measurements of CDOM, for example, in concert with glider observations in order to characterize CDOM source (i.e. allochthonous vs. autochthonous) throughout a given region of the water column. These methods are applicable to other regions of the world. Additional applications could also include gliders equipped with hyperspectral sensors tuned to frequencies associated with hydrocarbons in order to monitor weathering and transport dynamics of oil spill material. Clustering methods could also be used to identify areas where new CDOM formation

occurs or to draw connections to spatiotemporal variability in primary production or hypoxic events.

3. MODELING SOURCES OF GULF OF MEXICO DOM USING HYPERSPECTRAL CDOM FLUORESCENCE TO DIAGNOSE COASTAL WATER QUALITY: APPLICATION TO OCEAN BUOYANCY GLIDER DATA

3.1. Introduction

Characterizing the sources and sinks of dissolved organic matter (DOM) in the marine environment, particularly in river-dominated continental shelves, is critical for advancing the understanding of global carbon flux (Osburn et al., 2016), ocean acidification (Cai et al., 2011; Kealoha et al., 2020), harmful algal blooms (Bargu et al., 2019; D'Sa, 2008; Hetland & Campbell, 2007), and hypoxia (Bianchi et al., 2009; Rowe & Chapman, 2002). The relative degree of contribution of rivers to the regional DOM budget can be estimated through investigation of the chromophoric (or "colored") portion of DOM (CDOM). CDOM is a component of DOM that absorbs and/or fluoresces light in the ultraviolet-visible (UV-VIS) portion of the electromagnetic spectrum (200-700 nm). CDOM can be measured *in-situ* using fluorometers, which themselves can be deployed aboard autonomous underwater vehicles (e.g. ocean buoyancy gliders), towed profiling systems, or on shipboard CTD instrument panels. High-resolution multispectral analysis of discrete water samples additionally allows for the qualification of the CDOM source (i.e. allochthonous vs. autochthonous, marine vs. terrestrial, etc.) as well as its chemical composition (Boyd & Osburn, 2004; Coble, 1996, 2007) and relative molecular weight (Fichot & Benner, 2012; Helms et al., 2008). In the Greenland Sea, Stedmon & Markager (2001) demonstrated the utility of optical

measurements for differentiating between terrestrial and marine CDOM. Analysis of CDOM dynamics has also been used: to estimate water mass mixing (D'Sa & DiMarco, 2009; Gao et al., 2019; Tehrani et al., 2013); to research the potential impacts of climate change on the global DOM pool (Zhou et al., 2018); as a proxy for dissolved organic carbon (Gao et al., 2020; Vodacek et al., 1995, 1997); and even for explaining habitat selection in sharks (Haulsee et al., 2015).

CDOM in the marine environment originates either from allochthonous (generally terrestrial) or autochthonous (locally created) sources. In addition to seasonal variation, the quantity and quality of allochthonous vs. autochthonous CDOM often varies as a function of proximity to rivers or estuaries, which can export substantial amounts of organic material to the coastal ocean. CDOM dynamics additionally vary because of microbial uptake and/or excretion; sedimentary and hydrothermal vent releases; and variations in the degree of solar irradiance. Solar photodegradation is the primary removal process of CDOM from the marine environment (Bianchi et al., 2009; Coble, 2007) but creation of CDOM by photochemical means can also occur (Stedmon & Nelson, 2015). Photodegradation acts preferentially on CDOM rather than colorless DOM (Yang et al., 2021) in large part because of its absorption of UV light from the sun (Helms et al., 2008).

The UV-VIS absorption spectrum of CDOM is essentially a featureless, exponential function that decreases with increasing wavelength. Following absorption (i.e. "excitation") at a particular wavelength, some CDOM compounds emit light (i.e. "fluoresce") at a wavelength longer than the corresponding excitation wavelength. The

resulting fluorescence spectrum also has broad, featureless peaks. The fluorescence properties of CDOM (i.e. "fluorescent DOM" or "FDOM") are commonly used to explore DOM dynamics. An excitation-emission matrix (EEM) compiles the fluorescence intensities of a water sample across a range of excitation wavelengths, typically from ~240-700 nm. The FDOM of a sample can be distinguished according to its approximate chemical composition (e.g. protein-like or humic-like) or source (e.g. freshwater or marine) based on its observed EEM fluorescence intensity (Coble, 1996, 2007).

Letter designations are often used to refer to the wavelengths of commonly observed FDOM fluorophores (Coble, 1996, 2007; Hansen et al., 2016). For example, peak C, observed at the excitation wavelength of ~340 nm and corresponding emission at ~440 nm, generally refers to FDOM of terrestrial origin. Similarly, peak M (found at excitation/emission of ~300/390 nm) identifies FDOM likely created within the marine environment, from either algal-derived material or other microbial processes, for example. The ratio of peak C:M (CMR) fluorescence intensity therefore provides an estimate of the relative amount of terrestrial to marine material present within a sample. Terrestrial FDOM would produce a CMR value > 1, marine FDOM would be < 1, and finally an equal share of FDOM from terrestrial and marine sources would yield a CMR value of exactly 1.

Another metric of FDOM fluorescence, the Biological Index (BIX), is the ratio of excitation/emission intensity at 310/380 nm (protein-like FDOM) to that at 310/430 nm (biological-like; Huguet et al., 2009). BIX measures the relative amount of recently

created ("fresh", not to be confused with "freshwater") autochthonous FDOM within a sample and is therefore sometimes referred to as FDOM "freshness" (Haywood et al., 2018). A larger BIX value signifies that the sample is comprised of a more sizable fraction of new material, thus implying recent autochthonous microbial activity, while a low BIX value suggests predominantly older and presumably more recalcitrant material is present. The BIX was originally developed to characterize DOM within an estuary (Huguet et al., 2009); however, the BIX has been applied to continental shelf environments research as well. For example, Carr et al. (2019) used BIX to quantify the microbial contribution of fresh material to the dissolved organic carbon pool in the Celtic Sea of the Northwest European Shelf. Chari et al. (2013) estimated phytoplankton production and associated grazing pressure through comparisons of BIX to fluorescence peak intensities; they also noted that CDOM fluorescence can be a more informative measure of phytoplankton production than chlorophyll-*a* (chl-a). In the northern Gulf of Mexico, the BIX has been used to quantify the rates of autochthonous DOM production triggered by diversion of Mississippi River water into the Lake Pontchartrain Estuary (Haywood et al., 2018; Kolic et al., 2014) and to explain the role of terrestrial DOM in air-sea carbon transfer (Bianchi et al., 2013).

The Texas-Louisiana shelf is a shallow, river-dominated region, situated within the northwestern region of the Gulf of Mexico. Seasonally variable discharge from freshwater sources, such as the Mississippi-Atchafalaya River System, fuels substantial phytoplankton blooms in the spring and autumn, which can in turn lead to large-scale fluctuations in regional water quality. However, the extent to which allochthonous

terrigenous material from coastal rivers and estuaries, as opposed to autochthonous production within the marine environment, modulates the biogeochemistry of the Texas-Louisiana shelf continues to be explored (Bianchi et al., 2010; DiMarco et al., 2010; Kim et al., 2019, 2020). Although many studies have underscored the significance of fluvial nutrient input as a control of northern Gulf of Mexico bottom-water hypoxia (Rabalais et al., 2007), other research has demonstrated that the physical structure of the water column is as critical to regulating hypoxia, if not more so (Feng et al., 2012; Fennel et al., 2011; Forrest et al., 2011). Although disputed by some researchers (Das et al., 2011), it has been suggested by others that northern Gulf of Mexico microbial processes are not simply a function of river discharge; instead, most riverine material is taken up near the source, while repeated sequences of microbial remineralization of DOM can fuel new production at increasing distances from the river mouth (Bianchi et al., 2009; Rowe & Chapman, 2002). While much of the research of CDOM dynamics in the northern Gulf of Mexico has been focused within or immediately adjacent to Galveston Bay (D'Sa et al., 2018; B. Liu et al., 2019; Gold-Bouchot et al. 2021) and the Mississippi-Atchafalaya basin (D'Sa & DiMarco, 2009; Fichot & Benner, 2012; Osburn et al., 2016), this study aims to explore CDOM across the Texas-Louisiana shelf.

Fluorometers are commonly used aboard research vessels and AUVs to measure ocean DOM concentrations. While such fluorometers provide an estimate of bulk CDOM concentration, they typically only measure excitation/emission at one wavelength pair and, as such, are limited in their ability to ascertain specifics of the nature of the CDOM that they measure. In this study, we combine the high-precision

capabilities of discrete spectrofluorometric sample analysis methodologies with the substantial data collection capacity of AUVs to demonstrate a more comprehensive regional mapping and understanding of marine FDOM dynamics. More specifically, we used EEM spectroscopy to derive the peaks C and M fluorescence intensity (from which we calculated CMR) and BIX for many northern Gulf of Mexico water samples. The resultant findings were compared to coincident measurements of other ocean parameters to derive a predictive algorithm for estimating BIX and CMR within existing glider datasets. This study is one of the first examples to use glider data to explain FDOM dynamics at a large scale in the northern Gulf of Mexico. The methods developed in this study could be applied to any other similar river-dominated continental shelf regions to support the determination of water mass or constituent material source and, therefore, lead to more precise quantification of riverine impacts to marine biogeochemistry and advance the understanding of the relative roles of transport and transformation processes of the continental shelf.

3.2. Data and Methods

3.2.1. In Situ FDOM Measurements

Water samples used as training data for FDOM algorithm development were collected aboard two separate cruises in 2019: from the July cruise, a total of 69 samples, and 187 from the October 2019 cruise (Figure 3.1). Dissolved oxygen concentration was analyzed according to the Winkler titration method while underway (Strickland & Parson, 1972). Samples were filtered at the point of collection from the CTD Niskin

bottle using a 0.2 μm syringe filter and stored in pre-combusted 20 mL disposable glass scintillation vials in a deep freezer aboard the ship. EEMS processing of the samples was accomplished following each cruise using a Horiba Aqualog model 800-UV spectrometer. The specific wavelengths extracted from the cruise data EEMS were, for BIX, excitation/emission 310/380 nm and 310/430 nm; and, for CMR, excitation emission 340/440 nm and 300/390 nm.

Fluorescence values were corrected by subtracting a blank (ultrapure water reference material Starna XXX), and correcting inner filter effect (self-absorption) using the Aqualog software. Fluorescence intensity results were transformed to Raman Units using the Raman peak area at λ_{exc}= 350 nm using the staRdom package (Reference) for the R statistical environment (R Core Team 2018).

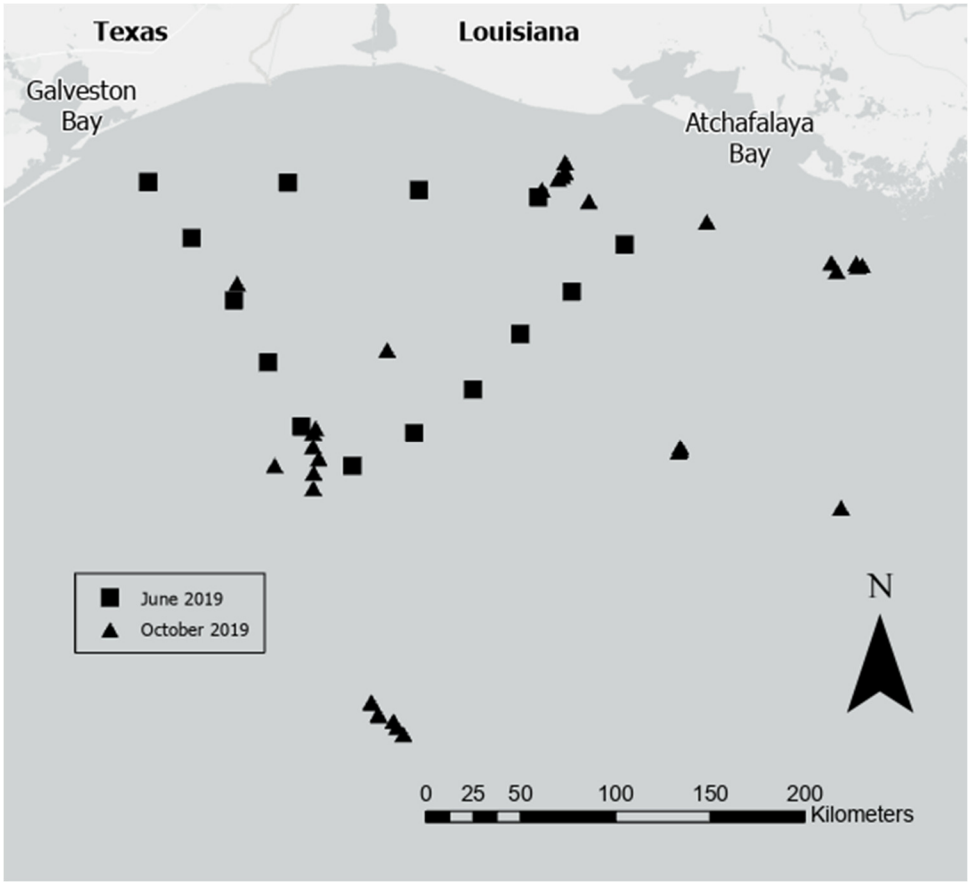

Figure 3.1: The northern Gulf of Mexico with study sampling stations from June 2019 (black squares) and October 2019 (black triangles).

3.2.2. Ocean Buoyancy Gliders

Glider data were supplied by the Texas A&M University Geochemical and Environmental Research Group (GERG). Each GERG glider mission is designated a numeric mission identifier. GERG gliders sample at 1 Hz frequency, undulating through the water column at 25-50 cm/s, and taking anywhere from 7-10 min to complete an undulation in 30 m water depth (Ramey et al., 2017). Ultimately, a 30-day glider mission could result in over half a million points of data collected. Glider mission deployment M13 took place from 1-20 July 2015 in the northern Gulf of Mexico using GERG glider

#540. M13 collected oceanographic data using a Seabird SBE41CP CTD and a WetLABS ECO Puck fluorometer. The ECO Puck measures chl-*a*, CDOM, and turbidity. In particular, the wavelengths used for CDOM fluorescence are 370/460 nm excitation/emission. Dissolved oxygen measurements were collected using an Aanderaa Optode.

3.2.3. FDOM Algorithm Development

Multiple linear regression analyses were performed on cruise FDOM data to create separate algorithms capable of predicting BIX (Figure 3.2) and CMR (Figure 3.3). Various terms were used for each model, most of which were collected using the shipboard CTD sensor instrument suite. The sole exception was that apparent oxygen utilization (AOU) was calculated using dissolved oxygen data from Winkler titrations. Additionally, "spiciness" (hereafter "spice") was calculated from salinity and temperature according to the methods of Flament (2002). All of the parameters used in the derived algorithms are either direct measurements or are derivable from direct observations that are routinely deployed as part of a CTD/rosette system and/or glider instrumentation package.

No defined pattern existed between either BIX or CMR and any other parameters, thus stepwise regression with interaction terms were used for modeling. Residuals showed no patterns (i.e., no heteroscedasticity) for either the BIX nor CMR algorithms (Figures B2 and B3).

The BIX predictive algorithm (Eq. 1) includes 8 terms derived via stepwise regression, resulting in an adjusted R-squared of 0.842 (P = 3.88e-32):

$$y = 0.83 + 0.06 * x1 + 0.04 * x2 + 0.03 * x3 + -0.03 * x4 + -0.007 * x5 + -0.01 * x6 + 0.002 * x7 + -0.009 * x8 \quad (1)$$

where y is estimated BIX, x1 is spice, x2 is bulk CDOM (excitation/emission 370/460), x3 is chl-a, x4 is AOU, x5 is spice*bulk CDOM, x6 is bulk CDOM*chl-a, x7 is depth*AOU, and x8 is spice*AOU.

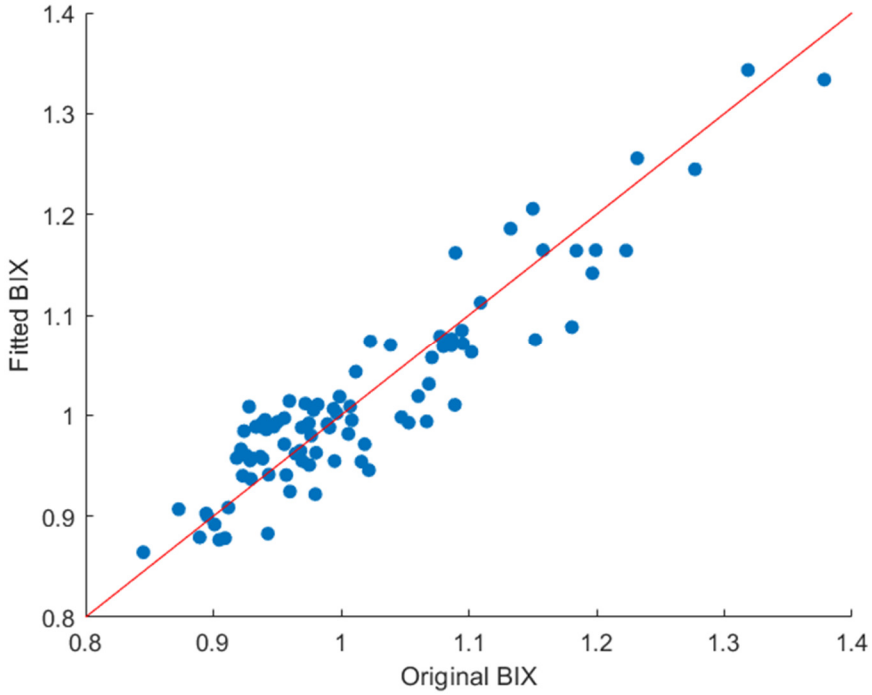

Figure 3.2: Direct BIX measurements from fluorometer readings of cruise samples vs. predicted BIX using the new algorithm (Eq. 1).

The CMR predictive algorithm (Eq. 2) includes 13 terms derived via stepwise regression, resulting in an adjusted R-squared of 0.64 (P = 3.18e-17):

$$y = 20.56 + -0.05 * x1 + -2.05 * x2 + -0.02 * x3 + 0.03 * x4 + -0.51 * x5 + -0.66 * x6 + 0.004 * x7 + -0.0018 * x8 + 0.001 * x9 + 0.012 * x10 + 0.057 * x11 + 0.016 * x12 + 0.015 * x13 \quad (2)$$

where y is estimated CMR, x1 is depth, x2 is bulk CDOM (excitation/emission 370/460), x3 is chl-a, x4 is AOU, x5 is salinity, x6 is temperature, x7 is depth*bulk CDOM, x8 is depth*AOU, x9 is depth*salinity, x10 is bulk CDOM*salinity, x11 is bulk CDOM*temperature, x12 is salinity*temperature, and x13 is bulk CDOM squared.

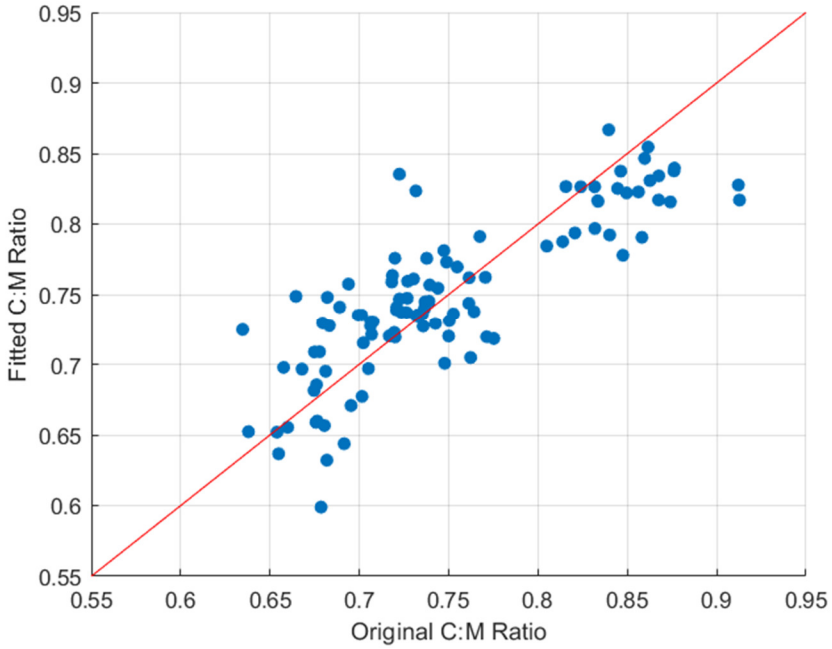

Figure 3.3: CMR values derived from direct measurements of Peak C and Peak M from fluorometer readings of cruise samples vs. predicted CMR using the new algorithm (Eq. 2)

3.3. Results

3.3.1. Observations of Modeled Glider FDOM

The newly derived FDOM algorithm was used to estimate BIX and CMR within the M13 glider dataset (Figure 3.4). Average modeled BIX for the M13 glider dataset was 1.1167 (σ = 0.0443). BIX values generally increased with depth and glider distance from the Texas coast. Modeled M13 BIX data reveal that the largest BIX values (i.e. freshest DOM) were generally found in the deepest waters of the glider dataset (> ~30m), in which were also found larger observations of AOU (Table 3.1), indicating active respiration and remineralized carbon. Conversely, while BIX was also found to increase with increasing salinity, the lowest BIX values were found in the intermediate salinity range of ~31-32. This region of lowest BIX was encountered at the beginning of the glider mission deployment, within the waters nearest to the Texas coast, but prior to entering the surface freshwater plume (salinity ~28-30) that was encountered from 6-11 July.

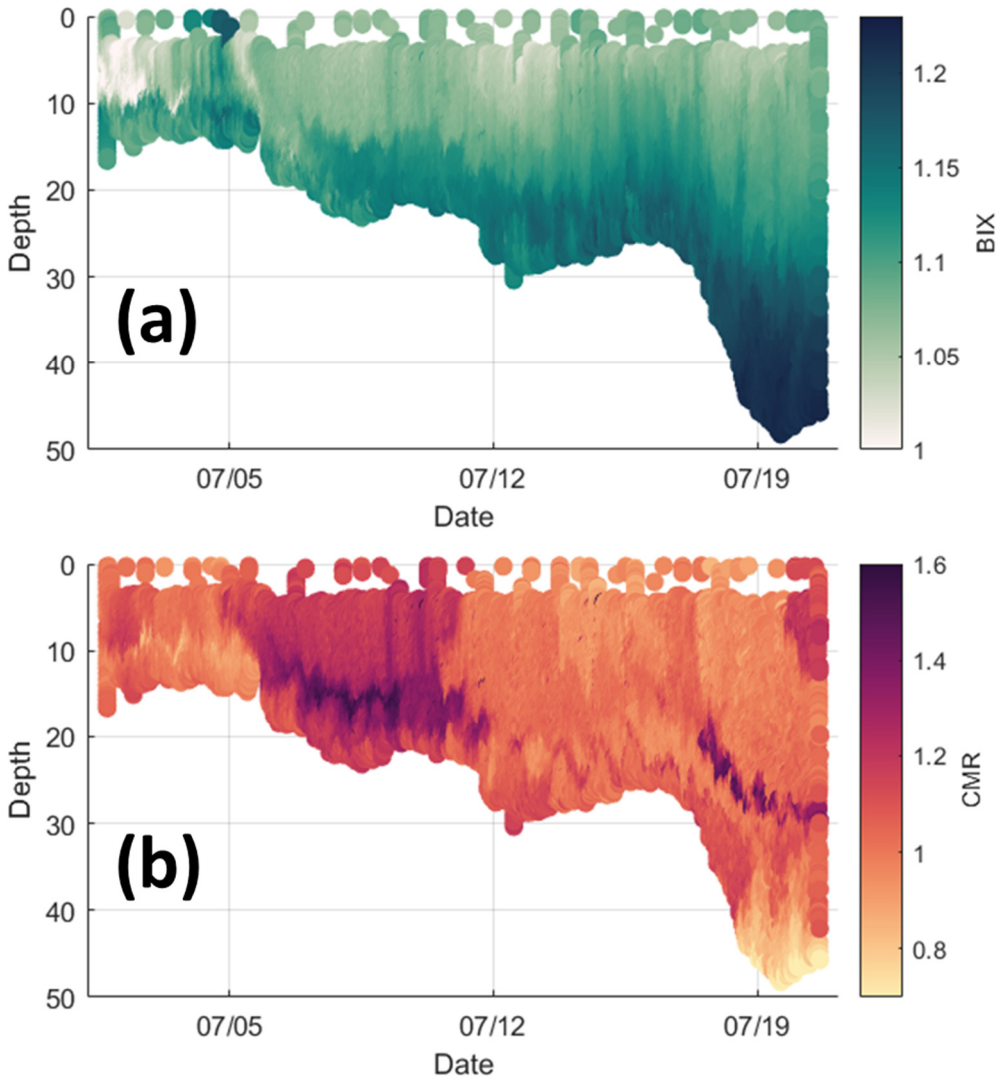

Figure 3.4 (a) Hovmöller time-series of M13 modeled BIX; (b) Time-series of M13 modeled CMR.

Table 3.1: Minimum and maximum modeled BIX and CMR values, and the corresponding depths at which and dates in 2015 during which they were found.

	M13 Modeled FDOM	**Depth (m)**	**AOU (ml L^{-1})**	**Date (2015)**
	BIX			
Max BIX	1.2305	42.41	1.4711	19 July
Min BIX	0.8415	10.20	0.2941	2 July
	CMR			

Max CMR	1.8844	19.09	0.3378	14 July
Min CMR	0.6066	45.84	2.1096	20 July

Mean M13 CMR was 1.0792 ($\sigma = 0.1374$). Minimum CMR values (i.e. implying less terrestrial, more marine FDOM) were observed in the deepest waters, concurrent with some of the highest AOU values observed (Table 3.1). The largest CMR values were found within the pycnocline at the base of the 6-11 July freshwater plume. An additional peak in CMR values was located within the pycnocline between 20-30m depth from 18-21 July.

Analysis of the modeled FDOM parameters becomes more illustrative when the M13 glider dataset is partitioned into clusters (Figure 3.5), as shown in Chapter 2. For example, the relationship between M13 dissolved oxygen and modeled BIX is low for most of the glider dataset; however, a positive correlation ($R^2 = 0.49$) was found between dissolved oxygen and BIX in the OSC. Similarly, a sharp decrease in BIX was observed with increasing chl-a within the NBC and PYC ($R^2 = 0.79$ and 0.52, respectively). The strongest positive relationships between salinity and BIX were found in the offshore clusters. While this positive relationship was mostly constant in the OSC, decreases in OBC BIX values resulted in a nonlinear relationship between salinity and BIX therein; a similar relationship was also observed in the NSC. Within the OBC, BIX decreased with increasing bulk CDOM ($R^2 = 0.36$), whereas a positive relationship between the two was found within the NSC ($R^2 = 0.31$). The only cluster in which BIX and turbidity were closely related was the NBC ($R^2 = 0.44$), indicating light availability

was not a significant control on freshness. Overall, the OBC had the highest mean BIX while the NSC had the lowest followed closely by the NBC (Table 3.2).

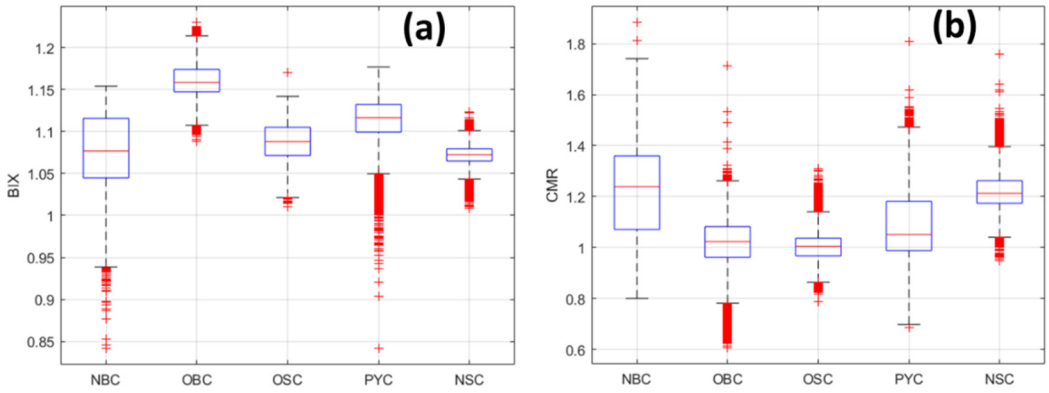

Figure 3.5: Boxplot of median and interquartile range of M13 modeled (a) BIX and (b) CMR, each grouped according to glider cluster.

Table 3.2: M13 cluster-specific average (and standard deviation) BIX and CMR.

M13 Cluster	Mean (σ) BIX	Mean (σ) CMR
NBC	1.074 (0.04)	1.23 (0.17)
OBC	1.16 (0.02)	1.02 (0.09)
OSC	1.09 (0.02)	1.00 (0.06)
PYC	1.11 (0.02)	1.08 (0.13)
NSC	1.07 (0.01)	1.22 (0.08)

Whereas Table 3.1 may seem to imply that BIX and CMR are simply a mirror of one another, the cluster-specific relationships between these modeled parameters and the other M13 observations do not show such a trend. For example, a relationship between CMR and either dissolved oxygen or salinity was not found within any of the clusters. Similarly to BIX, CMR decreased with respect to increasing chl-a within the NBC and PYC ($R^2 = 0.56$ and 0.18, respectively), albeit less significantly than with BIX. CMR increased positively with increasing bulk CDOM in all five clusters, most significantly within the OSC and NSC ($R^2 = 0.62$ and 0.59, respectively). Finally, likewise to BIX, the

only cluster in which CMR and turbidity were moderately related was the NBC (R^2 = 0.35). Among all clusters, CMR was highest in the NBC and NSC with the lowest CMR found in the OSC and OBC (Table 3.2).

3.3.2. Relationship between Modeled BIX and CMR

Inspecting the relationships of the modeled BIX and CMR parameters to one another can also yield further understanding of the processes measured by the M13 glider. A positive relationship exists between BIX and CMR within the NSC and NBC (R^2 = 0.67 and 0.50, respectively). Otherwise, no other significant relationships between BIX and CMR were observed in any of the other clusters. However, comparison of the standardized BIX and CMR values (Figure 3.6a) reveals an additional dynamic within the M13 dataset. Positive standardized BIX values were found almost exclusively beneath the pycnocline, whereas positive standardized CMR values were found both above and below the pycnocline. The positive BIX/negative CMR quadrant (Figure 3.6, purple) featured the highest mean AOU and lowest dissolved oxygen of the BIX/CMR standardized quadrants (Table 3.3). The highest mean chl-a observations were found within both negative BIX quadrants (Figure 3.6, orange and gold), with the lowest found in the positive BIX/positive CMR quadrant (Figure 3.6, blue). The glider bulk CDOM values somewhat mirrored CMR, with the largest mean bulk CDOM values found within the positive CMR quadrants (Figure 6, blue and gold), with the largest mean found within the positive CMR/negative BIX quadrant (Figure 6, gold).

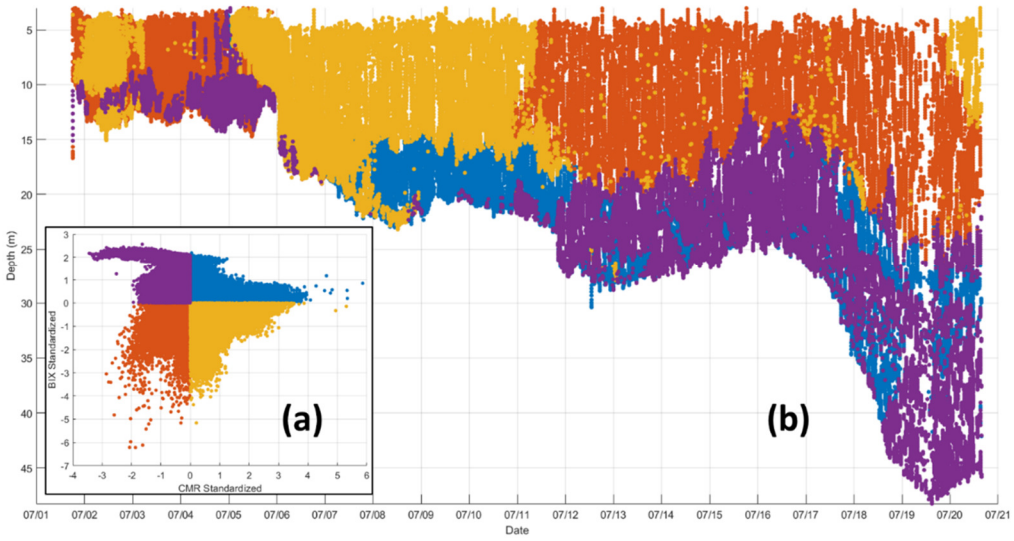

Figure 3.6 (a) Standardized CMR vs. standardized BIX, segmented into quadrants: purple colors correspond to positive BIX/negative CMR, blue colors to positive BIX/positive CMR, orange colors to negative BIX/negative CMR, gold colors to negative BIX/positive CMR. (b) M13 Hovmöller diagram where colors correspond to those described in (a).

Table 3.3: Mean values for M13 glider parameters within each of the standardized BIX vs. CMR quadrants

FDOM Characterization	Quadrant	Mean Salinity	Mean AOU	Mean DO	Mean Chl-a	Mean Bulk CDOM
New/Marine (Purple)	+BIX, -CMR	34.45	0.71	3.85	1.18	3.24
New/Terrestrial (Blue)	+BIX, +CMR	33.67	0.66	3.88	0.98	4.32
Older/Marine (Orange)	-BIX, -CMR	31.86	0.50	4.03	1.76	3.86
Older/Terrestrial (Yellow)	-BIX, +CMR	30.60	0.42	4.13	1.93	5.17

3.4. Discussion

M13 BIX values were largest beneath the pycnocline and further from the Texas coast, as evidenced by the fact that the OBC had the highest mean BIX among all clusters, suggesting that most of the DOM creation happened further from freshwater endmembers. It is also conceivable that DOM created above the pycnocline was photodegraded out of the system (Vodacek et al., 1997; Chen & Gardner, 2004; D'Sa

2008). However, the areas of larger BIX concentrations coincided with areas of greatest oxygen demand (larger AOU and perhaps where active respiration is occurring), whereas above-pycnocline waters were typified by lower or, at times, even negative AOU (i.e., supersaturated waters indicative of photosynthesis processes). The link between AOU and BIX suggests that the freshest DOM was produced by heterotrophic activity, rather than as a byproduct of presumably active primary production processes. Indeed, a positive correlation between BIX and AOU ($R^2 = 0.13$, $p << 0.05$), as compared to negative for BIX vs. chl-a ($R^2 = 0.12$, $p << 0.05$), is consistent with the idea of heterotrophic controls on BIX concentrations; although these determination coefficients are low, the large size ($n > 400{,}000$) of the M13 dataset suggests that even such small determination coefficients as these should be considered significant.

Whereas AOU and BIX values were positively correlated when reviewing the M13 dataset, different mechanisms are apparent when analyzing per cluster. For example, a negative relationship between BIX and AOU was found in both the OSC ($R^2 = 0.51$, $p << 0.05$) and NSC ($R^2 = 0.16$, $p << 0.05$), suggesting that fresh DOM (i.e. larger BIX) in the surface is more rapidly consumed than created. Although OSC chl-a concentrations were relatively low (mean chl-a = 0.32, $\sigma = 0.1211$), a positive relationship between chl-a and BIX was found in the OSC ($R^2 = 0.23$, $p << 0.05$), indicating that some of the surface DOM creation could have been associated with phytoplankton byproducts.

Deviations from the BIX vs. salinity conservative mixing line in the OBC (also the cluster with the highest mean AOU) hint that the deep offshore waters were possibly the location where fresher DOM was readily consumed by heterotrophs. Contrasting

relationships between bulk CDOM and BIX in the OBC and NSC speak to the potential nature of DOM in these clusters. Within the OBC, where more bulk CDOM was found, it was less fresh, perhaps older and more recalcitrant material as the freshest material was being more rapidly consumed. Within the NSC, more bulk CDOM corresponded to larger BIX, suggesting that much of the CDOM in the nearshore surface region was comprised of fresher material. Recalcitrant material was either sinking out of the surface or the freshest material had not yet had time to be photodegraded as had less fresh DOM. The inverse relationship between BIX and turbidity in the NBC implies that the freshest DOM was present in the least turbid waters, likely because more turbid waters were flush with inorganic sediments rather than FDOM.

Like M13 BIX, M13 CMR values were variable with respect to the pycnocline. Apart from higher values (implying a terrestrial DOM source) within the 6-11 July freshwater plume from the surface to ~15m, the largest CMR values were found about and within the pycnocline, mostly in the nearshore as supported by the higher mean CMR in the nearshore clusters. The summertime Gulf of Mexico pycnocline represents an especially strong, stable vertical density gradient which inhibits mixing (Zhang et al., 2020; Bianchi et al., 2010); therefore, sinking material accumulated at the pycnocline.

Photodegradation is the principal means of removal of FDOM from the ocean. Given the increased intensity and duration of sunlight in the summertime, photodegradative processes were accelerated. The DOM that was held aloft within the shallower and surface waters above the pycnocline was, therefore, likely previously photodegraded, resulting in decreased CMR values above the pycnocline. The two distinct CMR peaks

(at the base of the freshwater plume from 6-11 July and within the pycnocline from 18-21 July) suggest the glider encountered waters from distinct freshwater endmembers.

Cluster-specific analysis reveals additional mechanisms for CMR variability. . For example, negative relationships between CMR and chl-a in the NBC and PYC signify that photosynthesis was likely reduced in waters with larger CMR concentrations. Larger CMR values may imply the presence of DOM that absorbs light closer to the wavelengths required by primary producers; thus, photosynthesis was light-limited due to absorption by DOM. The positive correlations between bulk CDOM and CMR found within each cluster speak to the similarity in the glider fluorometer CDOM-measuring wavelength (excitation/emission at 370/460nm) and that of peak C fluorescence (excitation/emission at 340/440nm). Finally, as with BIX, the negative correlation between CMR and turbidity in the NBC indicates that turbid waters contained more marine material, presumably resuspended from the seafloor.

A reasonable expectation in the ocean is that the freshest DOM (larger BIX) would be associated with lower CMR values (Huguet et al. 2009; Salve et al. 2012) given that larger CMR values would imply a more terrestrial (and therefore relatively older in the ocean) DOM signal. However, the lowest BIX values were not coincident with the highest CMR values, which suggests that some of the terrestrial DOM that was observed was created more recently and had been introduced into the marine environment by the 2015 river flood pulses. Alternatively, a portion of the terrestrial DOM was possibly recalcitrant, whereas labile fresher material was consumed more rapidly, thus leading to the disparity. Ultimately, there was no correlation between BIX and CMR over the full

M13 glider dataset; however, different relationships were apparent within some of the various glider clusters. The strong positive relationship between BIX and CMR in the NSC and NBC clusters suggests that much of the DOM in the nearshore was comprised of fresh, terrestrial DOM. The river flooding that preceded the M13 glider mission caused fresher, terrestrial material to be exported to the ocean rapidly (within weeks) enough that, although it was more terrestrial (larger CMR), it was still relatively new (large BIX).

Positive standardized BIX values were concentrated almost exclusively beneath the pycnocline; the two quadrants featuring positive standardized BIX values also had increased oxygen demand (Table 3), again demonstrating the role of heterotrophic activity in producing the observed BIX concentrations. Similarly, the negative standardized BIX quadrants also featured greater mean chl-a concentrations, again echoing the notion that fresh DOM was not associated with byproducts of primary production. The NSC itself was devoid of any positive standardized BIX values or negative standardized CMR values. The OSC was dominated almost entirely by negative standardized BIX and CMR values, likely a combined result of photodegradation of DOM in the upper layers as well as advection of DOM-deficient waters from the south via the recently shifted TXLA current. Most of the positive standardized CMR values were found within and at the base of the NSC; however, additional agglomerations of positive standardized CMR found between 1-3 July and 20-21 July indicate the presence of waters potentially originating from other freshwater endmembers than that which was responsible for the 6-11 July freshwater plume (i.e. the NSC). The quadrants featuring

positive standardized CMR values also featured larger mean bulk CDOM concentrations, which indicates that the "CDOM" as measured by the glider fluorometer is more closely associated with terrestrial CDOM than DOM of other origins. The link between glider "CDOM" and peak C fluorescence is unsurprising given that the glider measures "CDOM" at 370/460 nm excitation/emission wavelengths, which is closer to the wavelengths representing peak C fluorescence (340/440 nm excitation/emission) than those of peak M (300/390 nm excitation/emission). Finally, the highest salinity quadrant was the positive standardized BIX/negative standardized CMR quadrant whereas the lowest salinity quadrant was the negative standardized BIX/positive standardized CMR quadrant, demonstrating that fresher, marine DOM was found in the saltiest water, whereas fresher water featured somewhat older, more terrestrial DOM.

The general trend of increasing BIX further from the Texas coast and within deeper waters highlights the riverine influences along the TXLA shelf. River-borne DOM was present in significant enough amounts either (a) to inhibit creation of new DOM in lower salinity waters or (b) to mask the presence of fresher DOM. Although the 6-11 July freshwater plume was located over ~30 nm SE of the mouth of Galveston Bay, river impacts are indicated by the highest CMR values (i.e. largest percentage of terrestrial material).

These observations indicate that the biogeochemistry of the TXLA shelf can be dominated by riverine sources, even during the summer, well after the peak river discharge of the spring months and after the summer current reversal (Cho et al. 1996; Cochrane and Kelly 1986). Furthermore, summer river influences are often intensified in

this region because it borders an area of the Gulf of Mexico coast that is particularly prone to tropical cyclone activity (DiMarco et al., 1995; Bender et al., 2007; Potter et al., 2019). Enhanced flooding from the 2015 May Texas-Oklahoma flood and Tropical Storm Bill likely resulted in many of the FDOM changes observed by glider M13.

3.5. Summary and Conclusion

Freshwater discharge into the marine environment can increase because of high precipitation levels from tropical cyclones, such as those observed in association with Tropical Storm Bill in 2015, which can in turn lead to significant alterations in salinity and biogeochemistry. This study used both discrete sample analysis of fluorescence and AUV data from an underwater buoyancy glider to describe FDOM dynamics across the TXLA shelf in the northern Gulf of Mexico following river flood events in July 2015. Discrete samples were used to produce algorithms for predicting BIX (a metric of FDOM freshness) and CMR (the ratio of terrestrial vs. marine DOM). Each of the derived algorithms were applied to the July 2015 glider dataset to compare the modeled FDOM with glider observations of other oceanographic data (e.g., salinity, dissolved oxygen, chl-a, etc.).

The results of this study showed that heterotrophic activity influenced autochthonous FDOM production offshore and beneath the pycnocline. Above the pycnocline, FDOM creation was linked to chl-a concentration, hinting that fresh FDOM in the upper water column was in part comprised of phytoplankton byproducts. Larger ratios of terrestrial to marine material were found in waters associated with plumes from freshwater

endmembers. Additionally, phytoplankton activity appeared to be hindered by the excess terrestrial material, which absorbs light at wavelengths nearer to that of chl-a pigments than does marine FDOM.

This study produced an enhanced analysis of FDOM dynamics that could benefit researchers interested in coastal DOM dynamics and transformations, and possible influences to coastal processes, including hypoxia. Our results demonstrate clear relationships between oxygen demand and FDOM in the northern Gulf of Mexico region. The negative correlation between chl-a and BIX in the NBC and PYC suggests that any DOM byproducts of photosynthetic activities may have been either: (a) rapidly consumed, although this is unsupported by the AOU data in these clusters; (b) heterotrophic consumption of phytoplankton both produced fresh DOM while simultaneously reducing chl-a; or (c) that waters replete with fresh DOM were otherwise resource-limited for primary producers (Quigg et al., 2011) such that photosynthetic activities were hindered in some way. In any case, the connections found between chl-a and modeled FDOM suggest the need for a more detailed examination of the role of DOM in processes related to primary production and photodegradation.

Future research could utilize these methods to derive similar FDOM algorithms of other coasts and oceans or apply these algorithms to other glider or towed system datasets. The data for this study came from two cruises in 2019: a summertime (July) cruise and an autumn (October) cruise. Therefore, future applications would benefit from collecting data from additional seasons to obtain a more complete picture of DOM dynamics. Additionally, AUV fluorometers could be equipped to measure fluorescence

at wavelengths that could directly measure the FDOM properties addressed in this study. Enhanced modeling of DOM dynamics throughout the ocean could help to resolve how DOM freshness varies temporally and relative to salinity gradients. For example, future applications could determine how DOM freshness differs about the pycnocline separately in the western vs. eastern regions of the Texas-Louisiana shelf. The regional disparity in DOM freshness could establish a quantitative approximation of the relative freshwater and organic matter contributions of the Mississippi, Atchafalaya, and major northern Texas rivers (e.g. Brazos and Trinity Rivers). The relationship of DOM freshness to dissolved oxygen would also show how respiration modulates DOM dynamics differently at the beginning versus at the end of the hypoxia season. While much of the research focus regarding river impacts to the TXLA shelf is concentrated around nutrients, this study has demonstrated the relative importance of multiple sources of DOM to the transformation and fate of biogeochemical processes of the northern Gulf of Mexico.

4. AN INVESTIGATION OF ORGANIC MATTER FRESHNESS AND SOURCE ON A RIVER-DOMINATED CONTINENTAL SHELF

4.1. Introduction

The shallow region of the northwestern Gulf of Mexico that borders the US states of Texas and Louisiana is commonly known as the Texas-Louisiana (TXLA) shelf. River influences, particularly from the Mississippi-Atchafalaya River System (MARS), play a critical role in modulating the biogeochemistry of the TXLA shelf. Seasonally-variable ocean currents, including the Texas Coastal Current, have also been shown to have a pronounced impact on TXLA shelf biogeochemistry (Cochrane & Kelly 1986; Nowlin et al. 2005; Dissertation Chapter 2). Resource-rich water from the MARS promotes seasonal algal blooms, the dissolved organic matter (DOM) byproducts of which have been recognized as one of the causes of hypoxia in the northern Gulf of Mexico region (Bianchi et al. 2010). However, the degree that allochthonous material, e.g., terrestrial material supplied directly from rivers to the marine environment, regulates TXLA shelf biogeochemistry, including in the development of hypoxia, has remained elusive (Bianchi et al., 2010; DiMarco et al., 2010; Kim et al., 2019, 2020). Outside of the estuarine and near-coastal environment, the dynamics of TXLA shelf DOM continue to be explored (Bianchi et al., 2009).

Colored DOM (CDOM) is DOM that absorbs light, whereas fluorescent DOM (FDOM) is the portion of the DOM pool that absorbs and subsequently emits light. Analyzing the wavelengths of emission of a water sample can be easily accomplished

using modern fluorometer technology and can divulge a host of information about the FDOM contained within a sample. For example, the peak excitation/emission wavelength pair of a sample can reveal where the FDOM originated (Coble 1996, 2007; McKnight et al. 2001; Para et al. 2010), how recently it was created (Huguet et al. 2009; Parlanti et al. 2000), its chemical composition (Coble 1996, 2007), how bioavailable it is (Coble 1996, 2007), and its relative molecular weight (Blough and Del Vecchio 2002).

Coble (1996, 2007) identified many of the excitation/emission wavelengths that correspond to certain properties of FDOM and ascribed to these wavelengths letter designations (e.g. C, M , T, A). For example, FDOM of terrestrial origin (peak C) is known to experience peak emission at ~440nm along excitation wavelength ~340nm. Conversely, marine FDOM (peak M), identified by peak excitation/emission at ~300/390nm, is DOM originating autochthonously in the ocean, mostly by microbial activity. Calculating the ratio of emission at peak C to the emission at peak M (C:M ratio = CMR) in a sample is a useful estimate of to what degree the sample is composed of terrestrial or marine FDOM (Burdige et al. 2004; Para et al. 2010, Helms et al. 2013) and has been explored in the mid-Atlantic and Chesapeake Bay area (Burdige et al. 2004), the northwestern Mediterranean Sea (Para et al. 2010), and the northern Pacific (Helms et al. 2013). A CMR value > 1 implies that a sample is comprised of FDOM originating primarily from a terrestrial source, whereas a CMR < 1 suggests mostly marine origin, and finally a value of exactly 1 implies an equivalent amount of terrestrial and marine FDOM within the sample.

Huguet et al. (2009) derived another fluorescence ratio, the Biological Index (BIX), which compares excitation/emission at ~310/380nm to the excitation/emission at 310/430nm to determine how much of a sample is composed of fresh (i.e. recently created, unrelated to "freshwater") FDOM. A BIX value of 1 or greater indicates that the FDOM in a sample is primarily of autochthonous bacterial origin, with values of 0.8 – 1, 0.7 – 0.8, and 0.6 – 0.7 implying strong, intermediate, and low amounts of autochthonous FDOM, respectively.

The BIX was developed to study estuarine DOM in the Gironde Estuary on the southwest coast of France (Huguet et al. 2009) and has also been used in other estuaries, such as the Lake Pontchartrain Estuary in the northern Gulf of Mexico (Haywood et al. 2018, Kolic 2014). However, other research has utilized BIX outside of the estuarine zone, for example within the Celtic Sea (Carr et al. 2019), the Bay of Bengal (Chari et al. 2013), as well as for estimating the contribution of terrestrial-derived DOM to the global carbon budget (Bianchi et al. 2013).

Modern submersible buoyancy gliders and towed systems utilize fluorometers that measure FDOM at one excitation/emission wavelength pair: 370/460 ("bulk CDOM"). The bulk CDOM wavelength pair is useful for approximating the amount of FDOM in a sample, but it does not reveal specifics about the nature of FDOM, such as its origin, relative age, etc. Therefore, in Dissertation Chapter 3, two algorithms were developed to be applied to glider or towfish fluorometer datasets: one algorithm for estimating BIX and another for estimation of CMR. In this study, the BIX and CMR algorithms were applied to a host of TXLA shelf towfish data from the Mechanisms Controlling Hypoxia

(MCH) project (DiMarco & Zimmerle, 2017) as well as from a submersible buoyancy glider. The goal of this study is to characterize the variability of FDOM on the TXLA shelf, both horizontally (i.e., the variability alongshelf with respect to proximity to the MARS endmembers moving downcoast towards Texas) as well as vertically (i.e., within the water column with respect to the pycnocline). Furthermore, we compared how these FDOM characteristics change from early summer to late summer, and how FDOM varied as a function of MARS discharge. Ultimately, the modeled FDOM results are compared to concurrent TXLA shelf oxygen data to determine whether FDOM of a particular freshness or source (terrestrial vs. marine) is more likely to coincide with the occurrence of hypoxia.

4.2. Data and Methods

4.2.1. MCH Acrobat Data

Data collected by a SeaSciences Acrobat cabled towfish during three separate years (Table 4.1, two cruises per year) of the Mechanisms Controlling Hypoxia (MCH) project (DiMarco & Zimmerle, 2017) were analyzed. The selected MCH data were collected along the shelf in June and August of each year from the Louisiana Bight to Galveston Bay, Texas. Multiple transect lines were collected from nearshore to offshore in each cruise. The selected MCH data were subsequently divided into five Sub-regions: A, B, C, D, and E (Figure 4.1). Sub-regions A, B, and C were first used by Kim et al. (2019), who built upon the northern Gulf of Mexico regionalization system (Rowe & Chapman, 2002; Belabbassi et al. 2006). In the present study, Sub-regions D and E were created to

fill the gap between Kim et al. (2019) Sub-region C and the western border of Region 3 from Belabbassi et al. (2006). Sub-region A corresponds to the Louisiana Bight region, encompassing the Mississippi River delta and the area due south of Grand Isle. Sub-region B is situated due south of and is the same approximate width as Terrebonne Bay; likewise for Sub-region C and Atchafalaya Bay. Finally, Sub-region D includes the vast region south of Cameron, Louisiana and White Lake, while Sub-region E is situated on the area of the shelf immediately southeast of Galveston Bay. In comparing data across Sub-regions or comparing above and below the pycnocline, Kruskal-Wallis ANOVA tests were used because each test had at least one group that was not normally distributed.

Table 4.1: Mission number and dates of Mechanisms Controlling Hypoxia project Acrobat data used in this study; data for each Sub-region per Cruise, with number of transects and data points in parentheses.

Mission	Year	Month	Sub-region (# Transects; # Data Points)				
			A	B	C	D	E
MS03	2011	June	(2; 53,042)	(3; 62,988)			
MS04	2011	August	(2; 67,233)	(3; 147,698)	(3; 105,828)	(2; 72,808)	(2; 57,615)
MS05	2012	June	(2; 73,950)	(3; 84,187)	(3; 78,280)	(3; 82,889)	(1; 33,551)
MS06	2012	August	(2; 86,810)	(2; 89,250)	(2; 62,976)		(1; 71,481)
MS07	2013	June	(2; 42,464)	(3; 130,803)	(3; 118,733)	(3; 143,359)	(2; 69,023)
MS08	2013	August		(2; 63,402)	(3; 92,524)	(3; 121,928)	(2; 78,029)

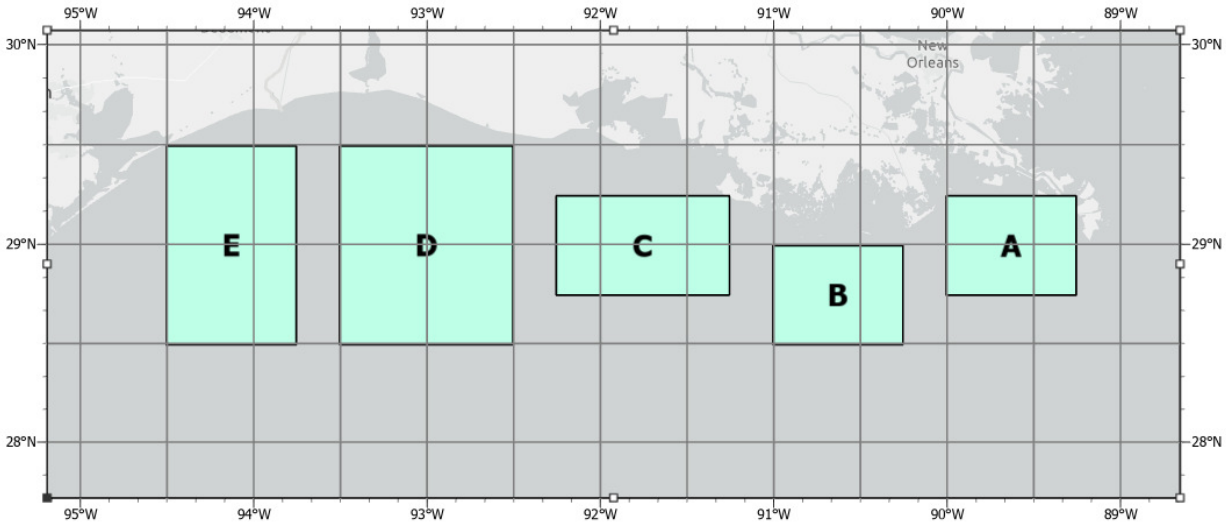

Figure 4.1: The five Sub-regions of the Texas-Louisiana shelf, adapted from Kim et al. (2019) and Belabbassi et al. (2006). Latitude/longitude borders for the top right and bottom left corners of each Sub-region are as follows: Sub-region A (29.25°N, 89.25°W; 28.75°N, -90°W), Sub-region B (29°N, 90.25°W; 28.25°N, 91°W), Sub-region C (29.25°N, 91.25°W; 28.75°N, 92.25°W), Sub-region D (29.5°N, 92.5°W; 28.5°N, 93.5°W), Sub-region E (29.5°N, 93.75°W; 28.5°N, 94.5°W).

Although stratification strength can be quantified by calculating Brunt-Vaisala frequency (Bianchi et al., 2010), for this study, the pycnocline density was calculated for each cruise Sub-region by comparing density, depth, salinity, and temperature. In most cases, the maximum density gradient was found to be ~1023 kg m^3. Thus, any data that are hereafter referred to as "sub-pycnocline" are those points at which the density was greater than or equal to the pycnocline density for their respective Sub-region; whereas "above-pycnocline" data are those for which density was less than the pycnocline density. MS08 Sub-region E was too well-mixed, therefore no sub- or above-pycnocline data were calculated for that Sub-region. Additionally, apparent oxygen utilization

(AOU) and spiciness were each calculated for the MCH data according to the methods of Weiss (1970) and Flament (2002), respectively.

4.2.2. Ocean Buoyancy Glider Data

Glider data from GERG glider M7 were analyzed to provide an estimate of TXLA shelf BIX and CMR in 2014. Glider M7 was deployed from 11 July to 12 August 2014, transiting eastward along the TXLA shelf from its deployment location south of Cameron Parish, Louisiana to its recovery south of the Isles Dernieres near Terrebonne Bay. 36% of the data collected by glider M7 were collected within one of the TXLA shelf Sub-region boxes, but only with Sub-regions C and D. It is important to note that, whereas the MCH Acrobat data were collected north-south, the M7 glider data were collected west-east. Glider FDOM data were not partitioned in the same way about the pycnocline as were the MCH Acrobat data; i.e. only the overall composite water column was analyzed for the glider data.

4.2.3. FDOM Algorithms

The development of the algorithms used to estimate FDOM freshness and source was discussed in Chapter 3:

The BIX predictive algorithm (Eq. 1) includes 8 separate terms. BIX values > 1 suggest that the majority of the FDOM in a sample is of aquatic bacterial origin, whereas BIX values < 1 indicate that a sample contains either a strong (BIX = 0.8 – 1), intermediate (BIX = 0.7 – 0.8), or low (BIX = 0.6 – 0.7) amount of autochthonous FDOM.

The CMR predictive algorithm (Eq. 2) includes 13 terms. A CMR value of exactly 1 would indicate that the marine and terrestrial DOM sources are present in equal parts volumetrically for that data point, whereas values > 1 reveal that terrestrial material is more dominant, while data points with < 1 CMR are chiefly comprised of marine material.

4.2.4. K-means Clustering

In many cases, comparison of bulk CDOM and salinity from individual MCH Acrobat tows (or "lines") represented a traditional conservative mixing line (i.e. decreasing CDOM with increasing salinity). Of the MCH lines for which the relationship between bulk CDOM and salinity was nonlinear (and thus mixing of CDOM was nonconservative), one line, MS03 Line 12 (the westernmost line of Sub-region B, flown due south of Isles Dernieres near Terrebonne Bay; hereafter "MS3L12") was selected for K-means clustering according to the methods developed in Dissertation Chapter 2. Clustering of the M7 glider dataset was also performed. The clustering techniques are used in this chapter both to demonstrate again the efficacy of the clustering technique for analysis of these types of data as well as to explore how the nature of FDOM (i.e. larger or smaller BIX or CMR) influenced the tendency of DOM to be mixed nonconservatively. As with the MCH Acrobat data (and clusters in Dissertation Chapter 2), the Kruskal-Wallis ANOVA test was used to compare M7 glider clusters because of heteroscedasticity in the cluster data.

4.3. Results

The results section of this chapter is organized in the following way:

- Sections 4.3.1 and 4.3.2 detail, respectively, BIX and CMR shelf-wide seasonal, interannual, geographic (along the shelf), and vertical water column variability;
- Section 4.3.3 presents a comparison of modeled BIX vs. CMR observations;
- Section 4.3.4 investigates the potential role of modeled FDOM characteristics on shelf oxygen observations;
- Section 4.3.5 supplements the MCH Acrobat data with analysis of modeled FDOM observations from a submersible buoyancy glider; and,
- in Section 4.3.6, the clustering techniques from Chapter 2 were used to determine whether modeled FDOM characteristics could explain variability in water column clusters and whether FDOM variability could be explained by clustering.

4.3.1. Observations of TXLA Shelf DOM Freshness

The seasonality and interannual variability of the TXLA shelf Sub-regions FDOM freshness were examined. The estimated values for BIX (Figure 4.2), the measure of FDOM freshness, were found to be significantly different within all TXLA shelf Sub-regions of each cruise (i.e. all MS03 Sub-regions were statistically distinct, all MS4 Sub-regions were statistically distinct, etc.) according to a Kruskal-Wallis ANOVA. Mean BIX was > 1 within all Sub-regions across all cruises (Table A1), indicating that the majority of the FDOM measured by the Acrobat was of an autochthonous biological or aquatic bacterial origin (Huguet et al., 2009). The largest observed mean BIX for any cruise Sub-region, 1.41 ($\sigma = 0.04$), was found in MS04 Sub-region E, the westernmost

Sub-region in the year with the greatest MARS discharge. The lowest mean BIX, 1.06 (σ = 0.13), was also found in cruise MS04, but in Sub-region B.

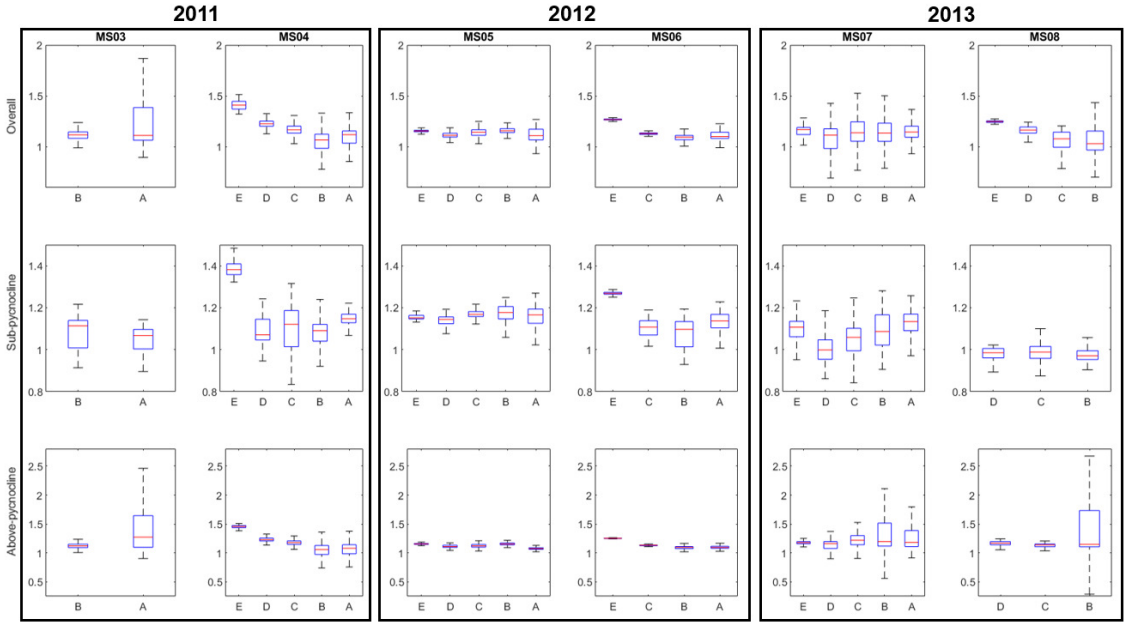

Figure 4.2: Boxplots of the estimated BIX values within each cruise (columns), with each boxplot grouped by Sub-region. The top row is overall BIX, the middle row is sub-pycnocline BIX, and the bottom row is above-pycnocline BIX.

While BIX in all Sub-regions was > 1, means and σ varied from cruise to cruise or Sub-region to Sub-region. In the August cruises, mean BIX generally increased from east to west, with Sub-region E featuring the largest mean BIX in MS04 and MS06 (August 2011 and 2012, respectively). In the other August cruise, MS08, Sub-region E mean BIX was a near second to that of Sub-Region B. Although significantly different according to the ANOVA, mean BIX was least variable across Sub-regions of the June cruise from the drought year (MS05). Differences between Sub-regions in one of the other June cruises, MS07, were more pronounced: e.g. Sub-Region B mean BIX of 1.38 (σ = 0.88) vs. Sub-region E mean BIX of 1.15 (σ = 0.06). The other June cruise (MS03),

only collected data in Sub-regions A and B, with mean and σ BIX found to be larger in Sub-region A. In all six cruises, regardless of month, BIX σ were largest in the easternmost Sub-regions and lowest in the west.

Sub-pynocline BIX trends were similar to the overall BIX observations (i.e. the entire water column composite, both above and below the pynocline) reported above. All Sub-regions in all but one of the cruises, MS08, featured mean BIX > 1; even for the MS08 Sub-regions, mean BIX was greater than or equal to 0.97, which still indicates a strong autocthonous DOM component (Huguet et al., 2009). The August cruises again had increased BIX in the westernmost Sub-region, while June cruises were again less variable across Sub-regions. One of the main differences in the sub-pycnocline vs. overall BIX was that sub-pycnocline BIX variability was low across all cruises and Sub-regions, with all but one Sub-region featuring BIX σ < 0.1 (MS04 Sub-Region C BIX σ = 0.10). The lowest sub-pycnocline mean BIX, 0.97 (σ = 0.03; BIX < 1 indicates strong autochthonous component), was found in MS08 Sub-region B, while the largest, 1.39 (σ = 0.03), was, as with the overall BIX, found within MS04 Sub-region E.

BIX conditions above the pycnocline varied widely from cruise to cruise, but all above-pycnocline mean BIX values were > 1. The drought year cruises, MS05 and MS06, featured both mean and σ above-pycnocline BIX values similar to those in the sub-pycnocline, although only MS06 Sub-Region B had BIX values above and below the pycnocline that were not significantly different, according to the ANOVA. Some Sub-regions from other cruises also featured similar conditions in the sub- and above-pycnocline, but with greater disparity in σ BIX (e.g. MS04 Sub-Regions A and B).

In all cruises for which Sub-region E was found to be well-stratified (i.e. sufficiently stratified to allow for sub- and above-pycnocline comparisons), the Sub-region E sub- and above-pycnocline mean and σ BIX were always significantly different from one another. Additionally, Sub-region E again featured the largest mean BIX in the August cruises. The largest above-pycnocline mean BIX, 1.69 (σ = 1.19), was found in MS07 Sub-region B, but comparatively large mean BIX were also found in MS03 Sub-region A (1.59, σ = 0.82), MS07 Sub-region A (1.59, σ = 1.34), and MS08 Sub-region B (1.57, σ = 0.81). The largest difference in sub- and above-pycnocline mean BIX values were found in Sub-region B of the 2013 cruises (above-pycnocline 0.60 > sub-pycnocline in MS07, and 0.59 > sub-pycnocline in MS08). The lowest above-pycnocline mean BIX value, 1.05 (σ = 0.14), was found in MS04 Sub-region B, which was one of 7 cruise Sub-regions in which the above-pycnocline mean BIX was lower than that below the pycnocline. The phenomenon of lower mean BIX above the pycnocline vs. below was also observed in MS04 Sub-region A, MS05 Sub-regions A, B, C, and D, and MS06 Sub-region A.

Whereas sub-pycnocline σ BIX were all uniformly low, significant variability in σ BIX was observed in some of the above-pycnocline cruise Sub-regions, most notably in MS07, wherein Sub-regions A and B featured σ BIX of 1.34 and 1.19, respectively. MS03 Sub-region A and MS08 Sub-region B both also observed σ BIX > 0.80. In the Sub-regions nearest the MARS endmembers (A, B, and C), mean BIX values were larger in June than August, except in the 2012 drought year, during which mean BIX were nearly equivalent or were actually slightly larger in August than June. The

differences from June to August were most pronounced in 2011, the MARS flood year, with above-pycnocline mean and σ BIX larger for MS03 than MS04 Sub-regions A and B; however, the observations were reversed in the sub-pycnocline Sub-region A, with MS04 mean BIX greater than that of MS03.

4.3.2. Observations of TXLA Shelf Terrestrial vs. Marine DOM

The seasonal and interannual differences of TXLA shelf CMR was examined. In 11 of the 25 cruise Sub-regions analyzed, overall mean CMR values (i.e. composite of both sub- and above-pycnocline) were < 1, indicating a greater component of marine material than terrestrial (Figure 4.3). The remaining 14 cruise Sub-regions saw mean CMR values > 1, but often with large σ. For example, in all Sub-regions with mean CMR > 1.5, σ CMR was > 1. Kruskal-Wallis ANOVA revealed that overall CMR were significantly different within each cruise Sub-region, except between MS07 Sub-regions C and D. Additionally, the ANOVA showed that Sub-region C in August 2011 (MS04) was not significantly different from August of 2013 (MS08). The MS07 Sub-regions also observed some of the highest mean CMR values (all were > 1.25); MS03 Sub-region A, MS04 Sub-regions A and B, and MS08 Sub-region B also observed mean CMR values above the same threshold. MS08 Sub-region B, in particular, contained the highest mean CMR, 2.30 (σ = 1.92), while the lowest mean CMR, 0.69 (σ = 0.08) was found within the drought year cruise MS06 Sub-region E. Whereas August cruise mean BIX values were higher in Sub-region E, mean CMR was always lowest in Sub-region E in each cruise, with August cruises each showing a decrease in Sub-region E mean CMR as

compared to their concurrent year June cruise counterparts (save for MS03, for which no data were recorded in Sub-region E). The changes in mean CMR from June to August were not consistent each year: in 2011, mean CMR values decreased from June to August in Sub-region A, but increased in Sub-region B; the 2012 Sub-regions, save for Sub-region E, observed an increase; while the 2013 Sub-regions decreased, except for in Sub-region B, which saw the greatest change from June to August, wherein mean CMR increased by 0.47. The largest mean CMR was observed in either Sub-region A or B for each cruise, except for MS05, in which Sub-region D saw the largest mean CMR.

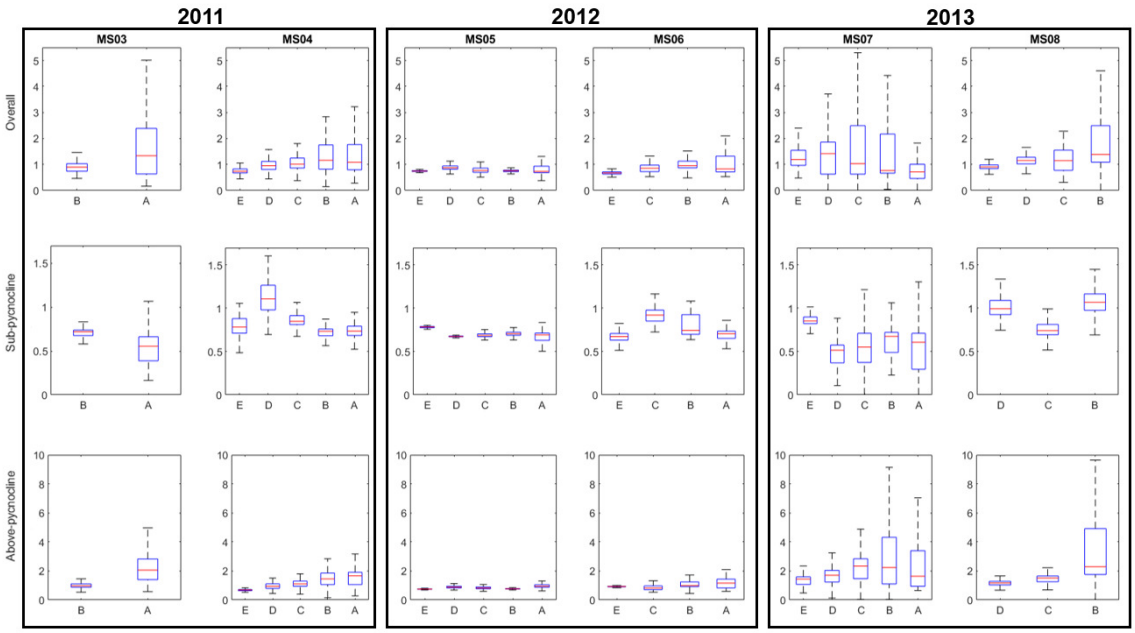

Figure 4.3: Boxplots of the estimated CMR values within each cruise (columns), with each boxplot grouped by Sub-region. The top row is overall CMR, the middle row is sub-pycnocline CMR, and the bottom row is above-pycnocline CMR.

Sub-pycnocline mean CMR values were < 1 in all but three cruise Sub-regions: MS08 Sub-regions B and D and MS04 Sub-region D, the latter of which featured the largest mean CMR value, 1.11 ($\sigma = 0.19$). The lowest mean CMR in the sub-pycnocline,

0.49 ($\sigma = 0.17$) was found in MS07 Sub-region D, which also saw the greatest change from June to August, increasing to 1.02 ($\sigma = 0.13$) in MS08 Sub-region D. In contrast to the overall CMR results, sub-pycnocline mean CMR values were largest in Sub-region E during the 2012 and 2013 June cruises (MS05 and MS07). Also contrasting many of the overall CMR Sub-regions, sub-pycnocline σ CMR values were all < 0.24, with 14 cruise Sub-regions observing σ CMR < 0.1. Within each cruise, all Sub-regions were significantly different from one another; however, the mean CMR values in Sub-region E were not significantly different between cruise MS04 and MS05.

ANOVA revealed that CMR values above and below the pycnocline were significantly different from one another in all cruise Sub-regions (e.g. MS04 Sub-region A sub-pycnocline CMR vs. above-pycnocline CMR). Mean CMR values were larger above the pycnocline than in their sub-pycnocline counterparts in all but 3 cruise sub-regions: MS04 and MS05 Sub-region E, and MS06 Sub-region C. Of the 24 (MS08 Sub-region E was not sufficiently stratified, thus it was not included in the sub- or above-pycnocline analysis) cruise Sub-regions, 15 observed above-pycnocline mean CMR > 1, with 7 of those Sub-regions featuring mean CMR > 1.5. Of the 9 sub-regions with above-pycnocline mean CMR < 1, only 3 saw mean CMR < 0.84. As with the overall CMR, the largest above-pycnocline mean CMR, 3.34 ($\sigma = 2.10$) was found in MS08 Sub-region B, with MS04 Sub-region E featuring the lowest above-pycnocline mean CMR, 0.69 ($\sigma = 0.06$). In all cruises, either Sub-region A or B contained the largest observed above-pycnocline mean CMR. The June 2012 and 2013 (MS05 and MS07) Sub-region E above-pycnocline mean CMR values were lowest among their respective

cruise Sub-regions; Sub-region E was also the lowest in August 2011 (MS04). The largest change from June to August was encountered in 2013 Sub-region C, wherein above-pycnocline mean CMR decreased from 2.27 ($\sigma = 0.99$) in MS07 to 1.45 ($\sigma = 0.26$) in MS08. A decrease in above-pycnocline mean CMR was also observed from June to August in Sub-region A in 2011 and Sub-region D in 2013. Otherwise, above-pycnocline mean CMR increased from June to August in the other 2011 and 2013 Sub-regions and within all Sub-regions in 2012. Finally, CMR in all cruise Sub-regions was significantly different above the pycnocline, except when comparing Sub-regions A and B in MS06.

4.3.3. Comparisons of DOM Freshness and Source

BIX and CMR were compared to calculate the determination coefficient (R^2) between the two in each Sub-region in the sub-pycnocline, above-pycnocline, and in each Sub-region overall (i.e. composite of sub- and above-pycnocline). Considering each Sub-region overall, the strongest BIX/CMR relationships ($R^2 > 0.7$) were found in the eastern Sub-regions in cruises MS03, MS07, and MS08, with the largest having been encountered in MS07 Sub-region A ($R^2 = 0.85$). The only other cruise Sub-regions with an overall BIX/CMR $R^2 > 0.5$ were western Sub-regions: MS06 Sub-region E ($R^2 = 0.64$), MS07 Sub-region C ($R^2 = 0.507$), and MS08 Sub-Region C ($R^2 = 0.627$).

In the sub-pycnocline, only cruise MS06 featured any Sub-regions with $R^2 > 0.5$ for BIX/CMR, the largest correlation being found in MS06 Sub-region B ($R^2 = 0.81$) and the least in MS06 Sub-region A ($R^2 = 0.538$). As with the overall Sub-regions, the

strongest above-pycnocline BIX/CMR correlations ($R^2 > 0.7$) were again found in Sub-regions A or B in cruises MS03, MS07, and MS08, the largest R^2 having been found in MS07 Sub-region A ($R^2 = 0.89$). Otherwise, BIX/CMR correlations were not well-correlated above the pycnocline.

4.3.4. Relationships between TXLA Shelf FDOM and Oxygen

BIX and CMR were each compared to dissolved oxygen concentration and AOU to calculate their respective determination coefficients (R^2) in each cruise Sub-region in the sub-pycnocline, the above-pycnocline, and overall (i.e. sub- and above-pycnocline composite). Determination coefficients between BIX and oxygen were generally similar to those between BIX and AOU; the only instances in which BIX vs. oxygen R^2 differed by > 0.1 from those of BIX vs. AOU were in the sub-pycnocline in MS03 Sub-region A and MS07 Sub-regions A and B. Similarly, CMR and oxygen determination coefficients were generally similar to those of CMR and AOU, with R^2 of each differing by > 0.1 again only in the sub-pycnocline in three cruise Sub-regions: Sub-region A in both MS03 and MS04, and MS06 Sub-region B.

BIX and oxygen correlations were generally higher in the western Sub-regions than in the east. The strongest correlations (Figure 4.4) between BIX and oxygen ($R^2 > 0.9$) in the overall cruise Sub-regions (i.e. composite of sub- and above-pycnocline values) were found in MS04 Sub-region E (BIX/oxygen $R^2 = 0.93$; BIX/AOU $R^2 = 0.95$) and MS08 Sub-region C (BIX/oxygen $R^2 = 0.917$; BIX/AOU $R^2 = 0.914$). Strong BIX vs. oxygen correlations ($R^2 > 0.65$) were also found in MS04 Sub-regions C and D, MS07 Sub-

regions C and E, and MS08 Sub-region D. The 2012 drought year cruises (MS05 and MS06) only featured one Sub-region, MS06 Sub-region E, in which the BIX and oxygen determination coefficient was > 0.5. The CMR and oxygen overall Sub-region determination coefficients were generally lower than those of BIX and oxygen, with CMR/oxygen or CMR/AOU R^2 values > 0.7 found only in MS08 Sub-region C (CMR/oxygen $R^2 = 0.70$; CMR/AOU $R^2 = 0.71$). Otherwise, only MS03 Sub-region A, MS05 Sub-region E, MS07 Sub-region D, and MS08 Sub-region B featured CMR and oxygen $R^2 > 0.5$.

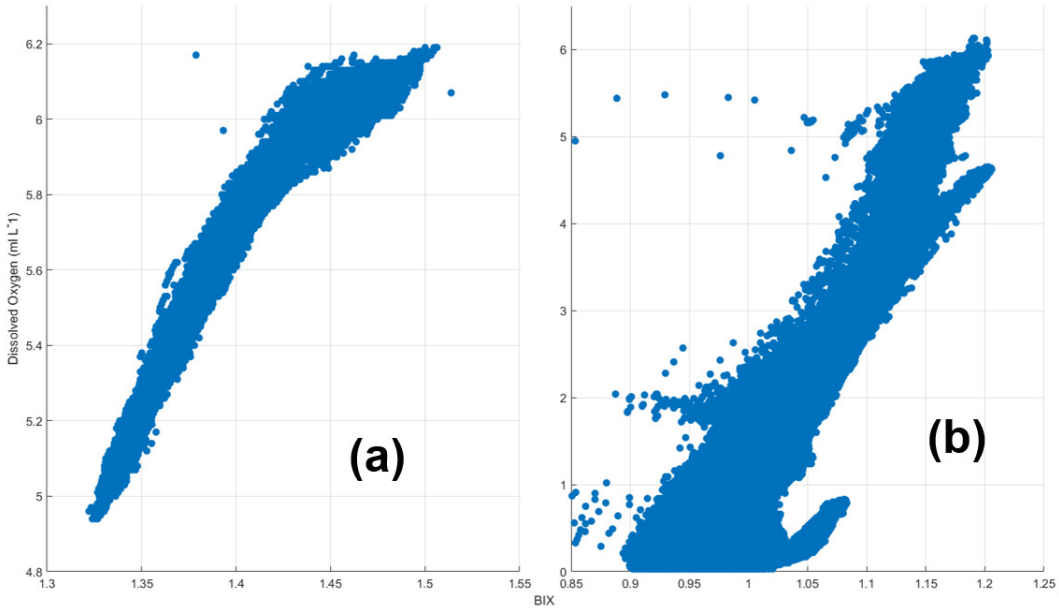

Figure 4.4: BIX (x-axis) vs. dissolved oxygen (y-axis) for (a) MS04 Sub-region E (R^2 = 0.93) and (b) MS08 Sub-region C (R^2 = 0.917)

Sub-pycnocline relationships between BIX and oxygen were often stronger than those found in the composite Sub-region analysis, with 7 cruise Sub-regions featuring BIX and oxygen $R^2 > 0.9$ and a further 8 cruise Sub-regions with $R^2 > 0.65$. As with the

overall Sub-regions, BIX and oxygen were more closely related in the western Sub-regions than in the east. The strongest relationships ($R^2 > 0.95$) between BIX and oxygen were found in the sub-pycnocline in MS04 Sub-regions C and E, and MS07 Sub-region B. Sub-pycnocline BIX and oxygen relationships were strong ($R^2 > 0.6$) in Sub-regions B and C in all cruises, except MS08 Sub-region C (BIX/oxygen $R^2 = 0.557$; BIX/AOU $R^2 = 0.54$). Sub-pycnocline CMR and oxygen were not well-correlated beneath the pycnocline, with only MS06 Sub-region C featuring CMR and oxygen $R^2 > 0.5$ (CMR/oxygen $R^2 = 0.567$; CMR/AOU $R^2 = 0.507$).

BIX and oxygen were not as well-correlated above the pycnocline. Only 2 Sub-regions, MS04 Sub-region E and MS08 Sub-region C, observed BIX and oxygen $R^2 > 0.85$. Otherwise, only MS04 Sub-region C, MS07 Sub-region E, and MS08 Sub-region D contained BIX and oxygen $R^2 > 0.5$; however, BIX/AOU R^2 was also > 0.5 in MS04 Sub-region D (BIX/oxygen $R^2 = 0.48$; BIX/AOU $R^2 = 0.515$). As with the sub-pycnocline, CMR and oxygen were not well-correlated above the pycnocline, with only MS05 Sub-region E featuring CMR and oxygen $R^2 > 0.5$ (CMR/oxygen $R^2 = 0.54$; CMR/AOU $R^2 = 0.545$). Otherwise, only 4 other Sub-regions featured above-pycnocline CMR and oxygen $R^2 > 0.3$: MS05 Sub-region D, MS07 Sub-regions A and B, and MS08 Sub-region B.

The MCH Acrobat data also provided an opportunity to compare FDOM within and without of hypoxic waters. In all Sub-regions in which hypoxia was observed, mean BIX values were lower in hypoxic waters than in non-hypoxic waters (Table 4.2). Similarly, in all but 2 Sub-regions (MS04 Sub-region D and MS06 Sub-region C), mean CMR was

lower in hypoxic waters than non-hypoxic waters (Table 4.3). Finally, in all Sub-regions, Kruskal-Wallis ANOVA indicated that both BIX and CMR were significantly different in hypoxic vs. non-hypoxic waters.

Table 4.2: Mean BIX within each Sub-region (rows) of each cruise (columns), grouped by BIX within hypoxic waters and BIX within non-hypoxic waters. Green and red cells indicate the largest and smallest mean BIX value, respectively, in that grouping. Standard deviation (σ) in parentheses.

Mean BIX in Hypoxic Waters (σ)							
	MS03	**MS04**	**MS05**	**MS06**	**MS07**	**MS08**	**Region**
Sub-Region A	1.04 (0.07)	1.08 (0.09)	1.06 (0.06)	**1.10 (0.07)**	1.05 (0.08)	ND	**East**
Sub-Region B	1.02 (0.09)	0.99 (0.09)	1.08 (0.01)	1.05 (0.06)	1.03 (0.13)	0.98 (0.04)	
Sub-Region C	ND	0.98 (0.05)	No Hypoxia	1.05 (0.01)	1.00 (0.05)	0.98 (0.03)	**Middle**
Sub-Region D	ND	1.04 (0.03)	No Hypoxia	ND	0.98 (0.06)	0.98 (0.03)	**West**
Sub-Region E	ND	No Hypoxia	No Hypoxia	No Hypoxia	**0.82 (0.07)**	No Hypoxia	

Mean BIX in non-Hypoxic Waters (σ)							
	MS03	**MS04**	**MS05**	**MS06**	**MS07**	**MS08**	**Region**
Sub-Region A	1.57 (0.79)	1.14 (0.33)	1.13 (0.06)	1.12 (0.04)	1.37 (0.94)	ND	**East**
Sub-Region B	1.19 (0.27)	**1.08 (0.13)**	1.16 (0.03)	1.09 (0.03)	1.48 (0.97)	**1.66 (0.83)**	
Sub-Region C	ND	1.17 (0.05)	No Hypoxia	1.13 (0.02)	1.19 (0.10)	1.13 (0.04)	**Middle**
Sub-Region D	ND	1.23 (0.05)	No Hypoxia	ND	1.13 (0.14)	1.16 (0.04)	**West**
Sub-Region E	ND	No Hypoxia	No Hypoxia	No Hypoxia	1.15 (0.06)	No Hypoxia	

Table 4.3: Mean CMR within each Sub-region (rows) of each cruise (columns), grouped by CMR within hypoxic waters and CMR within non-hypoxic waters. Green and red cells indicate the largest and smallest mean CMR value, respectively, in that grouping. Standard deviation (σ) in parentheses.

Mean CMR in Hypoxic Waters (σ)							
	MS03	**MS04**	**MS05**	**MS06**	**MS07**	**MS08**	**Region**
Sub-Region A	0.68 (0.48)	0.85 (0.34)	0.59 (0.12)	0.73 (0.16)	**0.52 (0.32)**	ND	**East**
Sub-Region B	0.98 (0.62)	1.02 (0.37)	0.62 (0.01)	0.85 (0.12)	0.88 (0.94)	**1.17 (0.33)**	
Sub-Region C	ND	0.88 (0.09)	No Hypoxia	1.00 (0.03)	0.69 (0.52)	0.81 (0.18)	**Middle**
Sub-Region D	ND	1.15 (0.21)	No Hypoxia	ND	0.55 (0.28)	1.15 (0.23)	**West**
Sub-Region E	ND	No Hypoxia	No Hypoxia	No Hypoxia	0.60 (0.07)	No Hypoxia	

Mean CMR in non-Hypoxic Waters (σ)							
	MS03	**MS04**	**MS05**	**MS06**	**MS07**	**MS08**	**Region**
Sub-Region A	2.09 (1.15)	1.35 (0.63)	0.82 (0.20)	1.07 (0.35)	1.64 (2.74)	ND	**East**
Sub-Region B	1.13 (0.75)	1.38 (0.53)	**0.75 (0.05)**	1.09 (0.31)	2.10 (2.35)	**3.56 (2.16)**	
Sub-Region C	ND	1.08 (0.24)	No Hypoxia	0.85 (0.16)	1.86 (1.15)	1.43 (0.31)	**Middle**
Sub-Region D	ND	0.95 (0.18)	No Hypoxia	ND	1.69 (0.68)	1.15 (0.19)	**West**
Sub-Region E	ND	No Hypoxia	No Hypoxia	No Hypoxia	1.26 (0.34)	No Hypoxia	

4.3.5. Glider FDOM Observations

Glider M7 data were only available for Sub-regions C and D. Mean BIX and mean CMR were all > 1 for each of the Sub-regions, with Sub-region C featuring larger mean values for both BIX and CMR. Although mean and σ BIX were similar for both Sub-regions, Kruskal-Wallis ANOVA indicated that BIX was significantly different between the two Sub-regions whereas CMR was not significantly different between Sub-regions C and D. Additionally, sub-region C was the only M7 Sub-region in which BIX and CMR were somewhat correlated ($R^2 = 0.421$). Oxygen and AOU comparisons revealed that BIX and oxygen were well-correlated ($R^2 > 0.6$ in both Sub-regions). CMR and oxygen were only correlated in Sub-region C (CMR/oxygen $R^2 = 0.508$; CMR/AOU $R^2 = 0.64$).

Considering the M7 dataset as a whole, rather than only within the Sub-region boxes, the waters with largest BIX values were found in the easternmost portion of the glider deployment, east of -91.5 W and in waters shallower than ~25 m depth. This area of greater BIX values corresponds to waters immediately south of, but not included within, Sub-region B. A patch of lower M7 BIX concentrations (BIX < 1) were found near the ocean bottom from -91.8 W to -91.2 W, the western portion of which is included in Sub-region C. The patch of low BIX corresponded to observations of hypoxia. Similar observations of low BIX values and hypoxic waters were found near the bottom from -93.2 W to -92.6 W, within the boundaries of Sub-region D. A patch of larger CMR values (mean CMR = 1.42, $\sigma = 0.31$) was found in the upper waters (i.e. not near the bottom) from -92.5 W to -91.5W, the region of the TXLA shelf due south of Atchafalaya Bay. Bottom waters east of -92.5 W featured lower CMR values (mean CMR = 0.82, $\sigma = 0.35$) than in the upper waters, especially in the deep waters (> 25m) east of -91.5 W (mean CMR = 0.45, $\sigma = 0.15$), immediately south of Sub-region B. West of -92.5 W, within the confines of Sub-region D, CMR values were more uniform throughout the water column.

4.3.6. TXLA Shelf Clustering

Silhouette scoring of the bulk CDOM vs. salinity plot for MS3L12 recommended 5 clusters (Figure B4) according to the k-means clustering algorithm: the bottom cluster (Figure 4.5 gold), the pycnocline cluster (Figure 4.5 purple), the bloom cluster (Figure 4.5 blue) containing the largest mean chl-a concentration of the five clusters, the mid-

depth cluster (Figure 4.5 green) above the pycnocline but below the surface, and the surface cluster (Figure 4.5 orange).

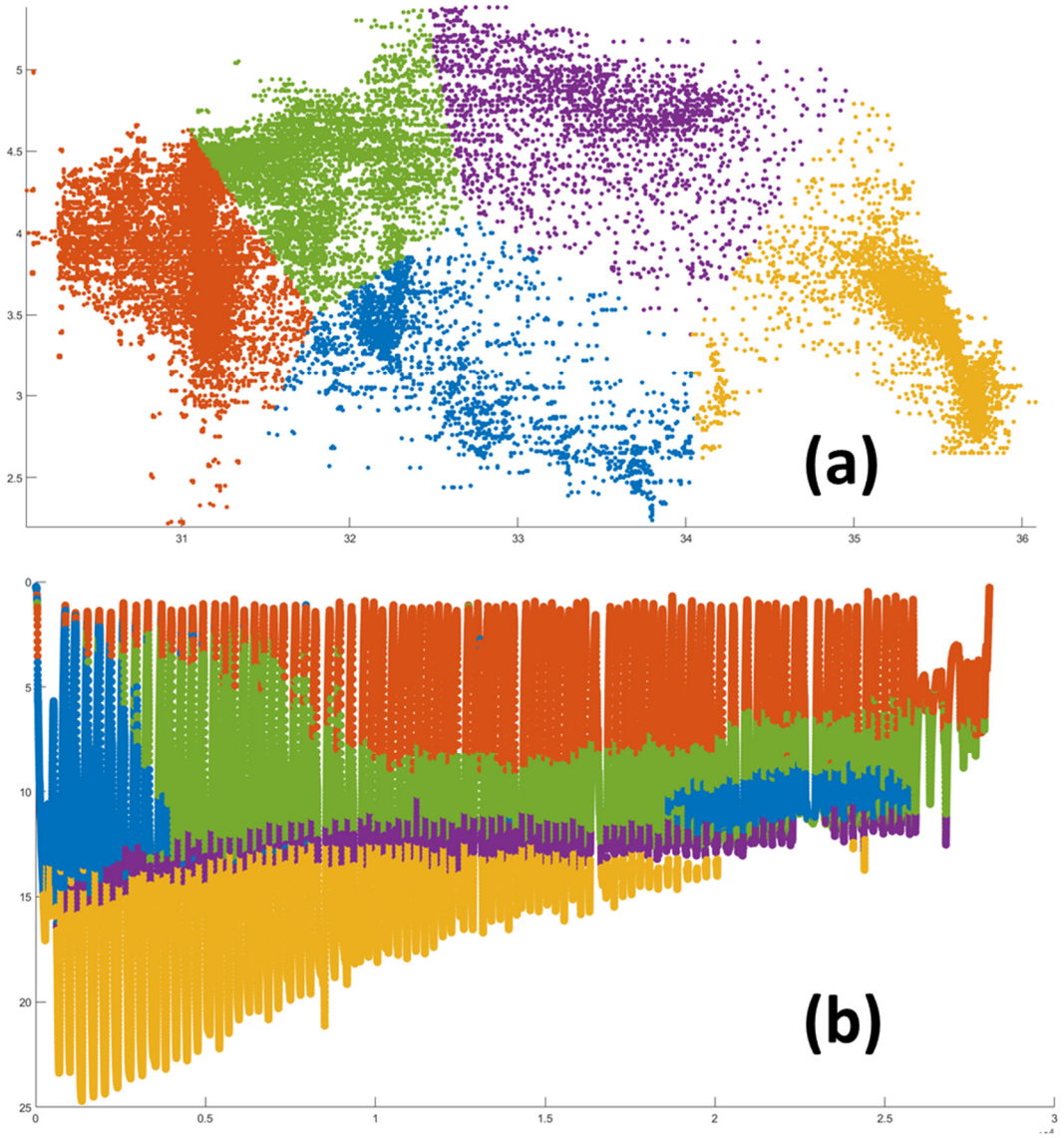

Figure 4.5 (a) MS03 Line 12 CDOM vs. salinity plot, with k-means cluster colors; (b) MS03 Line 12 Hovmöller diagram with colors corresponding to cluster colors from (a).

Of the MS3L12 clusters, the surface, bloom, and mid-depth clusters all observed mean BIX values > 1, with the largest means found in the surface cluster (mean BIX = 1.13, σ = 0.01) and the bloom cluster (mean BIX = 1.12, σ = 0.02). Kruskal-Wallis

ANOVA determined that the BIX values within the bloom and surface clusters were not significantly different from one another. Mean BIX < 1 were found in the pycnocline cluster (mean BIX = 0.97, σ = 0.05) and bottom cluster (mean BIX = 0.99, σ = 0.03). Mean CMR was largest in the mid-depth cluster (mean CMR = 1.09, σ = 0.08). Otherwise, mean CMR values were < 1 in the remaining MS3L12 clusters, with the lowest mean value, 0.69 (σ = 0.06), being found in the bottom cluster. Kruskal-Wallis ANOVA determined that, of the CMR values within the MS3L12 clusters, only the pycnocline and surface clusters were not significantly different from one another. BIX and CMR were not well-correlated in any of the MS3L12 clusters, with only the pycnocline and surface clusters featuring BIX vs. CMR $R^2 > 0.3$.

Significant differences in dissolved oxygen concentration and oxygen demand were noted among the MS3L12 clusters. 99.7% of the hypoxia observed by the Acrobat in MS3L12 was found within the bottom and pycnocline clusters, with the bottom cluster featuring mean oxygen concentration of just 0.39 ml L^{-1} (σ = 0.36). Although the mean oxygen concentration was larger in the pycnocline cluster (mean oxygen = 0.844, σ = 1.06), Kruskal-Wallis ANOVA determined that oxygen values were not significantly different between the bottom and pycnocline clusters. Mean oxygen values in the remaining clusters were all well above the hypoxic threshold (2 ml L^{-1}), with the largest mean oxygen concentration (4.87, σ = 0.17) being found in the surface cluster. Oxygen demand data reflect the range of oxygen concentrations found within the MS3L12 clusters, with the largest levels of AOU found in the bottom cluster (mean AOU = 4.32, σ = 0.38) and pycnocline cluster (mean AOU = 3.78, σ = 1.09). Mean AOU was < 1 in

the remaining clusters, with the lowest mean AOU (-0.38, $\sigma = 0.19$) being found in the surface cluster. Kruskal-Wallis ANOVA determined that AOU was significantly different in all clusters.

Apart from within the MS3L12 bottom cluster, dissolved oxygen concentration and AOU were each strongly correlated with BIX. The R^2 determination coefficient between BIX and oxygen were > 0.54 for all non-bottom clusters and > 0.61 for BIX and AOU in the non-bottom clusters. The BIX and oxygen relationships were strongest in the mid-depth cluster (BIX/oxygen $R^2 = 0.856$, BIX/AOU $R^2 = 0.866$) and the pycnocline cluster (BIX/oxygen $R^2 = 0.979$, BIX/AOU $R^2 = 0.978$). MS3L12 oxygen and CMR were not well-correlated, with only the pycnocline, surface, and bottom cluster (for AOU only) observing CMR/oxygen or CMR/AOU $R^2 > 0.3$.

Of the MS3L12 clusters, mean chl-a values ranged from ~1.3 to ~2.1, with the largest mean values found in the bloom cluster (mean chl-a = 2.13, $\sigma = 0.57$) and the mid-depth cluster (mean chl-a = 1.99, $\sigma = 0.74$); the pycnocline cluster observed the lowest mean chl-a (1.31, $\sigma = 0.77$). Turbidity values were generally opposite of chl-a, with the bloom, mid-depth, and surface clusters featuring the lowest mean turbidity values and the pycnocline and bottom clusters observing the largest mean turbidity. Kruskal-Wallis ANOVA found that chl-a values differed significantly among all clusters, as did turbidity.

The M7 glider data were divided into 6 clusters according to Silhouette score recommendation (Figure B5): 3 surface clusters and 3 bottom clusters (Figure 4.6). The

bottom clusters are designated according to the approximate maximum depth of water in which they are located: the “shallow bottom” cluster (SBC, Figure 4.6 dark blue) found in waters generally < 25m deep, the “deep bottom” cluster (DBC, Figure 4.6 light blue) in waters reaching up to ~40m depth, and the “max CDOM” cluster (MXC, Figure 4.6 purple) in which are found the maximum CDOM values of any of the clusters. The MXC is enveloped almost entirely by the SBC. The surface clusters identified by the k-means algorithm coincide approximately with the shift in the Texas Coastal Current from downcoast prior to ~28 July to upcoast thereafter; this observation mirrors the change in cluster designations seen in glider M13 (Dissertation Chapter 2), although M7 was located further east than M13. Therefore, the surface clusters are designated the “downcoast surface” (DCS, Figure 4.6 green), the “upcoast surface” (UCS, Figure 4.6 gold), and the “deep surface” (DPS, Figure 4.6 orange) corresponding to the surface cluster which was situated primarily atop the DBC. While the M7 cluster designation are generally descriptive of the cluster locations within the water column, it is important to note that the surface clusters do at times intrude to the bottom. Additionally, portions of the “deep” clusters are found within the shallower water depths otherwise dominated by the “shallow” clusters earlier in the glider deployment. Therefore, it is most appropriate to consider each of the clusters as distinct water masses and also to remember that the designations are given simply as a matter of reference for convenience of their description rather than strict boundaries of their location along the shelf or within the water column.

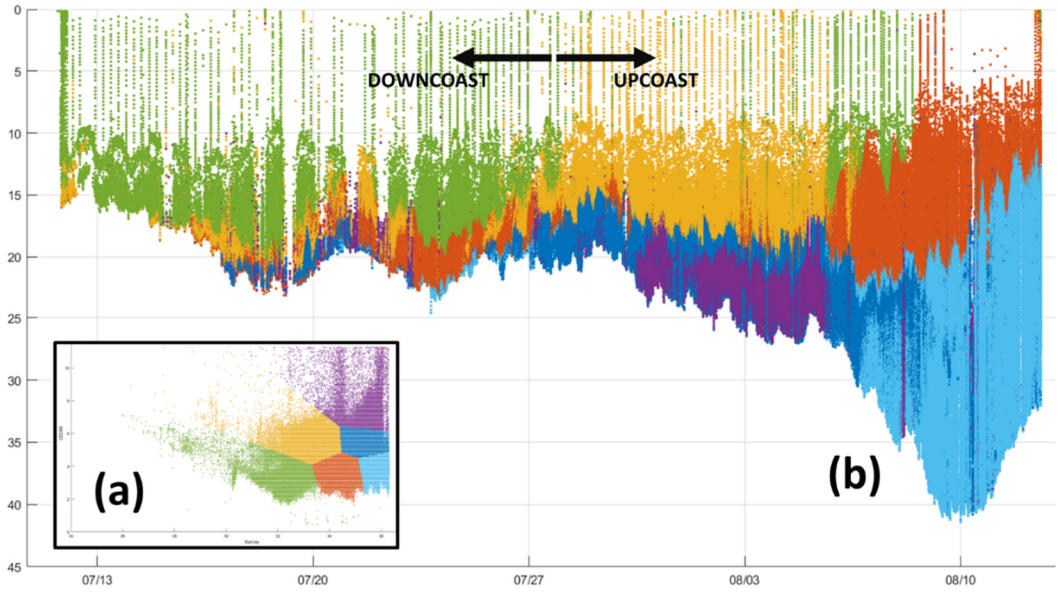

Figure 4.6 (a) M7 CDOM vs. salinity plot, with k-means cluster colors; (b) M7 Hovmöller diagram with colors corresponding to cluster colors from (a), arrows showing the approximate date at which the Texas Coastal Current shifted from downcoast to upcoast. Note that, apart from intermittent surfacing to transmit data, the glider was kept below 10m depth in order to avoid ship impacts.

As the name implies, the MXC observed the largest mean CDOM value (7.117 mg L^{-1}, $\sigma = 0.826$) while the DCS and DPS had the lowest, each with mean CDOM values of 3.345 mg L^{-1} ($\sigma = 0.725$) and 3.75 mg L^{-1} ($\sigma = 0.575$), respectively. Mean salinity values were > 35 in all the bottom clusters and < 35 in the surface clusters, with the DCS featuring the lowest mean salinity (32.31, $\sigma = 0.79$). Kruskal-Wallis ANOVA demonstrated that CDOM values were significantly different for each cluster, as were salinity values, which indicates that the k-means algorithm was successful in segmenting the M7 dataset into partitions all distinct from one another.

M7 mean BIX values were > 1 in all clusters, indicating that the majority of the FDOM encountered by the glider was of aquatic bacterial origin (Huguet et al. 2009).

Mean CMR values, however, were > 1 (primarily terrestrial origin) only within the surface clusters. The bottom clusters all featured mean CMR < 1 (marine origin), with the DBC featuring the lowest mean (0.62, $\sigma = 0.28$). BIX and CMR were each significantly different among all clusters, according to Kruskal-Wallis ANOVA. Comparing BIX and CMR to each other in each of the M7 clusters, through calculation of the determination coefficient, produced $R^2 > 0.4$ in only the SBC ($R^2 = 0.447$); the two FDOM properties were otherwise not correlated with one another in the remaining clusters.

Mean dissolved oxygen concentrations were > 3 ml L^{-1} in all clusters but the MXC (mean oxygen = 2.98, $\sigma = 0.68$). Apart from the MXC and SBC (mean oxygen = 3.08, $\sigma = 0.76$), the remaining clusters all observed mean dissolved oxygen > 3.4 ml L^{-1}. AOU reflected the oxygen concentration distribution among clusters, with the largest mean AOU being found in the MXC (mean AOU = 1.74, $\sigma = 0.726$) and the SBC (mean AOU = 1.6, $\sigma = 0.83$). The surface clusters all contained mean AOU values < 1.05 ml L^{-1}, the lowest of which was found in the DCS (mean AOU = 0.829, $\sigma = 0.289$). No negative AOU values were encountered by the M7 glider. Finally, the Kruskal-Wallis ANOVA revealed that the UCS and DCS were not significantly different with respect to oxygen concentration, but that AOU of all clusters differed significantly.

Hypoxia was encountered in all 6 of the MS7 clusters, with most of the hypoxia (~82%) found within the bottom clusters; of all the observations of hypoxia made by the M7 glider, ~39% were found within the SBC, ~24% within the DBC, and ~19% within the MXC. The remaining ~18% of hypoxic observations were usually found in instances

when the surface clusters intruded toward the bottom. ~15%, ~12%, and ~10% of the SBC, MXC, and DBC, respectively, were hypoxic, whereas hypoxic observations comprised only ~6%, ~4%, and ~0.5% of the UCS, DPS, and DCS clusters, respectively.

As with the MCH Acrobat data, comparisons were made between the M7 modeled FDOM parameters and oxygen/AOU. BIX and oxygen were strongly correlated ($R^2 > 0.7$) in the DPS and UCS, but less so in the DCS, which had the lowest R^2 among all clusters (DCS BIX/oxygen $R^2 = 0.42$, BIX/AOU $R^2 = 0.41$). M7 CMR and oxygen were not strongly correlated, with only the UCS finding CMR/AOU $R^2 > 0.5$ (UCS CMR/oxygen $R^2 = 0.48$, CMR/AOU $R^2 = 0.517$).

4.4. Discussion

4.4.1. Inferences from FDOM Modeling

Of the data from the 6 MCH cruises analyzed in this study, both BIX and CMR values were more variable in 2011 (MARS flood year) and 2013 (average MARS discharge year) than in 2012 (MARS drought year), especially in the Sub-regions nearest to the Mississippi River delta (Sub-regions A and B). The fact that standard deviations were largest in the east TXLA shelf and lowest in the west reflects the more turbulent nature of the biogeochemistry in the Sub-regions that lie in closer proximity to the Mississippi River. While Sub-region C is located due south of Atchafalaya Bay, Sub-region C BIX variability was similar to the variability found within Sub-regions D and E. The same was true of CMR in Sub-region C, except in June 2013, in which year mean CMR values for the western Sub-regions were more pronounced and more variable than

for those Sub-regions in other years. Mean CMR was usually larger in the eastern Sub-regions than in the west; therefore, the westernmost Sub-regions are less influenced by the Mississippi than are the eastern Sub-regions and, thus, the FDOM conditions in the west TXLA shelf are usually more homogenous. The Mississippi River, therefore, appears most impactful to FDOM freshness variability in the TXLA shelf region south of Terrebonne Bay and east to Southwest Pass, while outflow from Atchafalaya Bay is not as strong of a control on FDOM variability, in general.

August mean BIX values were largest in Sub-region E in 2011 and 2012 (and second largest in 2013), which can be explained by various possibilities or a combination thereof. Firstly, Sub-region E is well-ventilated enough in August that remineralization of DOM (and subsequent creation of new DOM) is not impeded by lack of oxygen (otherwise anoxic or hypoxic conditions would inhibit microbial activity). Among the August cruises, Sub-region E always had the largest mean dissolved oxygen concentration and was the only Sub-region with no observed hypoxia. Additionally, Sub-region E is the shallowest of the Sub-regions of this study, with a maximum depth of only ~17m (as compared to maximum depths > 20m for Sub-regions C and D, and > 35m for Sub-regions A and B); therefore, ventilation of Sub-region E is easily accomplished by even moderate atmospheric wind conditions. Another possibility is that the shift in the Texas Coastal Current from downcoast to upcoast in the late summer results in fresher FDOM being transported to Sub-region E from the regions southwest of Galveston Bay; this possibility is linked to Sub-region E's distant proximity from the MARS endmembers, which are potential sources of less fresh DOM. Indeed, Sub-region

E featured the lowest mean CMR among all Sub-regions in each cruise, which demonstrates that MARS FDOM is not present in as sizable amounts in the westernmost reaches of the TXLA shelf as compared to the east, either because MARS FDOM is mostly consumed or mixed out before it can reach Sub-region E. It is also notable that Sub-region E mean CMR for the August cruises was always lower than in June, a reflection of the fact that river discharge is decreased enough by August and/or that the upcoast shift in the Texas Coastal Current impedes the flow of MARS-derived terrestrial material that far west in the late summer.

The June 2013 cruise was the only cruise for which Sub-region E mean CMR was >1, suggesting that the downcoast Texas Coastal Current was able to transport MARS DOM further in June 2013 than in any of the other cruises. While Sub-region E data were unavailable for June 2011, it is likely that Sub-region E mean CMR values would also have exceeded 1 that year, given that MARS discharge was even larger in 2011 than in 2013. While Sub-region E is near Galveston Bay, the downcoast Texas Coastal Current should carry most Galveston-derived terrestrial material away from Sub-region E in the early summer. Even after the upcoast shift, lower average CMR values in August suggest that Galveston Bay does not usually contribute a significant amount of terrestrial material to the TXLA shelf (except during anomalous flood events, as in 2015, see Dissertation Chapter 1). Therefore, terrestrial DOM originating directly from the MARS endmembers, most of which likely derive from Atchafalaya Bay, can reach the western TXLA shelf and be present in concentrations exceeding marine DOM in the early summer months (see Chen et al. 2000 and Quigg et al. 2011 for similar results).

Sub-pycnocline mean BIX values were highest in Sub-region E in the August 2011 and 2012 cruises, suggesting that fresh FDOM is consumed more readily in the eastern Sub-regions throughout the summer and/or that more fresh FDOM creation happens in the western TXLA shelf in the summer. The upcoast shift in the Texas Coastal Current, which would have occurred not long before the August cruises, could also have brought fresh organic matter from other regions of the Gulf and would meet Sub-region E first. The relatively shallow depth of Sub-region E also allows for FDOM to be transported more easily to the sub-pycnocline from the upper water column than in the other deeper Sub-regions. Additionally, any FDOM being propelled westward by MARS plumes would accumulate in Sub-region E as the upcoast current in August inhibits its transit further downcoast. Sub-region E did feature the largest CMR values in the June 2012 and 2013 cruises, prior to the upcoast shift. However, mean CMR values were larger (suggesting more terrestrial-derived FDOM) in Sub-regions C and D than in Sub-region E in August 2011 and 2012, suggesting that such a "pileup" of terrestrial material, caused by the upcoast current, actually occurs in Sub-regions C and D. No pycnocline was observed in Sub-region E in August 2013, but CMR values were still > 1 in Sub-regions B and D in that year. Therefore, it is most likely that the upcoast current in August transports terrestrial FDOM from Sub-region E to its neighboring Sub-regions to the east and that the upcoast current inhibits the flow of new terrestrial material from the MARS endmembers further downcoast than Sub-region D. This "pileup" effect is supported by the observation that the largest increase in sub-pycnocline mean CMR between June and August occurred in Sub-region D in 2013. Additionally, while no data

were available for Sub-region D in 2012, Sub-region C in 2012 also saw a significant increase in mean CMR from June to August, whereas Sub-region E mean CMR decreased in 2012. In the case of both BIX and CMR, variability was much lower in the sub-pycnocline than above-pycnocline. The low variability in sub-pycnocline FDOM is indicative of a more stable biogeochemical environment therein, whereas above the pycnocline, more processes remove or create fresh organic matter (e.g. photodegradation, primary production and associated grazing activities, influx of material from buoyant, less dense freshwater plumes, etc.) causing increased variability. Low BIX variability could also be evidence of a near-equilibrium environment in the sub-pycnocline, with FDOM creation occurring at rates close to and/or in tandem with FDOM consumption rates.

In the June 2011 and 2013 cruises, large mean BIX values above the pycnocline in Sub-regions A and B imply that fresh FDOM was supplied directly by the Mississippi River and/or that resource-rich waters from the river encouraged creation of fresh FDOM in the eastern TXLA shelf; this determination is also supported by the large mean CMR values (i.e. more terrestrial material) in those Sub-regions for the same cruise years, with Sub-regions A or B featuring the largest mean CMR in each cruise. Above-pycnocline BIX values were also more variable in the eastern Sub-regions, again signifying the importance of the Mississippi River in modulating TXLA FDOM. Larger mean BIX values above the pycnocline in Sub-region E in August of 2011 and 2012 again hint at the "pileup" of fresh FDOM therein due to the upcoast current. Lower above-pycnocline mean CMR values in the western Sub-regions (sometimes lower even

than sub-pycnocline CMR) indicates that most terrestrial material is photodegraded, consumed, or mixed out before it reaches the western TXLA shelf. The photodegradation and consumption of terrestrial material would lead to creation of fresh FDOM, explaining the larger BIX values in Sub-region E. The observation of similar sub- and above-pycnocline mean BIX values (with Sub-region B not significantly different) in the 2012 drought year reiterates the importance of the MARS endmembers in regulating TXLA shelf FDOM conditions, as the below-average MARS discharge apparently led to decreased BIX values. River discharge in 2011 and 2013 corresponded to larger and more variable above-pycnocline mean BIX in the June cruises of those years, because fresh FDOM from the rivers is carried in more buoyant freshwater and/or because production of new material was increased in the resource-rich plume water. Whereas sub-pycnocline BIX increased from June to August in 2011, above-pycnocline BIX in that year dropped off significantly from June to August, either due to photodegradation (which would occur more readily in the upper water column due to increased sunlight exposure), remineralization, sinking of material beneath the pycnocline, or mixing of FDOM elsewhere throughout the summer. It is important to reiterate that, of all cruise Sub-regions, ANOVA showed that only BIX in MS06 Sub-region B had values above and below the pycnocline not significantly different from one another; therefore, in every Sub-region, sub-pycnocline CMR always significantly differed from above-pycnocline CMR and, outside of MS06 Sub-region B, so did BIX. The reason for the homogeneity of fresh FDOM throughout the water column of MS06 Sub-region B was that decreased MARS discharge in the 2012 drought year did not

introduce sufficient amounts of fresh FDOM directly to the above-pycnocline region of the water column or provide sufficient resources to stimulate fresh FDOM production therein. The lower August 2012 above-pycnocline mean CMR values for Sub-regions A and B, as compared to the other August cruises, also confirms the lack of Mississippi River influence that year.

Strong positive correlations between BIX and CMR in the overall water column (i.e. composite of sub- and above-pycnocline) as well as within the above-pycnocline itself of Sub-regions A and B in June 2011 and both 2013 cruises confirms that the Mississippi River is responsible for directly introducing terrestrial DOM that is "fresher" and/or that the resources found within waters laden with terrestrial DOM are sufficient to encourage new FDOM production in the marine environment. The same effect can be observed with DOM from the Atchafalaya plume, as Sub-region C, the Sub-region nearest to Atchafalaya Bay, also observed convincing positive relationships between BIX and CMR in the 2013 cruises. A relatively strong negative relationship between BIX and CMR in Sub-region E in August 2012 implies that consumption of terrestrial material, perhaps supplied by Texas rivers due to the upcoast current, was responsible for creation of fresh FDOM on the west TXLA shelf. The August 2012 cruise was the only cruise that saw any strong relationships between BIX and CMR in the sub-pycnocline, with all MS06 Sub-regions observing strong negative relationships between the two. The upcoast shift in the Texas Coastal Current in late summer prevented what little MARS-derived material there would have been in the 2012 drought year from transporting downcoast, which compounded with introduction of terrestrial material from Texas rivers, both of

which occurrences resulted in consumption of terrestrial material from these sources and concurrent creation of fresh FDOM.

In the overall Sub-regions (i.e. composite water column, regardless of pycnocline) CMR and oxygen/AOU were more strongly related (i.e. larger R^2) than were BIX and oxygen/AOU in Sub-region A in all cruises and in Sub-region B in all cruises but June 2011. The positive relationship between CMR and oxygen/AOU in the Sub-regions nearest the Mississippi River suggests that terrestrial material is preferred for remineralization (and concurrent respiration) therein and that the DOM source is more important than its freshness in regulating oxygen consumption and concentration in the eastern TXLA shelf. A somewhat strong relationship between CMR and oxygen/AOU in Sub-region C of the 2014 glider dataset suggests that terrestrial-derived material from Atchafalaya Bay can also encourage oxygen consumption in that region of the TXLA shelf. Otherwise, the relationship between BIX and oxygen/AOU was usually stronger than CMR and oxygen/AOU in Sub-regions C, D, and E; the strong positive correlation between BIX and oxygen/AOU, and lack of a strong correlation between CMR and oxygen/AOU, in these Sub-regions indicates that FDOM freshness, rather than FDOM source, is more impactful to controlling oxygen consumption and concentration in the middle and western TXLA shelf.

Beneath the pycnocline, BIX and oxygen/AOU were always more strongly correlated than were CMR and oxygen/AOU, signifying the importance of FDOM freshness rather than FDOM source on oxygen controls in the lower water column. The strongest correlations, all of which were positive, between BIX and oxygen/AOU were

generally found outside of Sub-region A, suggesting that the oxygen consumption in Sub-region A is related to other factors than FDOM freshness or source; the biogeochemistry in Sub-region A is subject to the turbulent conditions brought on by proximity to the Mississippi River outflow. Elsewhere on the TXLA shelf, oxygen consumption occurs more readily when fresher FDOM is present, regardless of the fresh FDOM source. If marine-derived fresh DOM (e.g. from an algal bloom) were solely responsible for observed sub-pycnocline oxygen conditions, then both BIX and CMR would have strong correlations with oxygen/AOU in the sub-pycnocline. Instead, sub-pycnocline oxygen conditions can be accounted either to marine or terrestrial sources of FDOM, with more oxygen being consumed where the FDOM is fresher. Above the pycnocline, the same dynamic holds true, with stronger relationships between BIX and oxygen/AOU than for CMR most of the time; however, there are fewer instances overall of significant correlations for either FDOM property. FDOM in the upper water column is more subject to removal by photodegradation or mixing than by remineralization/respiration processes, which explains the lack of a relationship between FDOM properties and oxygen above the pycnocline.

The observed relationships between FDOM and oxygen also has implications for TXLA shelf hypoxia. The observation of lower BIX in hypoxic waters indicates that hypoxia is indeed fueled by fresher DOM, as is commonly believed to be the case on the TXLA shelf. However, the presence of lower CMR in hypoxic waters implies that hypoxia is also fueled at times by terrestrial material directly from freshwater endmembers, rather than by algal bloom byproducts exclusively. However, the

observation of lower CMR in hypoxic waters could alternatively be explained by the fact that non-hypoxic waters are typically those above the pycnocline and that terrestrial material (i.e. larger CMR) should coincide with lower salinity waters, also generally above the pycnocline. Finally, yet another explanation is that the high rates of remineralization below the pycnocline create an excess of marine-like DOM byproducts such that CMR is skewed heavily away from the terrestrial end of the spectrum.

4.4.2. Inferences from Dataset Clustering

K-means clustering was successful in dividing the MS3L12 Acrobat data into subdivisions that accurately represent distinct areas of the water column within which different biogeochemical processes occur. BIX values were larger in the upper water column and lower within and beneath the pycnocline, where nearly all hypoxia was found; the lower BIX values corresponded to lower oxygen values, reflecting again the importance of FDOM freshness in regulating TXLA shelf oxygen consumption. Although BIX and oxygen/AOU were not strongly correlated in the bottom cluster, this observation is explained by the fact, while bottom-water respiration does occur due to the presence of fresh FDOM, hypoxic conditions ultimately inhibit consumption of the remaining fresh FDOM, which otherwise would be respired if oxygen conditions were sufficient. A slight positive correlation between CMR and oxygen (negative for CMR/AOU) in the bottom and pycnocline MS3L12 clusters suggests that terrestrial-derived material was also consumed. Conversely, a negative CMR/oxygen correlation (positive for CMR/AOU) in the surface cluster indicates that marine FDOM is more

readily consumed in the upper water column and/or that other removal processes (e.g. photodegradation and mixing) are a more impactful control upon terrestrial material in the surface layers than is microbial consumption. Oxygen concentration was also largest, while AOU was lowest, in the surface cluster, which is a reflection of ventilation of more oxygen from the atmosphere; low chl-a values in the surface cluster do not suggest the oxygen values were a result of oxygen creation by phytoplankton. While BIX and oxygen/AOU were correlated in the surface cluster, the relative lower rates of respiration therein (as evidenced by low AOU values) hint that other processes (e.g. photodegradation) were also responsible for removing FDOM from the surface such that there was not enough labile DOM present in the surface cluster to be consumed.

The different clusters of the summer 2014 data from the M7 glider are also representative of unique areas of the water column. However, the clusters were at times themselves segmented in time. For example, a portion of the DSC was identified even after the Texas Coastal Current shift to upcoast. Each of these separate regions of the DSC could represent different water masses or, conversely, could each have been Atchafalaya water, for example, that was impeded by the upcoast shift after ~28 July but was then later encountered by the glider again from ~5-9 August. The MXC being found following the shift in Texas Coastal Current implies that the current impeded flow of high-CDOM waters downcoast, essentially causing a buildup of CDOM-laden water south of Atchafalaya Bay from ~30 July to ~5 August, an observation that was also noted in the MCH Acrobat data. Similarly, the UCS observed mean CDOM values greater than the other surface clusters, which is explained by the same accumulation of

MARS waters being impeded by the upcoast current. The "pileup" effect, induced by the upcoast current, is also apparent in the CMR and salinity observations, both of which are highest in the UCS. While oxygen was not necessarily impacted by the currents in the same way, the UCS did observe the strongest correlation between BIX and oxygen as well as CMR and oxygen/AOU, suggesting that microbial remineralization of FDOM in that cluster was an important process controlling FDOM therein, perhaps more so than photodegradation or mixing. Still, the largest AOU values were found in clusters corresponding to the ocean bottom, showing that respiration processes where more active in the lower water column. Of the bottom clusters, BIX was correlated with oxygen/AOU only in the SBC and DBC; however, the MXC was also the cluster with lowest mean BIX, largest mean AOU, and lowest mean oxygen concentration, all of which observations combine to suggest that the low oxygen conditions encountered in the MXC were a result of consumption of fresh FDOM, but that hypoxic conditions prevented consumption of more fresh FDOM that otherwise would have been consumed had hypoxia not been present.

4.5. Summary and Conclusion

The relative degree of FDOM freshness and approximate FDOM source (i.e. terrestrial or marine) was estimated over a 3-year period in the summertime along the TXLA shelf in the northern Gulf of Mexico. K-means clustering of CDOM/salinity from an underwater buoyancy glider and a line of towfish observations was also performed. FDOM on the TXLA shelf was strongly influenced by a variety of processes and by

MARS discharge, particularly on the eastern regions of the shelf nearer to the Mississippi River delta; the impacts to shelf FDOM are lessened when MARS discharge is lower (e.g. during the drought year of 2012). Changes in FDOM characteristics from early summer to late summer in the middle TXLA shelf, south of Atchafalaya Bay, and in the western TXLA shelf also coincided with the shift in the Texas Coastal Current from downcoast to upcoast. FDOM was at times observed to be simultaneously "fresh" (according to the Huguet et al. 2009 definition) and of terrestrial origin. Fresh FDOM was often positively related to oxygen concentration and simultaneously negatively related to AOU, regardless of whether the FDOM was derived from terrestrial or marine sources. Relationships between FDOM freshness and occurrences of hypoxia were also found; i.e. MARS endmembers can supply terrestrial DOM that is "fresh" (according to the Huguet et al. 2009 definition) and that encourages respiration and leads to hypoxia. Therefore, not all hypoxia on the TXLA shelf necessarily occurs as a function of algal bloom derivatives; shelf processes and marine origin material have more influence further from the MARS endmembers. Stratification, however, is an enduring necessity for the formation of bottom-water hypoxia, regardless of proximity to freshwater sources. Stratified waters generally lead to bottom hypoxia, but the organic fuel for respiration is increasingly marine (rather than terrestrial) on the western shelf. Interannual variability in environmental conditions (e.g. river discharge, winds, nutrient loading, etc.) also play a role in determining what types of DOM drive TXLA shelf hypoxia. Vertically, significant differences in the nature of FDOM were also observed above and below the pycnocline: in the lower water column, microbial remineralization

and respiration were the primary control of FDOM, and vice versa (i.e., the nature of FDOM was a control on the rates of respiration), whereas photodegradation and mixing were more likely the chief mechanisms for regulating FDOM in the upper layers. Ultimately, this research provides a foundation for a quantitative determination of what percentage of the total hypoxic area is directly attributable to the MARS effluent and what percentage is not.

5. CONCLUSIONS

5.1. Summary and Conclusions

The concentration and quality of dissolved organic matter (DOM) on the Texas-Louisiana (TXLA) shelf is an important regulator of both heterotrophic and autotrophic processes. Fresh DOM (i.e. DOM that was created more recently) in the ocean can be created as a byproduct of microbial processes or introduced into the ocean directly by rivers. The light-absorbing characteristics of chromophoric DOM (CDOM) also impede photosynthetic processes that require light at similar wavelengths as the wavelengths at which CDOM absorbs. Fluorescent DOM (FDOM) is a subset of CDOM that absorbs and subsequently emits light. The wavelengths at which FDOM emissions occur can be used to distinguish when and where the FDOM was created. Comparisons of FDOM and salinity are commonly used to gauge the degree of conservative mixing in a system. This dissertation used an established machine learning technique, k-means clustering, in a novel application upon FDOM/salinity matrices. Clustering in this way can be used to formulate a precise representation of water column structure, revealing the locations of discrete water masses, as confirmed by analytical techniques such as ANOVA and PCA, and clarifying the biogeochemical or physical processes by which different water masses can be distinguished. For example, some clusters are isolated due to increased oxygen demand therein, even to the point of hypoxia, whereas other clusters are identified due to the influence of a river plume or wind-induced upwelling. Similarly, clustering can be used to determine the extent of and mechanisms by which shelf DOM is mixed

nonconservatively vs. conservatively. Ultimately, the k-means clustering approach of TXLA shelf data produced a representation of the water column that is consistent with both the known climatological patterns of the region, including shifts in the Texas Coastal Current, as well as concurrent observations of meteorological and oceanographic data.

Furthermore, this dissertation utilized spectrofluorometric methods of discrete water sample analysis to devise algorithms for estimating FDOM source and freshness in existing subsurface ocean datasets, such as underwater buoyancy gliders and towed systems. Fresh FDOM is typically more labile and, as a result, leads to increased oxygen demand. On the TXLA shelf, density gradients induce strong stratification in the summer months such that the bottom ocean layers are not well-ventilated. FDOM modeling results determined that, when fresh DOM is present in sufficient quantities in the lower water column, the increased oxygen demand can lead to hypoxia, regardless of the source of the DOM. Within TXLA shelf river plumes, low light availability, either due to absorption by CDOM or scattering by riverine inorganic material, resulted in decreased chl-*a* concentrations. Despite the lack of algal blooms in the river plumes, hypoxic conditions were still observed. This dissertation, therefore, showed that the fresh DOM that induces hypoxia on the TXLA shelf can originate either from microbial sources or DOM supplied directly from a freshwater endmember, such as the Mississippi River. While microbial activity was the primary control on lower water column FDOM variability, FDOM in the upper layers was controlled more by photodegradation by sunlight and conservative mixing.

5.2. Scientific Hypotheses Addressed

Three hypotheses were addressed by this dissertation in the following ways:

- **H1**: The results of Chapters 2 and 4 demonstrated that hierarchical clustering by the k-means approach can isolate distinct water masses within subsurface buoyancy glider and towed system datasets. The clusters represent water masses that are distinguished by differences in density and also by their respective ongoing biogeochemical processes. For example, in lower water column clusters, oxygen demand is generally increased. Additionally, shifts in ocean currents and winds also bring about changes in the ocean structure that are reflected by the clustering results.
- **H2**: The freshest FDOM was indeed found on the western side of the TXLA shelf, both above and below the pycnocline, but typically only in the late summer. In the early summer, fresher FDOM was found above the pycnocline on the eastern TXLA shelf, nearest to the Mississippi River, whereas sub-pycnocline FDOM freshness was more homogenous in early summer.
- **H3**: TXLA shelf FDOM freshness was always positively correlated to oxygen concentration and negatively correlated to AOU. Strong correlations could be found both in early summer and late summer, thus indicating that fresh FDOM continues to be introduced to the TXLA shelf, either autochthonously or allochthonously, and can fuel oxygen demand throughout the summertime. However, weak correlations between oxygen and FDOM freshness were also observed in other locations in early and late summer, indicating that oxygen demand is not always related to how recently FDOM was created.

5.3. Future Research Directions

This dissertation clarified FDOM dynamics on a continental shelf, the results of which could benefit researchers interested in how DOM modulates other marine processes, including hypoxia. The connections that were found between chl-*a* and modeled FDOM also suggest the need to examine the role of light-absorbed DOM in regulating primary production. The data used in this study were exclusively summertime data; therefore, future applications would benefit from collecting data from other times of year to understand how DOM varies seasonally. While TXLA shelf river impacts have often been nutrient-centric, this study has demonstrated the importance of considering the role of both marine and terrestrial DOM in the biogeochemistry of the northern Gulf of Mexico.

Clustering via the k-means algorithm is a definitive method by which large subsurface oceanographic datasets should be analyzed. Other more complex machine learning methods of hierarchical clustering, such as DBSCAN, should also continue to be explored to maximize the information gained from clustering glider, towfish, and other marine technology platform data. Future research could utilize the FDOM algorithms that were developed in this study to estimate FDOM freshness and source in other existing datasets or to create similar FDOM algorithms in other regions. Additionally, FDOM research would benefit from tuning field-based fluorometers to additional wavelengths associated with FDOM emission in order to measure FDOM freshness and source continuously *in situ*.

REFERENCES

Ahmad, H. (2019). Machine Learning Applications in Oceanography. *Aquatic Research*, 161–169. https://doi.org/10.3153/AR19014

Bauer, P., Thorpe, A., & Brunet, G. (2015). The quiet revolution of numerical weather prediction. *Nature*, *525*(7567), 47–55. https://doi.org/10.1038/nature14956

Bender, L. C., Guinasso, N. L., Walpert, J. N., Lee, L. L., Martin, R. D., Hetland, R. D., Baum, S. K., & Howard, M. K. (2007). Development, Operation, and Results From the Texas Automated Buoy System. Gulf of Mexico Science, 25(1). https://doi.org/10.18785/goms.2501.04

Bianchi, T.S., DiMarco, S. F., Smith, R. W., & Schreiner, K. M. (2009). A gradient of dissolved organic carbon and lignin from Terrebonne–Timbalier Bay estuary to the Louisiana shelf (USA). Marine Chemistry, 117(1–4), 32–41. https://doi.org/10.1016/j.marchem.2009.07.010

Bianchi, T. S., DiMarco, S. F., Cowan, J. H., Hetland, R. D., Chapman, P., Day, J. W., & Allison, M. A. (2010). The science of hypoxia in the Northern Gulf of Mexico: A review. *Science of The Total Environment*, *408*(7), 1471–1484. https://doi.org/10.1016/j.scitotenv.2009.11.047

Bianchi, T. S., Wysocki, L. A., Schreiner, K. M., Filley, T. R., Corbett, D. R., & Kolker, A. S. (2011). Sources of Terrestrial Organic Carbon in the Mississippi Plume Region: Evidence for the Importance of Coastal Marsh Inputs. Aquatic Geochemistry, 17(4–5), 431–456. https://doi.org/10.1007/s10498-010-9110-3

Bianchi, T. S., Garcia-Tigreros, F., Yvon-Lewis, S. A., Shields, M., Mills, H. J., Butman, D., Osburn, C., Raymond, P., Shank, G. C., DiMarco, S. F., Walker, N., Reese, B. K., Mullins-Perry, R., Quigg, A., Aiken, G. R., & Grossman, E. L. (2013). Enhanced transfer of terrestrially derived carbon to the atmosphere in a flooding event: ENHANCED TRANSFER OF TDOC TO THE ATMOSPHERE. Geophysical Research Letters, 40(1), 116–122. https://doi.org/10.1029/2012GL054145

Blough, N. V., and R. Del Vecchio. (2002). Chromophoric DOM in the coastal environment, p. 509–546. In D. A. Hansell and C. A. Carlson [eds.], Biogeochemistry of marine dissolved organic matter. Academic Press.

Boyd, T. J., & Osburn, C. L. (2004). Changes in CDOM fluorescence from allochthonous and autochthonous sources during tidal mixing and bacterial degradation in two coastal estuaries. Marine Chemistry, 89(1–4), 189–210. https://doi.org/10.1016/j.marchem.2004.02.012

Burdige, D. J., Kline, S. W., & Chen, W. (2004). Fluorescent dissolved organic matter in marine sediment pore waters. Marine Chemistry, 89(1–4), 289–311. https://doi.org/10.1016/j.marchem.2004.02.015

Cai, W.-J., Hu, X., Huang, W.-J., Murrell, M. C., Lehrter, J. C., Lohrenz, S. E., Chou, W.-C., Zhai, W., Hollibaugh, J. T., Wang, Y., Zhao, P., Guo, X., Gundersen, K., Dai, M., & Gong, G.-C. (2011). Acidification of subsurface coastal waters enhanced by eutrophication. Nature Geoscience, 4(11), 766–770. https://doi.org/10.1038/ngeo1297

Carr, N., Davis, C. E., Blackbird, S., Daniels, L. R., Preece, C., Woodward, M., & Mahaffey, C. (2019). Seasonal and spatial variability in the optical characteristics of DOM in a temperate shelf sea. Progress in Oceanography, 177, 101929. https://doi.org/10.1016/j.pocean.2018.02.025

Chari, N., Rao, P. S., & Sarma, N. S. (2013). Fluorescent dissolved organic matter in the continental shelf waters of western Bay of Bengal. Journal of Earth System Science, 122(5), 1325–1334. https://doi.org/10.1007/s12040-013-0349-0

Chen, R. F., & Gardner, G. B. (2004). High-resolution measurements of chromophoric dissolved organic matter in the Mississippi and Atchafalaya River plume regions. Marine Chemistry, 89(1–4), 103–125. https://doi.org/10.1016/j.marchem.2004.02.026

Cho, K., Reid, R. O., & Nowlin, W. D. (1998). Objectively mapped stream function fields on the Texas-Louisiana shelf based on 32 months of moored current meter data. *Journal of Geophysical Research: Oceans*, *103*(C5), 10377–10390. https://doi.org/10.1029/98JC00099

Coble, P. G. (1996). Characterization of marine and terrestrial DOM in seawater using excitation-emission matrix spectroscopy. Marine Chemistry, 51(4), 325–346. https://doi.org/10.1016/0304-4203(95)00062-3

Coble, P. G. (2007). Marine Optical Biogeochemistry: The Chemistry of Ocean Color. 17.

Cochrane, J. D., & Kelly, F. J. (1986). Low-frequency circulation on the Texas-Louisiana continental shelf. *Journal of Geophysical Research*, *91*(C9), 10645. https://doi.org/10.1029/JC091iC09p10645

D'Sa, E. (2008). Colored dissolved organic matter in coastal waters influenced by the Atchafalaya River, USA: Effects of an algal bloom. Journal of Applied Remote Sensing, 2(1), 023502. https://doi.org/10.1117/1.2838253

D'Sa, E. J., & DiMarco, S. F. (2009). Seasonal variability and controls on chromophoric dissolved organic matter in a large river-dominated coastal margin. Limnology and Oceanography, 54(6), 2233–2242. https://doi.org/10.4319/lo.2009.54.6.2233

D'Sa, E. J., Joshi, I., & Liu, B. (2018). Galveston Bay and Coastal Ocean Optical-Geochemical Response to Hurricane Harvey From VIIRS Ocean Color. Geophysical Research Letters, 45(19), 10,579-10,589. https://doi.org/10.1029/2018GL079954

DiMarco, S. F., Kelly, F. J., Zhang, J., & Guinasso, N. L. (1995). Directional Wave Spectra on the Louisiana-Texas Shelf During Hurricane Andrew. Journal of Coastal Research, S1(21), 217-233.

DiMarco, S. F., & Reid, R. O. (1998). Characterization of the principal tidal current constituents on the Texas-Louisiana shelf. *Journal of Geophysical Research: Oceans*, *103*(C2), 3093–3109. https://doi.org/10.1029/97JC03289

DiMarco, S. F., Chapman, P., Walker, N., & Hetland, R. D. (2010). Does local topography control hypoxia on the eastern Texas–Louisiana shelf? *Journal of Marine Systems*, *80*(1–2), 25–35. https://doi.org/10.1016/j.jmarsys.2009.08.005

DiMarco, S. F., & Zimmerle, H. M. (2017). MCH Atlas: Oceanographic Observations of the Mechanisms Controlling Hypoxia Project. (TAMU-SG-17-601). Texas A&M University, Texas Sea Grant.

Feng, Y., DiMarco, S. F., & Jackson, G. A. (2012). Relative role of wind forcing and riverine nutrient input on the extent of hypoxia in the northern Gulf of Mexico: WIND FORCING ON GULF HYPOXIA. Geophysical Research Letters, 39(9). https://doi.org/10.1029/2012GL051192

Feng, Y., Fennel, K., Jackson, G. A., DiMarco, S. F., & Hetland, R. D. (2014). A model study of the response of hypoxia to upwelling-favorable wind on the northern Gulf of Mexico shelf. *Journal of Marine Systems*, *131*, 63–73. https://doi.org/10.1016/j.jmarsys.2013.11.009

Fennel, K., Hetland, R., Feng, Y., & DiMarco, S. (2011). A coupled physical-biological model of the Northern Gulf of Mexico shelf: Model description, validation and analysis of phytoplankton variability. Biogeosciences, 8(7), 1881–1899. https://doi.org/10.5194/bg-8-1881-2011

Fichot, C. G., & Benner, R. (2012). The spectral slope coefficient of chromophoric dissolved organic matter (S 275-295) as a tracer of terrigenous dissolved organic carbon in river-influenced ocean margins. Limnology and Oceanography, 57(5), 1453–1466. https://doi.org/10.4319/lo.2012.57.5.1453

Flament, P. (2002). A state variable for characterizing water masses and their diffusive stability: Spiciness. Progress in Oceanography, 54(1–4), 493–501. https://doi.org/10.1016/S0079-6611(02)00065-4

Forgy, Edward W. (1965). Cluster analysis of multivariate data: efficiency versus interpretability of classifications. Biometrics. 21 (3): 768–769.

Forrest, D. R., Hetland, R. D., & DiMarco, S. F. (2011). Multivariable statistical regression models of the areal extent of hypoxia over the Texas–Louisiana continental shelf. Environmental Research Letters, 6(4), 045002. https://doi.org/10.1088/1748-9326/6/4/045002

Gao, L., Gao, Y., Zong, H., & Guo, L. (2019). Elucidating the Hidden Nonconservative Behavior of DOM in Large River-Dominated Estuarine and Coastal Environments. *Journal of Geophysical Research: Oceans*, *124*(6), 4258–4271. https://doi.org/10.1029/2018JC014731

Gao, L., Gao, Y., Song, S., & Zhang, F. (2020). Non-conservative behavior of dissolved organic carbon in the Changjiang (Yangtze River) Estuary and the adjacent East China Sea. Continental Shelf Research, 197, 104084. https://doi.org/10.1016/j.csr.2020.104084

Glenn, S., Schofield, O., Kohut, J., McDonnell, J., Ludescher, R., Seidel, D., Aragon, D., Haskins, T., Handel, E., & Haldeman, C. (2011). The Trans-Atlantic Slocum glider expeditions: A catalyst for undergraduate participation in ocean science and technology. *Marine Technology Society Journal*, *45*(1), 52–67.

Gold-Bouchot, G., Polis, S., Castañon, L. E., Flores, M. P., Alsante, A. N., & Thornton, D. C. O. (2021). Chromophoric dissolved organic matter (CDOM) in a subtropical estuary (Galveston Bay, USA) and the impact of Hurricane Harvey. Environmental Science and Pollution Research. https://doi.org/10.1007/s11356-021-14509-x

Hansen, A. M., Kraus, T. E. C., Pellerin, B. A., Fleck, J. A., Downing, B. D., & Bergamaschi, B. A. (2016). Optical properties of dissolved organic matter (DOM): Effects of biological and photolytic degradation: DOM optical properties following degradation. Limnology and Oceanography, 61(3), 1015–1032. https://doi.org/10.1002/lno.10270

Haulsee, D., Breece, M., Miller, D., Wetherbee, B., Fox, D., & Oliver, M. (2015). Habitat selection of a coastal shark species estimated from an autonomous underwater vehicle. Marine Ecology Progress Series, 528, 277–288. https://doi.org/10.3354/meps11259

Haywood, B. J., White, J. R., & Cook, R. L. (2018). Investigation of an early season river flood pulse: Carbon cycling in a subtropical estuary. Science of The Total Environment, 635, 867–877. https://doi.org/10.1016/j.scitotenv.2018.03.379

Helms, J. R., Stubbins, A., Ritchie, J. D., Minor, E. C., Kieber, D. J., & Mopper, K. (2008). Absorption spectral slopes and slope ratios as indicators of molecular weight,

source, and photobleaching of chromophoric dissolved organic matter. Limnology and Oceanography, 53(3), 955–969. https://doi.org/10.4319/lo.2008.53.3.0955

Helms, J. R., Stubbins, A., Perdue, E. M., Green, N. W., Chen, H., & Mopper, K. (2013). Photochemical bleaching of oceanic dissolved organic matter and its effect on absorption spectral slope and fluorescence. Marine Chemistry, 155, 81–91. https://doi.org/10.1016/j.marchem.2013.05.015

Hetland, R. D., & Campbell, L. (2007). Convergent blooms of Karenia brevis along the Texas coast. Geophysical Research Letters, 34(19), L19604. https://doi.org/10.1029/2007GL030474

Hetland, R. D., & DiMarco, S. F. (2008). How does the character of oxygen demand control the structure of hypoxia on the Texas–Louisiana continental shelf? *Journal of Marine Systems*, *70*(1–2), 49–62. https://doi.org/10.1016/j.jmarsys.2007.03.002

Hong, X., Chang, S. W., Raman, S., Shay, L. K., & Hodur, R. (2000). The Interaction between Hurricane Opal (1995) and a Warm Core Ring in the Gulf of Mexico. *MONTHLY WEATHER REVIEW*, *128*, 19.

Huguet, A., Vacher, L., Relexans, S., Saubusse, S., Froidefond, J. M., & Parlanti, E. (2009). Properties of fluorescent dissolved organic matter in the Gironde Estuary. Organic Geochemistry, 40(6), 706–719. https://doi.org/10.1016/j.orggeochem.2009.03.002

Johnson, K., Berelson, W., Boss, E., Chase, Z., Claustre, H., Emerson, S., Gruber, N., Körtzinger, A., Perry, M. J., & Riser, S. (2009). Observing Biogeochemical Cycles at Global Scales with Profiling Floats and Gliders: Prospects for a Global Array. *Oceanography*, *22*(3), 216–225. https://doi.org/10.5670/oceanog.2009.81

Joshi, S. R., Kukkadapu, R. K., Burdige, D. J., Bowden, M. E., Sparks, D. L., & Jaisi, D. P. (2015). Organic Matter Remineralization Predominates Phosphorus Cycling in the Mid-Bay Sediments in the Chesapeake Bay. *Environmental Science & Technology*, *49*(10), 5887–5896. https://doi.org/10.1021/es5059617

Kämpf, J., & Chapman, P. (2016). Upwelling Systems of the World. A Scientific Journey to the Most Productive Marine Ecosystems. Springer International Publishing. https://doi.org/10.1007/978-3-319-42524-5

Kim, J., Chapman, P., Rowe, G., & DiMarco, S. F. (2019). Implications of different nitrogen input sources for potential production and carbon flux estimates in the coastal Gulf of Mexico (GOM) and Korean coastal waters. Numerical Models/Biological Processes/Shelf-sea depth/Shelf Seas. https://doi.org/10.5194/os-2019-46

Kim, J., Chapman, P., Rowe, G., & DiMarco, S. F. (2020). Categorizing zonal productivity on the continental shelf with nutrient-salinity ratios. *Journal of Marine Systems*, *206*, 103336. https://doi.org/10.1016/j.jmarsys.2020.103336

Kolic, P. E., Roy, E. D., White, J. R., & Cook, R. L. (2014). Spectroscopic measurements of estuarine dissolved organic matter dynamics during a large-scale Mississippi River flood diversion. Science of the Total Environment, 10.

Kourou, K., Exarchos, T. P., Exarchos, K. P., Karamouzis, M. V., & Fotiadis, D. I. (2015). Machine learning applications in cancer prognosis and prediction. *Computational and Structural Biotechnology Journal*, *13*, 8–17. https://doi.org/10.1016/j.csbj.2014.11.005

Libbrecht, M. W., & Noble, W. S. (2015). Machine learning applications in genetics and genomics. *Nature Reviews Genetics*, *16*(6), 321–332. https://doi.org/10.1038/nrg3920

Liu, B., D'Sa, E. J., & Joshi, I. D. (2019). Floodwater impact on Galveston Bay phytoplankton taxonomy, pigment composition and photo-physiological state following Hurricane Harvey from field and ocean color (Sentinel-3A OLCI) observations. Biogeosciences, 16(9), 1975–2001. https://doi.org/10.5194/bg-16-1975-2019

Lloyd, Stuart P. (1982). Least squares quantization in PCM. IEEE Transactions on Information Theory. 28 (2): 129–137.

Mahmud, S., DiMarco, S. F., & Knap, A. (2020). *Glider Observations of the Loop Current and LC Eddy in the western Gulf of Mexico: Comparison of buoyancy mixing and relationships with oxygen variability*. Ocean Sciences Meeting 2020.

Moore, A. M., Arango, H. G., Di Lorenzo, E., Cornuelle, B. D., Miller, A. J., & Neilson, D. J. (2004). A comprehensive ocean prediction and analysis system based on the tangent linear and adjoint of a regional ocean model. *Ocean Modelling*, *7*(1–2), 227–258. https://doi.org/10.1016/j.ocemod.2003.11.001

Nakagawa, S., & Cuthill, I. C. (2007). Effect size, confidence interval and statistical significance: A practical guide for biologists. *Biological Reviews*, *82*(4), 591–605. https://doi.org/10.1111/j.1469-185X.2007.00027.x

National Oceanic and Atmospheric Administration, Goni, G., Todd, R., Jayne, S., Halliwell, G., Glenn, S., Dong, J., Curry, R., Domingues, R., Bringas, F., Centurioni, L., DiMarco, S., Miles, T., Morell, J., Pomales, L., Kim, H.-S., Robbins, P., Gawarkiewicz, G., Wilkin, J., … Sanabia, E. (2017). Autonomous and Lagrangian Ocean Observations for Atlantic Tropical Cyclone Studies and Forecasts. *Oceanography*, *30*(2), 92–103. https://doi.org/10.5670/oceanog.2017.227

Nowlin, W. D., Jochens, A. E., DiMarco, S. F., Reid, R. O., & Howard, M. K. (2005). Low-Frequency Circulation Over the Texas-Louisiana Continental Shelf. In W. Sturges & A. Lugo-Fernandez (Eds.), *Geophysical Monograph Series* (pp. 219–240). American Geophysical Union. https://doi.org/10.1029/161GM17

Nowlin, W. D., Jochens, A. E., Reid, R.O., DiMarco, S.F. (1998). Texas-Louisiana Shelf Circulation and Transport Processes Study: Synthesis Report Volume I: Technical Report. 486.

Oey, L.-Y., Ezer, T., Forristall, G., Cooper, C., DiMarco, S., & Fan, S. (2005). An exercise in forecasting loop current and eddy frontal positions in the Gulf of Mexico: FORECASTING LOOP CURRENT AND EDDY FRONTS. *Geophysical Research Letters*, *32*(12), n/a-n/a. https://doi.org/10.1029/2005GL023253

Osburn, C. L., Boyd, T. J., Montgomery, M. T., Bianchi, T. S., Coffin, R. B., & Paerl, H. W. (2016). Optical Proxies for Terrestrial Dissolved Organic Matter in Estuaries and Coastal Waters. Frontiers in Marine Science, 2. https://doi.org/10.3389/fmars.2015.00127

Para, J., Coble, P. G., Charrière, B., Tedetti, M., Fontana, C., & Sempéré, R. (2010). Fluorescence and absorption properties of chromophoric dissolved organic matter (CDOM) in coastal surface waters of the northwestern Mediterranean Sea, influence of the Rhône River. Biogeosciences, 7(12), 4083–4103. https://doi.org/10.5194/bg-7-4083-2010

Parlanti, E. (2000). Dissolved organic matter ¯uorescence spectroscopy as a tool to estimate biological activity in a coastal zone submitted to anthropogenic inputs. Organic Geochemistry, 17.

Pauer, J. J., DePetro, P. A., Anstead, A. M., & Lehrter, J. C. (2014). Application of a one-dimensional model to explore the drivers and lability of carbon in the northern Gulf of Mexico. Ecological Modelling, 294, 59–70. https://doi.org/10.1016/j.ecolmodel.2014.09.007

Potter, H., DiMarco, S. F., & Knap, A. H. (2019). Tropical Cyclone Heat Potential and the Rapid Intensification of Hurricane Harvey in the Texas Bight. *Journal of Geophysical Research: Oceans*, *124*(4), 2440–2451. https://doi.org/10.1029/2018JC014776

Quigg, A., Sylvan, J. B., Gustafson, A. B., Fisher, T. R., Oliver, R. L., Tozzi, S., & Ammerman, J. W. (2011). Going West: Nutrient Limitation of Primary Production in the Northern Gulf of Mexico and the Importance of the Atchafalaya River. *Aquatic Geochemistry*, *17*(4–5), 519–544. https://doi.org/10.1007/s10498-011-9134-3

R Core Team (2018) R: A language and environment for statistical computing. Vienna, Austria. https://www.R-project.org

Rabalais, N. N., Turner, R. E., Sen Gupta, B. K., Boesch, D. F., Chapman, P., & Murrell, M. C. (2007). Hypoxia in the northern Gulf of Mexico: Does the science support the Plan to Reduce, Mitigate, and Control Hypoxia? *Estuaries and Coasts*, *30*(5), 753–772. https://doi.org/10.1007/BF02841332

Ramey, E., DiMarco, S. F., Dreger, K., & Zimmerle, H. M. (2017). Performance of an Ocean Buoyancy Glider in a Coastal Region: Application to the Gulf of Mexico Hypoxic Zone. *Marine Technology Society Journal*, *51*(4), 41–51.

Reichstein, M., Camps-Valls, G., Stevens, B., Jung, M., Denzler, J., Carvalhais, N., & Prabhat. (2019). Deep learning and process understanding for data-driven Earth system science. *Nature*, *566*(7743), 195–204. https://doi.org/10.1038/s41586-019-0912-1

Rouet-Leduc, B., Hulbert, C., & Johnson, P. A. (2019). Continuous chatter of the Cascadia subduction zone revealed by machine learning. *Nature Geoscience*, *12*(1), 75–79. https://doi.org/10.1038/s41561-018-0274-6

Rousseeuw, P. J. (1987). Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *Journal of Computational and Applied Mathematics*, *20*, 53–65. https://doi.org/10.1016/0377-0427(87)90125-7

Rowe, G. T., & Chapman, P. (2002). Continental Shelf Hypoxia: Some Nagging Questions. *Gulf of Mexico Science*, *20*(2). https://doi.org/10.18785/goms.2002.08

Ruiz, S., Renault, L., Garau, B., & Tintoré, J. (2012). Underwater glider observations and modeling of an abrupt mixing event in the upper ocean: ABRUPT MIXING IN THE UPPER OCEAN. *Geophysical Research Letters*, *39*(1), n/a-n/a. https://doi.org/10.1029/2011GL050078

Salve, P. R., Lohkare, H., Gobre, T., Bodhe, G., Krupadam, R. J., Ramteke, D. S., & Wate, S. R. (2012). Characterization of Chromophoric Dissolved Organic Matter (CDOM) in Rainwater Using Fluorescence Spectrophotometry. Bulletin of Environmental Contamination and Toxicology, 88(2), 215–218. https://doi.org/10.1007/s00128-011-0424-7

Scharroo, R., Smith, W. H. F., & Lillibridge, J. L. (2005). Satellite altimetry and the intensification of Hurricane Katrina. *Eos, Transactions American Geophysical Union*, *86*(40), 366. https://doi.org/10.1029/2005EO400004

Shchepetkin, A. F., & McWilliams, J. C. (2005). The regional oceanic modeling system (ROMS): A split-explicit, free-surface, topography-following-coordinate oceanic model. *Ocean Modelling*, *9*(4), 347–404. https://doi.org/10.1016/j.ocemod.2004.08.002

Stedmon, C. A., & Markager, S. (2001). The optics of chromophoric dissolved organic matter (CDOM) in the Greenland Sea: An algorithm for differentiation between marine and terrestrially derived organic matter. Limnology and Oceanography, 46(8), 2087–2093. https://doi.org/10.4319/lo.2001.46.8.2087

Stedmon, Colin A., & Nelson, N. B. (2015). The Optical Properties of DOM in the Ocean. In Biogeochemistry of Marine Dissolved Organic Matter (pp. 481–508). Elsevier. https://doi.org/10.1016/B978-0-12-405940-5.00010-8

Stommel, H. (1989). The Slocum Mission. *Oceanography*, *2*(1), 22–25. https://doi.org/10.5670/oceanog.1989.26

Strickland, J.D.H., & Parsons, T.R. (1972). A practical handbook of seawater analysis. Bulletin 167, Journal of the Fisheries Research Board of Canada, Ottawa, Canada

Tehrani, N., D'Sa, E., Osburn, C., Bianchi, T., & Schaeffer, B. (2013). Chromophoric Dissolved Organic Matter and Dissolved Organic Carbon from Sea-Viewing Wide Field-of-View Sensor (SeaWiFS), Moderate Resolution Imaging Spectroradiometer (MODIS) and MERIS Sensors: Case Study for the Northern Gulf of Mexico. Remote Sensing, 5(3), 1439–1464. https://doi.org/10.3390/rs5031439

Testor, P., de Young, B., Rudnick, D. L., Glenn, S., Hayes, D., Lee, C. M., Pattiaratchi, C., Hill, K., Heslop, E., Turpin, V., Alenius, P., Barrera, C., Barth, J. A., Beaird, N., Bécu, G., Bosse, A., Bourrin, F., Brearley, J. A., Chao, Y., … Wilson, D. (2019). OceanGliders: A Component of the Integrated GOOS. *Frontiers in Marine Science*, *6*, 422. https://doi.org/10.3389/fmars.2019.00422

Vodacek, A., Blough, N. V., DeGrandpre, M. D., DeGrandpre, M. D., & Nelson, R. K. (1997). Seasonal variation of CDOM and DOC in the Middle Atlantic Bight: Terrestrial inputs and photooxidation. Limnology and Oceanography, 42(4), 674–686. https://doi.org/10.4319/lo.1997.42.4.0674

Vodacek, A., Hogel, F. E., Swift, R. N., Yungel, J. K., Peltzer, E. T., & Blough, N. V. (1995). The use of in situ and airborne fluorescence measurements to determine UV absorption coefficients and DOC concentrations in surface waters. Limnology and Oceanography, 40(2), 411–415. https://doi.org/10.4319/lo.1995.40.2.0411

Walker, N. D. (2005). Wind and Eddy-Related Shelf/Slope Circulation Processes and Coastal Upwelling in the Northwestern Gulf of Mexico. In *Circulation in the Gulf of Mexico: Observations and Models* (pp. 295–313). American Geophysical Union (AGU). https://doi.org/10.1029/161GM21

Wang, Z., DiMarco, S. F., & Polzin, K. (2018). Turbulence Observations in the Northern Gulf of Mexico from a Slocum Glider. *2018 OCEANS - MTS/IEEE Kobe Techno-Oceans (OTO)*, 1–6. https://doi.org/10.1109/OCEANSKOBE.2018.8559161

Yang, L., Zhang, J., Yang, G.P. (2021). Mixing behavior, biological and photolytic degradation of dissolved organic matter in the East China Sea and the Yellow Sea. Science of the Total Environment, 19.

Zeng, X., Li, Y., He, R., & Yin, Y. (2015). Clustering of Loop Current patterns based on the satellite-observed sea surface height and self-organizing map. Remote Sensing Letters, 6(1), 11–19. https://doi.org/10.1080/2150704X.2014.998347

Zhang, X., Hetland, R. D., Marta-Almeida, M., & DiMarco, S. F. (2012). A numerical investigation of the Mississippi and Atchafalaya freshwater transport, filling and flushing times on the Texas-Louisiana Shelf: FRESHWATER CHARACTERISTICS ON A SHELF. Journal of Geophysical Research: Oceans, 117(C11). https://doi.org/10.1029/2012JC008108

Zhang, W., Hetland, R. D., Ruiz, V., DiMarco, S. F., & Wu, H. (2020). Stratification duration and the formation of bottom hypoxia over the Texas-Louisiana shelf. *Estuarine, Coastal and Shelf Science*, *238*, 106711. https://doi.org/10.1016/j.ecss.2020.106711

Zhou, Y., Davidson, T. A., Yao, X., Zhang, Y., Jeppesen, E., de Souza, J. G., Wu, H., Shi, K., & Qin, B. (2018). How autochthonous dissolved organic matter responds to eutrophication and climate warming: Evidence from a cross-continental data analysis and experiments. Earth-Science Reviews, 185, 928–937. https://doi.org/10.1016/j.earscirev.2018.08.013

APPENDIX A

TABLES

Table A1: Mean BIX within each Sub-region (rows) of each cruise (columns), grouped by overall, sub-pycnocline, and above-pycnocline Sub-regions. Green and red cells indicate the largest and smallest mean BIX value, respectively, in that grouping. Standard deviation (σ) in parentheses.

Mean Overall BIX (σ)							
	MS03	**MS04**	**MS05**	**MS06**	**MS07**	**MS08**	**Region**
Sub-Region A	1.40 (0.71)	1.14 (0.32)	1.12 (0.07)	1.11 (0.05)	1.29 (0.85)	ND	**East**
Sub-Region B	1.17 (0.26)	**1.06 (0.13)**	1.16 (0.03)	1.09 (0.04)	1.38 (0.88)	1.29 (0.67)	
Sub-Region C	ND	1.16 (0.07)	1.14 (0.04)	1.13 (0.02)	1.15 (0.13)	1.07 (0.08)	**Middle**
Sub-Region D	ND	1.22 (0.06)	1.11 (0.02)	ND	1.09 (0.14)	1.15 (0.06)	**West**
Sub-Region E	ND	**1.41 (0.04)**	1.16 (0.01)	1.27 (0.01)	1.15 (0.06)	1.25 (0.01)	

Mean Sub-pycnocline BIX (σ)							
	MS03	**MS04**	**MS05**	**MS06**	**MS07**	**MS08**	**Region**
Sub-Region A	1.05 (0.06)	1.15 (0.03)	1.15 (0.07)	1.13 (0.05)	1.12 (0.07)	ND	**East**
Sub-Region B	1.08 (0.08)	1.08 (0.06)	1.17 (0.04)	1.08 (0.07)	1.09 (0.09)	**0.97 (0.03)**	
Sub-Region C	ND	1.09 (0.10)	1.17 (0.02)	1.11 (0.04)	1.05 (0.07)	0.99 (0.05)	**Middle**
Sub-Region D	ND	1.09 (0.06)	1.14 (0.02)	ND	1.00 (0.05)	0.98 (0.03)	**West**
Sub-Region E	ND	**1.39 (0.03)**	1.15 (0.01)	1.27 (0.01)	1.09 (0.06)	No Pycnocline	

Mean Above-pycnocline BIX (σ)							
	MS03	**MS04**	**MS05**	**MS06**	**MS07**	**MS08**	**Region**
Sub-Region A	1.59 (0.82)	1.13 (0.39)	1.08 (0.04)	1.09 (0.04)	1.59 (1.34)	ND	**East**
Sub-Region B	1.20 (0.29)	**1.05 (0.14)**	1.16 (0.03)	1.09 (0.03)	**1.69 (1.19)**	1.57 (0.81)	
Sub-Region C	ND	1.17 (0.05)	1.13 (0.03)	1.13 (0.01)	1.21 (0.11)	1.12 (0.05)	**Middle**
Sub-Region D	ND	1.23 (0.05)	1.11 (0.02)	ND	1.12 (0.15)	1.16 (0.05)	**West**
Sub-Region E	ND	1.45 (0.02)	1.16 (0.01)	1.25 (0.01)	1.17 (0.05)	No Pycnocline	

Table A2: Mean CMR within each Sub-region (rows) of each cruise (columns), grouped by overall, sub-pycnocline, and above-pycnocline Sub-regions. Green and red cells indicate the largest and smallest mean CMR value, respectively, in that grouping. Standard deviation (σ) in parentheses.

Mean Overall CMR (σ)							
	MS03	**MS04**	**MS05**	**MS06**	**MS07**	**MS08**	**Region**
Sub-Region A	1.65 (1.19)	1.31 (0.62)	0.80 (0.21)	0.99 (0.35)	1.40 (2.48)	ND	**East**
Sub-Region B	1.11 (0.73)	1.29 (0.52)	0.75 (0.05)	1.04 (0.29)	1.83 (2.18)	**2.30 (1.92)**	
Sub-Region C	ND	1.06 (0.23)	0.78 (0.09)	0.86 (0.16)	1.56 (1.15)	1.17 (0.40)	**Middle**
Sub-Region D	ND	0.96 (0.18)	0.89 (0.11)	ND	1.38 (0.78)	1.15 (0.19)	**West**
Sub-Region E	ND	0.76 (0.10)	0.75 (0.02)	**0.69 (0.08)**	1.26 (0.34)	0.92 (0.13)	

Mean Sub-pycnocline CMR (σ)							
	MS03	**MS04**	**MS05**	**MS06**	**MS07**	**MS08**	**Region**
Sub-Region A	0.53 (0.19)	0.73 (0.09)	0.68 (0.09)	0.69 (0.07)	0.51 (0.23)	ND	**East**
Sub-Region B	0.71 (0.05)	0.72 (0.08)	0.70 (0.04)	0.79 (0.11)	0.60 (0.16)	1.07 (0.15)	
Sub-Region C	ND	0.87 (0.08)	0.69 (0.04)	0.92 (0.08)	0.55 (0.23)	0.75 (0.09)	**Middle**
Sub-Region D	ND	**1.11 (0.19)**	0.69 (0.08)	ND	**0.49 (0.17)**	1.02 (0.13)	**West**
Sub-Region E	ND	0.79 (0.10)	0.78 (0.01)	0.68 (0.07)	0.86 (0.08)	No Pycnocline	

Mean Above-pycnocline CMR (σ)							
	MS03	**MS04**	**MS05**	**MS06**	**MS07**	**MS08**	**Region**
Sub-Region A	2.26 (1.05)	1.59 (0.58)	0.99 (0.16)	1.15 (0.33)	2.94 (3.59)	ND	**East**
Sub-Region B	1.27 (0.81)	1.49 (0.47)	0.77 (0.04)	1.11 (0.29)	3.15 (2.55)	**3.34 (2.10)**	
Sub-Region C	ND	1.12 (0.23)	0.84 (0.08)	0.85 (0.17)	2.27 (0.99)	1.45 (0.26)	**Middle**
Sub-Region D	ND	0.95 (0.18)	0.91 (0.10)	ND	1.73 (0.64)	1.16 (0.19)	**West**
Sub-Region E	ND	**0.69 (0.06)**	0.75 (0.02)	0.93 (0.08)	1.38 (0.29)	No Pycnocline	

APPENDIX B

FIGURES

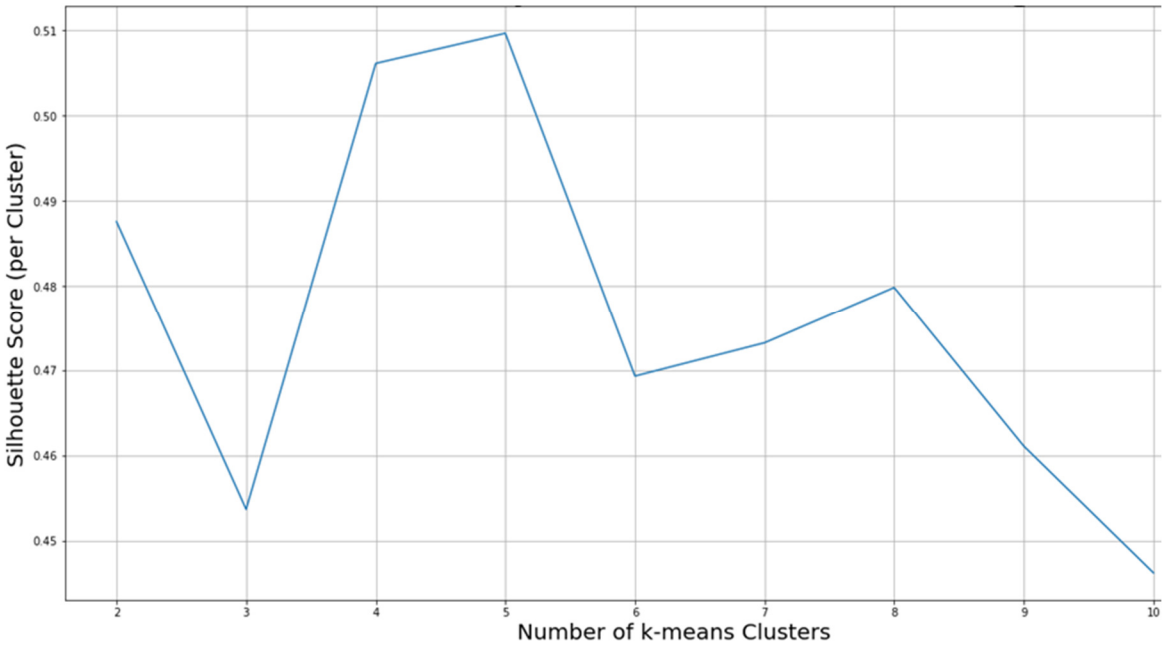

Figure B1: Results of silhouette scoring of the M13 glider dataset to determine optimal number of k-means clusters. Note that the peak is at k = 5 clusters.

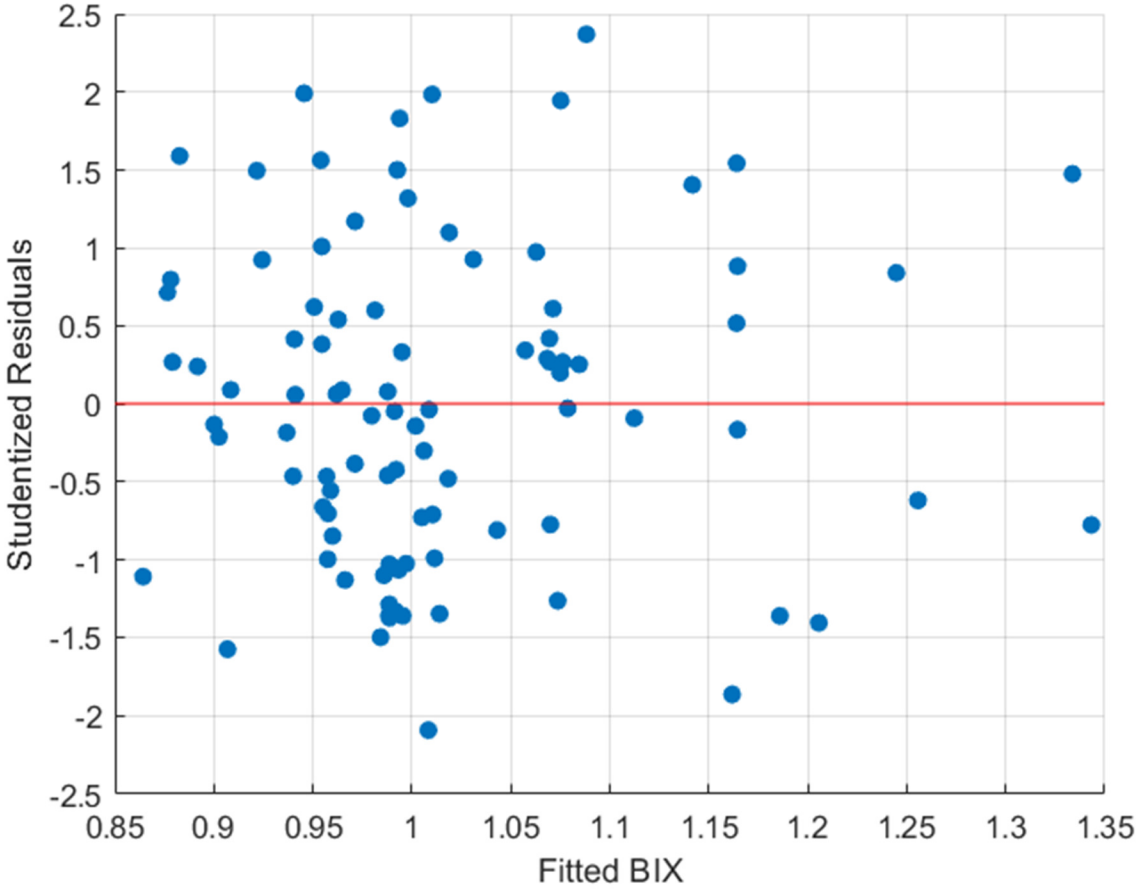

Figure B2: Studentized residuals for new Biological Index (BIX) algorithm developed in Chapter 3.2.3.

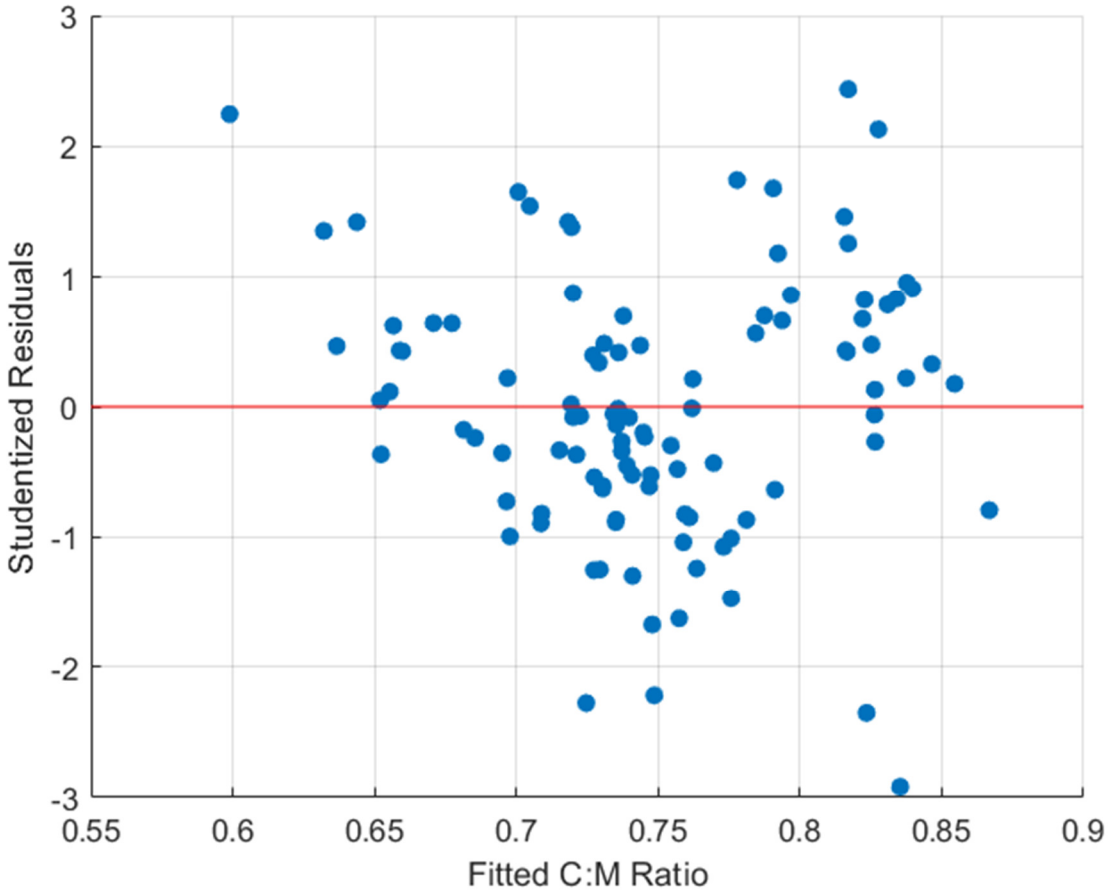

Figure B3: Studentized residuals for new Peak C:Peak M Ratio (CMR) algorithm developed in Chapter 3.2.3.

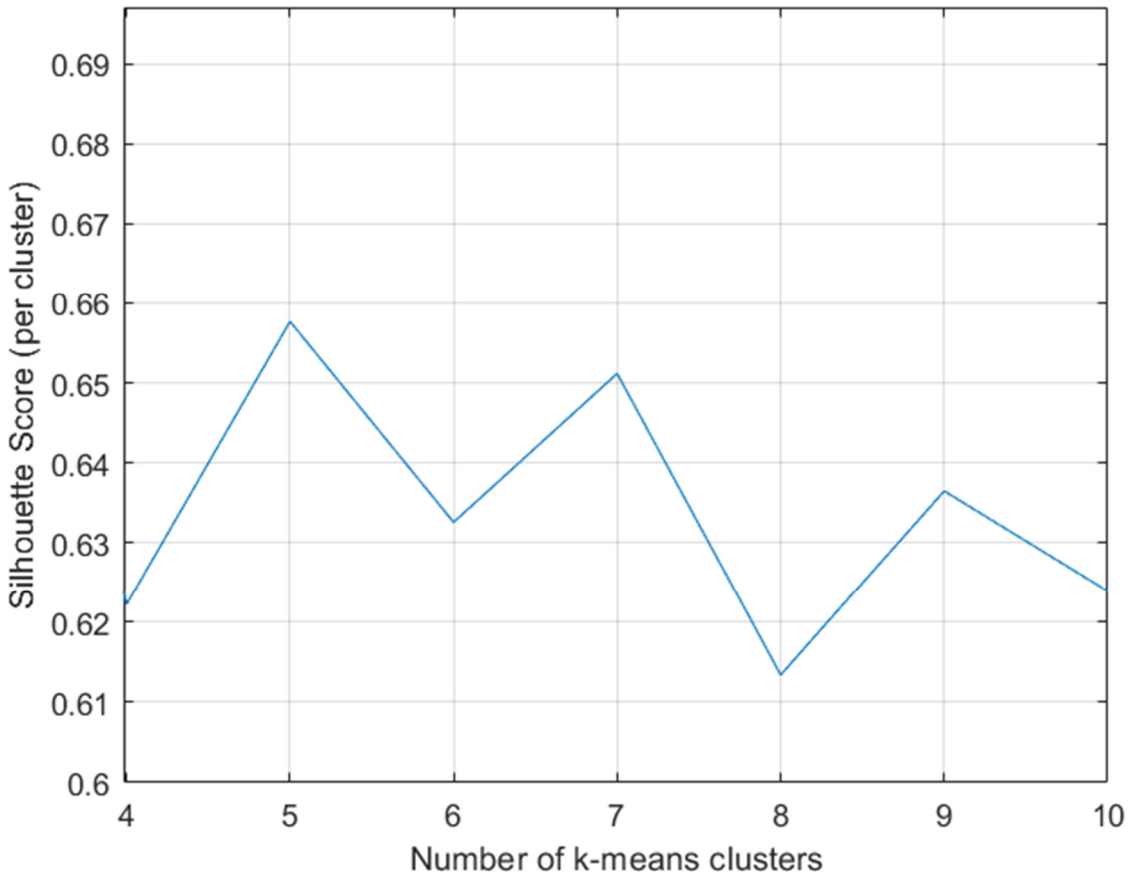

Figure B4: Results of silhouette scoring of MS03L12 to determine optimal number of k-means clusters. Note that the peak is at k = 5 clusters.

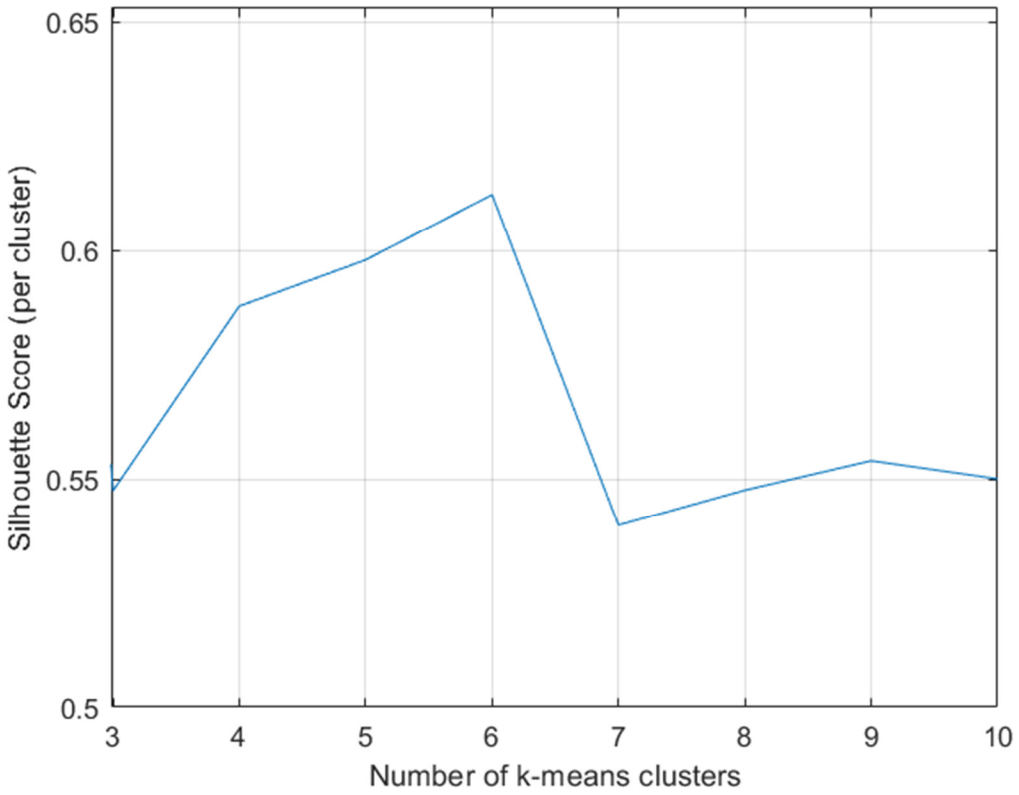

Figure B5: Results of silhouette scoring of glider M7 dataset to determine optimal number of k-means clusters. Note that the peak is at k = 6 clusters.

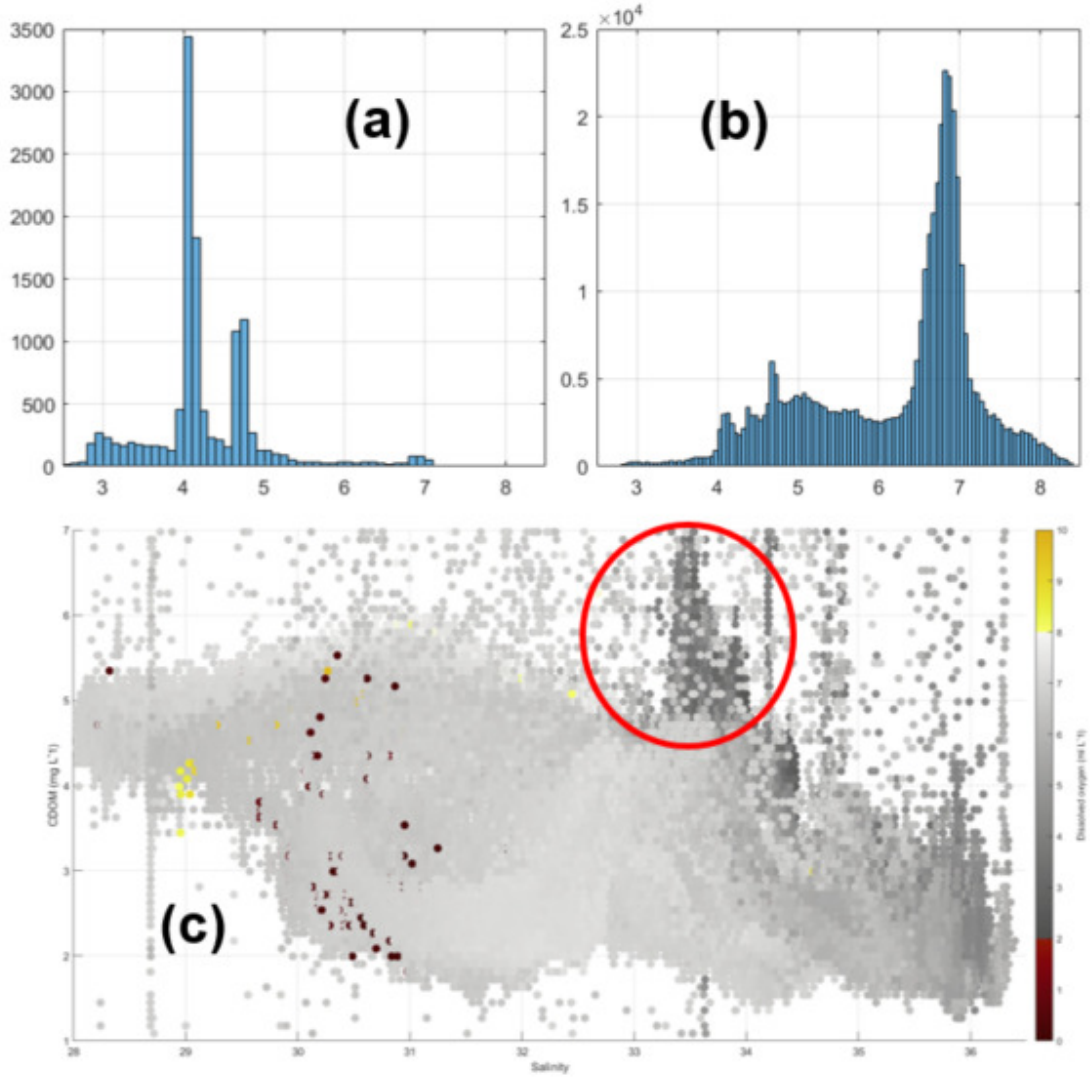

Figure B6: (a) Dissolved oxygen concentration (mean = 4.25 ml L^{-1}) within the CDOM pulse (c, red circle) encountered by glider M14; (b) dissolved oxygen concentration (mean = 6.19 ml L^{-1}) within the entire M14 dataset; (c) M14 salinity vs. CDOM, colored by dissolved oxygen concentration, red circle indicating the CDOM pulse mentioned in Section 2.4.1.