

Mesoscale Eddy Dynamics and Scale in the Red Sea

Dissertation by

Michael F. Campbell, Jr.

In Partial Fulfillment of the Requirements

For the Degree of

Doctor of Philosophy

King Abdullah University of Science and Technology,

Thuwal, Kingdom of Saudi Arabia

December 2020

EXAMINATION COMMITTEE PAGE

The dissertation of Michael F. Campbell, Jr. is approved by the examination committee.

Committee Chairperson: Prof. Burton Jones

Committee Co-Chair: Prof. Joanne Ellis

Committee Members: Prof. Michael Berumen, Prof. Ibrahim Hoteit, Prof. Luc Rainville

ABSTRACT

Mesoscale Eddy Dynamics and Scale in the Red Sea

Michael F. Campbell, Jr.

Recent efforts in understanding the variability inherent in coastal and offshore waters have highlighted the need for higher resolution sampling at finer spatial and temporal resolutions. Gliders are increasingly used in these transitional waters due to their ability to provide these finer resolution data sets in areas where satellite coverage may be poor, ship-based surveys may be impractical, and important processes may occur below the surface. Since no single instrument platform provides coverage across all needed spatial and temporal scales, Ocean Observation systems are using multiple types of instrument platforms for data collection. However, this results in increasingly large volumes of data that need to be processed and analyzed and there is no current “best practice” methodology for combining these instrument platforms. In this study, high resolution glider data, High Frequency Radar (HFR), and satellite-derived data products (MERRA_2 and ARMOR3D NRT Eddy Tracking) were used to quantify: 1) dominant scales of variability of the central Red Sea, 2) determine the minimum sampling frequency required to adequately characterize the central Red Sea, 3) discriminate whether the fine scale persistency of oceanographic variables determined from the glider data are comparable to those identified using HFR and satellite-derived data products, and 4) determine additional descriptive information regarding eddy occurrence and strength in the Red Sea from 2018-2019. Both Integral Time Scale and Characteristic Length Scale analysis show that the persistence time frame from glider data for

temperature, salinity, chlorophyll- α , and dissolved oxygen is 2-4 weeks and that these temporal scales match for HFR and MERRA_2 data, matching a similar description of a "weather-band" level of temporal variability. Additionally, the description of eddy activity in the Red Sea also supports this 2-4-week time frame, with the average duration of cyclonic and anticyclonic eddies from 2018-2019 being 22 and 27 days, respectively. Adoption of scale-based methods across multiple ocean observation areas can help define "best practice" methodologies for combining glider, HFR, and satellite-derived data to better understand the naturally occurring variability and improve resource allocation.

ACKNOWLEDGEMENTS

I would like to say thank you to my supervisor Burt Jones for providing me the opportunity to apply the skills that I learned from my master's degree in Geography to Marine Science in the Red Sea. I am thankful for the lessons that I have learned from him on how to be a better scientist and leader, and especially for his patience while I have improved my writing.

I am also thankful for my co-supervisor Jo Ellis, and her timely arrival at KAUST. I am thankful for her direction and support with developing the statistical methods described in this dissertation and for pointing out an analysis method to answer the questions that I had.

I would like to thank the other committee members, Mike Berumen, Ibrahim Hoteit, and Luc Rainville for their comments and support during my time at KAUST. I have learned both from their discussions as well as from their comments to this document.

I am thankful for the members from Integrated Ocean Processes for their ongoing support during all stages of my work here. I am thankful for their willingness to listen to my ideas and for their help in developing my communication skills.

Finally, I would like to thank my wife Colleen and my son Topher for all their love and patience while finishing my degrees. I am thankful for Topher's enthusiasm and encouragement. I am especially thankful for Colleen being the first person I discuss new ideas with, and her help both in keeping me on track and grounded.

TABLE OF CONTENTS

EXAMINATION COMMITTEE PAGE	2
ABSTRACT.....	3
ACKNOWLEDGEMENTS	5
TABLE OF CONTENTS.....	6
LIST OF ABBREVIATIONS.....	8
LIST OF SYMBOLS	9
LIST OF ILLUSTRATIONS.....	10
LIST OF TABLES	12
Chapter 1: Introduction	13
1.1 The Red Sea	13
1.2: Characteristic Length Scale and variability	16
1.3: Research Goals	22
Chapter 2: Characterizing temporal scales of temperature, salinity, and chlorophyll variability using high-resolution glider data in the Red Sea.....	24
2.1: Introduction.....	24
2.1.1 Study Area	30
2.2: Methods	34
2.2.1: Measurements used in this study	34
2.2.2: Time series analysis used in this study	41
2.3: Results.....	45
2.4: Discussion.....	54
2.4.1: Temporal Scales.....	55
2.4.2: Comparisons between Glider, HFR, and MERRA_2	56
2.4.3: Conclusion	57
Chapter 3: Localized influence of two counter-rotating eddies in the north central Red Sea as seen by two parallel AUVs	60
3.1: Introduction.....	60
3.1.1: Context of eddies in the Red Sea	61
3.1.2: Study Area	68
3.1.3: Single glider line results	70
3.2: Methods	71
3.2.1: Glider deployment and calibration	71
3.2.2: HFR calibration.....	76
3.2.3: Short term persistence through Error X analysis	77

3.3: Results:	83
3.3.1: Water column time series plots (glider)	83
3.3.2: Geostrophic Velocities from glider relative to 400 m	84
3.3.3: Comparison between section glider GSV and ARMOR3D GSV at north and south line	88
3.4: Discussion	92
3.4.1: Meaning for the Red Sea	92
3.4.2: Conclusions	94
Chapter 4: Sea surface temperature and eddy center occurrence cluster analysis in the Red Sea	96
4.1: Introduction	96
4.1.1: Physical effects of Red Sea eddies	98
4.2: Methods:	101
4.2.1: AVISO+ NRT 3.0	101
4.2.2: Spatial processing and statistics	103
4.2.3: Permutation Distribution Clustering of SST	106
4.3: Results	107
4.3.1: Description of Principal Components	108
4.3.2: Getis-Ord Gi* clustering	111
4.3.3: Machine Learning clustering	113
4.4: Discussion	117
Chapter 5: Summary	120
5.1: Future Work	125
REFERENCES:	126
Appendix A: Autocorrelation and Error X figures, Chapter 2	133
Appendix B: Error X figures, Chapter 3	143

LIST OF ABBREVIATIONS

ADT	Absolute Dynamic Topography
AUV	Autonomous Underwater Vehicle
AVISO+	Archiving, Validation, and Interpretation of Satellite Oceanographic data
CI	Confidence Interval
CLS	Characteristic Length Scale
CTD	Conductivity, Temperature, Depth
GEBCO	General Bathymetric Chart of the Ocean
HFR	High Frequency Radar
ITS	Integral Time Scale
KAUST	King Abdullah University of Science and Technology
kNN	k Nearest Neighbor
MERRA_2	Modern Era Retrospective Analysis for Research and Applications-2
NASA	National Aeronautics and Space Administration
NOAA	National Oceanic and Atmospheric Administration
NRT	Near Real Time
OC-CCI	Ocean Color Climate Change Initiative
OI	Optimum Interpolation
PC	Principle Component
PCA	Principle Component Analysis
PDC	Permutation Distribution Clustering
SLA	Sea Level Anomaly
SST	Sea Surface Temperature
SSTA	Sea Surface Temperature Anomaly
TEOS-10	Thermodynamic Equation of Seawater

LIST OF SYMBOLS

$\sigma\text{-}\theta$	Sigma Theta, normalized density of seawater, actual density – 1000 kg/m ³
τ	Tau, wind stress on the surface of seawater
u	east (+)/west (-) directional velocity vector
v	north (+)/south (-) directional velocity vector

LIST OF ILLUSTRATIONS

Figure 1.1: A map of Red Sea circulation.....	16
Figure 1.2: Stommel diagram	17
Figure 2.1: A map of the Red Sea.....	31
Figure 2.2: Temperature-salinity diagrams.....	33
Figure 2.3: Chlorophyll-density diagram.....	34
Figure 2.4: Glider transect	38
Figure 2.5: Glider transect	39
Figure 2.6: Glider time series data	47
Figure 2.7: Glider time series data.....	47
Figure 2.8: High Frequency Radar time series data.....	49
Figure 2.9: MERRA_2 time series data.....	49
Figure 2.10: MERRA_2 time series data.....	50
Figure 2.11: Glider ITS and CLS results	52
Figure 2.12: HFR and MERRA_2 ITS and CLS results.....	53
Figure 3.1: A map of Red Sea surface circulation	63
Figure 3.2: A map of Red Sea chlorophyll a	67
Figure 3.3: A map of the Red Sea.....	69
Figure 3.4: Glider transect	74
Figure 3.5: Glider transect	75
Figure 3.6: Glider time series data.....	78
Figure 3.7: Glider time series data.....	79
Figure 3.8: Glider time series data.....	80
Figure 3.9: HFR time series data	81
Figure 3.10: Glider CLS results	82
Figure 3.11: Glider geostrophic velocities	85
Figure 3.12: Glider average geostrophic velocities	87
Figure 3.13: Glider and ARMOR3D geostrophic velocities	89
Figure 3.14: Glider and ARMOR3D geostrophic velocities	90
Figure 3.15: A map of glider and ARMOR3D geostrophic velocities	90
Figure 3.16: A map of glider and ARMOR3D geostrophic velocities	91
Figure 3.17: A map of glider and ARMOR3D geostrophic velocities	91
Figure 4.1: A map of the Red Sea.....	97
Figure 4.2: A map of Red Sea eddies	102
Figure 4.3: A map of Red Sea bathymetry	104
Figure 4.4: A map of Red Sea eddy occurrence	105
Figure 4.5: Principle Component Analysis results	109
Figure 4.6: Hot Spot Analysis results	112
Figure 4.7: Sea Surface Temperature Anomaly clustering results	114
Figure 4.8: A map of HSA and SSTA clustering	116
Figure A.1: Autocorrelation and Characteristic Length Scale.....	133
Figure A.2: Autocorrelation and Characteristic Length Scale.....	134
Figure A.3: Autocorrelation and Characteristic Length Scale.....	135
Figure A.4: Autocorrelation and Characteristic Length Scale.....	136
Figure A.5: Autocorrelation and Characteristic Length Scale.....	137

Figure A.6: Autocorrelation and Characteristic Length Scale	138
Figure A.7: Autocorrelation and Characteristic Length Scale	139
Figure A.8: Autocorrelation and Characteristic Length Scale	140
Figure A.9: Autocorrelation and Characteristic Length Scale	141
Figure A.10: Autocorrelation and Characteristic Length Scale	142
Figure B.1: Characteristic Length Scale	143
Figure B.2: Characteristic Length Scale	144
Figure B.3: Characteristic Length Scale	145
Figure B.4: Characteristic Length Scale	146
Figure B.5: Characteristic Length Scale	147
Figure B.6: Characteristic Length Scale	148
Figure B.7: Characteristic Length Scale	149
Figure B.8: Characteristic Length Scale	150

LIST OF TABLES

Table 2.1: Summary of the glider equipment used in this study	36
Table 3.1: Glider deployment dates	69
Table 3.2: Summary of the equipment used in this study	76
Table 3.3: V velocity results	92
Table 4.1: Result of PCA for bathymetry and eddy count information	110
Table 4.2: A table of summarized characteristics of the SSTA subregions.....	115

Chapter 1: Introduction

1.1 The Red Sea

The Red Sea is a semi-enclosed basin lying between the Asian and African continents.

The basin is about 2200 km long, has a maximum depth of about 2900 m, and an average width of 280 km (Raitsos et al., 2013; Triantafyllou et al., 2014). The only major input of water in the Red Sea is through the Bab el Mandab (160 m deep) at the southern end of the Red Sea (Murray and Johns, 1997). Surface water from the Gulf of Aden enters the Red Sea at a temperature of 25-30° C and salinity of 36-36.5 while intermediate water from the Gulf of Aden enters the basin at temperature of 16.5° C and salinity at 37.5 (Sofianos and Johns, 2015). By the time the waters re-enter the Gulf of Aden salinity reaches above 40 (Sofianos and Johns, 2007; Yao et al., 2014a; Yao et al., 2014b). The Red Sea is very warm, with Red Sea Deep Water reaching 21.4-21.5° C and salinity ~ 40.6. The Red Sea is also shallow, with approximately 41 percent of the total surface covering depths less than 100 m with extensive reef systems on both sides of the basin despite the warm temperatures (Churchill et al., 2014a; Gerges, 2002; Rasul and Stewart, 2015; Silverman et al., 2007; Triantafyllou et al., 2014).

The overall circulation in the Red Sea is based on a reverse estuarine system, with salinity increasing steadily with distance from Bab el Mandab (Berumen et al., 2019b; Sofianos and Johns, 2003; Sofianos and Johns, 2007). Two main modes of interchange between the Red Sea and the Gulf of Aden exist, a two-layer system which occurs during the winter months (October-April) and dominates the overall annual circulation regime for the region, and a three-layer system that occurs during the summer months (June-

August) (Sofianos and Johns, 2015). Typically, the months of May and September are transitional months that have lighter winds (Langodan et al., 2014; Smeed, 1998). The timing is linked to the interchange between the Red Sea and the Gulf of Aden and the change in direction of the monsoon winds blowing through Bab el Mandab. The two-layer winter-time circulation occurs when the monsoon winds blow into the Red Sea from the Gulf of Aden, which drive surface flow into the southern Red Sea (Langodan et al., 2014; Sofianos and Johns, 2003; Sofianos and Johns, 2007; Yao et al., 2014a; Yao et al., 2014b). As the northeast monsoon winds weaken and reverse to the southwest monsoon, the interchange converts to a three-layer system in which surface water from the Red Sea flows into the Gulf of Aden, Gulf of Aden Intermediate Water flows underneath the surface water into the basin, and Red Sea Outflow Water flows underneath that into the Gulf of Aden (Sofianos and Johns, 2015; Sofianos and Johns, 2003; Yao et al., 2014a). Red Sea Outflow Water is an important component of the seawater throughout the region and can be found as far south as Mozambique and as far east as the Bay of Bengal (Beal et al., 2000; Jain et al., 2017).

Recent research into circulation within the Red Sea tends to focus on four types of energy transfer: surface buoyancy loss due to evaporation, wind-driven circulation, boundary currents, and friction caused by the bathymetry of the Red Sea (Triantafyllou et al., 2014; Wahr et al., 2014; Yao et al., 2014a; Yao et al., 2014b; Zhai et al., 2015; Zhan et al., 2016). Each of these processes have been evaluated in models as contributors to energy transfer within the Red Sea, with each method recreating the same type of eddies observed in satellite imagery, including Sea-Level Anomaly (SLA) data sets. These papers identify surface buoyancy loss and winds as the primary methods of energy

transfer (Triantafyllou et al., 2014; Wahr et al., 2014; Yao et al., 2014a; Yao et al., 2014b; Zhan et al., 2016).

Eddies are considered an important factor to the short-term circulation patterns within the Red Sea, and each individual eddy typically lasts from two to six weeks (Karimova and Gade, 2014; Zhan et al., 2014). When active, eddies are a dominant mode of kinetic energy transfer, with kinetic energy transfer being an order of magnitude higher than other modes of kinetic energy transfer (Zhan et al., 2016). However, the recurrent eddies have a small temporal extent, raising the question of what other mechanisms account for energy transfer over longer time frames.

When multi-year averages of surface circulation are examined, the short-term eddies are no longer seen, but instead a clear boundary current system is in effect (Figure 1.1) (Sofianos and Johns, 2003; Yao et al., 2014b; Zhai et al., 2015). This long-term boundary system is also supported by recent observations from gliders and shore-based HF-radar systems (<https://portus.kaust.edu.sa/portus/>), (Zarokanellos et al., 2017a). This interchange between an eddy dominated system and a boundary current system plays an important role in determining the length of a study period needed to answer questions related to energy dynamics in the Red Sea, with shorter time frames more focused on mesoscale eddy events. One of the important factors in research design is to understand the relevant scale of dynamic processes.

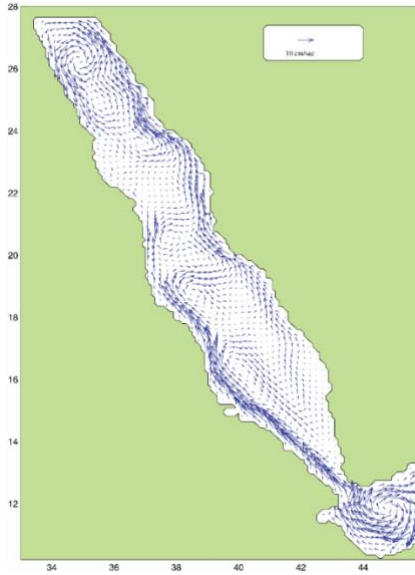


Figure 1.1: Nine-year composite map of sea surface circulation resulting from thermo-haline fields only in the Red Sea. This figure is copied from Sofianos and Johns (2003) and is a result from the Miami Isopycnic Coordinate Ocean Model simulation experiment discussed in their paper.

1.2: Characteristic Length Scale and variability

Scale is not a new concept within marine science. Initial discussions about the role of scale within ecosystems were introduced with the Stommel sea level variability plot, (Figure 1.2) and included discussion of other parameters, including temperature, tides, and current velocities (Stommel, 1963). Many different types of plots have been based on this initial discussion, covering ideas from effective sampling techniques to instrument limitations and other topics (Hedley et al., 2016; Phinn et al., 2010; Turner et al., 2001; Vance, 2007; Wu, 1999). The many variations of these plots show a general positive trend with longer time scale events extending over greater spatial distances. The Stommel diagram also indicates how sampling needs to match the spatio-temporal

extents of the processes involved, as the variation can shift drastically between different spatio-temporal extents.

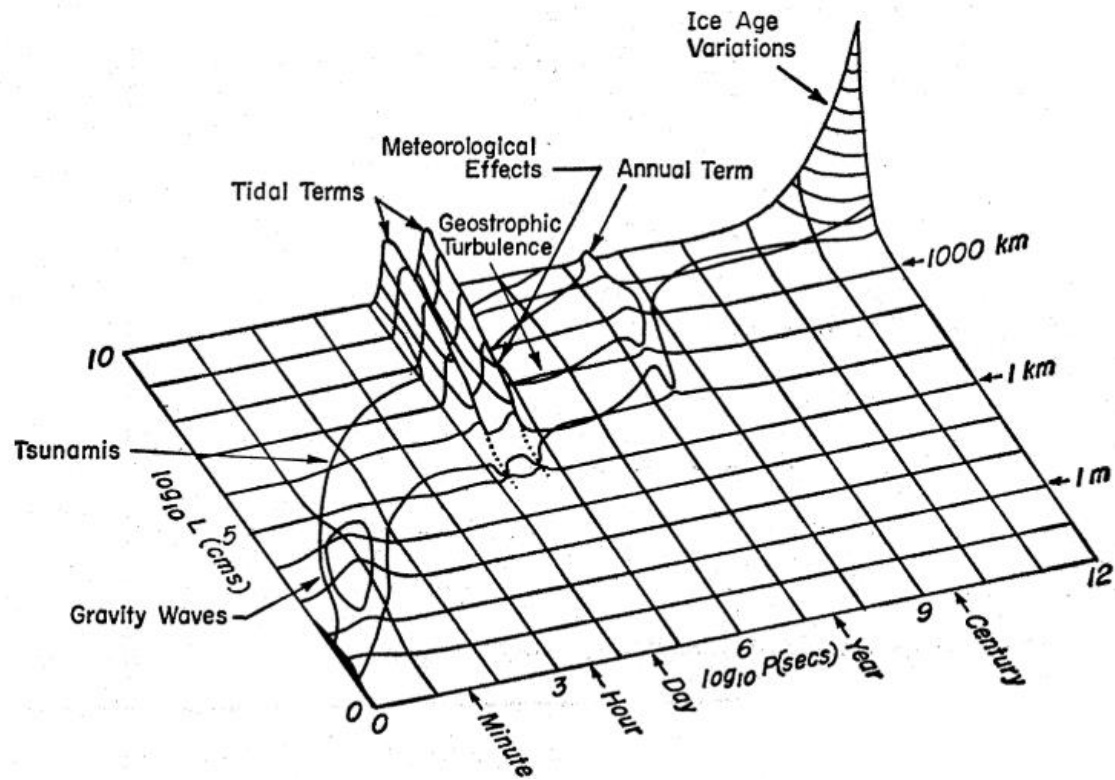


Figure 1.2: Stommel Diagram of sea level variability over time and space. Copied from Stommel (1963).

This research utilizes a variety of measurement techniques, focusing on the spatio-temporal constraints applicable to each method, and how these constraints affect the research and analysis of the Red Sea. These results aim to establish that the data needed for analysis matches up with the logistical limitations of the data collection methods (Ellis and Schneider, 2008; Hewitt et al., 2007; Holland et al., 2004). When data can only be collected at fixed intervals, for example on a weekly basis, then processes that occur at a higher frequency cannot be accurately resolved. For time series data, this

limitation is known as the Nyquist frequency, in which time series data can only be analyzed accurately when the observations frequency is at least twice the desired analysis frequency due to the typical periodic nature of time series data sets (Thomson and Emery, 2014). In the case of data collected weekly, the Nyquist frequency enabling correct interpretation would be every two weeks or greater. Spatio-temporal data of a periodic nature has a similar trend. Observations must be made at a sufficiently high resolution to fully resolve features, such as measuring both the crest and trough of a wave. However, the data has an additional limitation due to the concept of autocorrelation, in which observations that are made at closer distances tend to be similar to each other, which can create a non-standard distribution (Getis and Ord, 1992; Mitchell, 2005; Quattrochi and Goodchild, 1997). The most efficient data collection occurs when collected at resolutions that either avoid or explain spatio-temporal autocorrelation while still sufficiently detailed to fully describe any feature of interest.

One of the goals of initial planning of an ocean observation system-based study is to determine the scales at which the variables in the system are either homogenous or heterogenous, where homogeneity indicates the sameness or persistency of the system and heterogeneity describes the variability of the system (Rahbek, 2005; Thompson and McGarigal, 2002; Turner et al., 2001). The degree of either homogeneity or heterogeneity affects the variety of statistical analyses that can be performed to describe and summarize the data. The typical approach to improve efficient utilization of time and money is to establish a spatio-temporal sampling methodology that ensures collected samples will maximize heterogeneity (Hedley et al., 2016; Hewitt et al., 2015; Hewitt et al., 2007). Heterogeneity has two different defined components, the measured scale of

heterogeneity and the functional scale of heterogeneity. The measured scale of heterogeneity is the easier to define of the two components, as it is simply the “grain size” of the collected data (Turner et al., 2001). Grain is the resolution of the data or the spatio-temporal region represented by each data unit. Any data collected at a finer scale than the grain is considered homogenous within that unit and can be considered as a part of the natural bounds of variability. The functional scale of heterogeneity is the spatio-temporal range that is important and affects changes across the entire system and has been described as the Characteristic Length Scale (CLS), which will be used throughout this research (Keeling et al., 1997; Pascual and Levin, 1999; Thompson and McGarigal, 2002; Turner et al., 2001; Ward et al., 2018). CLS can be calculated for individual variables within a system with sufficient data (Crowder and Norse, 2008; Habeeb et al., 2005; Holland et al., 2004; Hurlbert and White, 2005). CLS for an entire system rather than for a few variables is typically determined by examining the interactions and importance of each of the variables as a part of the whole.

The relationship between grain and CLS has three main aspects. Data collected below the range of the grain can be considered a threshold (homogenous) within the examined system, indicating a minimum aggregation of data needed within the system. An example of this is seen in the working definition of the mixed layer in oceanography. In the mixed layer, the density changes only slightly with depth until the pycnocline is reached indicating a high degree of similarity (Zervakis et al., 2016). The mixed layer depth is important because data recorded within the mixed layer is considered homogenous, and presents the threshold values of parameters necessary for

phytoplankton production, which in turn is the basis for several other aspects of the mixed layer habitat (Brainerd and Gregg, 1995).

The second case is examining data at a scale between the grain and the CLS, where the grain size has an adequate resolution to measure the variability of a natural system, the grain needs to match the same spatial and temporal scales being studied, as is the case when using Synthetic Aperture Radar to measure eddy size and rotations to determine variability in sub-mesoscale eddies, where the images had a spatial grain size of 150 m (Dreschler-Fischer et al., 2014; Karimova and Gade, 2014). These sub-mesoscale eddies occur at spatial scales typically less than 70 km and for time periods less than 2 weeks, which limits detection in current global eddy tracking data products. In these studies, the spatial and temporal scale of the collected datasets fall within the spatial and temporal scale of the eddies, which allows for the characterization of the events of interest.

The final case is when grain is above the CLS. In this final situation, the descriptive information collected from the grain establishes the available limits within the system, or the carrying capacity. Measurements to determine carrying capacity do not have to occur as frequently as other types of measurements and are typically used as a part of the initial descriptions of a new system. An example of this can be seen in the Red Sea through the series of numerical simulations that have been run and continued to be developed in the Red Sea. Initial descriptions of Red Sea circulation have included various details of the interactions of eddies within the basin, going from permanent eddies to reoccurring to semi-permanent as the grain of the numerical simulations has reached finer spatio-temporal resolutions (Patzert, 1974; Sofianos and Johns, 2003; Yao et al., 2014a; Yao et al., 2014b; Zarokanellos et al., 2017b; Zhan et al., 2014). Both threshold and carrying

capacity values do not have to be monitored as closely as variables that fall within the functional heterogeneity (Turner et al., 2001).

This study addresses the CLS examining mesoscale activity in the Red Sea, with a primary focus on the north central part of the basin. For marine systems, the CLS is defined as a volume of water in which measured variables are at a steady state, with less variability within the defined spatio-temporal unit of water than between adjacent, equivalent sized units. The series of topics discussed within this dissertation bring a better understanding of the CLS of the spatio-temporal variables including temperature, salinity, chlorophyll- α , dissolved oxygen, and eddy occurrence and strength that are currently being measured as a part of an oceanic observation system in the Red Sea. This project also establishes a framework to determine the impact that new technology will have on the ability to understand circulation within the Red Sea and develop methods to re-assess CLS as new data sources become available. As circulation is a major component of marine habitat health, these variables play an important role in understanding processes that affect the entire Red Sea ecosystem. A recent trend in economic development within the Red Sea as a part of Saudi Arabian nation policy associated with Vision 2030 also indicates the need for a deeper understanding of the overall marine system (Almahasheer and Duarte, 2020). As development and utilization of the Red Sea increases due to projects related to Vision 2030, the need to understand in greater detail the interactions that occur both in the localized areas of the Red Sea and the entire basin as a whole will increase.

1.3: Research Goals

The goal of this dissertation is to examine the effects of the introduction of data collected at finer scale to an overall understanding of the Red Sea to provide an initial framework for planning future studies in the Red Sea, with the goal of establishing an adaptive methodology for oceanographic observations, where sampling locations and frequencies can be adjusted while maintaining continuity with previously existing data collection. In addition, effort must be made to ensure that analysis methods for the different types of data are comparable, indicating the need for highly flexible and accurate statistical analysis. Recent efforts in the Red Sea have introduced two tools for examining spatio-temporal scale in the Red Sea. The first tool is buoyancy-propelled, autonomous underwater vehicles (gliders), which collect oceanographic data across the entire water column along user-defined transect lines. The second tool is surface current mapping through High Frequency Radar (HFR), which provides hourly information on the speed and direction of surface currents over an extended area on an hourly basis. The addition of these tools allows for the collection of data at sufficient detail to describe events occurring on a spatial scale of several kilometers and a temporal scale of one to five hours. One goal of this dissertation is to examine the feasibility of using a non-linear, non-uniform CLS calculation with glider and HFR data sets in the Red Sea. A second goal is to examine some of the potential processes involved with variability in sea surface temperature, eddy occurrence, and physical structure in the Red Sea. Oceanographic information will be analyzed at both mesoscale (70-300 km) and basin wide levels of scale. Chapters 2 and 3 are focused on the application of CLS to short-term data sets collected by gliders and HFR. Chapter 4 is focused on basin wide dynamics, examining

the link between sea surface temperature and eddies. The successful application of the methods can then be used for further analysis over longer time series and as a starting point for developing an adaptive observation system in the Red Sea.

Chapter 2: Characterizing temporal scales of temperature, salinity, and chlorophyll variability using high-resolution glider data in the Red Sea

2.1: Introduction

The importance of scale across both time and space was introduced by Henry Stommel for marine research in the early 1960s. His initial discussion focused on the logistical processes that were relevant to sampling the spatio-temporal variation of sea level. Of course, other physical variables such as temperature and current velocities are additionally associated with sea level fluctuations (Stommel, 1963). The central concept was to demonstrate how oceanic processes varied across scales and that sampling efforts had to be planned with consideration of this variation. Issues of scale have become increasingly important in both seascape ecology and ocean observation systems, two areas of research that aim to characterize the marine environment (Ellis and Schneider, 2008; Kavanaugh et al., 2016; Nickols et al., 2015). The Stommel diagram has been adapted many times since the initial discussion, providing the key paradigms for understanding the interaction between length and time within marine systems.

Derivatives of Stommel's initial diagram have focused on pelagic trophic levels (Dickey, 2003), predictive capabilities of marine observational platforms at various scales (Hedley et al., 2016), and have even extended to topics outside of marine science (Rose et al., 2017; Turner et al., 2001). One general pattern apparent in the derivatives of the original Stommel diagram is a positive trend of increasing time scale associated with increasing spatial scales. As new technologies for observation in the marine environment emerge, a range of processes can be measured and analyzed at increasingly finer spatio-temporal

resolutions. However, an increasing volume of data does not *a priori* result in a better understanding of the underlying marine system (Little et al., 2018). Marine systems are a complex interplay between physical processes and biological response, which can vary dramatically across multiple levels of scale (Hidalgo et al., 2016; Kavanaugh et al., 2016; Mazloff et al., 2018). With this increase in available data, it is important to define the appropriate spatio-temporal scales required to resolve the key physical and biological processes. Identifying and defining these discrete units allows for direct analysis of processes measured by different instruments at a uniform spatio-temporal scale (Davis et al., 2019; Hedley et al., 2016; Hewitt et al., 2015).

As the density of spatio-temporal data for an area increases, conclusions can be drawn about the important, naturally occurring scales that affect the study area, as can be seen in the CalCOFI project (McClatchie, 2016; Rudnick et al., 2017), the global Argo program (Holte et al., 2017), and BGC Argo program (Terzic et al., 2019). The CalCOFI project has been ongoing with regular observations since 1951, and is the longest running, geographically extensive ocean observation project. The data collected has been important for identifying and refining regionality and variability in the California Current system (McClatchie, 2016). The addition of autonomous vehicles to the project through the California Underwater Glider Network has increased the availability of fine-scale data and revealed new oceanographic features in the project area (McClatchie, 2016; Rudnick et al., 2017). The Argo program has been influential in understanding the features of the global ocean. As of 2017, over 1,250,000 Argo profiles were used to characterize the temperature, salinity, and mixed layer depth of all the oceans, providing a global climatology as well as monthly data (Holte et al., 2017). This data has provided

insight into circulation and large-scale oceanographic structures. The Argo Biogeochemical program has similar aims to the Argo program, but with instrumentation specifically designed for biogeochemical measurements. Until recently, optical properties have largely been excluded or simplified in biogeochemical models due to a lack of *in situ* sampling (Terzic et al., 2019). Additional data from biogeochemical floats in the Mediterranean Sea describe the vertical, spatial, and temporal variability of zonal gradients in the basin. These studies highlight the need for high-resolution sampling at fine spatial (meso- to sub-mesoscale) and temporal scales (intra-seasonal) to understand longer-term variability of oceanographic processes.

One of the important aspects of each of these projects has been the amount of information collected to differentiate between different regions of the study areas. While many different methods of classification can be used to determine regionality, most forms of classification tend to rely on defining both the absolute value and variance associated with each measurement location over time compared to spatially close sample locations. Typically, a difference metric is used to organize the data, with locations that are statistical dependent typically classified into the same region. Several different terms are used to characterize this naturally occurring organization of variables, including Integral Time Scale (ITS), Characteristic Length Scale (CLS), correlation length, scale of variability and others, each with a slightly different associated calculation (Habeeb et al., 2005; Keeling et al., 1997; Thomson and Emery, 2014; Ward et al., 2018). Each of these terms seek to determine the distance at which consecutively collected data points remain statistically dependent to each other, which for the purposes of this research will be defined as “persistency” and considered in the context of persistent monitoring of areas

of interest while maintaining as much of the variability in the system as possibly (Davis et al., 2002; Davis et al., 2019; Habeeb et al., 2005; Rudnick, 2016; Rudnick et al., 2017).

Defining persistency for a study area is an important step in understanding the logistics necessary to provide continuous monitoring. By better understanding the natural, dynamic processes in an area, monitoring efforts can be adjusted to provide comparable data quality optimizing the allocation of observing resources (Rudnick et al., 2017).

Uniform time series data tend to be sinusoidal and to properly characterize events sampling must occur at a frequency at least twice as often as the event. This sampling frequency is known the Nyquist frequency. As long as the Nyquist frequency for each defined station in the study area is maintained, monitoring resources can be adjusted without losing data quality (Thomson and Emery, 2014). In the past, accurate analysis into persistency has only occurred over large-scale studies, either covering a large spatial extent or an extensive time frame, but seldom both at the same time (Habeeb et al., 2005; Keeling et al., 1997; Pascual and Levin, 1999; Ward et al., 2018).

In recent years, autonomous platforms have greatly increased in capability and deployment duration and are able to provide continuous data sets for extended periods of time (Davis et al., 2019; Testor et al., 2019). Depending on the mission parameters, the onboard instrumentation, and the frequency of data collection; these platforms sustain observations for extended durations with minimal or infrequent human interaction necessary (Rudnick et al., 2004). These persistent monitoring missions have enabled resolution of much larger spatial and temporal scales than previously available, which in turn permits understanding kilometer length scales of variability within the study area (Rudnick et al., 2017). While many types of autonomous platforms currently exist for

oceanic research, this discussion only focuses on observations from passively powered, buoyancy-propelled autonomous underwater vehicle Seaglider (gliders). Gliders use changes in buoyancy and hydrodynamic lift to produce forward motion. While in motion, gliders collect subsurface oceanographic data along a sawtooth path, providing information that cannot be collected by other autonomous platforms that measure surface values. Gliders are particularly useful for long-term studies because the battery life can extend from 6 weeks to 3 months, with some missions exceeding over a year for a single deployment (Pelland et al., 2013; Rudnick, 2016). These passively powered platforms are particularly effective for oceanic research in the Red Sea due to their effective duration and their ability to access areas or conditions that would limit ship based scientific surveys.

Gliders have been effective in examining the water column in the Red Sea, a semi-enclosed basin lying between the Asian and African continents. The basin is about 2200 km long, has a maximum depth of about 2800 m, and an average width of 280 km (Raitsos et al., 2013; Triantafyllou et al., 2014). The basin is considered very saline and warm, and like most tropical waters is oligotrophic (Brewin et al., 2015; Gittings et al., 2019). The overall circulation in the Red Sea is based on a reverse estuarine system, with salinity increasing steadily with distance from Bab el Mandab (Sofianos and Johns, 2015; Sofianos and Johns, 2003; Sofianos and Johns, 2007). Primary energy transfer in the Red Sea is due to surface buoyancy loss and winds (Triantafyllou et al., 2014; Wahr et al., 2014; Yao et al., 2014a; Yao et al., 2014b; Zhan et al., 2016). The energy transfer is often expressed through semi-permanent eddies that occur throughout the basin (Karimova and Gade, 2014; Zhan et al., 2019; Zhan et al., 2014; Zhan et al., 2016).

Additional research has also described three characteristics of both wind forcing and sea level variability in the Red Sea and the associated temporal scales with each process (Churchill et al., 2018; Churchill et al., 2014b; Sultan et al., 1995). The first process is seasonal variability linked to the shift in the monsoonal winds over the Gulf of Aden flowing into the Red Sea. The second described component of sea level variability is the semi-diurnal tide. The third source of sea level variability, and the highest in magnitude of variability, is described as a weather system, with a duration of two to four weeks (Churchill et al., 2018).

As gliders are continually used as a tool within marine research, the primary roles and capabilities of the platform are expanding across several additional research fields. Several different practices in deployment patterns, data analysis, and intended targets are under development for operating gliders due to the context-dependent nature of sampling (Davis et al., 2019; Rudnick, 2016; Testor et al., 2019). One method for glider deployments is in conjunction with a marine observations system, which often includes a variety of shore-based data collection methods, modelling, and periodic ship-based research (Chao et al., 2017; Rudnick et al., 2017; Zarokanellos et al., 2017a; Zarokanellos et al., 2017b). Additionally, gliders are often used in conjunction with satellite derived data, typically to extend patterns observed on the surface deeper into the water column (Frajka-Williams et al., 2009; Little et al., 2018; Testor et al., 2019). These different oceanographic platforms collect data at different spatio-temporal resolution, which can confuse a clear analysis with mismatched spatio-temporal binning. Given the increasing use of gliders to compliment other forms of marine observation systems this study aims to identify the frameworks that are able to quantify the finest temporal range of data to

resolve oceanographic features while maintaining data independence across multiple spatio-temporal resolutions. We present a method to integrate gliders, High Frequency Radar (HFR, Chapter 2.2.1.4), and an atmospheric reanalysis product MERRA_2 (Modern-Era Retrospective Analysis for Research and Applications-2, Chapter 2.2.1.3) data to define short-term persistency, within the north central Red Sea. Notably this study used high resolution glider data, HFR and MERRA_2 data to test methods to quantify: 1) dominant scales of variability of the glider time series, 2) determine the minimum sampling frequency required to adequately characterize the glider time series and 3) discriminate whether the temporal variations measured from the glider are similar to variations determined from the HFR and MERRA_2 data.

2.1.1 Study Area

The chosen study area for this analysis is a portion of the north central Red Sea offshore from King Abdullah University of Science and Technology (KAUST, Figure 2.1). This study area was chosen for several reasons, including the frequency of eddies near the area, a semi-persistent boundary current, the sustained mapping of surface currents using HFR, coverage of recently validated data from the MERRA_2 data set (Al Senafi et al., 2019), and proximity to the university for access. The glider was deployed through the month of October 2017 with the intent to resolve the fine scale persistency of oceanographic variables in the region. October is an important time as it is part of the transitional period between the southwest monsoonal forcing in summer and the winter northeast monsoonal conditions for the Red Sea (Bower and Farrar, 2015; Sofianos and Johns, 2015; Yao et al., 2014a; Yao et al., 2014b). Because both the glider and the water it is sampling are moving both in time and space, this data set will be compared to two

data sets at fixed locations from the region. The first is the hourly recorded surface current velocities from HFR and the second is the hourly reported values from the MERRA_2 dataset (Rienecker et al., 2011) . Since these data sets are at fixed locations, they provide resources to cross validate the results from the glider collected data.

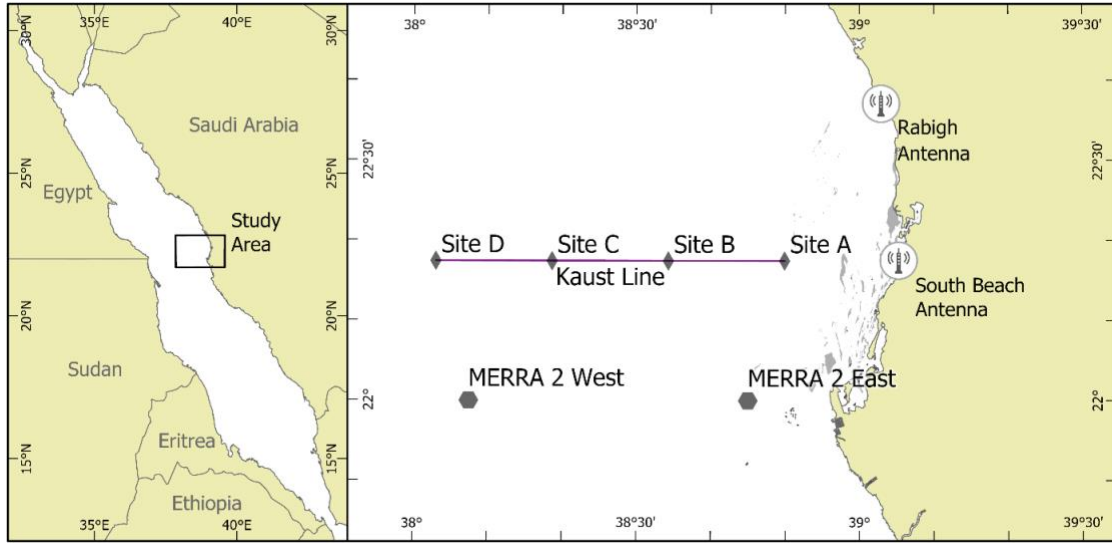


Figure 2.1: Location of the study area within the Red Sea. The glider transect (KAUST Line) is the solid line, the sites used from the HFR data are the diamonds (Sites A-D), and the locations for the MERRA_2 East and West sites are the hexagons, slightly south of the glider line. All three data sets run from October 1 -31, 2017.

Previous work utilizing *in situ* observations, satellite remote sensing, and modelling has shown that eddies are an important characteristic of the circulation in the Red Sea (Zarokanellos et al., 2017a; Zhan et al., 2018; Zhan et al., 2014; Zhan et al., 2016). These works have provided a basic description of the size, location, frequency, and duration of eddies in the Red Sea, and are indicative of the magnitude of eddies resulting primarily from thermohaline circulation and wind forcing. Eddies in the Red Sea typically last from two to six weeks, with smaller diameter eddies tending to occur more frequently, yet dissipate more quickly than larger eddies. However, when multi-year averages of

surface circulation are examined, the short-term eddies are no longer seen, but instead a clear boundary current system is in effect (Sofianos and Johns, 2003). This eastern boundary current has been observed in numerical models and with *in situ* observations and is under continuing investigation using both glider and HFR observations. These observations indicate that the eastern boundary current can typically be defined by a sharp decrease in values for salinity (<39.5), indicating water that recently entered the Red Sea from Bab el Mandeb. The water in the boundary current also has increased values for temperature and chlorophyll- α when compared to the surrounding water, as can be seen in Figures 2.2 and 2.3, along the western part of the transect line. The exact location that the eastern boundary current crosses the KAUST transect (Figure 2.1) can shift between the nearshore and offshore region depending on the influence of local eddies. Over time, persistency measurements can show when a major shift in the underlying dynamics of a system occurs. Within the context of the north central Red Sea, it is expected that the transition between the eastern boundary current and its interruption by local eddies will have an impact on the persistence in the study area, providing an ideal system to develop methodologies to measure the dominant scales of variability.

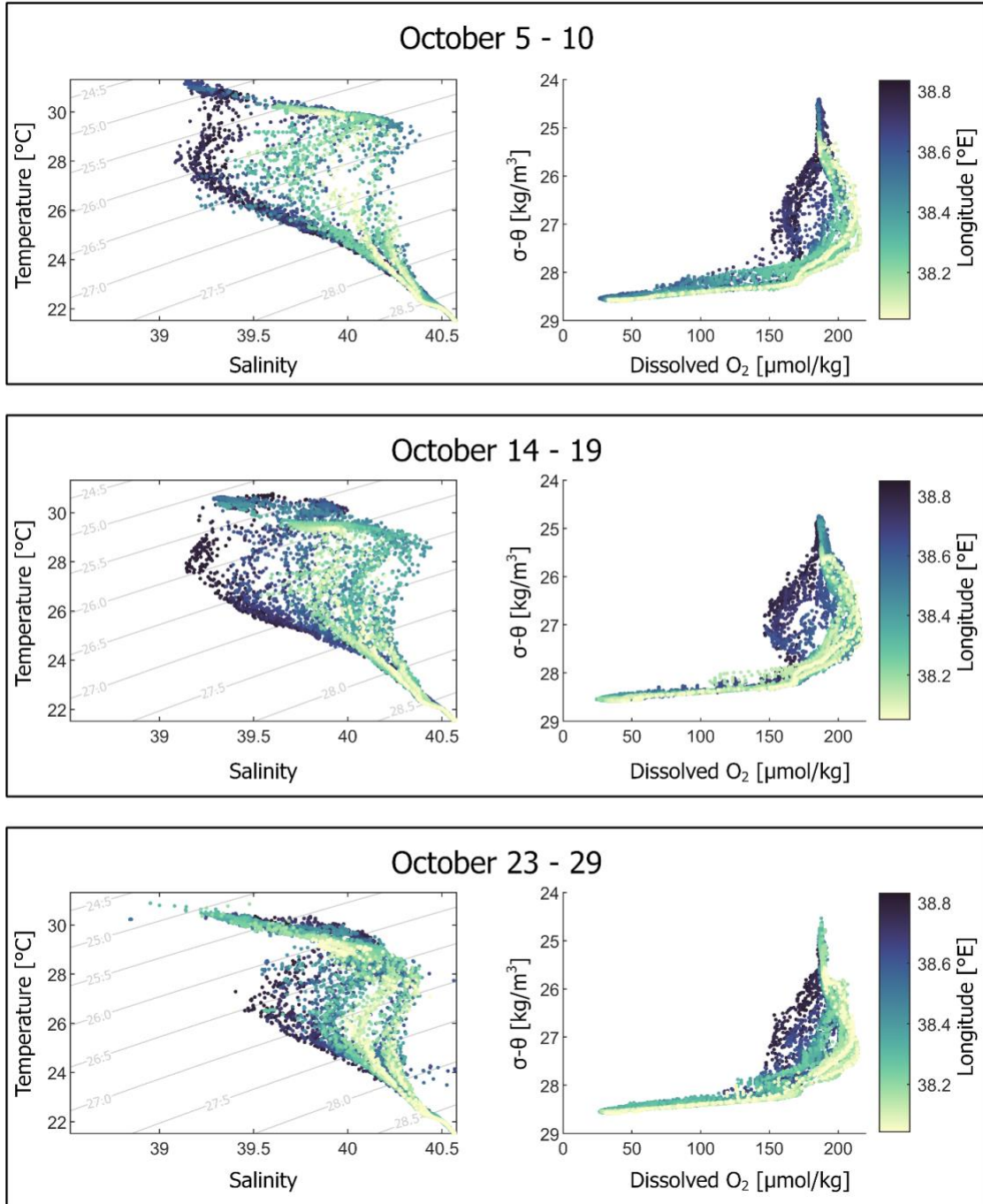


Figure 2.2: Temperature-salinity and σ - θ dissolved oxygen values for selected times of the glider mission. The data shows two diverging water masses for the middle range of the water column based on longitude in the early part of the month. Over time, these two water masses start to intermix.

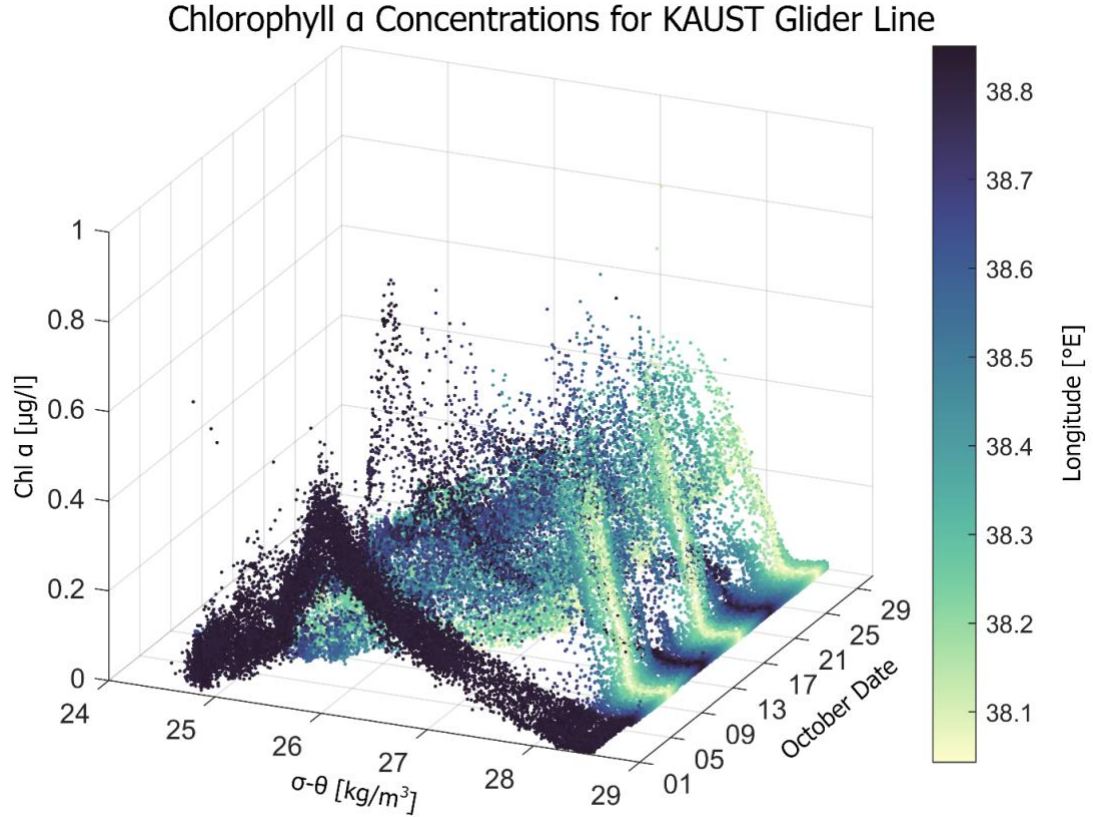


Figure 2.3: Chlorophyll- α concentrations for the KAUST glider line over the month of October. The isopycnal location of the peak in chlorophyll- α is different for the nearshore region (higher longitude) than the offshore region. Towards the end of the month, the two maxima appear to be merging.

2.2: Methods

2.2.1: Measurements used in this study

2.2.1.1: Glider data

The primary data source for this research is from a glider (Seaglider®) deployment that occurred from September 2017 into November 2017. The glider was equipped with a CTD, a dissolved oxygen sensor, a 3-wavelength fluorometer, and a 3-wavelength

backscatter sensor (Table 2.1). The glider traversed an 80 km east-west transect between a starting waypoint approximately 20 km offshore to an offshore waypoint approximately 100 km from the coast. The glider was set to dive to 500 m depth which represents a safe dive depth along the transect where the bottom depth varied from less than 600 m to more than 1000 m, and a dive depth which optimized horizontal resolution at about 2 km per dive with a dive interval of about 2.33 hours along the majority of the transect. Data from the glider were collected from the entire dive cycle. However, the sampling frequency varied as a function of depth. The sampling frequency was once every 10 seconds in the upper 100 meters, then once every 50 seconds between 100 and 250 m, and was reduced to once every 100 seconds between 250 m and 500 m. The sampling frequency corresponds to data collected once every 1.3 m for the upper 100 m, once every 6.5 m from 100 to 250 m, and one every 13 m from 250 to 500 m. For this analysis three representative depths (6 m, 75m, and 150 m) and isopycnals (25.75 kg/m^3 , 26.75 kg/m^3 , and 27.75 kg/m^3) were chosen for analysis. Since the goal of this research is to compare spatio-temporal patterns from glider collected data to existing data sets at the surface, the selected depth and isopycnal layers were selected to be in the upper portion of the water column extending into the upper portion of the subpycnocline layer. The depth of 6 m represents the upper limit of consistent measurements from the glider since the vehicle can take several meters to reach steady flight, where the different response of the sensors (temperature, salinity, chlorophyll α , dissolved oxygen) can be accurately accounted and corrected for. The depth layer of 75 m represents the typical depth of the chlorophyll- α maximum from the collected data and the 150 m depth layer is below the deepest recorded mixed layer depth from the data set. The isopycnals were selected for a

different reason, the 25.75 kg/m³ isopycnal is the shallowest isopycnal that extends through the entire collection period, and the other two isopycnals were selected as an even incremental increase until the denser waters located below 28.00 kg/m³. These depths and isopycnals bracket the upper and lower extents of the water column experiencing greatest variability.

Equipment	Parameters	Excitation [nm]	Emission [nm]
CTD (Seabird CT-sail, unpumped)	Temperature Conductivity		
WET Labs	Chlorophyll- α	470	695
ECO Puck	CDOM	370	460
(FL3)	Phycocyanin	630	680
WET Labs	Optical	532	532
ECO Puck	backscatter at 3	650	650
(BB3)	wavelengths	880	880
Oxygen sensor (Aanderaa optode, model 4331)	Dissolved Oxygen		

Table 2.1: Summary of the glider equipment used in this study.

2.2.1.2: Processing and quality control of glider data

Raw instrument measurements (counts) for each parameter were transformed into geophysical quantities by applying the manufacturer-provided scaling factor and dark count. Then, each profile was quality controlled by applying methods that have been specifically developed for each parameter (Organelli et al., 2017; Thierry et al., 2018)

following Argo's 'real-time' quality control procedure and data management (Wong et al., 2020).

Following Schmechtig et al. (2014), vertical profiles of chlorophyll- α were adjusted for non-zero deep values and corrected for non-photochemical quenching according to Xing et al. (2012) (Schmechtig et al., 2014; Xing et al., 2012; Xing et al., 2017). Furthermore, the chlorophyll- α values were then divided by a factor of two to correct the overestimation observed by standard Wet Lab fluorometers as described in Roesler et al. (2017) (Roesler et al., 2017). Spikes were removed from the chlorophyll- α measurements using a low-pass median filter. Correction of O_2 measurements were performed by applying a factor deduced from the comparison between the O_2 values obtained via Winkler titration and analysis (Carpenter, 1965; Winkler, 1888) and those from the glider. During the glider deployment, water samples were collected at nominal depths (5, 10, 30, 50, 100, 150, 200, 300, 400, and 600 m) and stored in 125 mL iodine titration flasks following standard operating procedures (Langdon, 2010). The samples within the surface (5 and 10 m) were collected in triplicate. The amount of O_2 in each sample volume was converted in $\mu\text{mol}\cdot\text{kg}^{-1}$ using the measured density of seawater. Finally, temperature, salinity, chlorophyll- α and O_2 quality-controlled vertical profiles were binned in 2 m depth intervals and interpolated onto a grid with a 2.33-hour temporal spacing in this study, a regular data set needed for statistical analysis as described in 2.2.2. Additionally, these same quality-controlled files were binned in 0.125 kg/m^3 density intervals at the same 2.33-hour temporal spacing. The time series for the upper 250m from the glider deployment is shown in Figure 2.4 and the density-based profiles from 24.5 to 28.5 kg/m^3 are presented in Figures 2.5.

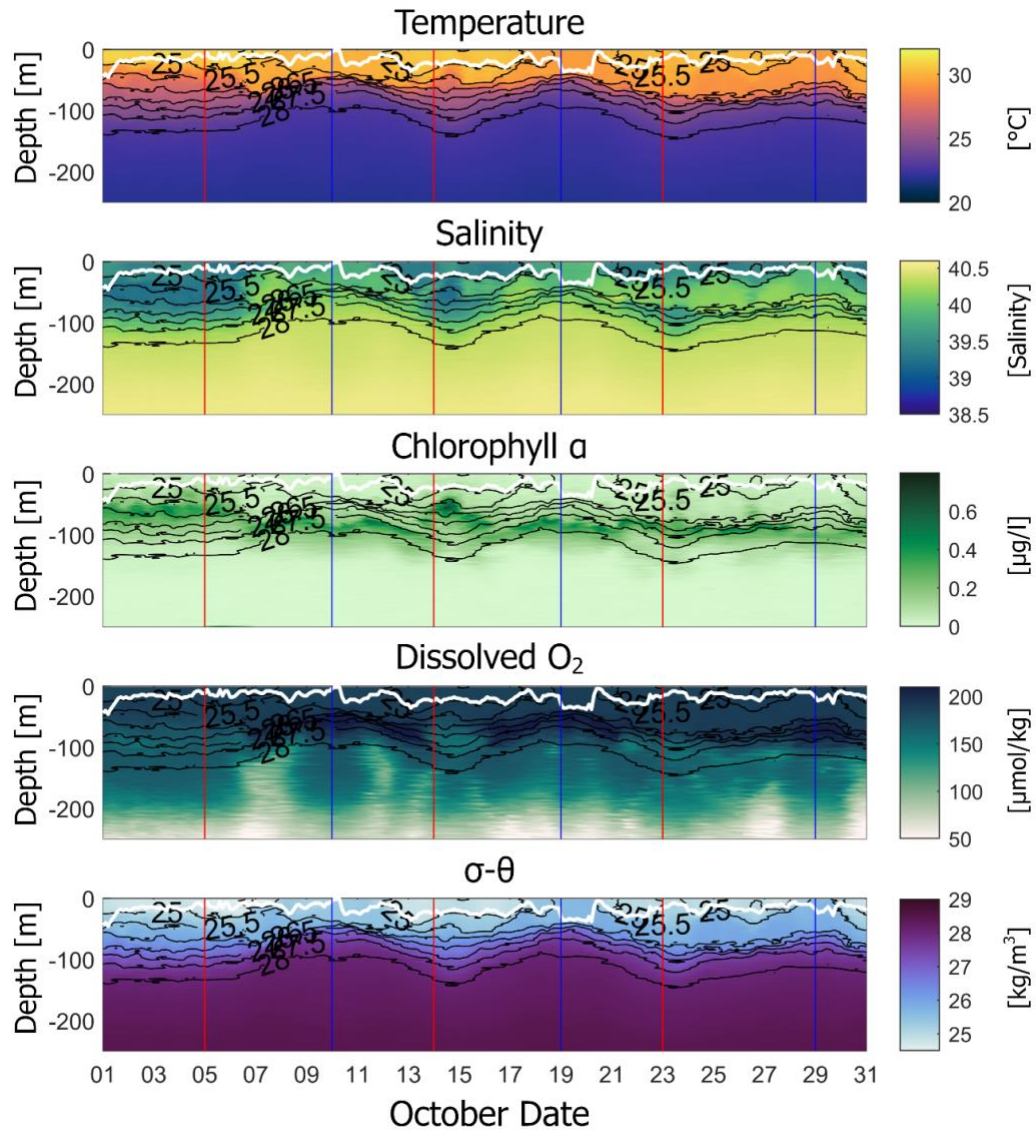


Figure 2.4: Kaust glider line time series plots of temperature, salinity, chlorophyll- α , dissolved oxygen, and sigma-theta from 0-250 m. The white line indicates the mixed layer depth based on the de Boyer index (de Boyer Montégut et al., 2004). The red vertical lines indicate when the glider is inshore, the blue lines indicate when the glider is offshore. All color ramps used for the figures in this research are using color ramps from cmocean (Thyng et al., 2016).

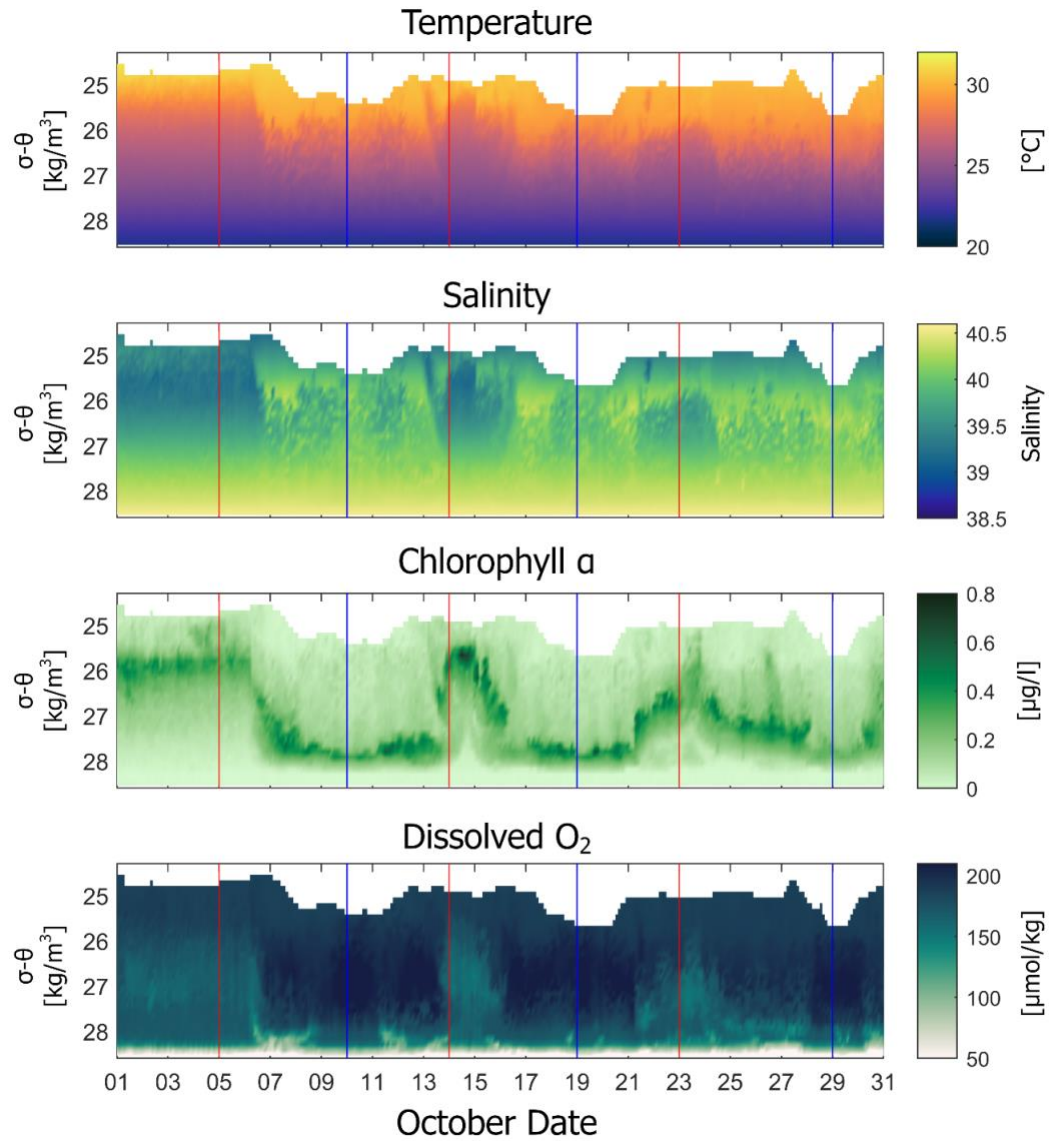


Figure 2.5: Kaust glider line time series plots of temperature, salinity, chlorophyll- α , and dissolved oxygen from 24.5 - 28.5 kg/m³. The red vertical lines indicate when the glider is inshore, the blue lines indicate when the glider is offshore.

2.2.1.3: MERRA_2 data

Since the introduction of NASA's Modern-Era Retrospective Analysis for Research Applications version 2 (MERRA_2) (Gelaro et al., 2017; Rienecker et al., 2011) several

publications have examined the application of this new reanalysis data set to provide additional understanding of the seasonality occurring in the Red Sea. The reanalysis data from MERRA_2 is especially useful for the region due to the lack of long-term meteorological measurements within the region. Three recent publications have described in detail the degree of correlation between the reanalysis data from MERRA_2 and data from a meteorological buoy that was deployed for two years near the study area. Al Senafi et al. (2019) demonstrated that MERRA_2 data shows a high degree of correlation ($r = 0.97-0.98$) for heat flux when compared with buoy data (Al Senafi et al., 2019). Menezes et al. (2019) describes the correlation ($r = 0.89 - 0.92$) for wind speed, zonal velocity, meridional velocity, and evaporation rates (Menezes et al., 2019). Sun et al. (2019) has used MERRA_2 as an independent data set for the verification of a new predictive model called SKRIPS (Sun et al., 2019). Based on the validation of the MERRA-2 data by other investigators, the data set is used here as a complementary data set that will facilitate interpretation of our results.

MERRA_2 data (MERRA2_400.tavg1_2d_ocn_Nx files) from two locations nearest to the KAUST glider line (38.75 E, 22.000 N and 38.125 E, 22.000 N, Figure 2.1) are used to compare persistency calculations from of the glider and HFR data set from October 2017. The parameters selected for analysis are air temperature and wind velocity at 10 m, wind curl, and the water skin temperature (3-day average). This data set provides a resource to understand additional forcing components that contribute to eddy driven dynamics.

2.2.1.4: CODAR data

The HFR system is located on the central western coast of Saudi Arabian and consists of two CODAR Seasonde® sites in operation since July 2017 (Solabarrieta et al., in review). The HFR system transmits at the 16.12 MHz frequency with hourly data provided over a 120 km range with a spatial resolution of 3 km. The received backscatter signal was converted into radial velocities using the MUltiple SIgnal Classification algorithm (Schmidt, 1986). The MATLAB package HF Progs (<https://cencalarchive.org/~cocpmb/COCMPwiki>) was then used to combine radial currents and generate gap-filled total, two dimensional currents using Open Modal Analysis (Kaplan and Lekien, 2007) for the entire month of October 2017. Both before and after October 2017, errors in the HFR limited the amount of data available to create gap filled currents using Open Modal Analysis. Four evenly spaced locations across the KAUST glider line were selected for time series analysis, as seen in Figure 2.1 (Sites A-D).

2.2.2: Time series analysis used in this study

One of the difficulties of working with data collected during glider missions is that the glider is likely not sampling the system synoptically (Rudnick and Cole, 2011; Thomson and Emery, 2014). The lack of synopticity increases the amount of inherent dependency that occurs within the study area, which then increases the minimum length of time data is needed to understand the system. While an increase in available information in an area can greatly improve the understanding of the physical processes, many areas across the world have limited time series data due to a lack of previous observation opportunities

(Davis et al., 2002; Davis et al., 2019; Rudnick et al., 2004). Many of the new observation platforms were specifically designed to be deployed in areas that lack long-term time series data, and the presented analysis methods were designed to operate in systems that may not be data rich (Lee et al., 2012; Lermusiaux et al., 2017; Webster et al., 2014). Because gliders travel at approximately 0.25 m/s (20km/day), it can be difficult to separate out which portion of the observed variation is due to a change in time, longitude, latitude, density, or depth. In the case of this study, latitudinal variation was limited by maintaining the same latitude within a data set as closely as possible. While mission planning is potentially able to minimize the observed variation of some of these dimensions, minimizing these features *a priori* risks an incorrect characterization of the natural variability in the system, especially when the sampling locations are located too close to each other in time or space (Ellis and Schneider, 2008; Hewitt et al., 2007; Holland et al., 2004). This study presents two statistical methods that can be used to determine the natural framework of time analysis within the study area. This study seeks to define the optimal sampling frequency to provide both an effective characterization of the system and ensure that the full range of natural variation is also described.

2.2.2.1: Autocorrelation analysis

Autocorrelation is defined as the normalized cross covariance of a time series data set at increasing distant time lags, and can be calculated from non-linear but uniform data sets (Habeeb et al., 2005; Thomson and Emery, 2014). Autocorrelation analysis identifies the time lags at which the time series data set has significant dependencies or correlations. Stochastic processes are time dependent and show a decreasing correlation with time (Thomson and Emery, 2014). One method of eliminating stochastic processes in time

series data sets is to average the data set over regular intervals. If the selected interval is short enough to be highly correlated, then the assumption of data independence is violated. The time that this data dependency occurs over can be determined by comparing the correlation at increasingly distant data points within the time series data (Thomson and Emery, 2014). At a lag of zero, data is perfectly correlated with itself. In a typical system, the magnitude of autocorrelation tends to decrease as the time lag increases, and the correlation remains close to zero and no longer crosses the threshold into significant dependency. Data collected at such lags are considered independent for statistical analysis (Thomson and Emery, 2014). In situations in which data follows a cyclical pattern, like a diurnal cycle, then the autocorrelation analysis will typically oscillate between regions of positive and negative correlation until the correlation values fall below significant levels, with each peak in the correlation indicating the duration of the cycle. The time lag at which the last significant autocorrelation peak is called the Integral Time Scale (ITS). ITS defines the scale where lag transitions from dependency to independency, and indicates lag at which the time series data no longer has a significant level of either positive or negative correlation (Thomson and Emery, 2014).

2.2.2.2: Characteristic Length Scale (Error X) analysis

The Characteristic Length Scale (CLS) analysis for this research is an adaptation of the method published in Ward et al. (2018), which described Error X as a way to calculate CLS from existing transects and further suggested that the method could be applied to glider and tow vehicle data sets (Habeeb et al., 2005; Keeling et al., 1997; Ward et al., 2018). The layers selected are from the upper portion of the water column which are most influenced by surface conditions, and thus relevant for comparisons with

MERRA_2 and HFR data. This focus allows for a better analysis of the relevant time-space scales for the upper portion of the water column by incorporating the surface forcing (MERRA_2) and the surface current maps recorded at finer spatial resolution.

Two circumstances were described in which Error X can be used to calculate CLS (Habeeb et al., 2005; Ward et al., 2018), where both methods are using changes in the spatial dimension as a replacement for time. The first circumstance is a short-time series approach, and the second is adapting measurements over space as a replacement for measurements over time. Even though the transect line developed for this research was set up to examine both transitions through time and space, this paper will only focus on the short time series approach, which includes a combination of relatively few time steps coupled with the spatial displacement across a single time step (Habeeb et al., 2005). This is limited due to the relatively small spatial extent of the glider deployment mission, which prevented determining only the spatial CLS. CLS calculation is an iterative process comparing the difference between the recorded value and the predicted value for increasing larger separation distances. Predictive values are calculating by using a k Nearest Neighbor (kNN) approach. examining the difference between a single point in the data set and the average of values located both before and after the data point at increasing distances. This difference is calculated at all distances up to half of the overall length of the data set. The error curve is then calculated by using a Monte Carlo chain simulation (100 runs) averaging twenty randomly selected error measurements at each distance bin. These 100 Error X results are then averaged together to produce a final mean CLS calculation with an associated 95% confidence interval (Habeeb et al., 2005; Keeling et al., 1997; Ward et al., 2018). As measurements are further removed in either

space or time (or both) from a point of interest, the error between the prediction method and the actual values is expected to steadily increase until the error becomes asymptotic. CLS is typically defined as the distance at which the error curve becomes asymptotic. While other prediction methods instead of kNN moving mean have been described since Keeling et al. 1997, this method was selected as a proof of concept in extending the CLS approach into glider related studies.

2.3: Results

The KAUST glider line shows evidence of a transitional period within the month-long data set presented. As seen in Figure 2.2, in early October the KAUST line shows a sharp difference between the nearshore and offshore regions in both the temperature-salinity diagrams and dissolved oxygen. These figures have the longitudinal location of each sample shaded to highlight the difference between the two regions. The two sections of the line show different characteristics through the water column, with the nearshore region exhibiting less saline water with lower dissolved oxygen. Over the course of the month, the ends of the line become less distinctive. Figure 2.3 shows the time series data for chlorophyll- α and density across the full time of the glider mission. This figure is also shaded by longitudinal location, using the same range as Figure 2.2. This figure also shows a similar result when comparing between the density location of the peak chlorophyll- α concentration in the water column. The nearshore region shows a shallower peak, located between 26.00 – 26.25 kg/m³, while the offshore region shows a deeper peak at the 28.00 kg/m³ isopycnal. Both Figures 2.2 and 2.3 show that these differences become less distinct towards the end of the month.

Both the 0-250 m and the 24.5 – 29.00 kg/m³ time series figures (Figures 2.4-2.5) for the KAUST transect line show a difference between the nearshore region and the offshore region during the start of the month of October. These differences indicate that the two regions of the transect line are experiencing different conditions. The time series figures also show a transitional region between the ends of the transect, indicating the potential of interchange between the nearshore and offshore region. This idea is further supported by the decrease in the difference of the nearshore and offshore regions towards the end of the month of October. The variance seen in temperature, salinity, chlorophyll- α , and dissolved oxygen over the course of this single month supports the idea presented by Churchill et al. of the weather band (2-4 weeks) playing an important role in understanding the natural variance occurring within the Red Sea (Churchill et al., 2018). The time series data set for each variable at each of the three depths (6, 76, and 150 m) and the three density layers (25.75, 26.25, 28.00 kg/m³) are presented in Figures 2.6 and 2.7.

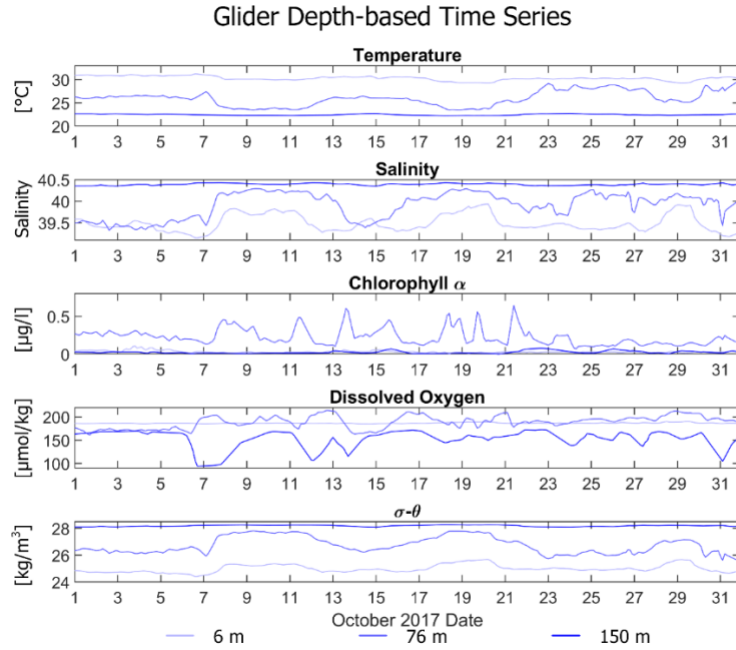


Figure 2.6: KAUST glider line time series profiles for temperature, salinity, dissolved oxygen, chlorophyll- α , and density anomaly at 6, 76, and 150 m depth.

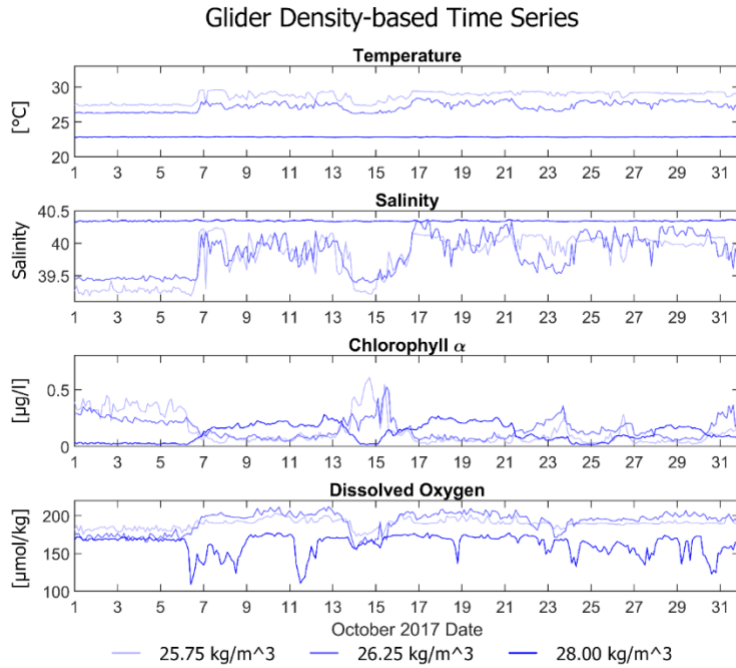


Figure 2.7: Kaust glider line time series profiles for temperature, salinity, dissolved oxygen, and chlorophyll- α at the 25.75, 26.25, and 28.00 kg/m³ isopycnals.

The HFR data set (Figure 2.8) presents a similar story with differences between the nearshore region and the offshore region along the KAUST transect. As with the glider data, the HFR data in the nearshore region (Site A) is different from the rest of the HFR sites (B, C, and D). Velocity from the HFR is recorded as two directional velocity vectors, the u vector (- west, + east), and the v vector (- south, + north). Overall, the surface currents recorded are predominantly northward over the transect line, with both higher magnitudes and variance in the north/south direction. The magnitude of northward surface currents also increases going from east to west. Diel variation is also apparent throughout the time series dataset, typically more noticeable in the east/west direction. A final item of note is a relatively extreme event that occurred on October 29th, in which the typical northward currents shifted southward over the course of a few hours. This shift occurred at sites B, C, and D, missing the nearshore region of the transect.

The time series figures for the MERRA_2 data (Figures 2.9 and 2.10) also show a difference between the east and west site. In this data, the east site has a higher degree of variability in the diel signal, especially with regards to the 10 m air temperature. The east site, which is closer to the shore, shows that the wind flows are predominately eastward and southward during the study period, with higher variation in the magnitude of the eastward flow. The western site overall shows less variability through the study period, indicating that the location has higher stability than the east site potentially due to increased distance from the shore.

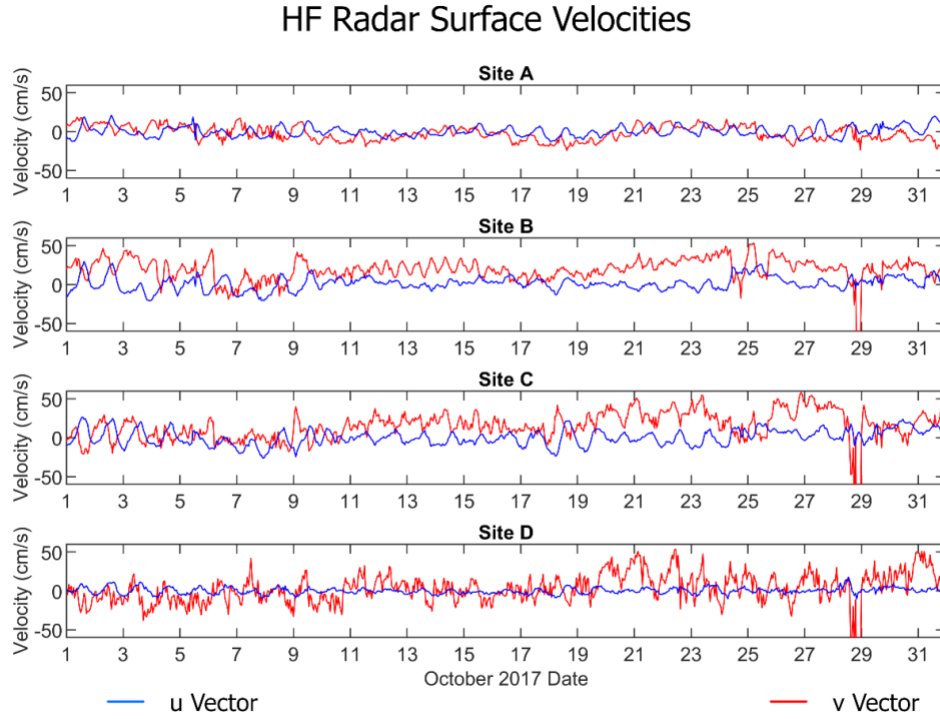


Figure 2.8: HFR surface velocities time series at the four sites located across the KAUST glider line. Location of the sample sites can be found in table 2.

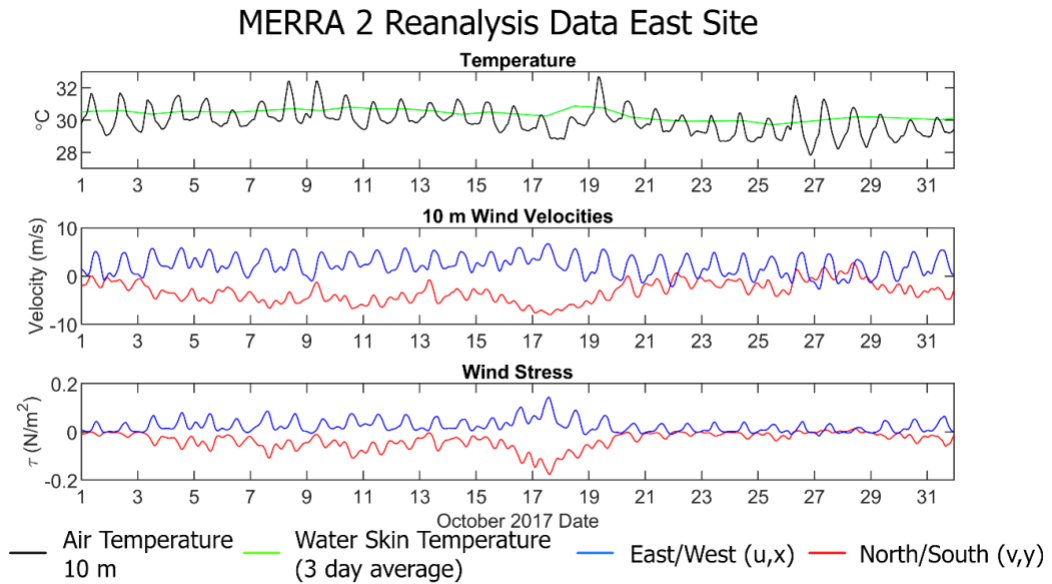


Figure 2.9: MERRA_2 time series for East Site (38.75 E, 22.000 N).

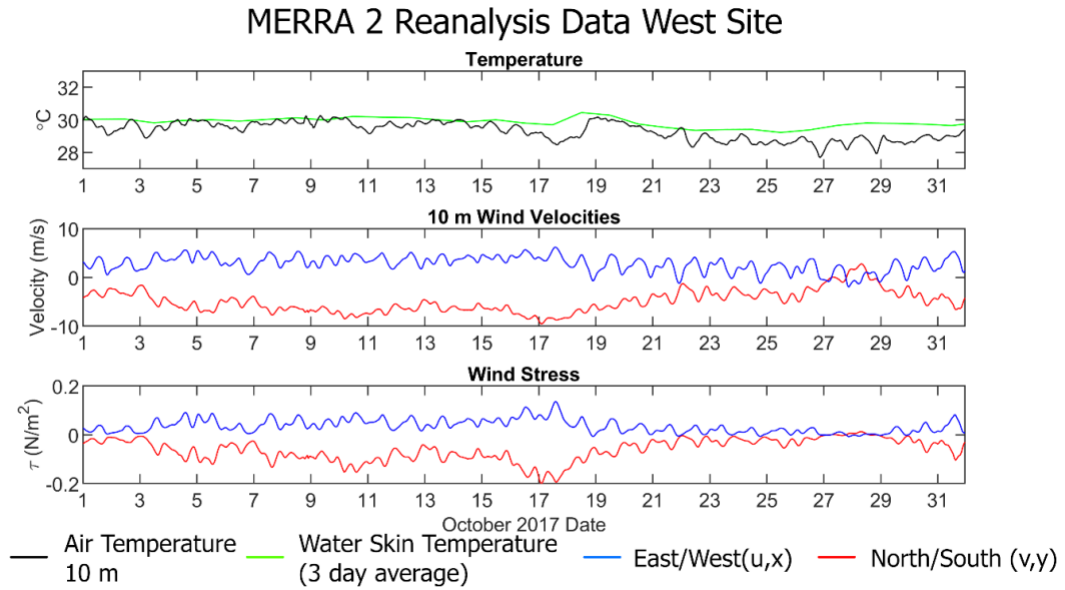


Figure 2.10: MERRA_2 time series for West Site (38.125 E, 22.000 N).

Figure 2.11 presents the results of both the ITS and CLS calculations for the two glider derived data sets, and Figure 2.12 presents the results for the ITS and CLS calculations for the HFR and MERRA_2 data sets. The full results from the ITS and CLS (with 95% confidence intervals) from this chapter can be found in Appendix A. Each variable was used directly for both the ITS and CLS calculations. The ITS was calculated by selecting the last significant (95% confidence interval) positive or negative autocorrelation peak. The CLS was calculated for each variable by visually estimating the time distance that the Error X plot became asymptotic as a function of time lag. The results for ITS fall in the 2-4-week range, with a mean time lag of 19.1 days (std 4.6, median 16) for the glider data and 21.9 days (std 2.8, median 21.5) for the HFR and MERRA_2 data sets. The CLS calculation also matches the 2-4-week range expected, with a mean time lag of 19.6 days (std 4.0, median 19) for the glider data and 21.5 days (std 4.1, median 22) for the

HFR/MERRA_2 data. The glider ITS and CLS do not show any significant difference between the different depth or density layers. Four of the variables did not have a sufficiently long time series to reach an asymptote. These four variables are chlorophyll- α and dissolved oxygen at 76 m, chlorophyll- α at 6 m, and temperature at the 26.25 kg/m³ isopycnal. Three of these variables are a part of biogeochemical processes. Salinity, temperature, dissolved oxygen, and chlorophyll- α show variations that coincide with the position of the glider in the study area, with these differences decreasing over the course of the study period. This time frame matches with the expected weather system time period as described and is within the typical duration of eddies within the Red Sea (Churchill et al., 2018; Karimova and Gade, 2014; Zhan et al., 2014).

Figure 2.12 shows the comparison for ITS and CLS for the MERRA_2 and HFR data sets. One of the key items to note is the lack of a CLS estimate for the MERRA_2 data at the west site. This is due to the Error X plots not converging to an asymptote, which indicates that the time series data was not sufficiently long to resolve the CLS. Aside from the non-resolved time series, both methods show that the typical persistence time is in the 3 – 4-week range, with only one value, u surface velocity at HFR site B showing a persistence value of 15 days.

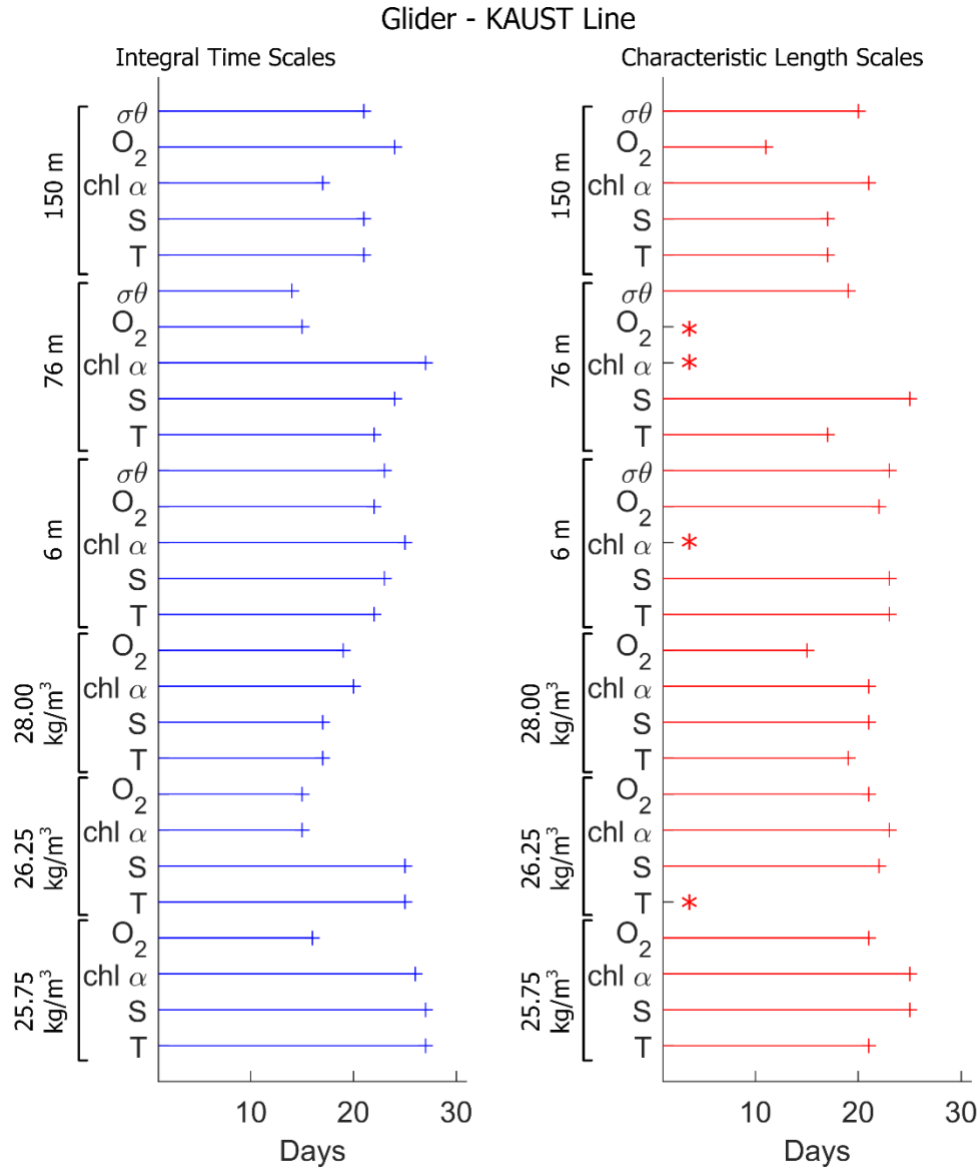


Figure 2.11: Integral Time Scale (ITS) and Characteristic Length Scale (CLS) results from the glider time series data. The average time lag is 19.1 days for ITS and 19.6 days for CLS, which is within the 2 – 4-week weather band period. While some results between the two methods are contradictory, the averages are similar. CLS values marked with an asterisk did not reach an asymptote.

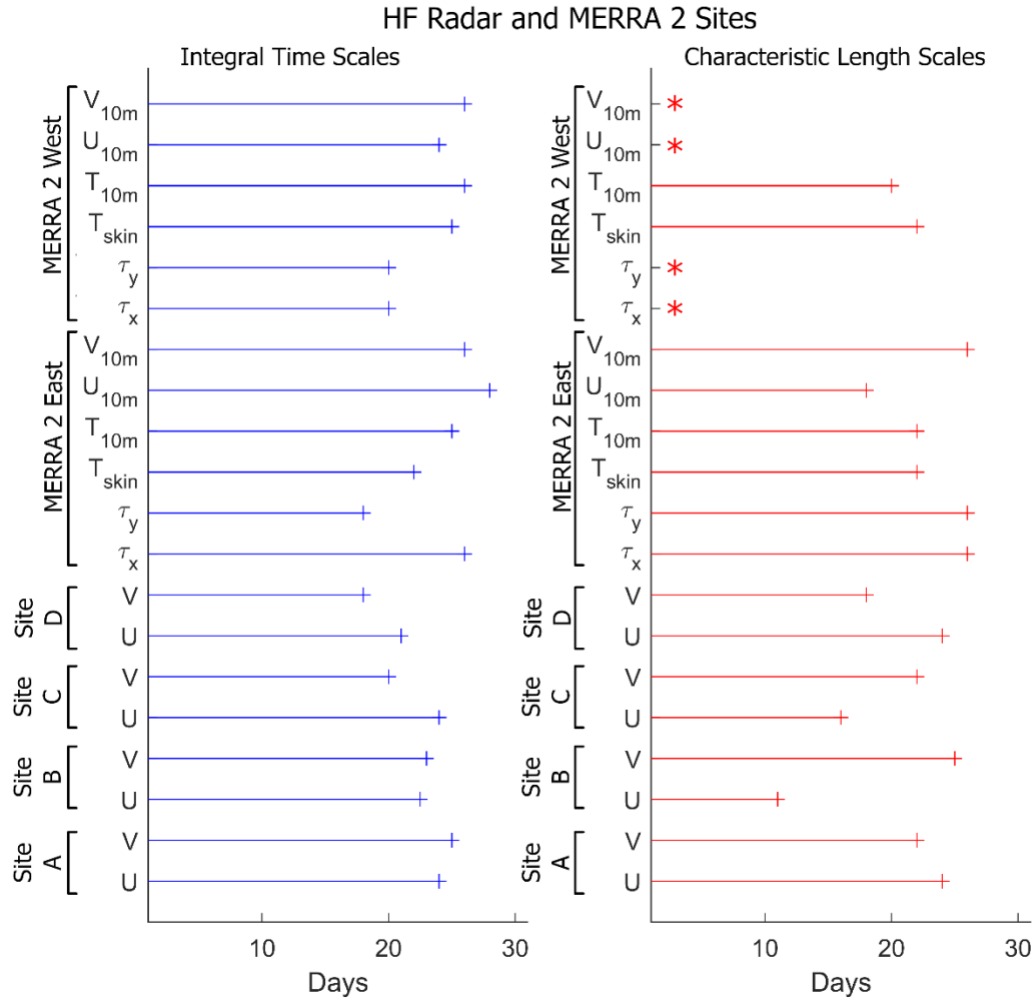


Figure 2.12: Integral Time Scale and Characteristic Length Scale results from the HFR and MERRA_2 data. The average time lag is 21.9 days for ITS and 21.5 days for CLS, which is within the 2 – 4-week weather band period. While some results between the two methods are contradictory, the averages are similar. CLS values marked with an asterisk did not reach an asymptote.

2.4: Discussion

This analysis of glider, HFR, and MERRA_2 data has successfully trialed two different analytical methods to determine the dominant scales of temporal variability (persistence) in the central Red Sea. Both methods, the Integral Time Scale (ITS) and Characteristic Length Scale (CLS) calculations, were successfully used to determine the short-term persistence from transect based marine data across several important platforms for ocean observation systems. When considered together, both the ITS and CLS methods arrived at similar values for the persistence scale over the month-long data set, with a mean time lag of 19.06 (ITS) and 19.59 (CLS) days for the glider data and 21.93 (ITS) and 21.5 (CLS) days for the HFR/MERRA_2 data. Differences between some of the physical variables (surface velocity, wind speed, salinity, temperature) and biogeochemical variables (dissolved oxygen and chlorophyll- α) potentially do exist, but the presented data set is too small to draw any conclusions at this point. However, the two methods were able to achieve results that support previous analysis into persistency within the study area and that match with our current understanding of the dominant physical processes occurring with the study area. Furthermore, this analysis was able to show that even processes that started out under different initial conditions were still influenced at the same time scales previously identified. These results show that these two analysis methods can be further applied to other similar ocean observation data sets that cover different spatial and temporal extents. The CLS calculation will be useful moving forward, as the method is designed to handle higher dimensional data sets. An additional benefit of the CLS calculations is that the method functions even with small data gaps, which is not true of some other time lagged variance measurements, including ITS. As a

part of mission planning, the persistence scales identified using CLS can be used to ensure that planned transect lines are able to fully meet the Nyquist frequency requirement for time series data sets, by establishing at what time frame a transect line can be considered synoptic.

2.4.1: Temporal Scales

One of the more difficult aspects of working with glider data is the inherent spatio-temporal blurring due to the movement of both the glider and the water it is sampling. Gliders provide detailed information across their entire range of operation, however specific events seen from glider data can be difficult to match with data from other platforms. This difficulty in matching data from different platforms indicates the importance of developing adaptive systems that can maintain data quality while adjusting sampling frequencies or locations. This fact has been a topic of research for many ocean observation systems, trying to find the best way to tie in the multiple instrument platforms to produce a highly detailed and accurate representation of the spatial and temporal processes occurring within the area of interest. The persistence measurements presented here have been able to corroborate with previous studies using different instrument platforms, indicating the importance of the 2-4 week-long weather system. This information is important to resolve some of the issues with spatio-temporal blurring between instrument platforms by providing an initial time bin to ensure independence between measurements. While this initial approach is only applicable for the short time frame considered for a small portion of the Red Sea, it does provide an initial framework to be able to compare multiple platforms, by setting a uniform initial scale to examine the varied data sets. As more data is both collected and analyzed, this persistence

measurement can be updated as needed. Within the context of this study, the transect line showed differences at the beginning of the study period between the inshore and offshore sections. However, over the course of the month both sections show similar persistence values, indicating that external forcing on the two sections is producing the same degree of variation. Additionally, evidence suggests that towards the end of the month of October consistency increases in the water column along the entire transect, suggesting either that the transect line has become more well mixed or that the differences noted earlier in the month had moved away from the transect line. While the conditions appeared to be different at the start of the study period, the entire transect was influenced by the same factors, indicating a degree of dependency between the inshore and offshore regions.

2.4.2: Comparisons between Glider, HFR, and MERRA_2

As more variables are included in an ocean observation system, it becomes increasingly difficult to incorporate the available data sets into a single unified snapshot of the oceanographic condition. Even when only considering surface-based measurements, there are limitations to the level that satellite and *in situ* measurements can be integrated. These integration methods require either spatial or temporal averaging or applying an offset to either the *in situ* data or satellite derived products (Bouffard et al., 2012; Liblik et al., 2016; Little et al., 2018; Pascual et al., 2017; Testor et al., 2019). Trying to integrate surface conditions to processes occurring at depth is an equally important challenge. The addition of gliders to ocean observation systems is specifically targeting the need to extend surface-based measurements into the water column at high temporal resolution. Recent work with ocean observation systems has been focused on comparing

glider derived data with surface measurements including comparing geostrophic velocities from satellite derived altimetry, numerical simulations, and gliders to reduce errors in coastal waters. In coastal areas, satellite derived data sets tend to have reduced quality due to different factors, such as cloud coverage, complicated sub surface structuring of the water column, or interference from land. The Red Sea presents an ideal location to test new methodologies in multi-platform integration due to the relatively simple conditions present. This aspect of the Red Sea has led it to be of interest in many numerical simulation studies and is applicable for this study as well. Given our current understanding of the forcing processes in the Red Sea, it is expected and confirmed by this study that the typical processes occurring in the north central Red Sea should be consistent throughout the at least the upper 150 m of the water column, as is demonstrated in both the ITS and CLS analysis. We have shown that the previously described persistence time frames are replicated not just at the surface, but also extend deeper into the water column, with the weather band time frame consistent down to 150 m and the 28.00 isopycnal, with no significant difference between the two methods or the different depth and density layers. We also see that the persistence time frame of eddies matches with the persistence seen through the ITS and CLS calculations, providing further evidence of the importance of eddies in the circulation patterns in the central Red Sea.

2.4.3: Conclusion

While the study is small in both spatial and temporal extent, we were able to find a persistence time frame for many of the measured variables. This persistence time frame is the distance between two individual time steps that we can ensure independence

between the data points. Both the ITS and CLS analysis can be used to give an understanding of the typical time frame or cycle of natural variation within a system. In this study area, previous research has shown that most of the circulation energetics are based on eddy dynamics. Both the ITS and CLS calculations fall within the expected time frame based on eddy persistence in the Red Sea (2 – 6 weeks) with a mean time lag of 19.1 (ITS) and 19.6 (CLS) days for the glider data, and 21.9 (ITS) and 21.5 (CLS) for the HFR/MERRA_2 data. While this initial study encompasses a relatively short time series over a small spatial extent, this method of analysis still matches the current understanding of the local dynamics and thus shows potential for examining larger spatio-temporal data sets.

As seen from this research, marine data sets are typically both non-linear and non-regular, as both the natural movement of water and the logistics with collecting data often preventing fully systematic data collection plans. The non-linear and non-regular nature of marine data limits the types of statistical analysis that can be performed, either by averaging data to remove non-linear patterns or using gap filling methods to ensure regularity. In this research, the ITS analysis requires that data been both detrended and set to a uniform time step, which introduces a degree of information loss. In contrast, the CLS method was specifically developed with non-linear and non-regular data in mind, increasing the overall information retained through analysis. One of the typical considerations for any spatio-temporal analysis is finding the right balance between resolving fine scale details and managing to fill in gaps in coverage by either averaging or interpolation. One of the important findings from this research is that the CLS calculation was more robust to gaps than the ITS calculation, giving a greater degree of

flexibility over data that includes gaps. This can be an important factor within marine systems, where often even with the best efforts result in occasional gaps.

Local and regional conditions play an important role on marine systems across the world. The Red Sea itself has a unique set of conditions, especially when considering the temperature and salinity ranges recorded in the basin. Even within the Red Sea, conditions are very different depending on the location within the basin. While the information presented for this study only directly applies to this glider mission, the methods presented can be adopted to other similar data sets both within the Red Sea and across the world.

Chapter 3: Localized influence of two counter-rotating eddies in the north central Red Sea as seen by two parallel AUVs

3.1: Introduction

In recent years interest has increased in the Red Sea as a resource for scientific investigation. This recent increase has been driven in part by an increase in the availability of remotely sensed data from satellite, accessibility of the Red Sea for *in situ* research, and the relative simplicity of the basin as an entire system (Jones and Kattan, 2017; Raitsos et al., 2013; Sofianos and Johns, 2015; Yao et al., 2014a; Yao et al., 2014b). The basin has often been described as a reverse estuary, in which the salinity gradually increases to a maximum of 40.6 in the main basin with distance from the interchange with the Gulf of Aden at Bab el Mandab, which has been an area of interest for research for many years (Asfahani et al., 2020; Berumen et al., 2019b; Chaidez et al., 2017; Sofianos and Johns, 2015; Zarokanellos et al.). The Red Sea is also very warm, representing some of the highest known temperatures that coral reefs survive at (Chaidez et al., 2017; Ellis et al., 2019). With increasing global pressure on coral systems coming from multiple factors, developing a deeper understanding of the context in which coral reefs are surviving in the Red Sea can drive a better understanding of coral reefs in other parts of the world (Ellis et al., 2019; Hewitt et al., 2015; Raitsos et al., 2017). The recent trend in economic development within the Red Sea as a part of Saudi Arabian nation policy associated with Vision 2030 also indicates the need for a deeper understanding of the overall marine system (Almahasheer and Duarte, 2020).

This understanding comes when we can characterize both the short-term and long-term variability within the system and its effect on the underlying biological and physical processes (Jones and Kattan, 2017; Raitso et al., 2013; Thomson and Emery, 2014). Indeed, reaching a better characterization of system variability across multiple levels of scale is one of the goals of a typical marine observation system (Guinehut et al., 2012; Holte et al., 2017; McClatchie, 2016; Rudnick et al., 2017). While this characterization is easier to accomplish for areas that have a long running time series data set, recent oceanic questions require comparisons between areas of the oceans that have not had the benefit of a consistently maintained long-term observation system (Centurioni et al., 2017; Hernández-Carrasco et al., 2018; Jones and Kattan, 2017). There is an increased need to characterize short-term and long-term variability with relatively recently establish ocean observation systems (Borrione et al., 2019; Liblik et al., 2016; Pascual et al., 2017). At this time, the Red Sea is developing a consistent long-term time series data set focused on the central part of the basin. Included in the development of this data set is an integrated ocean observation system using satellite-derived data, ship-based measurements, High Frequency Radar, biogeochemical floats, and Autonomous Underwater Vehicles (gliders) (Jones and Kattan, 2017).

3.1.1: Context of eddies in the Red Sea

The importance of eddies in the Red Sea as a key feature of circulation and energy transfer have been understood for many years. Past studies have focused on describing many of the basic characteristics of Red Sea eddies by using several different methods, including numerical simulations, ship-based surveys, remotely sensed data, and gliders. Several conclusions have been drawn from this research. Mesoscale eddies occur

frequently within certain regions of the Red Sea (Karimova and Gade, 2014; Yao et al., 2014a; Yao et al., 2014b; Zhan et al., 2014; Zhan et al., 2016). Since these eddies tend to have centers within a relatively small range, they have been described as semi-permanent, with an eddy typically lasting from 2-6 weeks with an average diameter of 100-260 km (Karimova and Gade, 2014; Zhan et al., 2014). Figure 3.1 shows the calculated geostrophic velocities over the entire Red Sea for the month of January 2018, which shows some of the characteristic patterns of eddies within the basin. These velocities are from the Multi Observation Global Physical Near Real Time data product called ARMOR3D, available from Copernicus Marine Services (<https://marine.copernicus.eu/>). The ARMOR3D data product consist of observations collected both by satellite and *in situ* instrumentation to produce global coverage of oceanographic variables (Guinehut et al., 2012).

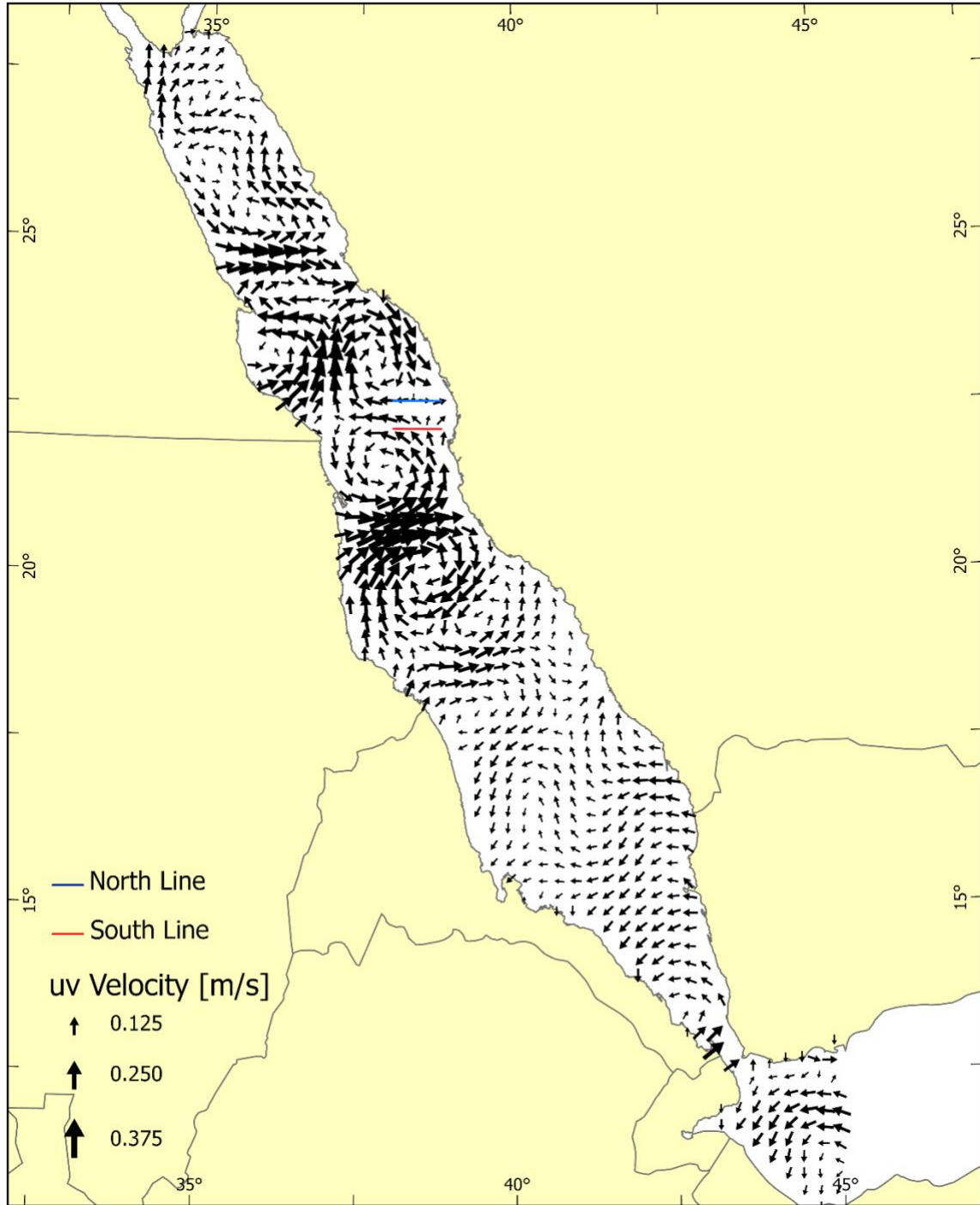


Figure 3.1: Surface geostrophic velocities of the entire Red Sea for January 2018 from the ARMOR3D data set. The size and direction of the arrow indicates the magnitude and direction of the u - v velocity vector. The alternating eddy structure often present within the Red Sea is visible with a series of counter rotating eddies across the central and northern parts of the basin.

One aspect of Red Sea eddies is that depending on the location of the eddy center within the basin, they can be deformed due either to limited basin width or the presence of another eddy in proximity (Karimova and Gade, 2014). Eddies in the Red Sea occur more frequently in the central and northern portions of the Red Sea at latitudes typically above 19° N (Karimova and Gade, 2014; Zarokanellos et al., 2017a; Zhan et al., 2014).

A second aspect of eddies within the Red Sea is that they contribute to a net flow throughout the entire basin over extended periods of time (Sofianos and Johns, 2003).

While this net flow may be obscured in at any given time, the net flow for the Red Sea typically starts at Bab el Mandab, where it continues along the western side of the basin.

The flow continues northward along the west side of the basin until it reaches anywhere from 17-20° N, depending on local wind intensity (Bower and Farrar, 2015). At this point, the net flow crosses the basin and proceeds northward along the eastern side until it reaches the north Red Sea, where evaporation and cooling increase the surface density of the water (Asfahani et al., 2020; Yao et al., 2014a; Yao et al., 2014b). While this general pattern is understood, at any given time this overall pattern can locally be either interrupted or in some cases reversed due to the location and rotation of eddies

(Zarokanellos, 2018). A third aspect of eddies within the Red Sea is their potential role in species distributions for the Red Sea, especially when considering coral,

phytoplankton, and coral reef fish. From a biodiversity standpoint, the Red Sea is

typically broken into either two, dependent on the species composition within the regions.

The southern-most region, located below 17-19° N has a species composition that is more similar to the Gulf of Aden than to the rest of the Red Sea (Berumen et al., 2019a;

Roberts et al., 2016; Wang et al., 2019). Numerical simulations have shown that drifting

particles originating from the Southern Red Sea are more likely to remain in the region or end up in the Gulf of Aden, with a low portion of particles drifting into the more northern portions of the Red Sea (Wang et al., 2019). In contrast, species in the Red Sea north of 19° N tend to remain within the northern part of the basin and the species composition is similar throughout the basin (Berumen et al., 2019a; Raitzos et al., 2017). It has been noted in previous research that these differences are potentially linked to different oceanographic processes between the two regions, but that more research into both oceanographic and species dynamics are needed before drawing any further conclusions (Berumen et al., 2019a).

Eddy dynamics also play an important role in the distribution of chlorophyll throughout the Red Sea. Based on chlorophyll productivity in the Red Sea, the basin is typically split into either three or four categories, typically arranged along the central axis (Gittings et al., 2019; Kheireddine et al., 2017; Raitzos et al., 2013). Chlorophyll values are highest in the southern Red Sea, where water from the Gulf of Aden advects into the Red Sea basin. Upwelling and mixing in the gulf contribute to a flux of nutrients into the euphotic zone increasing phytoplankton biomass (Dreano et al., 2016; Fratantoni et al., 2006; Gittings et al., 2017; Yao and Hoteit, 2015). As this water moves northward across the basin, the available nutrients gradually are utilized, and by the time the water from the Gulf of Aden has reached the central Red Sea, the overall productivity is greatly lowered (Gittings et al., 2019; Raitzos et al., 2013). Typically, the central Red Sea is considered as either one or two phytoplankton bioregions, with each half of the central region associated with its geographic neighbor (Kürten et al., 2019). In the northern part of the basin, phytoplankton concentration gradually increases due to an increase in nutrients in

the photic zone from mixing driven by evaporation and cooling (Asfahani et al., 2020; Kheireddine et al., 2017). Figure 3.2, also from the month of January 2018, is a composite satellite image for January 2018 of the Red Sea from the Ocean Colour Climate Change Initiative, with higher chlorophyll in the southern and northern parts of the basin, with some eddy structures seen in the central and northern parts of the basin (<https://marine.copernicus.eu/>).

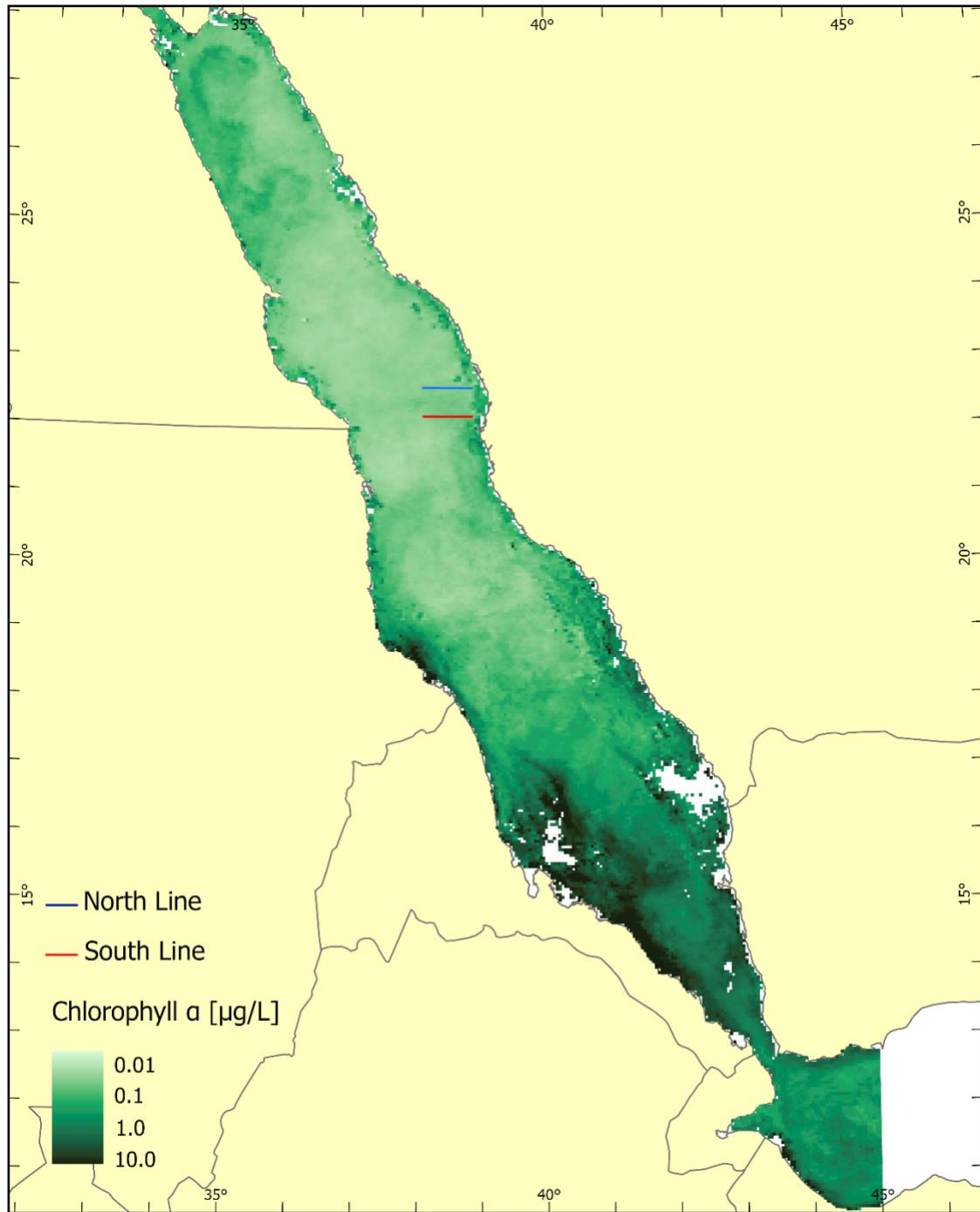


Figure 3.2: Satellite chlorophyll- α monthly composite for January 2018 using the OC-CCI algorithm. The color has been adjusted to a logarithmic scale to better highlight details across the entire Red Sea. A higher concentration of chlorophyll- α can be seen between the two glider lines, a possible effect of the convergent eddies.

3.1.2: Study Area

The study area for this research is a section of the north central Red Sea located at 22° N on the Saudi Arabian side of the Red Sea (Figure 3.3). Several factors led to the selection of this study area, including the presence of High Frequency Radar (HFR) sites, proximity to King Abdullah University of Science and Technology (KAUST), and the presence of semi-permanent eddies in the area. The study composes of two parallel observation lines spaced 40 km apart that run from about 20 km offshore to 100 km offshore. These two lines, located at 22.125° N for the South Line and 22.625° N for the North Line, were placed in a way to provide the safest, consistent coverage for glider data within the study area. The inshore portion of the line was placed within 5 km of the reefs located along the shoreline to ensure that the glider had sufficient space for navigation around waypoints. The offshore portion of each line was placed to within 5 km of the major shipping lanes running along the central axis of the Red Sea. The latitudinal spacing of 40 km was selected to ensure that the two glider lines would not overlap, since previous unpublished glider missions in the area have shown that local current conditions can cause drift of about 15 km from a given transect line of 80 km in length. The placement of these lines also coincides with data available from both HFR and the Global ARMOR3D L4 Reprocessed (ARMOR3D) data set, which are collocated either on or proximate to the glider lines (Guinehut et al., 2012; Mulet et al., 2012). Table 3.1 shows the deployment dates of the three gliders used to cover the two transect lines.

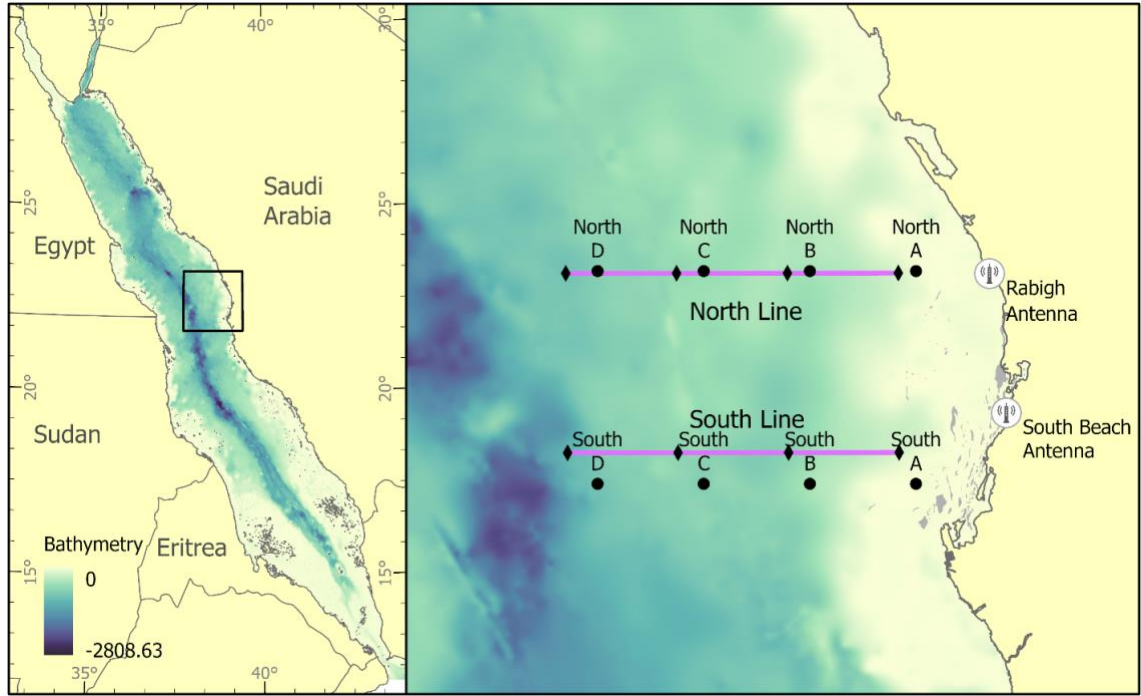


Figure 3.3: Location of the study area within the Red Sea. The two glider lines (North and South Line) is the solid line, the eight locations for the HFR data are the diamonds (North A-D, South A-D) and the circles are the locations used for the ARMOR 3D data. The glider and HFR data were analyzed from December 31, 2017 to January 31, 2018. The ARMOR 3D data is a four-week average centered on January 15, 2018.

Glider	Transect	Start Date	End Date
SG 213	North	12/25/2017	2/14/2018
SG 215	South	1/21/2018	2/25/2018
SG 216	South	11/13/2017	1/30/2018

Table 3.1: Glider deployment dates.

January was selected due to several aspects of Red Sea climatology. The winter months for the north central Red Sea typically show a decrease in the overall stratification typical of the summer months in this area (Yao et al., 2014b; Zarokanellos et al., 2017b). The winter months for this part of the basin show the greatest variability and depth of the mixed layer, and also have the greatest likelihood for multiple eddies to be close by (Yao et al., 2014a; Zarokanellos et al., 2017b). As seen in Figure 3.1, during this study period,

two counter-rotating eddies interface in the study area, with an anticyclonic eddy located to the north of the study area and a cyclonic eddy located to the southwest. The coldest seawater temperatures in this region typically occur in late January to mid-February.

3.1.3: Single glider line results

All ocean systems are highly dynamic and operate at several different key time scales of variability (Thomson and Emery, 2014). When multiple variables coming from a variety of sensors need to be compared, as is an increasingly important part of Ocean Observation Systems, it is important to understand the effect that sampling frequency can have on the data sets being examined (Borrione et al., 2019; Liblik et al., 2016; Pascual et al., 2017). Each variable can have a different natural time scale and mismatches between both temporal and spatial sampling frequency can confuse the complex interplay of oceanic systems and lead to potential mischaracterization of events (Holte et al., 2017; Rudnick, 2016; Terzic et al., 2019; Testor et al., 2019). One of the methods for resolving these potential mismatches is to identify the periodicity of each of the variables for each platform, and then to apply a temporal filter that averages out each variable to the same time scale (Thomson and Emery, 2014; Ward et al., 2018). One of the limitations to this approach is that this method works best with an extensive time series record for the area of interest. New areas of interest or recently placed instrument platforms do not immediately have access to a long time series to perform highly detailed temporal characterization, which limits the techniques that can be used to determine an appropriate temporal filter (Habeeb et al., 2005; Keeling et al., 1997; Ward et al., 2018). Currently, techniques are under development to provide appropriate temporal filters over smaller time frames until long-term, multiyear data sets can be collected. Recent work within the

study area has examined a method of cross comparison between different types of instrument platforms focusing on short-term persistence (Chapter 2). Two statistical methods were used to determine the short-term persistency of biological and physical variables in the north central Red Sea to find the dominant scale of variability for each variable. The two methods, Characteristic Length Scale and Integral Time Scale, found that the dominant scale of variability during the study period was 19.1 and 19.6 days (Chapter 2). This persistence time frame has also been described for other work in the area, where the highest magnitude of variation in the north central Red Sea was described as a weather band lasting two to four weeks, rather than seasonal or interannual variability (Churchill et al., 2018).

3.2: Methods

3.2.1: Glider deployment and calibration

Successful maintenance of the two glider lines over the time frame of the study period required the deployment of three gliders (Seaglider®). These deployments were handled in a similar manner as described in Chapter 2, with some slight adjustments due to multiple gliders in the water. These adjustments occurred as a variation to a longer-term goal to have a perpendicular glider transect in operation over the course of an entire year from September 2017 to September 2018. Partway through the month of January the glider that was on station for the transect line needed to have batteries replaced, which meant the addition of a third glider to the system. After the month-long multi-glider deployment, the remaining glider was left on station to continue observations. The additional gliders were always launched in proximity to a glider already in the water, and

a ship based CTD cast was used to compare the quality of the glider collected data. Each glider was equipped with a CTD, a dissolved oxygen sensor, a 3-wavelength fluorometer, and a 3-wavelength backscatter sensor (Table 3.2). The gliders traversed an 80 km east-west transect between a starting waypoint approximately 20 km offshore to an offshore waypoint approximately 100 km from the coast for both the north and south lines. The glider was set to dive to 500 m depth along the south line which represents a safe dive depth along the transect where the bottom depth varied from less than 600 m to more than 1000 m. The glider along the north line was set to dive to 450 m depth where the bottom depth varied from less than 500 m to more than 1000 m. The dive depth for both lines had an optimized horizontal resolution at about 2 km per dive with a dive interval of about 2.33 hours along the majority of the transect. Data from the gliders were collected from the entire dive cycle. However, the sampling frequency varied as a function of depth. The sampling frequency was once every 10 seconds in the upper 100 meters, then once every 50 seconds between 100 and 250 m, and was reduced to once every 100 seconds between 250 m and 500 m. The sampling frequency corresponds to data collected once every 1.3 m for the upper 100 m, once every 6.5 m from 100 to 250 m, and one every 13 m from 250 to 500 m. The data from the glider deployments can be seen in Figures 3.4 (North Line) and 3.5 (South Line). For this analysis three representative depths (6 m, 76m, and 150 m) and isopycnals (27.00 kg/m^3 , 27.50 kg/m^3 , and 28.00 kg/m^3) were chosen for analysis. Since the goal of this research is to contrast the spatio-temporal patterns of gliders at a relatively short latitudinal distance, the selected depth and isopycnal layers were selected to be in the upper portion of the water column extending into the upper portion of the subpycnocline layer. As a difference

from the representative depths and isopycnals as described in Chapter 2, the month of January shows an increase in the surface density of both the north and south line, with the 27.00 kg/m^3 isopycnal being the shallowest isopycnal that is present throughout the study period. Further justification for these representative layers can be found in Chapter 2, but these depths and isopycnals bracket the upper and lower extents of the water column experiencing greatest variability. All further processing and quality control for the glider data followed the same process as described in Chapter 2. Geostrophic velocities were calculated referenced to a depth 400 m for an assumed level of zero motion both in the north and south line to directly compare the geostrophic velocities across the two transects (McDougall et al., 2009). A 32-day mean of each of the two lines was used to calculate the dynamic height and then geostrophic velocities.

North Line

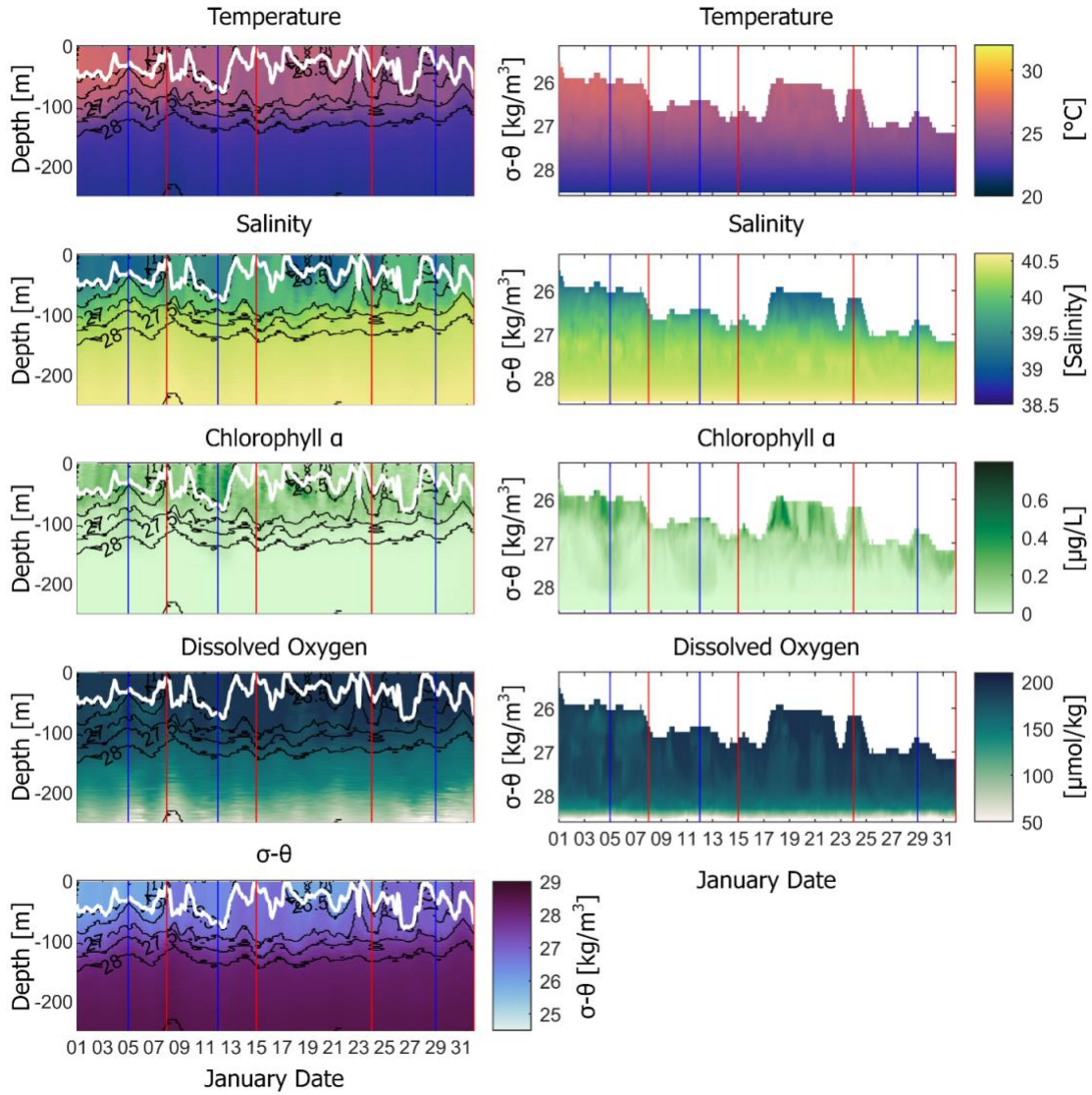


Figure 3.4: North Line glider time series plots for Temperature, Salinity, chlorophyll- α , dissolved oxygen, and $\sigma-\theta$ from 0-250 m and 25.5 – 28.5 kg/m³. The white line indicates the mixed layer depth based on the de Boyer index. The red vertical lines indicate when the glider is inshore, the blue lines indicate when the glider is offshore.

South Line

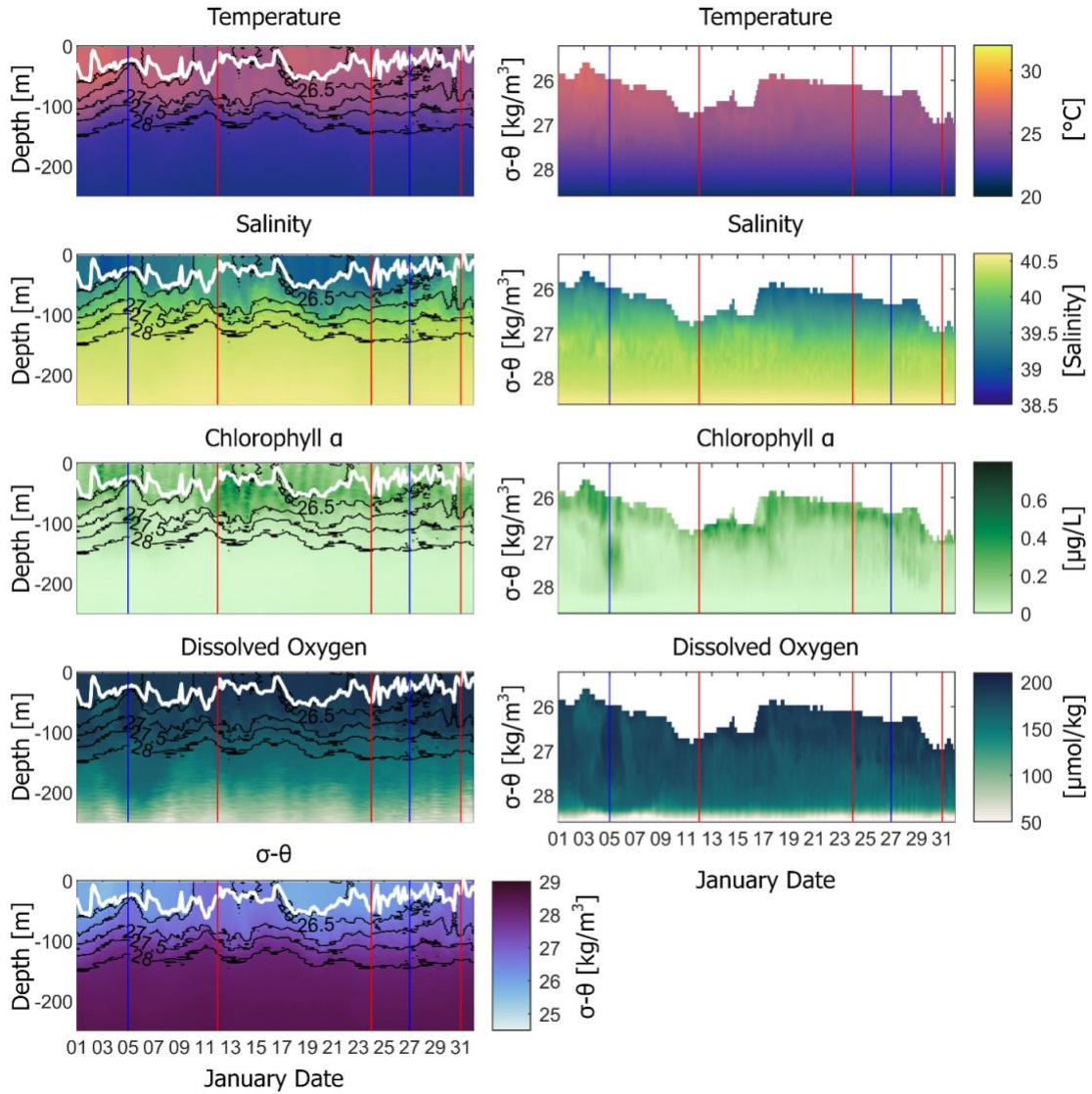


Figure 3.5: South Line glider time series plots for Temperature, Salinity, chlorophyll- α , dissolved oxygen, and $\sigma-\theta$ from 0-250 m and 25.5 – 28.5 kg/m³. The white line indicates the mixed layer depth based on the de Boyer index. The red vertical lines indicate when the glider is inshore, the blue lines indicate when the glider is offshore.

Equipment	Parameters	Excitation [nm]	Emission [nm]
CTD (Seabird CT-sail, unpumped)	Temperature Conductivity		
WET Labs	Chlorophyll- α	470	695
ECO Puck	CDOM	370	460
(FL3)	Phycocyanin	630	680
WET Labs	Optical	532	532
ECO Puck	backscatter at 3	650	650
(BB3)	Wavelengths	880	880
Oxygen sensor (Aanderaa optode, model 4331)	Dissolved Oxygen		

Table 3.2: Summary of the equipment used in this study.

3.2.2: HFR calibration

The HFR system is located on the central western coast of Saudi Arabian and consists of two CODAR Seasonde sites in operation since July 2017 (Solabarrieta et al., in review).

The HFR system transmits at the 16.12 MHz frequency with hourly data provided over a 120 km range with a spatial resolution of 3 km. The received backscatter signal was converted into radial velocities using the Multiple Signal Classification algorithm (Schmidt, 1986). The MATLAB package HF Progs

(<https://cencalarchive.org/~cocpmb/COCMPwiki>) was then used to combine radial

currents and generate gap-filled total, two dimensional currents using Open Modal Analysis (Kaplan and Lekien, 2007) for the entire month of January 2018. Four evenly spaced locations across both the north and south glider line were selected for time series analysis, as seen in Figure 3.3.

3.2.3: Short term persistence through Error X analysis

Deployment conditions for oceanographic instrument platforms are typically far from ideal, which leads to serious complications when trying to compare multiple types of data. Due to the dynamic nature of seawater, even slight differences in positioning and sampling frequency can cause a high degree of spatio-temporal blurring of recorded events (Bouffard et al., 2012; Rudnick, 2016). To simplify comparisons between data sets, a typical practice is to apply a spatio-temporal filter to the data set of interest, which characterizes and simplifies variability that occurs within the data set (Thomson and Emery, 2014). Ideally, an appropriate filter can be established detailed time series analysis, where details such as diel patterns, seasonal shifts, and interannual variability can all be accounted for. In the absence of long time series data sets, other methods can be employed to provide a spatio-temporal filter for data comparison. In the case of this research, Error X analysis was performed as described in (Ward et al., 2018). In order to determine the correct filter to compare the glider and HFR data from both the north and south lines, Error X analysis was performed on the glider time series at the three representative depths and isopycnals and the HFR time series at each of the eight locations as described in Chapter 2. Figure 3.6-3.8 show each of the individual time series used for the Error X analysis and Figure 3.9 shows the time series data from the HFR. The full results from the CLS (with 95% confidence intervals) analysis can be

found in Appendix B. Figure 3.10 shows the Characteristic Length Scale for each of the glider variables after the analysis. The north line has a mean CLS of 22.2 days (std 3.46, median 20) and the south line has a mean CLS of 25.2 days (std 3.6, median 26). Of note for this analysis is that some of the variables listed do not actually reach a characteristic length scale within the time frame of the study. This is especially apparent with the HFR data, in which none of the variables achieved a CLS for the month of January. For the variables that did achieve a measurable CLS, the length is close to previously reported values (Chapter 2) (Churchill et al., 2018) for the weather band period of the Red Sea.

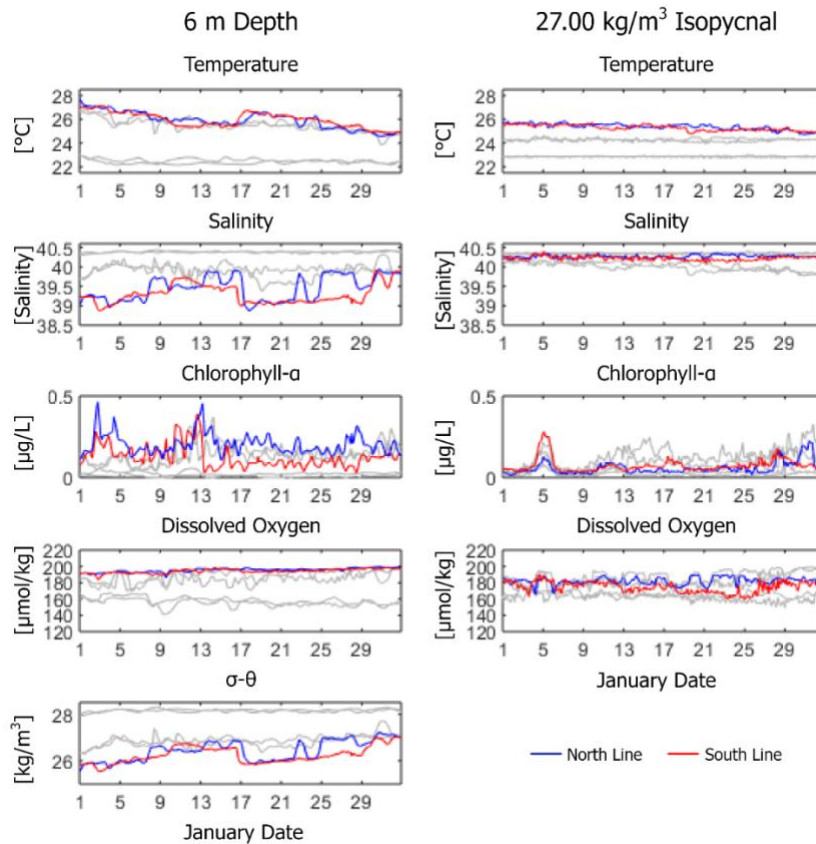


Figure 3.6: North and South Line glider time series data set at 6 m and 27.00 kg/m³ isopycnal for the month of January 2018 used in the Characteristic Length Scale calculation. The light gray shows the values from 76, 150 m and 27.50, 28.00 kg/m³

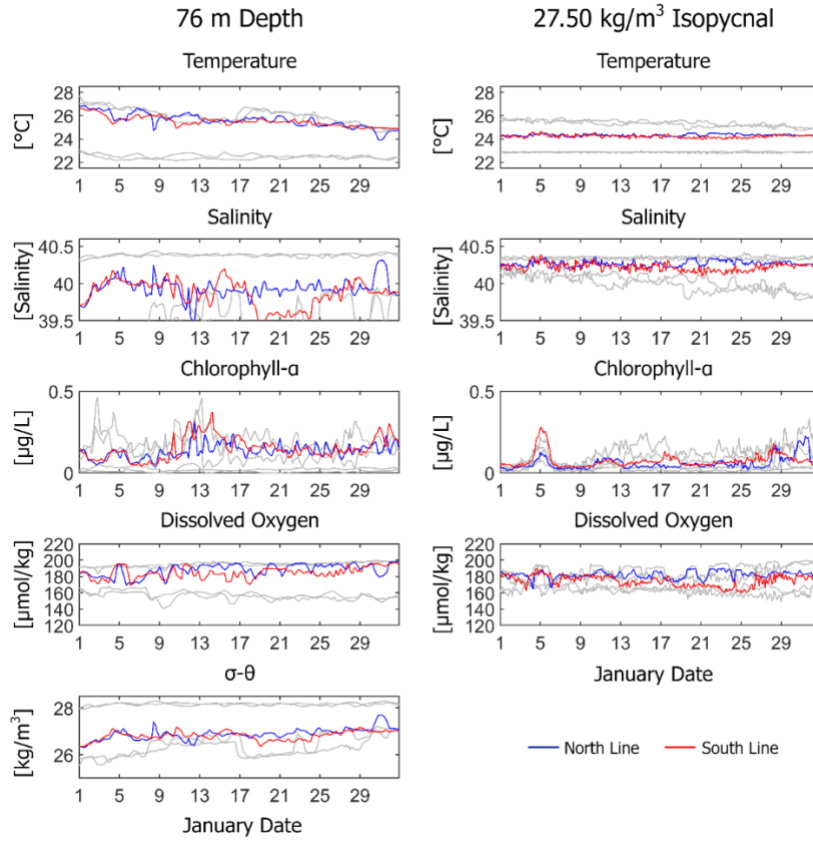


Figure 3.7: North and South Line glider time series data set at 76m and 27.50 isopycnal for the month of January 2018. These data sets were used in the Characteristic Length Scale calculation. The light gray shows the values from 6 m and 27.00 kg/m³ as well as 150 m and 28.00 kg/m³ isopycnal. Salinity has been rescaled from Figure 3.4.

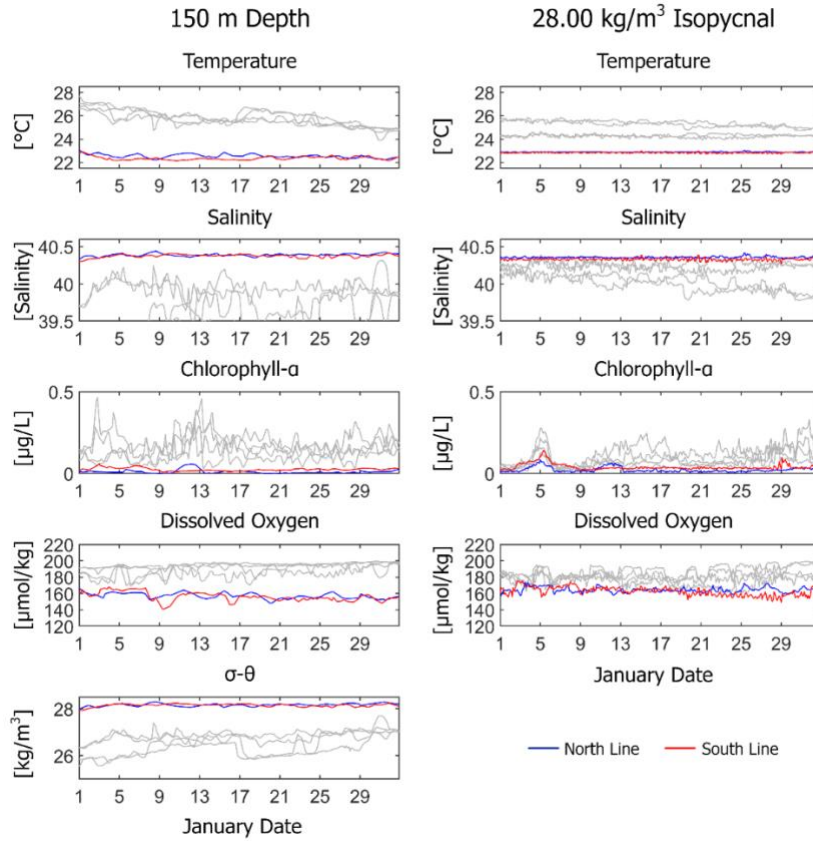


Figure 3.8: North and South Line glider time series data set at 150m and 28.00 isopycnal for the month of January 2018. These data sets were used in the Characteristic Length Scale calculation. The light gray shows the values from 6 m and 27.00 kg/m³ as well as 76 m and 27.50 kg/m³ isopycnal. Salinity has been rescaled from Figure 3.4.

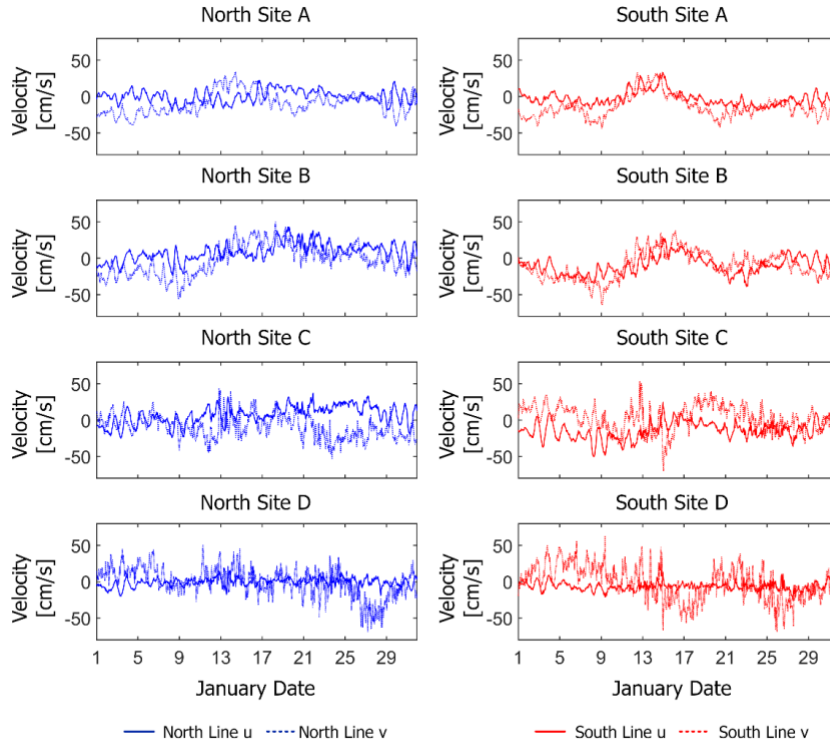


Figure 3.9: HFR time series data across the study area for the month of January 2018. These data sets were used in the Characteristic Length Scale calculation. Over the time frame of the study period, the CLS calculation did not reach an asymptote for any of the HFR variables measured.

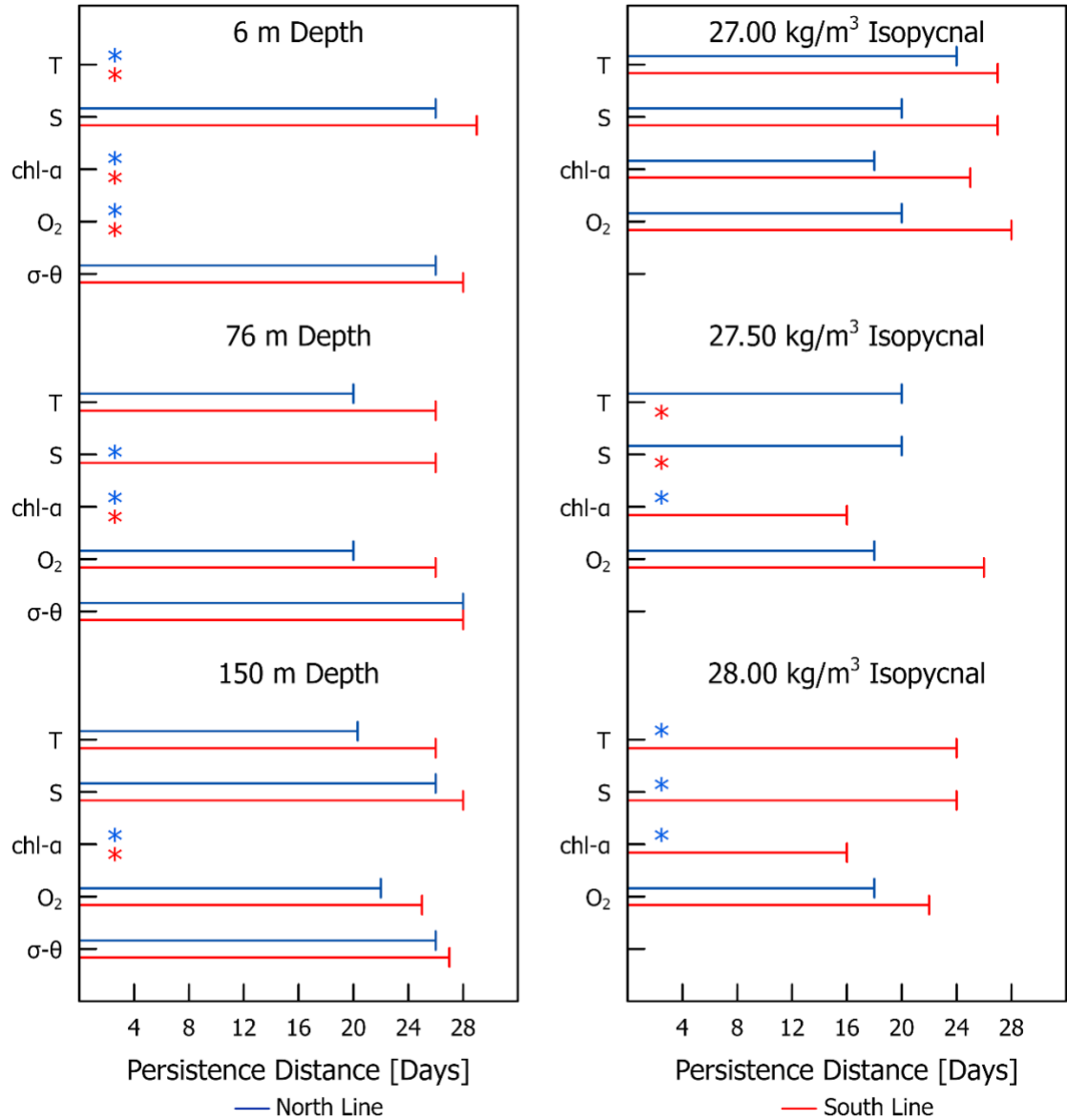


Figure 3.10: Short term persistence of glider data from Characteristic Length Scale analysis. Variables labelled with an asterisk did not reach an asymptote. The north line has a mean CLS of 22.2 days (std 3.5) while the south line has a mean CLS of 25.2 days (std 3.6). These time frames fall within the expected 2-4-week weather band period that has previously been described for the study area.

3.3: Results:

3.3.1: Water column time series plots (glider)

Both the north and south glider line show clean signs of winter cooling and increased salinity over the month of January (Figure 3.4-3.8). Both cooling and an increase in salinity contribute to the overall increase in density of water across all depths, especially since both glider lines exhibit a patch of less saline water at the surface between the 17-21 of January, which decreases the density of water at the surface to nearly the same values as at the start of the month (Figure 3.4, 3.5). This patch of lower salinity water occupies the offshore portion of the glider line and could be related to the transport of water from the Gulf of Aden. The mixed layer depth for the northern line varies between 20 m and 80 m (Figure 3.4). One of the times that the mixed layer depth reaches 80 m is on January 12, associated with a patch of low salinity water in the upper water column. The other time the mixed layer depth deepens to 80 m occurs on January 27th and is associated with a shallowing of the 27.00 kg/m³ isopycnal. The mixed layer depth for the southern line shows less variability overall, ranging from 20 to 60 m over the month. The maximum depths for the southern line occur on January 11th and the 18th-22nd (Figure 3.5). The mixed layer depth of January 11th occurs between two patches of lower salinity water, and the event on the 18th-22nd closely follows the 26.00 kg/m³ isopycnal. In general, the northern line is more dense than the southern line, as is expected as a part of the general latitudinal trends in the Red Sea, in which latitude has been shown to correlate positively with salinity and negatively with temperature (Chaidez et al., 2017; Yao et al., 2014a; Yao et al., 2014b). Also as expected, the water column at or below 150

m tends to show little variation over the course of the month, indicative of relatively stable conditions at depth.

3.3.2: Geostrophic Velocities from glider relative to 400 m

As with many aspects of ocean observation systems, specific instrument platforms and data sets are best designed to handle specific questions. While the instantaneous state of the ocean is important for many types of question, some questions are best understood over a longer time frame, including information about geostrophic velocities. The CLS analysis performed indicates the length of an appropriate averaging filter, in this case the mean of the glider transects for the north and south line over the month of January. The geostrophic velocities were calculated in spatial bins of 0.08° longitude. Figure 3.11 shows the calculated geostrophic velocities relative to an assumed depth of zero motion at 400 m perpendicular to the glider transects. The north velocity field shows 2 flow regimes, with the first one mainly southward (negative v) west from 38.55° E, with a mainly northward flow (positive v) east of this location. At the same time, the southern line shows a more complex pattern, with apparent banding throughout the entire transect. In comparing the two, the north line shows a slightly lower magnitude of velocity overall. Both transects show a southward flow from $38.17 - 38.27^\circ$ E and a northward flow from $38.55-38.67^\circ$ E, indicating the potential for some features of the geostrophic velocity to transmit across the 40 km distance between the two transects. The geostrophic velocities from individual transects were not considered due to each transect lasting from 3-4 days, which may include tidal and inertial motion that are not a part of this analysis.

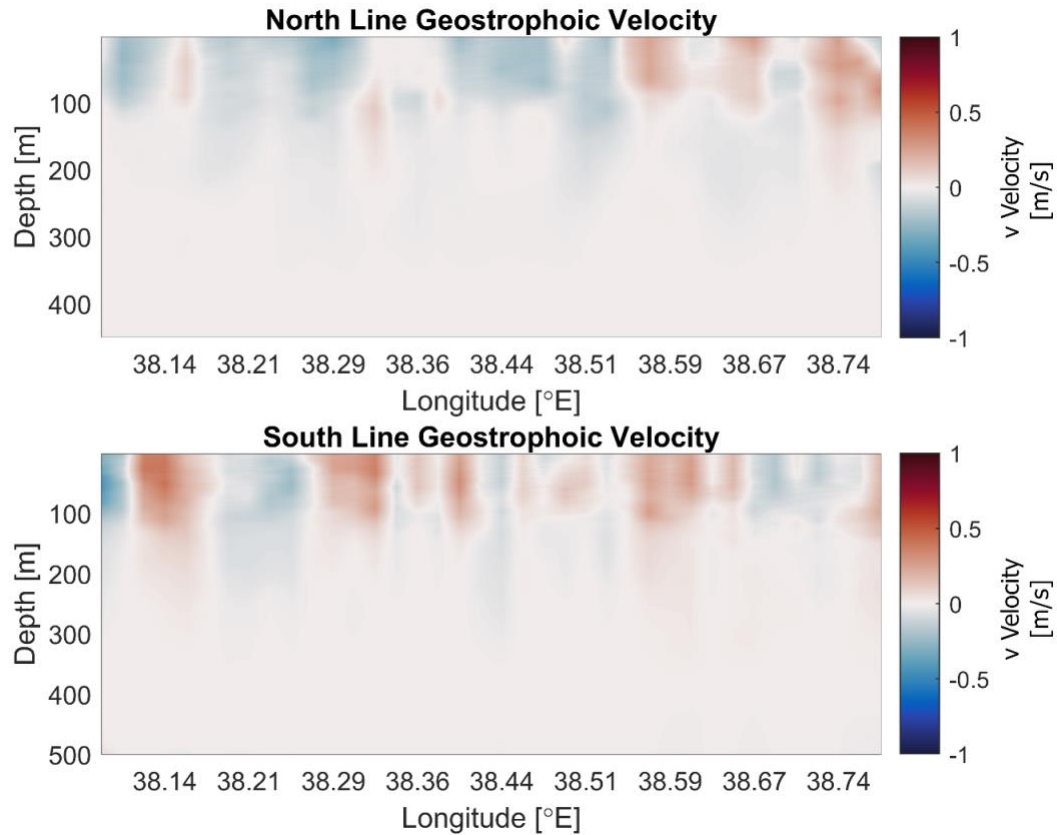


Figure 3.11: Geostrophic velocities for averaged transects for January from the North and South glider lines. The geostrophic velocities were calculated for a reference depth of 400 m using the TEOS-10 toolbox for Matlab. Positive velocities are flow towards the north, and the upper 100 m of the water column shows a potential convergence in the offshore region of the two lines.

From Figure 3.11, it appears that the overall geostrophic flows between the north and south transects are roughly balanced, however, Figure 3.12 shows a more detailed view of the mean values of the geostrophic velocity for each transect. The overall means for the two transect lines extending across the entire transect line and down to 100 m in the water column shows a difference between the two mean geostrophic flows. The north transect has a mean velocity of 0.0337 m/s flowing southward, while the south transect has a mean velocity of 0.0416 m/s flowing northward. The two transect lines also have

different minimum and maximum ranges with the north transect between -0.32 and 0.33 m/s and the south transect between -0.45 and 0.42 m/s. Both glider transects show a gradual decrease in the velocity magnitude as depth increases across the entire transects. In contrast, when each latitudinal bin is examined the patterns follow along with the results seen in Figure 3.11, with the difference in magnitudes easier to visualize. Overall, the picture presented from the mean structure of the geostrophic velocities indicates a northward flow on the southern transect and a southward flow on the northern transect, indicative of convergence of flows between the two glider lines. This convergence should set up a corresponding offshore flow between the two glider transects, since the proximity of reefs on the inshore side of the transects would prevent flow in that direction. However, the nature of geostrophic velocity calculations and the setup of these transects only provide velocity information perpendicular to the transects lines, which means that they are only able to provide the north-south component of the geostrophic velocities. To further resolve the flow conditions, additional data are needed for the study area.

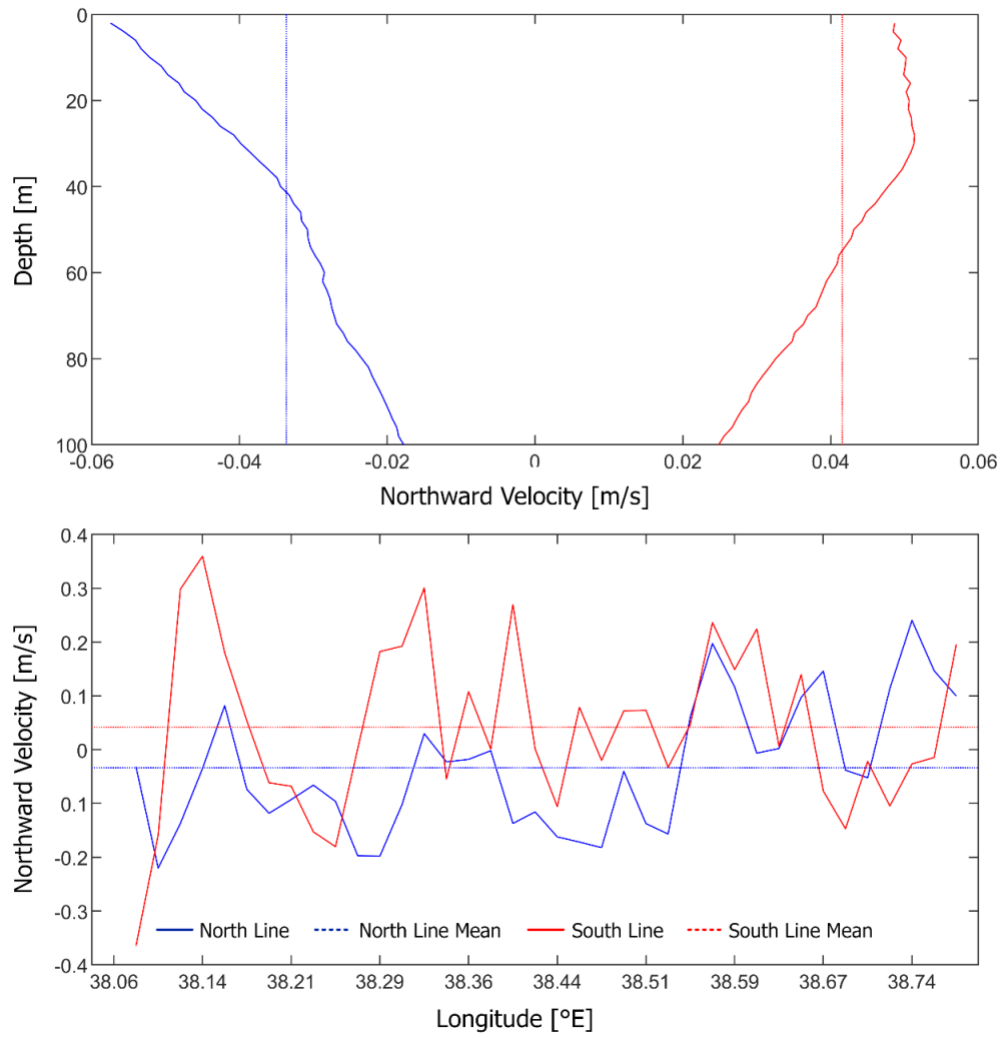


Figure 3.12: Average geostrophic velocities from the north and south glider line. The top is the average velocity for each two-meter depth bin down to 100 m across each entire line. The bottom is the average velocity from 0-100 m for each latitudinal bin. For both parts of the figure the blue and red dashed lines is the overall mean across each of the glider lines.

3.3.3: Comparison between section glider GSV and ARMOR3D GSV at north and south line

Since the data provided from the glider transects are only able to provide geostrophic velocities perpendicular to the transect, they are insufficient to fully resolve the geostrophic velocities within the study area. Additional data are needed to confirm the presence and directionality of the potential convergent flow seen from the glider data and the geostrophic velocities provided from the ARMOR3D data set helps to offset some of the limitations from a single data source. However, before any comparisons can be made, it is important to verify that the glider and ARMOR3D data sets do match up, and to what degree they correspond in both space and time. The glider transects were summarized to a monthly value due to the CLS analysis. To match this temporal scale from the glider data, a monthly average was obtained for the ARMOR3D data set for the entire Red Sea (Figure 3.1). The spatial mismatch between the two datasets, since the data from ARMOR3D has a spatial separation of 25.75 km for the study area, was accounted for by creating four spatial bins of equal size for each glider transect. The geostrophic velocities for each quarter transect were then averaged and compared to the corresponding v vector ARMOR3D data, as seen in Figures 3.13 and 3.14. The differences between the geostrophic velocities from ARMOR3D and glider data can be seen in Table 3.3. The greatest difference in velocity between the two data sets occurred at the surface on the North Line at Site C, with a difference of 0.1169 m/s. The smallest difference occurred at 150 m on the North Line at Site C, with a difference of 0.0004 m/s. The results from the glider geostrophic velocities match reasonably well with the v vector values from ARMOR3D at the surface, 75 m, and 150 m depths. Figures 3.15-3.17 show

a combination of the geostrophic velocities from the gliders and ARMOR3D, with both the v vectors and combined u, v vectors displayed. With the addition of the ARMOR3D data set, we can conclude that the convergence seen in the glider data does contribute to an offshore flow between the two transect lines, likely resulting from the cyclonic/anticyclonic eddies seen in Figure 3.1.

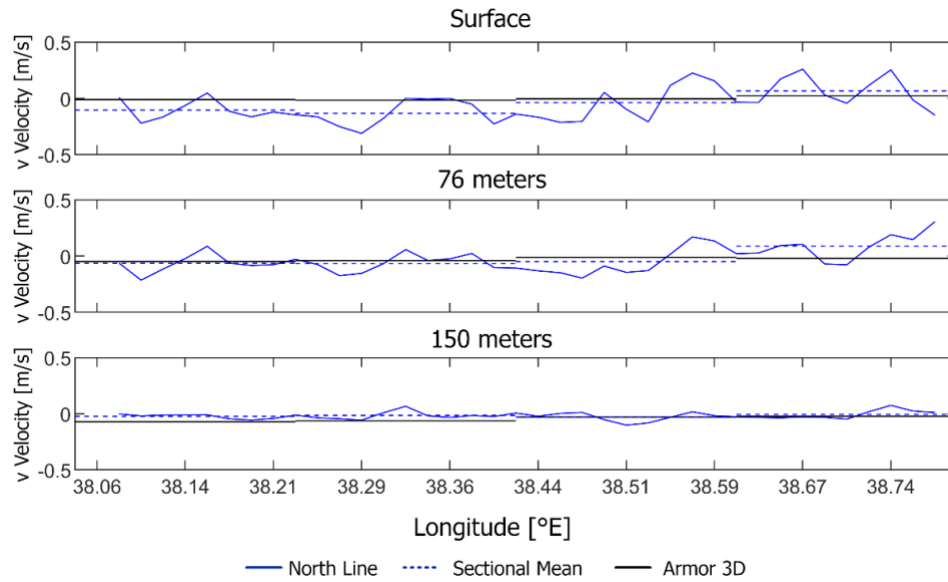


Figure 3.13: Geostrophic velocities for the north line at the representative depths used for CLS analysis with the v vector of the ARMOR3D data set. As an aid to compare to the ARMOR3D data set, the glider line was split into four roughly equal sections that correspond to the geographic coordinates to find the mean value of the geostrophic velocity for that region. While the average values are close, there is a degree of spatial variability that is lost due to the averaging.

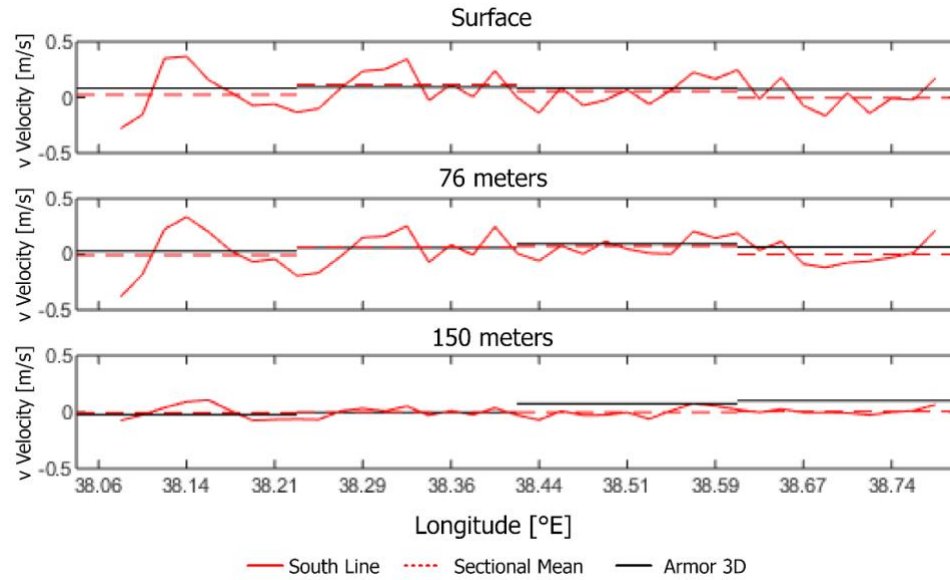


Figure 3.14: The same as Figure 3.13, but for the South Line.

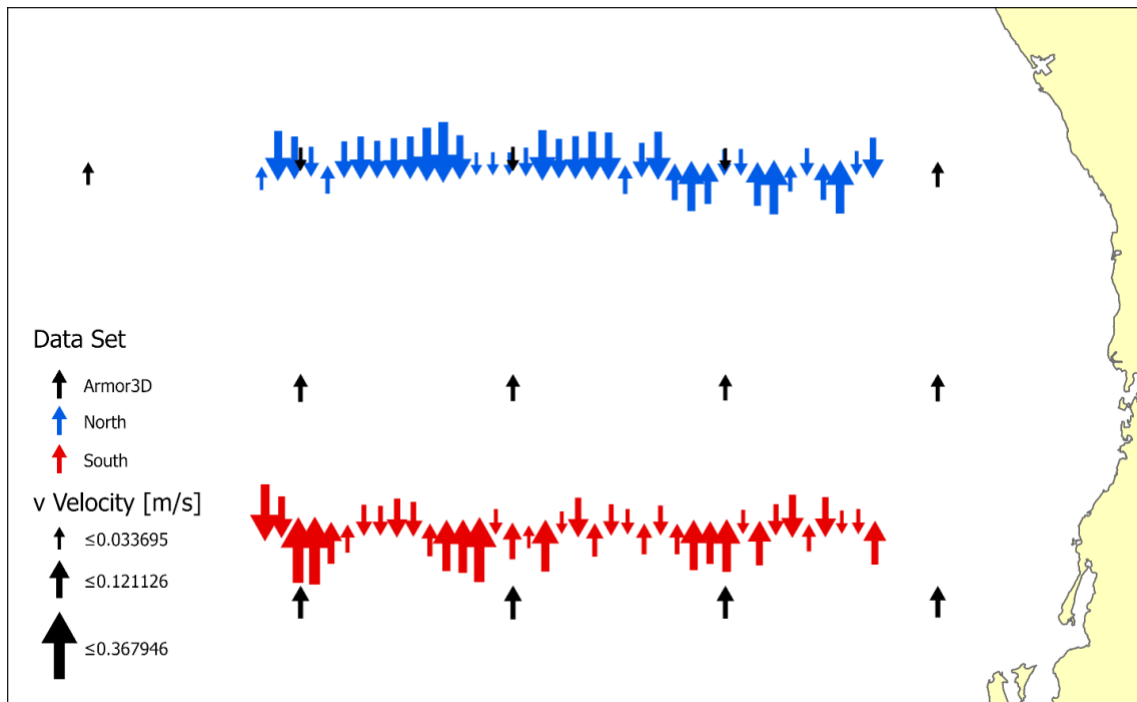


Figure 3.15: Geostrophic velocities for both the North and South Line at the surface (6m) and v vector velocities from ARMOR3D across the study area.

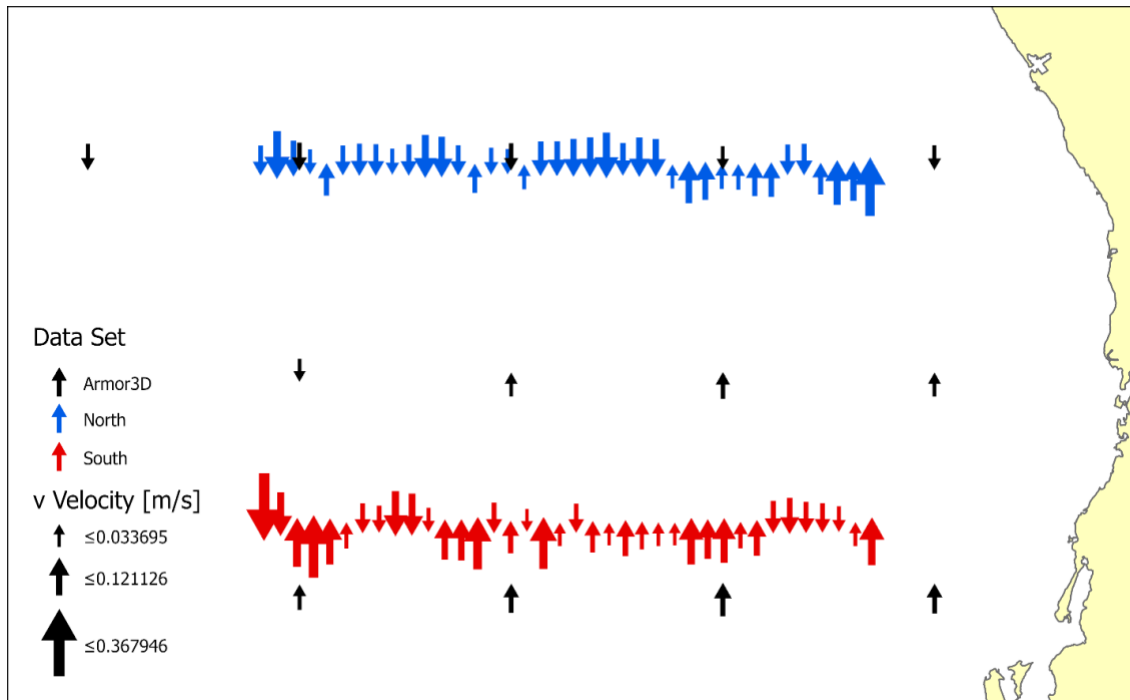


Figure 3.16: As in Figure 3.15, but at 75 m depth.

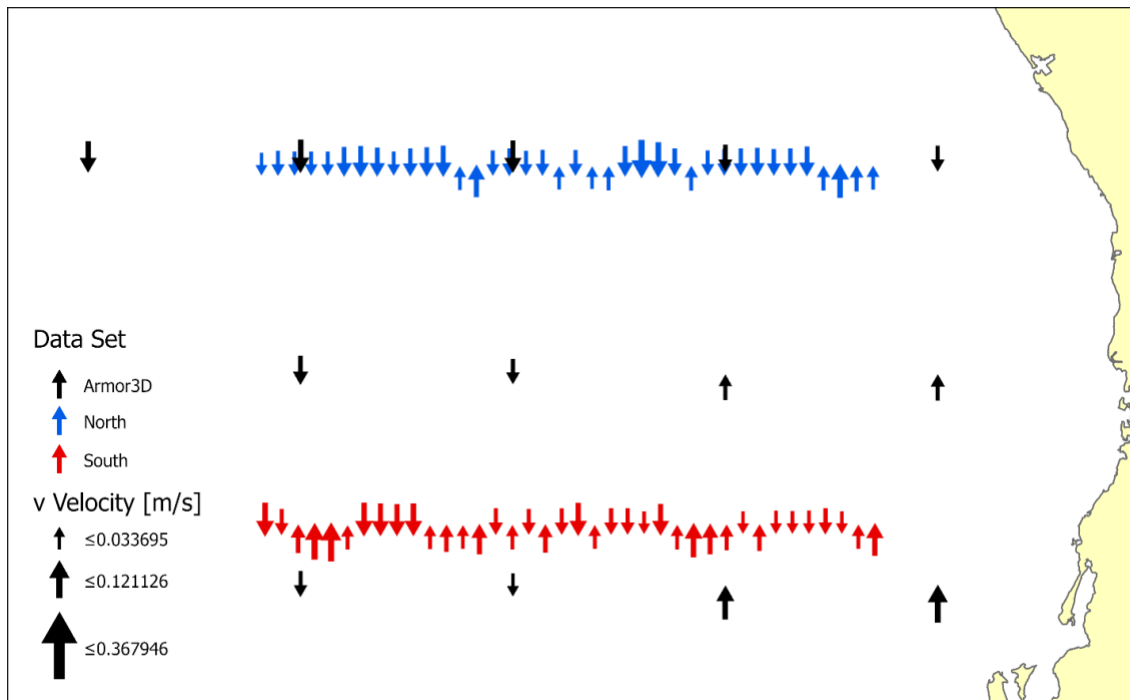


Figure 3.17: As in Figures 3.15 and 3.16, but at 150 m.

Depth	Line		Site D	Site C	Site B	Site A
Surface	North	ARMOR3D	-0.011	-0.016	-0.004	0.022
		Glider	-0.1044	-0.1329	-0.0375	0.0654
		Difference	0.0934	0.1169	0.0335	0.0434
	South	ARMOR3D	0.051	0.093	0.085	0.073
		Glider	0.024	0.1144	0.0558	-0.0041
		Difference	-0.027	0.0214	-0.0292	-0.0771
75 m	North	ARMOR3D	-0.049	-0.041	-0.012	-0.022
		Glider	-0.0651	-0.0669	-0.0502	0.0881
		Difference	0.0161	0.0259	0.0382	-0.1101
	South	ARMOR3D	0.028	0.06	0.095	0.065
		Glider	-0.0078	0.0646	0.0739	0.0008
		Difference	0.0358	0.0046	-0.0211	-0.0642
150 m	North	ARMOR3D	-0.072	-0.063	-0.03	-0.023
		Glider	-0.0235	-0.0152	-0.0304	-0.0059
		Difference	-0.0485	-0.0478	0.0004	0.0171
	South	ARMOR3D	-0.023	-0.004	0.075	0.1
		Glider	-0.006	0.0009	-0.0011	0.0067
		Difference	-0.017	-0.0049	0.0761	0.0933

Table 3.3: V velocity results. Difference in v velocities between the ARMOR3D values and the mean values calculated from the quartered glider transects. Velocity units are in m/s.

3.4: Discussion

3.4.1: Meaning for the Red Sea

The combination of the presented glider data and ARMOR3D results provides information regarding variability dynamics for the north central Red Sea. First, we are able to provide a brief description of the variability along the central axis of the Red Sea. We were able to observe over the course of the month two distinct circulation patterns affecting the northern transect line and the southern transect line, primarily due to the influence of two counter-rotating eddies near the study area. This indicates the importance of setting up an adaptive sampling system when trying to resolve local scale

events within the basin. During the study period, the two glider transects showed little interaction from one transect to the next. However, it is possible that this is due primarily to the counter-rotating eddies, and it is expected that conditions would be different if one or both eddies were not present. With this knowledge, a more efficient sampling system could be established, where gliders are moved closer together during eddy events and further away during non-eddy events.

Eddy dynamics are an important factor to understanding the overall processes governing the Red Sea. Eddies are directly involved at many different scales across the Red Sea, from providing some of the structure of the overall circulation, to cross-basin transport, and down to providing contrast at a distance of a few tens of kilometers. Due to the variety of scales important in the Red Sea, it is vital to have analytical techniques available that can compare across multiple levels of scale. As seen from this research, a significant level of detail that can be lost in the transition from 2 km data to 25.75 km data, as seen from the transition from the geostrophic velocities calculated from the gliders and from ARMOR3D. The geostrophic velocities present from the glider data demonstrate a much higher level of variability, including reversals in the overall mean flow for a given segment. For certain fine-scale processes, this loss of variability may be acceptable, but for others it may not. It is through the combination of multi-scale data that we can start to seriously answer questions about the scale that we need to sample at to collect the correct type of data. However, once these disparate data sets can be accurately compared, we are able to better understand the interactions between the coarse and fine scale processes occurring.

As presented in this research, the integration of multiple data sets has led to a better understanding of current conditions in the north central Red Sea with just a brief time series data set. Using the CLS analysis has provided an initial understanding of important temporal scales within the basin, matching temporal scales with previous research in the area. As data collection expands in spatio-temporal extent across the Red Sea, this analysis can be used to help determine longer relevant temporal and spatial scales for the Red Sea, extending into characterization of seasonal and interannual patterns.

3.4.2: Conclusions

Outside of the Red Sea, this research can help address several important issues. Spatio-temporal patterns are highly influenced by local conditions, with characteristic length scales developed at one location rarely applicable outside of the context of the immediate study area, let alone to geographically distant areas. Often conditions between adjacent coastal and open water areas are sufficiently distinct to complicate understanding the connections that exist between the two regions. The need to determine these connections even at a local scale, let alone for the global system, has led to the development of new techniques. This is actually a cyclical aspect of oceanographic research, where more complicated questions lead to the development of new instrument platforms and deployment strategies, and the increase in data quantity and quality then results in the development of new analytical techniques as well as the improvement of established methods. Novel methods to incorporate numerical simulations, satellite data, ship-based observations and *in situ* instruments are increasingly used to develop a better understanding of complex marine systems. Even with few direct connections between

geographically distant areas of the global marine system, the oceans are all interconnected, so increased understanding in any given area of the ocean provides at least in part to a better understanding of global patterns.

Chapter 4: Sea surface temperature and eddy center occurrence cluster analysis in the Red Sea

The work presented in this chapter regarding the Permutation Distribution Clustering analysis was performed by Nabila Bounceur.

4.1: Introduction

The Red Sea is one of the major shipping passages in the world, providing a connection between the Mediterranean Sea and the Indian Ocean from a marine standpoint and a connection between Africa and Asia from a terrestrial standpoint (Figure 4.1). In addition to socio-economics reasons to study the Red Sea, the basin provides a relatively simple study area when compared to other basins of similar size or ecological importance (Berumen et al., 2019b; Jones and Kattan, 2017; Langodan et al., 2014). One aspect of the simplicity of the basin is the existence of only one major source of inflow across a relatively shallow sill (Sofianos and Johns, 2015; Sofianos and Johns, 2003). This sill has allowed for detailed measurements of both the inflow and outflow of the basin, providing information on the overall energy budget of the Red Sea (Zhan et al., 2016). Another important aspect of the Red Sea is the existence of coral reefs at higher latitude, temperature, and salinity than elsewhere in the world (Berumen et al., 2019a). These two aspects of the Red Sea have led to an interest in understanding circulation within the basin through remote sensing, *in situ* observations, and numerical simulations.

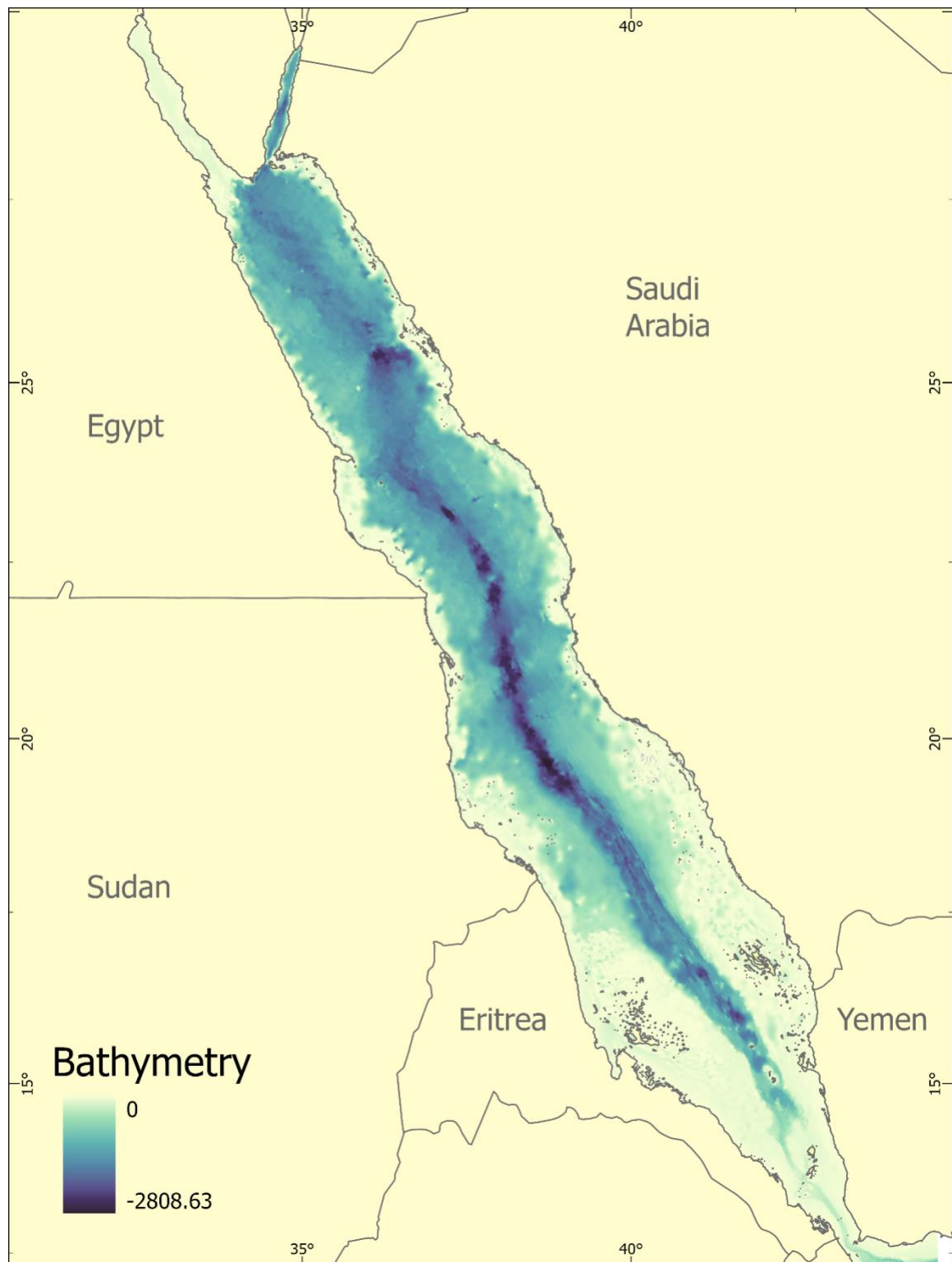


Figure 4.1: Bathymetric map of the Red Sea using the GEBCO worldwide digital elevation model.

4.1.1: Physical effects of Red Sea eddies

Circulation in the Red Sea has been described at two major characteristic scales. The first level that has been described is the basin-wide circulation pattern. Overall, the Red Sea can be considered as a reverse estuary system, with salinity gradually increasing with increasing distance from Bab el Mandab (Sofianos and Johns, 2015; Yao et al., 2014a; Yao et al., 2014b). This increase in salinity, along with the latitudinal temperature gradient, helps set up the thermohaline circulation in the basin. In addition to thermohaline circulation, the overall wind regime also influences basin-wide circulation (Bower and Farrar, 2015). The Red Sea has two major wind patterns that have been described. The first is a continuous flow from the north across the entire basin, typically present from May to September, during the southwest monsoon over the Arabian Sea. The second major wind pattern is a reversal of the winds over the southern portion of the Red Sea due to the influence of the Indian Ocean northeast monsoon over the Gulf of Aden and Indian Ocean which occurs during the months of October to April (Bower et al., 2013; Sofianos and Johns, 2015; Yao et al., 2014a; Yao et al., 2014b). These combined forces of thermohaline circulation and an alternating wind regime are the major contributors to the long-term boundary current in the Red Sea (Sofianos and Johns, 2007; Yao et al., 2014a; Yao et al., 2014b). In this current, water flows into the Red Sea from Bab el Mandab, where modelling indicates it initially flows along the western coast of the basin. At about 18-20 °N, this flow then crosses over to the eastern side of the basin (Wang et al., 2019; Zarokanellos et al., 2017a; Zhai et al., 2015). The water then continues along the eastern side of the basin until reaching the northern Red Sea, where

the formation of Red Sea Intermediate Water occurs (Yao et al., 2014a; Yao et al., 2014b). This water, now subsurface, then flows back to the Gulf of Aden through Bab el Mandab. In contrast to the this flow system described by modelling, *in situ* observations have indicated that at least some of the water from the Gulf of Aden is transported along the eastern side of the basin as well, causing cooling in inshore waters (Churchill et al., 2014a; DeCarlo et al., 2020).

While this first major characteristic scale has been well described, recent research has been focused on the importance of mesoscale eddies in Red Sea circulation. One of the key descriptive features noted about the Red Sea is the occurrence of eddies, which have been described as either permanent, semi-permanent, recurrent, or semi-persistent (Jones and Kattan, 2017; Karimova and Gade, 2014; Yao et al., 2014a; Yao et al., 2014b; Zhan et al., 2014). Regardless of the exact term used, research has highlighted that eddies occur very often and in specific locations in the Red Sea. These eddies occur across the Red Sea at sufficiently high frequency to obscure the long-term mean boundary current pattern previously described (Sofianos and Johns, 2003; Sofianos and Johns, 2007; Zarokanellos et al., 2017b; Zhan et al., 2014). A recent study of eddies in the Red Sea, derived from AVISO+ Sea Level Anomaly (SLA) data from 1992-2012, described the characteristics of close to 1,000 separate mesoscale eddy tracks for the basin over the course of the 20 years (Zhan et al., 2014). These eddies, including both cyclonic and anticyclonic eddies, have an average diameter range of 70-200 km and typically lasted for 6 weeks. This study showed that the rotation of the eddy is highly dependent on the location of the eddy within the basin. Additional research has also shown that when active, these mesoscale eddies can account for kinetic energy transfer up to ten times

greater than that of mean flow (Zhan et al., 2014). These eddies have a direct effect on the localized variability of salinity, temperature, and biogeochemical particles in the Red Sea, indicating a strong link between SST conditions and eddy location (Yao et al., 2014a; Yao et al., 2014b; Zarokanellos et al., 2017a; Zhan et al., 2014). Additionally, recently described scales of temporal variability are linked closely to the typical duration of Red Sea eddies, lasting on the order of 2-4 weeks as described in the previous chapters (Churchill et al., 2018). Recent *in situ* observations of Red Sea eddies from both ship-based observations and autonomous underwater vehicles have provided a high level of detail at depth describing the interactions between mesoscale eddies and salinity, temperature, and chlorophyll- α and these interactions can extend to a depth of 150 m (Asfahani et al., 2020; Gittings et al., 2019; Kürten et al., 2019; Zarokanellos et al., 2017a).

The physical structure of the Red Sea has also been suggested as an important factor for the recurrent positioning of eddies (Karimova and Gade, 2014; Zhan et al., 2014). In part, this is due to the nature of the land around the basin, which has been discussed as an influencing factor for the consistent wind across the northern part of the basin (Bower and Farrar, 2015; Zhai and Bower, 2013). Also, given that a significant portion of the described mesoscale eddies are considered cross-basin due to their diameter, recognition exists of topographical constraints to the formation and propagation of eddies in the Red Sea (Yao et al., 2014a; Yao et al., 2014b). This research seeks to understand the spatial interconnectedness of bathymetry, SST, and eddy formation within the Red Sea, and to compare the results to previous research in the basin.

4.2: Methods:

4.2.1: AVISO+ NRT 3.0

The AVISO+ eddy tracking product has long been used as a primary source of information about eddy occurrence for many years (Faghmous et al., 2015). The recent addition of new sea surface altimetry data due to new satellites in orbit has allowed for the addition of a new eddy tracking product with data starting from 1 January 2018, the Near Real Time (NRT) 3.0 eddy tracking product. This new version of the AVISO+ eddy tracking experimental product includes several processing steps to produce a global record of daily eddy occurrence. The NRT 3.0 eddy tracking product starts by incorporating all available satellite altimetry to produce a gridded global absolute dynamic topography (ADT) (Duacs/AVISO+, 2017). This global data set then uses the algorithms described in Mason et al. (Mason et al., 2014) to identify eddies and track their progression through time. In the eddy detection stage, a low-pass filter is first applied to the full dataset to remove any large-scale variability for processes greater than 300 km, outside of the mesoscale range (Duacs/AVISO+, 2017; Mason et al., 2014). This filter is then subtracted from the original gridded data to highlight only the mesoscale variability. The individual eddies are then identified by finding closed contours of -100 - +100 cm according to criteria established (Mason et al., 2014). A best-fit circle of this closed contour is then created, and the center of the circle is identified as the eddy center. These eddies are then tracked by comparing gridded data from subsequent days in an increasing search radius derived from latitude over the course of 4 following days (Duacs/AVISO+, 2017). If the eddy is lost between the initial

identification and day four, then an interpolation is applied to account for any missing days. In the case that a portion of the projections may cross over a land boundary, the two affected eddies are separated into different tracks. The final step is to filter out any eddies that have a duration less than 14 days (Duacs/AVISO+, 2017). This new product does have some limitations. The data set analyzed for this research includes 684 days of eddy activity, from 1 January 2018 to 12 November 2019. An additional limitation is the NRT product is not directly comparable to the delayed time product, due primarily to differences between eddy tracking from ADT and SLA. The eddy center occurrences can be seen in Figure 4.2.

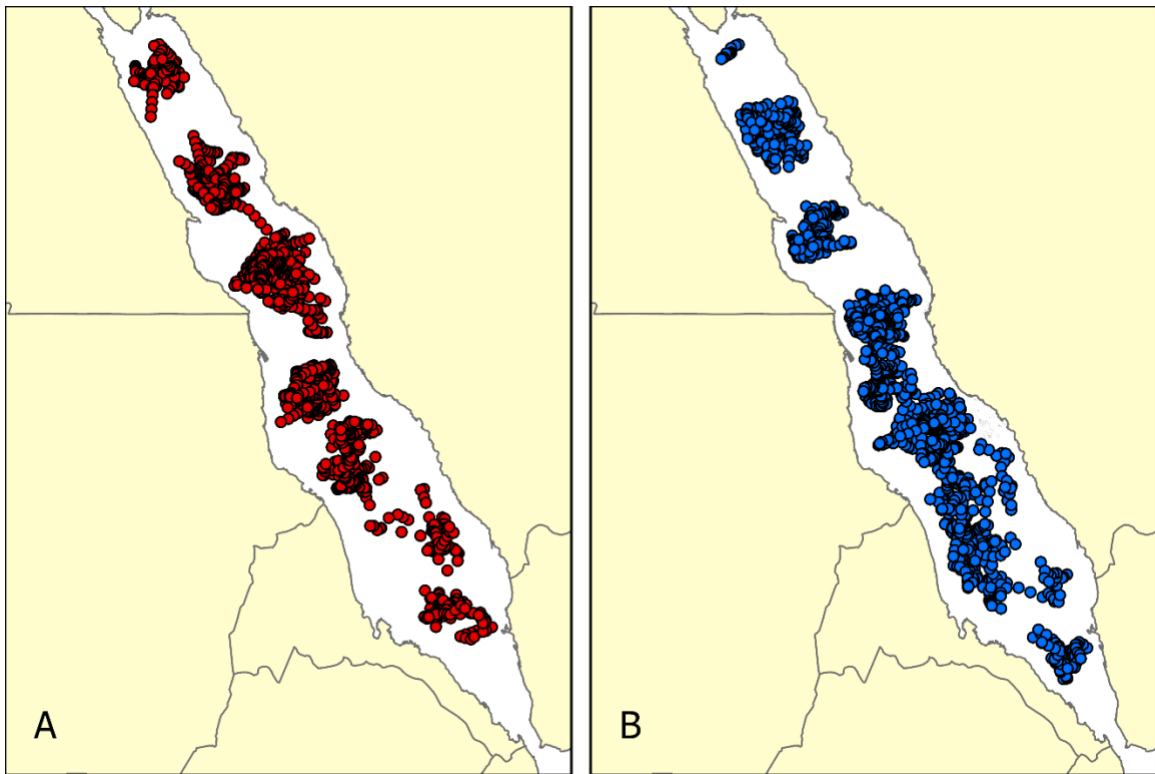


Figure 4.2: Anticyclonic (A) and cyclonic (B) daily eddy center occurrences from 1 January 2018 to 12 November 2019. Eddy centers are from the AVISO+ NRT 3.0 data set. 1623 cyclonic and 1724 anticyclonic eddy occurrence days were recorded, for a total of 3347 eddy occurrence days

4.2.2: Spatial processing and statistics

Several types of spatial analysis were performed on the NRT eddy tracking data, which included multiple steps. The first step was the creation of incremental contours for the entire basin of the Red Sea using bathymetry downloaded from the GEBCO global data set (Weatherall et al., 2015). These contours were created at 50 m increments from 0-500 m using ArcGIS Pro 2.6, as seen in Figure 4.3. A central axis for the Red Sea was created running from the Sinai Peninsula to Bab el Mandab. The midpoints for these two lines were then calculated and connected to create the central axis. Next, the central axis was separated into 100 equal length segments of approximately 21 km each. A set of 100 equidistance bins were created perpendicular to the central axis of the Red Sea to provide a regular distance from Bab el Mandab. Cross-basin distances for each bin were calculated. A count was collected for every day an eddy center (cyclonic and anticyclonic) occurred in the bin and the mean diameter calculated. The count of cyclonic and anticyclonic eddy occurrence day per bin can be seen in Figure 4.4. This created a data set of six variables, distance from Bab el Mandab, cross-basin distance, cyclonic eddy count, cyclonic eddy mean radius, anticyclonic eddy count, and anticyclonic eddy mean radius for each of the 50 m depth contours. As depth decreased for the Red Sea, the area covered by each cross-basin bin shrank, with some of the bins in the southern portion of the Red Sea either greatly shortened or not existing. The data extracted from each depth interval was then analyzed using a normalized Principle Component Analysis (PCA) within the software R.

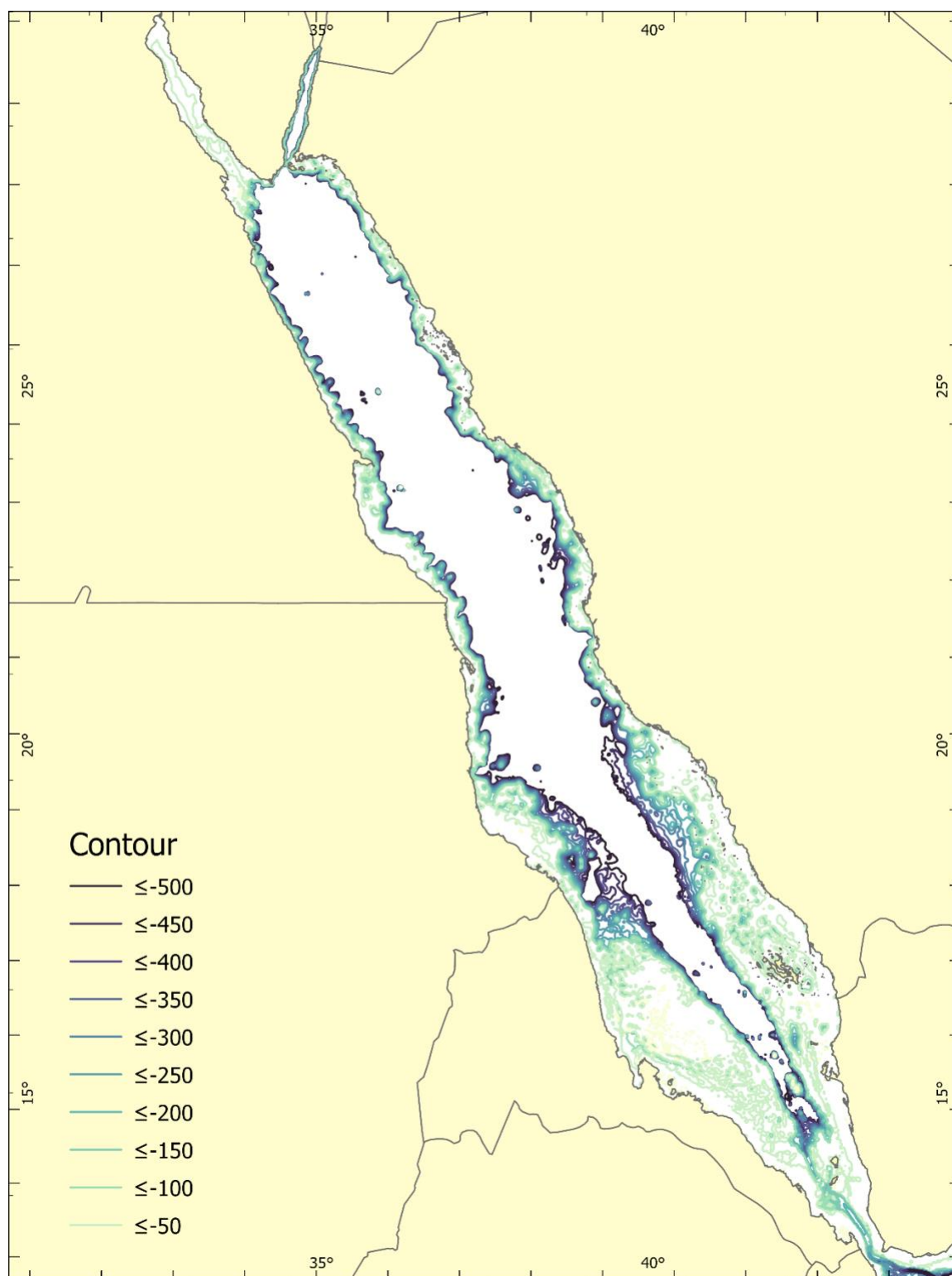


Figure 4.3: 50 m contour lines derived from the GEBCO elevation data for the entire Red Sea, starting from the surface and terminating at 500 m.

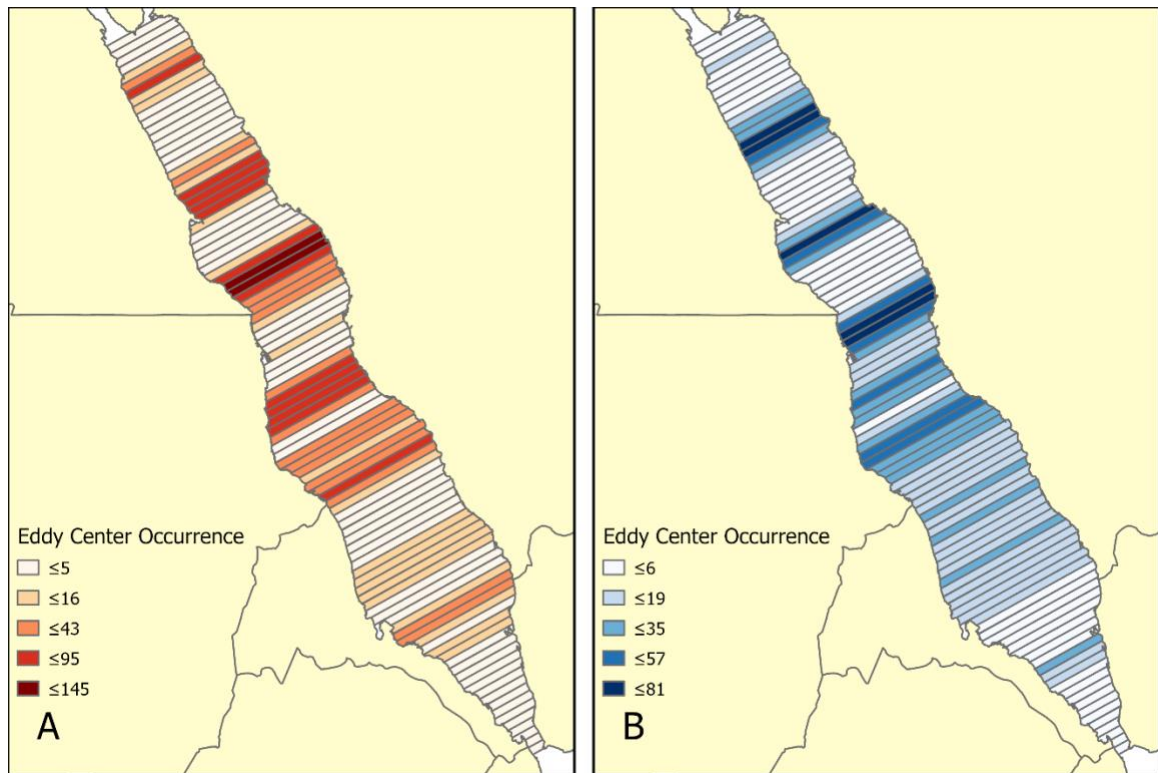


Figure 4.4: Number of daily eddy center occurrences for each of the 100 cross-basin bins for anticyclonic (A) and cyclonic (B) eddies.

A separate cluster analysis was performed on the eddy center occurrences using the Optimized Hot Spot Analysis (Getis-Ord G_i^*) tool within ArcGIS Pro 2.6 (Getis and Ord, 1992). Hot spot analysis is a way to extract information in addition to localized clustering within a data set, while at the same time testing for global clustering. The analysis proportionally compares the sum of a feature and its neighbors to the sum of all the features within the data set, and then compares the difference to a random chance distribution. This analysis provides not only on where a given data set is clustered, but if high or low values are clustered together (Getis and Ord, 1992). In the case of occurrence data, high values indicate areas where the data is significantly clustered, while

low values will indicate areas where data does exist, but that it is significantly dispersed. The optimized version of hot spot analysis examines multiple resolution sizes to find a best fit for the existing data.

4.2.3: Permutation Distribution Clustering of SST

SST data are observations from the National Oceanic and Atmospheric Administration (NOAA) Optimum Interpolation (OI) SST data set (Noted OI.v2 NOAA OI SST V2) (Banzon et al., 2016; Reynolds et al., 2007) for 38 years over the period 1981-2019. This data set is high resolution using optimum interpolation (OI), *in situ* (i.e. ships and buoys), and satellite observations combined on a daily and $1/4^\circ$ (0.25° latitude by 0.25° longitude) spatial grid to form a spatially complete SST field of a global 1440 by 720 grid. Permutation Distribution Clustering (PDC) is a complexity-based approach to clustering time series (Brandmaier, 2011). The SST subregions are identified by applying a PDC to weekly SST anomalies where the mean climatology has been removed through.

This approach combines a hierarchical clustering based on the complete agglomerative scheme to merge the time series into distinct clusters (Wilks, 2011), a permutation test as a proxy of the difference in time series complexity, and an alpha-divergence criteria as a dissimilarity measure. Using the divergence between the permutation distribution (Amari and Nagaoka, 2007) leads to a measure of dissimilarity based on the relative complexity of the time series. Alpha divergence is chosen as it is a generalization of a set of common dissimilarity measure, therefore theoretically motivated by the kullback leibler divergence (Kullback and Leibler, 1951). It has many advantages including invariance to

monotonic transformations and linear time complexity to the time-series length. The permutation distribution is obtained by counting the frequency of distinct order patterns in an embedded dimension of the original time series. The number of embedding dimensions is determined automatically by generalized information criteria based on a minimum entropy heuristic which is parameter free.

4.3: Results

Within the Red Sea during this almost two-year study period a total of 1623 cyclonic and 1724 anticyclonic eddy occurrence days, for a total of 3347 eddy occurrence days (with 684 days) were detected. Figure 4.2 indicates the center of each of the individual eddy occurrence days, separated by rotation. This includes a total of 76 individual cyclonic eddies and 65 anticyclonic eddies. The cyclonic eddies reported had a minimum duration of 10 days, with a maximum duration of 90 days, and a mean duration of 21.6 days and a maximum of 5 distinct eddies occurring in the basin on a single day. The minimum recorded diameter for cyclonic eddies was 59.5 km, with a maximum of 308.6 km and a mean diameter of 133.9 km. Anticyclonic eddies show slightly different information, with a minimum duration of 10 days, a maximum duration of 118 days, and a mean duration of 26.8 days, with a maximum of 6 concurrent eddies on a single day. The minimum and maximum diameters were slightly smaller than with the cyclonic eddies, at 55.6 and 291.5 km respectively; however, the mean diameter was slightly larger at 138.6 km. These numbers vary from those previously reported with a 180 km mean, but this may be due to the previous omission of eddies with a diameter smaller than 70 km due to potential misidentification and the removal of eddies lasting less than 14 days (Zhan et

al., 2014). Smaller diameter eddies tend to have a shorter life span, so the difference in diameters and duration may be due to these spatio-temporal filters.

4.3.1: Description of Principal Components

A PCA was performed for the six variables at each depth separately due to the variations in the total number of cross-basin bins available, ranging from 100 bins at the surface down to 85 bins by the 500 m contour. The variance explained by the first two principle components is visually shown in Figure 4.5 with the full results shown in Table 4.1. The strongest first component occurred at the surface (40.20% variance) with the two strongest response variables being cross-basin distance and the diameter of anticyclonic eddies. The combined variance explained for the first two components at the surface is 65.62%, an additional 25.42% explained by the second component. The two strongest response variables for the second component had the occurrence of anticyclonic eddies with a positive relationship and mean cyclonic eddy diameter as the strongest negative relationship. The 150 m depth peak in the PCA matches with previously described vertical structuring of eddies in the Red Sea as well as from currently unpublished eddies in the Red Sea, where 150 m is a typical effective depth (Asfahani et al., 2020; Bower and Farrar, 2015; Quadfasel and Baudner, 1993; Zarokanellos, 2018).

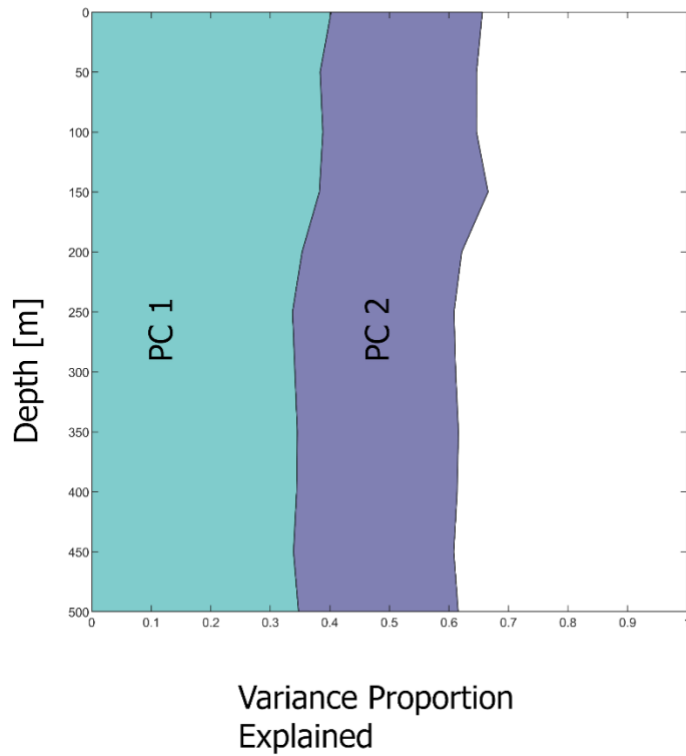


Figure 4.5: Total variance described by the first two principle components for each of the 50 m depth bins. The highest explained variance for the first principle component occurred at the surface, with 40.20% of variance explained. The highest explained variance for the first two components occurs at 150 m depth, with 66.58% of the variance explained.

Depth	PC 1 Standard Deviation	PC 1 Proportion	PC 1 Cumulative	Distance Bin	Cross Basin Distance	Cyclonic Eddy Count	Cyclonic Eddy Mean Radius	Anticyclonic Eddy Count	Anticyclonic Eddy Mean Radius	PC 2 Proportion	PC 2 Cumulative	Distance Bin	Cross Basin Distance	Cyclonic Eddy Count	Cyclonic Eddy Mean Radius	Anticyclonic Eddy Count	Anticyclonic Eddy Mean Radius
0	1.5530	0.4020	0.4020	0.1439	0.5759	0.3405	0.5004	0.1172	0.3394	0.2228	0.5172	1.2349	1.2557	0.2628	0.6465	0.3352	0.0683
50	1.5172	0.3836	0.3836	0.4360	0.6160	0.2980	0.4205	0.3394	0.6465	0.2228	0.5172	1.2349	1.2557	0.2628	0.6465	0.3352	0.0683
100	1.5261	0.3882	0.3882	0.4343	0.6153	0.3592	0.2650	0.2867	0.6465	0.2582	0.5109	1.2446	1.2446	0.2582	0.6465	0.3352	0.0683
150	1.5145	0.3823	0.3823	0.4625	0.6141	0.2781	0.1750	0.3894	0.6658	0.2835	0.6658	0.0425	0.1149	0.5646	0.5279	0.5143	0.3382
200	1.4559	0.3533	0.3533	0.4658	0.6454	0.3397	0.1390	0.3774	0.6216	0.2683	0.6216	0.1673	0.0107	0.5394	0.6295	0.5281	0.3762
250	1.4223	0.3372	0.3372	0.5078	0.6550	0.2711	-0.0178	0.3860	0.6086	0.2715	0.6086	-0.0866	0.1158	0.6000	0.6446	0.4501	0.0249
300	1.4111	0.3413	0.3413	0.5312	0.6548	0.2723	-0.0423	0.3814	0.6113	0.2700	0.6113	-0.0672	0.1177	0.6075	0.6404	0.4488	0.0327
350	1.4396	0.3454	0.3454	0.5165	0.6607	0.3440	0.1103	0.3231	0.6161	0.2707	0.6161	-0.1740	-0.0093	0.5507	0.6451	0.5001	0.2486
400	1.4371	0.3442	0.3442	0.5254	0.6601	0.3689	0.1370	0.2833	0.6138	0.2695	0.6138	-0.2221	-0.0473	0.5267	0.6469	0.5016	0.0305
450	1.4259	0.3389	0.3389	0.5768	0.6696	0.3444	0.0642	0.2830	0.6082	0.2694	0.6082	-0.1836	-0.0144	0.5527	0.6660	0.4566	0.0925
500	1.4437	0.3474	0.3474	0.5936	0.6622	0.3338	0.0343	0.2868	0.6156	0.2682	0.6156	-0.1422	0.0104	0.5629	0.6571	0.4792	0.1197

Table 4.1: Result of PCA for bathymetry and eddy count information. Of particular note from this table is that the ranking for the variables at the surface is different from all other depths, leading to potentially different dynamics in the expression of variability.

The strongest combined score for the first two principle components occurred at the 150 m depth contour, explaining a combined variance of 66.58%. The strongest responses from first component (38.42%) showed a positive association with cross-basin distance and distance from Bab el Mandab, while the second component (28.35%) had a positive association with occurrence of cyclonic eddies and mean radius of cyclonic eddies.

Across all depths excluding the surface, the first principle component ranges from 33.89 – 38.23%, and consists of the variables cross-basin distance and distance from Bab el Mandab.

4.3.2: Getis-Ord Gi* clustering

The hot spot analysis shows not only that clustering occurs for both the cyclonic and anticyclonic eddies, but also defines some of the spatial characteristics of the clustering. Since hotspot analysis compares local or neighborhood clustering with global values, one of the first terms that the analysis provides is an optimized neighborhood size for the data set. For the two eddy categories the size of the local neighborhood is similar at 15.93 by 15.93 km for cyclonic eddies and 14.93 by 14.93 km for anticyclonic eddies. Figure 4.6 shows the location of each of these eddy occurrence neighborhood cells and the level of clustering or dispersion of eddy centers in that neighborhood in any clustering or dispersion is occurring. Empty boxes on Figure 4.6 represent areas where eddies have occurred, but there is no significant clustering or dispersion, and eddy center placement can be explained by random factors. Gray scale colors for both eddy rotations indicate an increasing amount for dispersion with darker gray. Neighborhood cells with either a more intense blue (cyclonic) or red (anticyclonic) represent areas with an increasingly significant degree of clustering. For the cyclonic eddies, Figure 4.6 shows four cluster

areas within the Red Sea, located within the northern and central sections of the basin. Anticyclonic eddies also show four cluster areas, mostly located in the northern and central regions, but one cluster area is located at the northern extreme of the southern Red Sea. Figure 4.6 also demonstrates the highly alternating nature of the eddy center clustering. This confirms the alternating eddy pattern for the Red Sea over the course of the nearly two-year data set. With first a cyclonic eddy cluster in the northern Red Sea, followed by an anticyclonic cluster. The rotation of the eddy center clusters continues to alternate through the northern and central Red Sea. In the southern portion of the basin there is almost no clustering or dispersion, with most of the cluster neighborhoods falling into the random placement category.

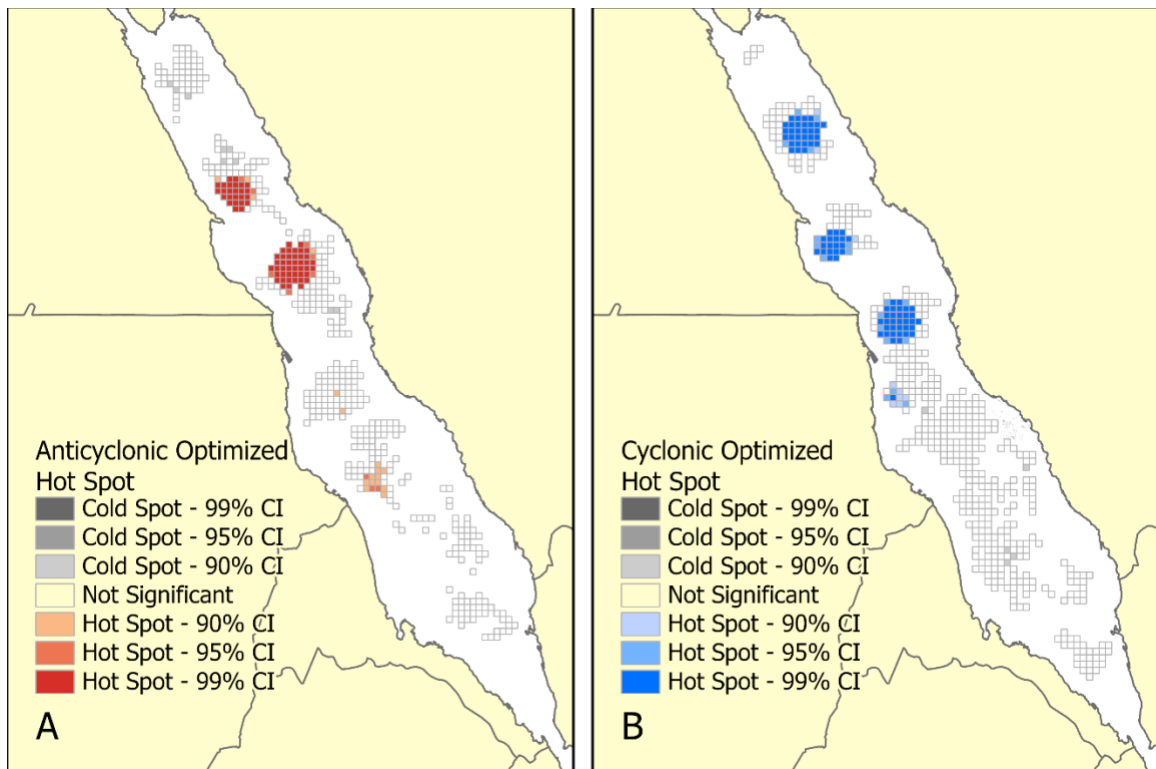


Figure 4.6: Optimized Getis-Ord G_i^* Hot Spot Analysis for anticyclonic (A) and cyclonic (B) eddies in the Red Sea. Of note is the alternating rotation of daily eddy occurrence clusters in the northern Red Sea.

4.3.3: Machine Learning clustering

Patterns relative to the dynamic of sea surface temperature variability uncovered in Bounceur and Knio (in preparation) by using a statistical clustering approach on the long-term weekly SST time series, match with a high degree of correspondence the spatial distribution of the eddy centers *ergo* demonstrating a correspondence between mesoscale eddy dynamics and SST variability over the Red Sea.

According to the minimum entropy heuristic, we identified six interconnected subregions in SST which displayed distinct climate variability over the Red Sea (Figure 4.7 and Table 4.2). Each of the six subregions show a different eddy characteristic, *videlicet* the northern bin shows no clustering, the next south shows an eddy pair cluster, the next south shows alternating two cyclonic and one anticyclonic, the south west one shows a cluster of cyclonic eddies, the south east shows a cluster of anticyclonic eddies, and the southern-most shows no clustering again, with some significant dispersion to the eddies (Figure 4.8). This suggests that eddy center location has an impact on the clustering of SST data with a high degree of spatial correlation as eddies affect the SST values and their spatial distributions.

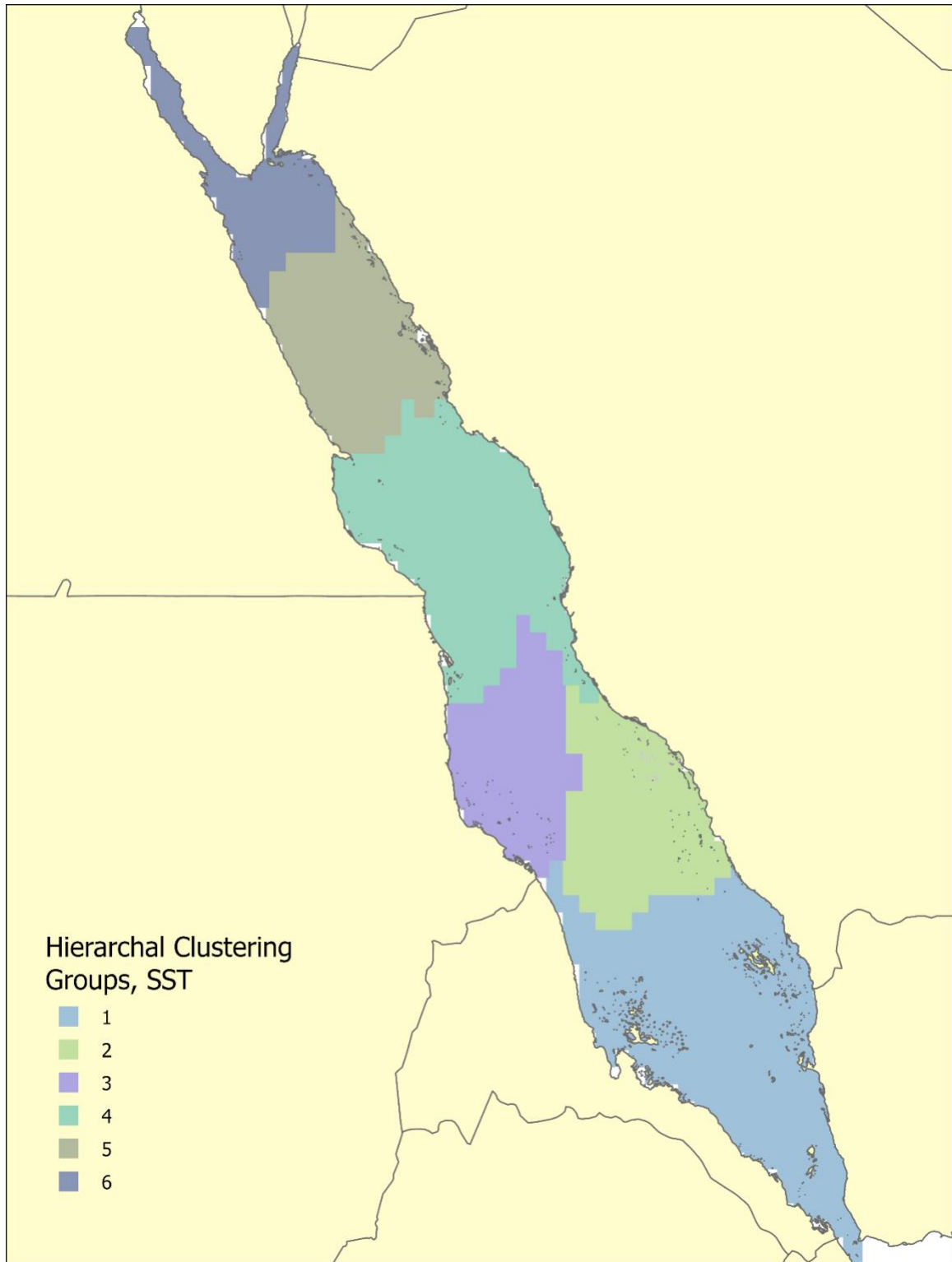


Figure 4.7: Spatial extents of the 6 clustered regions of weekly sea surface temperature anomaly over the Red Sea using a Permutation Distribution Hierarchical Clustering approach.

Cluster	Total grid points	Variable	Median	Q1	Q2	min	max
Region 1	204	SST					
		SST Anomaly	0.1186	-0.0429	0.2712	-0.4873	0.7483
Region 2	94	SST					
		SST Anomaly	0.0617	-0.0941	0.2028	-0.5777	0.7351
Region 3	75	SST					
		SST Anomaly	0.0547	-0.1566	0.2371	-0.5971	0.9688
Cluster	Total grid points	Variable	Median	Q1	Q2	min	max
Region 4	147	SST					
		SST Anomaly	-0.0167	-0.1723	0.1311	-0.5836	0.9037
Region 5	90	SST					
		SST Anomaly	-0.0061	-0.1557	0.1897	-0.5484	0.8856
Region 6	63	SST					
		SST Anomaly	0.0033	0.1235	0.1208	-0.7031	0.5351

Table 4.2: A table of summarized characteristics of the SSTA subregions. Summary of environmental variability in different clustered regions along the Red Sea for the period 1981-2019 showing mean climatology (i.e. long-term mean) and main quantiles for both SST and SSTA. Regions details are displayed from the south to the north to show the spatial variability of SST characteristics across latitudinal gradients of the Red Sea.

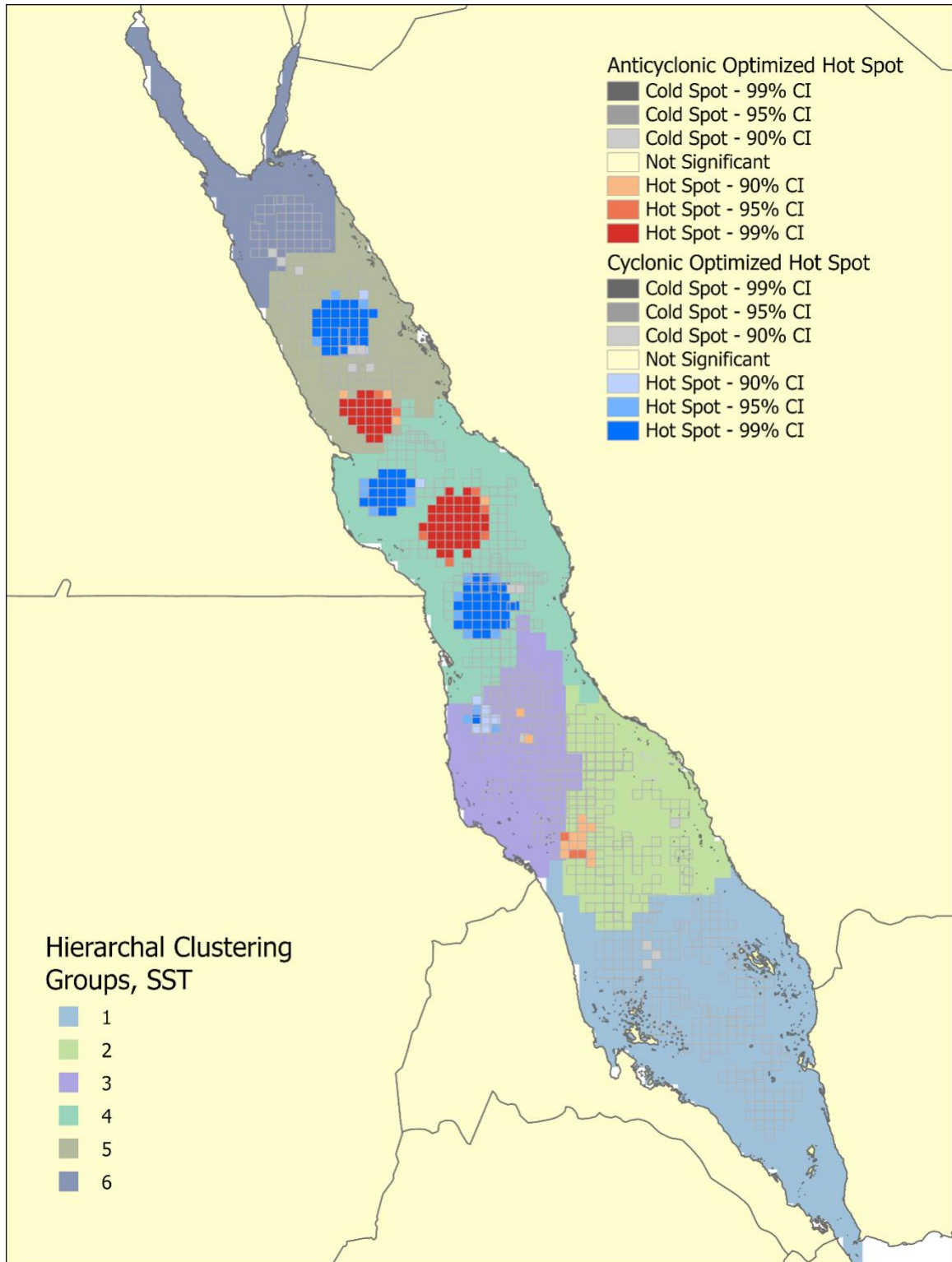


Figure 4.8: The results of the Permutation Distribution Hierarchical Clustering and the Optimized Getis-Ord G_i^* Hot Spot Analysis. Each PDC group has a unique characteristic eddy pattern, indicating the link between SST and eddy distribution.

4.4: Discussion

The linkage between the eddy locations in the Red Sea and the Sea Surface Temperature clustering as presented here provides a unique description of the eddy dynamics occurring within each identified SSTA cluster. Starting from the southern Red Sea, the first cluster no significant clustering of either cyclonic or anticyclonic eddies. The second SSTA cluster shows a significant occurrence of anticyclonic eddies, indicating this portion of the Red Sea is more influenced by anticyclonic dynamics. The third cluster is defined by a cyclonic eddy cluster, again located near the boundary, and contrasting with the second cluster. The fourth SSTA cluster has two cyclonic eddy clusters and one anticyclonic cluster, indicating the potential for significant transport between both sides of the basin, especially at times when counter-rotating eddies are present. The fifth SSTA cluster has an anticyclonic cluster and a cyclonic cluster, still creating the potential for transfer between the two sides of the basin, but at a potentially lower frequency than is the fourth SSTA cluster. The final SSTA cluster in the northern-most part of the basin doesn't have any eddy rotational preferences like the southern-most part of the Red Sea, however it can be differentiated due to the lack of a major inflow from outside of the Red Sea.

The increase in temporal resolution seen through the NRT 3.0 eddy tracking product has introduced smaller detectable eddies over the course of the nearly two years presented here, providing additional information on eddy characteristics in the Red Sea. While a direct comparison to previous mesoscale eddy data sets may not be appropriate given the differences in satellite availability, the time length of the two studies, and changes in

eddy detection processing, some general comparisons can be made. First, the general pattern of alternating eddies is replicated even with the smaller temporal time frame covered for this study, and matches with the pattern described in the 20 year data set. This indicates that eddy dynamics are important both at short and long-term temporal scales. While the cluster analysis examined all the eddy occurrences together, it should be noted that within the eddy tracking data set there are instances of simultaneous alternating eddies occurring, again demonstrating the importance of eddy dynamics at multiple time scales. Second, the southern Red Sea is shown to operate under different eddy mechanics than the rest of the basin, likely due to a combination of factors including seasonal wind patterns, extended shelves and islands, and reduced cross-basin distances at depth. This difference has been demonstrated through previously mentioned research and will continue to be an important area for further study. The different eddy mechanisms described in the southern Red Sea have been linked to species regionality for a variety of organisms. Third, the mesoscale eddy characteristics described in this research do deviate from previously reported values, decreased in both magnitude and duration, while the general spatial patterns described do match well. Additional research is warranted to further examine whether this decrease is due to the reduced sample size in this data set or if it is due to the finer resolution available with the new eddy tracking product. Fourth, the temporal variability described in this research matches with other temporal scales reported for temperature and salinity both at the surface and at depth.

The addition of multiple data sources and analysis methods has a cooperative effect on increasing our understanding of mesoscale processes in the Red Sea. Each conclusion drawn in this research supports the explanation of the overall physical processes

described, with each of the three methods building and describing the spatial patterns in the basin. The eddy center occurrences have a high degree of spatial correlation with the SST clustering, providing a potential physical process differentiating the SST clusters. This additional information on the mesoscale processes in the Red Sea provide a background for understanding the conditions that affect submesoscale processes, providing a future avenue to associate multiple levels of scale within the basin from a few kilometers up to basin wide processes. Developing a multiscale observation methodology within a relatively simple portion of the global marine system will help future research to better understand the complex interactions that occur across different levels of scale.

Chapter 5: Summary

As new methods for data collection emerge within marine science, there is a greater opportunity to understand the inter-connected global marine system. Two main approaches to the development of new data collection techniques exist, one is a focus on examining the inherent variability and patterns of global concerns, often accomplished through new advances in satellite technology, *in situ* sampling like the ARGO and biogeochemical ARGO programs, or improved numerical simulations. The second approach is to improve detailed research in local systems, simultaneously trying to better understand the fine scale patterns and how those patterns interact beyond local scale conditions into regional and global issues. An initial disconnect exists between these two approaches, in part due to the grain size of measurements playing a non-trivial role in system variability. Whether a temporal or spatial variation in grain size, there typically does not exist a simple, linear relationship between an increase in variability and an increase in grain size. The relationship typically has to be developed by examining both new and existing data sets and methods at multiple levels of scale. As an example of this disconnect within the Red Sea can be seen with the difference in circulation patterns for the basin when looking at a finer time step, where eddies dominate the system compared to a coarser time step, where the main flow is dominated by a boundary current system. Ideally, these system interactions can be determined with a sufficiently long time series data set taken at a sufficiently fine spatio-temporal scale. In areas where these detailed data sets exist, we have been able to draw an increased understanding of marine variability across multiple ranges of scale. However, the recent global interest in marine

economies has led to a push to responsibly develop marine systems that do not have an extensive historic record of data collection. This lack of long-term data can potentially lead to a lack of adequate information to ensure the sustainable utilization of marine resources. Each new location that deploys an observation system will help develop the overall understanding of global marine dynamics, making it important to ensure that data is collected at sufficient quality to draw comparisons to the global systems.

This research has examined several types of conclusions that can be drawn from the addition of finer spatio-temporal resolution data as well as highlighting a way to aid in the incorporation of this fine scale data to basin-wide considerations in the Red Sea. In many ways the basin is an ideal laboratory for analysis at multiple levels of scale due to the reduction of many global or local sources of natural variability. Both fresh water and tidal systems can greatly affect measured variability, just to name two, and both are greatly reduced in the Red Sea. This is one of the aspects of the Red Sea that makes it ideal for the development of new processes for analyzing marine data, as was capitalized on with this research. Another aspect of the Red Sea that makes it ideal as a natural laboratory is the reduced physical scale of the basin, making it very tractable as an experimental system.

Chapter Two examined the effectiveness of a recently discussed method to determine the Characteristic Length Scale using one-dimensional transect data. The Error X method was compared both to previous research within the basin and an analysis of temporal autocorrelation to determine the minimum grain size to apply to future data collected in the north-central Red Sea. Previous work has indicated the importance of eddy dynamics at a temporal scale of two to six weeks and the existence of a “weather band” variability

at two to four weeks. The results from the analysis of a single glider transect deployed in the north-central Red Sea and its comparison to High Frequency Radar and MERRA_2 data all confirm that the Error X method was able to identify a similar CLS in the three to four week range, matching with expected values. This shows that at least over this initial data set, the Error X method can provide useful information on the naturally occurring temporal grain.

Chapter Three starts to address the addition of spatial variability into the discussion of the minimum spatio-temporal grain for the Red Sea. With the addition of a second glider running a parallel transect, we examined the differences across a relatively short spatial separation of 40 km. At the time of the study, the north-central Red Sea was under the influence of two counter-rotating eddies, an anticyclonic eddy centered north of the study area and a cyclonic eddy centered to the southwest. CLS analysis determined that four weeks was again an appropriate temporal grain for both glider transects, while the same method showed no asymptote for the HFR data over the course of the month, implying that the HFR locations had a longer temporal grain than the length of the study period. Geostrophic velocities relative to an assumed level of zero motion at 400 m were calculated for each transect at a monthly average based on the temporal grain calculated by CLS. The geostrophic velocities showed that the two glider lines were influenced separately by the northern and south-western eddies. While not immediately apparent from the finer resolution time series data, the averaged transect lines showed a clear convergence zone between the two glider transects, inferring an offshore flow due to constraints from the near-shore reefs. This offshore flow was confirmed with a monthly average of geostrophic velocities from the ARMOR3D data set.

Chapter Four examines the typical characteristics of Red Sea eddies occurring from January 2018 to November 2019, using the recently developed AVISO+ NRT 3.0 Eddy Tracking data product. While this new product is not directly comparable with longer-term eddy tracking products, it does provide two important improvements for the Red Sea. First, the NRT 3.0 data product has a finer temporal resolution, increasing the detection of eddies of shorter duration. Eddy size tends to correlate with duration, so an increase in detection of shorter eddies can also mean an increase in detection of smaller eddies as well. Second, the product uses Absolute Dynamic Topography as opposed to Sea Level Anomaly, increasing eddy detection in areas that have consistent boundary currents or other long-term flows. In addition to describing the eddy characteristics, two separate cluster analyses were performed, one on eddy center location and the other on Sea Surface Temperature from 1981-2019. These two clustering methods showed a close spatial relationship between each other, with each of the SST clusters showing a different eddy characteristic.

As seen from the chapters presented, collecting data at increasingly finer scales has the potential to affect the conclusions drawn when compared to coarser scale data. One example of this can be seen in the comparison of the geostrophic velocities calculated from the glider data compared to the ARMOR3D data as seen in Chapter 3. While in general the patterns observed at the 2-kilometer scale and the 25-kilometer scale are the same, there is a quantifiable amount of variability lost at the coarse scale measurements. This information loss can be important for understanding modelling of events like larval dispersion, especially given the non-uniform nature of the geostrophic velocities from the glider line both across the horizontal and vertical extent described. This increase in

variability is important for larval dispersion because it shows that even a change in the position of the larva in the water column can have a drastic effect on the ultimate destination of the larva, with surface velocities flowing opposite to velocities at 76 m. This can introduce a new variable for modelling larval dispersion, where the daily behavior of the larvae in the water column need to be taken into consideration.

The results from this presented work indicates the importance of developing an adaptable ocean observation system as a part of the initial framework. The focus on an adaptable system is important for any new ocean observation systems initially due to a lack of available information. An adaptable system will be able to maintain data quality and understanding as conditions within the study area change or new instrumentation becomes available. As shown from these presented chapters, the addition of either new equipment or new data processing methods can affect the level of detail that is recorded, which in turn affects the type of questions that can be answered. Maintaining an adaptable ocean observation system requires flexibility in monitoring procedures as well as robust statistical methods that are able to account for time series data sets that may not be strictly uniform due to reallocation of research efforts and equipment. In consideration for the various projects currently in development across the Red Sea associated with Vision 2030, especially with the level of focus on marine research announced for both NEOM and the Red Sea Project, it is important to design these new ocean observation systems in an adaptable manner from the beginning of observations.

5.1: Future Work

The results presented have helped build a better understanding of fine scale dynamics within the Red Sea. Additionally, the results have also shown connections to basin-wide phenomena and indicate a potential path to better integration of local-scale dynamics with basin and potentially global dynamics. This work has also shown that in these specific cases, CLS seen at finer resolutions even over a short time frame match with values calculated from longer time series data. My interest going forward has two main directions. First, I would like to examine the effect of longer time series data sets on CLS calculation through the Error X method. I would like to achieve this through a longer multiple glider data set within the north-central Red Sea, as well as extending the analysis to other areas of the Red Sea. Another aspect of this line of research would be linking a more expansive glider data set with either new or existing numerical simulations. The second main direction I would like to take with this research is to move the analysis to similar ocean observation systems outside of the Red Sea. This comparison between regions will be useful in determining the applicability of the Error X CLS to potentially more complicated marine systems outside of the Red Sea.

REFERENCES:

- Al Senafi, F., Anis, A., Menezes, V., 2019. Surface Heat Fluxes over the Northern Arabian Gulf and the Northern Red Sea: Evaluation of ECMWF-ERA5 and NASA-MERRA2 Reanalyses. *Atmosphere*, 10(9): 504.
- Almahasheer, H., Duarte, C.M., 2020. Perceptions of Marine Environmental Issues by Saudi Citizens. *Frontiers in Marine Science*, 7: 600.
- Amari, S.-i., Nagaoka, H., 2007. *Methods of information geometry*, 191. American Mathematical Soc.
- Asfahani, K. et al., 2020. Capturing a Mode of Intermediate Water Formation in the Red Sea. *Journal of Geophysical Research: Oceans*, 125(4): e2019JC015803.
- Banzon, V., Smith, T.M., Chin, T.M., Liu, C., Hankins, W., 2016. A long-term record of blended satellite and in situ sea-surface temperature for climate monitoring, modeling and environmental studies. *Earth Syst. Sci. Data*, 8(1): 165-176.
- Beal, L.M., Ffield, A., Gordon, A.L., 2000. Spreading of Red Sea overflow waters in the Indian Ocean. *Journal of Geophysical Research: Oceans*, 105(C4): 8549-8564.
- Berumen, M.L. et al., 2019a. Fishes and Connectivity of Red Sea Coral Reefs. In: Voolstra, C.R., Berumen, M.L. (Eds.), *Coral Reefs of the Red Sea*. Springer International Publishing, Cham, pp. 157-179.
- Berumen, M.L. et al., 2019b. The Red Sea: Environmental gradients shape a natural laboratory in a Nascent Ocean, Coral reefs of the Red Sea. *Springer*, pp. 1-10.
- Borrione, I., Oddo, P., Russo, A., Coelho, E., 2019. Understanding altimetry signals in the Northeastern Ligurian Sea using a multi-platform approach. *Deep Sea Research Part I: Oceanographic Research Papers*, 145: 83-96.
- Bouffard, J. et al., 2012. Sub-surface small-scale eddy dynamics from multi-sensor observations and modeling. *Progress in Oceanography*, 106: 62-79.
- Bower, A.S., Farrar, J.T., 2015. Air-sea interaction and horizontal circulation in the Red Sea, The Red Sea. *Springer*, pp. 329-342.
- Bower, A.S. et al., 2013. New observations of eddies and boundary currents in the Red Sea, EGU General Assembly Conference Abstracts, pp. 9875.
- Brainerd, K.E., Gregg, M.C., 1995. Surface mixed and mixing layer depths. *Deep Sea Research Part I: Oceanographic Research Papers*, 42(9): 1521-1543.
- Brandmaier, A.M., 2011. Permutation distribution clustering and structural equation model trees.
- Brewin, R.J. et al., 2015. Regional ocean-colour chlorophyll algorithms for the Red Sea. *Remote Sensing of Environment*, 165: 64-85.
- Carpenter, J.H., 1965. The Chesapeake Bay Institute technique for the Winkler dissolved oxygen method. *Limnology and Oceanography*, 10(1): 141-143.
- Centurioni, L.R. et al., 2017. Northern Arabian Sea Circulation-Autonomous Research (NASCar) A RESEARCH INITIATIVE BASED ON AUTONOMOUS SENSORS. *Oceanography*, 30(2): 74-87.
- Chaidez, V., Dreano, D., Agusti, S., Duarte, C.M., Hoteit, I., 2017. Decadal trends in Red Sea maximum surface temperature. *Scientific Reports*, 7.
- Chao, Y. et al., 2017. Development, Implementation, and Validation of a California Coastal Ocean Modeling, Data Assimilation, and Forecasting System. *Deep Sea Research Part II: Topical Studies in Oceanography*.

- Churchill, J.H., Abualnaja, Y., Limeburner, R., Nellayaputhenpeedika, M., 2018. The dynamics of weather-band sea level variations in the Red Sea. *Regional Studies in Marine Science*, 24: 336-342.
- Churchill, J.H., Bower, A.S., McCorkle, D.C., Abualnaja, Y., 2014a. The transport of nutrient-rich Indian Ocean water through the Red Sea and into coastal reef systems. *Journal of Marine Research*, 72(3): 165-181.
- Churchill, J.H., Lentz, S.J., Farrar, J.T., Abualnaja, Y., 2014b. Properties of Red Sea coastal currents. *Continental Shelf Research*, 78: 51-61.
- Crowder, L., Norse, E., 2008. Essential ecological insights for marine ecosystem-based management and marine spatial planning. *Marine policy*, 32(5): 772-778.
- Davis, R.E., Eriksen, C.C., Jones, C.P., 2002. Autonomous buoyancy-driven underwater gliders. Taylor and Francis, London, pp. 37-58.
- Davis, R.E. et al., 2019. 100 Years of Progress in Ocean Observing Systems. *Meteorological Monographs*, 59: 3.1-3.46.
- de Boyer Montégut, C., Madec, G., Fischer, A.S., Lazar, A., Iudicone, D., 2004. Mixed layer depth over the global ocean: An examination of profile data and a profile-based climatology. *Journal of Geophysical Research: Oceans*, 109(C12).
- DeCarlo, T.M. et al., 2020. Patterns, drivers, and ecological implications of upwelling in coral reef habitats of the southern Red Sea.
- Dickey, T.D., 2003. Emerging ocean observations for interdisciplinary data assimilation systems. *Journal of Marine Systems*, 40: 5-48.
- Dreano, D., Raitsos, D.E., Gittings, J., Krokos, G., Hoteit, I., 2016. The Gulf of Aden Intermediate Water Intrusion Regulates the Southern Red Sea Summer Phytoplankton Blooms. *PloS one*, 11(12): e0168440.
- Dreschler-Fischer, L. et al., 2014. Detecting and tracking small scale eddies in the black sea and the Baltic Sea using high-resolution Radarsat-2 and TerraSAR-X imagery (DTeddie), *Geoscience and Remote Sensing Symposium (IGARSS)*, 2014 IEEE International. IEEE, pp. 1214-1217.
- Duacs/AVISO+, 2017. Mesoscale eddy trajectory Atlas product handbook.
- Ellis, J., Schneider, D., 2008. Spatial and temporal scaling in benthic ecology. *Journal of Experimental Marine Biology and Ecology*, 366(1): 92-98.
- Ellis, J.I. et al., 2019. Multiple stressor effects on coral reef ecosystems. *Global change biology*, 25(12): 4131-4146.
- Faghmous, J.H. et al., 2015. A daily global mesoscale ocean eddy dataset from satellite altimetry. *Scientific data*, 2(1): 1-16.
- Frajka-Williams, E., Rhines, P.B., Eriksen, C.C., 2009. Physical controls and mesoscale variability in the Labrador Sea spring phytoplankton bloom observed by Seaglider. *Deep Sea Research Part I: Oceanographic Research Papers*, 56(12): 2144-2161.
- Fratantoni, D.M., Bower, A.S., Johns, W.E., Peters, H., 2006. Somali Current rings in the eastern Gulf of Aden. *Journal of Geophysical Research: Oceans*, 111(C9).
- Gelaro, R. et al., 2017. The modern-era retrospective analysis for research and applications, version 2 (MERRA-2). *Journal of Climate*, 30(14): 5419-5454.
- Gerges, M.A., 2002. The Red Sea and Gulf of Aden action plan—facing the challenges of an ocean gateway. *Ocean & coastal management*, 45(11): 885-903.
- Getis, A., Ord, J.K., 1992. The analysis of spatial association by use of distance statistics. *Geographical analysis*, 24(3): 189-206.
- Gittings, J.A. et al., 2019. Remotely sensing phytoplankton size structure in the Red Sea. *Remote Sensing of Environment*, 234: 111387.

- Gittings, J.A. et al., 2017. Seasonal phytoplankton blooms in the Gulf of Aden revealed by remote sensing. *Remote Sensing of Environment*, 189: 56-66.
- Guinehut, S., Dhomps, A.-L., Larnicol, G., Le Traon, P.-Y., 2012. High resolution 3-D temperature and salinity fields derived from in situ and satellite observations. *Ocean Science*, 8(5): 845-857.
- Habeeb, R.L., Trebilco, J., Wotherspoon, S., Johnson, C.R., 2005. Determining natural scales of ecological systems. *Ecological Monographs*, 75(4): 467-487.
- Hedley, J.D. et al., 2016. Remote sensing of coral reefs for monitoring and management: A review. *Remote Sensing*, 8(2): 118.
- Hernández-Carrasco, I. et al., 2018. Impact of HF radar current gap-filling methodologies on the Lagrangian assessment of coastal dynamics.
- Hewitt, J.E., Ellis, J.I., Thrush, S.F., 2015. Multiple stressors, nonlinear effects and the implications of climate change impacts on marine coastal ecosystems. *Global change biology*.
- Hewitt, J.E., Thrush, S.F., Dayton, P.K., Bonsdorff, E., 2007. The effect of spatial and temporal heterogeneity on the design and analysis of empirical studies of scale-dependent systems. *The American Naturalist*, 169(3): 398-408.
- Hidalgo, M., Secor, D.H., Browman, H.I., 2016. Observing and managing seascapes: linking synoptic oceanography, ecological processes, and geospatial modelling. *ICES Journal of Marine Science: Journal du Conseil*: fsw079.
- Holland, J.D., Bert, D.G., Fahrig, L., 2004. Determining the spatial scale of species' response to habitat. *BioScience*, 54(3): 227-233.
- Holte, J., Talley, L.D., Gilson, J., Roemmich, D., 2017. An Argo mixed layer climatology and database. *Geophysical Research Letters*, 44(11): 5618-5626.
- Hurlbert, A.H., White, E.P., 2005. Disparity between range map-and survey-based analyses of species richness: patterns, processes and implications. *Ecology Letters*, 8(3): 319-327.
- Jain, V. et al., 2017. Evidence for the existence of Persian Gulf water and Red Sea water in the Bay of Bengal. *Climate dynamics*, 48(9): 3207-3226.
- Jones, B.H., Kattan, Y., 2017. KAUST's Red Sea Observing System. *Oceanography*, 30(2): 53-55.
- Kaplan, D.M., Lekien, F., 2007. Spatial interpolation and filtering of surface current data based on open-boundary modal analysis. *Journal of Geophysical Research: Oceans*, 112(C12).
- Karimova, S.S., Gade, M., 2014. Eddies in the Red Sea as seen by Satellite SAR Imagery, *Remote Sensing of the African Seas*. Springer, pp. 357-378.
- Kavanaugh, M.T. et al., 2016. Seascapes as a new vernacular for pelagic ocean monitoring, management and conservation. *ICES Journal of Marine Science*, 73(7): 1839-1850.
- Keeling, M., Mezić, I., Hendry, R., McGlade, J., Rand, D., 1997. Characteristic length scales of spatial models in ecology via fluctuation analysis. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 352(1361): 1589-1601.
- Kheireddine, M. et al., 2017. Assessing pigment-based phytoplankton community distributions in the Red Sea. *Frontiers in Marine Science*, 4: 132.
- Kullback, S., Leibler, R.A., 1951. On information and sufficiency. *The annals of mathematical statistics*, 22(1): 79-86.
- Kürten, B. et al., 2019. Seasonal modulation of mesoscale processes alters nutrient availability and plankton communities in the Red Sea. *Progress in Oceanography*, 173: 238-255.
- Langdon, C., 2010. Determination of Dissolved Oxygen in Seawater by Winkler Titration Using The Amperometric Technique.
- Langodan, S., Cavaleri, L., Viswanadhapalli, Y., Hoteit, I., 2014. The Red Sea: A Natural Laboratory for Wind and Wave Modeling. *Journal of Physical Oceanography*, 44(12): 3139-3159.

- Lee, C.M. et al., 2012. Marginal Ice Zone (MIZ) program: Science and experiment plan, WASHINGTON UNIV SEATTLE APPLIED PHYSICS LAB.
- Lermusiaux, P.F. et al., 2017. Optimal planning and sampling predictions for autonomous and lagrangian platforms and sensors in the northern Arabian Sea. *Oceanography*, 30(2): 172-185.
- Liblik, T. et al., 2016. Potential for an underwater glider component as part of the Global Ocean Observing System. *Methods in Oceanography*, 17: 50-82.
- Little, H., Vichi, M., Thomalla, S., Swart, S., 2018. Spatial and temporal scales of chlorophyll variability using high-resolution glider data. *Journal of Marine Systems*.
- Mason, E., Pascual, A., McWilliams, J.C., 2014. A new sea surface height-based code for oceanic mesoscale eddy tracking. *Journal of Atmospheric and Oceanic Technology*, 31(5): 1181-1188.
- Mazloff, M., Cornuelle, B., Gille, S., Verdy, A., 2018. Correlation Lengths for Estimating the Large-Scale Carbon and Heat Content of the Southern Ocean. *Journal of Geophysical Research: Oceans*, 123(2): 883-901.
- McClatchie, S., 2016. Regional fisheries oceanography of the California Current System. Springer.
- McDougall, T. et al., 2009. The International Thermodynamic Equation Of Seawater 2010 (TEOS-10): Calculation and Use of Thermodynamic Properties. Global Ship-based Repeat Hydrography Manual, IOCCP Report No, 14.
- Menezes, V.V., Farrar, J.T., Bower, A.S., 2019. Evaporative Implications of Dry-Air Outbreaks Over the Northern Red Sea. *Journal of Geophysical Research: Atmospheres*.
- Mitchell, A., 2005. The ESRI guide to GIS analysis, Volume 2: Spatial Measurements and Statistics. Redlands. CA: Esri Press.
- Mulet, S., Rio, M.-H., Mignot, A., Guinehut, S., Morrow, R., 2012. A new estimate of the global 3D geostrophic ocean circulation based on satellite data and in-situ measurements. *Deep Sea Research Part II: Topical Studies in Oceanography*, 77: 70-81.
- Murray, S.P., Johns, W., 1997. Direct observations of seasonal exchange through the Bab el Mandab Strait. *Geophysical Research Letters*, 24(21): 2557-2560.
- Nickols, K.J., White, J.W., Largier, J.L., Gaylord, B., 2015. Marine population connectivity: Reconciling large-scale dispersal and high self-retention. *The American Naturalist*, 185(2): 196-211.
- Organelli, E. et al., 2017. Bio-optical anomalies in the world's oceans: An investigation on the diffuse attenuation coefficients for downward irradiance derived from Biogeochemical Argo float measurements. *Journal of Geophysical Research: Oceans*, 122(5): 3543-3564.
- Pascual, A. et al., 2017. A Multiplatform Experiment to Unravel Meso- and Submesoscale Processes in an Intense Front (AlborEx). *Frontiers in Marine Science*, 4(39).
- Pascual, M., Levin, S.A., 1999. From individuals to population densities: searching for the intermediate scale of nontrivial determinism. *Ecology*, 80(7): 2225-2236.
- Patzert, W.C., 1974. Seasonal reversal in Red Sea circulation. *L'oceanographie Physique de la Mer Rouge*: 55-89.
- Pelland, N.A., Eriksen, C.C., Lee, C.M., 2013. Subthermocline eddies over the Washington continental slope as observed by Seaglidors, 2003–09. *Journal of Physical Oceanography*, 43(10): 2025-2053.
- Phinn, S., Roelfsema, C., Stumpf, R.P., 2010. Remote sensing: discerning the promise from the reality.
- Quadfasel, D., Baudner, H., 1993. Gyre-scale circulation cells in the Red-Sea. *Oceanologica Acta*, 16(3): 221-229.
- Quattrochi, D.A., Goodchild, M.F., 1997. Scale in remote sensing and GIS. CRC press.

- Rahbek, C., 2005. The role of spatial scale and the perception of large-scale species-richness patterns. *Ecology letters*, 8(2): 224-239.
- Raitsos, D.E. et al., 2017. Sensing coral reef connectivity pathways from space. *Scientific Reports*, 7(1): 1-10.
- Raitsos, D.E., Pradhan, Y., Brewin, R.J., Stenchikov, G., Hoteit, I., 2013. Remote sensing the phytoplankton seasonal succession of the Red Sea. *PLoS One*, 8(6): e64909.
- Rasul, N.M., Stewart, I.C., 2015. The Red Sea: the formation, morphology, oceanography and environment of a young ocean basin. Springer.
- Reynolds, R.W. et al., 2007. Daily high-resolution-blended analyses for sea surface temperature. *Journal of Climate*, 20(22): 5473-5496.
- Rienecker, M.M. et al., 2011. MERRA: NASA's modern-era retrospective analysis for research and applications. *Journal of climate*, 24(14): 3624-3648.
- Roberts, M.B. et al., 2016. Homogeneity of coral reef communities across 8 degrees of latitude in the Saudi Arabian Red Sea. *Marine Pollution Bulletin*, 105(2): 558-565.
- Roesler, C. et al., 2017. Recommendations for obtaining unbiased chlorophyll estimates from in situ chlorophyll fluorometers: A global analysis of WET Labs ECO sensors. *Limnology and Oceanography: Methods*, 15(6): 572-585.
- Rose, K.C. et al., 2017. Historical foundations and future directions in macrosystems ecology. *Ecology letters*, 20(2): 147-157.
- Rudnick, D.L., 2016. Ocean research enabled by underwater gliders. *Annual review of marine science*, 8: 519-541.
- Rudnick, D.L., Cole, S.T., 2011. On sampling the ocean using underwater gliders. *Journal of Geophysical Research: Oceans*, 116(C8).
- Rudnick, D.L., Davis, R.E., Eriksen, C.C., Fratantoni, D.M., Perry, M.J., 2004. Underwater gliders for ocean research. *Marine Technology Society Journal*, 38(2): 73-84.
- Rudnick, D.L., Zaba, K.D., Todd, R.E., Davis, R.E., 2017. A climatology of the California Current System from a network of underwater gliders. *Progress in Oceanography*, 154: 64-106.
- Schmechtig, C., Claustre, H., Poteau, A., D'Ortenzio, F., 2014. Bio-Argo quality control manual for the Chlorophyll-A concentration.
- Schmidt, R., 1986. Multiple emitter location and signal parameter estimation. *IEEE transactions on antennas and propagation*, 34(3): 276-280.
- Silverman, J., Lazar, B., Erez, J., 2007. Effect of aragonite saturation, temperature, and nutrients on the community calcification rate of a coral reef. *Journal of Geophysical Research: Oceans*, 112(C5).
- Smeed, D., 1998. Seasonal variation of the flow in the strait of Bab al Mandab. *Oceanographic Literature Review*, 6(45): 896.
- Sofianos, S., Johns, W.E., 2015. Water mass formation, overturning circulation, and the exchange of the Red Sea with the adjacent basins, *The Red Sea*. Springer, pp. 343-353.
- Sofianos, S.S., Johns, W.E., 2003. An oceanic general circulation model (OGCM) investigation of the Red Sea circulation: 2. Three-dimensional circulation in the Red Sea. *Journal of Geophysical Research: Oceans*, 108(C3).
- Sofianos, S.S., Johns, W.E., 2007. Observations of the summer Red Sea circulation. *Journal of Geophysical Research: Oceans*, 112(C6).
- Stommel, H., 1963. Varieties of oceanographic experience. *Science*, 139(3555): 572-576.
- Sultan, S., Ahmad, F., El-Hassan, A., 1995. Seasonal variations of the sea level in the central part of the Red Sea. *Estuarine, Coastal and Shelf Science*, 40(1): 1-8.

- Sun, R. et al., 2019. SKRIPS v1. 0: a regional coupled ocean–atmosphere modeling framework (MITgcm–WRF) using ESMF/NUOPC, description and preliminary results for the Red Sea. *Geoscientific Model Development*, 12(10): 4221-4244.
- Terzic, E. et al., 2019. Merging bio-optical data from Biogeochemical-Argo floats and models in marine biogeochemistry. *Biogeosciences*, 16(12): 2527-2542.
- Testor, P. et al., 2019. OceanGliders: a component of the integrated GOOS. *Frontiers in Marine Science*, 6: 422.
- Thierry, V. et al., 2018. Processing Argo oxygen data at the DAC level.
- Thompson, C.M., McGarigal, K., 2002. The influence of research scale on bald eagle habitat selection along the lower Hudson River, New York (USA). *Landscape Ecology*, 17(6): 569-586.
- Thomson, R.E., Emery, W.J., 2014. *Data analysis methods in physical oceanography*. Newnes.
- Thyng, K.M., Greene, C.A., Hetland, R.D., Zimmerle, H.M., DiMarco, S.F., 2016. True colors of oceanography: Guidelines for effective and accurate colormap selection. *Oceanography*, 29(3): 9-13.
- Triantafyllou, G. et al., 2014. Exploring the Red Sea seasonal ecosystem functioning using a three-dimensional biophysical model. *Journal of Geophysical Research: Oceans*, 119(3): 1791-1811.
- Turner, M.G., Gardner, R.H., O'Neill, R.V., 2001. *Landscape ecology in theory and practice*, 401. Springer.
- Vance, T.C., 2007. If you build it, will they come? Evolution towards the application of multi-dimensional GIS to fisheries-oceanography.
- Wahr, J., Smeed, D.A., Leuliette, E., Swenson, S., 2014. Seasonal variability of the Red Sea, from satellite gravity, radar altimetry, and in situ observations. *Journal of Geophysical Research: Oceans*, 119(8): 5091-5104.
- Wang, Y. et al., 2019. Physical connectivity simulations reveal dynamic linkages between coral reefs in the southern Red Sea and the Indian Ocean. *Scientific reports*, 9(1): 1-11.
- Ward, D.F., Wotherspoon, S., Melbourne-Thomas, J., Haapkylä, J., Johnson, C.R., 2018. Detecting ecological regime shifts from transect data. *Ecological Monographs*.
- Weatherall, P. et al., 2015. A new digital bathymetric model of the world's oceans. *Earth and Space Science*, 2(8): 331-345.
- Webster, S.E., Lee, C.M., Gobat, J.I., 2014. Preliminary results in under-ice acoustic navigation for Seagliders in Davis Strait, 2014 Oceans-St. John's. *IEEE*, pp. 1-5.
- Wilks, D.S., 2011. *Statistical methods in the atmospheric sciences*, 100. Academic press.
- Winkler, L.W., 1888. Die bestimmung des im wasser gelösten sauerstoffes. *Berichte der deutschen chemischen Gesellschaft*, 21(2): 2843-2854.
- Wong, A., Keeley, R., Carval, T., 2020. Argo quality control manual for CTD and trajectory data.
- Wu, J., 1999. Hierarchy and scaling: extrapolating information along a scaling ladder. *Canadian journal of remote sensing*, 25(4): 367-380.
- Xing, X. et al., 2012. Quenching correction for in vivo chlorophyll fluorescence acquired by autonomous platforms: A case study with instrumented elephant seals in the Kerguelen region (Southern Ocean). *Limnology and Oceanography: Methods*, 10(7): 483-495.
- Xing, X. et al., 2017. Correction of profiles of in-situ chlorophyll fluorometry for the contribution of fluorescence originating from non-algal matter. *Limnology and Oceanography: Methods*, 15(1): 80-93.
- Yao, F., Hoteit, I., 2015. Thermocline regulated seasonal evolution of surface chlorophyll in the Gulf of Aden. *PloS one*, 10(3): e0119951.

- Yao, F. et al., 2014a. Seasonal overturning circulation in the Red Sea: 2. Winter circulation. *Journal of Geophysical Research: Oceans*, 119(4): 2263-2289.
- Yao, F. et al., 2014b. Seasonal overturning circulation in the Red Sea: 1. Model validation and summer circulation. *Journal of Geophysical Research: Oceans*, 119(4): 2238-2262.
- Zarokanellos, N., 2018. The role of mesoscale processes controlling physical and biological variability in the oligotrophic Central Red Sea.
- Zarokanellos, N.D. et al., 2017a. Physical mechanisms routing nutrients in the central Red Sea. *Journal of Geophysical Research: Oceans*, 122(11): 9032-9046.
- Zarokanellos, N.D., Papadopoulos, V.P., Sofianos, S.S., Jones, B.H., Physical and biological characteristics of the winter-summer transition in the Central Red Sea. *Journal of Geophysical Research: Oceans*: n/a-n/a.
- Zarokanellos, N.D., Papadopoulos, V.P., Sofianos, S.S., Jones, B.H., 2017b. Physical and biological characteristics of the winter-summer transition in the Central Red Sea. *Journal of Geophysical Research: Oceans*, 122(8): 6355-6370.
- Zervakis, V., Kokkini, Z., Potiris, E., 2016. Estimating mixed layer depth with the use of a coastal high-frequency radar. *Continental Shelf Research*.
- Zhai, P., Bower, A., 2013. The response of the Red Sea to a strong wind jet near the Tokar Gap in summer. *Journal of Geophysical Research: Oceans*, 118(1): 421-434.
- Zhai, P., Pratt, L.J., Bower, A., 2015. On the crossover of boundary currents in an idealized model of the Red Sea. *Journal of Physical Oceanography*, 45(5): 1410-1425.
- Zhan, P., Gopalakrishnan, G., Subramanian, A.C., Guo, D., Hoteit, I., 2018. Sensitivity studies of the Red Sea eddies using adjoint method. *Journal of Geophysical Research: Oceans*, 123(11): 8329-8345.
- Zhan, P., Krokos, G., Guo, D., Hoteit, I., 2019. Three-Dimensional Signature of the Red Sea Eddies and Eddy-Induced Transport. *Geophysical Research Letters*, 46(4): 2167-2177.
- Zhan, P., Subramanian, A.C., Yao, F., Hoteit, I., 2014. Eddies in the Red Sea: A statistical and dynamical study. *Journal of Geophysical Research: Oceans*, 119(6): 3909-3925.
- Zhan, P. et al., 2016. The eddy kinetic energy budget in the Red Sea. *Journal of Geophysical Research: Oceans*.

Appendix A: Autocorrelation and Error X figures, Chapter 2

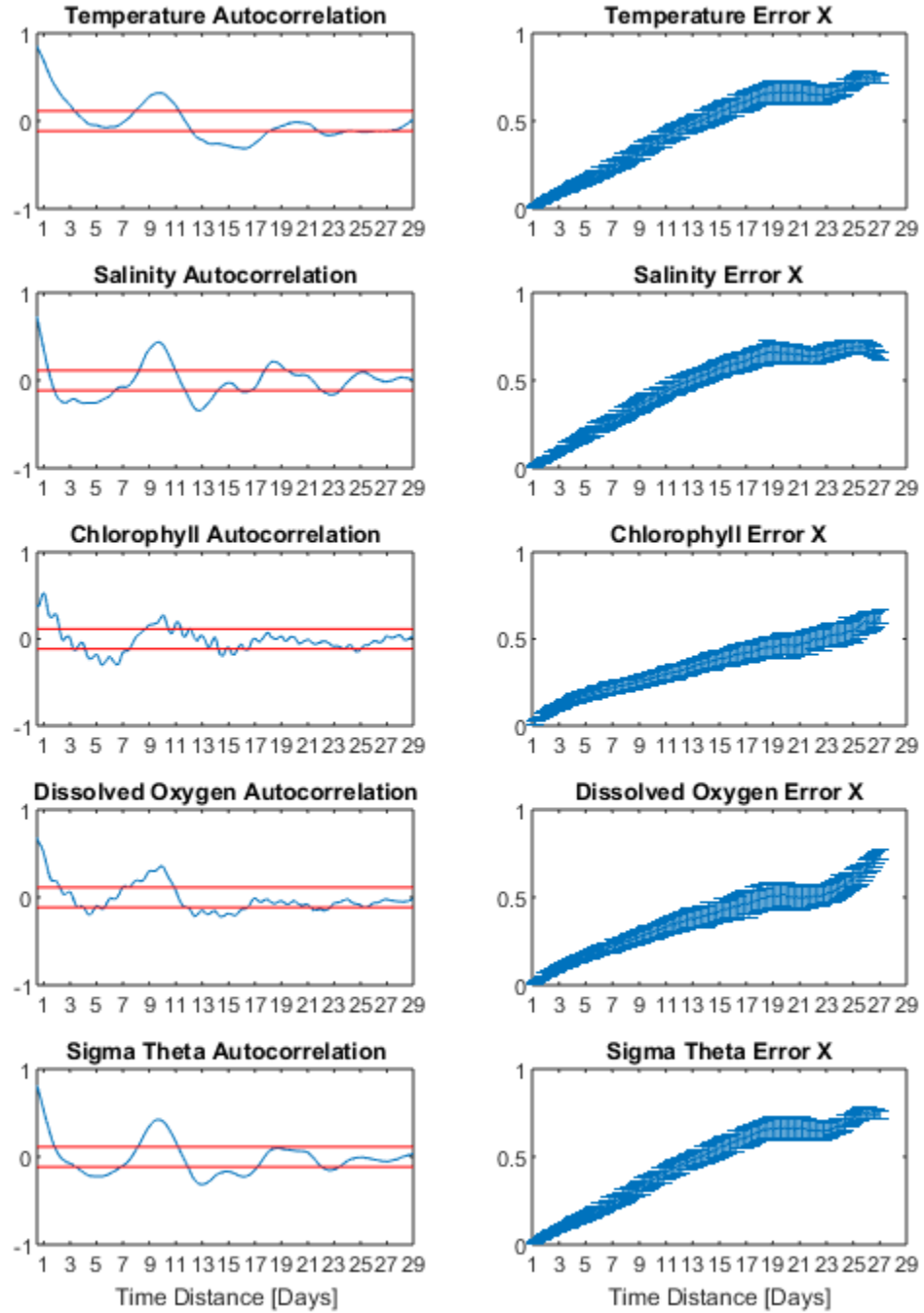


Figure A.1: Autocorrelation and Error X Monte Carlo simulation results from 6 m.

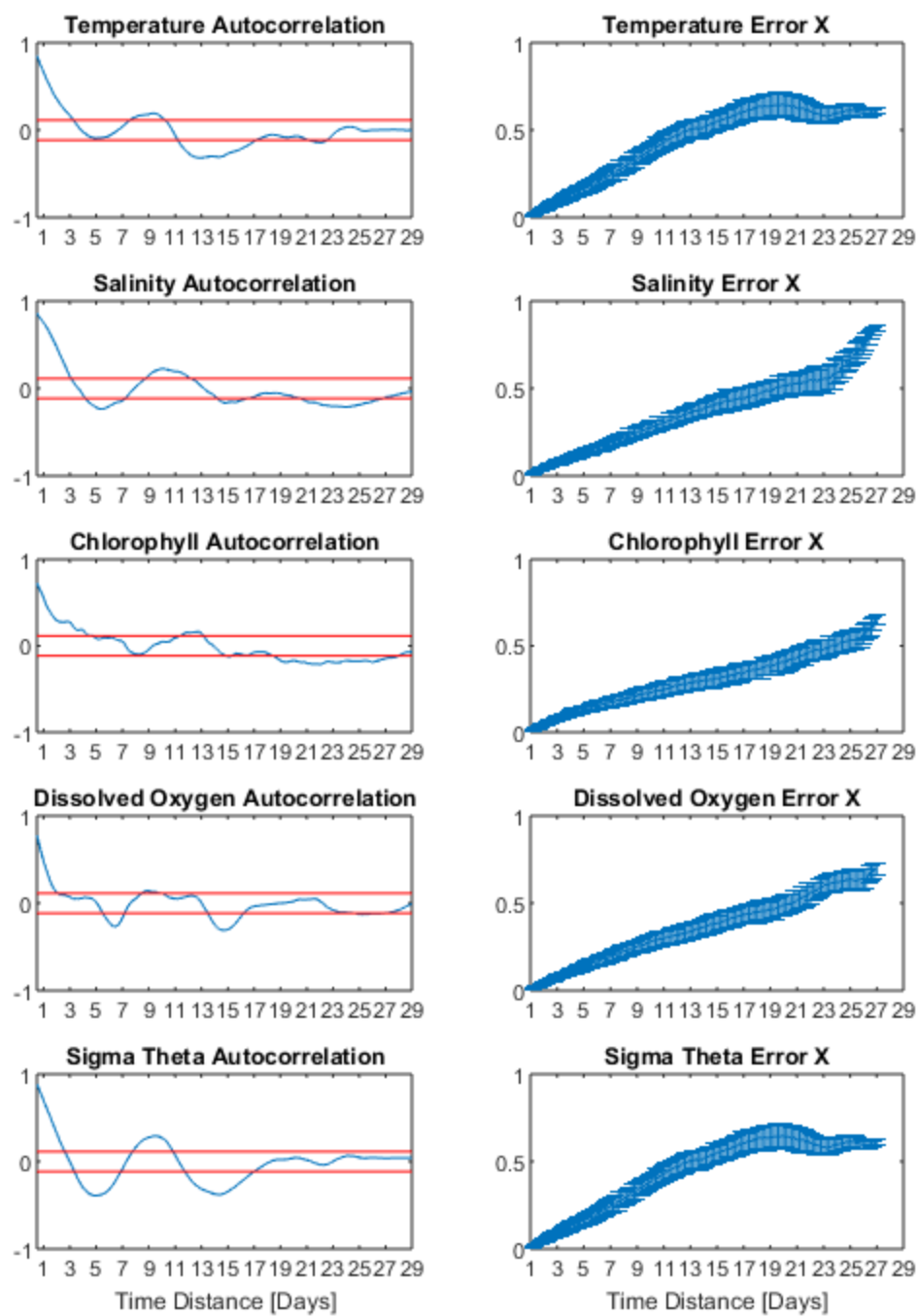


Figure A.2: Autocorrelation and Error X Monte Carlo simulation results from 76 m.

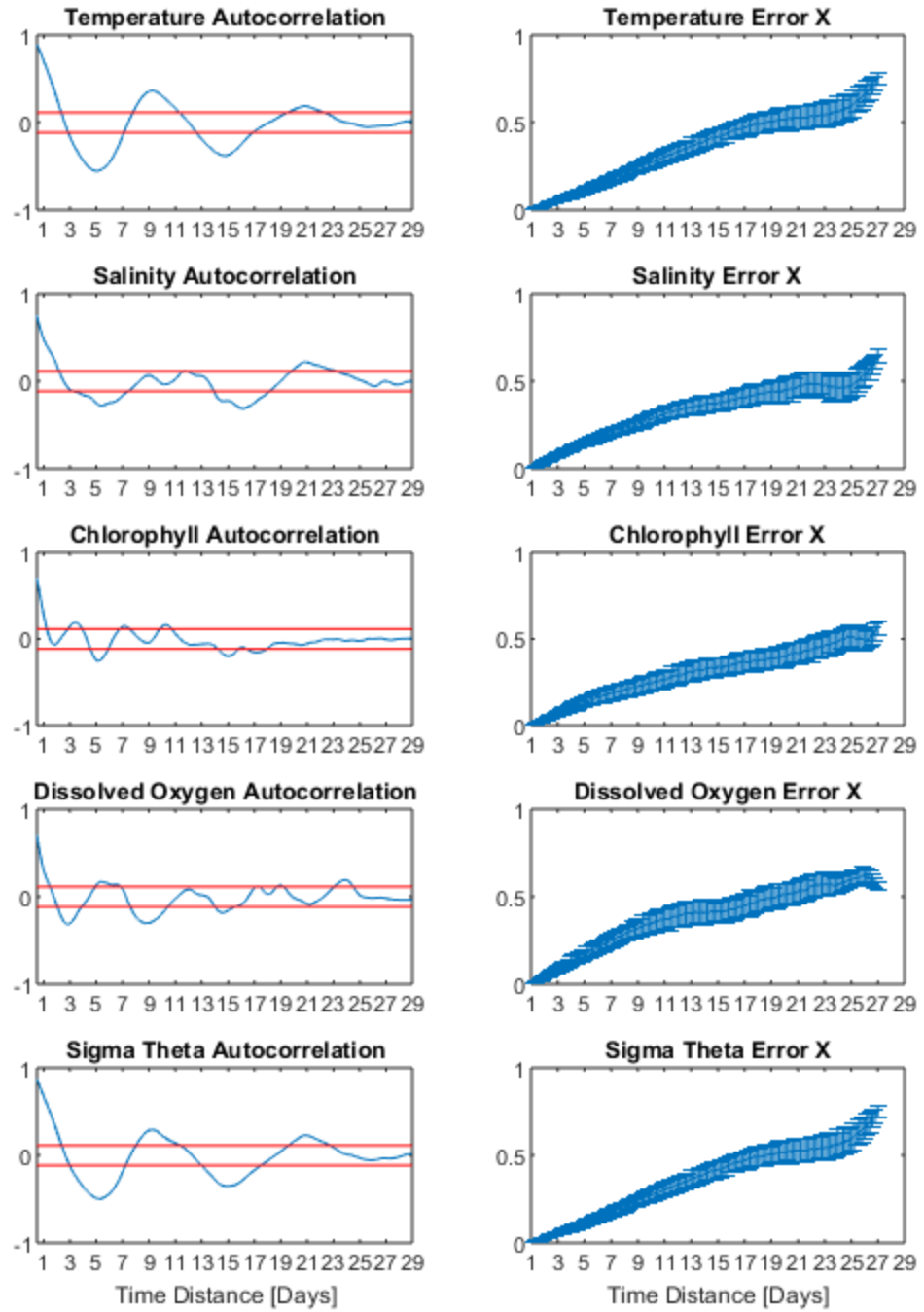


Figure A.3: Autocorrelation and Error X Monte Carlo simulation results from 150 m.

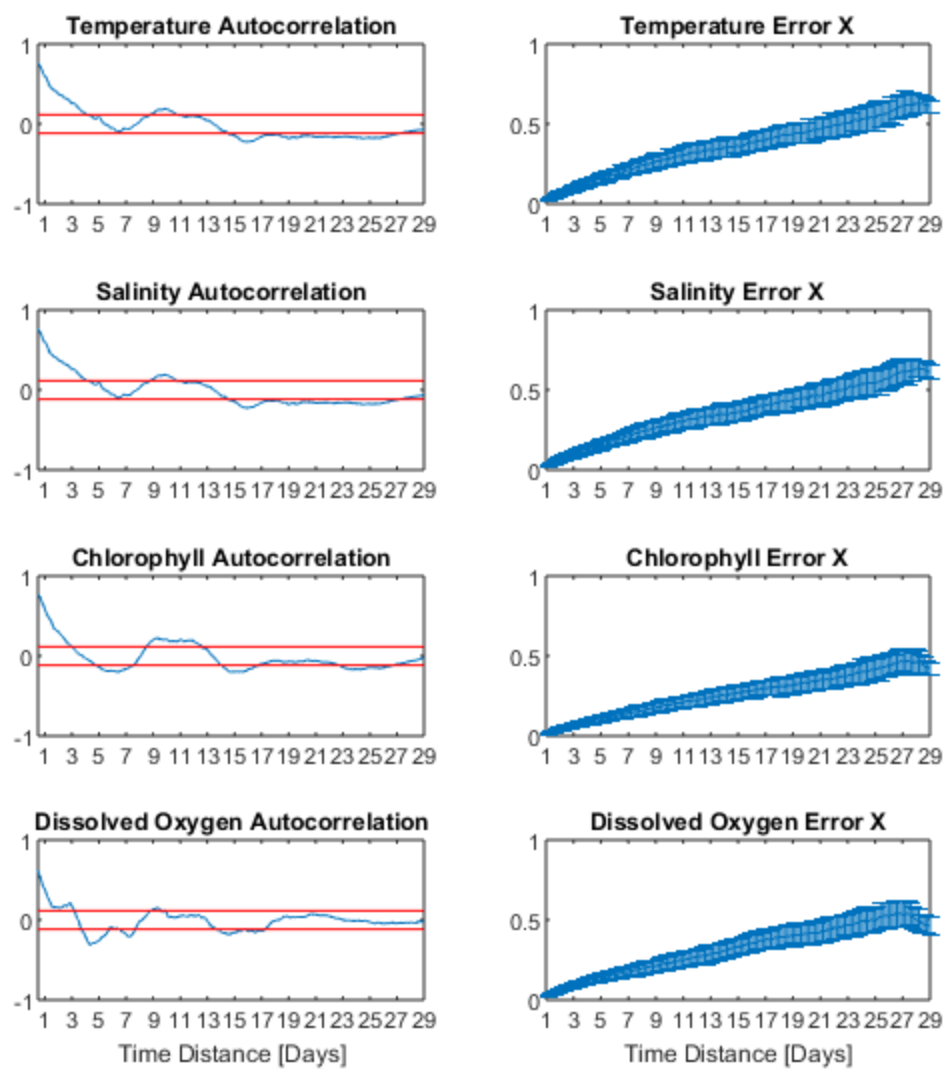


Figure A.4: Autocorrelation and Error X Monte Carlo simulation results from 25.75 kg/m³.

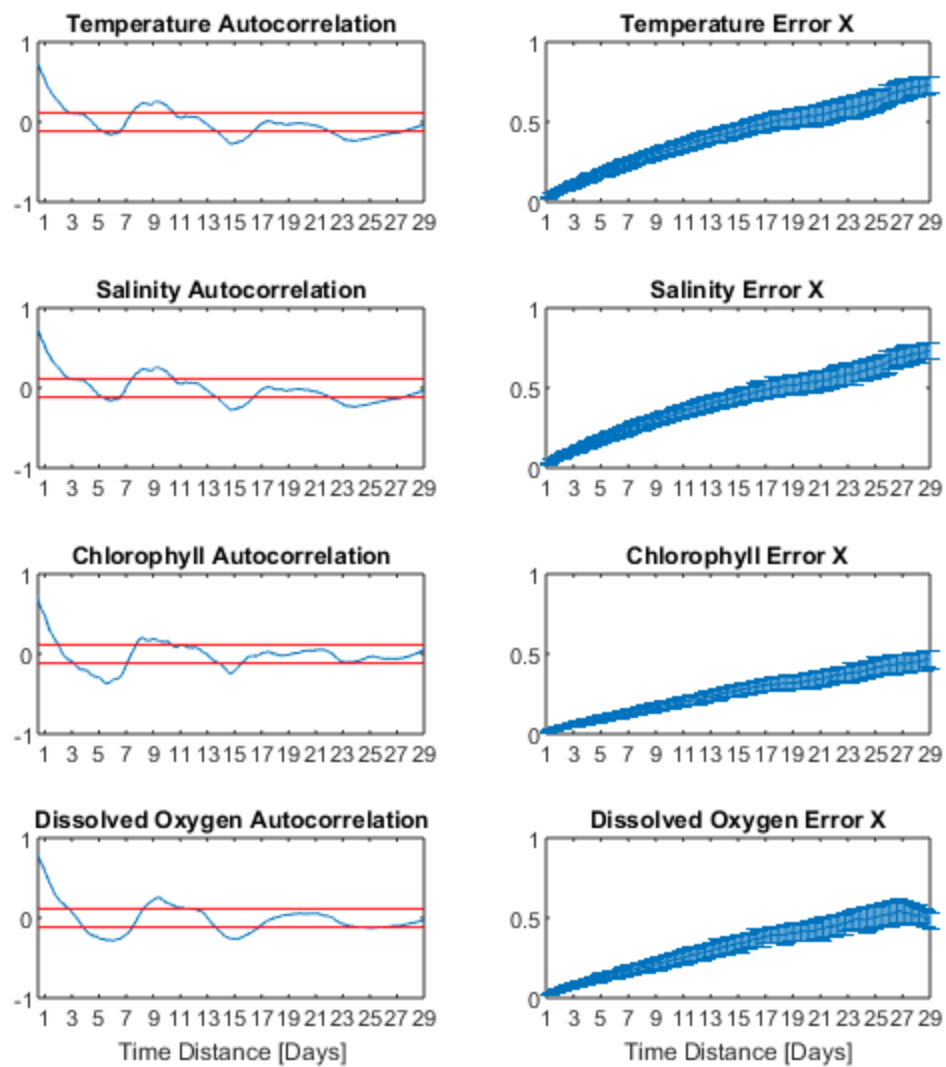


Figure A.5: Autocorrelation and Error X Monte Carlo simulation results from 26.25 kg/m³.

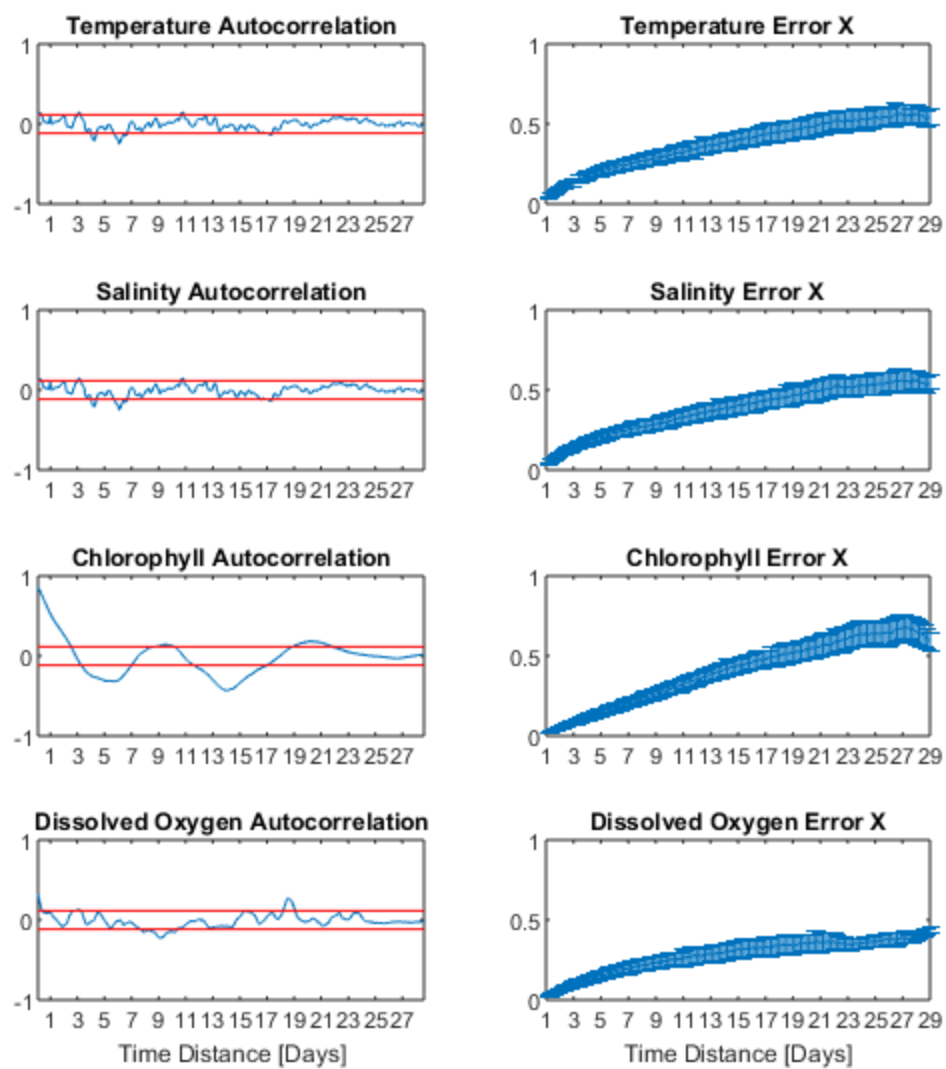


Figure A.6: Autocorrelation and Error X Monte Carlo simulation results from 28.00 kg/m³.

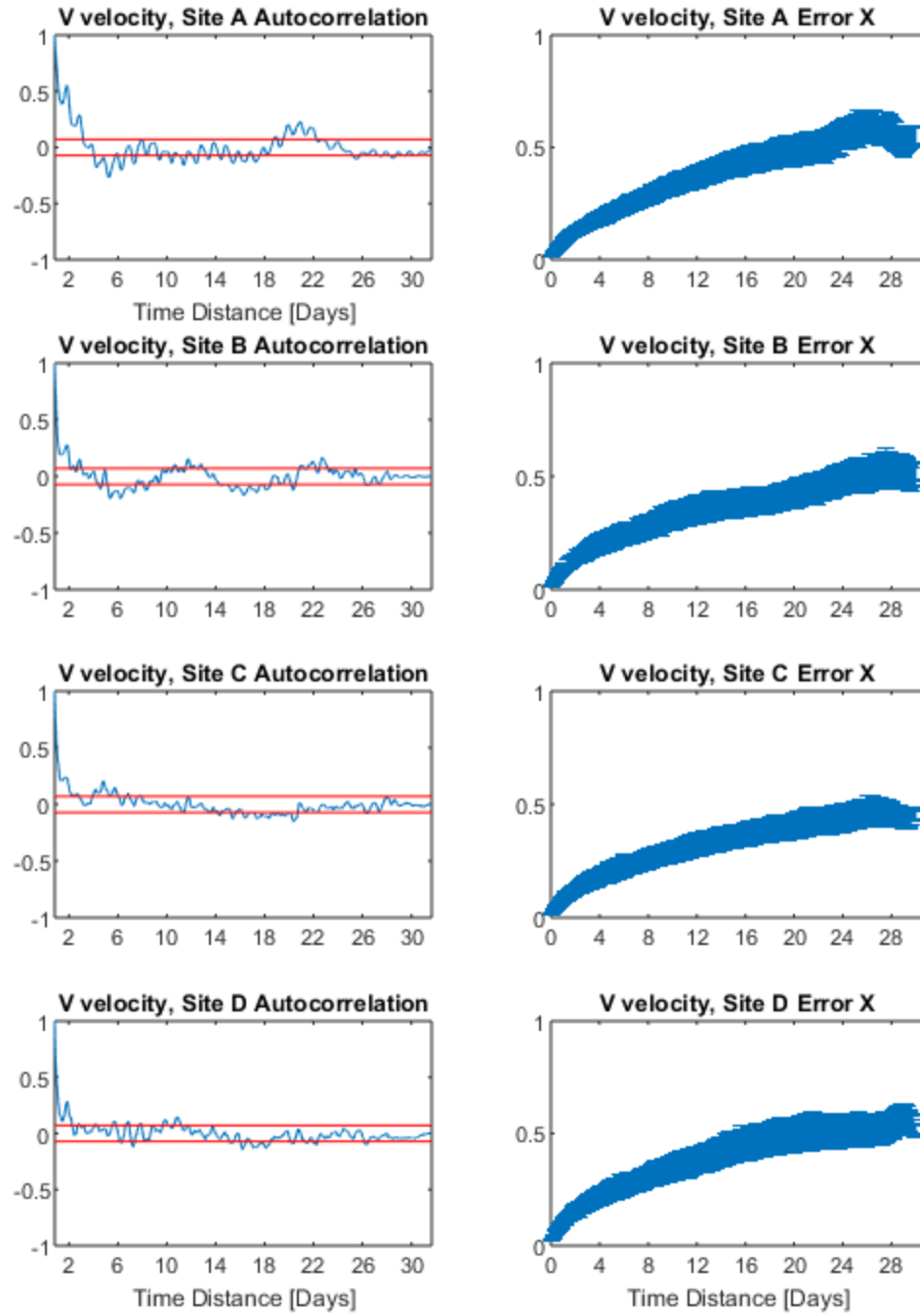


Figure A.7: Autocorrelation and Error X Monte Carlo simulation results from HFR v velocities.

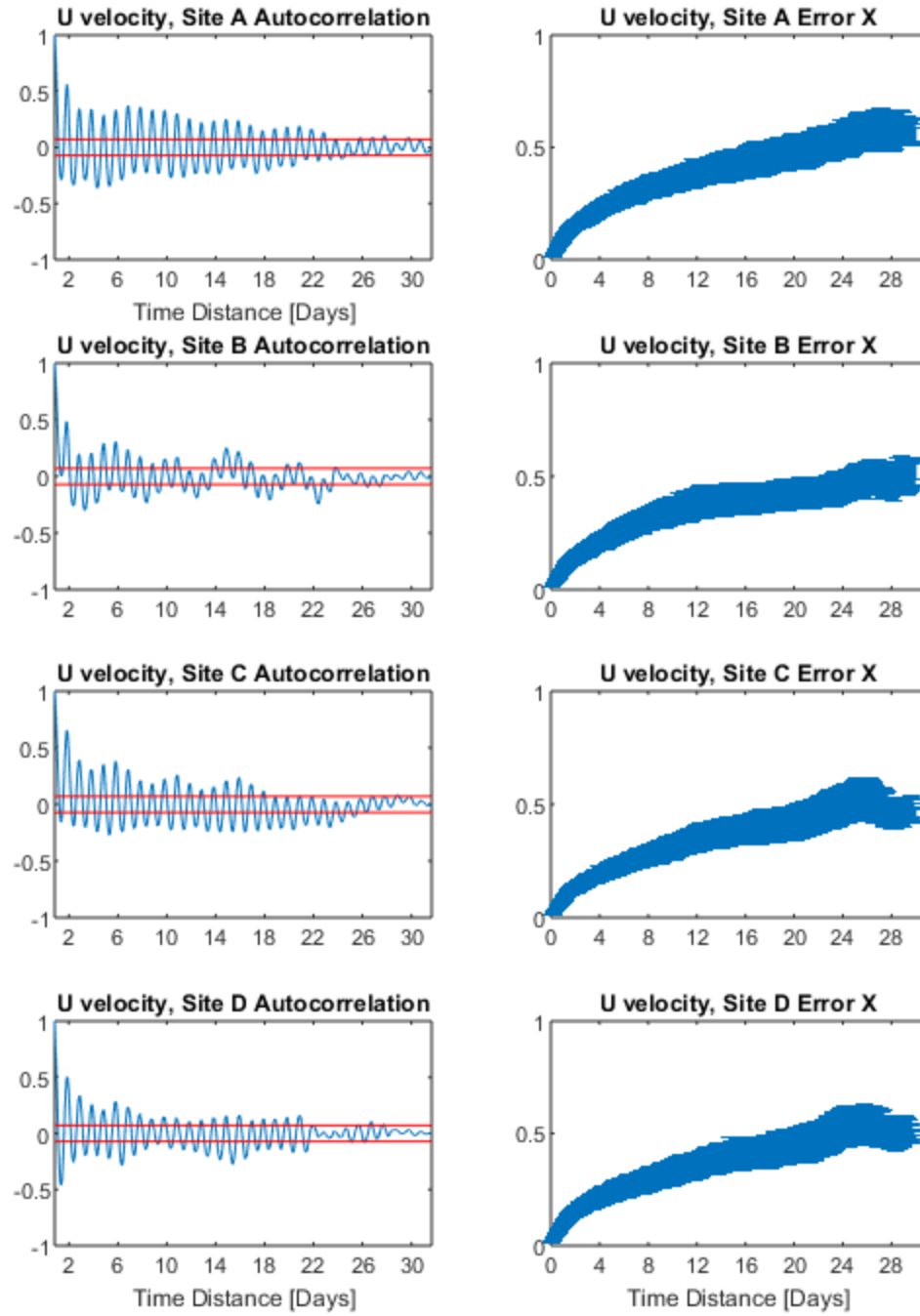


Figure A.8: Autocorrelation and Error X Monte Carlo simulation results from HFR u velocities.

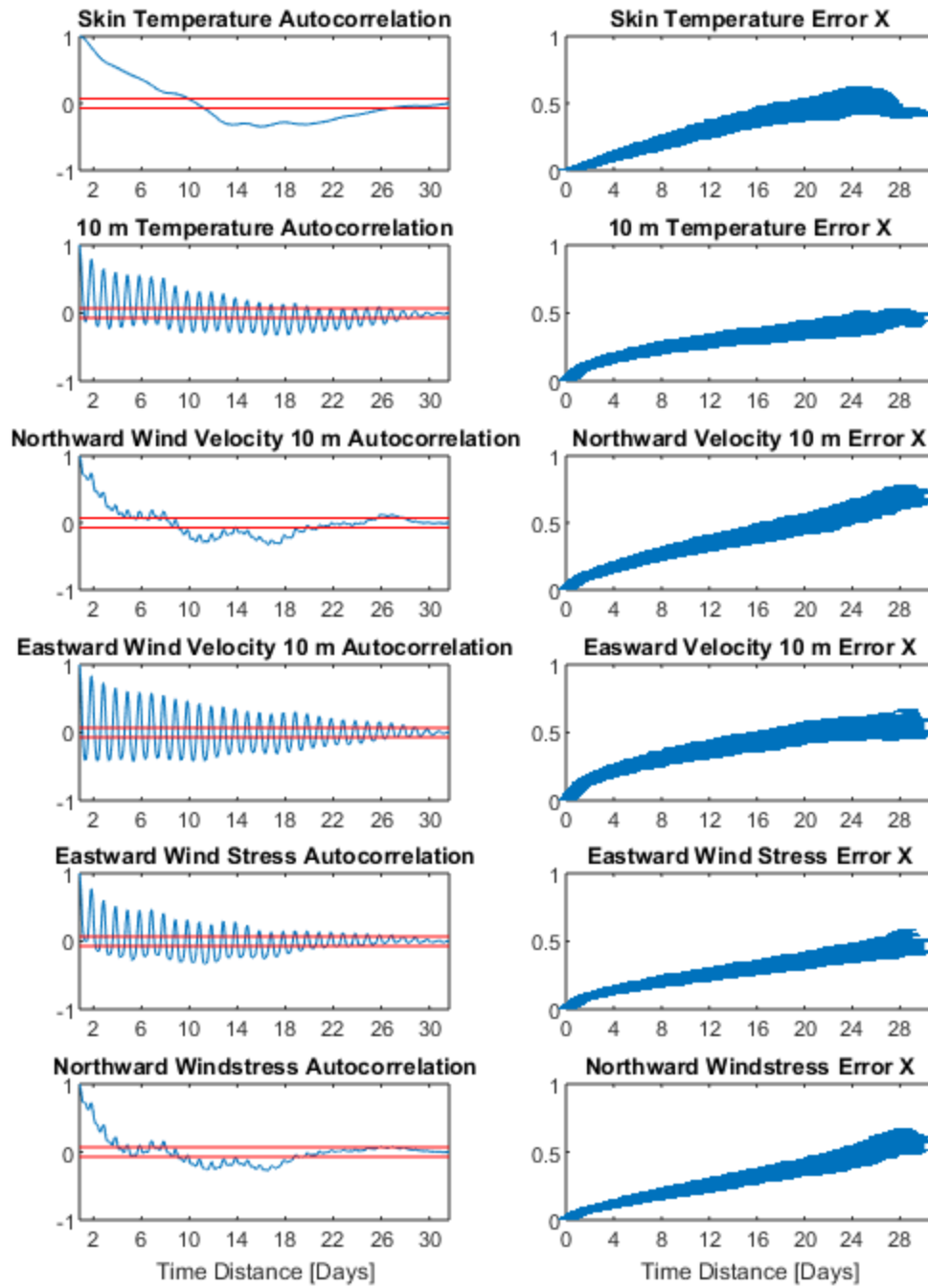


Figure A.9: Autocorrelation and Error X Monte Carlo simulation results from MERRA East Site.

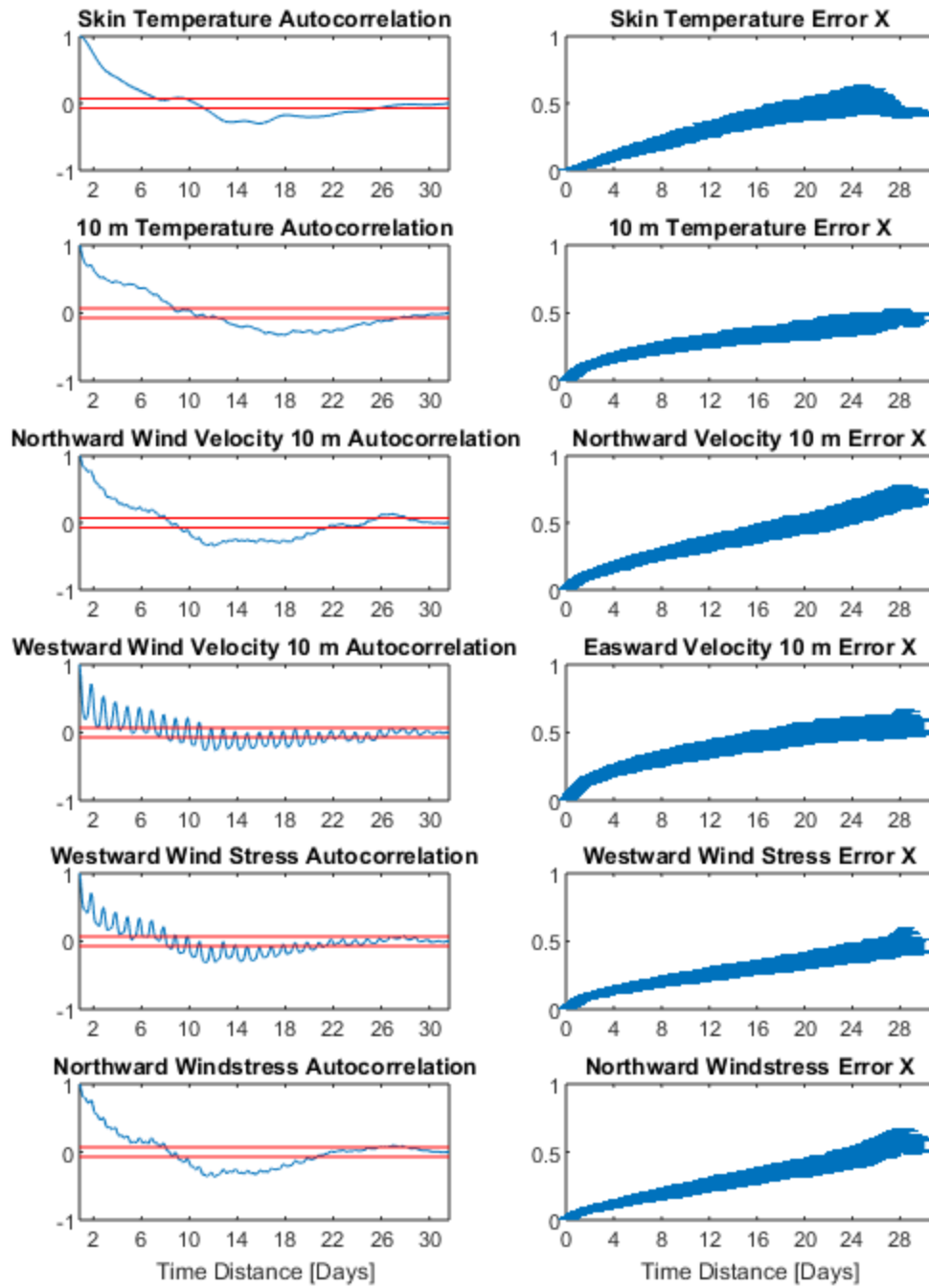


Figure A.10: Autocorrelation and Error X Monte Carlo simulation results from MERRA West Site.

Appendix B: Error X figures, Chapter 3

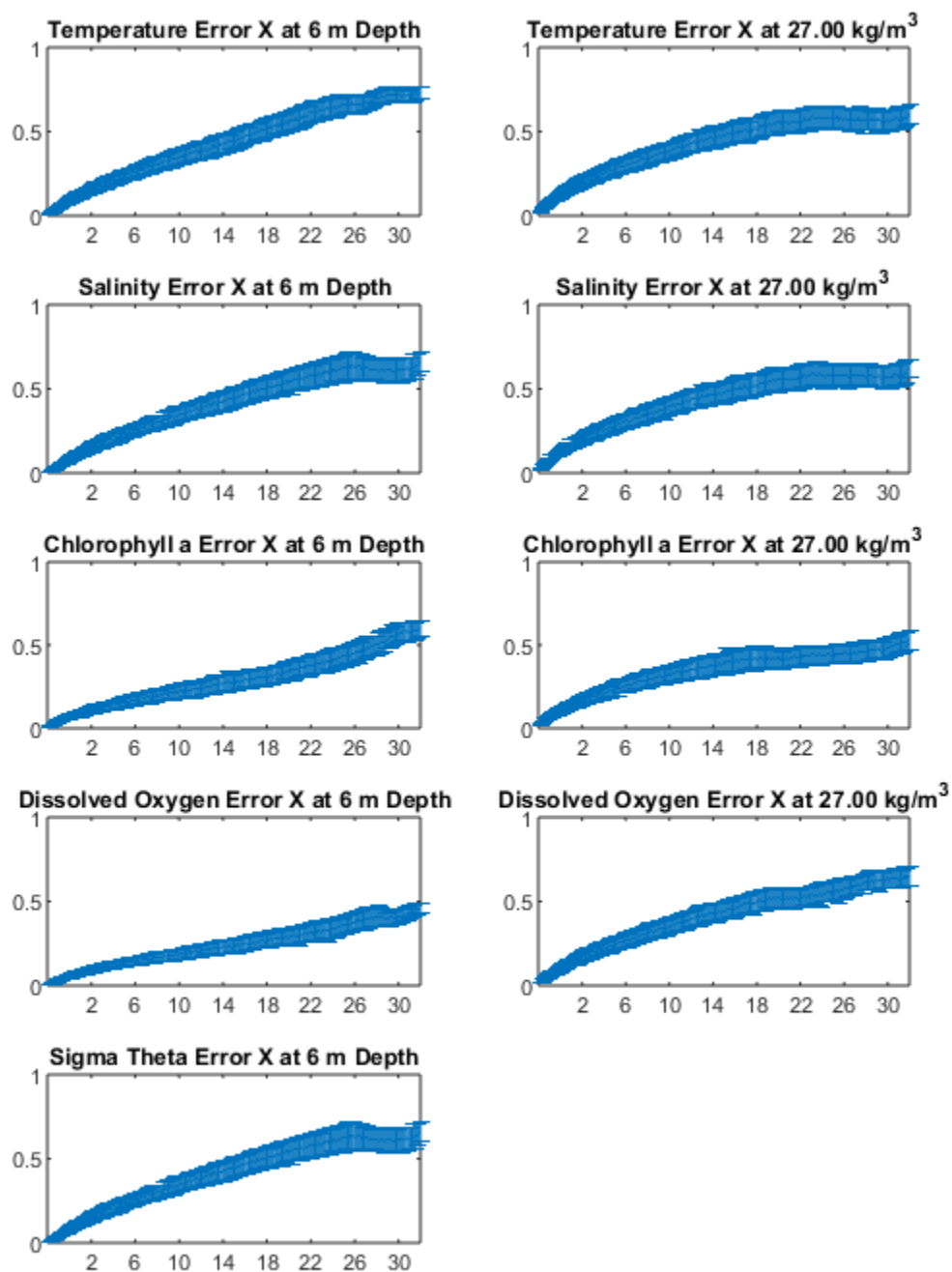


Figure B.1: North Transect Error X Monte Carlo simulation from 6 m and 27.00 kg/m³.

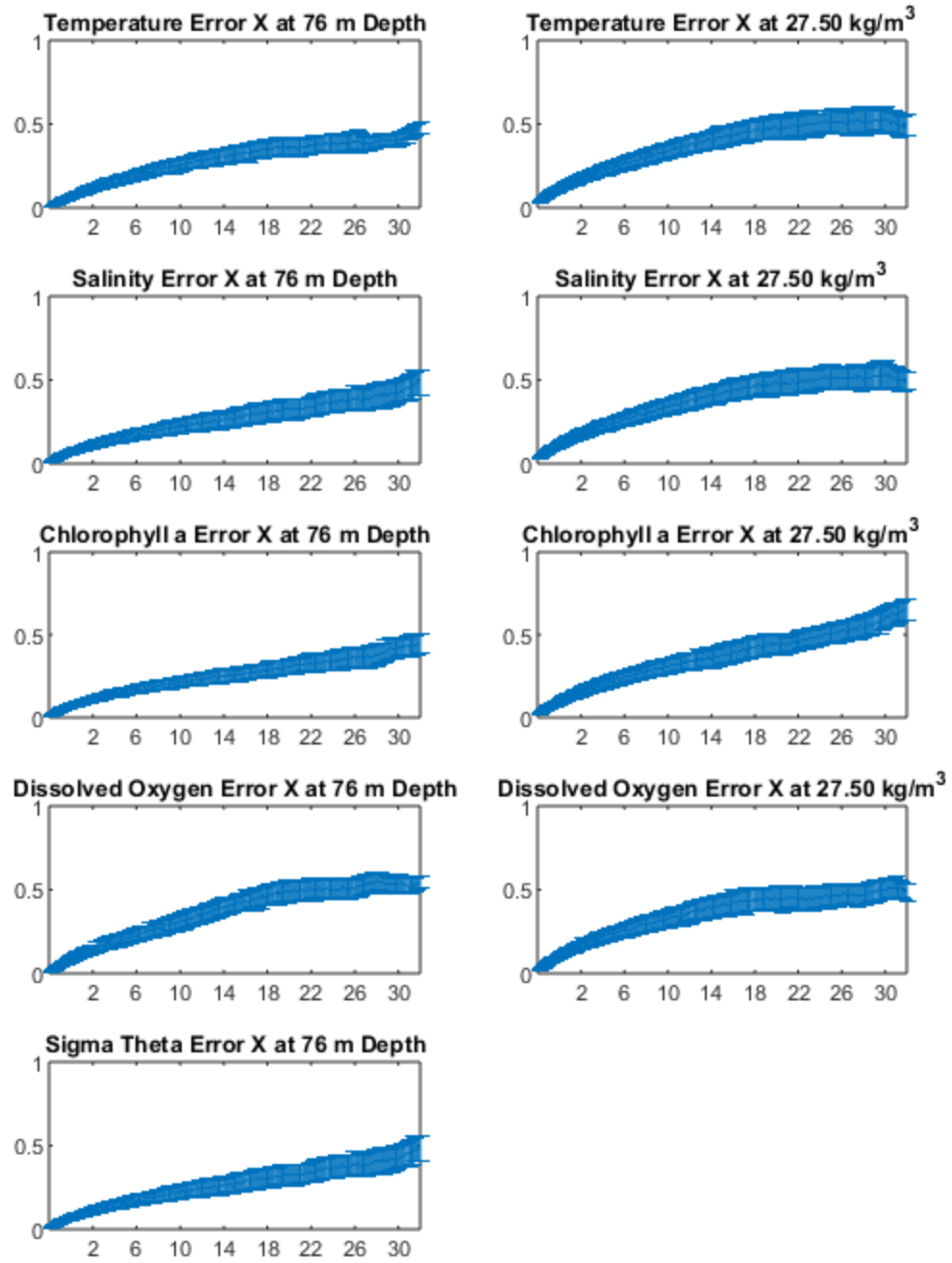


Figure B.2: North Transect Error X Monte Carlo simulation results from 76 m and 27.50 kg/m³.

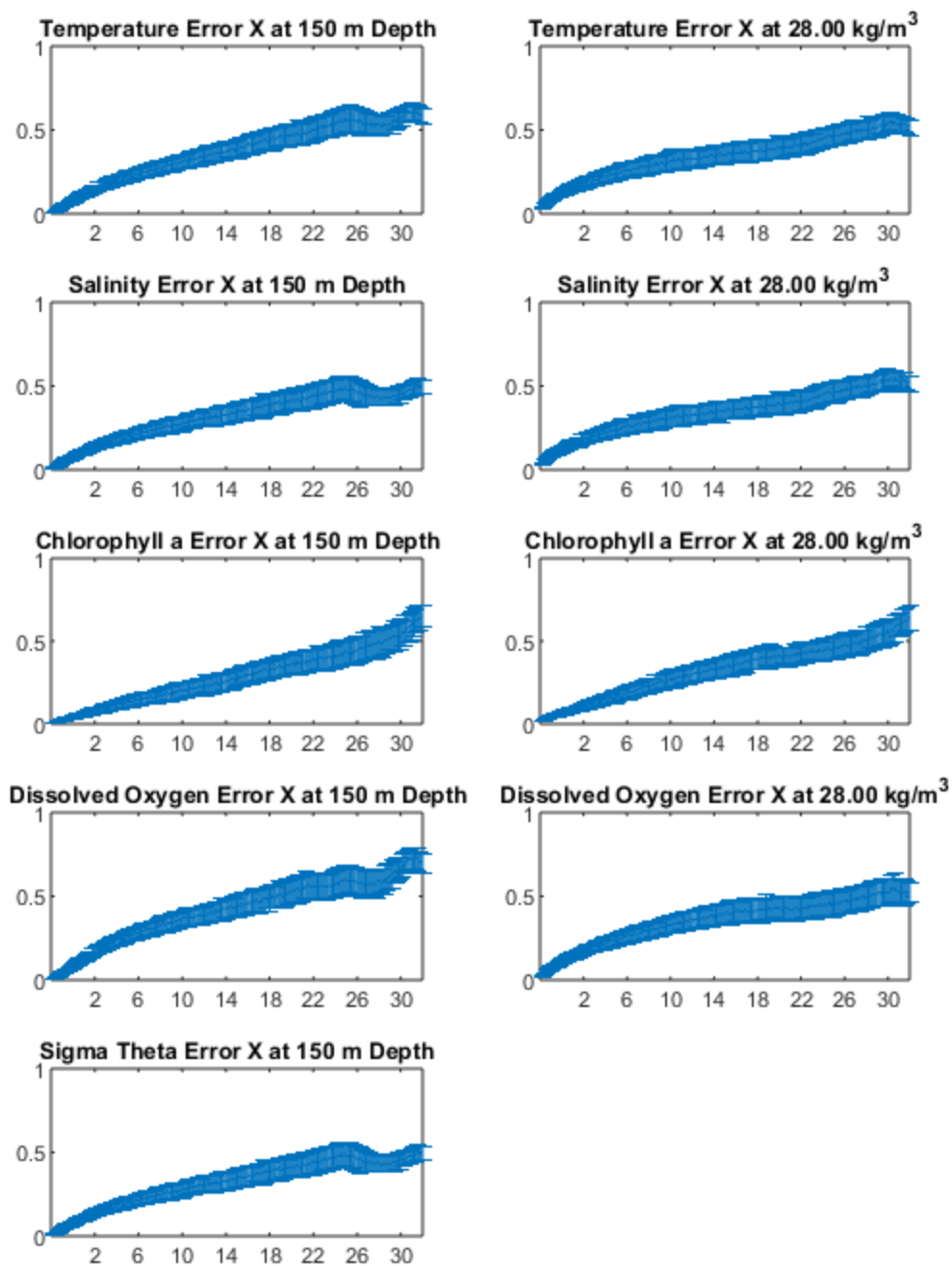


Figure B.3: North Transect Error X Monte Carlo simulation results from 150 m and 28.00 kg/m³.

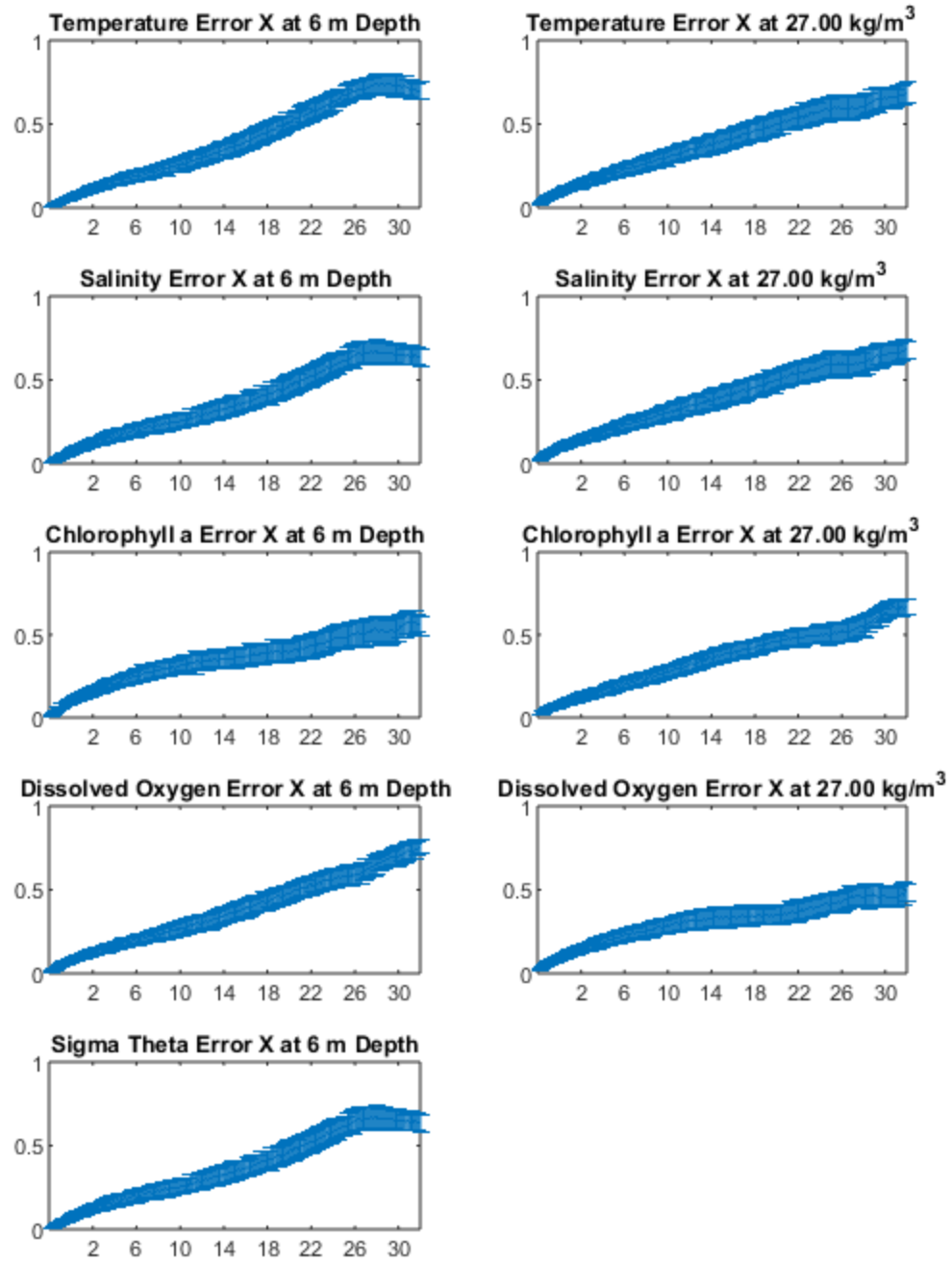


Figure B.4: South Transect Error X Monte Carlo simulation results from 6 m and 27.00 kg/m³.

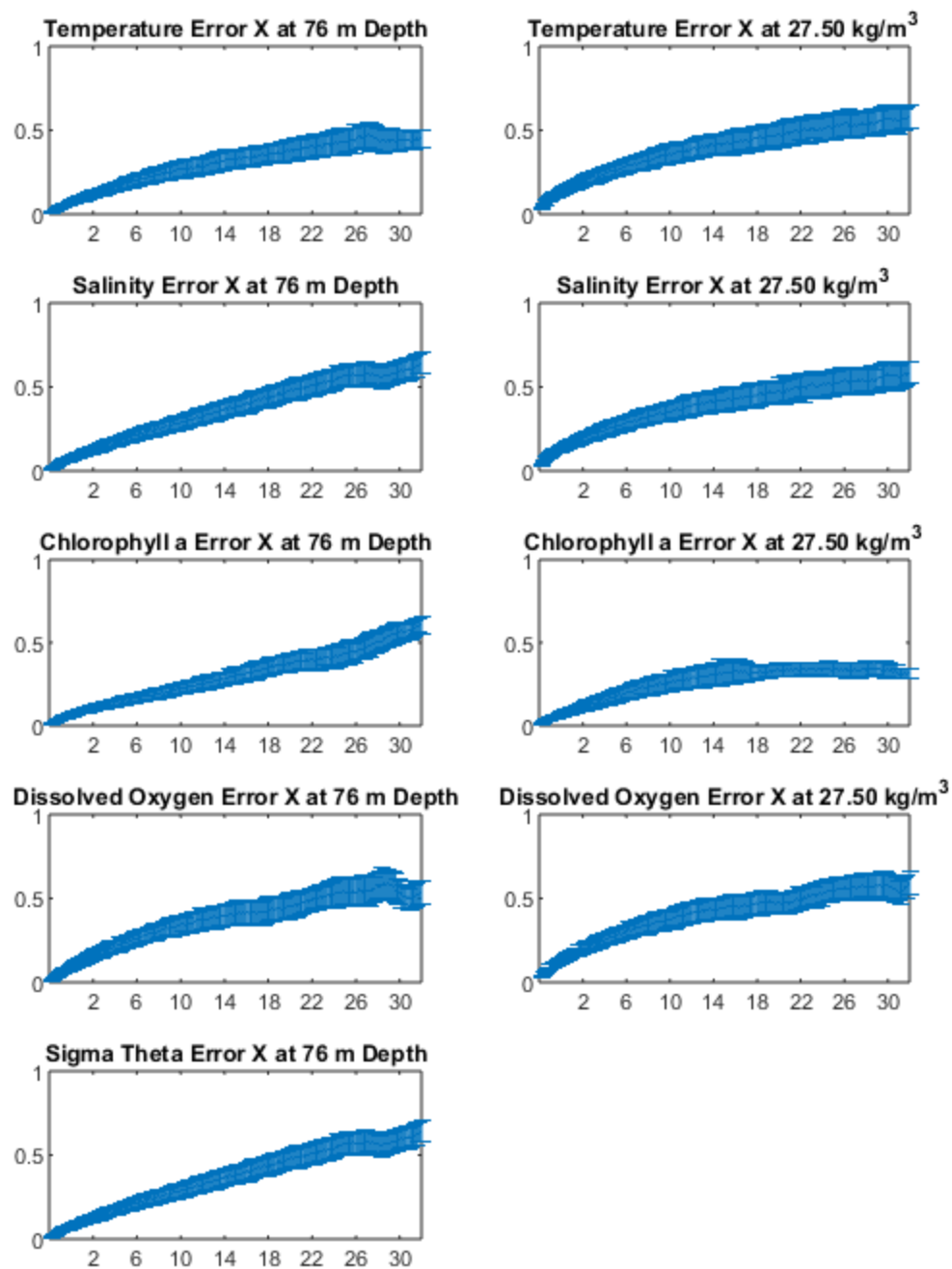


Figure B.5: South Transect Error X Monte Carlo simulation results from 76 m and 27.50 kg/m³.

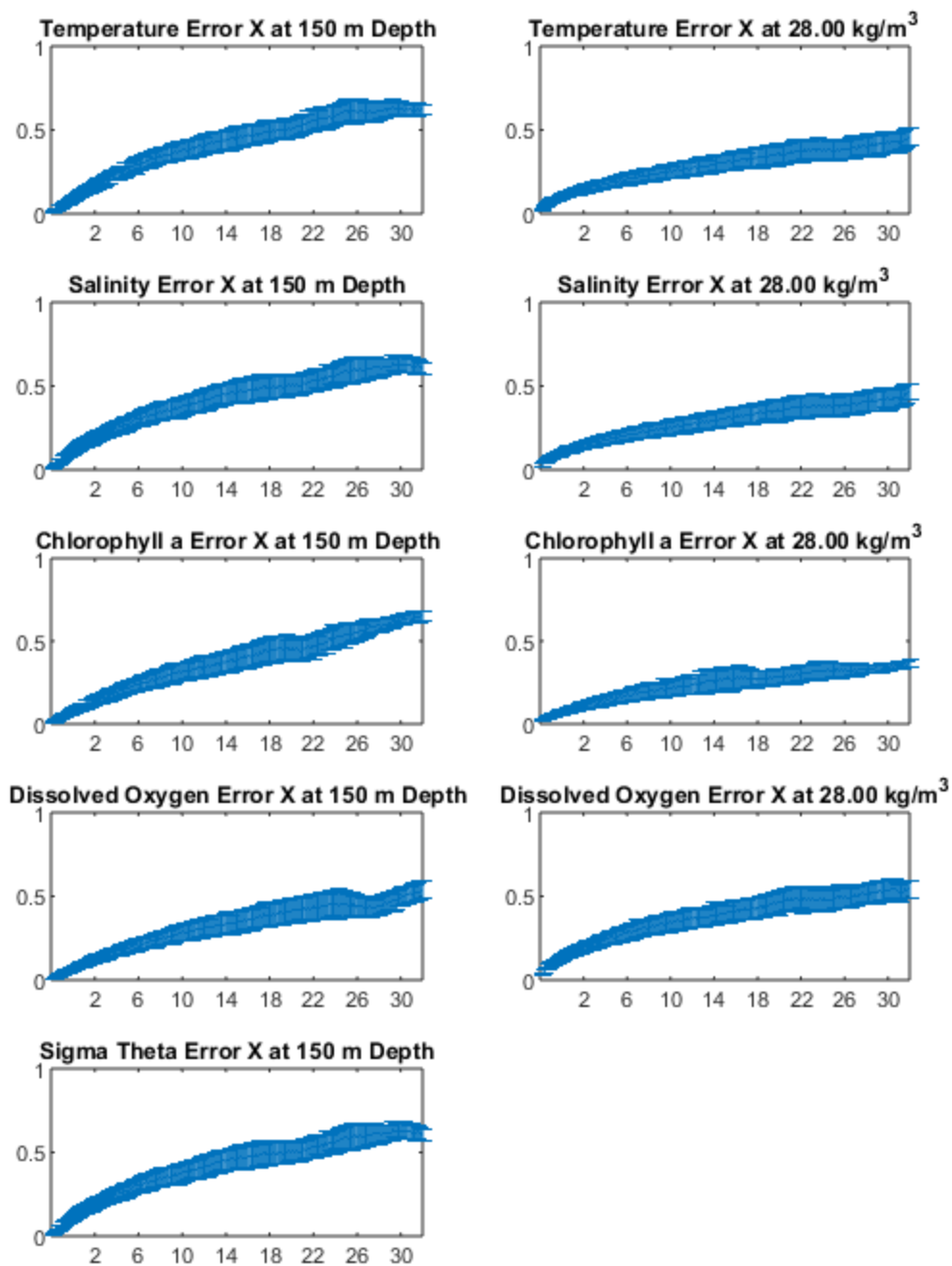


Figure B.6: South Transect Error X Monte Carlo simulation results from 150 m and 28.00 kg/m³.

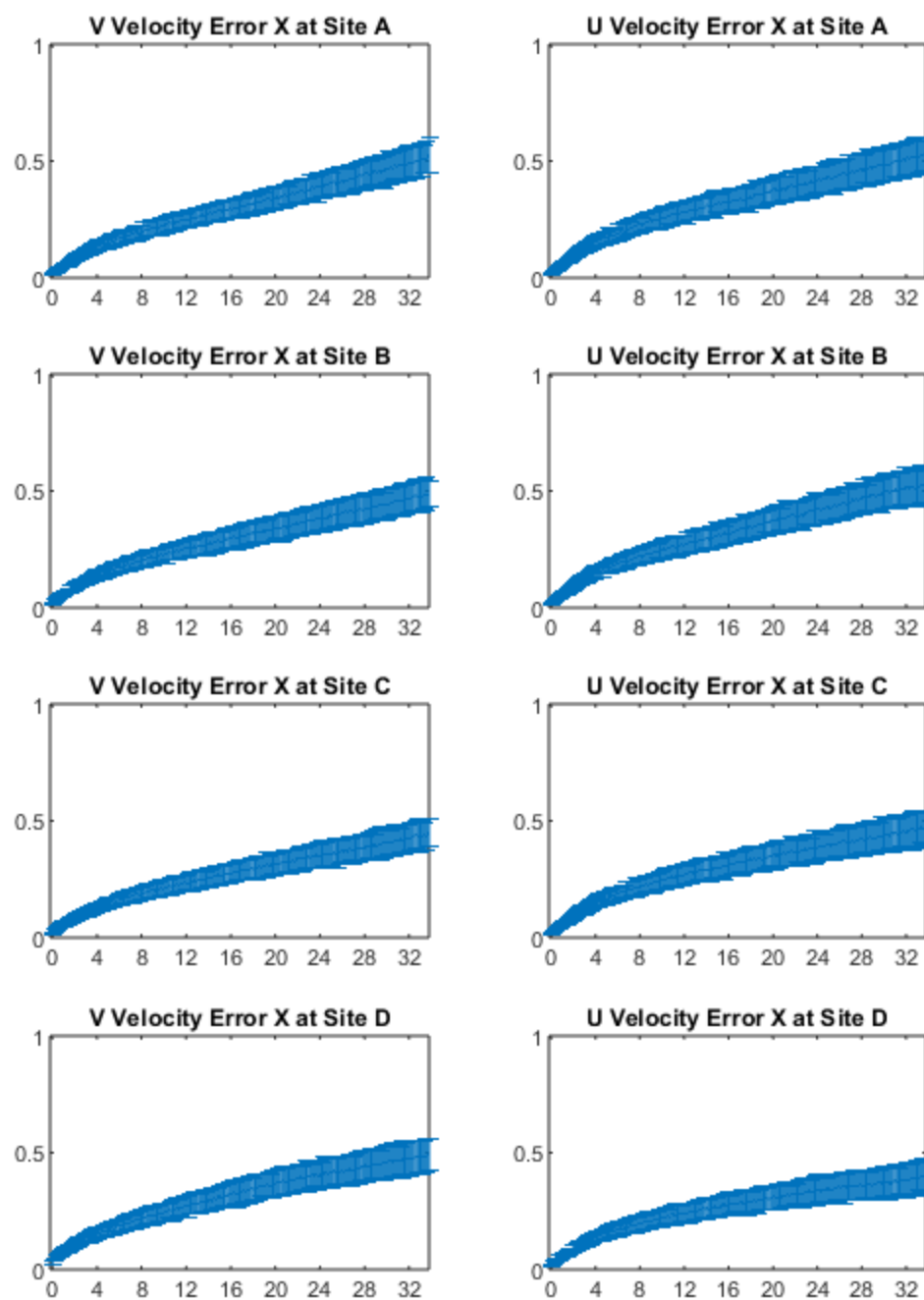


Figure B.7: North Transect Error X Monte Carlo simulation results from HFR.

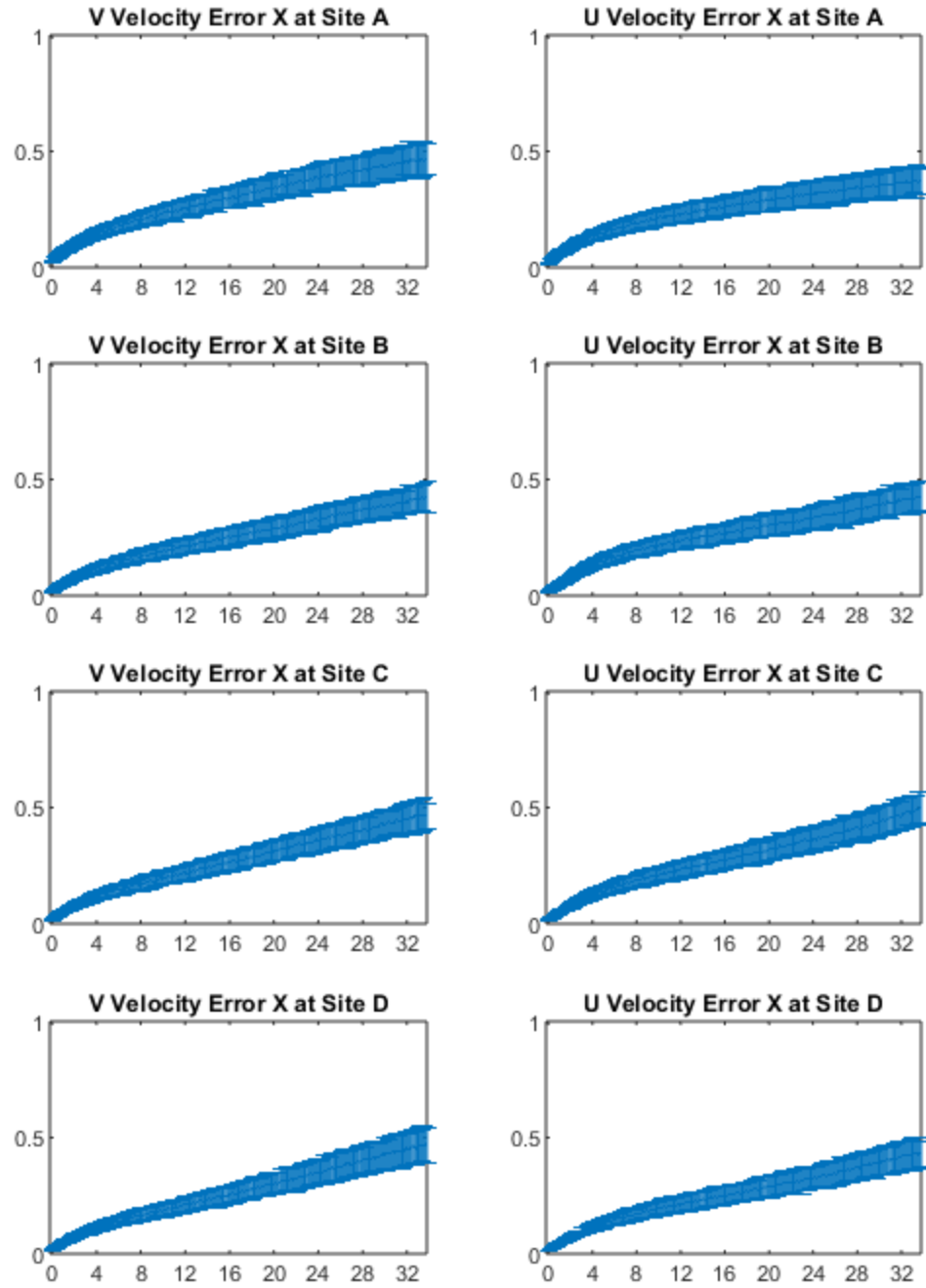


Figure B.8: South Transect Error X Monte Carlo simulation results from HFR.