Spring 5-1-2021

Assessing the Real-Time Lagrangian Predictability of the Operational Navy Coastal Ocean Model in the Gulf of Mexico

Lea Kristen Locke
University of Southern Mississippi

ASSESSING THE REAL-TIME LAGRANGIAN PREDICTABILITY OF THE OPERATIONAL NAVY COASTAL OCEAN MODEL IN THE GULF OF MEXICO

by

Lea Locke

A Dissertation
Submitted to the Graduate School,
the College of Arts and Sciences
and the School of Ocean Science and Engineering
at The University of Southern Mississippi
in Partial Fulfillment of the Requirements
for the Degree of Doctor of Philosophy

Approved by:

Dmitri Nechaev, Committee Chair
Maarten Buijsman
Stephan Howden
Igor Shulman
Michael Toner

May 2021

Published by the Graduate School

ABSTRACT

This study quantitatively assesses the drift predictive skill of Fleet Numerical Meteorology and Oceanography Center's (FNMOC's) operational ocean models which are used to support a wide range of military and civilian applications. Overall, the findings of this work support the recommendation of spatial filtering for regional-scale ocean model velocity fields used in deep-water drift applications. In conjunction with filtering, the use of a pure particle drift algorithm is suggested for short-term forecasts and a drift algorithm including a sub-grid scale, random flight, parameterization for predictions requiring extended forecast predictions.

Drift prediction skill is quantified through metrics of in-cloud percentage, distance error, and cloud size, which are used to assess the impact of different drift algorithms and underlying ocean models on the drift prediction capability. Through an exploration of parameterization additions to the drift algorithm, spatial filtering of model velocity fields, and increases in model horizontal resolution, drift prediction skill is shown to be counter-balanced on the accuracy of the model's dispersive characteristics along with the accuracy of the underlying model velocity field (i.e. data-constrained, predictable features). A regional scale model at a horizontal resolution typically employed by FNMOC (3-kilometers) is found to be grossly under dispersive, and derived drift predictions using a pure particle algorithm are not skillful in terms of in-cloud percentage beyond a 24-hour forecast. Parameterization additions (i.e. sub-grid scale velocity and Leeway), which enhance model dispersion, are shown to greatly improve the regional scale model's ability to predict a drift cloud that encompasses an object of interest at longer forecast lengths (> 24-hours) by increasing cloud size. Increasing the

model's horizontal resolution (500-meters) is likewise shown to improve in-cloud prediction performance at all forecast lengths, due to its more accurate representation of dispersion which results in much larger cloud size predictions compared to those from a regional scale model. Spatial filtering of regional scale velocity fields using a Gaussian filter removes uncertain, unpredictable features (i.e. submesocale) leaving behind a data-constrained velocity field. Even though spatial filtering suppresses dispersion further for an already under-dispersion regional scale model, filtering is shown to significantly improve drift prediction performance extending in-cloud skill farther into the forecast, reducing distance errors by 15-20%, and reducing cloud size predictions by 20-30%.

ACKNOWLEDGMENTS

First and foremost, I would like to offer my sincerest thanks to each of my committee members. Not only did they give freely of their time to read and re-read this document, but the substantive comments and suggestions they gave made this work so much better and stronger. A special thanks goes to my committee chair, Dr. Nechaev, who has acted as my teacher and advisor for more than 10 years now; to say I have learned an incredible amount from him would be an understatement.

I'd also like to acknowledge my Naval Oceanography community. As a long-time employee of the Naval Oceanographic Office and now Fleet Numerical, my job has offered me such incredible context in which to apply the knowledge that I've gained throughout my graduate studies. My work in operational oceanography has made me a better student, and my time at DMS has undoubtedly helped in my career. Additionally, my work family has offered nothing but encouragement throughout my (somewhat prolonged) journey through this Doctoral program. I will be forever grateful for their gentle (but persistent) inquiries and reassuring conversations!

DEDICATION

I dedicate this work to my family. To my husband, Shane, who has supported me through these LONG years. Thank you for always reminding me that I was more than capable of finishing this endeavor when I began to doubt myself. Thank you for helping me recognize that my mountain might really just be a mole-hill! I love you babe, and there's no one I'd rather have by my side. To my son, Cohen, you have sacrificed the most and I want to say thank you for giving up so much 'mama' time. I have been a student since the day you were born, and I know I've missed games, movies, and memories while I've researched, scripted, and typed these words. My hope is that from this degree you learn that hard-work and commitment really do pay off. To my baby girl, Layton. Even in today's time, people will make assumptions about your interests and capabilities just because you are a girl. I hope that you can look to this accomplishment and know that you can do anything you set your mind to! And lastly, to my mama – I can never thank you enough. Not once in my life did you ever doubt my ability to achieve, and I learned to believe in myself because I knew that would always be standing behind me telling me that I could do it.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF ILLUSTRATIONS

LIST OF ABBREVIATIONS

ADCP	Acoustic Doppler Current Profiler
ADT	Absolute Dynamic Topography
AOML	Atlantic Oceanographic and Meteorological Laboratory
CARTHE	Consortium for Advanced Research on Transport of Hydrocarbon in the Environment
CI	Confidence Interval
CMEMS	Copernicus Marine and Environmental Monitoring Service
COAMPS	Coupled Ocean Atmosphere Mesoscale Prediction System
CODE	Coastal Ocean Dynamics Experiment
CNCOM	Coastal NCOM
CWL	Crosswind Leeway
DBDB	Digital Bathymetry Database
DFT	Discrete Fourier Transform
DOF	Degrees of Freedom
DWL	Downwind Leeway
FNMOC	Fleet Numerical Meteorology and Oceanography Center
GCOOS	Gulf of Mexico Coastal Ocean Observation System
GDP	Global Drifter Program
GEBCO	General Bathymetric Chart of the Oceans
GLAD	Grand LAgrangian Deployment
GOM	Gulf of Mexico
GoMRI	Gulf of Mexico Research Initiative

HF	High-Frequency
HYCOM	HYbrid Coordinate Ocean Model
LHS	Left-Hand Side
LSB	Land Sea Breeze
LSGS	Lagrangian Sub-Grid Scale
LSM	Lagrangian Stochastic Models
MLD	Mixed Layer Depth
MY2	Mellor-Yamada Level 2 Turbulent Closure Scheme
NASA	National Aeronautics and Space Administration
NCODA	Navy Coupled Ocean Data Assimilation
NCOM	Navy Coastal Ocean Model
NDBC	National Data Buoy Center
NIO	Near-Inertial Oscillation
NIW	Near-Inertial Wave
NOAA	National Oceanographic and Atmospheric Administration
NRL	Naval Research Laboratory
OSU	Oregon State University
POM	Princeton Ocean Model
RHS	Right-Hand Side
RNCOM	Regional NCOM
SAR	Search-and-Rescue
SSH	Sea Surface Height
SSHA	Sea Surface Height Anomaly

SVP	Surface Velocity Program
SWOT	Surface Water and Ocean Topography
TKE	Turbulent Kinetic Energy
WMO	World Meteorological Organization

CHAPTER I - INTRODUCTION

1.1 Motivation of the Study

The United States Navy's Fleet Numerical Meteorology and Oceanography Center (FNMOC) runs operational ocean models that provide global coverage of the present and future physical state of the ocean, including current velocity. Hindcasts, nowcasts, and forecasts of ocean currents obtained from the Navy's operational models form the basis for many drift predictions needed to support real-world applications. For example, ocean currents from regional ocean models were used to make drift predictions in both the Deepwater Horizon oil spill in the Gulf of Mexico and to aid humanitarian efforts in the aftermath of the tsunami and subsequent nuclear disaster that devastated Fukushima, Japan. In addition to pollutant spills and search-and-rescue (SAR) applications, there are also many biological uses for drift prediction such as the spread of harmful algal blooms and larval transport. Thus, knowledge of the model accuracy of ocean currents and their associated drift predictions are of crucial importance.

This study develops a method to quantify the accuracy of real-time drift predictions from data-assimilating ocean models of differing resolutions run at FNMOC using an independent dataset provided by drifting buoys. Several parameterizations of physical processes not resolved by the model are also implemented in an effort to determine possible areas of improvement with respect to drift prediction. Regional bulk statistics are computed and analyzed to assess accuracy metrics.

The Gulf of Mexico (GOM) has been identified as the area of interest on which to quantify model drift accuracies, in large part due to an abundant amount of drifting buoy data available in the region. In addition to a robust in-situ dataset, the oil industry is also

very active in the region (Fig. 1.1), making drift prediction in the Gulf of utmost importance to the response and mitigation of adverse impacts of future spills to the Gulf's ecosystems and coastal communities. The oceanographic and atmospheric regime of the study area is discussed in more detail in the next section.

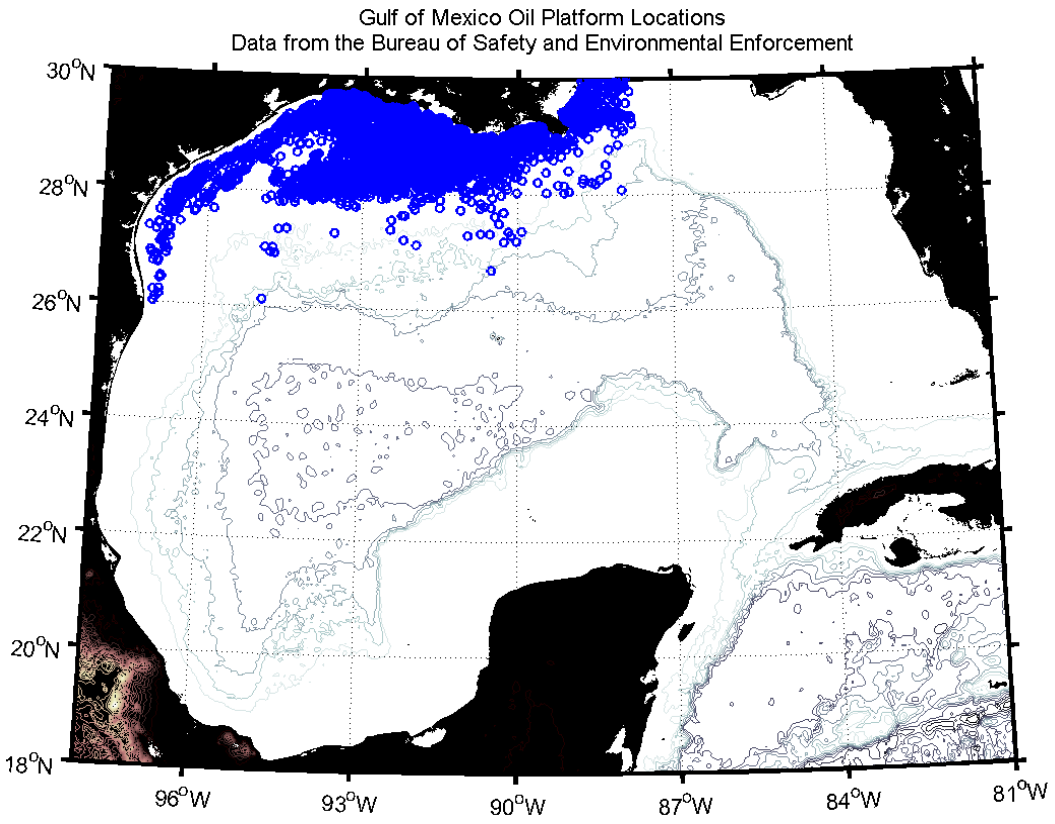

Figure 1.1 *Oil Platform Locations in the Gulf of Mexico*

Locations of drilling rigs and platforms in the Gulf of Mexico. Data was obtained from the Bureau of Safety and Environmental Enforcement and is current as of 10-01-2013 (http://www.data.bsee.gov/homepg/data_center/platform/platform.asp).

1.2 Background Information

1.2.1 Gulf of Mexico Oceanographic and Atmospheric Overview

The deep-water regions of the GOM are dominated by the Loop Current and the associated mesoscale eddies that are shed from this current, so called Loop Current eddies. Anti-cyclonic rings are shed from the Loop Current on average of once every 12

months; however, shedding events do exhibit considerable year-to-year variability. Once an eddy is separated from the main current, it then migrates westward across the basin where it interacts with other anti-cyclonic and cyclonic eddy features (Elliot 1982; Schmitz et al. 2005). Hence, the interior of the GOM can be characterized as a dynamic, eddy-rich region. A daily snapshot of the modeled mesoscale field shown in Fig. 1.2 highlights the dynamic nature of the eddy field in the GOM.

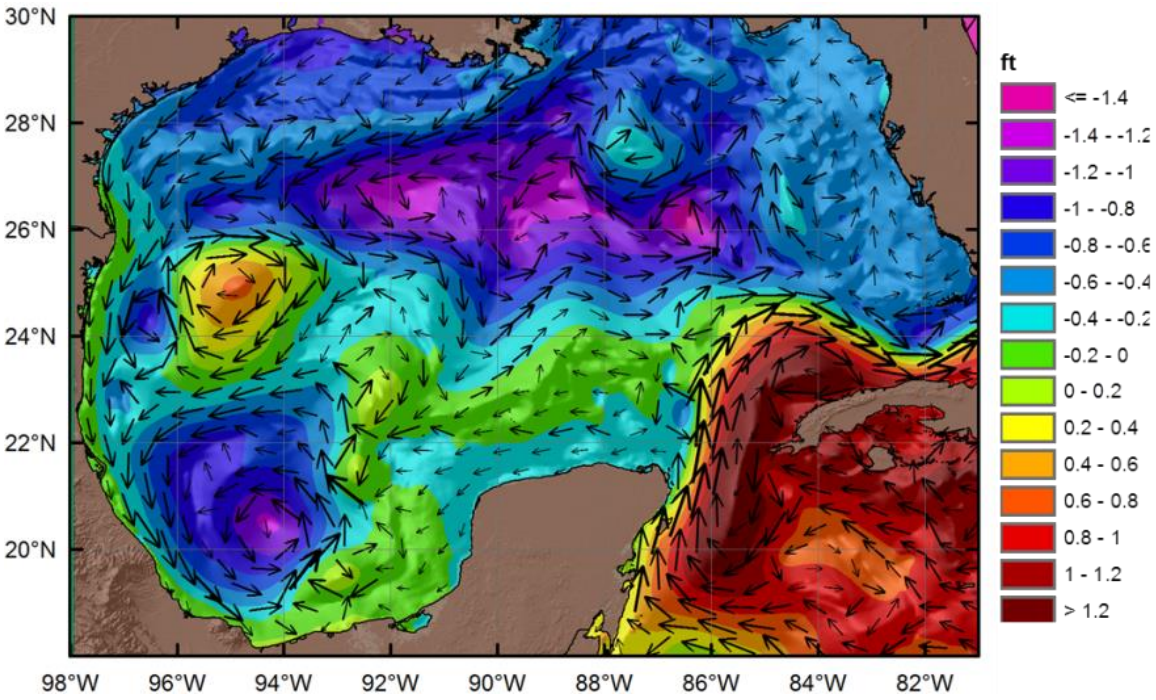

Figure 1.2 *Example of Modeled 24-hr Mean Sea Surface Height*

Taken from the 10 December 2013, 00-24-hour forecast from FNMOC's operational ocean model covering the Gulf of Mexico. The 24-hour averaged sea surface height anomaly (SSHA) is shown overlaid with 24-hour averaged surface currents.

The wind regime over the Northern Gulf Shelf, where the majority of oil platforms are located, is controlled by the strength and position of the Bermuda High pressure system. During the fall, the high-pressure system weakens and moves to the northeast supplying the shelf with winds that are generally easterly. During the spring and summer, this high strengthens and moves to the southwest, and southeasterly winds with

a strong southerly component dominate over the Gulf (Boicourt et al. 1998; De Velasco and Winant 1996). In general, currents over the Northern Gulf Shelf follow the seasonal wind pattern of wind forcing discussed above and are characterized by westward flow along bathymetric contours except during June and July, when currents reverse, becoming eastward (Boicourt et al. 1998; Ohlmann and Niiler 2005). In addition, local wind conditions can force reversals in long-shore currents on synoptic scales, and strong wind events can set up strong long-shore jets (Crout et al. 1984; Ohlmann and Niiler 2005). Eddy features can also intrude onto the shelf regions and have a significant impact on the outer shelf-circulation, driving cross-shelf flow and thus providing a mechanism for shelf and open-water interaction (Elliot 1982; Biggs et al. 2005; Ohlmann and Niiler 2005). Over the Northern Gulf Shelf, it is believed that the inner shelf is predominantly influenced by the weather, while the outer shelf is predominantly influenced by intrusion of mesoscale eddies (Ohlmann and Niiler 2005; Nowlin et al. 2005).

Tides in the GOM are largely diurnal in nature due to the near resonance of the Gulf with diurnal tidal forcing (Seim et al. 1987; Kantha 2005). Along the Northern Gulf shelf, diurnal tides dominate except in Apalachicola Bay where semi-diurnal tides prevail (Kantha 2005), and the dominant constituents are K_1, O_1, and M_2. Maximum M_2 and K_1 currents vary from 6 cm/sec to 1 cm/sec from Atchafalaya Bay to the shelf break, whereas maximum currents for O_1 vary from 5 cm/sec to 1 cm/sec (DiMarco and Reid 1998). Barotropic tidal currents in the deep regions of the Gulf are only a few cm/sec at best and can safely be ignored (Kantha 2005).

Near-inertial waves (NIWs) are common features observed on the shelf and slope in the Northern GOM (Chen 1996; DiMarco et al. 2000; Jarosz et al. 2007). These

oscillatory currents are usually wind-generated features, induced by sudden changes in the surface wind stress associated with local wind forcing (Pollard 1970). Velocities of 10-15 cm/s are typical for NIWs, but NIW velocities can reach as high as ~1 m/s due to the passage of extreme weather events such as hurricanes and typhoons (Hsu 2018). In the northern Gulf of Mexico where the latitude range is near 30°N, NIWs have a near diurnal local frequency. Near-inertial oscillations occurring around the critical latitudes, 30°S and 30°N, can resonate and amplify in the presence of any near-diurnal forcing, i.e. diurnal tides and land sea breezes (LSBs) (Jarosz et al. 2007; Zhang 2009; Gough et al. 2016). Likewise, stronger inertial oscillations are associated with the summer months and shallower mixed layers (DiMarco et al. 2000; Jarosz et al. 2007). Because oscillation amplitudes are tied to mixed layer depth, freshwater discharge can have a significant impact on NIWs (Gough et al. 2016). As the Northern GOM has several freshwater discharge inlets, this aspect is likely an important aspect of NIW dynamics for this region.

1.2.2 Ocean Model Overview

The Princeton Ocean Model (POM) based Navy Coastal Ocean Model (NCOM) is the numerical ocean model that forms the basis for modeled drift predictions. NCOM is based on the primitive equations and employs traditional approximation for Coriolis acceleration, Boussinesq, and hydrostatic approximations. Two implementations of NCOM are used to determine the Navy's operational drift predictive skill. A regional-scale implementation, Regional NCOM (RNCOM), encompasses the entire GOM and Caribbean Seas and has a horizontal resolution of 1/36° (~3 kilometers). A higher resolution Coastal NCOM (CNCOM) is nested within the RNCOM domain and has a

horizontal resolution of 500 meters. Domain boundaries of the RNCOM (blue) and CNCOM (red) used in the evaluation are shown in Fig. 1.3.

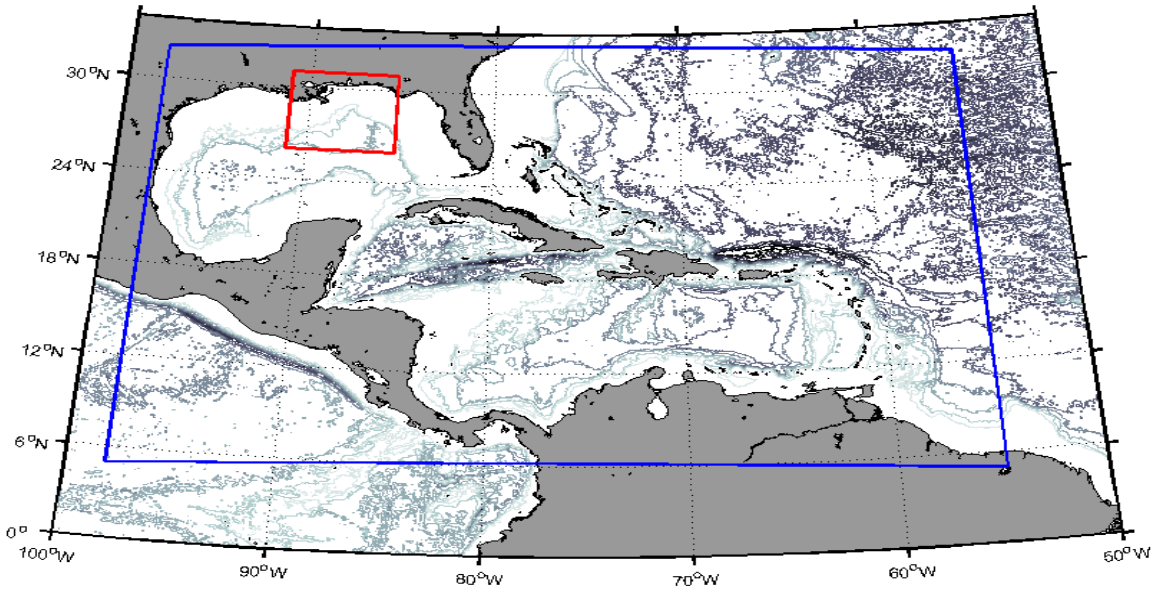

Figure 1.3 *Ocean Model Domain Coverage*

Model domain for the RNCOM covering the Gulf of Mexico region is shown by the outer blue box. The model domain for the CNCOM covering the northern Gulf shelf is shown by the inner red box.

NCOM uses a hybrid σ-z vertical coordinate system, with results calculated using 50 vertical levels, 35 of which are shallow water σ-surfaces and the remaining levels are z-levels. The σ-surfaces provide the higher vertical resolution needed to resolve important smaller scale processes. While computations are done using 50 vertical levels, model output is interpolated to 40 standard z-levels. Horizontal mixing coefficients are calculated following the Smagorinsky (1963) parameterization, and vertical mixing coefficients are calculated using the Mellor-Yamada Level 2 turbulence closure scheme (Barron et al. 2006). Both NCOM implementations assimilate data from in-situ and remotely sensed sources in near real-time through a data assimilation system, Navy

Coupled Ocean Data Assimilation (NCODA), employing a 3D variational method (Lunde and Coelho 2009).

RNCOM's bathymetry and coastline are from the Naval Research Laboratory's (NRL's) 2-minute Digital Bathymetry Database (DBDB) dataset. CNCOM's bathymetry and coastline are from the General Bathymetric Chart of the Oceans (GEBCO), which is 30-arc second resolution. FNMOC's Coupled Ocean Atmosphere Mesoscale Prediction System (COAMPS) provides the atmospheric forcing for both models and has a horizontal resolution of 15-kilometers. Fresh water input is provided through a monthly climatology of river runoff from the World Meteorological Office (WMO). At its open boundaries, RNCOM is forced with global HYbrid Coordinate Ocean Model (g-HYCOM) conditions, and RNCOM in turn provides the open boundary forcing for the nested CNCOM. Both models receive tidal forcing from the Oregon State University (OSU) global tidal database.

1.2.3 Objective of the Study

FNMOC routinely provides drift prediction support for a variety of real-world events, such as mine drift, search-and-rescue operations, and oil-spill cleanup and containment efforts (Fig. 1.4). These predictions play a vital role in successful mission planning and execution surrounding such real-world events; however, accuracy of FNMOC's drift prediction has not been thoroughly analyzed against available observations, and no robust metrics for evaluating the quality of model drift prediction currently exist. This study aims to provide a quantitative estimate of FNMOC's operational drift prediction accuracy and investigate ways in which this accuracy might be improved: 1) through the inclusion of offline (added after completion of the model

run) parameterizations in the drift algorithm, for sub-grid scale velocity, Stokes drift, and Leeway drift, processes known to be missing or inadequate in the model, 2) through filtering of model velocity fields, removing unconstrained features present in the model solution, and 3) through increasing the horizontal resolution of the model, and thereby shifting the numerically unresolved processes to smaller scales. Additionally, conclusions regarding possible improvements to model physics and forcing will be drawn from the study's findings.

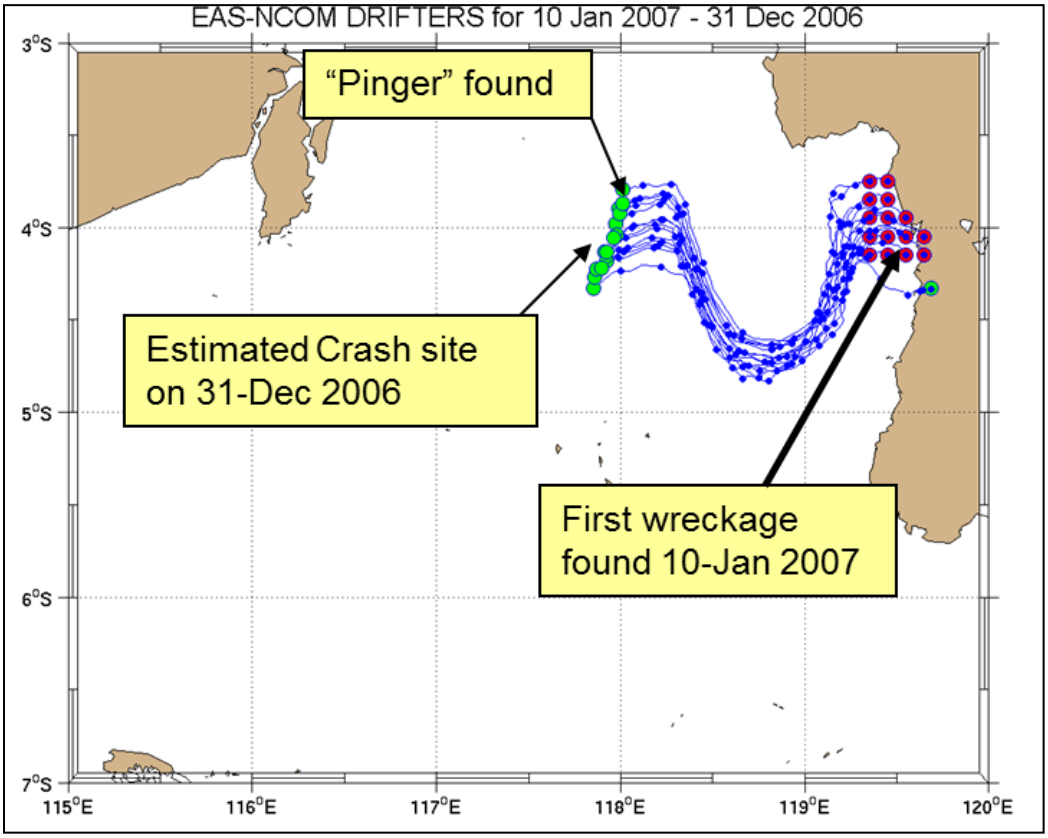

Figure 1.4 *Real-World Drift Prediction Support Product*

Real-world drift prediction used to support efforts to find the wreckage site of a commercial aircraft that crashed off the coast of Jakarta in 2006. East Asian Seas (EAS) NCOM was used to hindcast drift trajectories from where debris was found on January 10th, 2007 to the date of the crash on December 31st, 2006. The aircraft's pinger was found very near where EAS NCOM estimated the crash site to be.

CHAPTER II – RESEARCH TOOLS

2.1 Data Overview

2.1.1 Technical Description of Buoys

Two types of drifting buoys comprise the validation dataset employed in this study, Coastal Ocean Dynamics Experiment (CODE) drifters and Surface Velocity Program (SVP) drifters. CODE drifters have a cruciform-shaped drogue with four rectangular vanes extending radially from a vertical tube, and drogues are submerged at about 1-meter depth (Fig. 2.1). SVP drifters consist of a circular surface float, and a subsurface drogue, which is centered at 15-meters water depth (Fig. 2.2). Each drifter

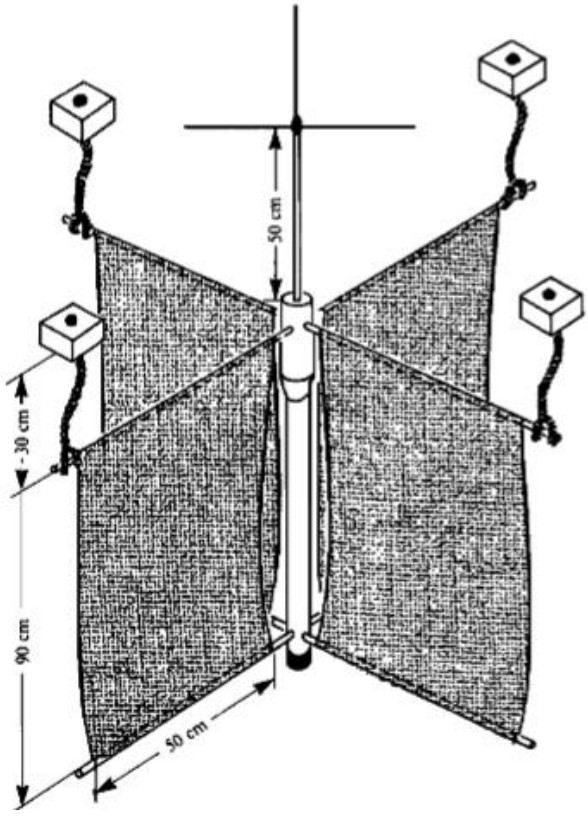

Figure 2.1 *Schematic of a CODE Surface Drifting Buoy*

The vertical tube contains batteries, the transmitter, and supports the antennae. Graphic taken from Davis (1985a).

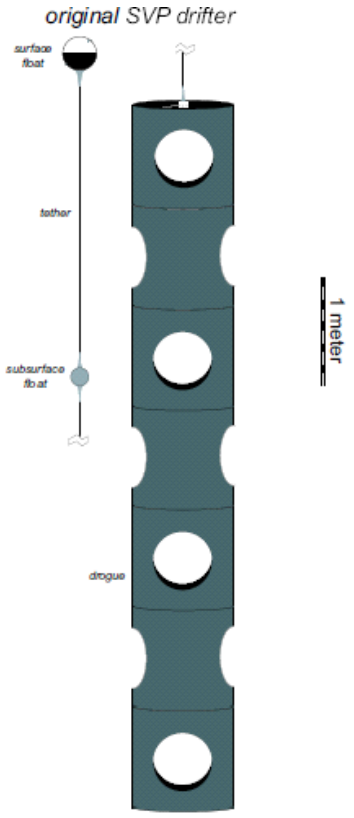

Figure 2.2 *Schematic of an SVP Surface Drifting Buoy*

Much of the tether length has been excluded. The drogue is centered at 15m depth. Graphic taken from Lumpkin and Pazos (2007).

type was designed to be as “water-following” as possible (Niiler et al. 1987; Davis 1985b); however, both designs are associated with some "slippage", which is defined as the horizontal motion of a drifter that differs from the average horizontal current taken over the drogue depth. Windage effects on the surface-exposed portion of the buoy, drag on the float and/or tether, and rectification of surface waves are all causes of drifter slippage (Niiler et al. 1987; Geyer 1989). Drifter slippage for SVP drifters was found to be 0.1% of the wind speed (Niiler et al. 1987; Niiler et al. 1995). CODE slippage was generally estimated by Davis (1985a) to be accurate to about 3 cm/s even under strong wind conditions. A more recent study (Poulain et al. 2019) was able to further refine the

CODE estimates and found slippage to be 0.1% of the wind speed, similarly to the SVP drifter.

2.1.2 Description of the Dataset

The Gulf of Mexico (GOM) drifting buoy data used in this study were obtained from two sources. CODE drifting buoy data are from the Consortium for Advanced Research on Transport of Hydrocarbon in the Environment's (CARTHE's) Grand LAgrangian Deployment (GLAD) experiment dataset downloaded from the Gulf of Mexico Research Initiative (GoMRI) site (Özgökmen 2013). SVP drifting buoy data are from the Global Drifter Program (GDP) Drifter Data Assembly Center maintained by the National Oceanographic and Atmospheric Administration's (NOAA's) Atlantic Oceanographic and Meteorological Laboratory (AOML) (Lumpkin 2019).

The GLAD dataset is comprised of 297 trajectories from CODE-type drifting buoys drogued at 1-meter depth, and the dataset spans July 22, 2012 through October 31, 2012. Most of the GLAD drifters were launched as triplets separated by roughly 100-meters at launch in an effort to measure multi-scale near-surface dispersion. Drifter positions in the GLAD dataset were low-pass filtered, using a 1-hour cutoff period, and interpolated to uniform 15-minute intervals starting on whole hours (Özgökmen 2013). From the AOML database, 11 SVP drifting buoy trajectories were retrieved in the Gulf of Mexico covering the same time period as the GLAD experiment data (Lumpkin 2019). AOML applies quality control procedures to drifter positions and interpolates them to 6-hour intervals using an optimum interpolation procedure called kriging (Hansen 1995). Both the GLAD and SVP datasets used in this study were quality controlled and post-

processed using the above described methods at the time of data download.

The time period for this study was chosen to coincide with the extensive dataset available from the GLAD deployment. However, during the middle of the GLAD experiment, Hurricane Isaac tracked through the region, entering the Gulf on August 27, 2012 and making landfall on August 29, 2012 (Fig. 2.3). Due to COAMPS underestimation of storm winds (Table 2.1) and the substantial windage effects that were surely felt by the surface drifters during storm passage, model to observation comparisons will exclude drifting buoy data collected from August 27 through September 14, 2012. While Hurricane Isaac made landfall on August 29, the exclusion dates were extended by two weeks after landfall in order to avoid the near inertial oscillations (NIOs), which are created by severe weather events and have been show to persist in the GOM for 6-8 days after the passage of a strong storm (Zhang 2013).

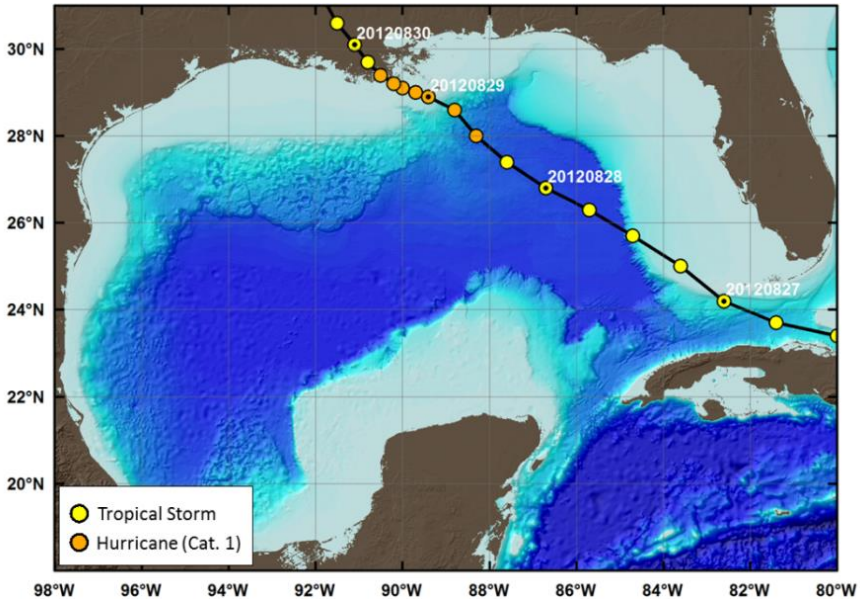

Figure 2.3 *Hurricane Isaac's Track through the Gulf of Mexico*

National Hurricane Center (NHC) best track information for Hurricane Isaac. http://www.nhc.noaa.gov/data/hurdat/hurdat2-1851-2015-070616.txt

Table 2.1 *Maximum Wind Speed during the passage of Hurricane Isaac*

Date	Time	Latitude	Longitude	Designation	Max Observed Wind Speed (kts)	Max Predicted Wind Speed (kts)
20120827	0000	24.2N	82.6W	Tropical Storm	50	46
20120827	1200	25.7N	84.7W	Tropical Storm	55	47
20120828	0000	26.8N	86.7W	Tropical Storm	60	52
20120828	1200	28.0N	88.3W	Hurricane	65	54
20120829	0000	28.9N	89.4W	Hurricane	70	58
20120829	1200	29.4N	90.5W	Hurricane	65	43
20120830	0000	30.1N	91.1W	Tropical Storm	55	34

Note: Information taken from: http://www.nhc.noaa.gov/data/hurdat/hurdat2-1851-2015-070616.txt

After removal of data points coinciding with the passage of Hurricane Isaac, drifter tracks showing data gaps of 12-hours or more were split into separate tracks. Subsequent tracks were retained for validation only if they were 48-hours in length or greater. A total of 453 drifter tracks from the GLAD dataset (Fig. 2.4) and 19 drifter tracks from AOML's SVP dataset (Fig. 2.5) were retained in the GOM for inclusion in this study.

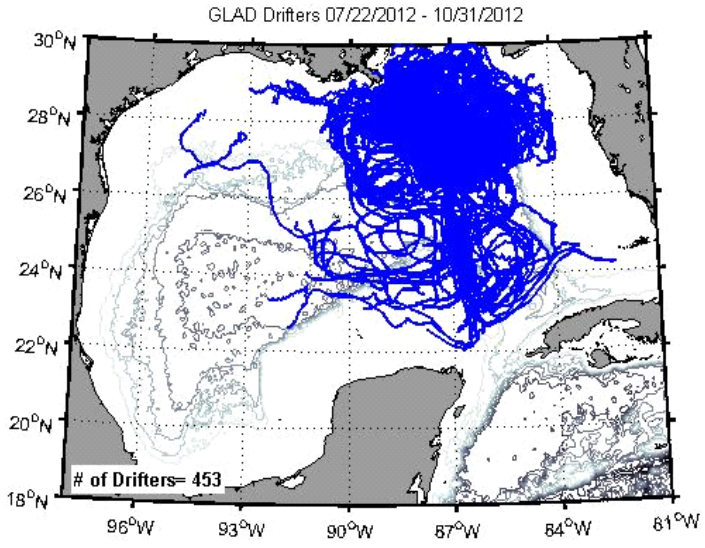

Figure 2.4 *GLAD Drifter Dataset Used in Study*

GLAD drifter deployment data in the Gulf of Mexico, 7/22/2012 through 10/31/2012.

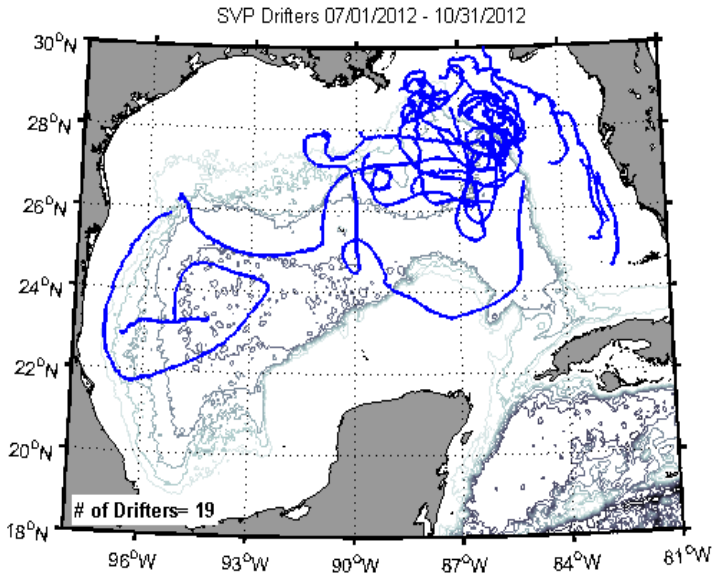

Figure 2.5 *SVP Drifter Dataset Used in Study*

SVP drifter data in the Gulf of Mexico, 07/01/2012 through 10/31/2012.

2.2 Methods

The operational ocean models that run at FNMOC are implemented using an Eulerian-based data assimilation scheme. However, the problem of drift prediction requires Lagrangian analysis of the Eulerian model fields. The difficulty of such a task arises from the vast amount of inherent uncertainty surrounding ocean model solutions and in turn modeled drift predictions (Griffa et al. 2013). Lagrangian motion, even for an idealized flow, can exhibit chaotic behavior (Aref 1984; Samuelson 1996). In the ocean, the combination of time dependence and a complex three-dimensional flow structure is likely to induce chaotic Lagrangian transport (Samuelson 1992; Meyers 1994; Yang and Liu 1997). A system exhibiting chaotic behavior is highly sensitive to initial conditions, yielding significantly different trajectories with only small perturbations to the initial condition field. In the ocean, this sensitivity is largely governed by mesoscale features such as eddies and jets (Haller and Poje 1998; Poje and Haller 1999), and the

characterization of eddy and jet features in ocean model solutions are highly uncertain due the sparse nature of oceanographic observations available to constrain feature strength and placement.

Unresolved features, inadequate parameterizations (e.g. horizontal and vertical mixing parameterizations, turbulent closure), and uncertainty in atmospheric or boundary forcing are other model related issues that can and do lead to more sources of errors in the ocean model's prediction fields. In the operational case, data-assimilation further compounds the uncertainty problem by introducing small and sometimes major adjustments to the placement and strength of eddies and jets, which certainly affects trajectory predictions. Additional uncertainty arises from the fact that most practical applications of drift prediction, such as contaminant drift and search-and-rescue operations, all deal with objects that are not passive and are subject to additional forces like windage and drag, leading to yet another source of uncertainty. Due to the above-mentioned issues, direct drifter to model comparisons tend to fare poorly (Barron et al. 2007; Price et al. 2006; Caballero 2008) and point to the necessity of a probabilistic approach to drift prediction. In what follows, a proposed methodology for one type of probabilistic approach, a drift probability map, is outlined, and this probability map is the product that will be used to evaluate model performance with respect to drift predictability in this study

2.2.1 Drift Probability Map Overview

To create a drift probability map, an initial array of modeled particle positions is integrated forward in time using a 4^{th} order Runge-Kutta scheme, with an 18-minute timestep (10 interpolations of velocity from one model time to the next) and linear time-

space interpolation. Particles are launched every 3-hours (a typical model output interval) over the duration of model availability using this procedure, and drift trajectory positions are saved at 24-hour increments for every initial particle in the array. FNMOC's operational models provide nowcasts and forecasts out to 96-hours on a daily basis, and multiple realizations of the solution are available for any given day containing perturbations of the initial conditions due to things such as changes in forcing and adjustments of the velocity field caused by data assimilation. For this study, the model duration over which particles are released includes the model run at the analysis date as well as model solutions from the previous three days (Fig. 2.6). The subsequent Lagrangian analysis is based on the concept of "hours after release." For a particular forecast scenario (e.g. 96-hours after release), the area occupied by modeled particles is divided into square bins of constant width, and particle positions located within each bin are counted. Each bin is then ranked accordingly as high or low probability areas (Fig. 2.7).

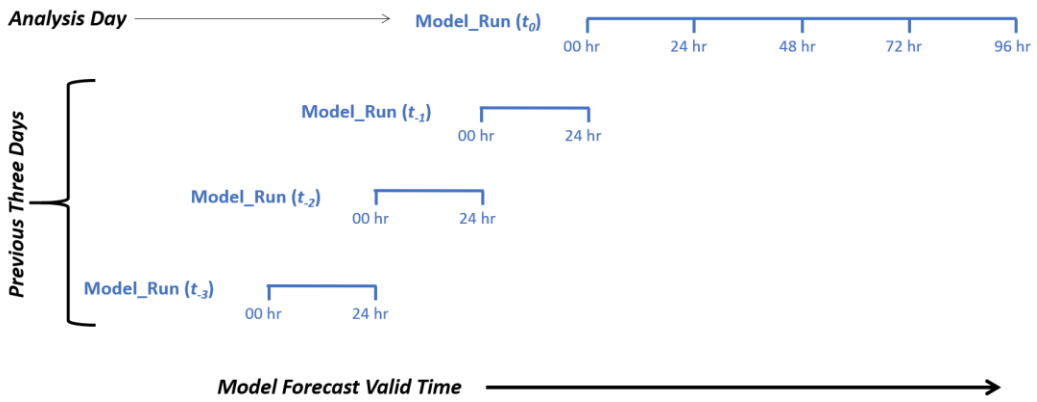

Figure 2.6 *Depiction of Model Data Used in Probability Map Creation*

Graphical depiction of the model data used to construct each probability map. Particles are launched every three hours and integrated out 96-hours using the first 24-hours of forecast from the previous three days of model runs, along with the entire model forecast from the analysis day.

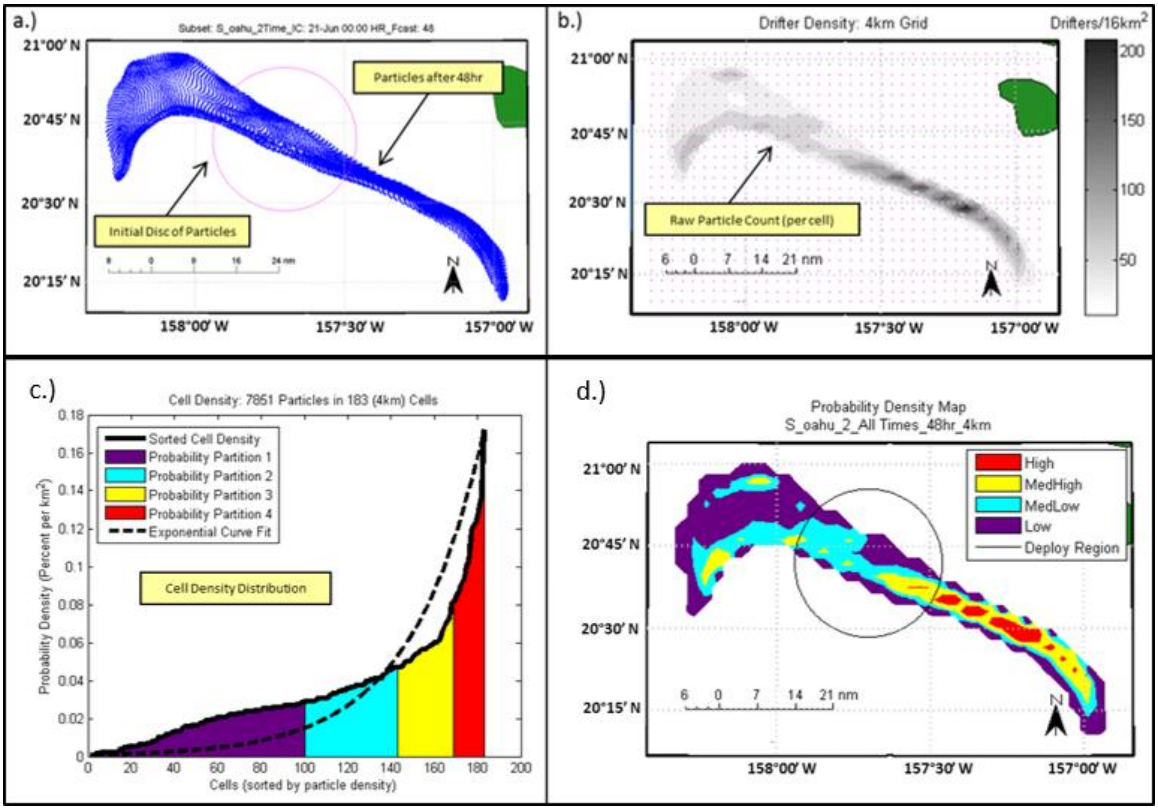

Figure 2.7 *Example of Drift Probability Map Creation Method*

Method used to go from final particle positions to a probability map. a.) Particle positions after a 48-hour drift integration from the initial disc of seeded particles (pink circle); b.) raw particle counts for each grid cell; c.) sorted and ranked grid cell information, where colors indicate the number of cells that constitute four consecutive 25% quantiles; d.) spatial depiction of sorted and ranked cells, where purple indicates the lowest probable grid cells (lowest density cells that together constitute 25% of particles) and red indicates the highest probable grid cells (highest density cells that together constitute 25% of particles).

To determine the accuracy of this probabilistic product, the initial array of modeled particles is centered on a position along a drifting buoy track so that the accuracy of the prediction can be assessed by comparing the position of the drifting buoy at the forecast time to the area covered by the probability map. This is done at multiple time moments along the same drifting buoy path to increase the number of validation comparisons, making the statistical measures more robust. For this study, a comparison is made every 24-hours along each available drifting buoy from both the GLAD and SVP datasets (Fig. 2.4 and 2.5). An example of this validation technique is shown for a

drifting buoy located in the central GOM, sampling an eddy feature, during August 2010 (Fig. 2.8). While the actual drifting buoy path falls outside of the predicted highest probable area, the model captures the same anti-cyclonic drift tendency seen in the observed drifter track.

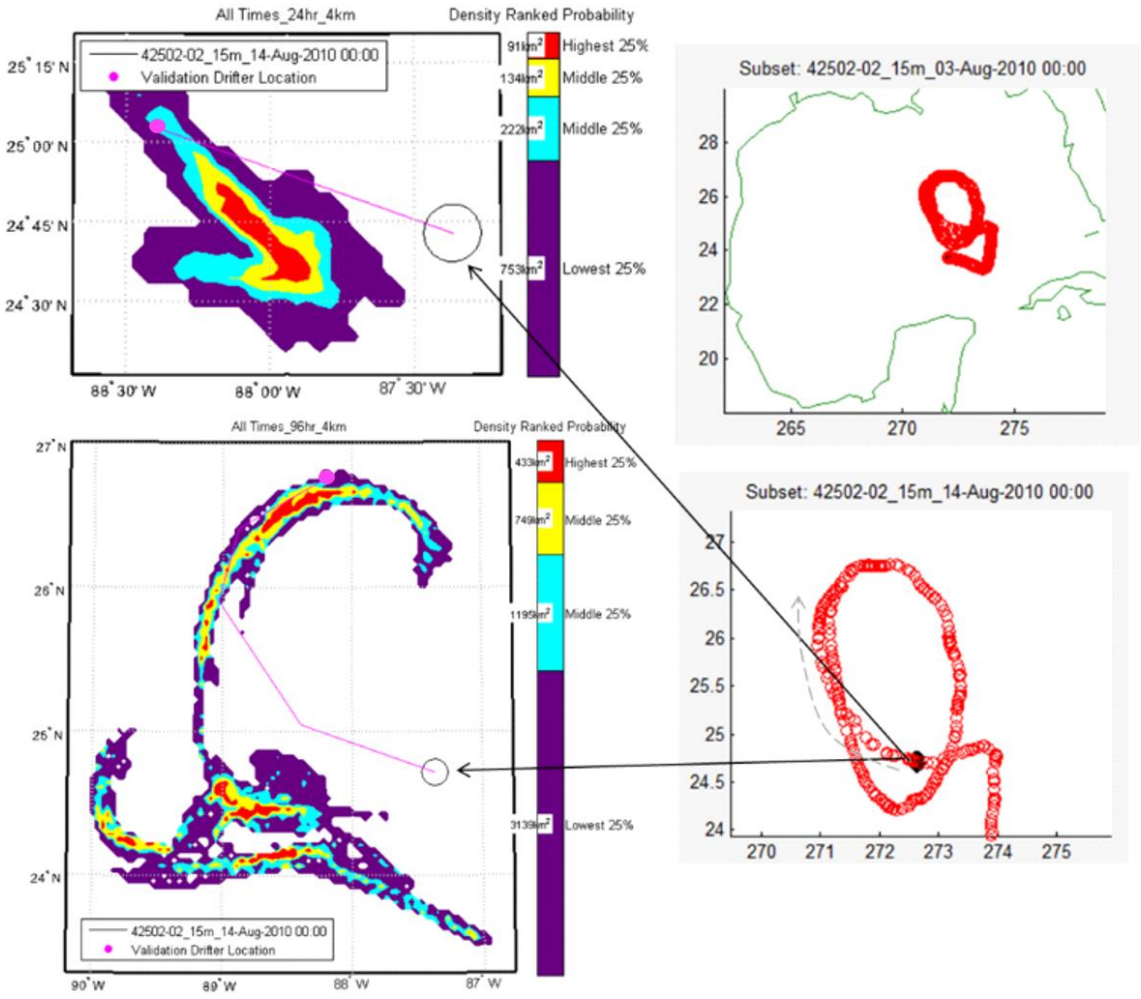

Figure 2.8 *Example of Drift Probability Map Validation*

Right hand panel: Initial launch locations for the modeled drift prediction are denoted in black and are centered on a position along the drifting buoy track (red). Left hand panel: Drift probability maps are shown for the 24-hour forecast (top) and 96-hour forecast (bottom) along with the initial (black circle) and ending drifting buoy position (pink dot) at the specified forecast hour. Drift predictions are produced from RNCOM.

2.2.2 Filtering of Model Velocity Fields

A recent study done by the Naval Research Laboratory (Jacobs et al. 2019) was able to show that spatial filtering of ocean model velocities leads to a decrease in drift trajectory errors. This enhanced performance is attributed to the filtering out of unconstrained features in the ocean model solution, which are defined as those that do not have sufficient observations for a skillful forecast. The threshold between constrained and unconstrained scales is determined by the observations that are available to the data assimilation. The principal source of consistent, large-scale observations for the ocean comes from altimeter measurements, which are the primary mechanism by which mesoscale features are corrected in FNMOC's operational ocean models. Jacobs et al. (2019) spatially filtered model velocity fields in the Gulf of Mexico using a Gaussian filter with ¼ power scale from 20 km to 300 km in 20 km increments and were able to demonstrate that removing features smaller than ~200 km in modeled velocity fields minimized predicted trajectory errors, reducing them by 20%, as compared to a large, surface drifting buoy dataset.

A similar filter as that used by Jacobs et al. (2019) is applied to the model data in this study to see if comparable trajectory error decreases can be obtained. The filter employed is a Gaussian filter with an e-folding scale of 58 km, corresponding to ¼ power scale of 220 km. In areas influenced by land, an average of velocity is taken using the kernel, the Gaussian convolution matrix, as a weighting matrix. An example of modeled velocity and corresponding filtered velocity fields is shown in Fig. 2.9, where it is readily observed that the major result of filtering is the removal of small-scale features in the model result. Likewise, the impact that filtered velocity has on drift prediction is shown

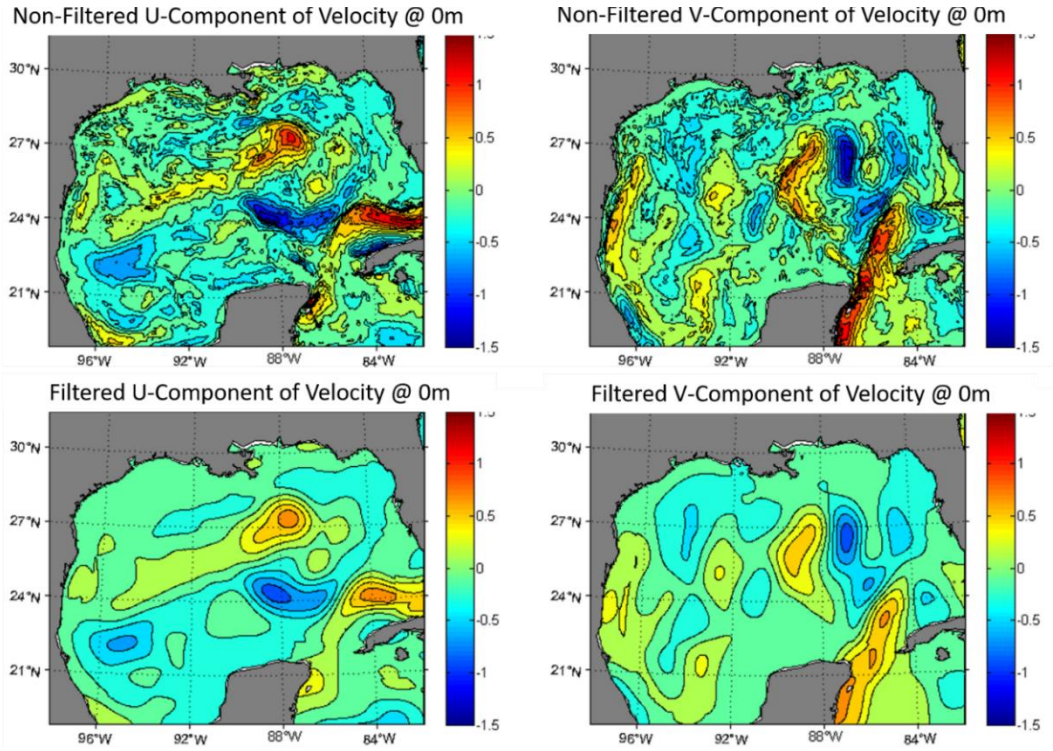

Figure 2.9 *Example of Filtering Applied to Model Velocities*

Top Left: Unfiltered, u-component of surface velocity from RNCOM model output. Top Right: Unfiltered, v-component of surface velocity from RNCOM model output. Bottom Left: Filtered, u-component of surface velocity from RNCOM model output. Bottom Right: Filtered, v-component of surface velocity from RNCOM model output. A Gaussian filter with e-folding scale of 58 km is used for filtering. Where data points are influenced by land, a weighted average of velocity values is taken where the Gaussian kernel acts as the weighting function.

in Fig 2.10. From this figure it can be seen that the major difference in the predictions using unfiltered and filtered velocity fields is that the stretching and filamenting behavior seen in the unfiltered case is greatly suppressed after filtering, which is an intuitive result considering the smoothed nature of the filtered velocity field, with no small-scale features present in the prediction. The suppression of drift cloud stretching, or dispersion as it is later defined, is shown quantitatively to be a direct result of filtering in the following section of this chapter (Fig. 2.17). Additional impacts of filtering on the accuracy of the model's current prediction is explored further in Chapter 3 (Fig. 3.17 – Fig. 3.20; Table

3.5). As the Jacobs et al. (2019) study only considers deep-water trajectory errors (drifters in water depths >500 m), filtering is only considered for the RNCOM model.

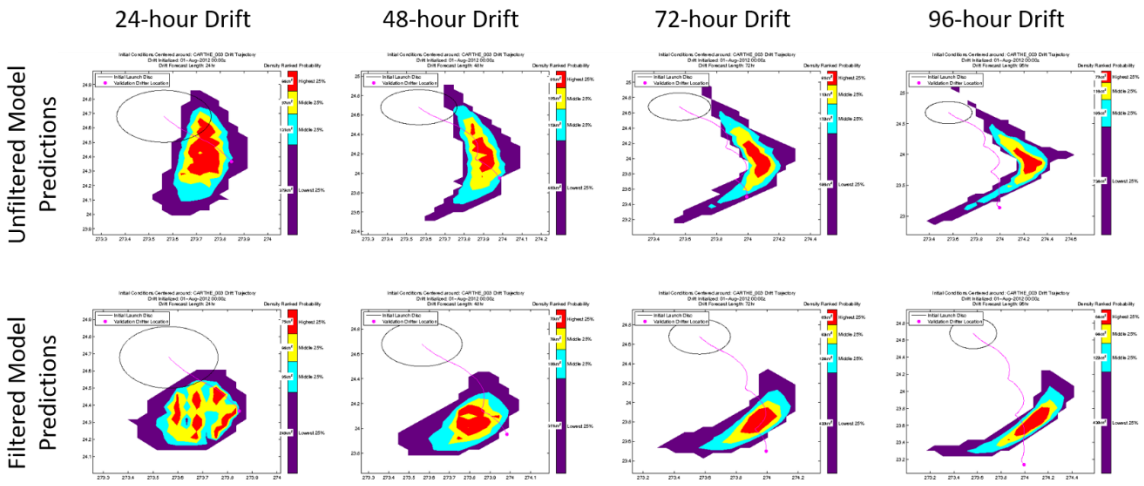

Figure 2.10 *Impact of Spatial Filtering on Drift Prediction*

Example of drift prediction derived from unfiltered (top row) versus filtered model velocity fields (bottom row), using a pure particle algorithm. Predictions are centered around a GLAD drifting buoy position on August 1, 2012. A 20-km disc of initial particles is released in each case and trajectories are integrated out 96-hours. Columns show the prediction at 24-intervals, with the left-most plot showing the 24-hour forecast and the right-most plot showing the 96-hour forecast.

2.2.3 Parameterizations to the Drift Algorithm

The vast majority of operational drift predictions are done at the surface or in near-surface layers, where currents are heavily influenced by turbulent mixing and wind waves (swell and locally generated wave field) – processes that are presently parameterized and neglected, respectively, in FNMOC's operational models. In order to understand the impact of these approximated and missing processes to Lagrangian applications, parameterizations of sub-grid scale velocity, Stokes drift, and Leeway velocity are derived and included in drift integration calculations. A spatially varying, time independent velocity field and a spatially, temporally invariant wind field (Fig. 2.11 and 2.12) are used to illustrate the effects of the different algorithms on drift prediction.

All algorithms are tested using the same launch locations and background velocity fields and atmospheric fields, where applicable. An example of the drift trajectories computed using a pure particle drift, meaning drift integrations computed using only the velocity fields from the native model output, is shown for reference in Fig. 2.13.

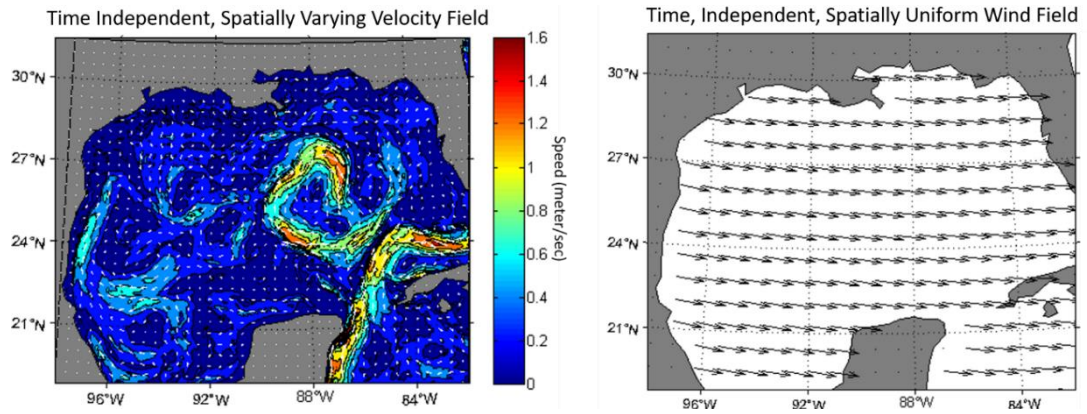

Figure 2.11 *Background Velocity and Atmospheric Fields used in Algorithm Testing*

Left: Spatially varying, temporally uniform velocity field used to compute drift trajectories for algorithm testing. Right: Spatially and temporally uniform atmospheric field used to compute drift trajectories for algorithm testing.

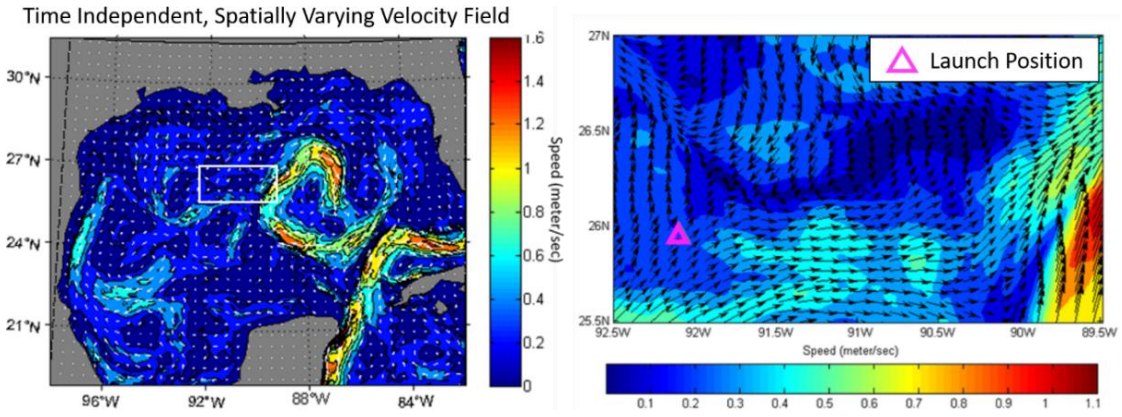

Figure 2.12 *Zoom-in of Velocity Structure near Launch Site for Algorithm Testing*

Left: Same velocity field as shown in Fig. 2.11 with extent of right-hand graphic shown in white box. Right: Zoom-in of velocity structure around the trajectory launch site (pink triangle).

Parameterization additions are included for RNCOM-derived trajectories and a summary for each parameterization is given in the subsequent sections of this chapter. Only pure particle trajectories are evaluated for CNCOM, because the primary purpose of that model run is to determine the impact model resolution has on the accuracy of the drift prediction.

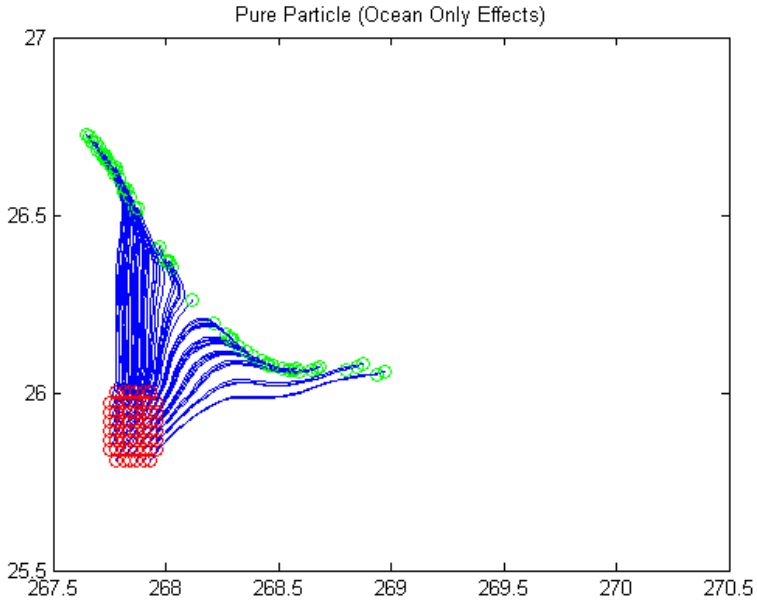

Figure 2.13 *Pure Particle Drift using Background Fields for Algorithm Testing*

A 20 km disc of particles was launched around the launch site (Fig. 2.12), and 96-hour drift trajectories were computed using the time invariant ocean model velocity fields shown in Fig, 2.12.

2.2.3.1 Sub-Grid Scale Velocity Parameterization

2.2.3.1.1 Relative Dispersion

In order to understand the impact of multi-scale flow on the dispersion of particles, the Lagrangian statistic of relative dispersion is discussed as a quantifying metric used to evaluate the performance of the drift prediction. Time dependent relative dispersion $D^2(t)$ is defined as

$$D^2(t) = \langle(\vec{r}_2(t) - \vec{r}_1(t))^2\rangle \quad (1)$$

where brackets indicate averaging over particle pairs. Relative dispersion measures the mean separation distance between two particles averaged over many pairs of particles (LaCasce 2008). Relative dispersion calculated from the GLAD drifter dataset is compared to relative dispersion calculated from modeled trajectories, and this comparison is used to determine the optimal sub-grid scale velocity parameterization needed to augment the drift prediction computed from RNCOM. Relative dispersion is also calculated and compared for predictions using single and multiple ensemble trajectory releases to determine the criteria for obtaining a prediction with reproducible drift characteristics. See Appendix A for further details concerning relative dispersion calculations.

Relative dispersion curves calculated from drifter pairs in the GLAD dataset and compared against dispersion calculated from RNCOM and CNCOM are shown in Fig. 2.14, and from them it can be seen at short times scales the 3 km RNCOM model significantly underestimates dispersion, while the higher resolution CNCOM model predicts dispersion characteristics much closer to those observed. Several studies have demonstrated that dispersion in the sub-mesoscale range is largely underestimated in ocean models of comparable resolution to the regional model implementation used in this study (Poje et al. 2010; Poje et al. 2014), which may account for the underestimation seen in RNCOM results. Conversely, it can be expected that the higher resolution of the CNCOM model allows more of the sub-mesoscale field to be resolved leading to a better representation of the dispersive nature of the near-surface velocity field. CNCOM dispersion results are shown only for reference and are not considered in further parameterization testing.

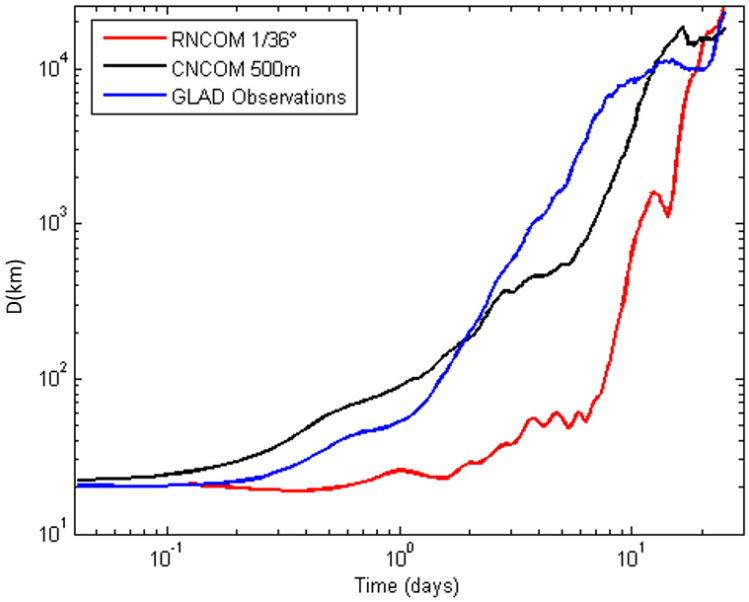

Figure 2.14 *Observed versus Modeled Relative Dispersion*

Relative dispersion curves calculated from GLAD drifter pairs that were located within 10 km and 1-hour of one another. For every drifter in the GLAD dataset, a model trajectory was initialized at the initial drifter deployment location, interpolated to the nearest model output time. Model trajectories were integrated out to 25-days. Model trajectories are taken from the native model velocity fields and not from the time invariant fields used in subsequent algorithm testing. No additional parameterizations were included in the model's drift trajectory calculations.

2.2.3.1.2 Parameterization Algorithm

The addition of a sub-grid scale parameterization through the use of Lagrangian Stochastic Models (LSMs) has shown promise in increasing pair separation rates in the sub-mesoscale range, where models of similar resolution to RNCOM show clear deficiencies (Haza et al. 2007; Haza et al. 2012). While more sophisticated Lagrangian Sub-Grid Scale (LSGS) parameterization models exist, a random flight model is chosen for simplicity and its well demonstrated effectiveness. The random flight model has been used extensively in the oceanographic community to incorporate turbulent motion below the grid scale of the model (Marinone et al. 2004; Paris et al. 2005; Cowen et al. 2006). The equations that describe particle motion, in the *i* direction, for the LSGS

parameterization can be written as

$$\delta x_i = (U_i + u_i)dt \tag{2}$$

$$\delta u_i = -\left(\frac{1}{\tau}\right) u_i dt + \sqrt{\frac{\sigma_i{}^2}{\tau}}\, dw_i \tag{3}$$

where δx_i is the total horizontal displacement of the particle during the time, *dt*. U_i is the resolved velocity field, u_i is the turbulent velocity fluctuation, σ_i^2 is the variance of the turbulent velocity fluctuation, τ is the characteristic Lagrangian decorrelation time-scale, dw_i is a random increment from a normal distribution with zero mean and variance of *2dt*, and $u_i(0) = \hat{u}$, where $\hat{u}$ is drawn from a Gaussian distribution with zero mean and variance of σ^2 (Griffa 1996).

Due to the stochastic component, *dw*, in the LSGS parameterization, dispersion calculated including this parameterization produces quite different curves when comparing multiple instances of the algorithm's results using only one iteration of releases. The differences seen in the dispersion curves imply that each instance of trajectories experiences an environment with different dispersive characteristics, due to the sensitivity of trajectory paths to the perturbations added to the velocity field. Small changes in velocities lead to different trajectories, which sample different parts of the flow field, and hence yield different dispersion characteristics (Fig. 2.15). Fig. 2.16 illustrates the need to compute dispersion using an ensemble of trajectory releases in order to obtain reproducible dispersion behavior. For each ensemble release, the initial velocity field is prescribed using an updated u_i field, which is drawn from a random distribution as noted in the preceding paragraph. From Fig. 2.16, it is seen that the 3-release, 5-release, 10-release, and 25-release runs all show similar variability in

dispersion curve separation, and overall similar behavior in curve shape, between the three ensemble instances. However, the 5 through 25-release runs show a smoother curve shape than in the 3-release comparison. While the 50-release runs show the least difference in curve shape, this amount of computation is prohibitive for operational turnaround timescales. Therefore, a 5-release run is chosen as the optimal ensemble to obtain a prediction with reproducible dispersion characteristics, and 5 consecutive runs are used in the creation of all drift probability maps derived from trajectories including the LSGS parameterization.

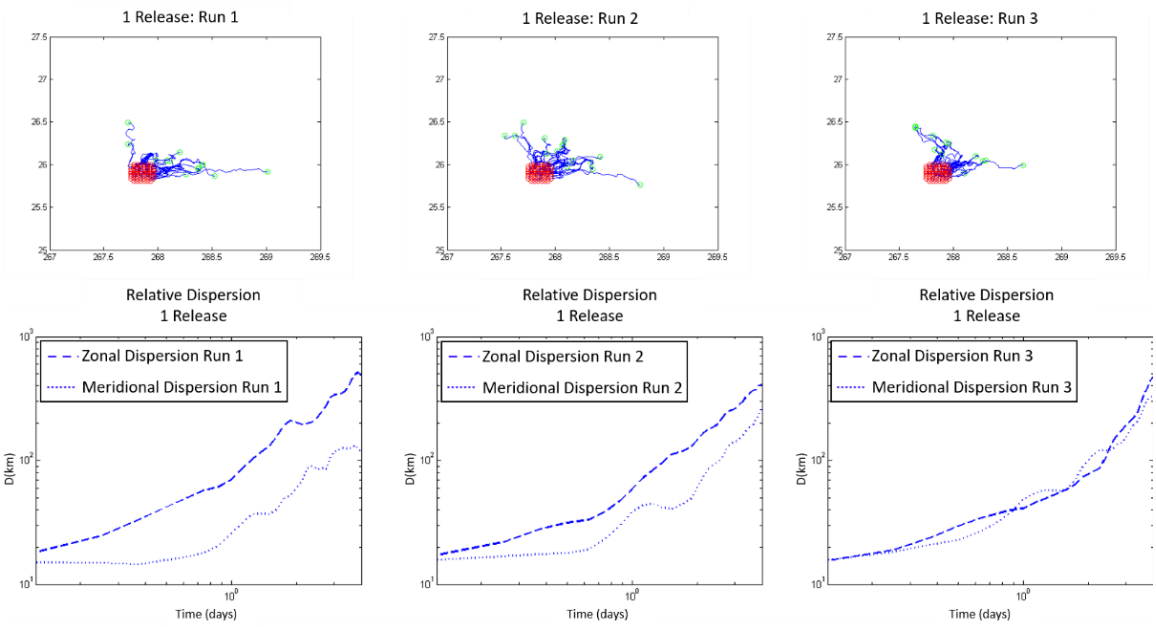

Figure 2.15 *Particle Distribution and Dispersion Characteristics using the LSGS Parameterization*

Top Panel: Modeled particle distribution for three consecutive runs including the LSGS parameterization. Trajectories were computed at the launch site and with the background fields shown in Fig.2.12. Bottom Panel: Zonal (dashed line) and meridional (dotted line) relative dispersion curves of model trajectories for each release. A 20 km disc of initial particles is used in all cases, and model pairs are defined as two initial particles that separated by no more than 10 km.

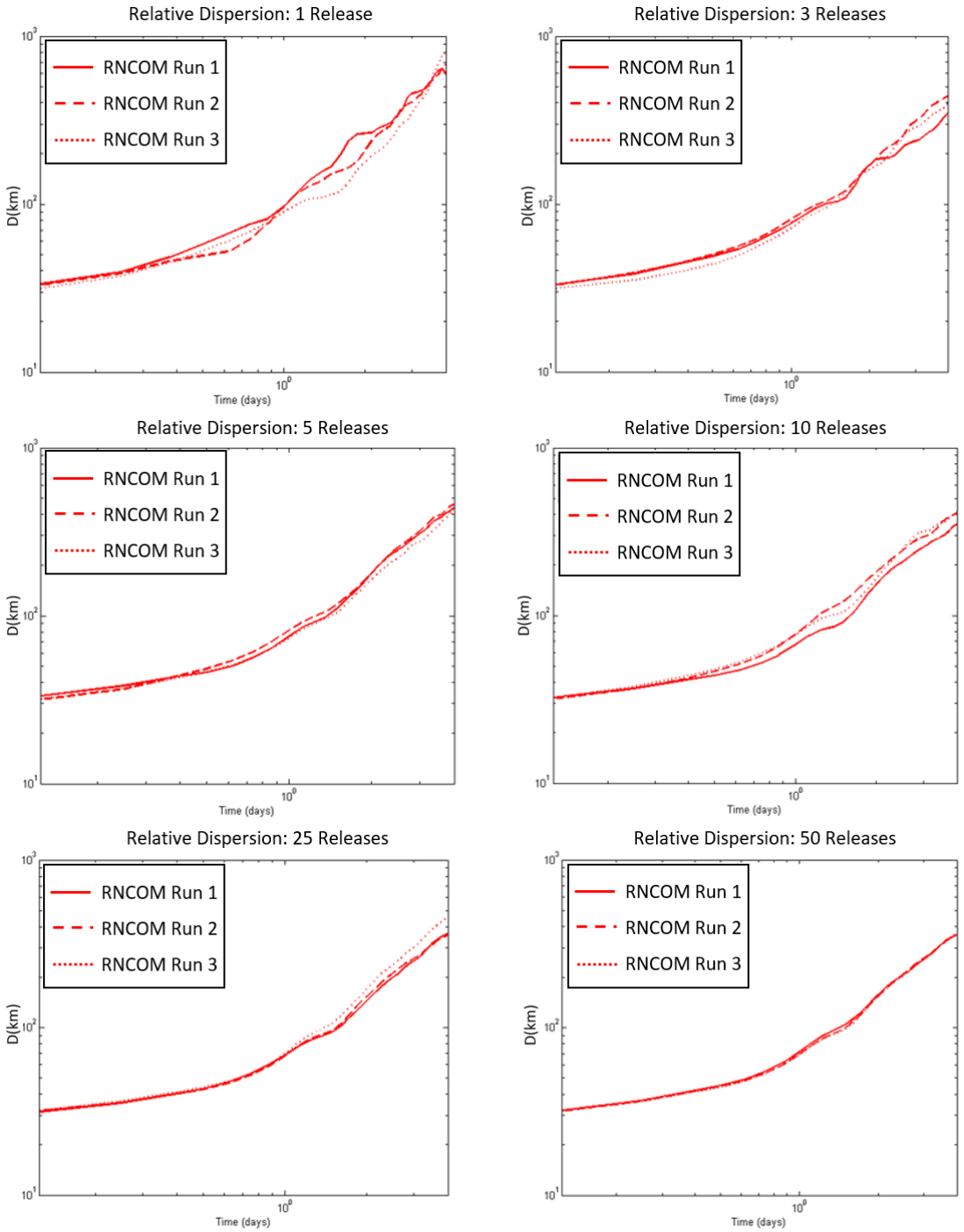

Figure 2.16 *Ensemble Testing for the Reproducibility of Dispersion Characteristics using the LSGS Parameterization*

Three consecutive runs were done for each iteration of ensemble-releases. Releases were performed at the same location and using the same background field shown in Fig. 2.12. The same criteria for model pairs as described for Fig. 2.15 was used in the calculation of relative dispersion shown here.

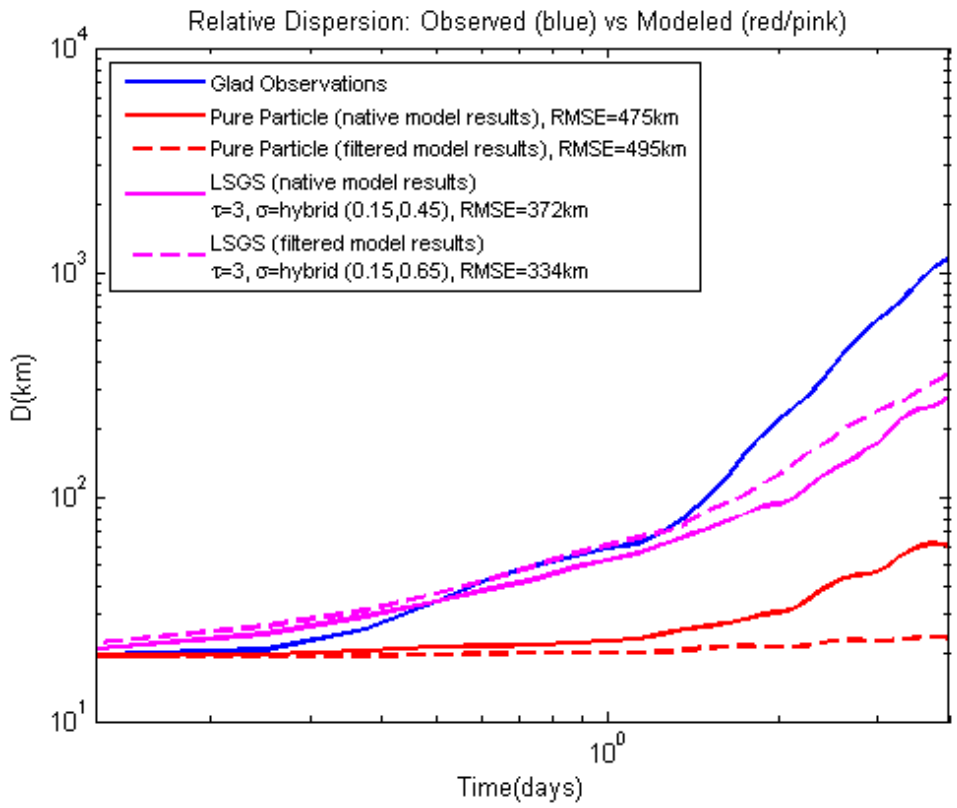

Figure 2.17 *Modeled Relative Dispersion using Optimized Parameters*

Comparison of observed relative dispersion (blue) of GLAD trajectories to modeled relative dispersion including the LSGS parameterization with optimized parameters (pink) calculated from the native model output (solid) as well as from the filtered model output (dashed). Dispersion for the modeled pure particle case is shown for reference (red) for both the native model output (solid) and the filtered model output (dashed). For these computations, a modeled trajectory was initialized at every initial GLAD deployment location and interpolated to the nearest model output time. Modeled trajectories were integrated out 4-days, using a 10-ensemble releases to ensure reproducible dispersion curve calculations. Only 4-days of drift integration was performed as that is the typical length of FNMOC's operational regional model forecasts.

Additional tests were also performed to determine optimal parameter values of τ and σ for the LSGS algorithm. It was found that a decorrelation timescale, τ, of 3-days and an initial turbulent fluctuation standard deviation, σ, of 0.15 ms^{-1} which increases to 0.45 ms^{-1} after 1-day into the drift integration, minimizes the error between modeled and observed dispersion curves. Increasing the turbulent fluctuation addition after 1-day of drift integration was necessary to replicate the behavior seen in the observed dispersion at

longer timescales. Since filtering is expected to eliminate small-scale features from model results, it is expected that relative dispersion will be reduced beyond the unfiltered model dispersion. Therefore, the LSGS parameterization is also optimized using the filtered model output as the background velocity field to the drift integration. As expected, the turbulent addition after 1-day into the drift prediction must be increased to 0.65 ms^{-1} in order to obtain similar error values as obtained for the unfiltered optimization. Modeled dispersion for the LSGS parameterization calculated from both the native model output as well as from the filtered model output are shown in Fig. 2.17, and it is seen that RNCOM dispersion curves including the parameterization are more comparable, at both short and long timescales, to the observed dispersion than for the pure particle case. An example of the impact that the LSGS parameterization has on drift trajectories is shown in Fig. 2.18; this example demonstrates that the net effect of including this parameterization is an increase in the spread of particle distribution.

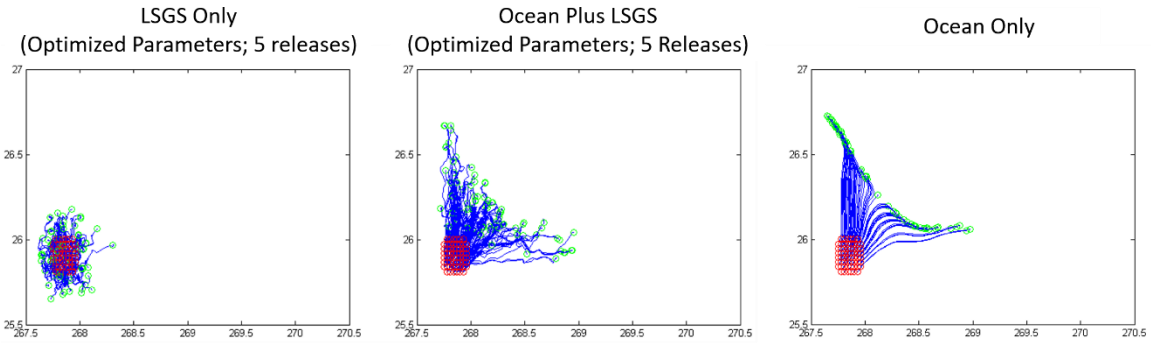

Figure 2.18 *Example of LSGS Parameterization Effect on Drift Trajectories*

Left: Drift trajectories computed including the LSGS parameterization. Velocity field is assumed to be zero. Center: Drift trajectories computed including the LSGS parameterization using a spatially varying, temporally constant velocity field (Fig. 2.12). Drift trajectories are shown from 5 ensemble releases and using the optimized parameters discussed in this section. Right: Drift trajectories computed using a pure particle algorithm, where drift is determined by predicted surface currents only. A 20 km disc of initial particles is used in all cases.

2.2.3.2 Stokes Drift Parameterization

Perhaps one the most demonstrative examples of the difference between Eulerian and Lagrangian descriptions of the current field is in Stokes drift. Stokes drift is the intrinsic forward motion of fluid particles associated with surface gravity waves (Stokes 1847). In a wave, fluid particles travel in nearly closed orbits, but since velocity fluctuations decay with depth, particles travel slightly farther forward beneath the wave crest than backward beneath the trough. This difference results in a net forward transport in the direction of wave propagation.

While several approaches exist for parameterizing Stokes drift, for simplicity sake, in this study the process is characterized using only wind speed and direction. The following relation taken from Wu (1983) is used to estimate the Stokes drift contribution,

$$\vec{V} = 0.0186 * \overrightarrow{W} \, (gL\left|\overrightarrow{W}\right|^{-2})^{0.03} \tag{4}$$

where $\vec{V}$ is the wave-induced surface current, $\overrightarrow{W}$ is the wind velocity at a 10-meter elevation above the mean sea surface, g is the gravity constant, and L is the wind fetch. An example of the impact this parameterization has on drift trajectories is shown in Fig. 2.19. A big assumption in parameterizing Stokes drift using only wind speed and direction is that the wave field always correlates with the local wind field, which is not always true, as is the case with swell. This is a definite weakness in the parameterization chosen. Improved results may be seen for Stokes drift calculated from a co-located wave model or directly incorporated into the ocean model solution by running a coupled ocean-wave model. Additionally, depth decrease of the Stokes contribution is not taken into account with this algorithm; introduction of an exponential decay for the derived Stokes

velocity component would be an area of improvement. Assessing the Stokes drift contribution through these more advanced methods is noted as an area for follow-on work.

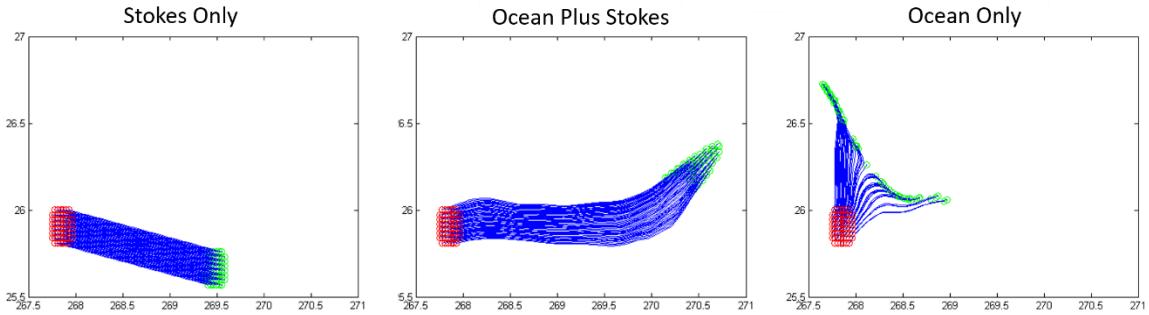

Figure 2.19 *Example of Stokes Parameterization Effect on Drift Trajectories*

Left: Drift trajectories computed including the Stokes parameterization. Velocity field is assumed to be zero and wind field is spatially and temporally uniform (Fig. 2.11). Center: Drift trajectories computed including the Stokes parameterization. In this example the velocity field is spatially varying but temporally constant (Fig. 2.12), and the wind field is spatially and temporally invariant (Fig. 2.11). Right: Drift trajectories computed using a pure particle algorithm, where drift is determined by predicted surface currents only. A 20 km disc of initial particles is used in all cases.

2.2.3.3 Leeway Parameterization

For most objects that need to be tracked in the ocean, drifting motion is a combination of several forces acting upon its surface (water currents, atmospheric wind, wave motion, and wave-induced currents) and its center of mass (buoyancy force) (Hackett et al. 2006). The total drift velocity of an object can be given by

$$\vec{V}_{drift} = \vec{V}_{curr} + \vec{V}_{rel} \quad (5)$$

where $\vec{V}_{curr}$ denotes the water velocity relative to earth, and $\vec{V}_{rel}$ denotes the object drift relative to the water. In general, $\vec{V}_{curr}$ contains contributions of the surface current, including Ekman drift, baroclinic motion, tidal and inertial currents, and Stokes drift

induced by waves (Hackett et al. 2006). Current velocities output by the numerical model also include these processes with the exception of Stokes drift. $\vec{V}_{rel}$, also widely known as Leeway drift, results from wind and wave forces acting on the object and is strongly dependent on the characteristics of the object. It is important to note that there are two mechanisms by which the wind causes a floating object to drift relative to the surface current. The first is the drag force exerted on the floating object as it is pushed through the water. If the profile of the object is asymmetric in any way, the drift will have both a downwind and crosswind drift component. The second is caused by surface gravity waves that are generated by the wind; these waves generally propagate in the downwind direction but may also move at some angle to the downwind direction. Wave induced Leeway drift is due to wave rectification and Stokes drift, when not Stokes is not explicitly accounted for in $\vec{V}_{curr}$. Since both physical mechanisms are wind driven, the combined effect of Leeway drift correlates well with surface wind speed and direction and is modeled as a function of wind speed using linear regression techniques (Allen 2005).

The Leeway model is based on extensive work by Allen and Plourde (1999) and a follow-up report done by Allen (2005) and is incorporated into several operational search and rescue (SAR) models (Breivik and Allen 2008; Cho et al. 2014). The Leeway coefficients for specific objects can be determined using direct methods, where the relative velocity between the drifting object and ocean currents is directly measured and compared to the wind velocity (Allen and Plourde 1999; Allen 2005). Direct method experiments were performed on both SVP drifters (Niiler et al. 1987; Niiler et al. 1995)

and CODE drifters (Poulain and Gerin 2019), and slippage for both drifter types was found to be very small, about 0.1% of the wind speed. Since the traditional approximation of Leeway drift has been shown to be minimal for this study's two drifter types, an alternate approach is taken for defining Leeway coefficients. In some very recent work done by Sutherland et al. (2020), Leeway coefficients for drifting buoys were estimated by comparing derived velocities from several drifting buoy trajectories to those predicted by an operational ocean model. The following method describes the process used for determining the coefficients that will be used in the Leeway parameterization; this method closely follows what was done by Sutherland et al. (2020). Drifting buoy trajectories are only considered from the GLAD dataset, as the reporting interval for this dataset is much finer than for SVP drifters, 15-minutes compared to 6-hours. Since slippage estimates are very similar for both types of drifters, it is expected that Leeway coefficients derived using GLAD drifters will also characterize SVP drifters well.

First, GLAD drifter trajectories are low-pass filtered to remove inertial oscillations, using a sixth-order Butterworth filter with a 60-hour cutoff frequency (Fig. 2.20). This was done primarily because both RNCOM and CNCOM are shown to have a weakness in the prediction of inertial oscillations, which is demonstrated in Chapter 3, Section 2 of this paper. Additionally, inertial oscillations have little effect on the overall drift prediction on timescales longer than the inertial period (i.e. approximately 24-hours in the GOM). After filtering, drifting buoy velocities are calculated using a center-difference of drifter locations, with the first and last values of velocity along the trajectories estimated using a forward- and backward-difference, respectively. COAMPS surface winds and RNCOM current velocity taken at the drifter's drogue depth (~2-

meters) are bilinearly interpolated to the drifter position at model output times. The timeseries of RNCOM current velocity along each drifter trajectory is also low-pass filtered to remove inertial oscillations, using the same filter applied to the buoy positions.

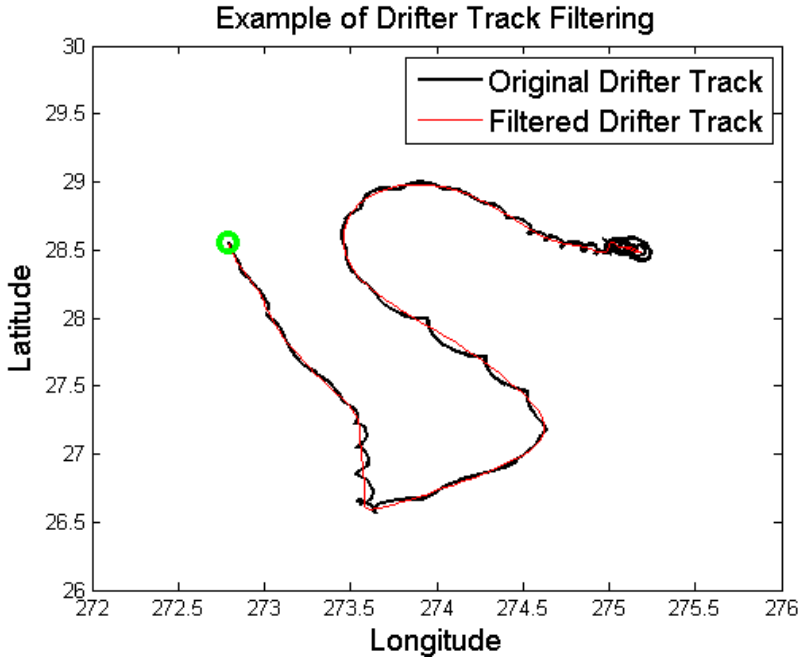

Figure 2.20 *Example of Low-Pass filtering done on Drifter Trajectories*

Comparison of a GLAD drifter trajectory before and after a low-pass filter was applied to drifter positions. The original drifter trajectory is shown in black and the low-pass filtered drifter trajectory is shown in red. The inertial oscillations seen in the original trajectory have been removed in the filtered case. The initial position of the drifter is indicated by the green circle. The filter used is a sixth-order low-pass Butterworth filter with a 60-hour cutoff frequency.

The Leeway velocity is then defined as the velocity contribution needed to make the model velocity perfectly match the observed velocity. An example of the drifter velocity, model velocity, and Leeway velocity along a drifter trajectory is shown in Figure 2.21. Spatially and temporally varying Leeway velocity, α, can then be determined along the drifter trajectory using the following formula

$$\alpha = \frac{u_d - u_o}{U_{10}} \tag{6}$$

where u_d is the drifting buoy velocity vector, u_o is the model current velocity vector, and

U_{10} is the 10-meter wind velocity vector. Here u_d, u_o and U_{10} are complex representations of velocity vectors, with the resulting leeway velocity vector, α, also complex. The real part of α is in the along-wind direction and the imaginary part of α in the cross-wind direction, with a negative value indicating a cross-wind component to the right of the wind. Downwind (i.e. real(α)) and crosswind (i.e. imaginary(α)) leeway velocity components were calculated in this way for every drifter trajectory in the GLAD dataset (Fig. 2.4).

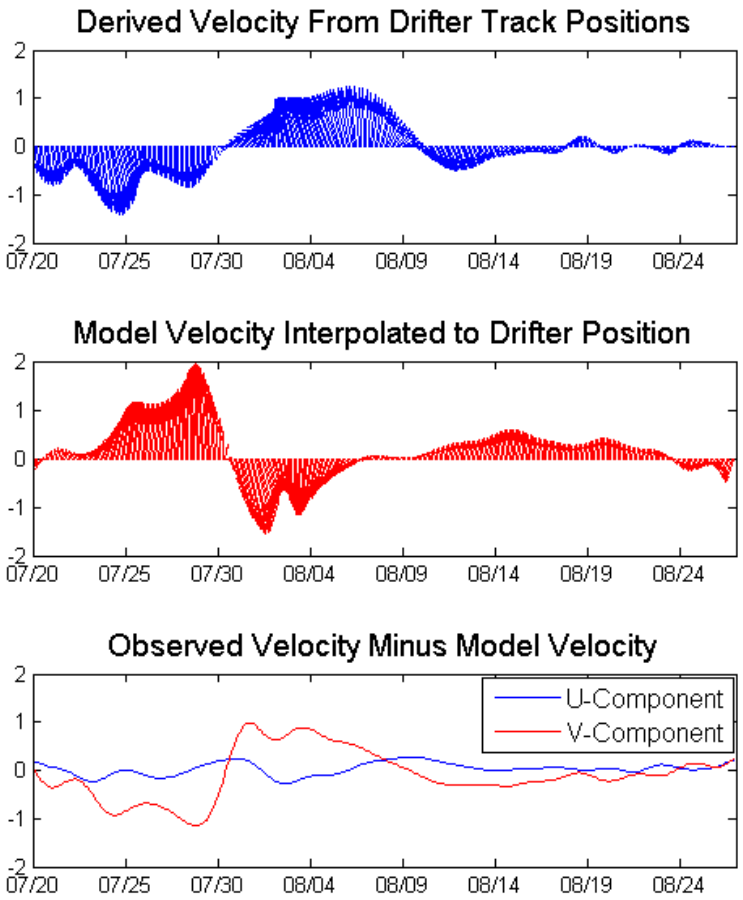

Figure 2.21 *Example of Leeway Velocity Computation*

Top-panel shows the low-pass filtered drifting buoy velocity taken along the drifter trajectory in Figure 2.20. Center-panel shows the low-pass filtered RNCOM velocity interpolated to the drifter position. Both drifter and model velocities are shown at model output times (3-hourly intervals). Bottom-panel shows the Leeway velocity (observed minus model velocity) for u and v velocity components.

Using a methodology as laid out by Sutherland (2020), the resulting distribution of α not only captures the time evolution of Leeway drift but also encompasses all sources of uncertainty in both the model prediction as well as in the observed drifter velocity. Since it is inherently difficult to unravel all the different sources of uncertainty that feed into the drift prediction, empirically determining a Leeway contribution that naturally takes into account these sources of uncertainty are desirable. The distribution of downwind Leeway (DWL) and crosswind Leeway (CWL) velocity components calculated using the above described method is shown in Fig. 2.22.

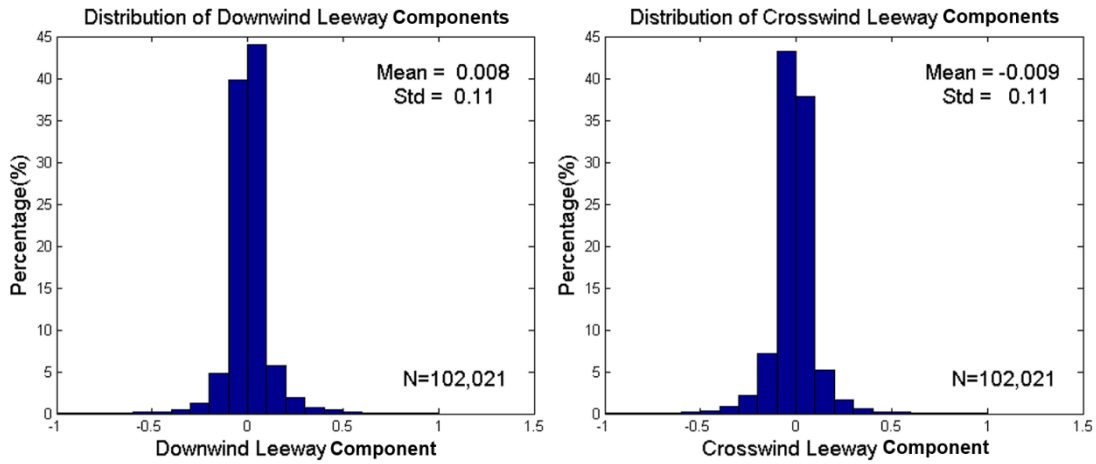

Figure 2.22 *Distribution of Derived Downwind and Crosswind Leeway Velocity Components*

Distribution of downwind (left) and crosswind (right) Leeway velocity components calculated using Eq. (6). Over 100,000 computations of Leeway velocity were derived from the GLAD drifter dataset and the regional scale ocean model (RNCOM) used in this study.

Mean and standard deviation statistics (Table 2.2) of DWL and CWL velocity components are then used to derive the linear regressions that determine the DWL and CWL velocity contribution according to wind speed (Allen and Plourde 1999; Allen 2005). The commonly used term *Leeway coefficients* refers to the slope and y-intercept of the regression line determined by the mean and standard deviation statistics of down- and

cross-wind Leeway velocity (Allen 2005). Unconstrained regressions are implemented in this study, as the Leeway velocity additionally includes model uncertainty, a non-zero Leeway drift even when wind speed is zero is acceptable. The regression slopes, a, and y-intercepts, b, are calculated using the following equations

$$Offset = Std_{DWL/CWL} * N(0,1) \tag{7}$$

$$a = mean_{DWL/CWL} + Offset/20 \tag{8}$$

$$b = Offset/2 \tag{9}$$

where $Std_{DWL/CWL}$ is the standard deviation of the DWL and CWL velocity components, *N(0,1)* indicates a random number drawn from a normal distribution, and $mean_{DWL/CWL}$ is the mean value of the DWL and CWL velocity components (Table 2.2). Eq. (7) – (9) are computed separately for DWL and CWL regressions. The *offset* is applied to both the slope (a) and y-intercept (b) perturbations to increase the variance expected at higher wind speeds and decrease the variance expected at lower wind speeds (Allen 2005). The DWL and CWL velocity contributions can then be computed from the linear regressions as follows

$$\overrightarrow{Leeway_{DW}} = \left(a * \overrightarrow{U_{10}}\right) + b \tag{10}$$

$$\overrightarrow{Leeway_{CW}} = \left(a * \left(orientation * \overrightarrow{U_{10}}\right)\right) + b \tag{11}$$

$$\overrightarrow{V_{drift}} = \overrightarrow{V_{curr}} + \overrightarrow{Leeway_{DW}} + \overrightarrow{Leeway_{CW}} \tag{12}$$

where a is the result of Eq. (8), b is the result of Eq. (9), U_{10} is the 10-meter wind velocity, V_{curr} is the ocean velocity vector, $Leeway_{DW}$ and $Leeway_{CW}$ are the downwind and crosswind leeway velocity vectors, and V_{drift} is the resulting ocean velocity including the leeway contribution. The *orientation* variable in Eq. (11) determines whether

$Leeway_{CW}$ will be to the left or to the right of the wind. Initially, this variable is set so that there is an equal probability of crosswind leeway to the left and right of the wind. During the drift integration, *orientation* is varied so that 4% of the particles are allowed to jibe in the opposite direction every hour during the drift integration, as is recommended by Allen (2005). An example of these linear regressions for $Leeway_{DW}$ and $Leeway_{CW}$ are shown in Fig. 2.23.

Table 2.2 *Mean and Standard Deviation of Leeway Components*

Coefficient Statistics	Mean	Standard Deviation
Downwind Leeway	0.008	0.11
Crosswind Leeway	-0.009	0.11

Note: Mean and standard deviation of the DWL and CWL velocity components distributions shown in Figure 2.22. Statistics are computed using 10,000 bootstrapped samples, with 453 (number of GLAD drifter trajectories) degrees of freedom and are significant at a 95% confidence level.

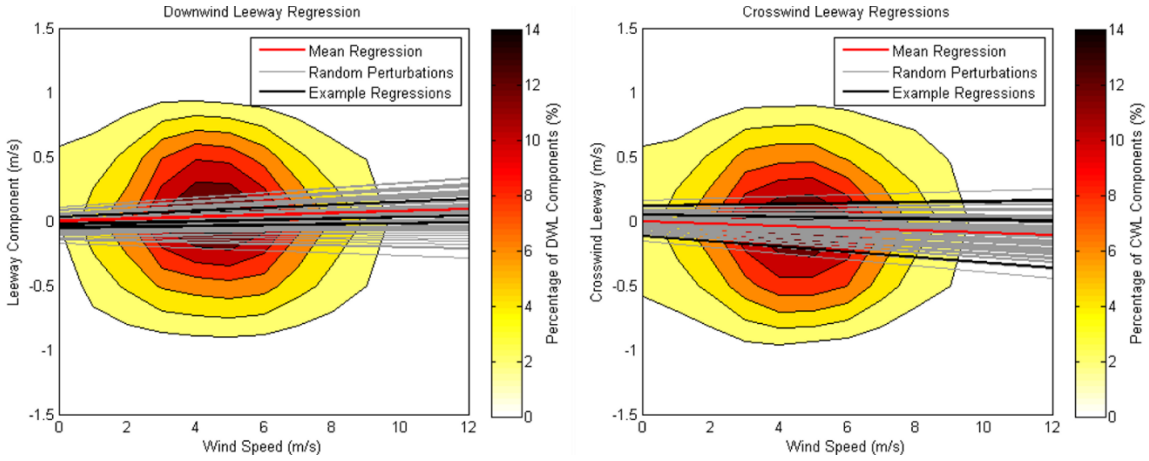

Figure 2.23 *Downwind and Crosswind Regressions used in Leeway Parameterization*

Downwind regressions are shown on the left. Crosswind regressions are shown on the right. The mean regression line is shown in red and perturbations around the mean regression are shown in gray. The bold black lines are the regressions that are used in the three example drift trajectories including the Leeway parameterization shown in Figure 2.24. The colored contours in the background show the percentage of downwind and crosswind Leeway velocity components calculated according to wind speed.

A basic assumption of the Leeway model is that a specific object has a unique set of Leeway equations and that it retains the same Leeway characteristics over the duration of the drift integration (Allen 2005). Therefore, a Monte Carlo approach is typically taken where *N* sets of Leeway equations are specified where *N* represents the number of ensemble replications for the drift integration, and each ensemble drift instance uses the same regression as a function of wind speed for the entire drift integration. Generally, *N* is quite high when implemented operationally. For example, Breivik and Allen (2008) cite an ensemble on the order of 500 when describing the operational SAR model employed by the Norwegian Joint Rescue Coordination Centres. Since an ensemble of

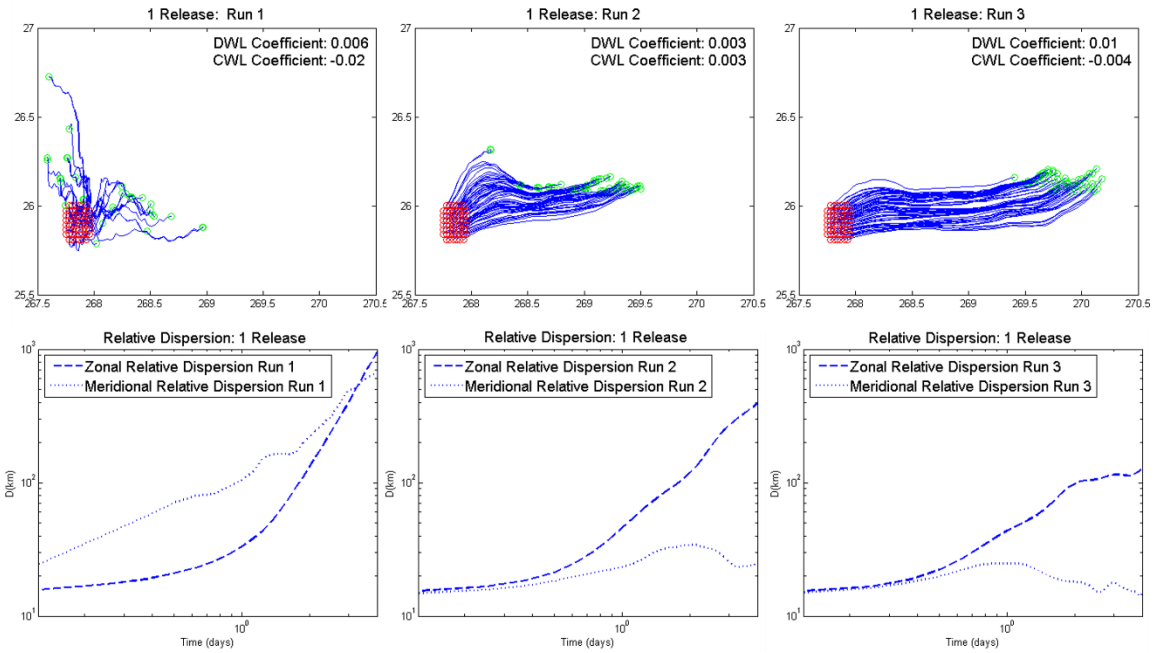

Figure 2.24 *Sensitivity of Particle Distribution to Leeway parameterization*

Top Row: Modeled particle distribution for three consecutive runs including the Leeway parameterization. Trajectories were computed at the launch site, using a 20 km disc of initial particles, and the background fields shown in Fig. 2.11. The DWL and CWL coefficients used for each run are shown at the top right of each figure; these coefficients correspond to the black regression lines shown in Figure 2.23. Bottom Row: Zonal (dashed line) and meridional (dotted line) relative dispersion curves computed from model trajectories for each release. Model pairs are defined as two initial particles that separated by no more than 10 km. See Appendix A for further details concerning relative dispersion calculations

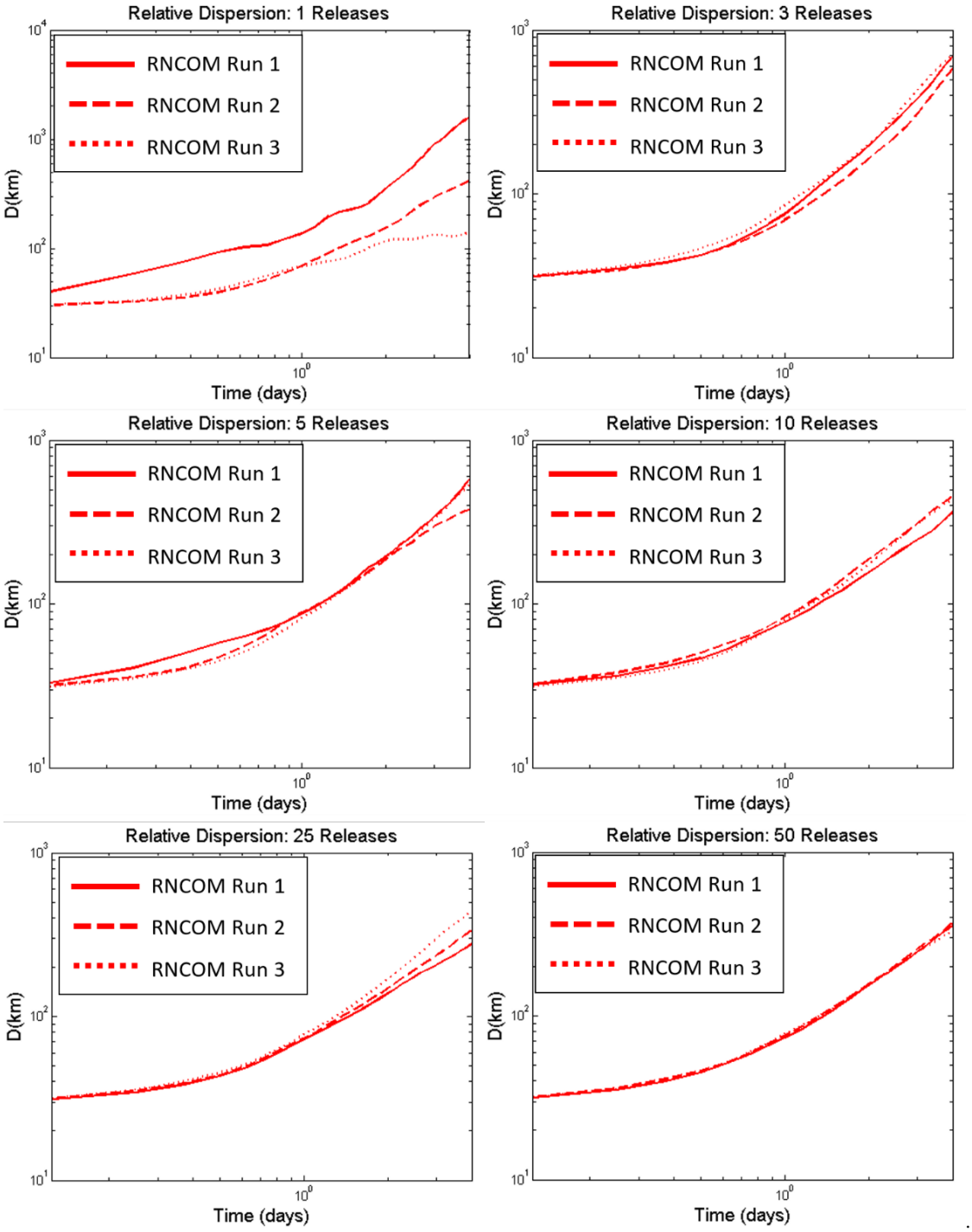

Figure 2.25 *Ensemble Testing for the Reproducibility of Dispersion Characteristics using the Leeway Parameterization*

Three consecutive runs were done for each iteration of ensemble-releases. Releases were performed at the same location and using the same background fields shown in Fig. 2.11. The same criteria for model pairs as described for Fig. 2.24 was used in the calculation of relative dispersion shown here.

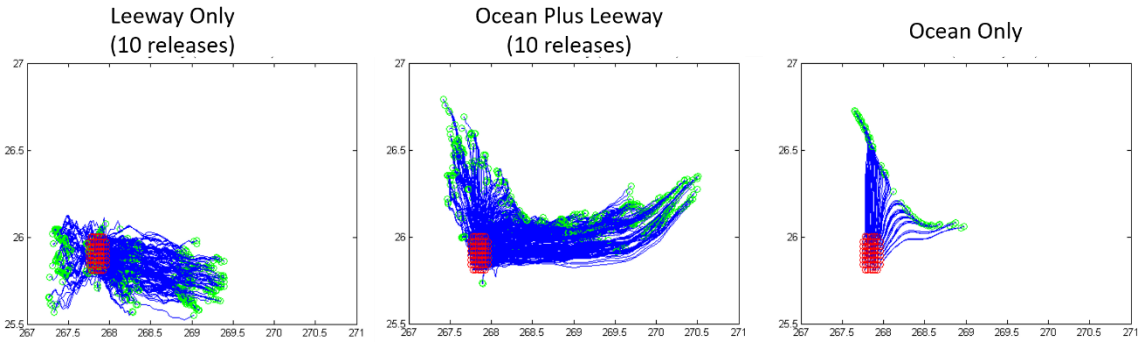

Figure 2.26 *Example of Leeway Parameterization Effects on Drift Trajectories*

96-hour drift trajectories computed including the Leeway parameterization and run using 10-ensemble releases. Left: Trajectories computed using only the Leeway drift. The background velocity field is assumed to be zero and wind field is spatially and temporally uniform (Fig. 2.11). Regression coefficients are time invariant and all trajectories use the same coefficients for each ensemble run. Center: Trajectories computed using both Leeway velocity and the background velocity and wind field shown in Fig. 2.11. Regression coefficients are time invariant and all trajectories use the same coefficients for each ensemble run. The same regression coefficients are applied for the Leeway only and ocean plus Leeway cases. Right: Drift trajectories computed using a pure particle algorithm, where drift is determined by predicted surface currents only.

that size is not feasible for this study, similar testing is done for Leeway as was done for the LSGS parameterization, where relative dispersion curves are used to estimate the ensemble size needed to reproduce dispersion characteristics.

Due to the large variance associated with the DWL and CWL regression coefficients, multiple instances of a single ensemble run can show quite significant variations in the resulting drift trajectories (Fig. 2.24), highlighting the necessity of a large ensemble size to adequately capture the spread in drift area associated with the range of DWL/CWL choices. Dispersion curves for different sized ensemble releases is shown in Fig. 2.25. While spread between curves is similar for 3, 5, 10, and 25-releases, a value of 10 is chosen as a compromise between a larger *N* value and the computational limitations of this study. It must be noted, however, that a much larger ensemble size is

likely needed in an operational setting than is chosen for this study, especially considering the large standard deviation values determined for the DWL and CWL coefficients using RNCOM (Table 2.2). An example of the impact the Leeway algorithm has on drifter trajectories is shown in Fig. 2.26 and demonstrates that the Leeway addition acts to increase the spread in particle positions in down and crosswind directions.

CHAPTER III – MODEL VALIDATION

Because accurate Eulerian current fields and wind-forcing fields are prerequisites for skillful Lagrangian drift predictions, validating the Eulerian fields is a necessary first step before advancing to accessing the accuracy of drift predictions. While the following model validation does not directly relate to the Lagrangian predictive skill which is the main objective of this study, the analyses provided in this chapter are meant to highlight the typical sources of error in NCOM's current prediction and provide some context for the interpretation of the Lagrangian bulk error statistics that follow in the results chapter of this work. Note that while wind velocity data is assimilated by the atmospheric forcing model, ocean current data is not yet an assimilated parameter by FNMOC's operational ocean models. Therefore, atmospheric model comparisons are not made with an independent dataset and validation metrics are, therefore, expected to be more skillful than the ocean current metrics.

3.1 Atmospheric Forcing Validation

Wind predictions over the GOM from the COAMPS atmospheric model are compared to observed winds at several National Data Buoy Center (NDBC) station locations (Fig. 3.1) to validate the accuracy of wind forcing provided to the ocean model. The time frame for the comparison coincides with the GLAD experiment, July 1, 2012 through October 31, 2012. While COAMPS wind predictions are output at a standard reference height of 10 meters, wind speed and direction in NDBC's historical data files are output at the buoy's anemometer measurement height, which varies widely according to buoy type. Therefore, all wind observations from NDBC stations are adjusted to a 10-meter reference height using the simple Power-Law method given by Hsu et al. (1994)

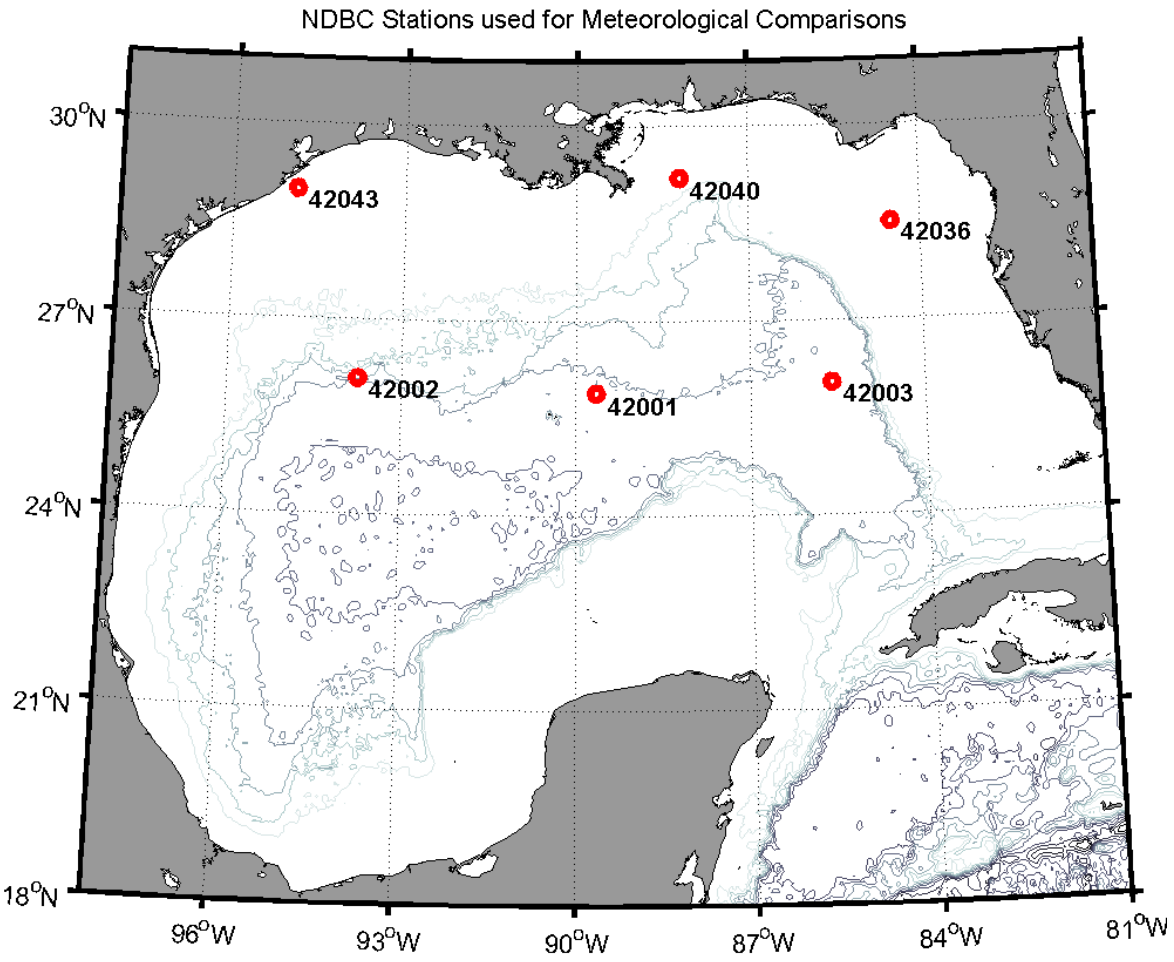

Figure 3.1 *Location of NDBC Stations Used in Atmospheric Model Validation*

Buoy information obtained from https://www.ndbc.noaa.gov/.

$$u_2 = \left(\frac{z_2}{z_1}\right)^P * u_1 \tag{13}$$

where u_2 is the wind speed measured at the desired reference height, z_2, and u_1 is the wind speed at the measurement height, z_1, and P=0.11. Although not as rigorous as other algorithms used to adjust wind speeds, Hsu's Power-Law method has been shown to compare favorably to more complex methods under near-neutral stability conditions, which prevail at sea (Hsu et al. 1994). Anemometer heights for the NBDC stations used in the GOM comparison are shown in Table 3.1.

Table 3.1 *NDBC Meteorological Station Anemometer Height*

Station	Anemometer Height
42001	5m
42002	5m
42003	5m
42036	5m
42040	4m
42043	3.4m

Note: Buoy information obtained from https://www.ndbc.noaa.gov/.

The two metrics used to validate wind predictions are wind speed correlations and vector correlations. Correlation coefficients calculated for wind speeds are Pearson's product moment correlations where 1.0 indicates perfect correlation, or complete linear dependence, between datasets and 0.0 indicates no correlation, or complete linear independence. Vector correlations follow Kundu (1976) and are defined as

$$R = \frac{\langle v_1'^{*}\, v_2' \rangle}{\sqrt{\langle v_1'^{*} v_1' \rangle \langle v_2'^{*} * v_2' \rangle}} \tag{14}$$

with v'= v -$\bar{v}$, where $\bar{v}$ denotes an average taken over the signal length and the asterisk indicates the complex conjugate; $v=W_x + iW_y$, where $W=(W_x, W_y)$ is a velocity vector, and the subscript denotes the two velocity datasets being correlated. The absolute value of R gives an overall measure of the correlation (1.0 indicates perfect correlation) between the two vector time series (Kundu 1976). All subsequent scalar and vector correlations appearing in this study follow the above cited formulations and are significant to the 95% level. The significance levels for both scalar and vector correlations are computed using 10,000 bootstrapped samples. The number of

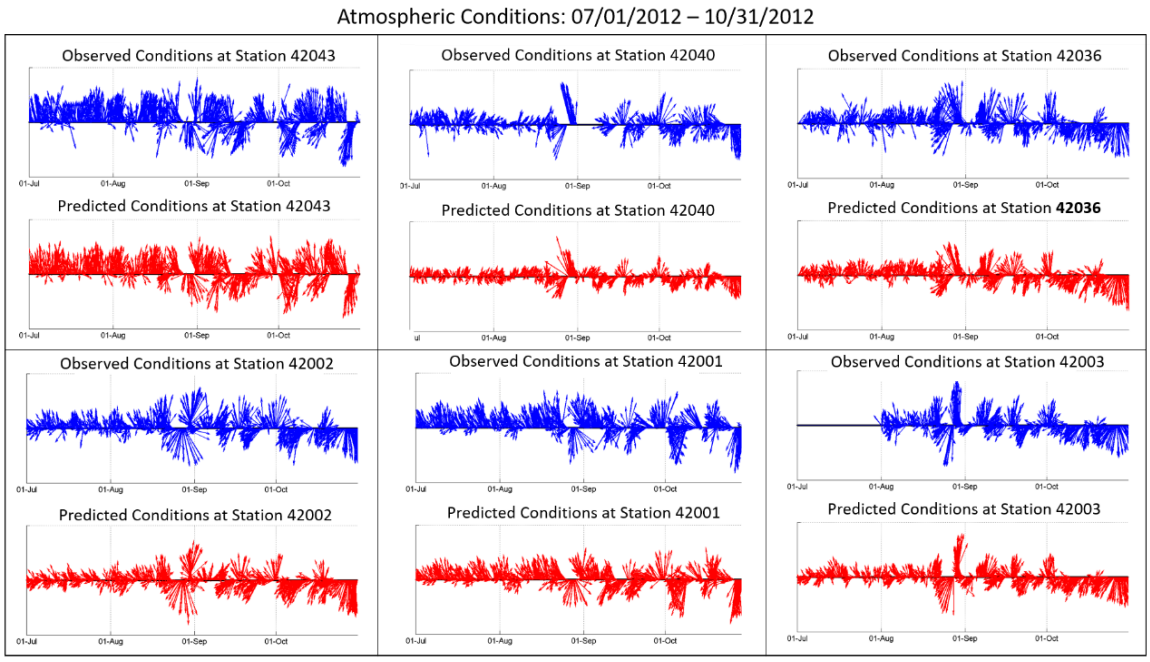

Figure 3.2 *Vector Time series of Observed versus Predicted Surface Winds*

Time series of observed (blue) and COAMPS predicted (red) wind direction at NDBC buoy locations in the GOM (Fig. 3.1) spanning the study period. Wind directions are converted to the "going-to" convention before decomposing into velocity components. Observed wind observations have been adjusted to a 10-meter reference height for comparison with modeled surface winds. Observations and model predictions are interpolated to a uniform timescale using a 1-hourly interval. Interpolation is not performed where data gaps of greater than 1-day are present in observed timeseries.

observations drawn for each sample is determined by using the effective degrees of freedom, obtained by dividing the length of each timeseries by its integral time scale, following Thomson (2014). The integral timescale for each buoy station is estimated from the lagged autocorrelation of the observed timeseries.

A qualitative look at the observed versus predicted winds over the GOM (Figs. 3.2 and 3.3) shows the predicted wind field compares well with the observed, capturing the overall character of the wind forcing. Vector correlations for all stations are above 0.8 (Table 3.2). Wind speed correlations are not as high as vector correlations, with values

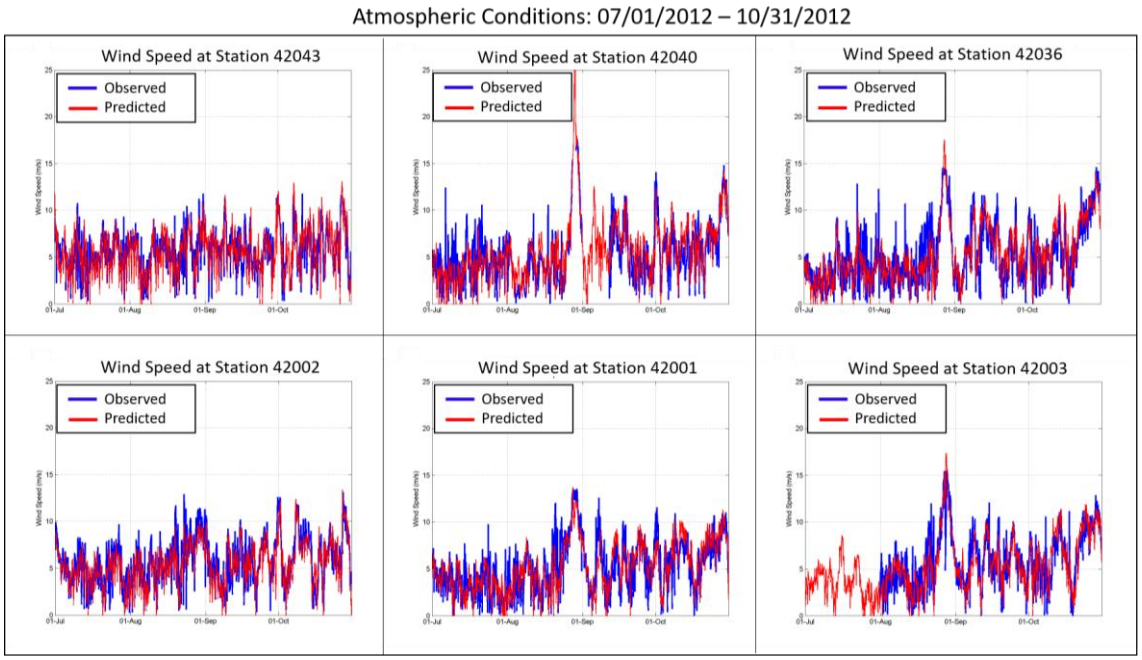

Figure 3.3 *Wind Speed Time series of Observed versus Predicted Surface Winds*

Time series of observed (blue) and COAMPS predicted (red) wind speed at NDBC buoy locations in the GOM (Fig. 3.1) spanning the study period. Observed wind observations have been adjusted to a 10-meter reference height for comparison with modeled surface winds. Observations and model predictions are interpolated to a uniform timescale using a 1-hourly interval. Interpolation is not performed where data gaps of greater than 1-day are present in observed timeseries.

Table 3.2 *Vector and Speed Correlations between Observed and Predicted Winds*

Station	Vector Correlation (95% CI)	Speed Correlation (95% CI)
42001	0.86 (0.85 -0.88)	0.76 (0.74 – 0.80)
42002	0.87 (0.86 – 0.89)	0.77 (0.74 – 0.80)
42003	0.89 (0.87 – 0.91)	0.83 (0.81 – 0.86)
42036	0.88 (0.86 – 0.90)	0.79 (0.76 – 0.84)
42040	0.81 (0.77 – 0.87)	0.76 (0.68 – 0.86)
42043	0.85 (0.84 – 0.87)	0.71 (0.68 – 0.75)

Note: Observations and model predictions are interpolated to a uniform timescale using a 1-hourly interval. Interpolation is not performed where data gaps of greater than 1-day are present in observed timeseries. These times are not included in correlation calculations. Correlations are significant to the 95% level and were calculated using 10,000 bootstrapped samples; confidence intervals (CI) are given in parenthesis.

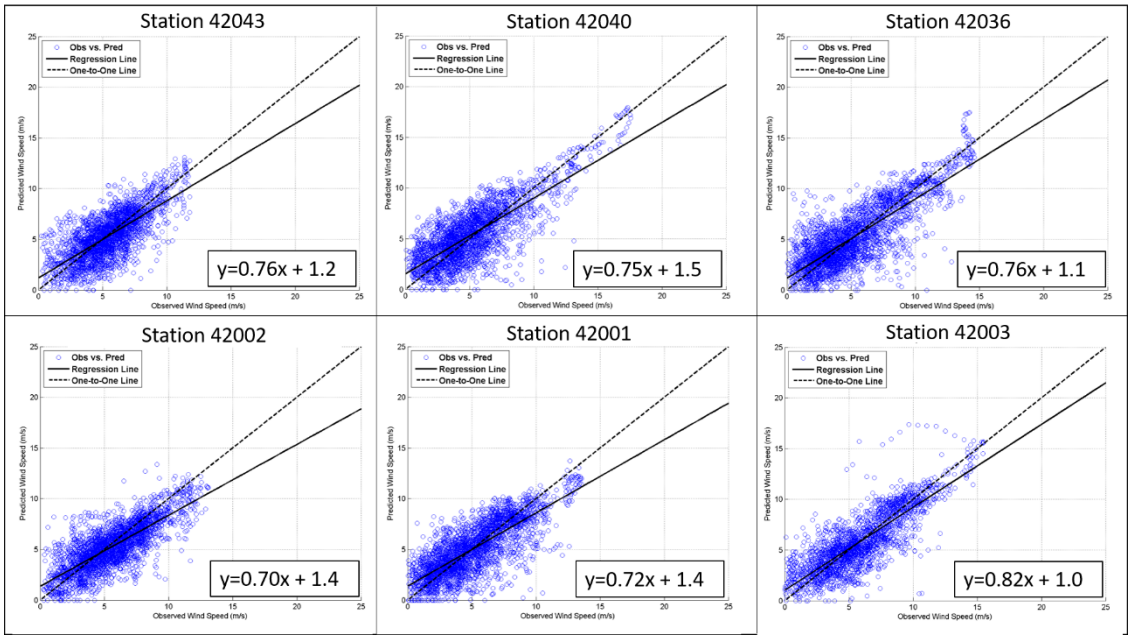

Figure 3.4 *Observed versus Predicted Wind Speed Comparison*

Linear regression of observed versus COAMPS modeled wind speeds at NDBC buoy locations in the GOM (Fig. 3.1) spanning the study period. Observed wind observations adjusted to a 10-meter reference height for comparison with modeled surface winds. Observations and model predictions are interpolated to a uniform timescale using 1-hourly interval. Interpolation is not performed where data gaps of greater than 1-day are present in observed timeseries; these times are therefore not included in regression analysis.

ranging from 0.71 to 0.83 (Table 3.2). In addition, it is seen from Fig. 3.4 that at all stations wind speeds are, in general, slightly under-predicted by the model.

3.2 Open-Ocean Eulerian Ocean Current Validation

In the open ocean, Eulerian accuracy will be determined, to a large extent, by the placement of mesoscale features, which is achieved mainly through the assimilation of sea surface height anomalies (SSHA) obtained from satellite altimeters (Metzger et al. 2014). Over the study period, operational ocean models were receiving SSHA data from three altimeters: Jason-1, Jason-2, and Cryosat-2. Altimeter data available to constrain the placement of features is shown in Fig. 3.5 and Fig. 3.6. Large gaps in altimeter coverage are apparent throughout the study period; therefore, it is expected that errors will exist in

the placement of RNCOM's mesoscale eddy field due to this lack of constraining data. It must be noted that while the spatial gaps in altimeter coverage are obvious, temporal gaps also exist. Jason-1 and Jason-2 satellites orbited on a 10-day repeat cycle, while Cryosat-2 orbited on a 369-day repeat cycle. Absolute Dynamic Topography (ADT) altimeter products covering the study area and the study period were produced and distributed by the Copernicus Marine and Environment Monitoring Service (CMEMS) (http://www.marine.copernicus.eu).

The altimeter-derived ADT field is computed using data from the same altimeter platforms to which operational ocean models have access and are assimilating; however, these multi-mission sea level products are produced with delayed-time data which is more precise than the near real-time along-track data that the model assimilates. It also must be noted that the ADT product is produced at a ¼ degree resolution, much coarser than the resolution of the NCOM models used in this study. However, at present this altimeter-derived product is the only observation-based data that is collected at a broad enough scale (both spatially and temporally) to provide some level of validation for the modeled mesoscale eddy field. As such, the altimeter-derived ADT field is compared to the sea surface height predicted by RNCOM. These comparisons are shown in Figs. 3.7 through 3.10 where it is seen that while there are differences in the shape and size of eddy features, the overall mesoscale eddy field is very similar between that derived from satellite altimeters and that predicted by RNCOM, as is expected.

The dominant features that will influence trajectory motion in the two drifter datasets are the large anti-cyclonic loop current eddy (W1), located at about 25°N, 89°W,

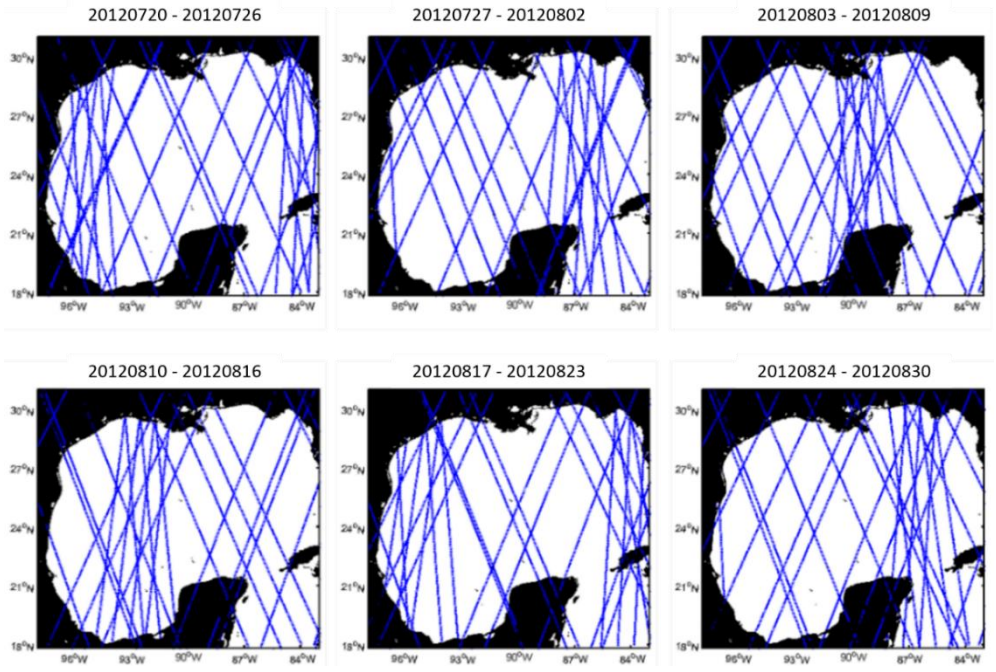

Figure 3.5 *Weekly Coverage of Altimeter Data: 20120720 through 20120830*

Plots include satellite track coverage from Jason-1, Jason-2, and Cryosat-2 missions. Time frame covers July through August 2012 and occurs before the passage of Hurricane Isaac.

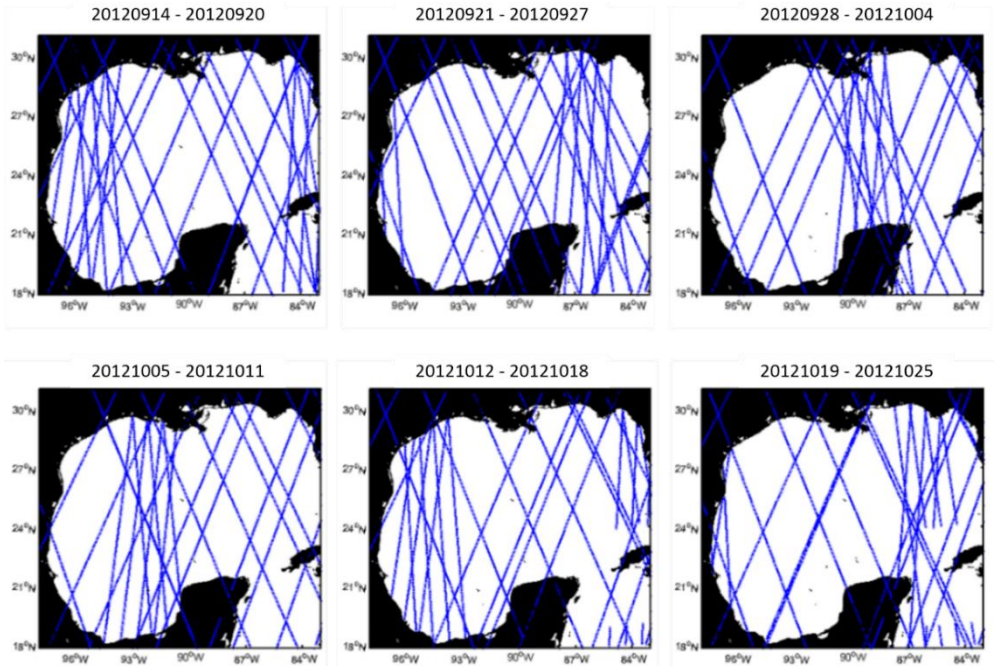

Figure 3.6 *Weekly Coverage of Altimeter Data: 20120914 through 20121025*

Plots include satellite track coverage from Jason-1, Jason-2, and Cryosat-2 missions. Time frame covers September through October 2012 and occurs after the passage of Hurricane Isaac.

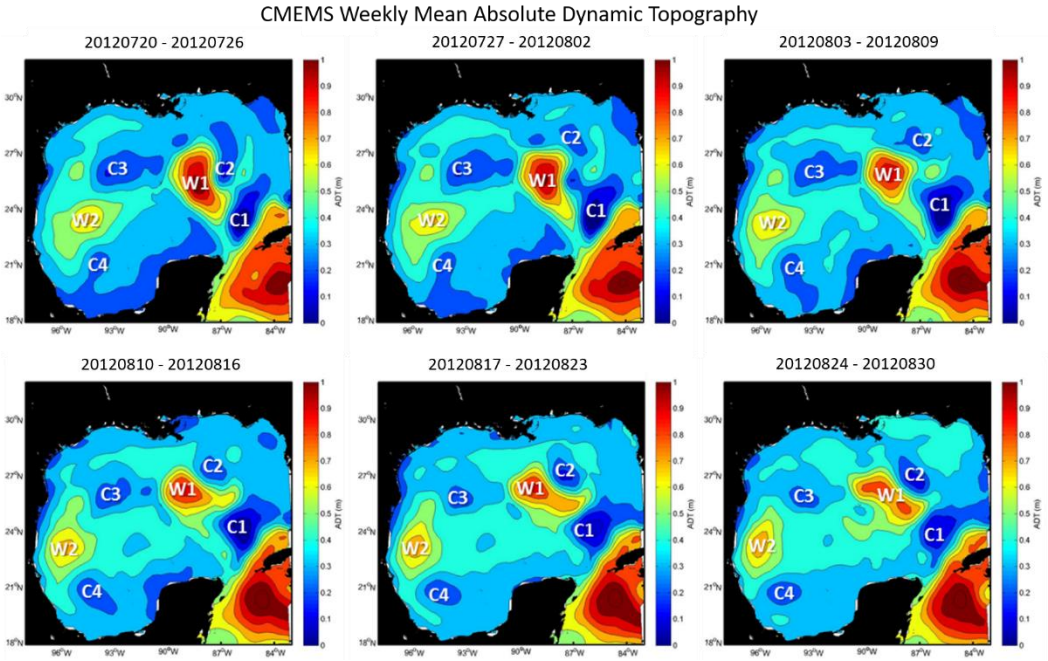

Figure 3.7 *CMEMS Dynamic Topography Fields: 20120720 through 20120830*

Weekly Mean Absolute Dynamic Topography from CMEMS. Time frame covers July through August 2012 and occurs before the passage of Hurricane Isaac.

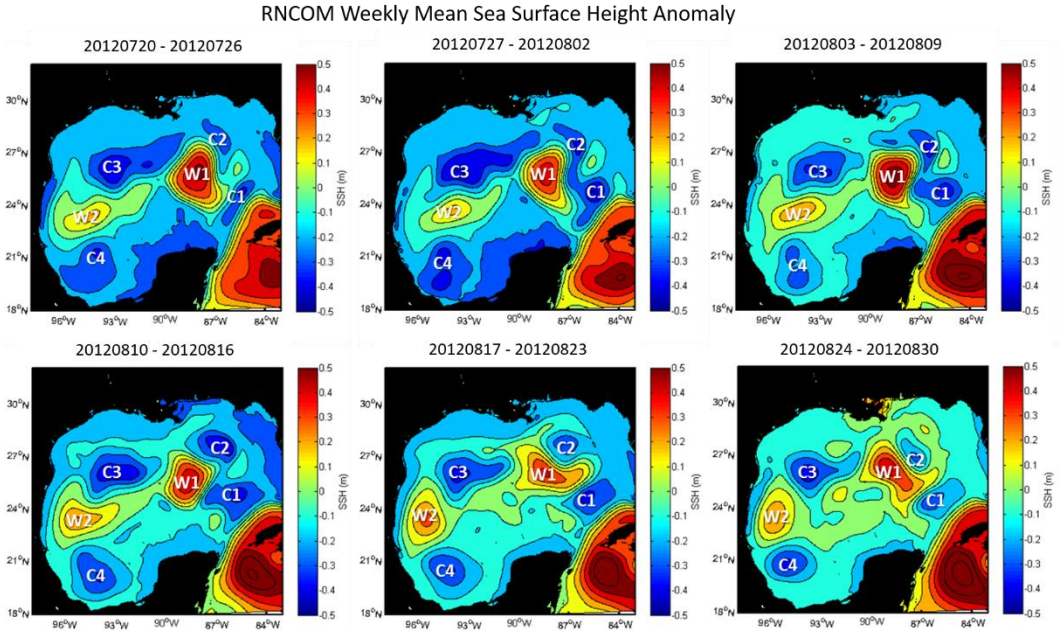

Figure 3.8 *RNCOM Mean Sea Surface Height Fields: 20120720 through 20120830*

Weekly Mean Sea Surface Height from RNCOM. Time frame covers July through August 2012 and occurs before the passage of Hurricane Isaac.

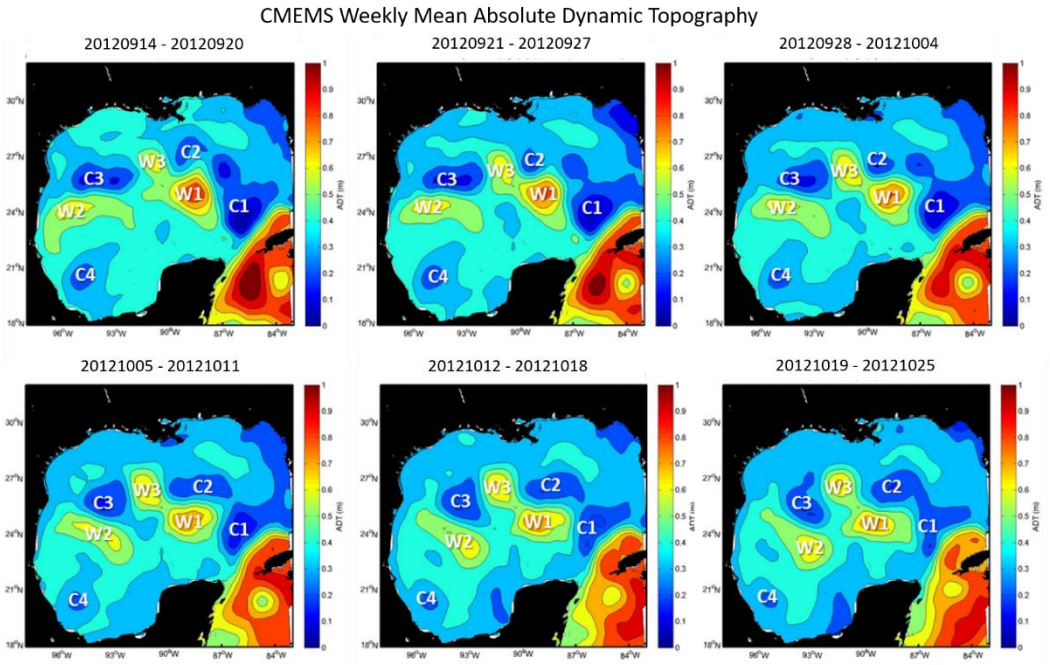

Figure 3.9 *CMEMS Dynamic Topography Fields: 20120914 through 20121025*

Weekly Mean Absolute Dynamic Topography from CMEMS. Time frame cover September through October 2012 and occurs after the passage of Hurricane Isaac.

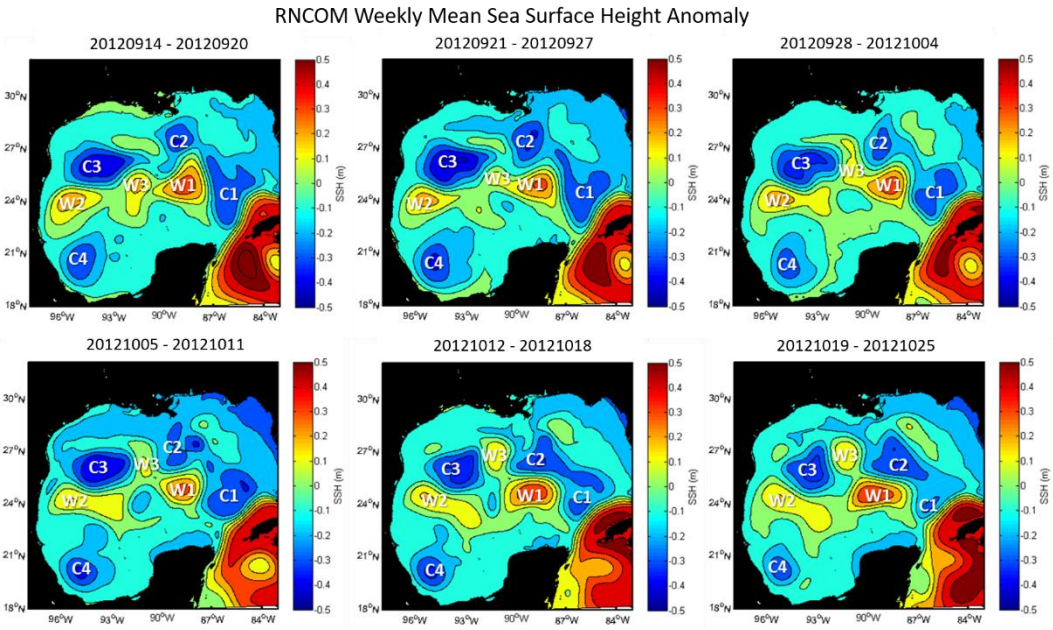

Figure 3.10 *RNCOM Mean Sea Surface Height Fields: 20120914 through 20121025*

Weekly Mean Sea Surface Height from RNCOM. Time frame cover September through October 2012 and occurs after the passage of Hurricane Isaac.

and two cyclonic features (C1 and C2), which are present to the east of W1. Two cyclonic features (C3 and C4) and one anti-cyclonic feature (W2) are also present in the western portion of the GOM basin but lie mostly outside the area sampled by the drifters (Figs. 2.4 and 2.5). It can be seen from the Absolute Dynamic Topography (ADT) fields in late August/early September, W1 begins to separate into two features (Fig. 3.7), shedding an anti-cyclonic eddy (W3) in mid-September (Fig. 3.9). RNCOM also shows the W3 feature (Fig. 3.8 and 3.10); however, the evolution of this event occurs differently than seen in the altimeter-derived ADT fields.

Differences in the evolution of the mesoscale eddy field has a huge impact on the Eulerian currents; this fact is highlighted by comparisons made between near-surface currents collected by two Acoustic Doppler Current Profilers (ADCP) mounted to NDBC buoy stations located near the shelf-break (Fig. 3.11). The ADCP data used were downloaded from NDBC's website. While the ADCP at station 42890, the inshore of the two stations, is located on a mobile offshore drilling unit, the ADCP at station 42360 is mounted to the buoy and is thus subject to buoy motion. ADCP depth bins used for each buoy are shown in Table 3.3. In order to remove the buoy's motion from current observations, NDBC ADCPs are set to collect data at one- to two-second intervals for five minutes each hour. The five-minute data record is then averaged to remove the motion of the buoy from the current observations. The data is also subject to a host of other quality control checks before being posted to NDBC's site (Crout and Wiley 2010). While these algorithms produce a buoy-mounted record that is largely free of buoy motion, a study done by Locke and Crout (2009) has shown that during high wind and

wave events buoy motion still likely degrades the quality of current measurements recorded.

The ADCP located farthest offshore, station 42360, captures the W3 shedding event, with the ADCP seeing northeastward currents of W3's western arm for most of the latter half of August and a change in rotation to southwestward currents around the 7th of September as W3 moves westward (Fig. 3.12). RNCOM shows the evolution of W3 less clearly than in the ADT fields. Two weaker anti-cyclonic areas (labeled N and S) to the east of W1 are seen in the RNCOM and ADT fields (Figs. 3.14 and 3.15). However,

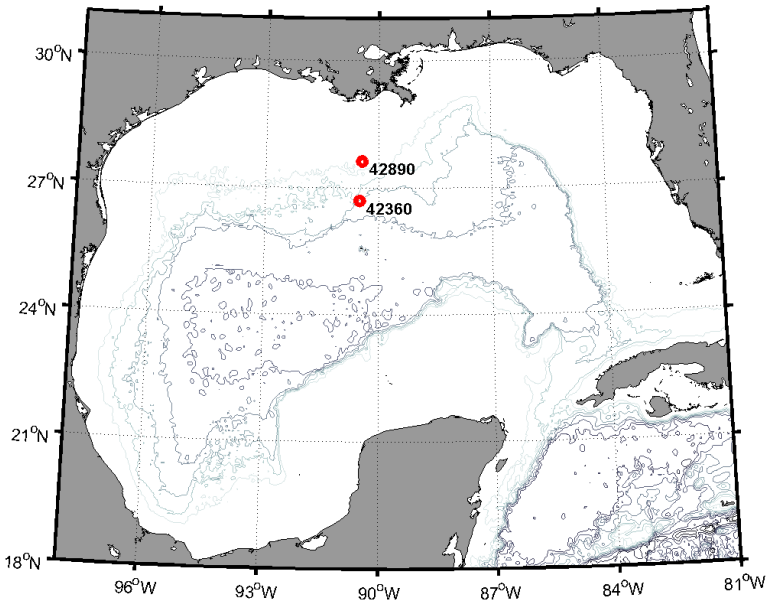

Figure 3.11 *Location of NDBC Stations Used in Open-Ocean Current Validation*

Buoy information obtained from https://www.ndbc.noaa.gov/.

Table 3.3 *NDBC Current Station Near-Surface ADCP Bin*

Station	ADCP Near-Surface Bin Depth
42360	2m
42890	15.2m

Note: Buoy information obtained from https://www.ndbc.noaa.gov/

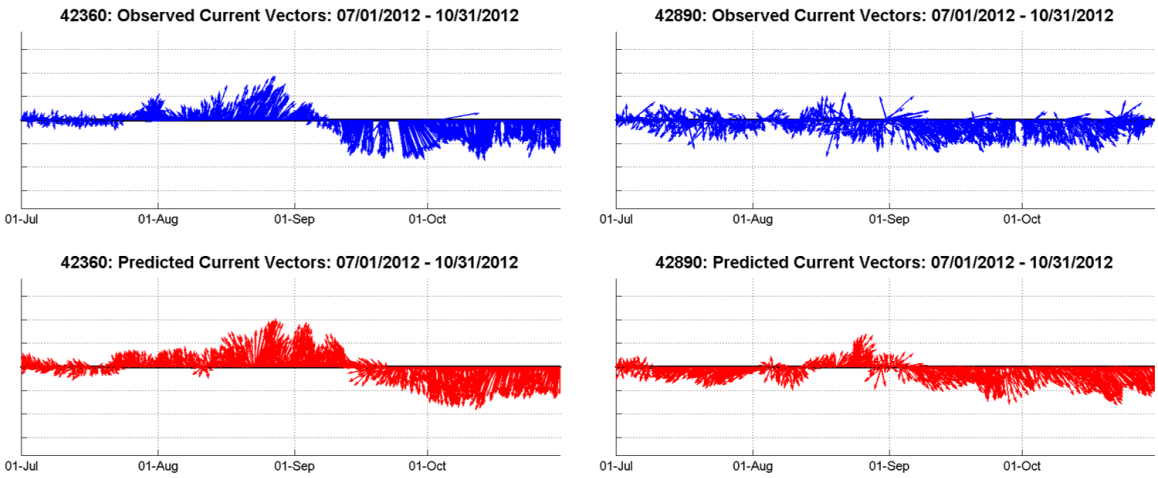

Figure 3.12 *Vector Time series of Observed ADCP versus RNCOM Predicted Near-Surface Currents*

Time series of observed ADCP current direction (blue) from NDBC buoys 42360 (left) and 42890 (right) and predicted RNCOM current direction (red) at the buoy location. The ADCP bin for 42360 is centered at 2 meters, and model currents are shown for the same depth. The ADCP bin for 42890 is centered at 15.2 meters, and model currents are shown for 15 meters; no spatial interpolation is done to account for this small difference. Observations and model predictions are interpolated to a uniform timescale using 1-hourly interval; however, interpolation is not performed where data gaps of greater than 1-day are present in the observed timeseries.

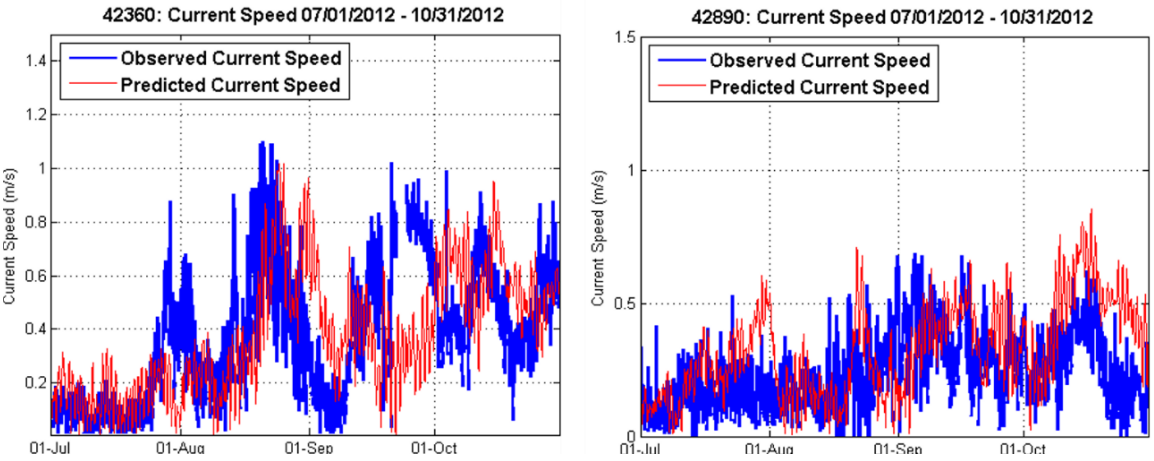

Figure 3.13 *Current Speed Time series of Observed ADCP versus RNCOM Predicted Near-Surface Currents*

Time series of observed ADCP current speed (blue) from NDBC buoys 42360 (left) and 42890 (right) and to predicted RNCOM current speed (red) at the buoy location. The ADCP bin for 42360 is centered at 2 meters, and model currents are shown for the same depth. The ADCP bin for 42890 is centered at 15.2 meters, and model currents are shown for 15 meters; no spatial interpolation is done to account for this small difference. Observations and model predictions are interpolated to a uniform timescale using 1-hourly interval; however, interpolation is not performed where data gaps of greater than 1-day are present in the observed timeseries.

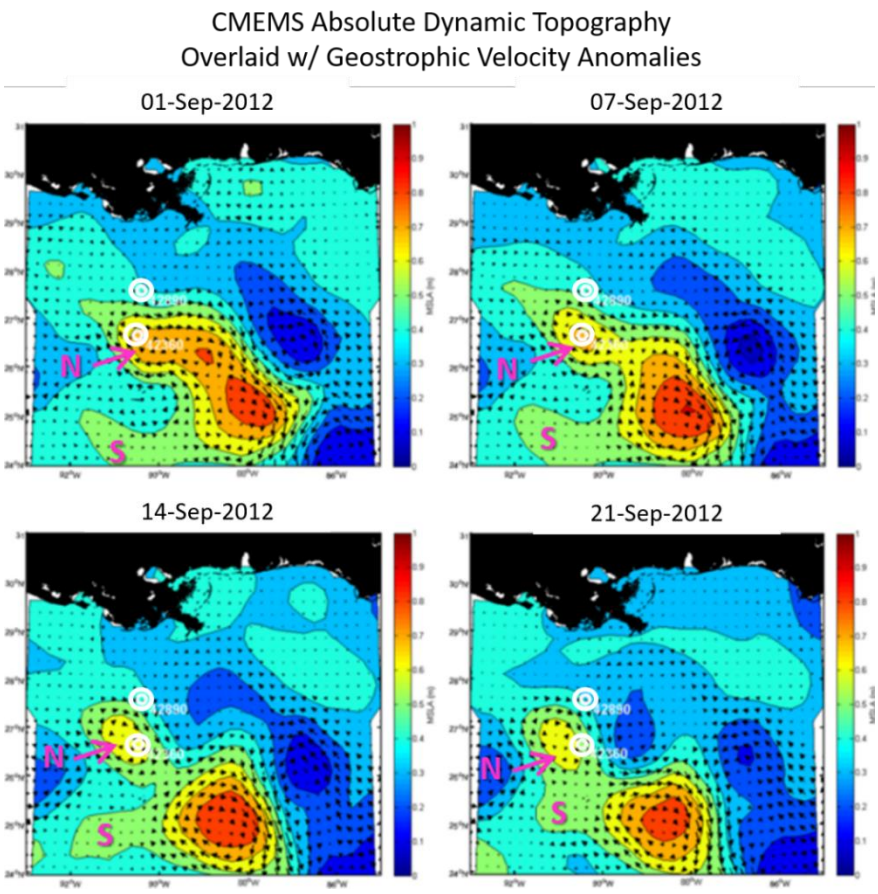

Figure 3.14 *CMEMS Observed Mesoscale Eddy Field Near Buoy Locations*

Daily snapshots of CMEMS Absolute Dynamic Topography (ADT) overlaid with geostropic velocity taken throughout the month of September showing the formation and shedding of W3. White dots indicate 42890 (northern station) and 42360 (southern station) locations.

where the northern feature, which eventually becomes W3, is strongest in the ADT fields, the southern feature is stronger in the RNCOM fields. Also, while the ADT fields clearly show W1 shedding W3 (Fig. 3.14), the RNCOM mean SSH fields do not show a clean shedding event or a continued coherent circulation for W3 over the same three-week period as the satellite-derived sea surface fields (Fig. 3.15). This difference in the evolution of the circulation at Station 42360 manifests itself as a delay in the rotation of currents from September 7th, as seen in the ADCP data, to sometime after the 14th in RNCOM (Fig. 3.12).

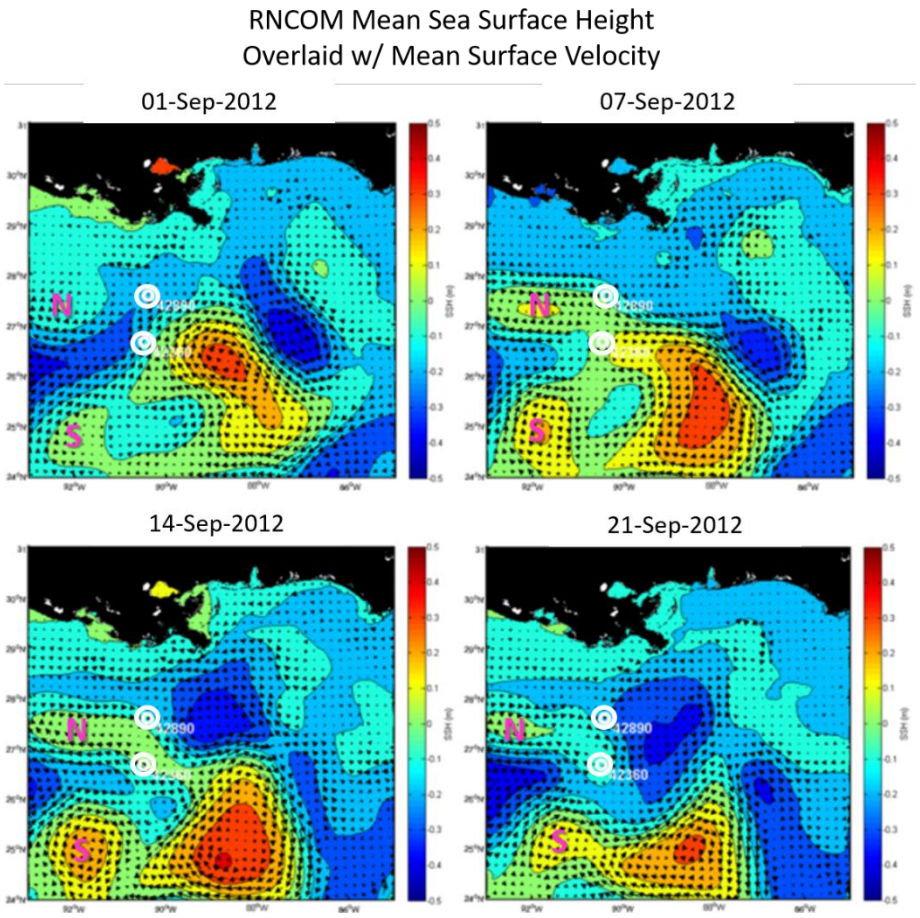

Figure 3.15 *RNCOM Predicted Mesoscale Eddy Field Near Buoy Locations*

Daily snapshots of RNCOM 24-hour averaged Sea Surface Height (SSH) fields overlaid with 24-hour averaged surface currents taken throughout the month of September showing the formation and shedding of W3. White dots indicate 42890 (northern station) and 42360 (southern station) locations.

Table 3.4 *Correlation Values for ADCP Observations and RNCOM*

Station	Vector Correlation (95% CI)	Speed Correlation (95% CI)
42360	0.66 (0.61-0.71)	0.43 (0.33-0.53)
42890	0.42 (0.36 – 0.48)	0.34 (0.25-0.44)

Note: Observations and model predictions are interpolated to a uniform timescale using 1-hourly interval. Interpolation is not performed where data gaps of greater than 1-day are present in the observed timeseries; these times are not included in correlation calculations. Correlations are significant to the 95% level and were calculated using 10,000 bootstrapped samples. Confidence intervals (CI) for correlations are given in parenthesis.

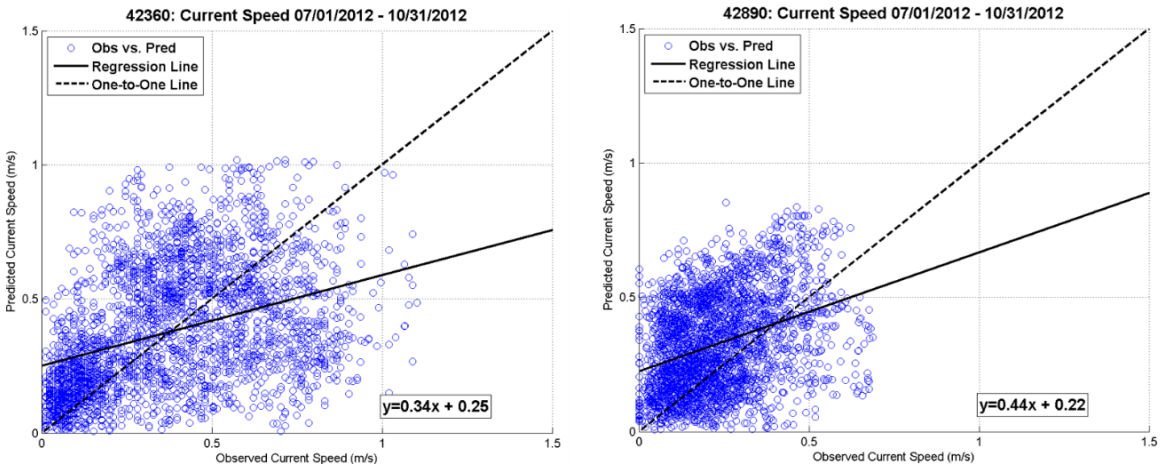

Figure 3.16 *Observed ADCP versus Predicted RNCOM Near-Surface Current Speed Comparison*

Linear regression of observed versus RNCOM modeled current speeds at NDBC buoy 42360 (left) and NDBC buoy 42890 (right). Observations and model predictions are interpolated to a uniform timescale using 1-hourly interval. Interpolation is not performed where data gaps of greater than 1-day are present in observed timeseries. These times are not included in regression analysis.

The ADCP at station 42890, located inshore of 42360, does not see any effects from the eddy field until early September, when W3 moves far enough north for the station to feel southeastward currents of W3's eastern arm (Fig. 3.12); RNCOM captures this current shift around the same time. However, while more variable currents are seen in this station's ADCP record during July, RNCOM predicted a much more coherent southwestward current, likely associated with C3 (Fig. 3.8). It is also worth noting that at several times during the W3 eddy shedding event, altimeter coverage over W1 and W3 is sparse (Fig. 3.6).

The discrepancies between the observed and modeled eddy field discussed above lead to poor vector and speed correlations between observed and RNCOM currents at the NDBC ADCP locations (Table 3.4). The scatter seen when plotting observed current speed against RNCOM modeled current speed (Fig. 3.16) is much greater than seen in the

wind speed comparison (Fig. 3.4), which is due to the differences noted between the observed and predicted eddy field.

Similar validation analyses are performed comparing ADCP-observed currents to filtered RNCOM predictions to better assess the impact that filtering has on velocity fields. Overall, very similar structure is predicted using filtered model velocities at both ADCP station locations (Fig. 3.12 and Fig. 3.17), but much lower current speeds are noted in general (Fig. 3.13 and Fig. 3.18) as is also evidenced by the much flatter regression lines for current magnitude (Fig. 3.19) than is seen in the unfiltered results (Fig. 3.16). Vector and speed correlations (Table 3.5) are not statistically different from those produced using unfiltered model velocities (Table 3.4). Current spectra from ADCP-observed currents were additionally compared to filtered and unfiltered model

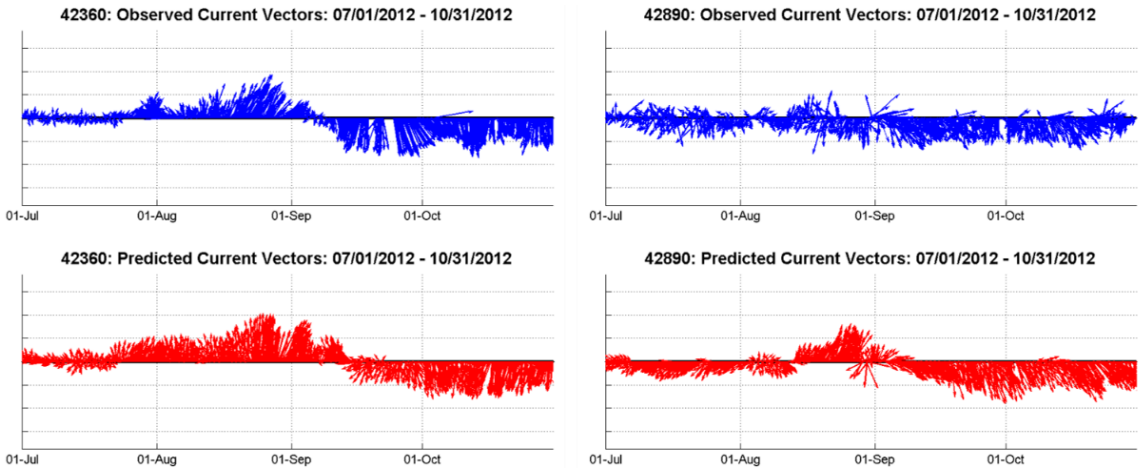

Figure 3.17 *Vector Time series of Observed ADCP versus Predicted Filtered-RNCOM Near-Surface Currents*

Time series of observed ADCP current direction (blue) from NDBC buoys 42360 (left) and 42890 (right) and filtered-RNCOM current direction (red) at the buoy location. The ADCP bin for 42360 is centered at 2 meters, and model currents are shown for the same depth. The ADCP bin for 42890 is centered at 15.2 meters, and model currents are shown for 15 meters; no spatial interpolation is done to account for this small difference. Observations and model predictions are interpolated to a uniform timescale using 1-hourly interval. Interpolation is not performed where data gaps of greater than 1-day are present in the observed timeseries.

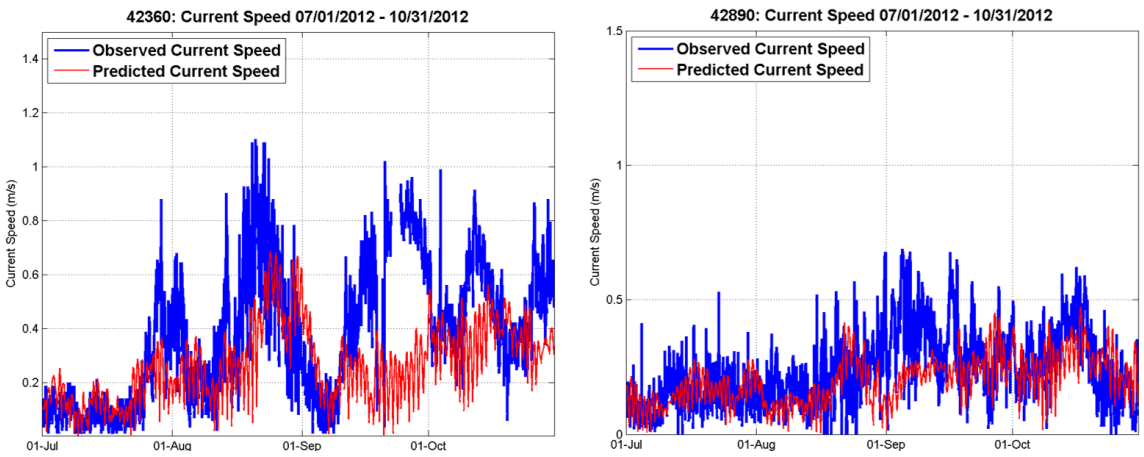

Figure 3.18 *Current Speed Time series of Observed ADCP versus Predicted Filtered-RNCOM Near-Surface Currents*

Time series of observed ADCP current speed (blue) from NDBC buoys 42360 (left) and 42890 (right) and to predicted filtered-RNCOM current speed (red) at the buoy location. The ADCP bin for 42360 is centered at 2 meters, and model currents are shown for the same depth. The ADCP bin for 42890 is centered at 15.2 meters, and model currents are shown for 15 meters; no spatial interpolation is done to account for this small difference. Observations and model predictions are interpolated to a uniform timescale using 1-hourly interval. Interpolation is not performed where data gaps of greater than 1-day are present in the observed timeseries.

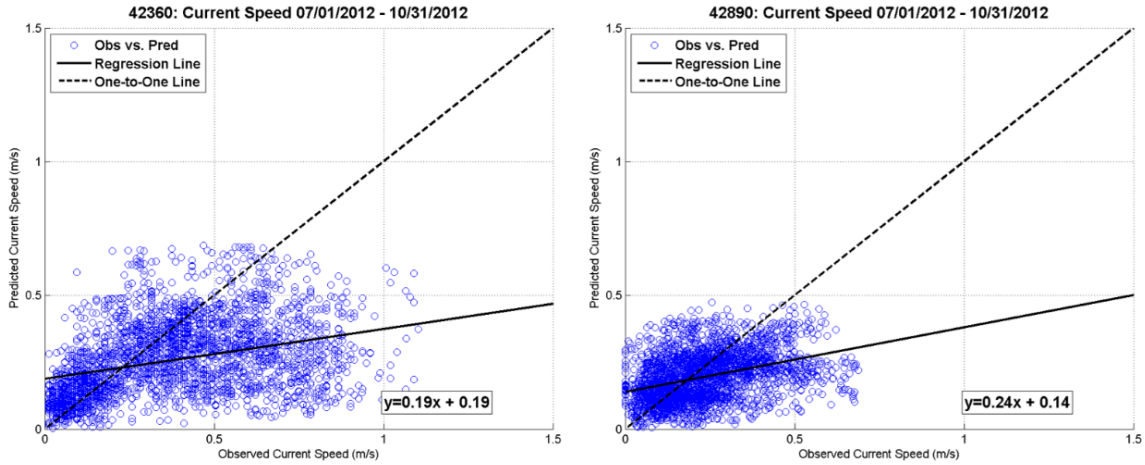

Figure 3.19 *Observed ADCP versus Predicted Filtered-RNCOM Near-Surface Current Speed Comparison*

Linear regression fit of observed versus filtered-RNCOM modeled current speeds at NDBC buoy 42360 (left) and NDBC buoy 42890 (right). Observations and model predictions are interpolated to a uniform timescale using 1-hourly interval. Interpolation is not performed where data gaps of greater than 1-day are present in the observed timeseries. These times are not included in regression analysis.

Table 3.5 *Correlation Values for ADCP Observations and Filtered-RNCOM*

Station	Vector Correlation (95% CI)	Speed Correlation (95% CI)
42360	0.70 (0.66 – 0.75)	0.42 (0.32 – 0.53)
42890	0.39 (0.33 – 0.46)	0.36 (0.27 – 0.45)

Note: Observations and model predictions are interpolated to a uniform timescale using 1-hourly interval; however, interpolation is not performed on observed timeseries where data gaps of greater than 1-day are present, and these times are therefore not included in correlation calculations. Correlations are significant to the 95% level and were calculated using 10,000 bootstrapped samples. Confidence intervals (CI) for correlations are given in parenthesis.

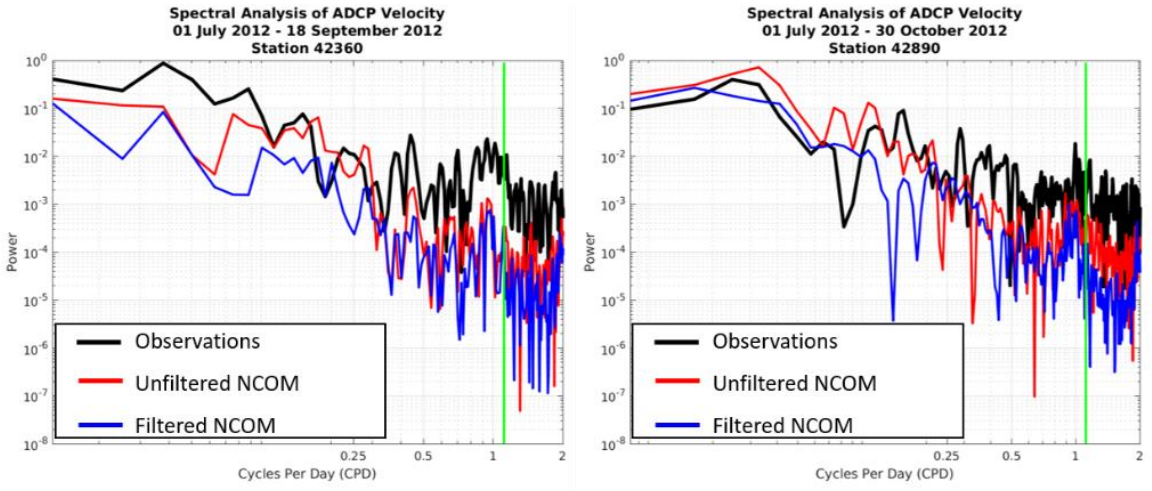

Figure 3.20 *Comparisons of Current Velocity Spectra for ADCP Observations, Unfiltered NCOM, and Filtered NCOM*

Discrete Fourier Transform (DFT) of observed and modeled currents at NDBC buoy locations. Spectra at NDBC buoy 42360 is shown on the left and NDBC buoy 42890 is shown on the right. Observed spectra are shown in black, unfiltered RNCOM spectra are shown in red, and filtered RNCOM spectra are shown in blue. The inertial frequency for the station's latitude is indicated by the green line. As temporal interpolation between large gaps in the observation data may introduce artifacts in the timeseries, for records where data gaps of 1-day or greater are present, spectra is computed using the largest continuous block of data (i.e. time gaps of less than 1-day).

predictions in order to ensure that the time-independent spatial filtering applied to model fields did not introduce any unexpected artifacts to the timeseries data. Fig. 3.20 shows

that observed currents and model predicted currents at the NDBC stations are dominated by low-frequency energy, which corresponds to mesoscale activity at station 42360 and mesoscale activity along with weather at station 42890. The spectra also show that spatial filtering results in an overall decrease in energy content of the velocity field, and no spectral artifacts are introduced as a result of filtering. The false energy spikes are suppressed in the filtered spectra bringing predictions more in line with observed, which could be a source of error reduction in the filtered modeled performance and a phenomenon that is not captured well by vector and speed correlation statistics.

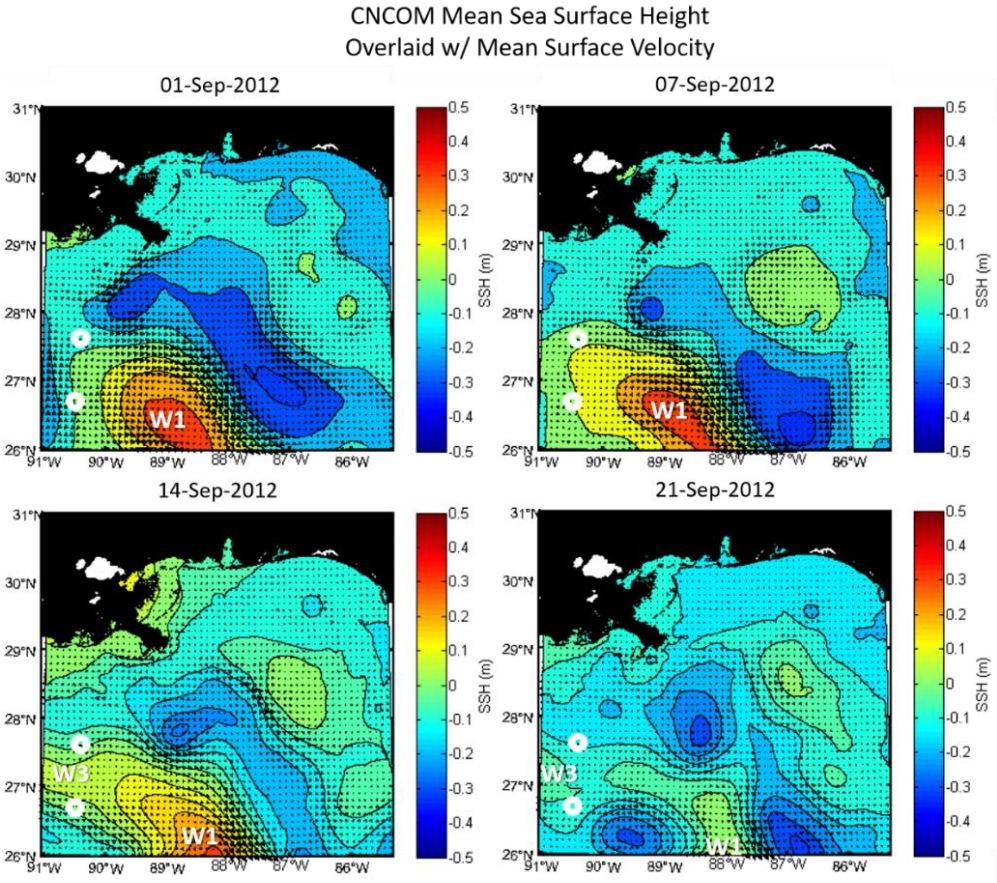

Figure 3.21 *CNCOM Predicted Mesoscale Eddy Field Near Buoy Locations*

Daily snapshots of CNCOM 24-hour averaged Sea Surface Height (SSH) fields overlaid with 24-hour averaged surface currents taken throughout the month of September. White dots indicate 42890 (northern station) and 42360 (southern station) locations.

Because most of the mesoscale activity lies outside the CNCOM domain as its southern boundary extends just over the shelf break to 26°N (Fig. 1.3), a comparison of the model's circulation pattern to the altimeter derived ADT field is more limited. W1 is seen in CNCOM current fields, but the shedding of W3 is very weak (Fig. 3.21). Similarly to the RNCOM result, the mismatch in the evolution of the eddy field in CNCOM leads the differences in the current field predicted by the high resolution model at the ADCP locations (Fig. 3.22 and Fig. 3.23) However, it must be noted that the difficulty of the CNCOM to predict these features may be due, in part, to the close proximity of the ADCP stations to the model's western boundary.

Where the ADCP at 42360 records the rotation of currents associated with the passage of W3, CNCOM only captures northwestern currents from late August to early September associated with western arm of W1 (Fig. 3.22). Modeled currents weaken after this when W1 shifts southward almost completely outside the CNCOM domain. Eventually CNCOM does form W3, and southwestward currents associated with W3's southeastern boundary are seen throughout much of October. Observed and CNCOM modeled currents at station 42890 show a more similar, variable current field through the first half of the study period, when the station location lies outside mesoscale activity (Fig. 3.22). However, throughout much of the latter half of the study period, southeastward currents are observed at 42890 and are associated with W3's northeastern boundary, but this feature is never located far enough to the north in CNCOM to affect modeled currents at the buoy's location. Just as for the RNCOM comparisons, discrepancies between the observed and modeled eddy field lead to poor vector and speed

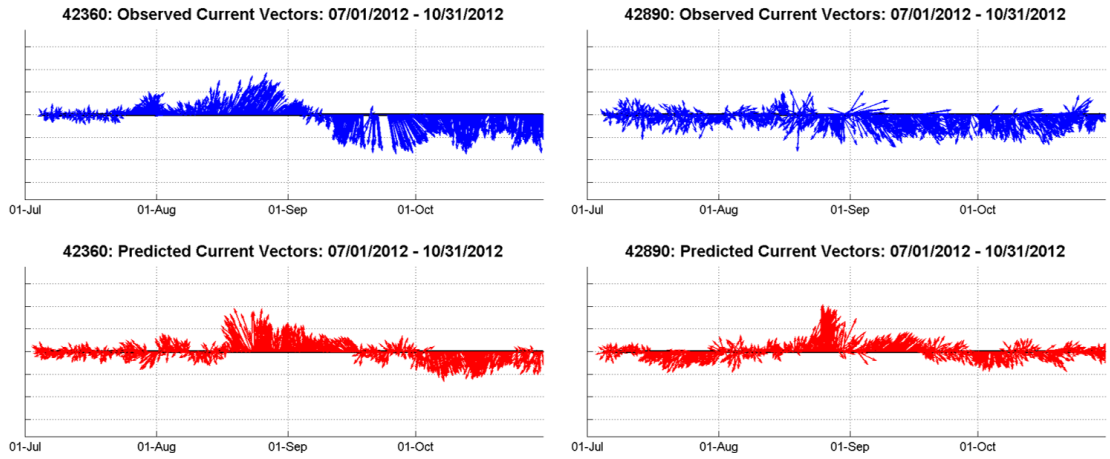

Figure 3.22 *Vector Time series of Observed ADCP versus Predicted CNCOM Near-Surface Currents*

Time series of observed ADCP current direction (blue) from NDBC buoys 42360 (left) and 42890 (right) and predicted CNCOM current direction (red) at the buoy location. The ADCP bin for 42360 is centered at 2 meters, and model currents are shown for the same depth. The ADCP bin for 42890 is centered at 15.2 meters, and model currents are shown for 15 meters; no spatial interpolation is done to account for this small difference. Observations and model predictions are interpolated to a uniform timescale using 1-hourly interval; however, interpolation is not performed on observed timeseries where data gaps of greater than 1-day are present.

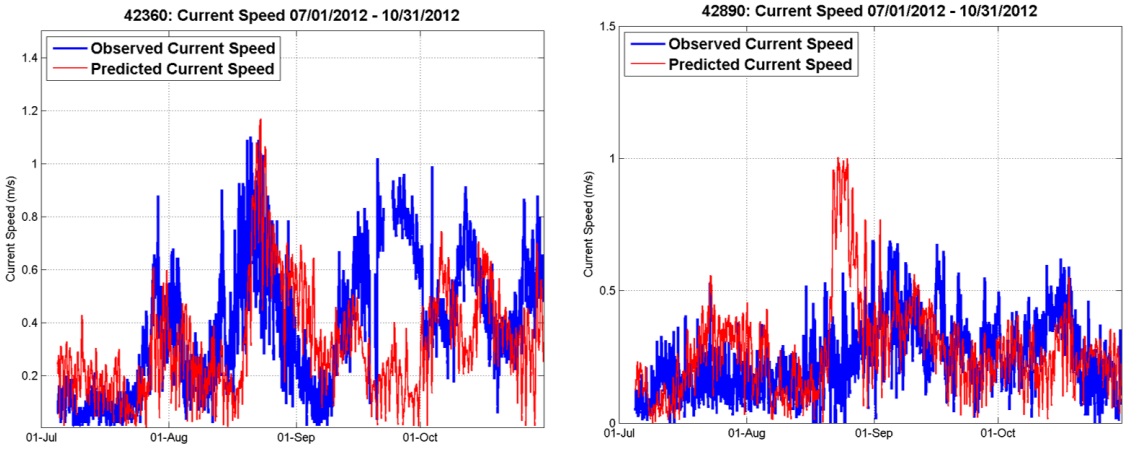

Figure 3.23 *Current Speed Time series of Observed ADCP versus Predicted CNCOM Near-Surface Currents*

Time series of observed ADCP current speed (blue) from NDBC buoys 42360 (left) and 42890 (right) and to predicted CNCOM current speed (red) at the buoy location. The ADCP bin for 42360 is centered at 2 meters, and model currents are shown for the same depth. The ADCP bin for 42890 is centered at 15.2 meters, and model currents are shown for 15 meters; no spatial interpolation is done to account for this small difference. Observations and model predictions are interpolated to a uniform timescale using 1-hourly interval. Interpolation is not performed where data gaps of greater than 1-day are present in the observed timeseries.

Table 3.6 *Correlation Values for ADCP Observations and CNCOM*

Station	Vector Correlation (95% CI)	Speed Correlation (95% CI)
42360	0.48 (0.41 – 0.56)	0.30 (0.18 – 0.43)
42890	0.26 (0.18 – 0.43)	0.22 (0.11 – 0.32)

Note: Observations and model predictions are interpolated to a uniform timescale using 1-hourly interval; however, interpolation is not performed where data gaps of greater than 1-day are present in the observed timeseries. These times are not included in correlation calculations. Correlations are significant to the 95% level and were calculated using 10,000 bootstrapped samples. Confidence intervals (CI) for correlations are given in parenthesis.

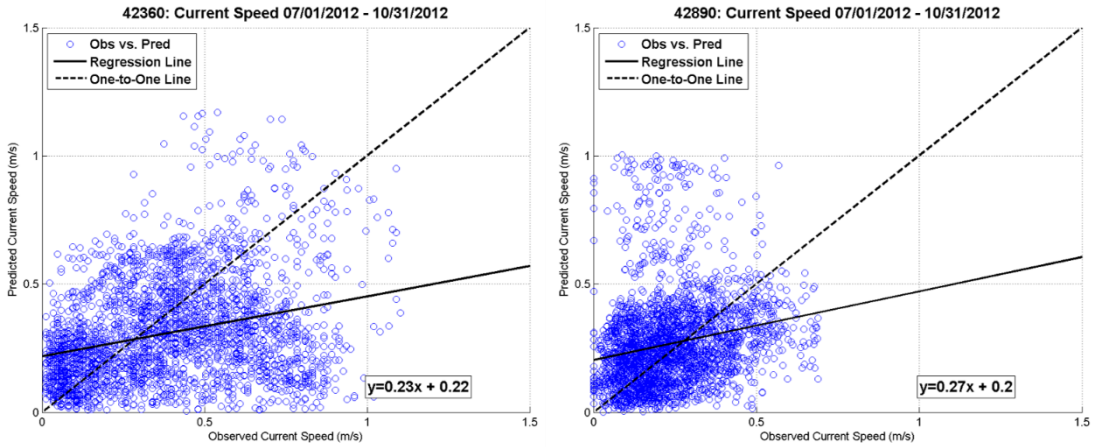

Figure 3.24 *Observed ADCP versus Predicted CNCOM Near-Surface Current Speed Comparison*

Linear regression fit of observed versus CNCOM modeled current speeds at NDBC buoy 42360 (left) and NDBC buoy 42890 (right). The ADCP bin for 42360 is centered at 2 meters, and model currents at the same depth are used in the comparison. The ADCP bin for 42890 is centered at 15.2 meters, and model currents at 15 meters are used in this comparison; no spatial interpolation is done to account for this small difference. Observations and model predictions are interpolated to a uniform timescale using 1-hourly interval; however, interpolation is not performed on observed timeseries where data gaps of greater than 1-day are present, and these times are therefore not included in regression analysis.

correlations at the NDBC ADCP locations as well as poor regression analysis (Table 3.6 and Fig. 3.24).

Overall, the ADCP comparisons between RNCOM and CNCOM highlight the difficulty FNMOC's ocean models have with predicting currents at a specific location. Without any doubt this be attributed to the data sparsity in the ocean and the data

assimilation's inability to more accurately constrain the placement of the boundaries of mesoscale features. Even though both NCOM models predict a similar circulation pattern with that of the altimeter derived product, variation in the movement and evolution of the mesoscale field leads to large, sometimes significant, differences in the current field than is observed. Additionally, currents are likely more uncertain than other state variables due to the fact that velocity is not an assimilated data field by FNMOC's operational data assimilation software. All of these points highlight the necessity of looking at current prediction in a more probabilistic manner with the present operational ocean modeling capability.

3.3 Shelf Eulerian Ocean Current Validation

Turning to the shelf, where currents are largely wind-driven (Ohlmann and Niiler 2005; Nowlin et al. 2005), accuracy of the current field will depend a great deal on the accuracy of the atmospheric forcing. High-frequency (HF) radar measurements of surface currents on the Northern Gulf Shelf from the Gulf of Mexico Coastal Ocean Observation System (GCOOS) monitoring stations are available for portions of the study period. HF radar coverage for the monitoring stations during July through August 2012 is shown in Fig. 3.25. Three points in the highest sampled regions were chosen for time series analysis. Fig. 3.26 shows observed and modeled shelf currents dominated by oscillatory events; however, the phasing of the oscillations in the predicted time-series is a notable issue. Velocity statistics indicate a very weak mean flow and much, much larger standard deviation values in both the observation and model data. Speed statistics are comparable between observed and modeled fields (Table 3.7).

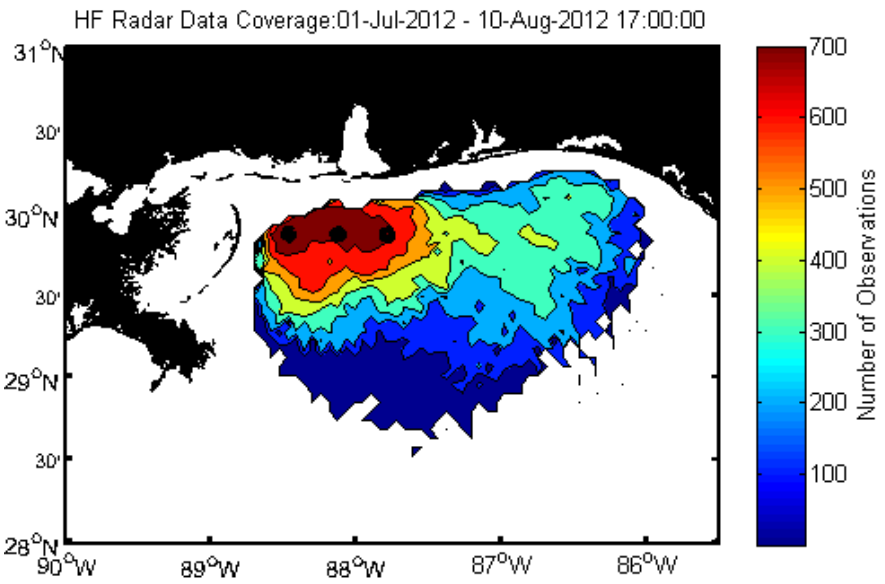

Figure 3.25 *HF Radar Station Coverage on the Northern Gulf of Mexico Shelf*

Black dots indicate locations chosen for time series analysis.

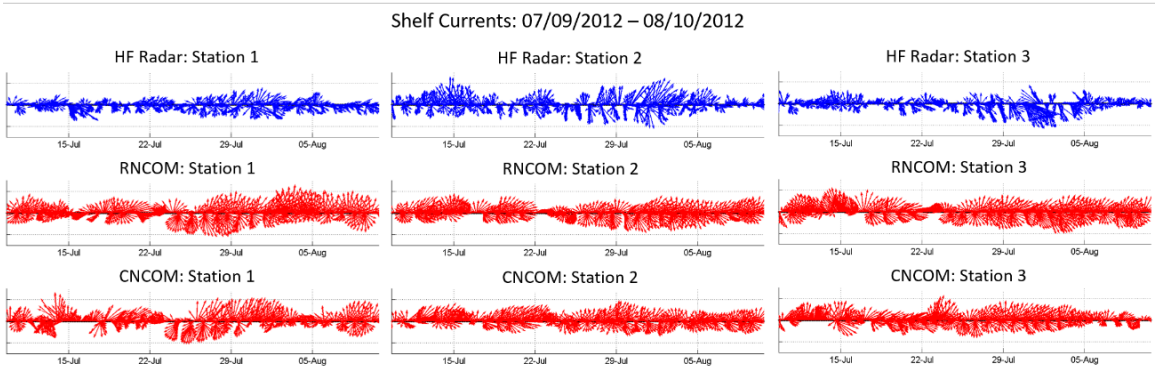

Figure 3.26 *Vector Time series of Observed HF-Radar versus Predicted Near-Surface Currents*

Time series of surface currents recorded the three HF Radar locations chosen (Fig. 3.25). Left Panel: shows the time series for the western station. Center Panel: shows the time series of the central station, and panel Right Panel: shows the time series of the eastern station. Observed currents (blue) are shown in the top row. RNCOM currents (red) are shown in the middle row, and CNCOM currents (red) are shown in the bottom row. Observations and model predictions are interpolated to uniform timeseries using at a 1-hourly interval.

Table 3.7 *Statistics of Observed and Predicted Surface Currents within HF Radar Coverage*

Station		Mean U-component (95% CI)	Std Dev U-component (95% CI)	Mean V-component (95%)	Std Dev V-component (95%)	Mean Speed (95%)	Std Dev Speed (95%)
1	HF-Radar	0.08 (0.07 – 0.09)	0.17 (0.17 – 0.18)	-0.05 (-0.05 - -0.04)	0.15 (0.14 – 0.15)	0.22 (0.21 – 0.22)	0.11 (0.10 – 0.11)
	RNCOM	0.02 (0.01 – 0.03)	0.13 (0.13 – 0.14)	0.00 (0.00 – 0.01)	0.13 (0.13 – 0.14)	0.18 (0.17 – 0.18)	0.08 (0.07 – 0.08)
	CNCOM	0.01 (0.00 – 0.02)	0.12 (0.12 – 0.12)	0.01 (0.00 – 0.02)	0.14 (0.13 – 0.14)	0.16 (0.16 – 0.17)	0.09 (0.08 – 0.09)
2	HF-Radar	0.04 (0.03 – 0.04)	0.14 (0.13 – 0.14)	0.01 (0.00 – 0.02)	0.20 (0.19 – 0.21)	0.22 (0.21 – 022)	0.12 (0.11 – 0.12)
	RNCOM	0.08 (0.07 – 0.09)	0.16 (0.16 – 0.16)	0.00 (0.00 – 0.01)	0.14 (0.14 – 0.15)	0.21 (0.20 – 0.21)	0.10 (0.09 – 0.10)
	CNCOM	0.10 (0.09 – 0.11)	0.20 (0.19 – 0.20)	0.03 (0.02 – 0.04)	0.18 (0.18 – 0.19)	0.27 (0.26 – 0.27)	0.11 (0.11 - 0.12)
3	HF-Radar	0.07 (0.06 – 0.08)	0.15 (0.14 – 0.16)	-0.05 (-0.05 - -0.04)	0.18 (0.18 – 0.19)	0.21 (0.20 – 0.21)	0.14 (0.14 – 0.15)
	RNCOM	0.07 (0.06 – 0.08)	0.15 (0.15 – 0.16)	0.00 (0.00 – 0.01)	0.14 (0.14 – 0.15)	0.20 (0.20 – 0.21)	0.09 (0.08 – 0.09)
	CNCOM	0.10 (0.09 – 0.11)	0.18 (0.17 – 0.18)	0.03 (0.02 – 0 .04)	0.18 (0.17 – 0.18)	0.24 (0.24 – 0.25)	0.12 (0.12 – 0.13)

Note: Model predictions are interpolated to observed times for mean and standard deviation calculations. Statistics are significant to the 95% level and were calculated using 10,000 bootstrapped samples, taken from over 700 HF-radar collections (N=771). Confidence intervals (CI) for statistics are given in parenthesis.

Spectral analysis for the three locations within the HF-radar coverage shows the oscillations are in fact near-inertial oscillations, with observed and predicted surface current spectra dominated by energy at near and sub-inertial frequencies (Fig. 3.27). Where observed spectra show near-inertial energy decreasing moving westward, modeled spectra show energy in the near-inertial range captured at the westernmost station, with this spectral peak decreasing moving eastward. While a great deal of variability exists in

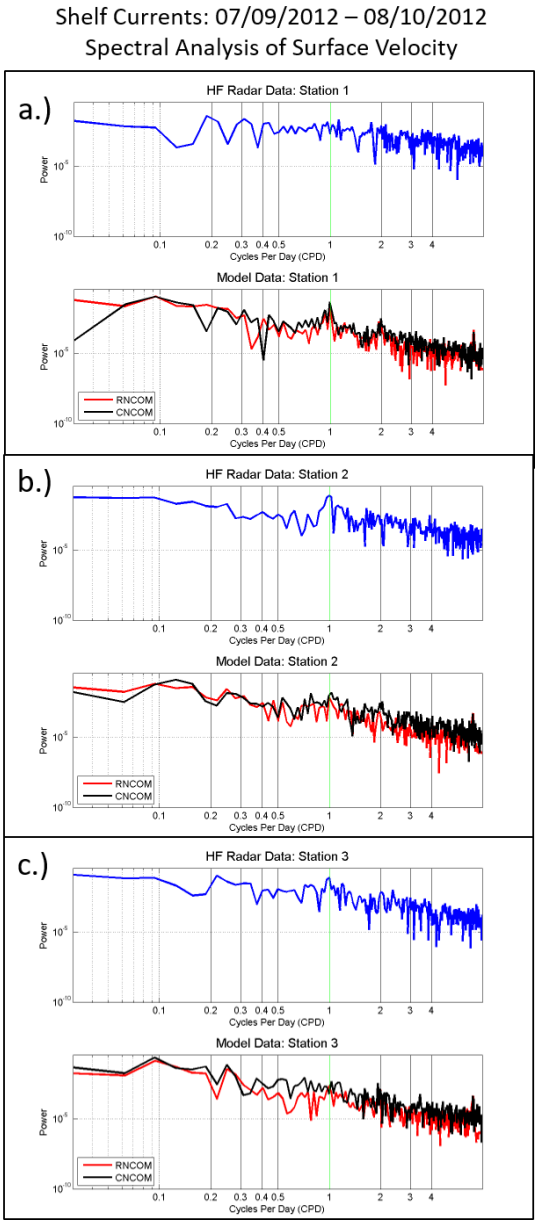

Figure 3.27 *Spectral Analysis of Surface Currents within HF Radar Coverage*

Discrete Fourier Transform (DFT) of current velocities at three different stationary locations within HF Radar coverage (Fig. 3.25). The westernmost station spectra are shown in the top panel (a), the central station spectra are shown in the middle panel (b), and the easternmost station spectra are shown in the bottom panel (c). Spectra from observed currents (blue) are shown in the top sub-plot. Spectra from RNCOM (red) and CNCOM (black) modeled currents are bottom sub-plot. The inertial frequency at the latitudinal location of the HF Radar station is indicated by the green vertical line in all cases.

the low-frequency energy content moving across the three locations, there is a consistent energy spike seen in the frequency band of ~3-11-day periodicity and is wind-driven in

nature, corresponding to the low-frequency signature seen in wind velocity observations recorded by a nearby NDBC station (Fig. 3.28 and Fig. 3.29). This result is consistent with previous findings that shelf currents on the northern GOM shelf are largely controlled by weather events (Ohlmann and Niiler 2005; Nowlin et al. 2005). Also noteworthy is the absence of a land sea breeze (LSB) in the observed wind data, which was inaccurately predicted by the COAMPS atmospheric model (Fig. 3.29). While overall the atmospheric forcing model accurately predicts large-scale weather events (Fig. 3.2 – 3.4), this nearshore comparison highlights a synoptic scale event that was not well-predicted and likely had adverse impacts on the ocean model's ability to forecast phenomena like near-inertial waves.

Because the strength of inertial oscillations is connected to the surface mixed layer depth (MLD), CTD profiles that were taken in the Northern Gulf of Mexico (Fig. 3.30) during the GLAD experiment are compared to model predicted MLD (Fig. 3.31) to determine if the deficiency in the model's ability to predict inertial oscillations is related to errors in the modeled MLD. Qualitatively, very shallow mixed layers are seen in

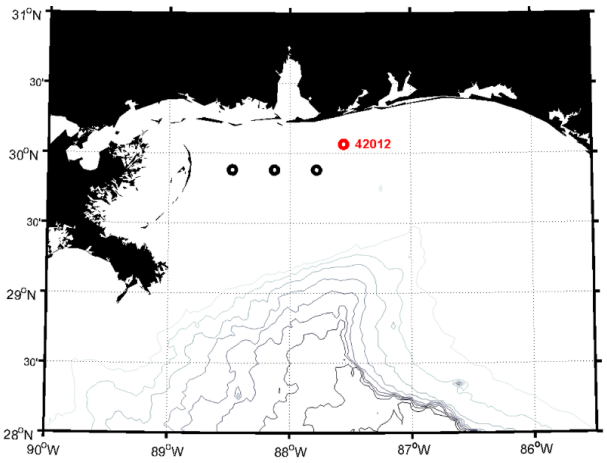

Figure 3.28 *NDBC Buoy closest to HF Radar Analysis Points*

HF radar time-series locations are shown by black circles. NDBC buoy 42012 is shown by the red circle.

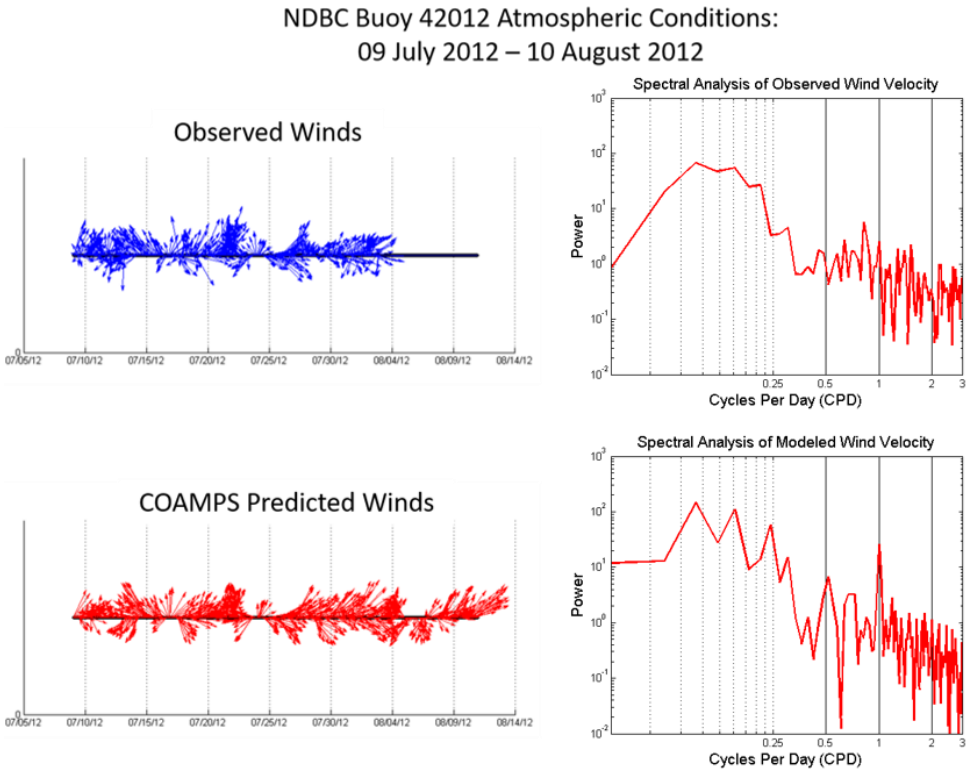

Figure 3.29 *Atmospheric Conditions Coinciding with HF Radar Current Analysis*

Atmospheric conditions shown over the same time period as the HF Radar current analysis (Fig. 3.26 and Fig. 3.27). Left-hand graphics show timeseries of wind velocity at NDBC Station 42012 (Fig. 3.28), with observed winds shown at the top and COAMPS predicted winds shown at the bottom. Right-hand graphics show the spectra of the wind velocity timeseries shown on the left, with observed shown on top and predicted shown on bottom.

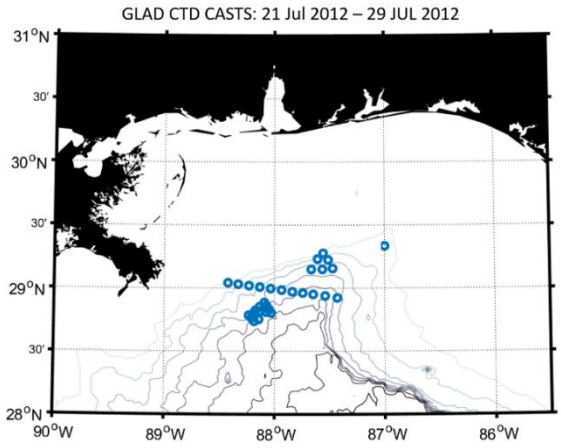

Figure 3.30 *GLAD In-Situ CTD Cast Locations*

Locations of 33 CTD observations taken coincidentally during the GLAD drifter experiment. In-situ observations were collected during an overlapping time period as the HF surface current analysis shown in Fig. 3.26 and Fig. 3.27.

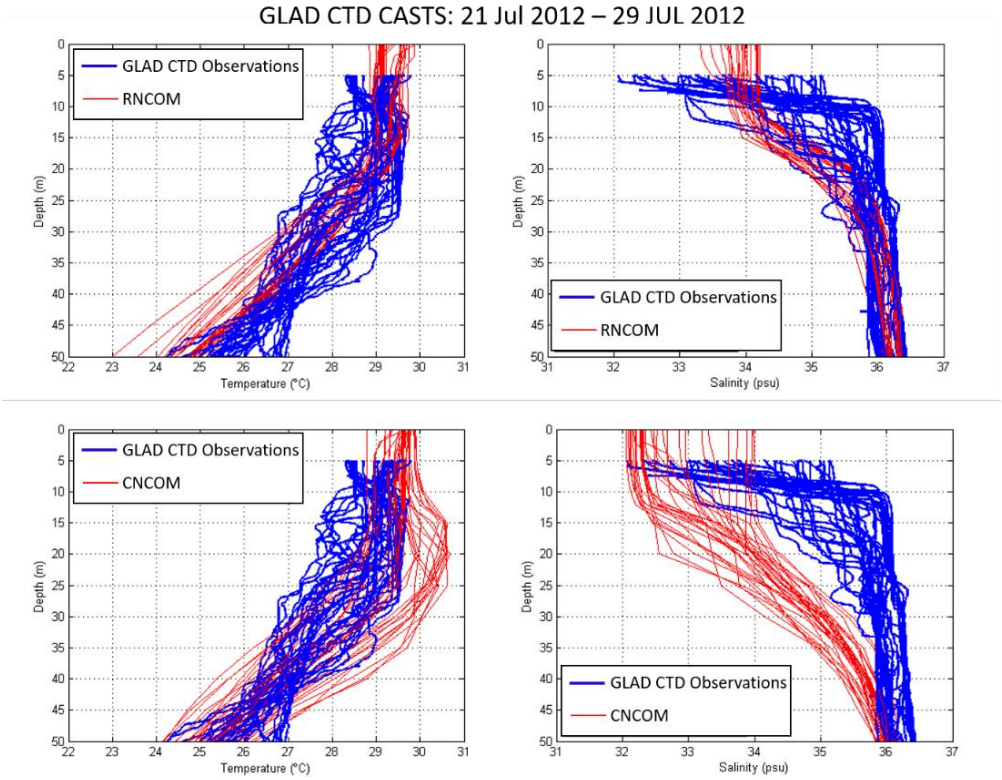

Figure 3.31 *GLAD In-Situ CTD Casts Temperature and Salinity Profiles*

Temperature and salinity observations taken during the GLAD experiment (blue) plotted along with predicted temperature and salinity profiles (red) at the same time and location as the observed profiles. Locations of the 33 observations are shown in Fig. 3.30.

observations and model predictions; however, much more structure is seen in observed temperatures in the upper 50-meters than the NCOM models show (Fig. 3.31). The near-surface, lower salinity signature seen in observations is also not predicted by either model (Fig. 3.31). The model's misrepresentation of near-surface salinity is likely due to the crude treatment of freshwater input via climatological river discharge values, as this area is impacted by outflow from the Mississippi River. While both RNCOM and CNCOM predict similar mean and standard deviation statistics for MLD (Table 3.8), error analysis shows that synoptic MLD differed from modeled MLD by at least 5-meters in over half of the comparisons (Fig. 3.32).

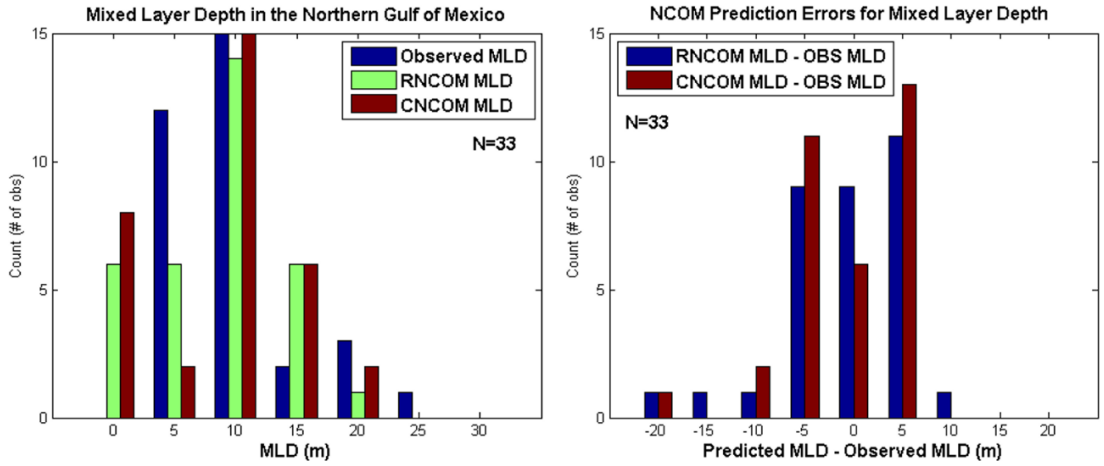

Figure 3.32 *Mixed Layer Depth Characteristics of GLAD CTD Observations and Modeled Profiles*

Mixed Layer Depth characteristics calculated from the 33 GLAD CTD observations shown in Fig. 3.30 and Fig. 3.31. The algorithm used defines mixed layer depth (MLD) as the depth in which all temperature values from the surface to that depth are within 0.1°C of the mean temperature of the layer. The distribution of the MLD for the in-situ profiles (blue) and corresponding RNCOM predicted profiles (green) and CNCOM predicted profiles (red) are shown on the left. Modeled MLD errors are shown on the right with RNCOM errors shown in blue and CNCOM errors shown in red.

Table 3.8 *Statistics of GLAD CTD Observations and Predicted Mixed Layer Depth*

	Mean MLD (95% CI)	**Std Dev MLD (95% CI)**
Observations	9.8 (9.2 – 10)	5.3 (4.0 – 5.3)
RNCOM	8.8 (8.2 – 9.9)	5.6 (4.9 – 5.9)
CNCOM	9.3 (8.6 – 9.9)	5.9 (5.2 – 5.9)

Note: Statistics are significant to the 95% level and were calculated using 10,000 bootstrapped samples, taken from 33 CTD observations and their corresponding modeled profiles (N=33). Degrees of freedom (DOF) was defined at the number of observations, in this case. Confidence intervals (CI) for statistics are given in parenthesis.

While not conclusive, the errors in NCOM's prediction of MLD along with errors in atmospheric forcing could be contributing factors to the mismatch seen between observed and modeled near-inertial signatures. Regardless, the model's weakness in predicting the near-inertial energy content as well as the incorrect prediction of the spatial

and temporal variability of these oscillations is noted as a potential area of concern for drift predictions in shallow water.

3.4 Upper-Ocean Velocity Profile Validation

Because the ocean's near-surface velocity is heavily influenced by boundary layer processes (Anis and Moum 1995), the last section of this chapter takes a closer look at the velocity profile of the near-surface layer. The need for analysis of the near-surface velocity profile is two-fold: 1.) to show that the NCOM models predict a near-surface velocity that is of comparable magnitude to that experienced in the real ocean, and 2.) to highlight the impact that the choice of the model velocity field's depth level has on the drift integration (i.e. would using the current field at a slightly shallower or deeper model depth produce significantly different drift results?). Both of these points will be demonstrated in the remainder of this section, accompanied by discussion of modeling considerations with respect to controlling processes of the near-surface layer.

Surface fluxes of heat and momentum from wind stress and those induced by surface waves are the dominant influences affecting the near-surface boundary layer (Anis and Moum 1995). Although surface fluxes of heat can affect the turbulent kinetic energy (TKE) in the near surface layer (i.e., diurnal cycle of surface warming acts to suppress TKE, and convective cooling acts to increase it), the focus here will be on those factors controlling the momentum flux and its subsequent vertical transfer. The momentum equations used by NCOM (Barron et al. 2006) are as follows

$$\frac{\partial u}{\partial t} + \nabla \cdot (\boldsymbol{v}u) = -\frac{1}{\rho_0}\frac{\partial p}{\partial x} + fv + Q_u + \nabla_H(A_M \nabla_H u) + \frac{\partial}{\partial z}(K_M \frac{\partial u}{\partial z}) \qquad (15)$$

$$\frac{\partial v}{\partial t} + \nabla \cdot (\boldsymbol{v}v) = -\frac{1}{\rho_0}\frac{\partial p}{\partial y} - fu + Q_v + \nabla_H(A_M \nabla_H v) + \frac{\partial}{\partial z}(K_M \frac{\partial v}{\partial z}) \qquad (16)$$

where u and v are the velocity components, f is the Coriolis parameter, $\boldsymbol{Q}$ is used to specify source and sink flows such as river and runoff inflows, ∇_H is the *nabla* operator with respect to horizontal coordinates, A_M is the horizontal mixing coefficient, and K_M is the vertical mixing coefficient. While injection of the momentum flux through wind stress occurs through surface boundary conditions, surface wave-induced momentum flux is neglected, since at present there is no coupling between the ocean and wave models. The last terms of Eqs. (15) and (16) control the vertical transfer of the surface momentum flux to the rest of the boundary layer. In order to specify K_M, a turbulence closure scheme is needed. Mellor-Yamada Level 2 (MY2) is the turbulent closure scheme currently implemented in operational versions of NCOM. In MY2, K_M is defined as

$$K_M = lqS_M \tag{17}$$

where l is the master length scale, q is the turbulent velocity scale, and S_M is an algebraic function of the Richardson number. The prognostic equation used to solve for q in the MY2 scheme is as follows

$$\frac{\partial b}{\partial t} = \frac{\partial}{\partial z}\left(lqS_q \frac{\partial b}{\partial z}\right) \tag{18}$$

where S_q is an empirical constant, and b is TKE ($q^2 = 2b; b = \frac{1}{2}\left(\overline{(u')^2} + \overline{(v')^2} + \overline{(w')^2}\right)$). The LHS of Eq. (18) represents the time rate of change of TKE, and the RHS represents vertical diffusion. MY2 assumes a strict balance between TKE generated by shear and TKE dissipation, in contrast to the Mellor-Yamada Level 2.5 turbulence closure scheme (MY2.5), where this balance is not assumed and instead terms for shear and dissipation appear in addition to what is shown in Eq. (18) (Mellor and Yamada

1982). Finally, the length scale, *l*, for MY2 is set to zero outside the surface layer. Within the surface layer, *l* is computed as

$$\boldsymbol{l = (\,(kz_t)^{-1} + (kz_b)^{-1}\,)^{-1}} \tag{19}$$

where z_t and z_b are the distances from the surface and bottom boundaries of the layer, respectively, and *k* is the von Kármán constant. The length scale, *l*, asymptotes to kz_o at the surface and bottom, where z_o is the roughness scale (Barron et al. 2006).

One possible weakness of MY2 is that it does not take into consideration the contribution breaking-surface waves have on TKE. Craig and Banner (1994) introduced an enhancement to MY2.5 that predicts the increased TKE dissipation rates observed under breaking waves (Kitaigorodskii et al. 1983, Thorpe 1984, Thorpe 1992, Osborn et al. 1992). When employed in MY2.5, this parameterization effectively reduces surface velocities (Craig and Banner 1994; Mellor and Blumberg 2004). This enhancement is not available in the MY2 scheme.

The absence of the wave-induced contribution to the momentum flux undoubtedly affects the prediction of the near-surface velocity profile (Craig and Banner 1994, Mellor and Blumberg 2004). However, comparisons between observed velocities derived from the study's two drifting buoy datasets and contemporaneous modeled near surface velocity profiles show that modeled velocities compare well with observed velocities from GLAD drifters (Fig. 3.33) and SVP drifters (Fig. 3.34). In both cases, velocity predictions from CNCOM compare better to observed in terms of mean and standard deviation than do RNCOM predictions. In addition, computation of the vertical shear $\left(\frac{du}{dz}\,,\ \frac{dv}{dz}\right)$ along drifter tracks for both RNCOM and CNCOM indicates there is little directional variability between layers in the upper 10-meters of the surface layer,

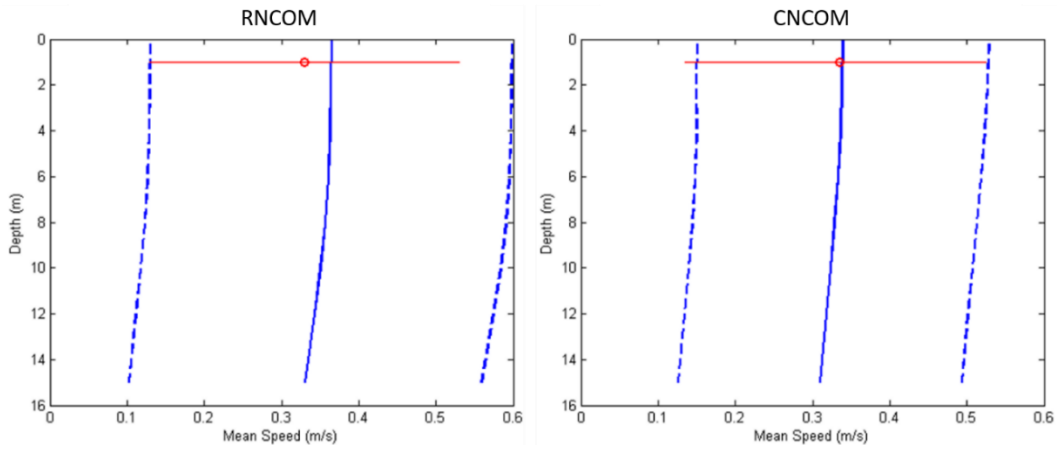

Figure 3.33 *Predicted Upper-Ocean Velocity Profile versus GLAD drifter derived velocities*

Mean observed current speed at 1-meter, derived from GLAD drifters, is shown in red, with the line indicating the standard deviation. Statistics for observed current speed were derived from the GLAD drifters using a centered-difference of position every 24 hours along the trajectories. Mean modeled current speeds are shown in blue, with the dotted lines indicating standard deviation. Modeled statistics were computed using velocity profiles taken at the same time and position as for the observed calculation. RNCOM results are shown on the left, and CNCOM results are shown on the right. 13,235 data points were available for RNCOM comparisons, and 11,628 data points were available for CNCOM comparisons.

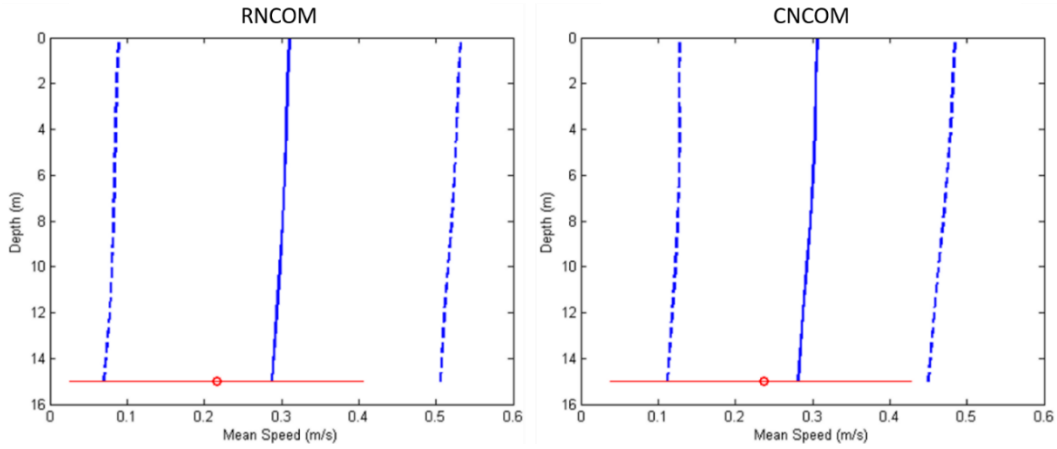

Figure 3.34 *Predicted Upper-Ocean Velocity Profile versus SVP drifter derived velocities*

Mean observed current speed at 15-meters, derived from SVP drifters, is shown in red, with the line indicating the standard deviation. Mean and standard deviation from the SVP drifters was computed by deriving a centered-difference speed from position measurements every 24 hours along SVP drifter trajectories. Mean modeled current speeds in the near-surface layer are shown in blue, with the dotted lines indicating standard deviation. Modeled statistics were computed using velocity profiles taken at the same time and position as for the observed calculation. RNCOM results are shown on the left, and CNCOM results are shown on the right. 807 data points were available for RNCOM comparisons, and 430 data points were available for CNCOM comparisons.

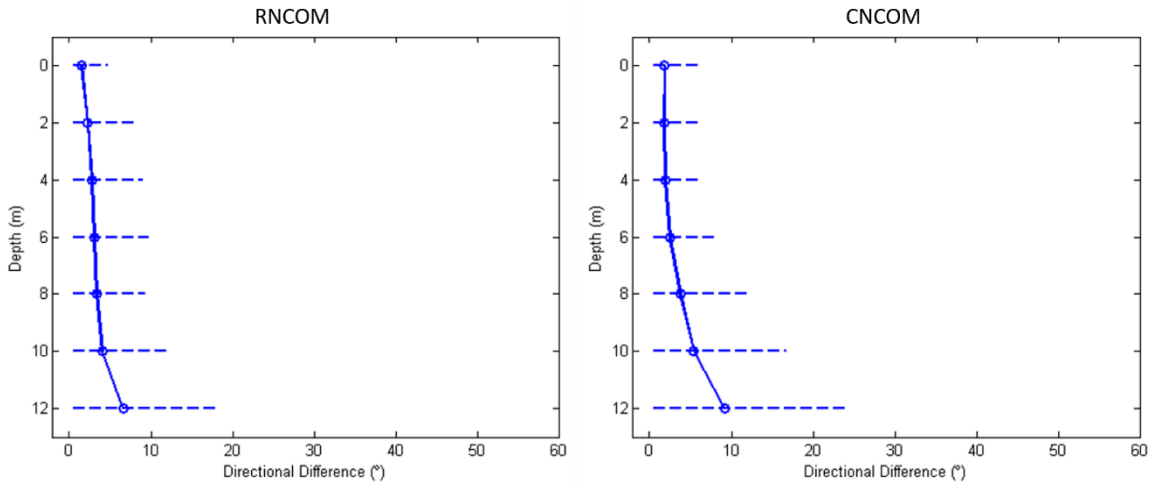

Figure 3.35 *Modeled Current Shear at GLAD Drifter Locations*

Mean difference in current direction in RNCOM (left) and CNCOM (right) between layers in the upper 15 meters of the water column; modeled current profiles are taken at the GLAD drifter's time and location every 24 hours. Dotted lines indicate the standard deviation at each depth. Directional differences are calculated as the absolute angle between modeled current directions between adjacent layers and is always ≤ 180°. 13,235 data points were available for RNCOM comparisons, and 11,628 data points were available for CNCOM comparisons.

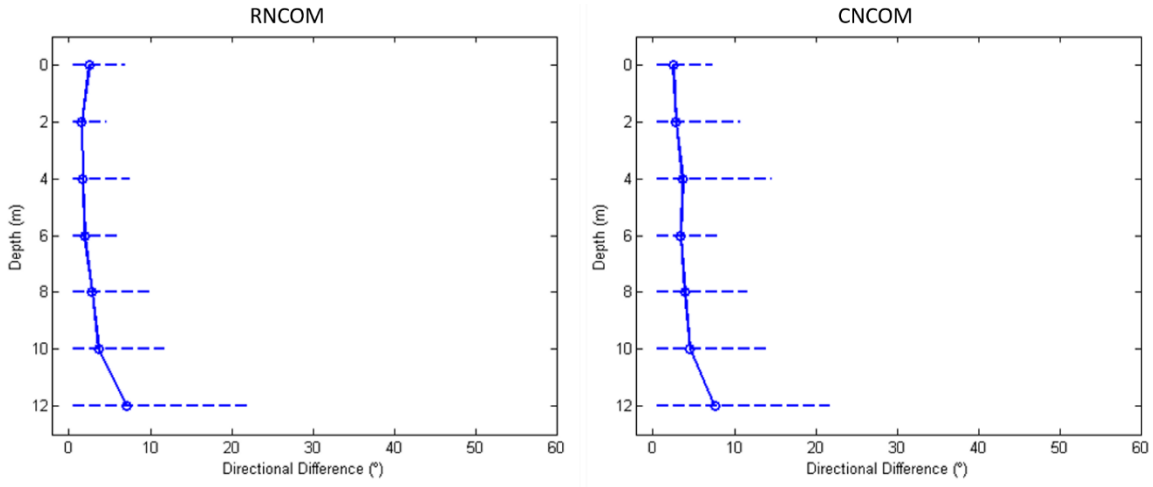

Figure 3.36 *Modeled Current Shear at SVP Drifter Locations*

Mean difference in current direction in RNCOM (left) and CNCOM (right) between layers in the upper 15 meters of the water column; modeled current profiles are taken at the SVP drifter's time and location every 24 hours. Dotted lines indicate the standard deviation at each depth. Directional differences are calculated as the absolute angle between modeled current directions between adjacent layers and is always ≤ 180°. 807 data points were available for RNCOM comparisons, and 430 data points were available for CNCOM comparisons.

whereas below 10-meters, the directional variability is much increased (Figs. 3.35 and Fig. 3.36). Therefore, model velocities taken at a single depth are likely adequate for drift prediction that influenced by only very shallow currents, as is the case for GLAD drifters. However, objects that sit lower in the water column may be better predicted by a depth-averaged velocity due to the inherent current variability that exists below 10-meters. SVP drifters, while drogued at 15-meters, are still approximated using a single depth level; further analysis is needed to show if a depth-averaged velocity would give a more skillful prediction for this case.

3.5 Summary of Validation Findings

The model validation analyses in this section characterize typical errors seen in model current predictions and identify their most common sources. In the GOM's deep-water environment, current prediction errors are most often due to errors in the placement of mesoscale features and these errors will have a large impact on drift predictive skill. For the shallow-water environment, it is seen that the model's major weakness in predicting the current velocity field lies in the model's inability to accurately forecast the phasing of NIWs that dominate the current signal, and this weakness is possibly linked to errors in the model's mixed layer depth prediction. However, since NIWs are oscillatory in nature, it is thought that the model's weakness in accurately predicting them will not adversely affect drift predictive skill, especially at forecast times longer than the inertial period (> 24-hours). Additionally, it is shown that very little directional variability exists in adjacent near-surface depth levels. This result ensures that the choice of the near-surface depth level used to compute drift trajectories will not significantly change skill metrics.

CHAPTER IV - RESULTS

As was introduced in Section 2.2.1, the validation statistics presented in this chapter are derived from probability map comparisons computed in the following way. An initial 20 km disc of modeled particles is centered along each drifting buoy track at daily 00z positions. Particles are launched every three hours and integrated out 96-hours using the first 24-hours of forecast from the previous three days of model runs, along with the entire model forecast from the analysis day (Fig. 2.6). Particle counts at every 24-hours of the forecast are binned for probability map creation. The position of the drifting buoy at the forecast time is then compared to the area defined by the drift probability map. An in-cloud metric is calculated by determining whether the observed drifter position at the forecast time falls within the model predicted cloud. Additionally, an error distance is determined by using a center-of-mass calculation, where the distance between the observed drifter's position and the center of every bin in the model predicted cloud is determined and the quantile (i.e. lowest to highest probable area definition) value of each bin is used as a weighting matrix. The cloud size for each probability map is also calculated since it is an important operational consideration, as large search area recommendations lead to longer search and recovery times and can dictate mission success or failure.

4.1 Probability Map Optimization Analysis and Construction Details

Sensitivity of the accuracy of probability maps with respect to the initial particle array size as well as the particle count bin size was investigated to optimize skill statistics. From Table 4.1, it is seen that larger discs of initial particle positions released leads to enhanced performance where the observed drifter falls within the modeled cloud

a higher percentage of the time. This outcome is likely due to the fact that the larger initial disc of particles samples more of the spatial variability associated with the current field, for instance, better capturing an eddy that may be misplaced by a few kilometers in the model solution.

Table 4.1 *Initial Condition Disc Optimization for Probability Maps*

In-Cloud Percentage Statistics for GLAD Dataset: Pure Particle Drift				
	24hr Forecast	**48hr Forecast**	**72hr Forecast**	**96hr Forecast**
IC Disc Diameter	**4km Bin**	**4km Bin**	**4km Bin**	**4km Bin**
10km Disc	47 %	24 %	17 %	13 %
15km Disc	62 %	34 %	24 %	18 %
20km Disc	75 %	43 %	31 %	25 %

Note: Probability maps for each initial condition disc size were computed using a 4km bin to determine final particle counts. Only pure particle drift results were used in initial condition disc size optimization analysis.

Table 4.2 *Bin Size Optimization for Probability Maps*

In-Cloud Percentage Statistics for GLAD Dataset: Pure Particle Drift				
	24hr Forecast	**48hr Forecast**	**72hr Forecast**	**96hr Forecast**
IC Disc Diameter	**4km Bin**	**5km Bin**	**6km Bin**	**7km Bin**
20km Disc	75 %	46 %	35 %	31 %

Note: Probability maps computed using a variable bin size with respect to forecast hour to determine final particle counts. Only pure particle drift results were used to in bin size optimization analysis.

Because small cell size, combined with the stretching and deformation of the Lagrangian volume expected at long forecast times, puts high and low probability cells in close proximity to one another, increasing the particle count bin size in tandem with forecast time was investigated in an effort to increase the robustness of the in-cloud metric. Furthermore, the patchiness and filaments present in model's predicted particle distribution at long time moments may or may not be a good representation of the true particle distribution and instead an artifact of the finite number of particles released to compute the probability maps. Table 4.2 shows that increasing the bin size for particle counts as the forecast time increases improves in-cloud percentages at longer drift

integration times. Therefore, an initial array size of 20 km as well as an initial 4 km bin size, which increases by 1 km per forecast day, is used in the creation of all subsequent probability maps underlying the presented statistical metrics. An example of the validation comparisons done using the above-described criteria for all drift algorithm parameterizations assessed in this study is shown in Fig. 4.1.

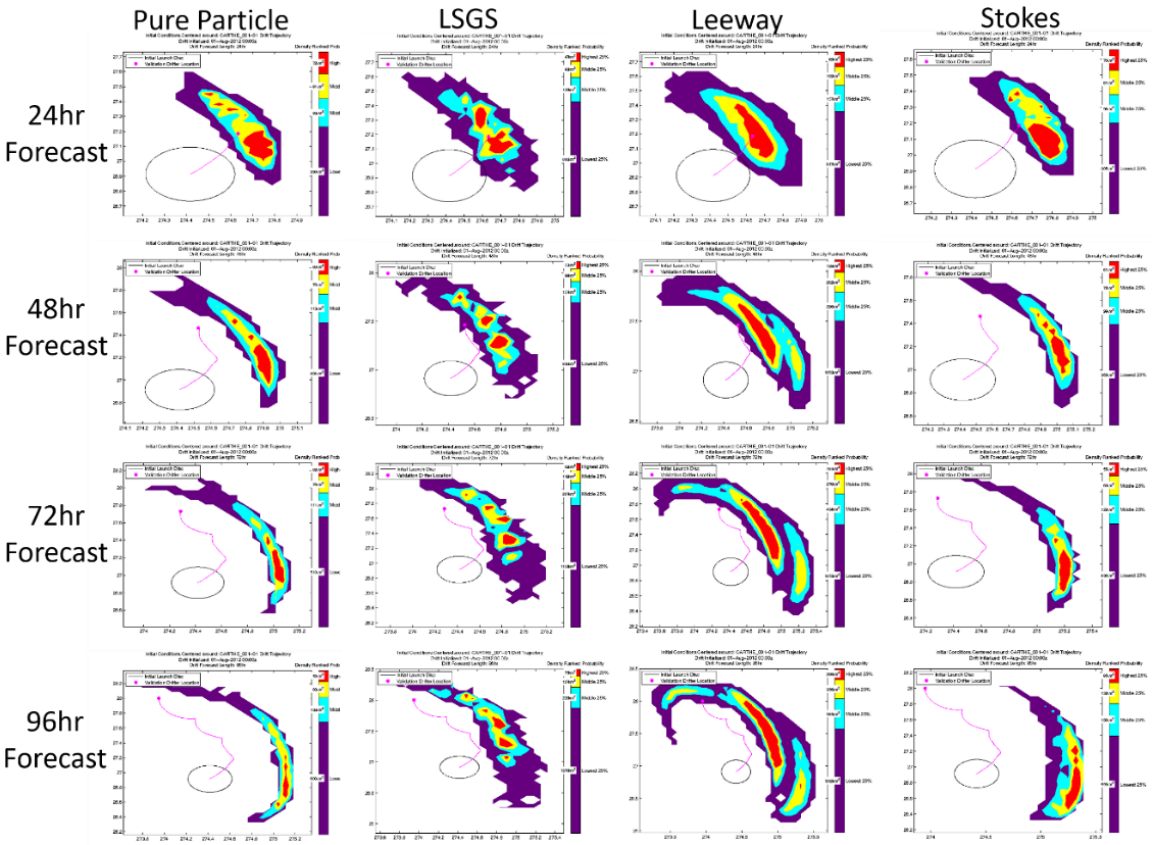

Figure 4.1 *Validation Example for Drift Algorithms*

Initial launch locations for the modeled drift prediction are denoted by the black circle and are centered on a position along a GLAD drifting buoy track on August 01, 2012 (pink line). Columns show drift predictions produced for the same initial conditions using each of the different drift algorithms tested in this study; from left to right: drift predictions using a pure particle algorithm, drift predictions including the LSGS parameterization, drift predictions including the Leeway parameterization, and drift predictions including the Stokes parameterization. Rows show the predictions at different forecast times; from top to bottom: 24-hour forecast prediction, 48-hour forecast prediction, 72-hour forecast prediction, and 96-hour forecast prediction. Drift predictions are produced from RNCOM.

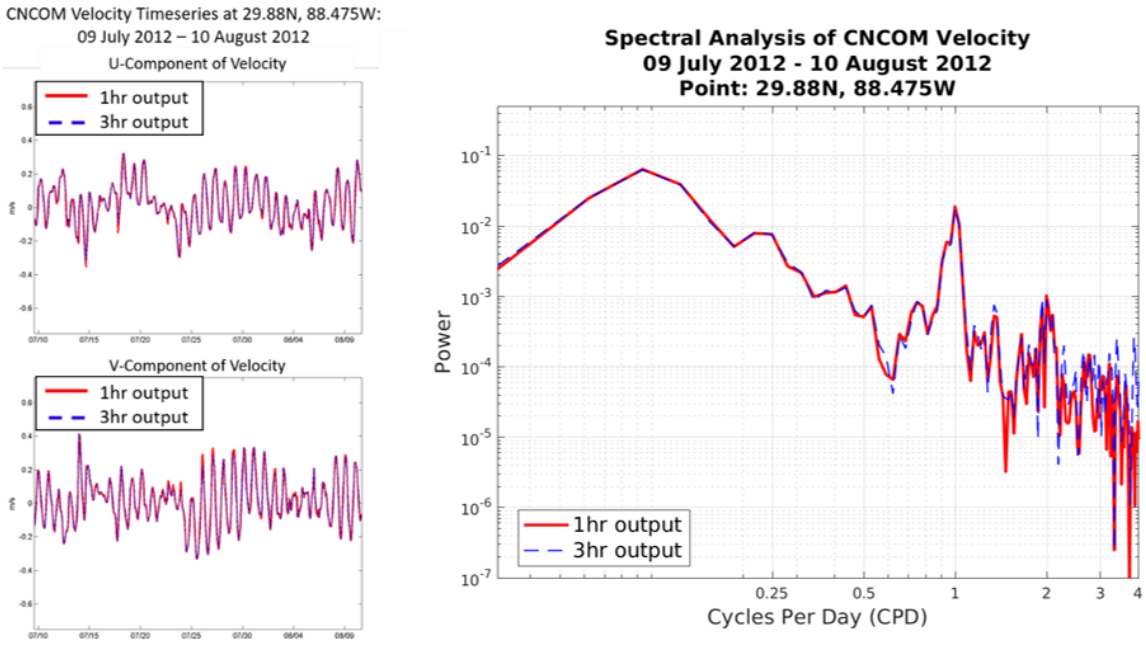

Figure 4.2 *CNCOM Output Frequency Velocity Comparison*

Right-hand plots show velocity time series of CNCOM model predictions, u-component shown on top and v-component shown on bottom, taken at HF Radar Station 1 (Fig. 3.24) inside the high-resolution model domain over the month of August, 2012. Left-hand plot shows the frequency spectra for both output datasets (1-hr and 3-hr) at the same HF Radar Station. In all cases the red solid line shows CNCOM velocity at a one-hour output interval, and the blue dashed line shows CNCOM velocity at a three-hour output interval.

Additionally, the necessity of using hourly output from the CNCOM model domain was considered as the increased temporal resolution combined with the increased spatial resolution of model output creates memory issues during drift integration calculations. CNCOM was sub-sampled to three-hourly output intervals and analysis on the velocity fields from the two datasets, hourly output and three-hourly output, was performed to determine the effect of decreasing the temporal frequency. Fig. 4.2 shows time series and frequency spectra for both datasets at the one of the HF Radar station locations analyzed in Fig 3.24. It is seen that sub-sampling the higher-resolution model's temporal output frequency yields almost identical time series and spectral characteristics.

From these results, it can be concluded that sub-sampling does not significantly alter the velocity structure resolved by CNCOM; therefore, three-hourly output is used to construct probability maps for the high-resolution model validation.

4.2 Validation Statistics for Whole Domain Datasets

The validation results presented in this section are computed using the entire GLAD (Fig. 2.4) and SVP (Fig. 2.5) datasets as ground truth. Metrics are presented first for the GLAD dataset and then for the SVP dataset. Each drifter dataset is compared to both unfiltered and filtered RNCOM derived probability maps, respectively, and each of the parameterization schemes discussed in Chapter 2 is considered. The three statistical metrics outlined previously, in-cloud percentage, error distance, and cloud size, are used throughout the remainder of this Chapter to assess the skill of the ocean models and the performance of each drift algorithm.

4.2.1 GLAD Dataset Validation Statistics

4.2.1.1 Unfiltered Results

In-cloud metric results are shown in Fig. 4.3 with histograms of the percentage of time a GLAD drifter falls within the area defined by drift probability maps for all forecast lengths (i.e., 24-hour, 48-hour, 72-hour, and 96-hour forecast times) and all drift algorithms considered in this study. For these in-cloud histograms, a score of zero indicates that the observed drifter falls outside of the model cloud of particles and a positive score indicates the observed drifter falls within the modeled cloud. Increasing positive values indicate that the observed drifter falls in areas of increasing predicted probability, with 1 indicating the lowest probable area and 4 indicating the highest probable area.

Defining skill for the in-cloud metric as scoring a positive value more than 50% of the time, it can be seen that skillful predictions do not extend to 48-hours when using the pure particle or Stokes drift algorithms (Fig. 4.3 and Table 4.3). Additionally, the Stokes algorithm provides no advantage over a pure particle drift in terms of in-cloud predictability, with no statistically significant differences seen between the performance of the two algorithms at any forecast length. Drift predictions including the LSGS and Leeway parameterizations show comparable performance, with both showing significantly improved skill at all forecast times. LSGS and Leeway produce a drift prediction that is ~15% more accurate with respect to in-cloud predictability at the 24-hour forecast over the pure particle case and ~30% more accurate at the 96-hour forecast over the pure particle case. LSGS and Leeway also extend skillfulness for in-cloud predictability out to the 96-hour forecast, with a 60% in-cloud percentage seen for LSGS and a 58% in-cloud percentage seen for Leeway (Table 4.3). However, examining the histogram shapes further (Fig. 4.3), it is seen that most of the in-cloud gain associated with the LSGS and the Leeway algorithm at forecast times beyond 24-hours occurs in the lowest probable area.

Mean distance errors show that a pure particle algorithm produces a result that gives the lowest distance errors for short-term 24-hour forecasts; however, all four algorithms produce similar errors at this forecast time showing a spread of less than 3 km (Fig. 4.4 and Table 4.4). For the 48-hour forecast and beyond, LSGS is the only algorithm that produces a more skillful prediction in terms of decreased distance errors with respect to a pure particle result. Both Stokes and Leeway produce higher distance

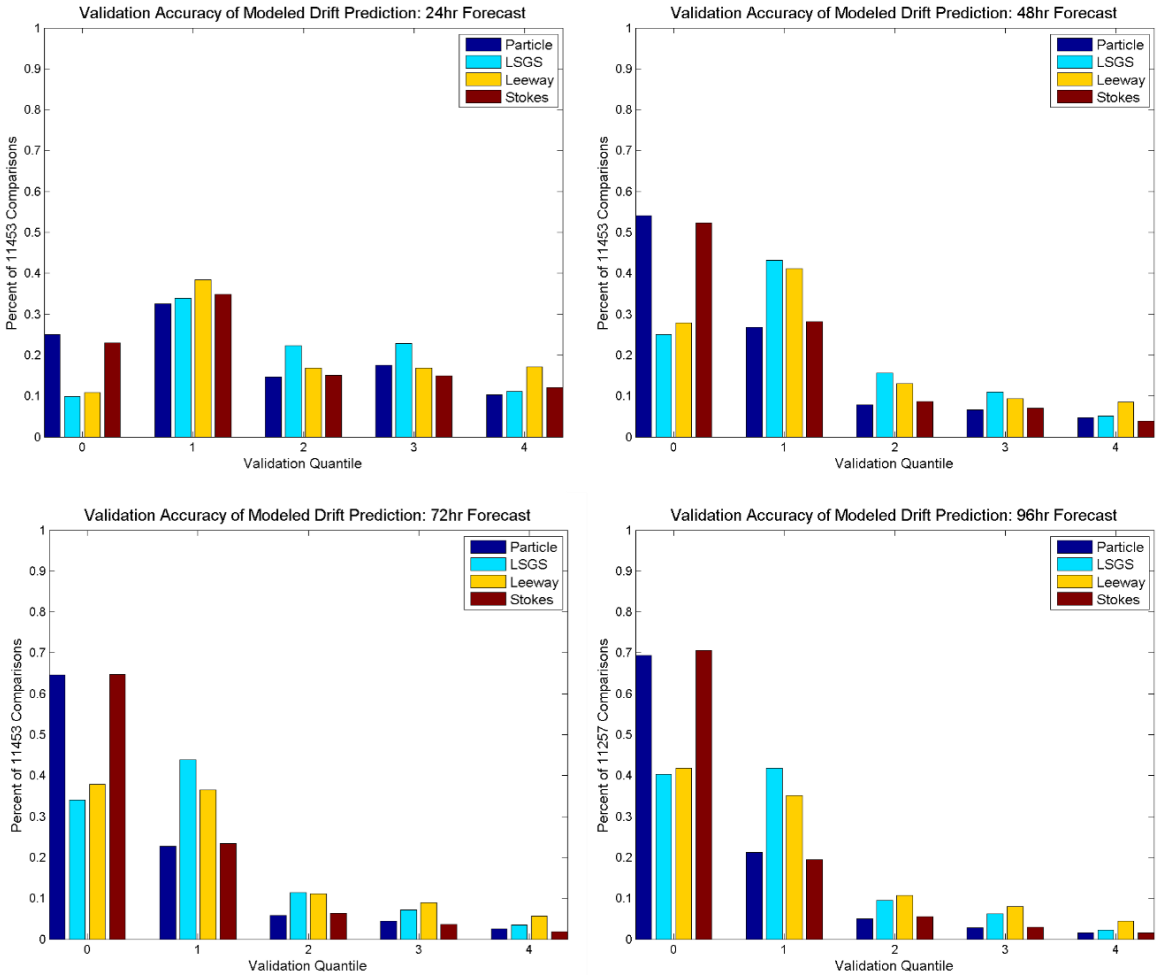

Figure 4.3 *GLAD Validation Accuracy Statistics for Whole Domain Dataset*

Histograms showing the percentage of time observed drifters falls within the area predicted by the drift probability map at the 24 hr (top left), 48 hr (top right), 72 hr (bottom left), and 96 hr (bottom right) forecast times. Pure particle distributions are shown in dark blue, Lagrangian Sub-Grid Scale (LSGS) distributions are shown in light blue, Leeway distributions are shown in yellow, and Stokes distributions are shown in red.

Table 4.3 *GLAD Validation In-Cloud Percentage Statistics: Whole Domain Results*

Drift Algorithm	**24hr Forecast**	**48hr Forecast**	**72hr Forecast**	**96hr Forecast**
Pure Particle	75 % (74 – 76)	46 % (45 – 47)	35 % (35 – 36)	31 % (30 – 31)
LSGS	90 % (90 – 91)	75 % (74 – 76)	66 % (65 – 67)	60 % (59 – 61)
Leeway	89 % (89 – 90)	72 % (71 – 73)	62 % (61 – 63)	58 % (57 – 59)
Stokes	77 % (76 – 78)	48 % (47 – 49)	35 % (34 – 36)	29 % (29 – 30)

Note: Statistics were derived from >10,000 validation comparisons. Confidence intervals are computed using 10,000 bootstrapped samples, with 11,453 degrees of freedom and are significant at a 95% confidence level.

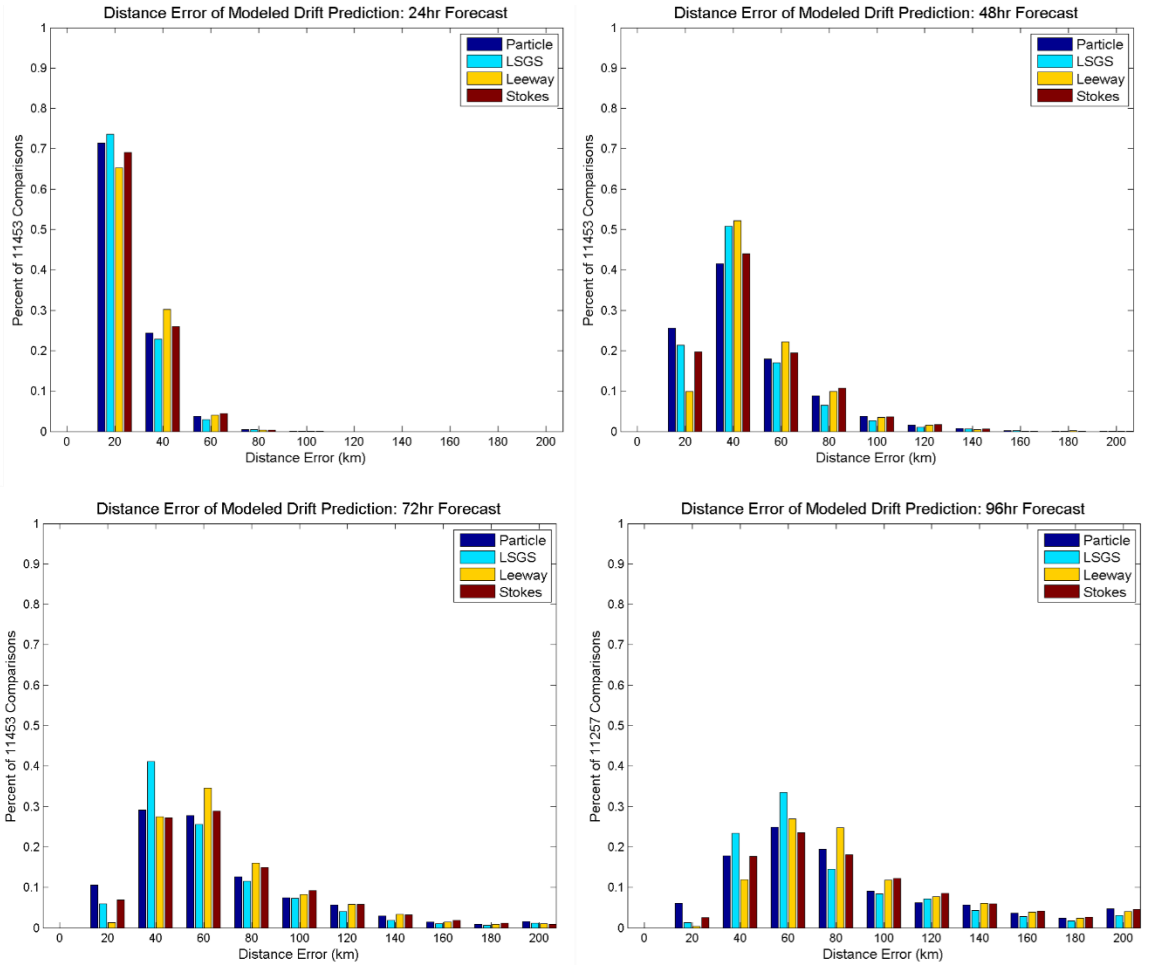

Figure 4.4 *GLAD Error Distance Statistics for Whole Domain Dataset*

Histograms showing the error distance between observed drifter positions and the center-of-mass of the predicted cloud for probability maps at 24 hr (top left), 48 hr (top right), 72 hr (bottom left), and 96 hr (bottom right) forecast times. Pure particle distributions are shown in dark blue, Lagrangian Sub-Grid Scale (LSGS) distributions are shown in light blue, Leeway distributions are shown in yellow, and Stokes distributions are shown in red.

Table 4.4 *GLAD Validation Mean Error Distance Statistics: Whole Domain Results*

Drift Algorithm	**24hr Forecast**	**48hr Forecast**	**72hr Forecast**	**96hr Forecast**
Pure Particle	27.1 km (26.9 – 27.3)	47.2 km (46.7 – 47.6)	66.6 km (65.9 – 67.3)	85.3 km (84.4 – 86.2)
LSGS	28.0 km (27.8 – 28.1)	45.6 km (45.2 – 46.0)	62.7 km (62.1 – 63.3)	79.3 km (78.5 – 80.2)
Leeway	29.8 km (29.6 – 30.0)	51.0 km (50.6 – 51.4)	71.3 km (70.7 – 71.9)	91.1 km (90.2 – 91.9)
Stokes	28.6 km (28.4 – 28.7)	49.5 km (49.1 – 49.9)	69.7 km (69.0 – 70.4)	89.6 km (88.7 – 90.6)

Note: Statistics were derived from >10,000 validation comparisons. Confidence intervals are computed using 10,000 bootstrapped samples, with 11,453 degrees of freedom and are significant at a 95% confidence level.

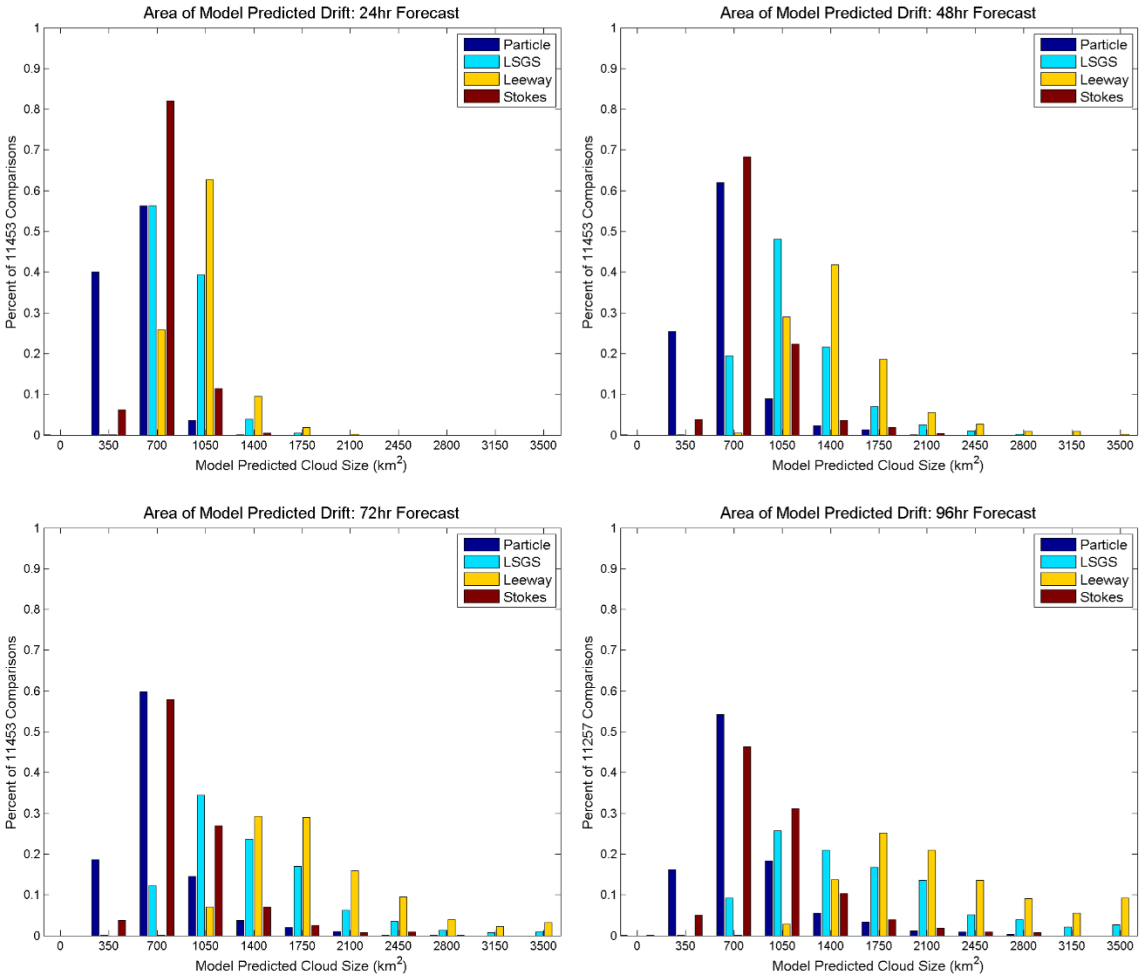

Figure 4.5 *GLAD Predicted Cloud Size Statistics for Whole Domain Dataset*

Histograms showing the predicted cloud size for probability maps at 24 hr (top left), 48 hr (top right), 72 hr (bottom left), and 96 hr (bottom right) forecast times. Pure particle distributions are shown in dark blue, Lagrangian Sub-Grid Scale (LSGS) distributions are shown in light blue, Leeway distributions are shown in yellow, and Stokes distributions are shown in red.

Table 4.5 *GLAD Validation Mean Cloud Size Statistics: Whole Domain Results*

Drift Algorithm	24hr Forecast	48hr Forecast	72hr Forecast	96hr Forecast
Pure Particle	575 km^2 (572 – 577)	671 km^2 (667 – 675)	750 km^2 (745 – 756)	829 km^2 (822 – 836)
LSGS	885 km^2 (882 – 888)	1149 km^2 (1143 – 1155)	1385 km^2 (1375 – 1395)	1608 km^2 (1595 – 1621)
Leeway	1004 km^2 (1001 – 1008)	1464 km^2 (1457 – 1472)	1868 km^2 (1857 – 1879)	2247 km^2 (2232 – 2262)
Stokes	687 km^2 (685 – 690)	808 km^2 (803 – 812)	881 km^2 (875 – 887)	956 km^2 (949 – 963)

Note: Statistics were derived from >10,000 validation comparisons. Confidence intervals are computed using 10,000 bootstrapped samples, with 11,453 degrees of freedom and are significant at a 95% confidence level.

errors than the pure particle case, with the greatest growth in error seen at the 96-hour forecast. The dichotomy of performance between LSGS and Leeway for the distance error metric is quite surprising, especially considering Leeway performs very comparably to LSGS for the in-cloud metric.

Cloud size metrics show growth in cloud size for all drift algorithms as forecast length increases, as expected (Fig. 4.5 and Table 4.5). The substantial increases in cloud size, particularly at longer forecast lengths, using both the LSGS algorithm and the Leeway algorithm are noted as an important area for consideration. Very large cloud size can significantly impact the utility of operational recommendations, as search operations over large areas becomes time-consuming and man-power intensive. The LSGS result produces a drift area about 55% larger than the pure particle case at the 24-hour forecast, increasing to a 95% larger drift area at the 96-hour forecast. Furthermore, the Leeway algorithm, as employed in this study, produces a cloud size about 13% larger than LSGS and about 75% larger than pure particle at the 24-hour forecast; with these percentages ballooning to 40% larger than LSGS and 170% larger than pure particle at the 96-hour forecast. Even with the significant improvements for in-cloud predictability, significantly larger cloud sizes combined with higher distance errors make the Leeway algorithm a less desirable option than LSGS at least for supporting SAR type operations. The area increases seen for LSGS may be an acceptable trade-off at longer forecast lengths given the greater skill for both the in-cloud and distance error metrics.

4.2.1.2 Filtered Results

The same metrics as described in 4.2.1.1 are used to assess the performance of drift predictions using filtered RNCOM model output. Looking at Table 4.6, it can be

seen that for almost all drift algorithms and at all forecast lengths, filtering of the RNCOM current field produces drift predictions that are more skillful for all three of the metrics utilized in this study. Increases of 6-10% due to filtering are seen in the in-cloud percentage statistics in the first 72-hours for the pure particle algorithm, with only marginal increases seen at the 96-hour forecast. Increases of 2-5% are seen in the in-cloud percentage statistics in the first 72-hours for both Leeway and Stokes parameterizations. Filtering also extends skillfulness (> 50% in-cloud) out to 48-hours for both the pure particle and Stokes algorithms. LSGS is the only algorithm that shows degraded performance for in-cloud predictability when using filtered velocity fields, with decreases ranging from 3% at the 48-hour forecast to 9% at the 96-hour forecast. Distance errors decrease by 11-22% for all drift algorithms (Table 4.7), with more modest decreases seen at shorter forecast lengths and greater decreases seen at longer forecast lengths. Additionally, filtering produces the added benefit of decreasing the cloud size of drift probability maps on average by a substantial amount, 14-49% (Table 4.7). Again, more modest cloud size decreases are seen at shorter forecast lengths and greatest cloud size decreases are seen at longer forecast lengths. Cloud size reduction is consequence of the filtering out of small-scale, sub-mesoscale features known to play an important role in increasing the dispersive nature of the current field (Poje et al. 2010; Poje et al. 2014). Because small cloud size is a desirable characteristic for operational search-and-rescue/search-and-recovery applications, reduction in cloud size combined with more skillful predictions for in-cloud performance as well as significant distance error decreases highlights the great utility of filtering for drift applications. While the degradation seen in the in-cloud performance for the LSGS algorithm in the filtered case

is likely a direct result of cloud size reduction, the benefit gained from smaller search-area recommendations along with decreased distance errors makes filtering advantageous even for LSGS. The addition of filtering makes the Leeway algorithm a more operationally palatable option, reducing significantly large cloud sizes to levels on par with those predicted by LSGS in the unfiltered case. However, distance errors and in-cloud performance for the filtered case of Leeway still does not outperform LSGS using unfiltered velocity fields.

Table 4.6 *GLAD Validation Metrics for Unfiltered versus Filtered RNCOM*

Forecast	**In-Cloud Percentage**		**Mean Error Distance**		**Mean Cloud Size**	
	Unfiltered	Filtered	Unfiltered	Filtered	Unfiltered	Filtered
Pure Particle						
24hr	75 % (74 – 76)	81 % (81 – 82)	27.1 km (26.9 – 27.3)	22.6 km (22.4 – 22.8)	575 km^2 (572 – 577)	447 km^2 (446 – 448)
48hr	46 % (45 – 47)	56 % (55 – 57)	47.2 km (46.7 – 47.6)	37.2 km (36.8 – 37.6)	671 km^2 (667 – 675)	440 km^2 (438 - 442)
72hr	35 % (35 – 36)	41 % (40 - 42)	66.6 km (65.9 – 67.3)	51.9 km (51.3 – 52.6)	750 km^2 (745 – 756)	430 km^2 (423 – 432)
96hr	31 % (30 – 31)	33 % (32 -34)	85.3 km (84.4 – 86.2)	66.6 km (65.7 – 67.5)	829 km^2 (822 – 836)	426 km^2 (423 – 428)
LSGS						
24hr	90 % (90 – 91)	90 % (91 – 92)	28.0 km (27.8 – 28.1)	25.0 km (24.8 – 25.1)	885 km^2 (882 – 888)	738 km^2 (736 – 739)
48hr	75 % (74 – 76)	72 % (71 – 73)	45.6 km (45.2 – 46.0)	39.5 km (39.1 – 39.9)	1149 km^2 (1143 – 1155)	832 km^2 (828 – 835)
72hr	66 % (65 – 67)	59 % (59 – 60)	62.7 km (62.1 – 63.3)	53.7 km (53.1 – 54.2)	1385 km^2 (1375 – 1395)	903 km^2 (898 – 908)
96hr	60 % (59 – 61)	51 % (50 – 52)	79.3 km (78.5 – 80.2)	67.7 km (67.0 – 68.5)	1608 km^2 (1595 – 1621)	973 km^2 (966 – 980)
Leeway						
24hr	89 % (89 – 90)	91 % (91 – 92)	29.8 km (29.6 – 30.0)	25.9 km (25.7 – 26.0)	1004 km^2 (1000 – 1008)	867 km^2 (865 – 870)
48hr	72 % (71 – 72)	76 % (75 – 77)	51.0 km (50.6 – 51.4)	43.0 km (42.7 – 43.4)	1464 km^2 (1457 – 1472)	1161 km^2 (1157 – 1165)
72hr	62 % (61 – 63)	66 % (65 – 67)	71.3 km (70.7 – 71.9)	59.9 km (59.3 – 60.4)	1868 km^2 (1857 – 1879)	1379 km^2 (1373 – 1386)
96hf	58 % (57 – 59)	59 % (58 – 60)	91.1 km (90.2 – 91.9)	76.5 km (75.7 – 77.3)	2247 km^2 (2232 – 2263)	1562 km^2 (1554 – 1570)
Stokes						
24hr	77 % (76 – 78)	82 % (82 – 83)	28.6 km (28.4 – 28.7)	24.7 km (24.5 – 24.9)	687 km^2 (685 – 690)	591 km^2 (589 – 594)
48hr	48 % (47 – 49)	54 % (53 – 55)	49.5 km (49.1 – 49.9)	41.5 km (41.1 – 41.9)	808 km^2 (803 – 812)	620 km^2 (617 – 622)
72hr	35 % (34 – 36)	38 % (38 – 39)	69.7 km (69.0 – 70.4)	57.8 km (57.2 – 58.4)	881 km^2 (875 – 887)	607 km^2 (604 – 611)
96hr	29 % (29 – 30)	30 % (29 – 31)	89.6 km (88.7 – 90.6)	73.8 km (72.9 – 74.6)	956 km^2 (949 – 963)	601 km^2 (598 – 605)

Note: Statistics were derived from >10,000 validation comparisons. Confidence intervals are computed using 10,000 bootstrapped samples, with 11,453 degrees of freedom and are significant at a 95% confidence level.

Table 4.7 *GLAD Validation Metrics Improvements seen Using Filtered RNCOM*

Forecast	In-Cloud Percentage Increase	Mean Error Distance Decrease	Mean Cloud Size Decrease
Pure Particle			
24hr	6 %	17 %	22 %
48hr	10 %	21 %	34 %
72hr	6 %	22 %	43 %
96hr	2 %	22 %	49 %
LSGS			
24hr	No Statistical Difference	11 %	17 %
48hr	-3 %	13 %	28 %
72hr	-7 %	14 %	35 %
96hr	-9 %	15 %	39 %
Leeway			
24hr	2 %	13 %	14 %
48hr	4 %	16 %	21 %
72hr	4 %	16 %	26 %
96hr	No Statistical Difference	16 %	30 %
Stokes			
24hr	5 %	14 %	14 %
48hr	6 %	16 %	23 %
72hr	3 %	17 %	31 %
96hr	No Statistical Difference	18 %	37 %

Note: Statistics were derived from >10,000 validation comparisons.

4.2.2 SVP Dataset Validation Statistics

Before results are presented for the SVP dataset comparisons, it must be noted that while the GLAD statistical metrics are derived using over 400 drifting buoy tracks yielding >10,000 validation comparisons, SVP statistical metrics are derived using a much smaller dataset. Only 19 SVP drifting buoys are available for comparison, yielding around ~700 validation comparisons. As a consequence of a much smaller sample size, some caution must be applied to the interpretation of statistics derived from the SVP comparisons as they are necessarily less robust than those achieved using the GLAD dataset. However, with that said, many of the same trends and tendencies noted in the GLAD results are also seen in the SVP results.

4.2.2.1 Unfiltered Results

Overall, higher in-cloud percentages are seen from the SVP comparisons than in the GLAD comparisons, with significantly better performance (7-9% more skillful) seen

for pure particle drift at all forecast times. Greater than 50% in-cloud performance is also extended to 48-hours for the pure particle algorithm vice the 24-hour limit seen in the GLAD results. Stokes drift shows no statistical difference in performance compared to pure particle in the first 48-hours of the prediction and shows degraded performance as compared to pure particle at the 72-hour and 96-hour forecast lengths. As was found in the GLAD analyses, the Stokes drift parameterization provides no advantage over a pure particle drift for SVP drifter trajectory predictions. Also, similarly to the GLAD results, the LSGS and Leeway algorithms remain the most skillful for in-cloud percentage, outperforming the pure particle and Stokes by 8-14% at the 24-hour forecast and by 20% or greater at longer forecast times. LSGS and Leeway also continue to be the only algorithms that extend skill for in-cloud prediction out to 96-hours (Table 4.8). Just as seen in the GLAD results, the histogram shape for LSGS and Leeway show that most of the in-cloud gain at forecast times beyond 24-hours occurs in the lowest probable area (Fig. 4.6).

Because of the smaller sample size, confidence intervals for distance errors are much wider for the SVP comparisons than seen in the GLAD comparisons. This results in a less conclusive performance metric, as LSGS does not show any statistically significant decreases in distance error over the pure particle case. Both Leeway and Stokes show degraded performance as compared to pure particle, with increased distance errors at all forecast lengths. It is worth noting that the Stokes algorithm produces the highest distance errors, and this poor performance of the Stokes algorithm for SVP comparisons is not a surprising result, as it is perhaps the most ill-suited of the parameterizations, particularly for SVP drifters. Since SVP drifters are drogued at 15 m,

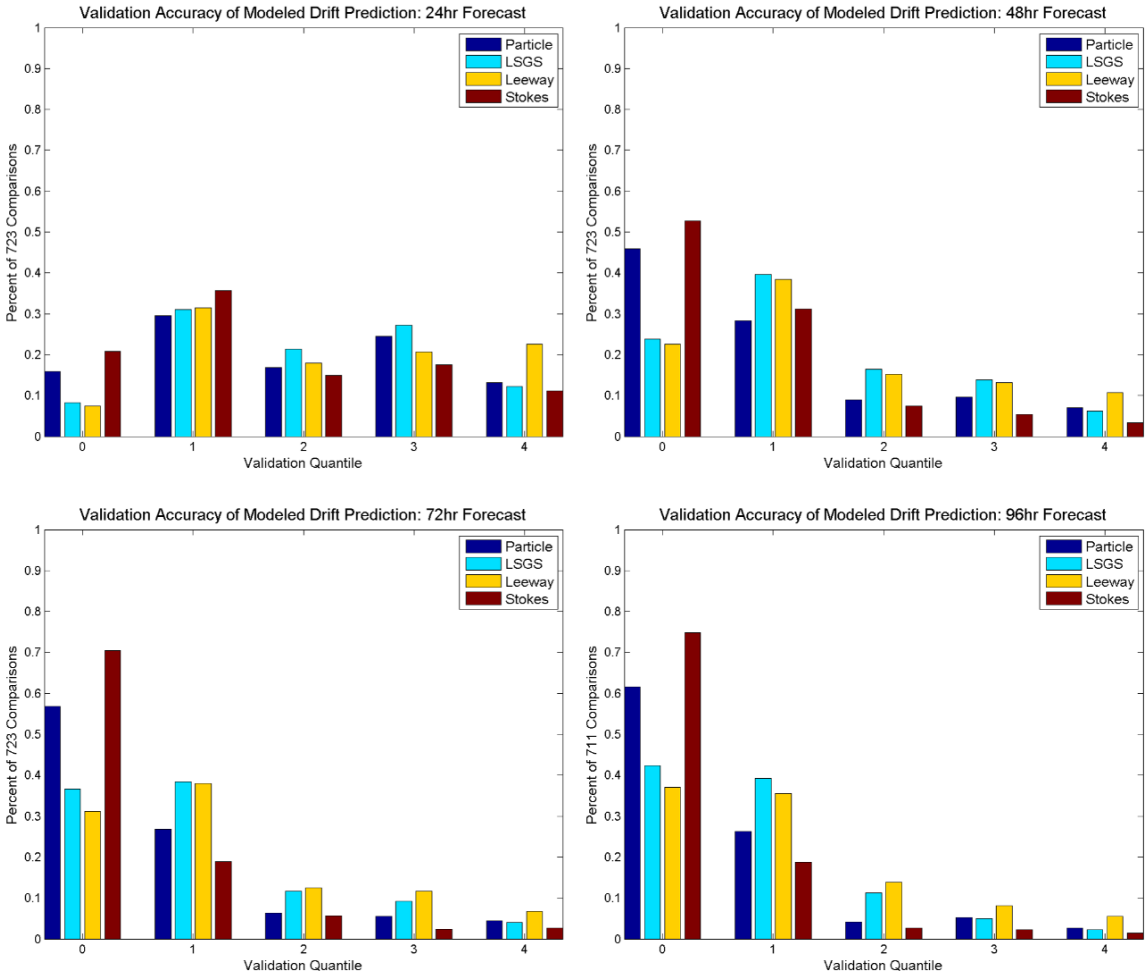

Figure 4.6 *SVP In-Cloud Percentage Statistics for Whole Domain Dataset*

Histograms showing the percentage of time observed drifters fell within the area predicted by the drift probability map at the 24 hr (top left), 48 hr (top right), 72 hr (bottom left), and 96 hr (bottom right) forecast times. Pure particle distributions are shown in dark blue, Lagrangian Sub-Grid Scale (LSGS) distributions are shown in light blue, Leeway distributions are shown in yellow, and Stokes distributions are shown in red.

Table 4.8 *SVP Validation In-Cloud Percentage Statistics: Whole Domain Results*

Drift Algorithm	**24hr Forecast**	**48hr Forecast**	**72hr Forecast**	**96hr Forecast**
Pure Particle	84 % (81 – 87)	54 % (50 – 58)	43 % (40 – 47)	38 % (35 – 42)
LSGS	92 % (90 – 94)	76 % (73 – 79)	63 % (60 – 67)	58 % (54 – 61)
Leeway	93 % (91 – 94)	77 % (74 – 80)	69 % (66 – 72)	63 % (59 – 66)
Stokes	79 % (76 – 82)	47 % (44 – 51)	30 % (26 – 33)	25 % (22 – 28)

Note: Statistics were derived from ~700 validation comparisons. Confidence intervals are computed using 10,000 bootstrapped samples, with 723 degrees of freedom and are significant at a 95% confidence level.

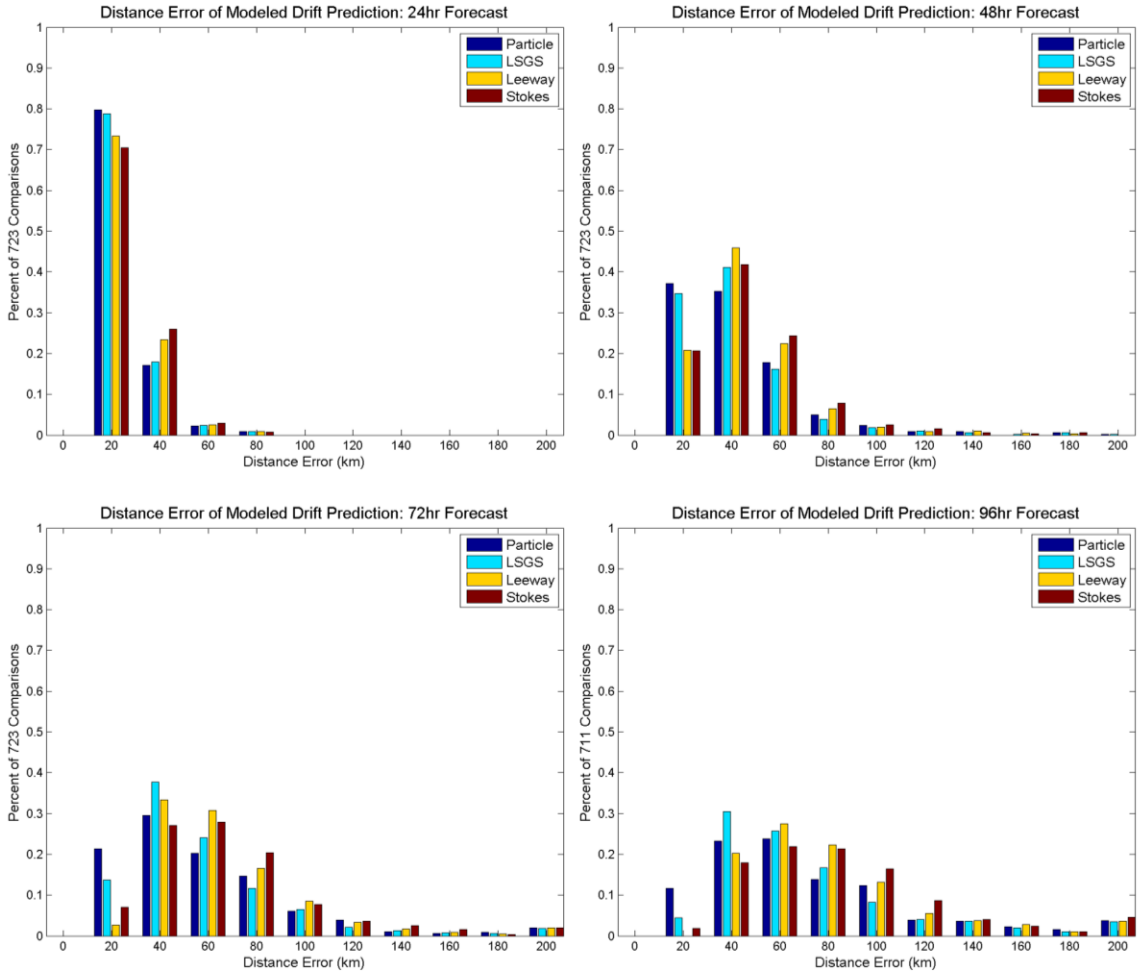

Figure 4.7 *SVP Error Distance Statistics for Whole Domain Dataset*

Histograms showing the error distance between observed drifter positions and the center-of-mass of the predicted cloud for probability maps at 24 hr (top left), 48 hr (top right), 72 hr (bottom left), and 96 hr (bottom right) forecast times. Pure particle distributions are shown in dark blue, Lagrangian Sub-Grid Scale (LSGS) distributions are shown in light blue, Leeway distributions are shown in yellow, and Stokes distributions are shown in red.

Table 4.9 *SVP Validation Mean Error Distance Statistics: Whole Domain Results*

Drift Algorithm	24hr Forecast	48hr Forecast	72hr Forecast	96hr Forecast
Pure Particle	24.5 km (23.7 – 25.3)	42.1 km (40.4 – 44.0)	59.8 km (57.0 – 62.8)	75.7 km (72.0 – 79.6)
LSGS	26.1 km (25.4 – 26.8)	42.7 km (41.1 – 44.4)	59.3 km (56.7 – 62.0)	74.3 km (71.0 – 77.9)
Leeway	27.9 km (27.3 – 28.7)	47.5 km (45.9 – 49.2)	66.5 km (64.0 – 69.1)	83.6 km (80.3 – 87.0)
Stokes	27.9 km (27.1 – 28.6)	48.8 km (47.1 – 50.6)	69.4 km (66.7 – 72.1)	87.4 km (84.0 – 91.0)

Note: Statistics were derived from ~700 validation comparisons. Confidence intervals are computed using 10,000 bootstrapped samples, with 723 degrees of freedom and are significant at a 95% confidence level.

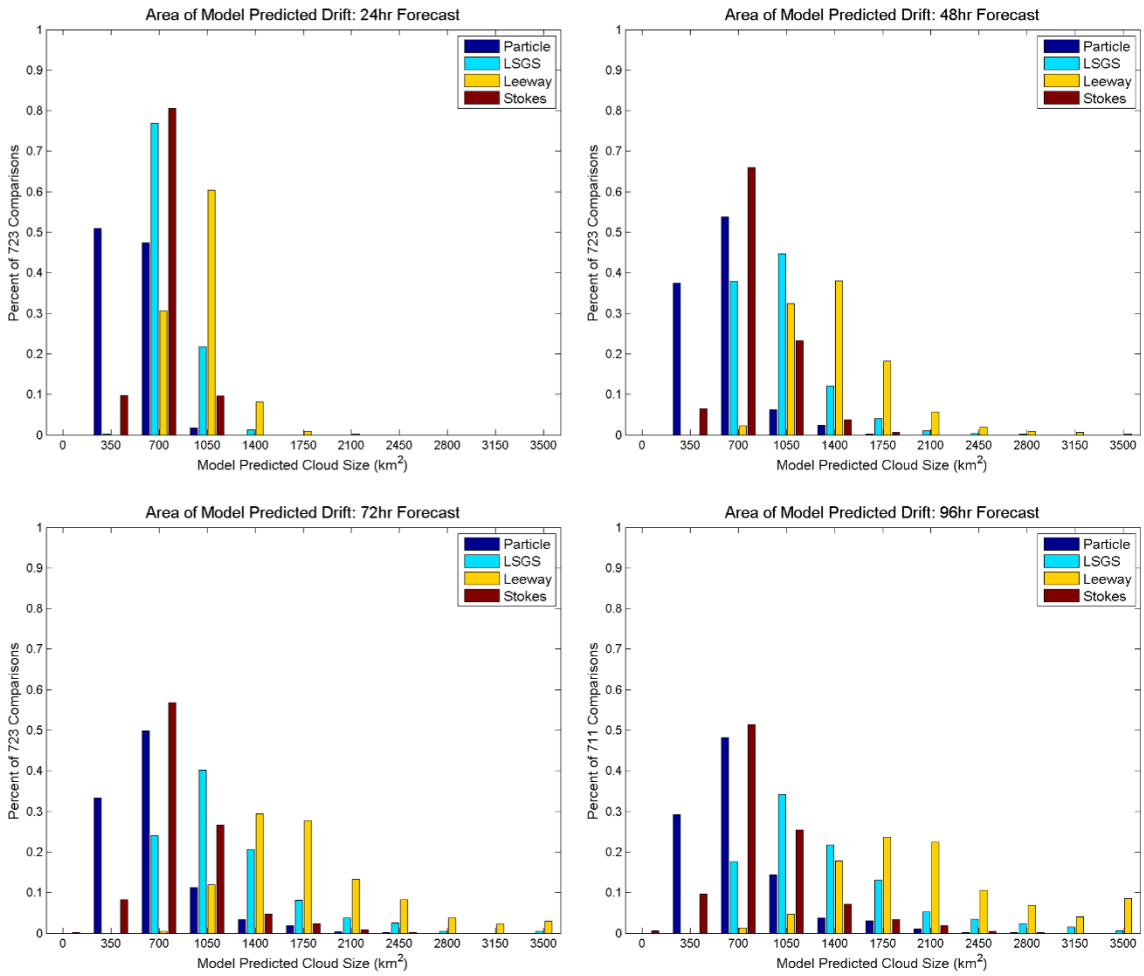

Figure 4.8 *SVP Predicted Cloud Size Statistics for Whole Domain Dataset*

Histograms showing the predicted cloud size for probability maps at 24 hr (top left), 48 hr (top right), 72 hr (bottom left), and 96 hr (bottom right) forecast times. Pure particle distributions are shown in dark blue, Lagrangian Sub-Grid Scale (LSGS) distributions are shown in light blue, Leeway distributions are shown in yellow, and Stokes distributions are shown in red.

Table 4.10 *SVP Validation Mean Cloud Size Statistics: Whole Domain Results*

Drift Algorithm	**24hr Forecast**	**48hr Forecast**	**72hr Forecast**	**96hr Forecast**
Pure Particle	545 km^2 (537 – 552)	613 km^2 (599 – 629)	675 km^2 (656 – 697)	726 km^2 (700 – 752)
LSGS	809 km^2 (800 – 819)	1015 km^2 (994 – 1036)	1193 km^2 (1160 – 1225)	1348 km^2 (1306 – 1391)
Leeway	980 km^2 (967 – 994)	1419 km^2 (1391 – 1448)	1796 km^2 (1752 – 1841)	2127 km^2 (2068 – 2189)
Stokes	677 km^2 (667 – 686)	780 km^2 (764 – 796)	830 km^2 (808 – 852)	874 km^2 (848 – 902)

Note: Statistics were derived from ~700 validation comparisons. Confidence intervals are computed using 10,000 bootstrapped samples, with 723 degrees of freedom and are significant at a 95% confidence level.

the Stokes contribution is likely minimal, if not negligible, at this depth which is not taken into account in the Stokes parameterization used in this study.

Because of the smaller sample size, confidence intervals for distance errors are much wider for the SVP comparisons than seen in the GLAD comparisons. This results in a less conclusive performance metric, as LSGS does not show any statistically significant decreases in distance error over the pure particle case. Both Leeway and Stokes show degraded performance as compared to pure particle, with increased distance errors at all forecast lengths. It is worth noting that the Stokes algorithm produces the highest distance errors, and this poor performance of the Stokes algorithm for SVP comparisons is not a surprising result, as it is perhaps the most ill-suited of the parameterizations, particularly for SVP drifters. Since SVP drifters are drogued at 15 m, the Stokes contribution is likely minimal, if not negligible, at this depth which is not taken into account in the Stokes parameterization used in this study.

Cloud size metrics show growth in cloud size for all drift algorithms as forecast length increases, as expected (Fig. 4.8 and Table 4.10). The substantial increases in cloud size for LSGS and Leeway, particularly for long forecast lengths, are again noted for the SVP derived metrics. Leeway produces a substantially bigger cloud size than all other algorithms, with an 80% bigger cloud than produced using a pure particle algorithm (21% bigger than LSGS) at the 24-hour forecast and an almost 200% bigger cloud than produced using pure particle algorithm (59% bigger than LSGS) at the 96-hour forecast.

4.2.2.2 Filtered Results

Overall, similarly as for the GLAD results, it can be said that filtering produces more skillful drift prediction than those achieved with unfiltered model fields. Though,

for the in-cloud metric, small sample size and wider confidence intervals lead to no conclusive performance increase seen with filtering for any of the algorithms or forecast lengths (Table 4.11 and Table 4.12). However, significant improvements are seen for all algorithms at all forecast lengths in distance error and cloud size metrics, with comparable decreases as those noted for the filtered case of the GLAD comparisons. Unlike GLAD though, even with the error decreases seen for filtering, no algorithm outperforms pure particle in terms of distance error or cloud size for the SVP case.

Table 4.11 *SVP Validation Metrics for Unfiltered versus Filtered RNCOM*

Forecast	**In-Cloud Percentage**		**Mean Error Distance**		**Mean Cloud Size**	
	Unfiltered	Filtered	Unfiltered	Filtered	Unfiltered	Filtered
Pure Particle						
24hr	84 % (81 – 87)	86 % (84 – 89)	24.5 km (23.7 – 25.3)	21.3 km (20.6 – 22.0)	545 km^2 (537 – 552)	439 km^2 (435 – 443)
48hr	54 % (50 – 58)	57 % (54 – 61)	42.1 km (40.4 – 44.0)	34.7 km (33.1 – 36.4)	613 km^2 (599 – 629)	426 km^2 (420 – 433)
72hr	43 % (40 – 47)	41 % (38 – 45)	59.8 km (57.0 – 62.8)	48.9 km (46.3 – 51.6)	675 km^2 (656 – 697)	414 km^2 (406 – 423)
96hr	38 % (35 – 42)	34 % (31 – 38)	75.7 km (72.0 – 79.6)	62.1 km (58.7 – 65.6)	726 km^2 (700 – 752)	406 km^2 (396 – 417)
LSGS						
24hr	92 % (90 – 94)	94 % (92 – 96)	26.1 km (25.4 – 26.8)	23.8 km (23.2 – 24.5)	809 km^2 (800 – 819)	732 km^2 (726 – 738)
48hr	76 % (73 – 79)	79 % (76 – 82)	42.7 km (41.1 – 44.4)	37.6 km (36.0 – 39.2)	1015 km^2 (994 – 1036)	824 km^2 (812 – 837)
72hr	63 % (60 – 67)	62 % (59 – 66)	59.3 km (56.7 – 62.0)	51.6 km (49.2 – 54.1)	1193 km^2 (1160 – 1225)	887 km^2 (869 – 906)
96hr	58 % (54 – 61)	53 % (50 – 57)	74.3 km (71.0 – 77.9)	64.7 km (61.5 – 68.0)	1348 km^2 (1306 – 1391)	974 km^2 (922 – 974)
Leeway						
24hr	93 % (91 – 94)	94 % (92 – 95)	27.9 km (27.3 – 28.7)	24.8 km (24.1 – 25.4)	980 km^2 (967 – 994)	863 km^2 (854 – 872)
48hr	77 % (74 – 80)	81 % (78 – 84)	47.5 km (45.9 – 49.2)	40.7 km (39.2 – 42.2)	1419 km^2 (1391 – 1448)	1157 km^2 (1141 – 1173)
72hr	69 % (66 – 72)	71 % (67 – 74)	66.5 km (64.0 – 69.1)	56.6 km (54.4 – 59.0)	1796 km^2 (1752 – 1841)	1379 km^2 (1356 – 1402)
96hr	63 % (59 – 66)	63 % (60 – 67)	83.6 km (80.3 – 87.0)	71.3 km (68.3 – 74.5)	2127 km^2 (2068 – 2189)	1554 km^2 (1523 – 1586)
Stokes						
24hr	79 % (76 – 82)	84 % (82 – 87)	27.9 km (27.1 – 28.6)	24.8 km (24.1 – 25.5)	677 km^2 (667 – 686)	603 km^2 (594 – 612)
48hr	47 % (44 – 51)	50 % (46 – 54)	48.8 km (47.1 – 50.6)	42.0 km (40.4 – 43.5)	780 km^2 (764 – 796)	642 km^2 (630 – 653)
72hr	30 % (26 – 33)	32 % (29 – 35)	69.4 km (66.7 – 72.1)	59.1 km (56.6 – 61.5)	830 km^2 (808 – 852)	634 km^2 (622 – 647)
96hr	25 % (22 – 28)	24 % (21 – 27)	87.4 km (84.0 – 91.0)	75.4 km (72.2 – 78.8)	874 km^2 (848 – 902)	621 km^2 (607 – 636)

Note: Statistics were derived from ~700 validation comparisons. Confidence intervals are computed using 10,000 bootstrapped samples, with 723 degrees of freedom and are significant at a 95% confidence level.

Table 4.12 *SVP Validation Metrics Improvements seen Using Filtered RNCOM*

Forecast	In-Cloud Percentage Increase	Mean Error Distance Decrease	Mean Cloud Size Decrease
Pure Particle			
24hr	No Statistical Difference	13 %	19 %
48hr	No Statistical Difference	18 %	31 %
72hr	No Statistical Difference	18 %	39 %
96hr	No Statistical Difference	18 %	44 %
LSGS			
24hr	No Statistical Difference	9 %	10 %
48hr	No Statistical Difference	12 %	19 %
72hr	No Statistical Difference	13 %	26 %
96hr	No Statistical Difference	13 %	28 %
Leeway			
24hr	No Statistical Difference	11 %	12 %
48hr	No Statistical Difference	14 %	18 %
72hr	No Statistical Difference	15 %	23 %
96hf	No Statistical Difference	15 %	27 %
Stokes			
24hr	No Statistical Difference	11 %	11 %
48hr	No Statistical Difference	14 %	18 %
72hr	No Statistical Difference	15 %	24 %
96hr	No Statistical Difference	14 %	29 %

Note: Statistics were derived from ~700 validation comparisons.

4.3 Validation Statistics for Coastal Domain Datasets

The validation results presented in this section are computed using subsets of the GLAD and SVP drifter datasets, confined to the CNCOM domain, as ground truth (Fig. 4.9). Metrics are presented first for the GLAD subset and then for the SVP subset. Each drifter subset is compared to CNCOM derived probability maps as well as to those derived from unfiltered and filtered RNCOM output. Validation results in this section only consider the pure particle drift algorithm, and the same three statistical metrics employed in Section 4.2 are used to assess the skill of the higher resolution ocean model for drift prediction as compared to the regional model. Assessments extend only to the 72-hour forecast as that was the extent of the CNCOM model run used in this study; 72-hours is also a typical forecast length employed for high-resolution operational model implementations at FNMOC.

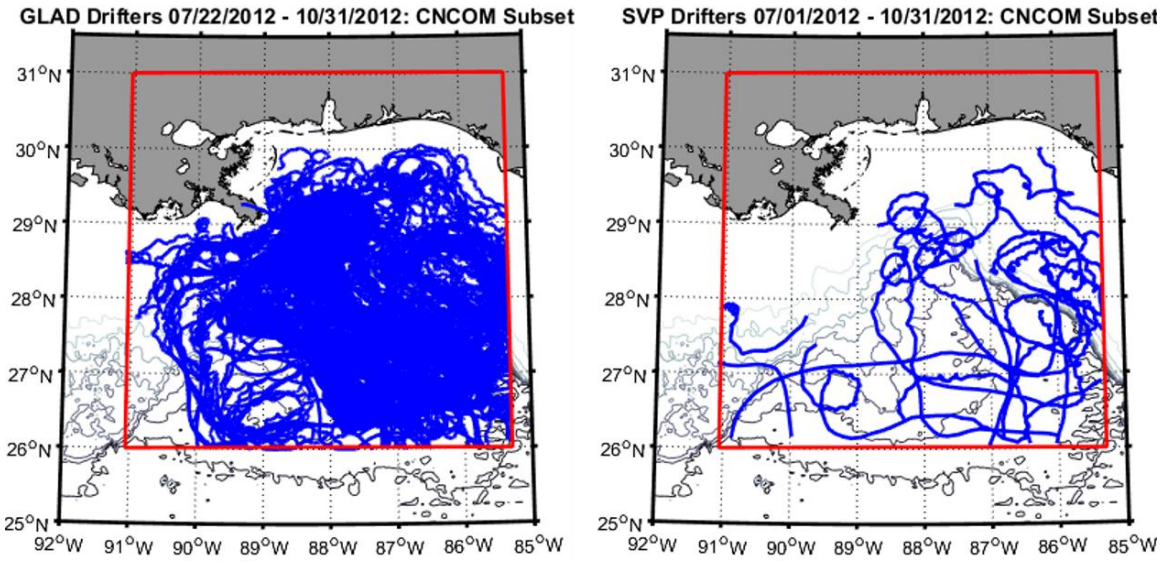

Figure 4.9 *Coastal Subset of Drifter Dataset Used to Validate High-Resolution Model Implementation*

4.3.2 GLAD Subset Validation Statistics

GLAD subset comparisons show that the higher resolution CNCOM model performs better than unfiltered and filtered RNCOM for the in-cloud percentage metric (Fig. 4.10 and Table 4.13). CNCOM outperforms unfiltered RNCOM by 13-22% and filtered RNCOM by 7-14%, with greater performance differentials seen at longer forecast times. Whereas RNCOM does not produce skillful in-cloud predictions past 48-hours (24-hours in the case of unfiltered RNCOM), CNCOM produces skillful predictions out to 72- hours. Examining histogram shapes for the in-cloud metric (Fig. 4.10), it is seen that increases in CNMOC's in-cloud skill occur mostly in the lowest probable areas.

Distance error metrics are comparable at 24-hour and 48-hour forecast times between CNCOM and unfiltered RNCOM showing only about a 3% difference in errors. At the 72-hour forecast, CNCOM outperforms unfiltered RNCOM showing distance errors that are about 8% less or ~5 km lower on average (Fig. 4.11 and Table 4.14).

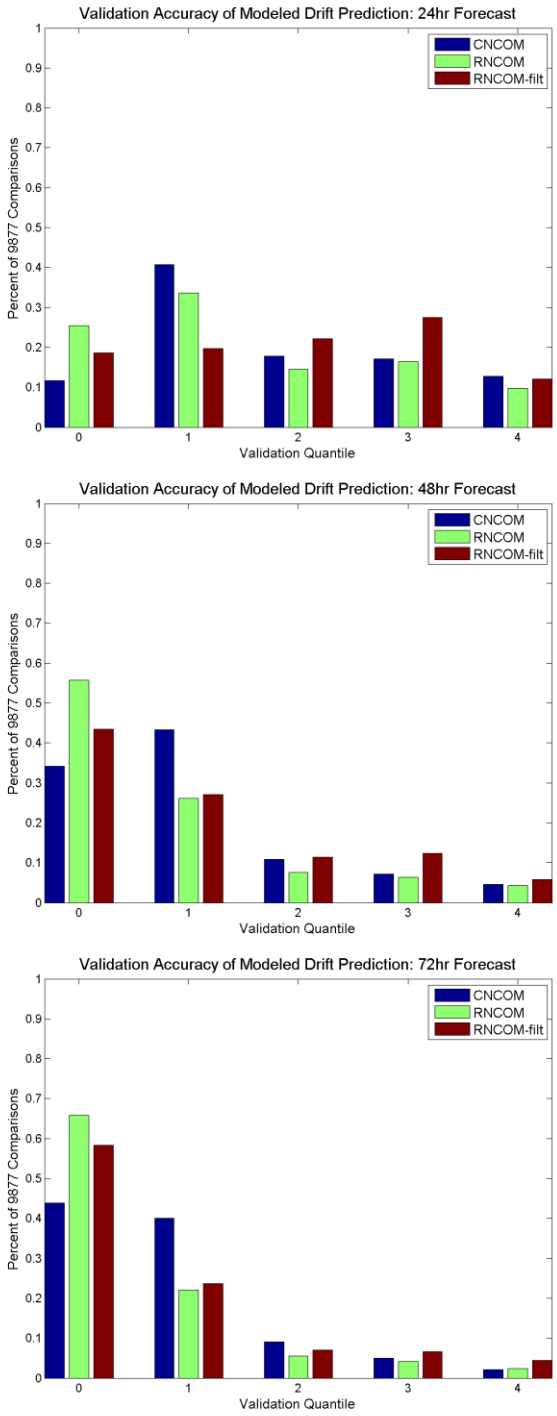

Figure 4.10 *GLAD In-Cloud Percentage Statistics for Coastal Domain Subset*

Histograms showing the percentage of time observed drifters fell within the area predicted by the drift probability map at the 24 hr (top), 48 hr (middle), and 72 hr (bottom) forecast times. Pure particle distributions from CNCOM are shown in dark blue, pure particle distributions from RNCOM are shown in green, and pure particle distributions from filtered RNCOM are shown in red.

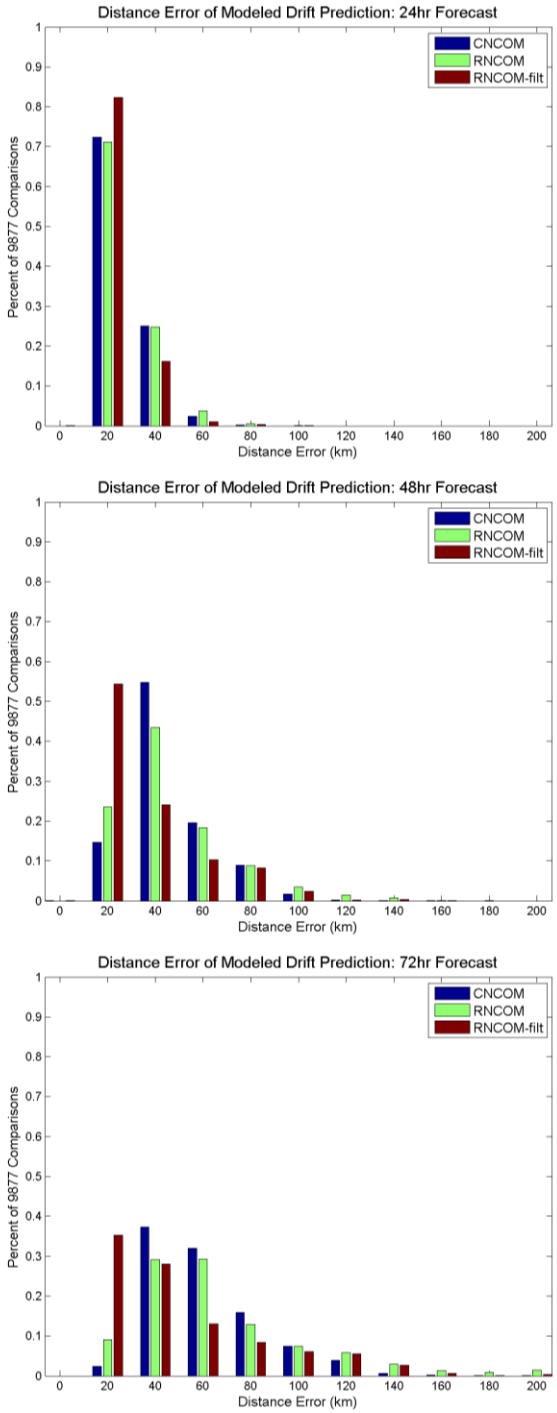

Figure 4.11 *GLAD Error Distance Statistics for Coastal Domain Subset*

Histograms showing the error distance between observed drifter positions and the center-of-mass of the predicted cloud for probability maps at 24 hr (top), 48 hr (middle), and 72 hr (bottom) forecast times. *Pure particle distributions from CNCOM are shown in dark blue, pure particle distributions from RNCOM* are *shown in green, and pure particle distributions from filtered RNCOM are shown in red.*

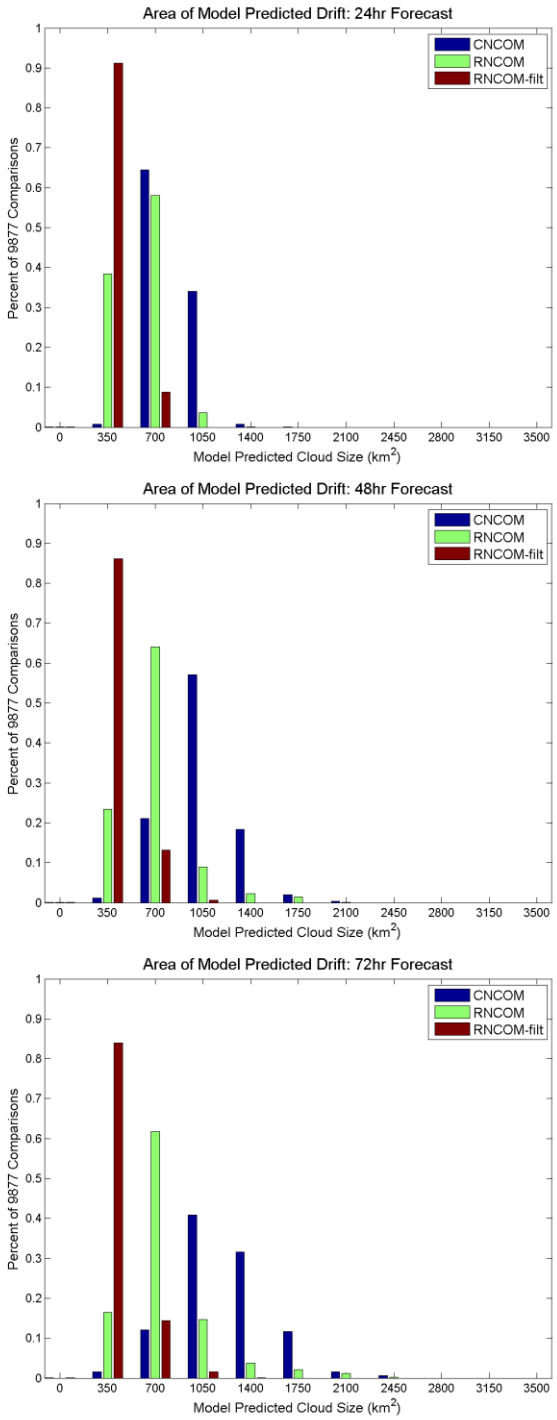

Figure 4.12 *GLAD Predicted Cloud Size Statistics for Coastal Domain Subset*

Histograms showing the predicted cloud size for probability maps at 24 hr (top), 48 hr (middle), and 72 hr (bottom) forecast times. Pure particle distributions from CNCOM are shown in dark blue, pure particle distributions from RNCOM are shown in green, and pure particle distributions from filtered RNCOM are shown in red.

Table 4.13 *GLAD Validation In-Cloud Percentage Statistics: Coastal Domain Results*

Pure Particle Drift	24hr Forecast (4km)	48hr Forecast (5km)	72hr Forecast (6km)
CNCOM	88 % (88 – 89)	66 % (65 – 67)	56 % (55 – 57)
RNCOM	75 % (74 – 75)	44 % (43 – 45)	34 % (30 – 35)
RNCOM-filtered	81 % (81 – 82)	57 % (56 – 58)	42 % (41 – 43)

Note: Statistics were derived from >9,800 validation comparisons. Confidence intervals are computed using 10,000 bootstrapped samples, with 9,877 degrees of freedom and are significant at a 95% confidence level.

Table 4.14 *GLAD Validation Mean Distance Error Statistics: Coastal Domain Results*

Pure Particle Drift	24hr Forecast (4km)	48hr Forecast (5km)	72hr Forecast (6km)
CNCOM	28.1 km (28.0 – 28.3)	46.1 km (45.7 – 46.4)	61.6 km (61.1 – 62.0)
RNCOM	27.2 km (27.0 – 27.4)	47.4 km (47.0 – 47.9)	66.9 km (66.2 – 67.6)
RNCOM-filtered	22.5 km (22.3 – 22.7)	37.0 km (36.6 – 37.5)	51.6 km (51.0 – 52.3)

Note: Statistics were derived from >9,800 validation comparisons. Confidence intervals are computed using 10,000 bootstrapped samples, with 9,877 degrees of freedom and are significant at a 95% confidence level.

Table 4.15 *GLAD Validation Mean Cloud Size Statistics: Coastal Domain Results*

Pure Particle Drift	24hr Forecast (4km)	48hr Forecast (5km)	72hr Forecast (6km)
CNCOM	838 km^2 (835 – 841)	1050 km^2 (1045 – 1055)	1212 km^2 (1206 – 1219)
RNCOM	577 km^2 (574 – 579)	677 km^2 (672 – 681)	760 km^2 (754 – 766)
RNCOM-filtered	445 km^2 (444 – 446)	439 km^2 (437 – 441)	430 km^2 (427 – 432)

Note: Statistics were derived from >9,800 validation comparisons. Confidence intervals are computed using 10,000 bootstrapped samples, with 9,877 degrees of freedom and are significant at a 95% confidence level.

Filtered RNCOM, however, shows distance errors 5 to 10 km (16-20%) lower than CNCOM at all forecast time.

Cloud size metrics show growth in cloud size for all models as forecast length increases, as expected (Fig. 4.12 and Table 4.15). However, CNCOM cloud size growth is significantly greater than for the RNCOM case. CNCOM cloud size is very similar to the cloud size numbers obtained for RNCOM drift including the LSGS parameterization (Table 4.5 and Table 4.10). CNCOM's higher resolution allows the model to resolve, at least partially, the sub-mesoscale field accounting for the increased dispersive nature of the model; whereas the only way this is achieved for RNCOM is through the addition of

the LSGS parameterization. Filtered RNCOM shows much reduced cloud size growth, as was seen in the previous section.

4.3.3 SVP Dataset Validation Statistics

The same cautions for the SVP subset must be noted as was done for the entire SVP dataset. While the GLAD statistical metrics for the CNCOM results are derived ~9,900 validation comparisons, SVP statistical metrics are derived from only ~350 validation comparisons. As a consequence of a much smaller sample size, the SVP-derived statistics are necessarily less robust than those achieved using the GLAD subset. However, as was the case in Section 4.2, many of the same trends and tendencies noted for the GLAD results are also seen for the SVP results.

CNCOM is shown to perform better than unfiltered RNCOM at all forecast lengths with respect to in-cloud prediction (Fig. 4.13 and Table 4.16), showing a 6% improvement over RNCOM at the 24-hour forecast and a 24% improvement over RNCOM at the 72-hour forecast length. Smaller sample size and wider confidence intervals lead to statistically insignificant differences for the in-cloud performance of CNCOM and filtered RNCOM at the 24-hour and 48-hour forecast lengths. However, at 72-hours CNCOM does clearly outperform filtered RNCOM, showing in-cloud percentages that are 18% higher than filtered RNCOM (Fig. 4.13 and Table 4.16). Skillful predictions follow the same trend as for the GLAD results, with CNCOM showing skillful predictions out to 72-hours, filtered RNCOM showing skillful predictions out to 48-hours, and unfiltered RNCOM showing skillful predictions out to 24-hours.

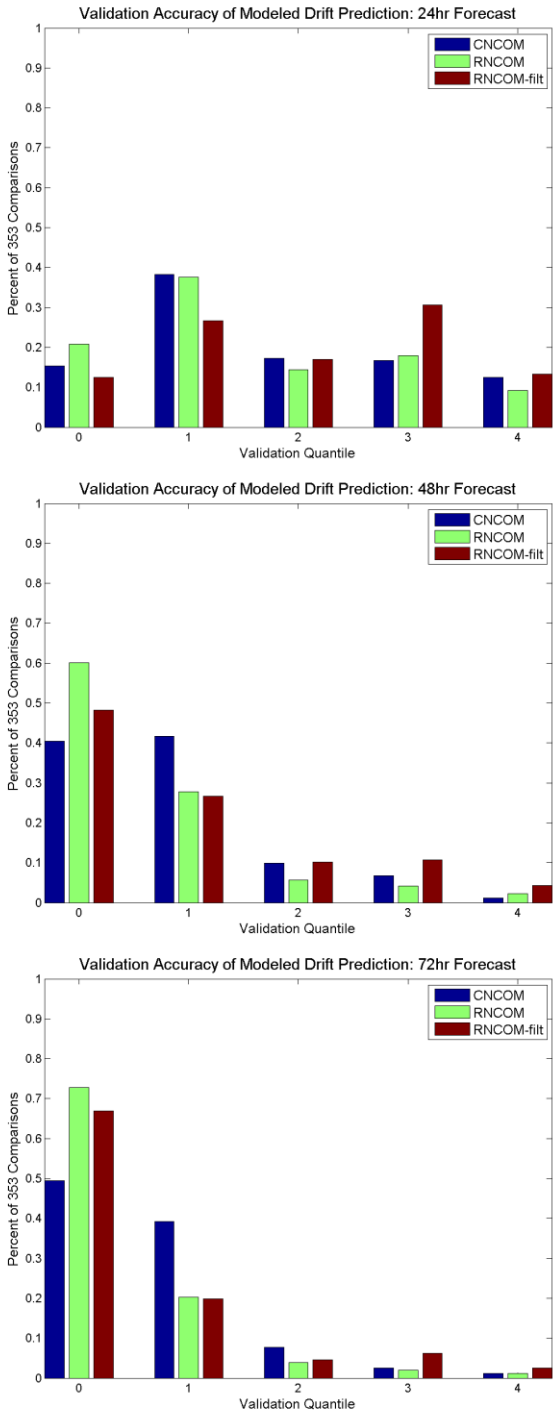

Figure 4.13 *SVP In-Cloud Percentage Statistics for Domain Coastal Subset*

Histograms showing the percentage of time observed drifters fell within the area predicted by the drift probability map at the 24 hr (top), 48 hr (middle), and 72 hr (bottom) forecast times. Pure particle distributions from CNCOM are shown in dark blue, pure particle distributions from RNCOM are shown in green, and pure particle distributions from filtered RNCOM are shown in red.

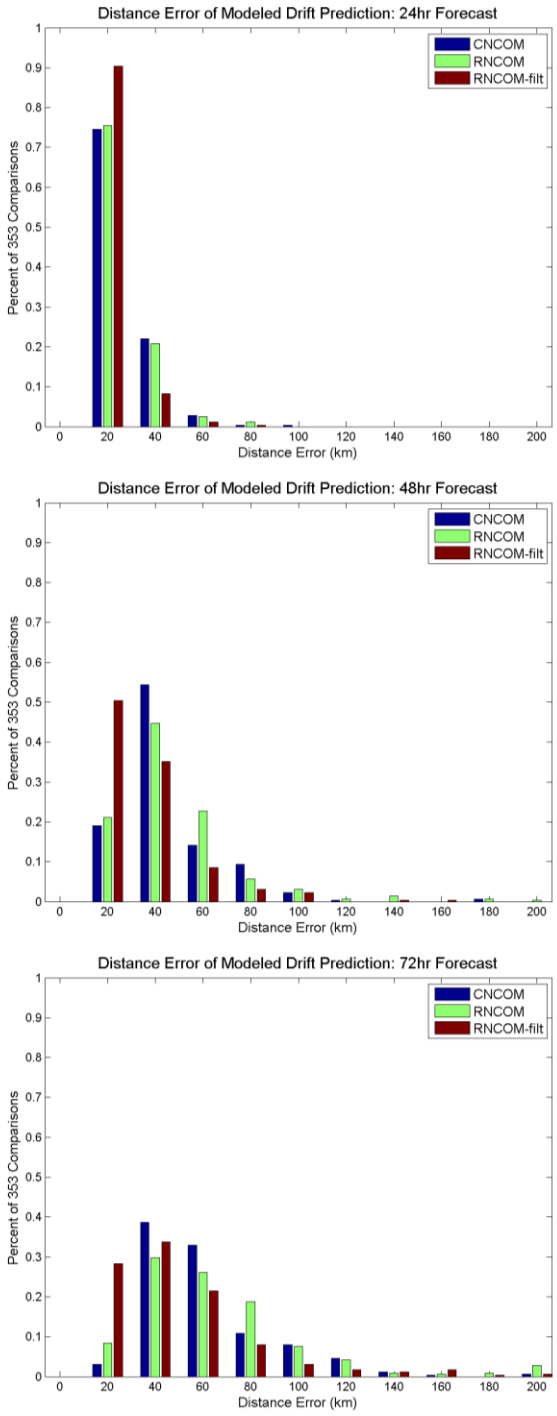

Figure 4.14 *SVP Error Distance Statistics for Coastal Domain Subset*

Histograms showing the error distance between observed drifter positions and the center-of-mass of the predicted cloud for probability maps at 24 hr (top), 48 hr (middle), and 72 hr (bottom) forecast times. Pure particle distributions from CNCOM are shown in dark blue, pure particle distributions from RNCOM are shown in green, and pure particle distributions from filtered RNCOM are shown in red.

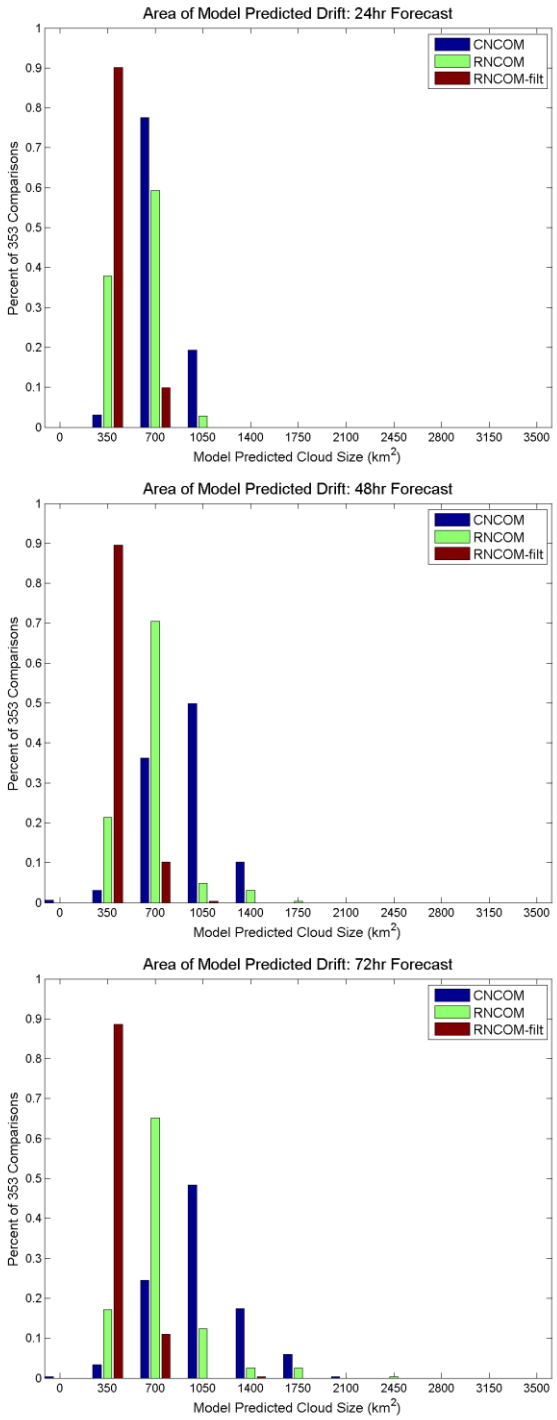

Figure 4.15 *SVP Predicted Cloud Size Statistics for Coastal Domain Subset*

Histograms showing the predicted cloud size for probability maps at 24 hr (top), 48 hr (middle), and 72 hr (bottom) forecast times. Pure particle distributions from CNCOM are shown in dark blue, pure particle distributions from RNCOM are shown in green, and pure particle distributions from filtered RNCOM are shown in red.

Table 4.16 *SVP Validation In-Cloud Percentage Statistics: Coastal Domain Subset Results*

Pure Particle Drift	24hr Forecast (4km)	48hr Forecast (5km)	72hr Forecast (6km)
CNCOM	85 % (81 – 88)	59 % (54 – 64)	51 % (45 – 56)
RNCOM	79 % (75 – 83)	40 % (35 – 45)	27 % (23 – 32)
RNCOM-filtered	88 % (84 – 91)	52 % (47 – 57)	33 % (28 – 38)

Note: Statistics were derived from ~350 validation comparisons. Confidence intervals are computed using 10,000 bootstrapped samples, with 353 degrees of freedom and are significant at a 95% confidence level.

Table 4.17 *SVP Validation Mean Error Distance Statistics: Coastal Domain Subset Results*

Pure Particle Drift	24hr Forecast (4km)	48hr Forecast (5km)	72hr Forecast (6km)
CNCOM	27.8 km (26.8 – 28.8)	45.8 km (43.7 – 48.0)	61.6 km (58.6 – 64.7)
RNCOM	26.8 km (25.7 – 28.0)	47.8 km (45.2 – 50.5)	68.0 km (64.0 – 72.3)
RNCOM-filtered	21.4 km (20.5 – 22.3)	35.5 km (33.4 – 37.7)	50.2 km (46.9 – 53.7)

Note: Statistics were derived from ~350 validation comparisons. Confidence intervals are computed using 10,000 bootstrapped samples, with 353 degrees of freedom and are significant at a 95% confidence level.

Table 4.18 *SVP Validation Mean Cloud Size Statistics: Coastal Domain Subset Results*

Pure Particle Drift	24hr Forecast (4km)	48hr Forecast (5km)	72hr Forecast (6km)
CNCOM	779 km^2 (766 – 792)	926 km^2 (902 – 949)	1040 km^2 (1009 – 1072)
RNCOM	570 km^2 (560 – 581)	652 km^2 (632 – 673)	721 km^2 (694 – 750)
RNCOM-filtered	446 km^2 (441 – 452)	430 km^2 (421 – 439)	413 km^2 (403 – 425)

Note: Statistics were derived from ~350 validation comparisons. Confidence intervals are computed using 10,000 bootstrapped samples, with 353 degrees of freedom and are significant at a 95% confidence level.

Distance error differences also produce inconclusive results between the performance of CNCOM and RNCOM, showing statistically insignificant differences at all forecast lengths (Fig. 4.14 and Table 4.17). However, it is shown even with this small sample size that filtered RNCOM produced a significant reduction in distance error at all forecast lengths, showing an improvement of 19-23%. Cloud size metrics show growth in cloud-size for CNCOM and unfiltered RNCOM (Fig. 4.15 and Table 4.18) and is significantly reduced by filtering, as seen in the GLAD results. CNCOM cloud size remains significantly greater than for RNCOM, with comparable values to RNCOM

including LSGS results calculated for the whole domain dataset (Table 4.5 and Table 4.10).

CHAPTER V - CONCLUSIONS AND NEXT STEPS

This study illustrates the absolute necessity of taking a probabilistic approach to drift prediction and validates a probabilistic methodology for creating an operational drift prediction product that can support a wide range of military and civilian applications at Fleet Numerical Meteorology and Oceanography Center (FNMOC). Drift prediction skill is quantified through metrics of in-cloud percentage, distance error, and cloud size. These metrics are used to assess the impact of different drift algorithms and underlying ocean models on the drift prediction capability, and clear limitations for modeled drift prediction are identified as well as substantive recommendations for the optimization of drift prediction skill.

Basin wide analysis shows that, using a pure particle drift algorithm, RNCOM produces a drift prediction with distance errors of ~25-kilometers at the 24-hour forecast, growing by roughly 20-kilometers per day. Additionally, pure particle RNCOM cannot predict a drifter cloud that contains the drifting buoy position more than 50% of the time beyond a 24-hour forecast. Statistics show that both the dispersion-enhancing parameterizations of LSGS and Leeway result in improved predictive skill, providing drift predictions that encompass actual drifting buoy positions ~90% of the time at the 24-hour forecast and ~60% of the time at the 96-hour forecast. Both of these algorithms increase the in-cloud performance of the prediction at all forecast lengths as well as extend the in-cloud skillfulness (> 50% probability) by 72-hours beyond what is possible using a pure particle drift algorithm. Increased in-cloud performance for both LSGS and Leeway is attributed to the increased cloud sizes produced by each algorithm. Since most of the gains in performance are seen in the areas defined by the model to be the lowest

probable, the implication is that the parameterizations do not improve the accuracy of the velocity prediction itself but instead are able to account for enough of the velocity field's uncertainty by enhancing dispersion (i.e. increasing cloud size), thereby improving the predictive performance. Ironically, large cloud size is also the greatest shortcoming for the LSGS and Leeway parameterizations. Large cloud sizes lead to much longer search times, which in a search and rescue situation could mean the difference between life and death. However, as Leeway produces significantly larger cloud sizes than LSGS, LSGS is the more desirable option especially considering that LSGS produces lower distance errors at all forecast times than is achieved with the Leeway parameterization. While pure particle produces the lowest distance errors at all forecast lengths, errors computed using the LSGS algorithm are comparable. Additionally, the Stokes parameterization, as defined in this study, is not shown to provide any advantage for drift prediction. However, weakness in the way the Stokes addition was assessed in this study (i.e. no exponential decay with depth accounted for in parameterization algorithm) and the existence of better methods for incorporating the Stokes contribution to the velocity field (i.e. direct calculation of Stokes drift from a co-located wave model, and/or implicit accounting of Stokes drift using a strongly coupled wave-ocean model) warrants further investigation of the Stokes impact on drift predictive skill.

The great utility of filtering for drift applications is demonstrated, with filtering increasing the skillfulness of predictions for pure particle drift to greater than 80% at the 24-hour forecast and greater than 50% at the 48-hour forecast. Distance errors are reduced by 13-22% over the forecast for the pure particle case. Lower reductions are seen for the parameterizations tested, but all see distance error reductions between 9-18%.

Additionally, filtering produces a significantly smaller cloud size (10-49%) than those predicted using unfiltered fields for all algorithms and at all forecast times. While filtering causes some degradation in skill for the LSGS algorithm, LSGS with filtering retains skillfulness out to the 96-hour forecast and has the added benefit of cloud size predictions that are ~30-40% smaller at 96-hours than those produced from unfiltered fields. Filtering also improves performance of the Leeway algorithm; however, even though filtering reduces cloud sizes, all metrics indicate the comparatively better performance with the addition of LSGS.

In light of these results, spatial filtering is a recommended post-processing step for regional-scale model velocity fields that are to be utilized for drift applications. As a pure particle algorithm produces the lowest distance errors, smallest cloud size, and an in-cloud statistic greater than 80% at the 24-forecast, it is recommended as the optimal choice for short-term forecast prediction of 24-hours or less. Even though distance errors and cloud sizes remain smallest for pure particle throughout the rest of the forecast, due to its significantly improved in-cloud performance at longer forecast lengths, LSGS is the recommendation for drift predictions beyond 48-hours.

Higher-resolution model fields show better performance over the regional-scale model. Just as with the LSGS and Leeway parameterizations, CNCOM's performance gains are due to its enhanced dispersive character rather than a more accurate current prediction, as the biggest gains for in-cloud percentage are seen in the areas defined by the model to be the lowest probable. However, CNCOM does not significantly outperform RNCOM after filtering has been applied. Additionally, due to its ability to resolve more of the sub-mesoscale field, the higher-resolution CNCOM suffers from the

same shortfall as RNCOM with LSGS and Leeway, producing drift predictions with substantially larger search areas. Further work is needed to determine if a model of comparable resolution to CNCOM would benefit from filtering and if it would then outperform a filtered regional-scale model. It must also be noted, that higher-resolution models come at greater computational costs, which necessarily limits their spatial extent in an operational environment. Therefore, often times the situational location of the prediction required dictates the use of regional-scale models (and in some instances, global-scale models) over higher-resolution, limited area models.

The high-resolution model results do, however, reveal the paradoxical relationship between drift performance and the presence of small-scale features. The inclusion of small-scale, sub-mesoscale features increases the dispersion of the model velocity field bringing it closer to observed values; this leads to improved drift predictive skill over a coarser resolution model that does not resolve these features as well. However, filtering model velocities and removing these small-scale features also leads to better drift predictive skill. This counter-intuitive finding highlights the fact that drift prediction depends on both the accuracy of the velocity field as well as the accuracy of the model's dispersive character. A higher resolution model allows for a more complete representation of the small-scale, sub-mesoscale field, which leads to greater dispersion that is much closer to what is observed. Higher dispersion leads to larger cloud size that naturally leads to a greater probability that the object of interest falls within the area predicted by the model. However, while sub-mesoscale features increase dispersion, they are also unconstrained by observations and thus are not predictable features. Therefore, filtering them out produces a more accurate velocity estimate, which also leads to

increased drift prediction skill. This counter-balance between the scale of features needed to accurately forecast dispersion (e.g. sub-mesoscale) and the scale of features that can be skillfully predicted (e.g. mesoscale) will remain until the observations exist that allow the intersection of these two scales.

While small cloud size is a desired characteristic for search-and-rescue or search-and-recovery operations, this may not be the case for other types of drift applications. For example, applications for pollutant or biologic (e.g., algae or larvae) drift, in some instances, may favor predictions with larger cloud sizes as they better capture uncertainty and highlight potentially impacted areas. Therefore, the drift application being supported must be considered when making choices with respect to the underlying ocean model and drift algorithm used for the prediction.

Lastly, looking forward, there are several upcoming advancements that are expected to improve upon the modeling capability for drift prediction. Data assimilation of velocity observations is expected to be integrated into the operational environment at FNMOC over the next few years, and inclusion of this field in the assimilation process has been shown to improve trajectory predictions for operational ocean models (Muscarella 2015). Additionally, the National Aeronautics and Space Administration's (NASA's) Surface Water and Ocean Topography (SWOT) satellite, which is expect to launch in 2021, is set to provide operational ocean models with a much higher spatial and temporal look at the SSHA field than current satellite altimetry platforms deliver and will resolve features down to a 15- to 30-kilometer scale (Morrow et al. 2019). Preliminary work by NRL to assess the impact of this new observational data stream to ocean model performance suggests that ocean models are likely to see significant improvements in

their ability to predict the mesoscale eddy field along with the surface velocity field (Carrier 2016). Both of these advancements will necessitate the re-evaluation of the drift predictability of the Navy's operational ocean modeling capability.

APPENDIX – RELATIVE DISPERSION CALCULATIONS

For all relative dispersion calculations, particle pairs are defined as two drifters with initial deployment locations within 10 km and 1-hour of one another. The relative dispersion calculation used to determine the optimal settings for the LSGS parameterization included initializing a single synthetic model drifter at the observed drifter's initial deployment location interpolated to the nearest model output time, with modeled trajectories subsequently integrated out to 25-days. Conversely, the relative dispersion calculation used to determine the ensemble size required for a final drift product with reproducible dispersion characteristics was based purely on modeled trajectories and drift integrations were only carried out to 4-days. Because there was a desire to compute zonal and meridional dispersion separately for ensemble testing, particle pairs were taken from a 20 km disc of initial particles. The definition of a particle pair is the same in all cases; however, for the case computed purely from modeled trajectories, the additional time constraint is not needed, as all trajectories are initialized simultaneously.

Since two single model trajectories necessarily sample less of the velocity field's spatial variabilitity than a 20 km disc of particles, the number of ensembles needed to give reproducible dispersion characteristics differs between the two cases. More ensemble releases are needed for reproducible dispersion curves in the case where a single model trajectory is initialized at each pair location found in the GLAD drifter dataset (Fig. A.1 and Fig. A.2). For this reason, an ensemble value of 10 vice 5 is cited for cases where particle pairs are defined based on the observational drifter dataset. It should be noted, however, the testing done using pairs taken from a 20 km disc of

particles is most appropriate for ensuring a reproducible drift product, as model trajectories are initialized from a disc of particles instead of from a single point to compute the drift probability maps being validated in this study.

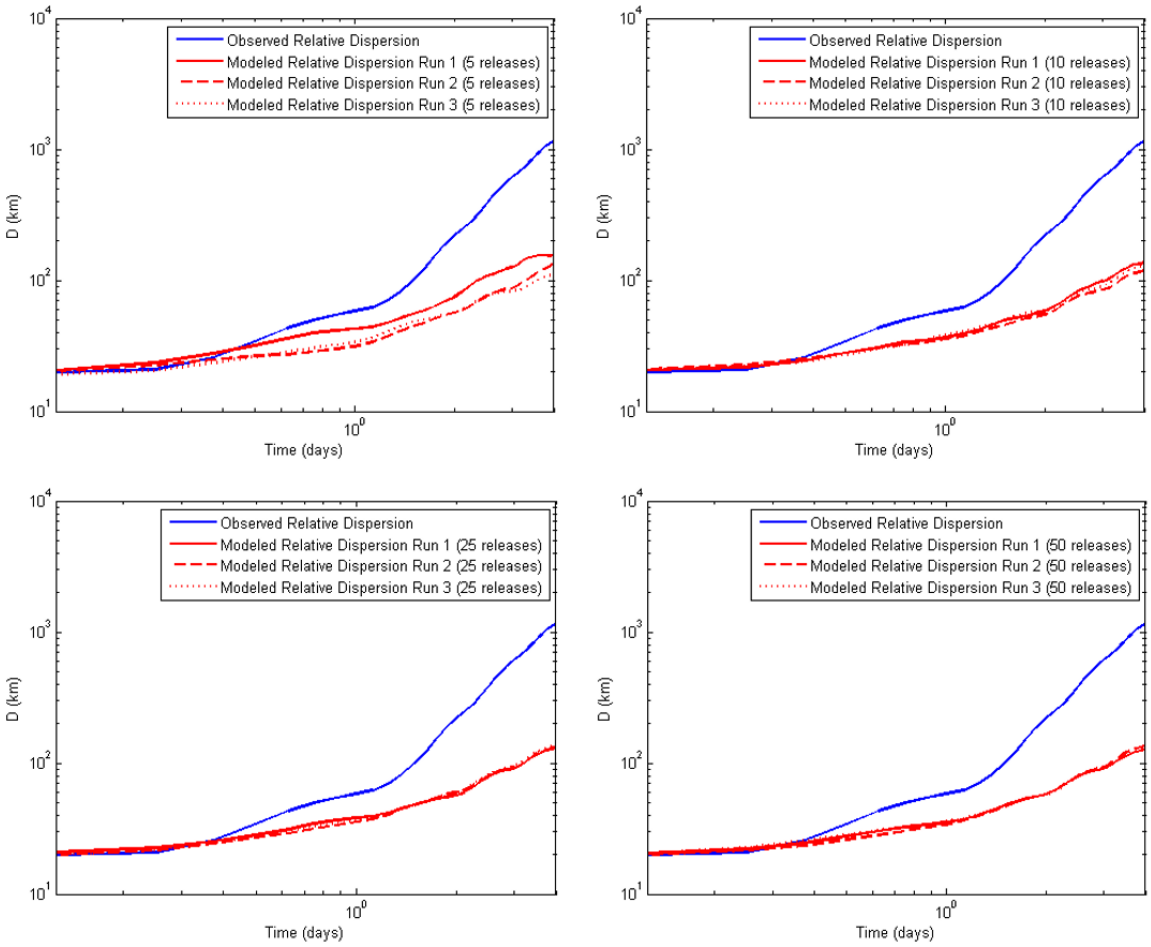

Figure A.1 *Reproducibility of Modeled Dispersion Characteristics using Particle Pair Locations Identified in the GLAD Drifter Dataset*

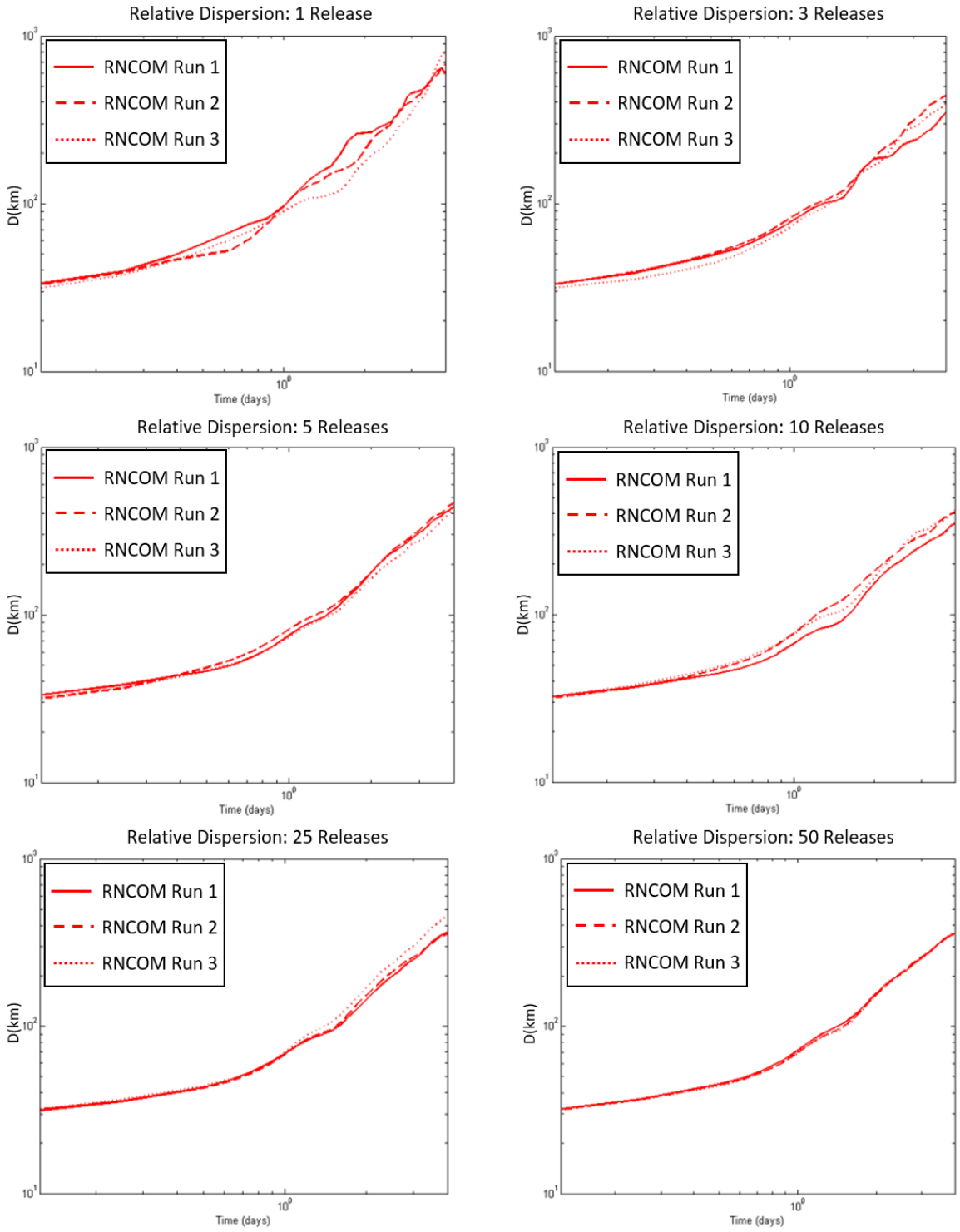

Figure A.2 *Reproducibility of Modeled Dispersion Characteristics using Particle Pairs taken from a 20-km Disc of Initial Particle Locations*

REFERENCES

Allen, A.J., 2005: Leeway divergence. U.S. Coast Guard Rep. CG-D-05-05, 128pp, http://www.stormingmedia.us/53/5345/A534534.html.

Anis, A., J.N. Moum, Surface wave-turbulence interactions: scaling ε(z) near the sea surface. *J. Phys. Oceanogr*., **25**, 2025-2045.

Aref, H., 1984: Stirring by chaotic advection. *J. Fluid Mech.*, **143**, 1-21.

Barron, C.N., A.B. Kara, P.J. Martin, R.C. Rhodes, and L.F. Smedstad, 2006: Formulation, implementation and examination of vertical coordinate choices in the Global Navy Coastal Ocean Model (NCOM). *Ocean Modell*., **11**, 347-375.

Barron, C.N., L.F. Smedstad, J.M. Dastugue, and O.M. Smedstad, 2007: Evaluation of ocean models using observed and simulated drifter trajectories: impact of seas surface height on synthetic profiles for data assimilation. *J. Geophys. Res*., **112** (C7), C07019

Biggs, D.C., A.E. Jochens, M.K. Howard, S.F. DiMarco, K.D. Mullin, R.R. Leben, F.E. Muller-Karger, and C. Hu, 2005: Eddy forced variations in on- and off-margin summertime circulation along the 1000-m isobaths of the northern Gulf of Mexico, 2000-2003, and the links with sperm whale distributions along the middle slope. *Circulation in the Gulf of Mexico: Observations and Models*, *Geophys. Monogr. Series*, No. 161, Amer. Geophys. Union, 71-85.

Boicourt, W.C, W.J. Wiseman, Jr., A. Valle-Levinson and L.P. Atkinson, 1998: Continental shelf of the southeastern United States and the Gulf of Mexico: In the shadow of the western boundary current. *The Sea, The Global Coastal Ocean:*

Regional Studies and Synthesis, A.R. Robinson and K.H. Brink, Eds., John Wiley & Sons, Inc., 135-182.

Breivik, O., and A.A. Allen, 2008: An operational search and rescue model for the Norwegian Sea and the North Sea. *J. Mar. Syst*., **69**, 599-610.

Caballero, A., M. Espino, Y. Sargarminaga, L. Ferrer, A. Uriarte, and M. Gonzalez, 2008: Simulating the migration of drifters deployed in the Bay of Biscay, during the Prestige crisis. *Mar. Pollut. Bull*., **56**, 475-482.

Carrier, M.J., H.E. Ngodock, S.R. Smith, I. Souopgui, and B. Bartels, 2016: Examining the potential impact of SWOT observations in an ocean analysis-forecasting system. *Mon. Wea. Rev*., **144**, 3767-3782.

Chen, C. S., R. O. Reid, and W. D. Nowlin, 1996: Near-inertial oscillations over the Texas Louisiana shelf, *J. Geophys. Res.*, **101**, 3509– 3524.

Cho, K., Y. Li, H. Wang, K. Park, J. Choi, K. Shin, and J. Kwon, 2014: Development and validation of an operational search and rescue modeling system for the Yellow Sea and the East and South China Seas. *J. Atmos. Ocean. Tech*., **31**, 197-215.

Cowen, R. K., C.B. Paris, and A. Srinivasan, 2006: Scaling of connectivity in marine populations. *Science*, **311**, 522-527.

Craig, P.D., M.L. Banner, 1994: Modeling wave-enhanced turbulence in the ocean surface layer *J. Phys. Oceanogr*., **24**, 2546-2559.

Crout, R.L., W.J. Wiseman, Jr., and W.S. Chuang, 1984: Variability of wind-driven currents, west Louisiana inner continental shelf: 1978-1979, *Contri. Mar. Sci*., **27**, 1-11.

Crout, R.L., and R. Wiley, 2010: Improved quality of National Data Buoy Center (NDBC) Acoustic Doppler Profiler (ADCP) Measurements. *Proc. of the MTS/IEEE OCEANS Conf. and Exhibition*. 2010, Seattle, OR, Institute of Electrical and Electronics Engineers, https://ieeexplore.ieee.org/document/5664596/.

Davis, R.E., 1985a: Drifter observations of coastal surface currents during CODE: the method and descriptive overview. *J. Geophys. Res*., **90**, 4741-4755.

Davis, R.E., 1985b: Drifter observations of coastal surface currents during CODE: the statistical and dynamical views. *J. Geophys. Res*., **90**, 4756-4722.

De Velasco, G.G. and C.D. Winant, 1996: Seasonal patterns of wind stress and wind stress curl over the Gulf of Mexico. *J. Geophys. Res. Oceans*, **101**, 18,127-18,140.

DiMarco, S.F. and R.O. Reid, 1998: Characterization of the principle tidal current constituents on the Texas-Louisiana shelf. *J. Geophys. Res*., **100**, 3093-3109.

DiMarco, S. F., M. K. Howard, and R. O. Reid, 2000: Seasonal variation of wind-driven diurnal current cycling on the Texas-Louisiana Continental Shelf, *Geophys. Res. Lett.*, **27**, 1017– 1020.

Elliott, B.A., 1982: Anticyclonic rings in the Gulf of Mexico. *J. Phys. Oceanogr*., **12**, 1292-1309.

Geyer, W.R., 1989: Field calibration of mixed layer drifters. *J. Atmos. Oceanic Technol*., **6**, 333-342.

Gough, M. K., A. J. H. M. Reniers, J.H. MacMahan, and S.D. Howden, 2016: Resonant near-surface inertial oscillations in the northeastern Gulf of Mexico, *J. Geophys. Res. Oceans*, 121, 2163– 2182.

Griffa, A., 1996: Applications of stochastic particle models to oceanographic problems. *Stochastic modelling in Physical Oceanography*, R. Adler, P. Mu¨ller, B. Rozovskii, Eds., Birkhäuser, 113–128.

Griffa, A., A. Haza, T.M. Özgökmen, A. Molcard, V. Taillandier, K. Schroeder, Y. Chang, and P.M. Poulain, 2013: Investigating transport pathways in the ocean. *Deep-Sea Res. II*, **85**, 81-95.

Hackett, B., and O. Breivik, C. Wettre, 2006: Forecasting the drift of objects and substances in the ocean. *Ocean Weather Forecasting: An Integrated View of Oceanography*, E.P. Chassignet and J. Verron, Eds. Springer, 507-523.

Hansen, D.V., P.P. Poulain, 1995: Quality Control and Interpolations of WOCE-TOGA Drifter Data, *J. Atmos. Oceanic Technol*., **13**, 900-909.

Haller, G., and A.C. Poje, 1998: Finite time transport in aperiodic flows. *Physica D*, **119**, 352-380.

Haza, A.C., L.I. Piterbarg, P. Martin, T.M., Özgökmen, and A. Griffa, 2007: A Lagrangian subgridscale model for particle transport improvement and application in the Adriatic Sea using the Navy Coastal Ocean Model. *Ocean Modell.*, **17**, 68–91.

Haza, A.C., T.M. Özgökmen, A. Griffa, Z.D. Garraffo, and L.I. Piterbarg, 2012: Parametrization of particle transport at submesoscales in the Gulf Stream region using Lagrangian subgridscale models. *Ocean Modell.,* **42**, 31– 49.

Hsu, C., 2018: Wind-Driven Near-Inertial Waves and Their Impact in the Gulf of Mexico. Doctoral dissertation, Texas A & M University, 90 pp. http://hdl.handle.net/1969.1/173605.

Hsu, S.A., E.A. Meindl, and D.B. Gilhousen, 1994: Determining the power-law wind-profile exponent under near-neutral stability conditions at sea, *J. Appl. Meteor*., **33**, 757 – 765.

Jacobs, G.A., J.M. D'Addezio, B. Bartels, P.L. Spence, 2019: Constrained scales in ocean forecasting. *Adv. Space Res*., https://doi.org/10.1016/j.asr.2019.09.018.

Jarosz, E., Z.R. Hallock, W.J. Teague, 2007: Near-Inertial Currents in the Desoto Canyon region, *Continental Shelf. Res*., **27**, 2407-2426.

Kantha, L., 2005: Barotropic tides in the Gulf of Mexico. *Circulation in the Gulf of Mexico: Observations and Models, Geophys. Monogr. Series*, No. 161, Amer. Geophys. Union, 159-163.

Kitaigorodskii, S.A., M.A. Donelan, J.L. Lumley, and E.A. Terray, 1983: Wave turbulence interactions in the upper ocean. Part II: Statistical characteristics of wave and turbulent components of the random velocity field in the marine surface layer. *J. Phys. Oceanogr*., **13**, 1988-1999.

Kundu, P.K., 1976: Ekman veering observed near the ocean bottom, *J. Phys. Oceanogr*., **6**, 238-242.

LaCasce, J.H., 2008: Statistics from Lagrangian observations. *Prog. Oceanogr*., **77**, 1-29.

Lipphardt, B. 2013: GLAD experiment CTD casts, northern Gulf of Mexico near DeSoto Canyon, July-August 2012. Distributed by: Gulf of Mexico Research Initiative Information and Data Cooperative (GRIIDC), Harte Research Institute, Texas A&M University–Corpus Christi. https://data.gulfresearchinitiative.org/data/R1.x134.073:0005. Accessed 2020-06-23.

Locke, L., and R.L. Crout, 2009: A study on the validity of buoy mounted Acoustic Doppler Current Profilers: A comparison of upward and downward looking systems in Onslow Bay, NC, *Proc. of the MTS/IEEE OCEANS Conf. and Exhibition*. 2009, Biloxi, MS, Institute of Electrical and Electronics Engineers, https://ieeexplore.ieee.org/document/5422277/.

Love, M.R., and C.J. Amante, B.W. Eakins, and L.A. Taylor, 2012: Digital Elevation Models of the Northern Gulf Coast: Procedures, Data Sources and Analysis, NOAA *Technical Memorandum NESDIS NGDC-59*, U.S. Dept. of Commerce, Boulder, CO, 43 pp.

Lumpkin, R., L. Centurioni, 2019: Global Drifter Program quality-controlled 6-hour interpolated data from ocean surface drifting buoys. Gulf of Mexico, July-October 2012. NOAA National Centers for Environmental Information. Dataset. https://doi.org/10.25921/7ntx-z961. Accessed 2017-03-27.

Lunde, B.N., and E.F. Coelho, 2009: Implementations of the Navy Coupled Ocean Data Assimilation system at the Naval Oceanographic Office, *Proc. of the MTS/IEEE OCEANS Conf. and Exhibition*. 2009, Biloxi, MS, Institute of Electrical and Electronics Engineers, https://ieeexplore.ieee.org/document/5422232.

Marinone, S.J., O.Q. Gutiérrez, and A. Parés-Sierra, 2004: Numerical simulation of larval shrimp dispersion in the northern region of the gulf of California, *Estuar Coast Shelf Sci*., **60**, 611−617.

Mellor, G.L., and A. Blumberg, 2004: Wave breaking and ocean surface layer thermal response. *J. Phys. Oceanogr.*, **34**, 693-698.

Mellor, G.L., and T. Yamada, 1982: Development of a turbulence closure model for geophysical fluid problems. *Rev. Geophys. Space Phys*., **20**, 851-875.

Metzger, E.J., O.M. Smedstad, P.G. Thoppile, H.E. Hurlburt, J.A. Cummings, A.J. Wallcraft, L. Zamudio, D.S. Franklin, P.G. Posey, M.W. Phelps, P.J. Hogan, F.L. Bub, and C.J. Dehaan, 2014: U.S. Navy operational global ocean and Arctic ice prediction systems, *Oceanograph.*, **27**, 32-43.

Meyers, S., 1994: Cross-frontal mixing in a meandering jet. *J. Phys. Oceanograph.*, **24**, 1641-1646.

Morrow R, L. Fu, F. Ardhuin, M. Benkiran, B. Chapron, E. Cosme, F. d'Ovidio, J.T. Farrar, S.T. Gille, G. Lapeyre, P. Le Traon, A. Pascual, A. Ponte, B. Qiu, N. Rascle, C. Ubelmann, J. Wang, E.D. Zaron, 2019: Global Observations of Fine-Scale Ocean Surface Topography With the Surface Water and Ocean Topography (SWOT) Mission. *Front. Mar. Sci*., **6**, 1-19.

Muscarella, P., M.J. Carrier, H. Ngodock, S. Smith, B.L. Lipphardt Jr., A.D. Kirwan, and H.S. Huntley, 2015: Do assimilated drifter velocities improve Lagragian predictability in an operational ocean model*? Mon. Wea. Rev.*, **143**, 1822-1832.

Niiler, P.P, R. Davis, and H. White, 1987: Water following characterization of a mixed layer drifter. *Deep Sea Res*., **34**, 1867-1882.

Niiler, P.P., A.S. Sybrandy, K. Bi, P.M. Poulain, and D. Bitterman, 1995: Measurements of the water-following capability of holey-sock and TRISTAR drifters. *Deep-Sea Res. I*, **42**, 1951-1964.

Nowlin, W.D., A.E. Jochens, S.F. DiMarco, R.O. Reid, and M.K. Howard, 2005: Low-frequency circulation over the Texas-Louisiana continental shelf. *Circulation in*

the Gulf of Mexico: Observations and Models, Geophys. Monogr. Series, No. 161, Amer. Geophys. Union, 219-240.

Ohlmann, J.C., and P.P. Niiler, 2005: Circulation over the continental shelf in the northern Gulf of Mexico. *Prog. Oceanogr.*, **64**, 45-81.

Osborn, T., D.M. Farmer, S. Vagle, S.A. Thorpe, and M. Cure, 1992: Measurements of bubble plumes and turbulence from a submarine. *Atmos. Ocean*., **30**, 419-440.

Özgökmen, T., 2013: GLAD experiment CODE-style drifter trajectories (low-pass filtered, 15-minute interval records), northern Gulf of Mexico near DeSoto Canyon, July-October 2012. Distributed by: Gulf of Mexico Research Initiative Information and Data Cooperative (GRIIDC), Harte Research Institute, Texas A&M University–Corpus Christi. doi:10.7266/N7VD6WC8

Paris, C.B., R.K. Cowen, R. Claro, and K.C. Lindeman, 2005: Larval transport pathways from Cuban snapper (Lutjanidae) spawning aggregations based on biophysical modeling, *Mar. Ecol. Prog. Ser*., **296**, 93−106.

Poje, A.C., and G. Haller, 1999: Geometry of cross-stream mixing in a double-gyre ocean model. *J. Physic. Oceanogr*., **29**, 1649-1665.

Poje, A., A. Haza, T. Özgökmen, M. Magaldi, and Z. Garraffo, 2010: Resolution dependent relative dispersion statistics in a hierarchy of ocean models. *Ocean Modell.*, **31**, 36–50.

Poje, A., T.M. Özgökmen, B.L. Lipphardt Jr., B.K. Haus, E.H. Ryan, A.C. Haza, G.A. Jacobs, A. J. H. M. Reniers, M.J. Olascoaga, G. Novelli, A. Griffa, F.J. Beron-Vera, S.S. Chen, E. Coelho, P.J. Hogan, A.D. Kirwan Jr., H.S. Huntley, and A.J. Mariano, 2014: Submesoscale dispersion in the vicinity of the

Deepwater Horizon spill. *Proc. Natl. Acad. Sci. USA*, **111**, 12,693-12,698.

Pollard, R. T., 1970: On the generation by winds of inertial waves in the ocean. *Deep Sea Res. Oceanogr. Abst.*, **17**, 795– 812.

Poulain, P., R. Gerin, 2019: Assessment of the water-following capabilities of CODE drifters based on direct relative flow measurements. *J. Atmos. Oceanic Technol*., **36**, 621-633.

Price, J.M., M. Reed, M.K. Howard, W.R. Johnson, Z.G. Ji, C.F. Marshall, N.L. Guinasso, and G.B. Rainey, 2006: Preliminary assessment of an oil-spill trajectory model using satellite-tracked, oil-spill simulating drifters. *Environ. Model. Softw*., **21**, 258-270.

Samuelson, R.M., 1992: Fluid exchange across a meandering jet, *J. Phys. Oceanogr.*, **22**, 431-440.

Samuelson, R.M., 1996: Chaotic transport by mesoscale motions. *Stochastic Modelling in Physical Oceanography*, R.J. Adler, P. Muller, and B.L. Rozovoskii, Eds., Birkhäuser, 423-438.

Schmitz, W.J., D.C. Biggs, A. Lugo-Fernandez, L.Y. Oey, and W. Sturges, 2005: A synopsis of the circulation in the Gulf of Mexico and on its continental margins. *Circulation in the Gulf of Mexico: Observations and Models*, *Geophys. Monogr. Series*, No. 161, Amer. Geophys. Union, 11-29.

Seim, H.E., B. Kjerfve, and J.E. Sneed, 1987: Tides of Mississippi Sound and the Adjacent Continental Shelf. *Estuar. Coast Shelf S.*, **25**, 143-156.

Smagorinsky, J., 1963: General circulation experiments with the primitive equations. I: The basic experiment. *Mon. Wea. Rev.*, **91**, 99-164.

Sutherland, G., N. Sootiens, F. Davidson, G.C. Smith, N. Bernier, H. Blanken, D. Schillinger, G. Marcotte, G. Röhrs, K, Dagestad, K.H. Christensen, O, Breivik, 2020: Evaluating the Leeway coefficient for different ocean drifters using operational models. Preprint available at http://arxiv.org/abs/2005.09527.

Stokes, G.G., 1847: On the theory of oscillatory waves. *Trans. Cambridge Philos. Soc*., **8**, 441-473.

Thomson, R.E., and W.J. Emery, 2014: Statistical Methods and Error Handling. *Data Analysis Methods in Physical Oceanography*, R.E. Thomson and W.J. Emery, Eds. Elsevier, 219-311.

Thorpe, S.A., 1984: On the determination of Kv in the near-surface ocean from acoustic measurements of bubbles. *J. Phys. Oceanogr*., **14**, 855-863.

Thorpe, S.A., 1992: Bubble clouds and the dynamics of the upper ocean. *Quart. J. Roy. Meteor. Soc*., **118**, 2498-2505.

Wu, J., 1983: Sea-surface drift currents induced by wind and waves. *J. Phys. Oceanogr*., **13**, 1441-1451.

Yang, H., and Z. Liu, 1997: The three-dimensional chaotic transport and the great ocean barrier. *J. Phys. Oceanogr.,* **27**, 1258-1273.

Zhang, X., S. F. Dimarco, D. C. Smith Iv, M. K. Howard, A. E. Jochens, and R. D. Hetland, 2009: Near-resonant ocean response to sea breeze on a stratified continental shelf, *J. Phys. Oceanogr.*, **39**, 2137– 2155.

Zhang, F., 2013: Inertial Oscillations in the Gulf of Mexico during 2005 Hurricane Season. M.S. thesis, Dept. of Oceanogr. and Coastal Sci., Louisiana State University, 79 pp.