https://doi.org/10.7289/V5/TM-NOS-CS-39

NOAA Technical Memorandum NOS CS 39

ASSESSMENT OF WATER LEVEL, SEA-SURFACE TEMPERATURE, AND SALINITY GUIDANCE FROM THREE NOAA MODELS OF THE WESTERN NORTH ATLANTIC

Silver Spring, Maryland
July 2017

noaa National Oceanic and Atmospheric Administration

U.S. DEPARTMENT OF COMMERCE
National Ocean Service
Coast Survey Development Laboratory

Office of Coast Survey
National Ocean Service
National Oceanic and Atmospheric Administration
U.S. Department of Commerce

The Office of Coast Survey (OCS) is the Nation's only official chartmaker. As the oldest United States scientific organization, dating from 1807, this office has a long history. Today it promotes safe navigation by managing the National Oceanic and Atmospheric Administration's (NOAA) nautical chart and oceanographic data collection and information programs.

There are four components of OCS:

The Coast Survey Development Laboratory develops new and efficient techniques to accomplish Coast Survey missions and to produce new and improved products and services for the maritime community and other coastal users.

The Marine Chart Division acquires marine navigational data to construct and maintain nautical charts, Coast Pilots, and related marine products for the United States.

The Hydrographic Surveys Division directs programs for ship and shore-based hydrographic survey units and conducts general hydrographic survey operations.

The Navigational Services Division is the focal point for Coast Survey customer service activities, concentrating predominately on charting issues, fast-response hydrographic surveys, and Coast Pilot updates.

https://doi.org/10.7289/V5/TM-NOS-CS-39

NOAA Technical Memorandum NOS CS 39

ASSESSMENT OF WATER LEVEL, SEA-SURFACE TEMPERATURE, AND SALINITY GUIDANCE FROM THREE NOAA MODELS OF THE WESTERN NORTH ATLANTIC

Philip Richardson and Zizang Yang
Office of Coast Survey, Coast Survey Development Laboratory,
Silver Spring, Maryland

July 2017

noaa National Oceanic and Atmospheric Administration

U. S. DEPARTMENT OF COMMERCE
Wilbur Ross, Secretary

National Oceanic and Atmospheric Administration
Benjamin Friedman, Acting Under Secretary

National Ocean Service
Dr. Russell Callender, Assistant Administrator

Office of Coast Survey
Rear Admiral Shepard Smith

Coast Survey Development Laboratory
Captain Edward J. Vandenameele
Division Chief

NOTICE

Mention of a commercial company or product does not constitute an endorsement by NOAA. Use for publicity or advertising purposes of information from this publication concerning proprietary products or the tests of such products is not authorized.

TABLE OF CONTENTS

LIST OF FIGURES v

LIST OF TABLES x

EXECUTIVE SUMMARY xi

1. INTRODUCTION 1
    1.1. G-RTOFS, ETSS, and ESTOFS 2
    1.2. Observed Data 9

2. METHODS OF MODEL SKILL ASSESSMENT 15
    2.1. Model and Data Time Series 15
    2.2. Forecast Cycle (FC) Based Method 19
    2.3. Forecast Hour (FH) Based Method 19

3. PERFORMANCE OF WATER LEVEL FORECASTS 23
    3.1. G-RTOFS 23
    3.2. ETSS 32
    3.3. ESTOFS 41
    3.4. Summary 50

4. PERFORMANCE OF G-RTOFS FOR SEA-SURFACE TEMPERATURE 53
    4.1. Compared with CO-OPS Observations 53
    4.2. Compared with NDBC Buoy Observations 62
    4.3. Compared with WOA09 Data 71
    4.4. Summary 74

5. PERFORMANCE OF G-RTOFS FOR SEA-SURFACE SALINITY 77
    5.1. Compared with NDBC Buoy Observations 77
    5.2. Compared with WOA09 Data 86
    5.3. Summary 89

6. SUMMARY 91

ACKNOWLEDGMENTS 93

REFERENCES 93

APPENDIX A. CO-OPS WATER LEVEL STATIONS USED FOR THE Water Level SKILL ASSESSMENT 95
APPENDIX B. CO-OPS PHYSICAL OCEANOGRAPHY OBSERVATION STATIONS USED for THE G-RTOFS SST SKILL ASSESSMENT 101
APPENDIX C. NDBC BUOY STATIONS USED for THE G-RTOFS SST SKILL ASSESSMENT 103
APPENDIX D. NDBC Stations Used for the G-RTOFS SSS Skill Assessment 105

LIST OF FIGURES

Figure 1.1. An illustration of the G-RTOFS grid with each cell representing 54 rows and 75 columns of the grid. This grid plot is obtained from http://polar.ncep.noaa.gov /global/about/... 3
Figure 1.2. The G-RTOFS forecasts of (a) subtidal water level, (b) SST, and (c) SSS fields at 0300 UTC on 1 January 2013 in U.S. East Coast and Gulf of Mexico waters 3
Figure 1.3. The ETSS model domains 5
Figure 1.4. Station locations of the ETSS point forecast guidance along the U.S. East Coast and Gulf of Mexico. This figure is a screen shot from http://www.nws.noaa.gov/mdl/ etsurge/index.php?page=map&region=ne&datum=msl&list=&map=0-48 &type=&stn=.. 6
Figure 1.5. The ESTOFS horizontal grid 8
Figure 1.6. Maps of CO-OPS water level stations used for evaluating performances of (a) G-RTOFS, (b) ETSS, and (c) ESTOFS subtidal water level forecasts.. 9
Figure 1.7. The map of the CO-OPS physical oceanography observation stations selected for evaluating the performance of G-RTOFS SST forecast.....11
Figure 1.8. The map of NDBC buoy stations selected for evaluating the performance of G-RTOFS SST forecast. 12
Figure 1.9. The map of NDBC buoy stations selected for evaluating the performance of G-RTOFS SSS forecast.....12
Figure 2.1. Subtidal water level time series of ETSS (blue) and CO-OPS (red) at CO-OPS water level station 8413320 during the ETSS forecast cycle on 10/01/2013 16
Figure 2.2. Subtidal water level time series of G-RTOFS (blue) and CO-OPS (red) at CO-OPS water level station 8413320 during the G-RTOFS forecast cycle on 10/01/2013.... 16
Figure 2.3. Subtidal water level time series of ESTOFS (blue) and CO-OPS (red) at CO-OPS water level station 8413320 during the ESTOFS forecast cycle on 10/31/2013 17
Figure 2.4. Sea-surface temperature time series of G-RTOFS (blue) and CO-OPS at CO-OPS station 9415020 (red) during the G-RTOFS forecast cycle 00z UTC 4/21/2013. 17
Figure 2.5. Sea-surface temperature time series of G-RTOFS (blue) and NDBC buoy 46247 (red) during the G-RTOFS forecast cycle on 4/21/2013. 18
Figure 2.6. Sea-surface temperature time series of G-RTOFS (blue) and WOA09 (at NDBC SST station 21) (red) (Table C.1 in appendix C) during the G-RTOFS forecast cycle on 10/21/2013. 18
Figure 2.7. Sea-surface salinity time series of G-RTOFS (blue) and WOA09 Station 13 (red) (Table D.1 in appendix D) during the G-RTOFS forecast cycle on 10/21/2013. 19
Figure 2.8. Data structure of the two-dimensional model-data time series array 20

Figure 2.9. Forecast hour based daily sub-tidal water level time series of G-RTOFS (blue) and CO-OPS observations (red) from 12/30/2012 to 02/03/2013 21
Figure 2.10. Forecast hour based daily SSH time series of ETSS (blue) and CO-OPS observations (red) from 9/28/2012 to 11/12/2012. 21
Figure 2.11. Forecast hour based daily SSH time series of ESTOFS (blue) and CO-OPS observations (red) from 1/01/2013 to 2/01/2013 22
Figure 2.12. Forecast hour based daily SST time series of G-RTOFS (blue) and CO-OPS observations (red) from 9/28/2012 to 11/12/2012. 22
Figure 3.1. Color coded RMSE maps of G-RTOFS water level forecasts in (a) January 2013, (b) April 2013, (c) July 2013, and (d) October 2013 at 52 CO-OPS water level stations (Table A.2).....24
Figure 3.2. Mean (red squares) and standard deviation (blue error bars) of the G-RTOFS water level forecast errors in: (a) January 2013, (b) April 2013, (c) July 2013, and (d) October 2013 at 52 CO-OPS water level stations (Table A.2).....25
Figure 3.3. Station-averaged root mean squared errors of G-RTOFS water level forecasts in (a) January 2013, (b) April 2013, (c) July 2013, and (d) October 2013 27
Figure 3.4. Color coded RMSE maps of G-RTOFS water level forecasts in January 2013. The six plots correspond to forecast guidance hours (a) 6, (b) 24, (c) 48, (d) 72, (e) 96, and (f) 144 28
Figure 3.5. Color coded RMSE maps of G-RTOFS water level forecasts in April 2013. The six plots correspond to forecast guidance hours (a) 6, (b) 24, (c) 48, (d) 72, (e) 96, and (f) 144 29
Figure 3.6. Color coded RMSE maps of G-RTOFS water level forecasts in July 2013. The six plots correspond to forecast guidance hours (a) 6, (b) 24, (c) 48, (d) 72, (e) 96, and (f) 144 30
Figure 3.7. Color coded RMSE maps of G-RTOFS water level forecasts in October 2013. The six plots correspond to forecast guidance hours (a) 6, (b) 24, (c) 48, (d) 72, (e) 96, and (d) 144.....31
Figure 3.8. Color coded RMSE maps of the ETSS water level forecast in (a) January 2013, (b) April 2013, (c) July 2013, and (d) October 2013 at 47 CO-OPS water level stations (Table A.1).....33
Figure 3.9. Means (red squares) and standard deviations (blue error bars) of the ETSS water level forecast errors in (a) January 2013, (b) April 2013, (c) July 2013, and (d) October 2013 at 47 CO-OPS water level stations (Table A.1) 34
Figure 3.10. The RMSE of the ETSS water level forecast at each forecast hour in (a) January 2013, (b) April 2013, (c) July 2013, and (d) October 2013……..36
Figure 3.11. Color coded RMSE maps of the ETSS water level forecast in January 2013. The four plots correspond to forecasts at hours (a) 12, (b) 24, (c) 48, and (d) 96.....37
Figure 3.12. Color coded RMSE maps of the ETSS water level forecast in April 2013. The four plots correspond to forecasts at hours (a) 12, (b) 24, (c) 48, and (d) 96.....38

Figure 3.13. Color coded RMSE maps of the ETSS water level forecast in July 2013. The four plots correspond to forecasts at hours (a) 12, (b) 24, (c) 48, and (d) 96 39
Figure 3.14. Color coded RMSE maps of the ETSS water level forecast in October 2013. The four plots correspond to forecasts at hours (a) 12, (b) 24, (c) 48, and (d) 96 40
Figure 3.15. Color coded RMSE maps of ESTOFS water level forecasts in (a) January 2013, (b) April 2013, (c) July 2013, and (d) October 2013 at 54 CO-OPS water level stations (Table A.3).....42
Figure 3.16. Mean (red squares) and standard deviation (blue error bars) of the ESTOFS water level errors for 2013 at 54 CO-OPS water level stations from (a) January 2013, (b) April 2013, (c) July 2013, and (d) October 2013 43
Figure 3.17. Station-averaged root mean squared errors of ESTOFS water level forecasts in (a) January 2013, (b) April 2013, (c) July 2013, and (d) October 2013 45
Figure 3.18. Color coded RMSE maps of ESTOFS water level forecasts in January 2013. The six plots correspond to forecast guidance hours (a) 6, (b) 24, (c) 48, (d) 96, (e) 144, and (f) 180 46
Figure 3.19. Color coded RMSE maps of ESTOFS water level forecasts in April 2013. The six plots correspond to forecast guidance hours (a) 6, (b) 24, (c) 48, (d) 96, (e) 144, and (f) 180 47
Figure 3.20. Color coded RMSE maps of ESTOFS water level forecasts in July 2013. The six plots correspond to forecast guidance hours (a) 6, (b) 24, (c) 48, (d) 96, (e) 144, and (f) 180 48
Figure 3.21. Color coded RMSE maps of ESTOFS water level forecasts in October 2013. The six plots correspond to forecast guidance hours (a) 6, (b) 24, (c) 48, (d) 96, (e) 144, and (f) 180 49
Figure 4.1. Color coded RMSE maps of the G-RTOFS SST forecast in (a) January 2013, (b) April 2013, (c) July 2013, and (d) October 2013 at 45 CO-OPS physical oceanography observation stations (Table B.1).....54
Figure 4.2. Means (red squares) and standard deviations (blue error bars) of the G-RTOFS SST errors in (a) January 2013, (b) April 2013, (c) July 2013, and (d) October 2013 at 45 CO-OPS physical oceanography observation stations (Table B.1) 55
Figure 4.3. Root mean squared errors (RMSE) of the G-RTOFS SST forecast at forecast hours 1-144 in (a) January 2013, (b) April 2013, (c) July 2013, and (d) October 2013 57
Figure 4.4. Color coded RMSE maps of the G-RTOFS SST forecast in January 2013. The six plots correspond to forecasts at hours (a) 6, (b) 24, (c) 48, (d) 72, (e) 96, and (f) 144 58
Figure 4.5. Color coded RMSE maps of the G-RTOFS SST forecast in April 2013. The six plots correspond to forecasts at hours (a) 6, (b) 24, (c) 48, (d) 72, (e) 96, and (f) 144 59

Figure 4.6. Color coded RMSE maps of the G-RTOFS SST forecast in July 2013. The six plots correspond to forecasts at hours (a) 6, (b) 24, (c) 48, (d) 72, (e) 96, and (f) 144 60
Figure 4.7. Color coded RMSE maps of the G-RTOFS SST forecast in October 2013. The six plots correspond to forecasts at hours (a) 6, (b) 24, (c) 48, (d) 72, (e) 96, and (f) 144 61
Figure 4.8. Color coded RMSE maps of the G-RTOFS SST forecast in (a) January 2013, (b) April 2013, (c) July 2013, and (d) October 2013 at 78 NDBC stations (Table C.1) 63
Figure 4.9. Means (red squares) and standard deviations (blue error bars) of the G-RTOFS SST errors in (a) January 2013, (b) April 2013, (c) July 2013, and (d) July 2013 at 78 NDBC stations (Table C.1) 64
Figure 4.10. Root mean squared errors (RMSE) of the G-RTOFS SST forecast at forecast hours 1-144 in (a) January 2013, (b) April 2013, (c) July 2013, and (d) October 2013 66
Figure 4.11. Color coded RMSE maps of the G-RTOFS SST forecast in January 2013. The model-data comparison was made against SST measured at 78 NDBC buoys (Table C.1). The six plots correspond to forecasts at hours (a) 6, (b) 24, (c) 48, (d) 72, (e) 96, and (f) 144 67
Figure 4.12. Color coded RMSE maps of the G-RTOFS SST forecast in April 2013. The model-data comparison was made against SST measured at 78 NDBC buoys (Table C.1). The six plots correspond to forecasts at hours (a) 6, (b) 24, (c) 48, (d) 72, (e) 96, and (f) 144 68
Figure 4.13. Color coded RMSE maps of the G-RTOFS SST forecast in July 2013. The model-data comparison was made against SST measured at 78 NDBC buoys (Table C.1). The six plots correspond to forecasts at hours (a) 6, (b) 24, (c) 48, (d) 72, (e) 96, and (f) 144 69
Figure 4.14. Color coded RMSE maps of the G-RTOFS SST forecast in October 2013. The model-data comparison was made against SST measured at 78 NDBC buoys (Table C.1). The six plots correspond to forecasts at hours (a) 6, (b) 24, (c) 48, (d) 72, (e) 96, and (f) 144 70
Figure 4.15. Color coded RMSE maps of the G-RTOFS SST forecast in (a) January 2013, (b) April 2013, (c) July 2013, and (d) October 2013 at 72 WOA09 data grid points (Table C.1) 72
Figure 4.16. Means (red squares) and standard deviations (blue error bars) of the G-RTOFS SST errors in (a) January 2013, (b) April 2013, (c) July 2013, and (d) October 2013 at 72 WOA09 data grid points (Table D.1) 73
Figure 5.1. Color coded RMSE maps of the G-RTOFS SSS forecast in (a) January 2013, (b) April 2013, (c) July 2013, and (d) October 2013 at 9 NDBC stations (Table D.1) 78

Figure 5.2. Means (red squares) and standard deviations (blue error bars) of the G-RTOFS SSS errors in (a) January 2013, (b) April 2013, (c) July 2013, and (d) October 2013 at 9 NDBC stations (Table D.1) 79
Figure 5.3. Root mean squared error (RMSE) of the G-RTOFS SSS forecast at forecast hours 1-144 in (a) January 2013, (b) April 2013, (c) July 2013, and (d) October 2013 81
Figure 5.4. Color coded RMSE maps of the G-RTOFS SSS forecast in January 2013. The model-data comparison was made against SSS measured at 9 NDBC buoys (Table D.1). The six plots correspond to forecasts at hours (a) 6, (b) 24, (c) 48, (d) 72, (e) 96, and (f) 144 82
Figure 5.5. Color coded RMSE maps of the G-RTOFS SSS forecast in April 2013. The model-data comparison was made against SSS measured at 9 NDBC buoys Table D.1). The six plots correspond to forecasts at hours (a) 6, (b) 24, (c) 48, (d) 72, (e) 96, and (f) 144 83
Figure 5.6. Color coded RMSE maps of the G-RTOFS SSS forecast in July 2013. The model-data comparison was made against SSS measured at 9 NDBC buoys Table D.1). The six plots correspond to forecasts at hours (a) 6, (b) 24, (c) 48, (d) 72, (e) 96, and (f) 144 84
Figure 5.7. Color coded RMSE maps of the G-RTOFS SSS forecast in October 2013. The model-data comparison was made against SSS measured at 9 NDBC buoys Table D.1). The six plots correspond to forecasts at hours (a) 6, (b) 24, (c) 48, (d) 72, (e) 96, and (f) 144 85
Figure 5.8. Color coded RMSE maps of the G-RTOFS SSS forecast in (a) January 2013, (b) April 2013, (c) July 2013, and (d) October 2013 at 72 WOA09 data grid points (refer to Appendix C) 87
Figure 5.9. Means (red squares) and standard deviations (blue error bars) of the G-RTOFS SSS errors in (a) January 2013, (b) April 2013, (c) July 2013, and (d) October 2013 at 72 WOA09 data grid points (refer to Appendix C) 88

LIST OF TABLES

Table 3.1. Forecast cycle-based statistics of the G-RTOFS and ETSS SSH errors.....51
Table 3.2. Forecast hour-based statistics of the G-RTOFS and ETSS SSH errors.....51
Table 4.1. Forecast cycle-based statistics of the G-RTOFS SST errors 75

EXECUTIVE SUMMARY

The Operational Nowcast/Forecast Systems (OFS), presently being developed by the National Ocean Service (NOS) of the National Oceanic and Atmospheric Administration (NOAA), make use of sea-surface height (SSH), sea-surface temperature (SST), and sea-surface salinity (SSS), forecast guidance from the Global Real-Time Ocean Forecast System (G-RTOFS), and SSH forecast guidance from the Extra-Tropical Storm Surge (ETSS) model and the Extratropical Surge and Tide Operational Forecast System (ESTOFS). The OFS use these forecast guidance to form their open ocean boundary forcings. To support future development of NOS OFS in the eastern U.S. coastal waters, we assessed the performance of the G-RTOFS forecast guidance for SSH, SST, and SSS, as well as the performance of the ETSS and ESTOFS forecast guidance for SSH.

Our intention is to gain insight into the model performance from two perspectives: (1) the model performance across a forecast cycle (FC) and (2) the evolution of performance associated with the forecast hour (FH) of the forecast cycle. Accordingly, we developed a FC based method and a FH based method, which are further described in Chapter 2. We applied the FC based method to estimate the bias, standard deviation, and root-mean-squared error of a forecast cycle over a series of cycles. The FH based method was applied to estimate the root-mean-squared error for each given forecast hour of the forecast cycle over a series of cycles.

We conducted model-data comparisons over four separate month long periods: January 2013, April 2013, July 2013, and October 2013. These months roughly correlate to winter, spring, summer, and fall, respectively. In this way, the analysis covers all four seasons of the year.

Table A lists the station averaged bias, standard deviation, and root-mean-squared error. These seasonally averaged values were calculated using the FC based method. As can be seen from the table, these model systems, especially G-RTOFS, produce forecast results with a very large negative bias.

Table A**.** Forecast cycle-based statistics of the G-RTOFS, ETSS, and ESTOFS SSH errors.

	SSH of G-RTOFS (144 hour Forecast Cycle)			**SSH of ETSS (96 hour Forecast Cycle)**			**SSH of ESTOFS (180 hour Forecast Cycle)**		
	bias (cm)	std (cm)	rmse (cm)	bias (cm)	std (cm)	rmse (cm)	bias (cm)	std (cm)	rmse (cm)
Mean	-49.1	7.9	51.0	-13.5	5.6	15.7	-12.0	9.6	17.5

Figure A depicts the performance of water level forecast guidance, in terms of RMSE (root mean square error) calculated by the FH based method and averaged over all stations, of the G-RTOFS, ETSS, and ESTOFS model systems through the first 48 hours of their forecast cycles. In general, the forecast guidance of SSH from the three systems demonstrates satisfactory agreement with

observations. Though the skill of each model system varies by season, the figure depicts values which are averaged over all four seasons. As can be seen from Figure A, the skill of each model system degrades as the forecast hour progresses through the cycle. The figure also demonstrates that for each of the systems, the rate of skill deterioration is greatest during the initial 24 hours. The rate of deterioration then levels out during the remainder of the forecast cycle. Figure A indicates that on average, ETSS has the highest skill (lowest RMSE) in forecasting water levels. Following ETSS are G-RTOFS, then ESTOFS.

The mean values which appear in Table B were obtained by averaging all the hourly values, hours 1 through 48, for each of the model systems. Note that the mean RMSE values which appear in Table B are significantly smaller than the mean RMSE values appearing in Table A. The values appearing in Table B were calculated using the FH based method. The FH based method includes an adjustment (initial model point – initial observed point) added to the entire forecast cycle for each day. This offset effectively removes the bias, resulting in a much smaller RMSE value.

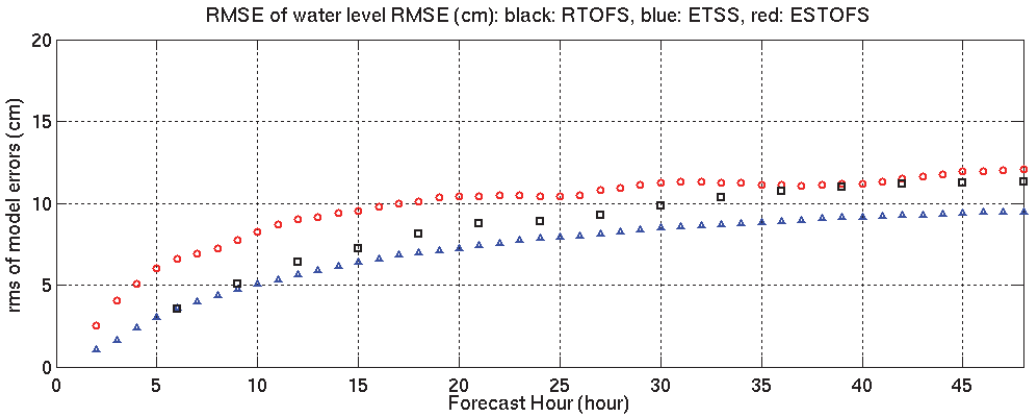


Figure A. RMSE of SSH forecasts from the ESTOFS (red), ETSS (blue) and G-RTOFS (black) models as a function of forecast hour. The plot depicts RMSE over the first 48 hours, using the Forecast Hour method.

Table B. RMSE of the ETSS, G-RTOFS, and ESTOFS SSH forecasts using the Forecast Hour method, averaged over the first 48 hours.

	ETSS (cm)	**G-RTOFS (cm)**	**ESTOFS (cm)**
Average RMSE	7.2	8.1	9.9

We evaluated the performance of the G-RTOFS SST forecast guidance by comparing the model results with observed data from CO-OPS physical oceanography observation stations, NDBC buoys, and the WOA09 database. As can be seen in Table C, the RMSE from the WOA09 comparison is the largest in magnitude, while the NDBC data fit best.

Table C. Forecast cycle-based RMSE errors (°C) of the G-RTOFS SST as compared to temperatures from three data sources.

Data sources	CO-OPS	NDBC	WOA09
Average RMSE	1.68	1.12	5.00

We evaluated the performance of the G-RTOFS SSS forecast guidance by comparing the model results with observed data from NDBC buoys and the WOA09 database. The SSS forecast guidance of G-RTOFS demonstrates poor agreement with data from the nine NDBC buoy stations. The station bias ranges between 0 and 20 psu. The large RMSE values occur at stations 4, 5, and 6 which are located within the estuaries of the mid-Atlantic. The station averaged RMSE is highest during the summer at about 10.5 psu and lowest during the fall at about 7.5 psu. The SSS forecast guidance of G-RTOFS demonstrates a satisfactory agreement with the WOA09 database. The station bias ranges between -4 and 4 psu through all four seasons, including outlier values.

1. INTRODUCTION

The National Ocean Service of NOAA has been developing operational forecast systems to produce nowcast/forecast guidance of ocean state variables including water levels, sea surface temperature, sea surface salinity, and three-dimensional (3-D) currents in the estuaries, coastal waters, and shelf waters of the U.S. These operational forecast systems produce valuable information to support the safety of maritime navigation, emergency response, and coastal environment management. The backbone of the various operational forecast systems are the hydrodynamic models. These models are forced with tidal and subtidal water levels, temperature, salinity, and currents on the model domain's open ocean boundary, as well as with meteorological forcing on the surface and with river discharge at the river entrances. Open boundary forcing plays a critical role in the accuracy of the OFS nowcast/forecast guidance.

The National Weather Service (NWS) of NOAA has developed the Global Real-Time Ocean Forecast System (G-RTOFS) and the Extra-Tropical Storm Surge (ETSS) Model. These two operational systems are run at the National Center for Environmental Prediction (NCEP) of NOAA. The OFS developed by NOS normally use subtidal water level forecast guidance from G-RTOFS and ETSS and 3-dimensional (3-D) temperature and salinity forecast guidance from G-RTOFS to drive their hydrodynamic model runs. The OFS use the ETSS subtidal water level output as backup when the G-RTOFS water level output is not available. Hence, it is worthwhile to evaluate the performance of the two models. As a first step in model evaluation, this project focuses on assessing the G-RTOFS skill in forecasting subtidal water levels, SST, and SSS, as well as on the skill of ETSS and ESTOFS in forecasting water levels.

In the present project, we are assessing the G-RTOFS, ETSS, and ESTOFS skill in the U.S. eastern coastal waters and we focus on four months: January, April, July, and October of 2013. These months roughly correspond to the seasons of winter, spring, summer, and fall, respectively. We evaluated the model performance by comparing the model results with in-situ observations (for water level and SST) as well as with climatological data from the monthly world ocean database (for SST and SSS).

The remainder of Section 1 introduces background information concerning the setups and operations of the G-RTOFS, ETSS, and ESTOFS as well as background information on the observational data sets used in this study. Section 2 describes the technical details of two methods, namely, the forecast cycle (FC) based method and the forecast hour (FH) based method, used to assess the model performance. Sections 3-5 discuss the model performance of SSH (for G-RTOFS, ETSS, and ESTOFS), SST (for G-RTOFS), and SSS (for G-RTOFS) guidance, respectively. Section 6 summarizes the model assessment results and the methods used to attain those results.

1.1. G-RTOFS, ETSS, and ESTOFS

This section provides an introduction to the model setup and operations of G-RTOFS, ETSS, and ESTOFS.

1.1.1. G-RTOFS

G-RTOFS is based on the Naval Oceanographic Office's (NAVO) configuration of the 1/12^{o} resolution eddy resolving global Hybrid Coordinates Ocean Model (HYCOM) (G-RTOFS, 2011). It is initialized daily with NAVO generated initial conditions using the Navy Coupled Ocean Data Assimilation (NCODA) system (Metzger, et al., 2014). The system assimilates in situ profiles of temperature and salinity from a variety of sources and remotely sensed SST, SSH and sea-ice concentrations. G-RTOFS is forced with 3-hourly momentum, radiation, and precipitation fluxes from NOAA'S operational Global Forecast System (GFS) fields.

The G-RTOFS ocean model has 32 vertical hybrid layers (isopycnal in the deep water, isolevel in the mixed layer, and sigma in shallow water). G-RTOFS has a horizontal grid of dimension 4500×3298. The grid has an Arctic bi-polar patch north of 47^{o}N and a Mercator projection south of 47^{o}N through 78.6^{o}S (Figure 1). The coastline is fixed at 10-m isobaths with the Bering Straits being open. The potential temperature is referenced to 2000 m depth and the first level is fixed at 1 m depth.

G-RTOFS became operational at the NWS NCEP Environmental Modeling Center (EMC) on October 24, 2011. It generates a forecast cycle four times a day at hours 00z, 06z, 12z, and 18z. In this study, we make use of the forecast cycle generated at hour 00z. Each forecast cycle produces, at three hour intervals, forecast output for sea surface values of SSH, SST, and SSS and at six hour intervals, forecast output for full volume parameters including 3-dimensional temperature, salinity, and currents. Each forecast cycle extends from hour 00 to hour 144. The period of operation ran from October 25 2011 through July 25 2013. On this date, the length of the forecast cycle was extended from 144 hours to 192 hours. G-RTOFS has run at NCEP, with this extended forecast cycle, ever since.

Figure 1.1 shows the G-RTOFS horizontal grid. Figures 1.2 (a-c), display snapshots of the G-RTOFS forecast of (a) subtidal water level, (b) SST, and (c) SSS fields in a region along the U.S. east coast and the U.S. Gulf of Mexico coast at 0300 UTC on 1 January 2013.

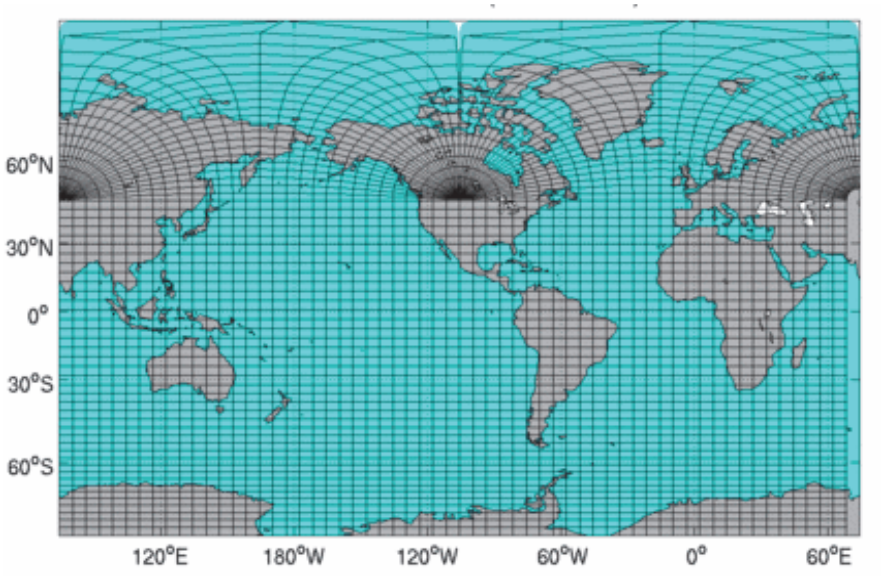
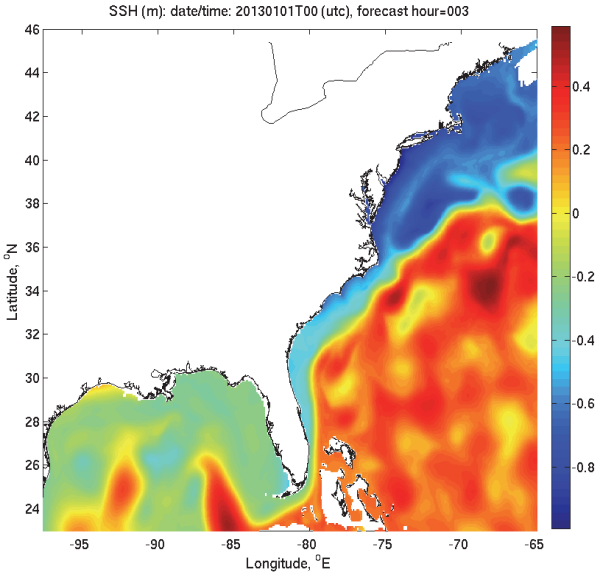
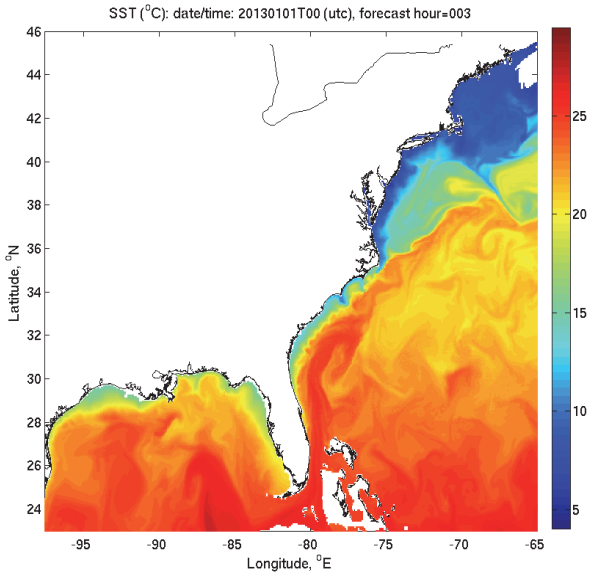


Figure 1.1. The G-RTOFS horizontal grid with each cell representing 54 rows and 75 columns of the entire grid. This grid plot is obtained from http://polar.ncep.noaa.gov/global/about/.

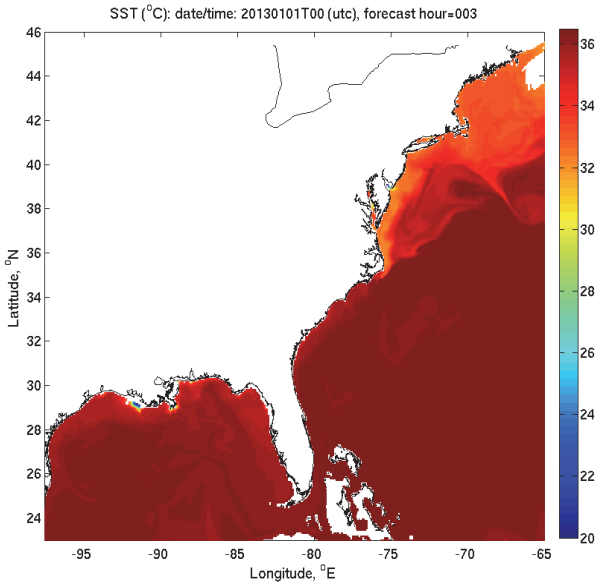


(a)

Figure 1.2. The G-RTOFS forecast of (a) SSH, (b) SST, and (c) SSS at 0300 UTC on 1 January 2013 showing U.S. East Coast and Gulf of Mexico waters.



(b)



(c)

Figure 1.2. (Continued)

1.2.2. ETSS

The ETSS model was developed by the Meteorological Development Laboratory (MDL) of the NWS. It is a variation of the Sea, Lake and Overland Surges from Hurricanes (SLOSH) model developed by NWS (ETSS, 1992). It is a prognostic, two-dimensional, barotropic model forced by real time output of winds and air pressure from the GFS run at NCEP. ETSS has been applied to

the continental shelf and coastal waters of multiple regions off the U.S. coast including Alaska (Arctic Ocean), the west coast (Pacific Ocean), the east coast (Atlantic Ocean), and the Gulf of Mexico. Figure 1.3 displays a map showing ETSS forecast guidance domains located in the Arctic Ocean and the Gulf of Alaska, in the Pacific Ocean off the U.S. west coast, in the Gulf of Mexico, and in the Atlantic Ocean off the southeast, mid-Atlantic, and northeast coasts of the U.S.

ETSS initially ran operationally at NCEP, in their Central Computing System (CCS), one time daily (00z) out to 96 hours producing numerical storm surge guidance (subtidal water levels) for extra-tropical systems. Beginning on August 26, 2012, the ETSS operation was transitioned onto the Weather and Climate Operational Supercomputer Systems (WCOSS).

NCEP disseminates the ETSS forecast guidance of subtidal water levels in both point and gridded formats. We used the point guidance outputs for the present study. Figure 1.4 shows the station locations along the U.S. East Coast and along the Gulf of Mexico. Since this study focuses on evaluating the ETSS performance in open coastal waters, some stations located within estuaries have been discarded. Table A.1 lists the names and geographical locations of the stations adopted for model-data comparisons.

Figure 1.3. Regions of ETSS model domains.

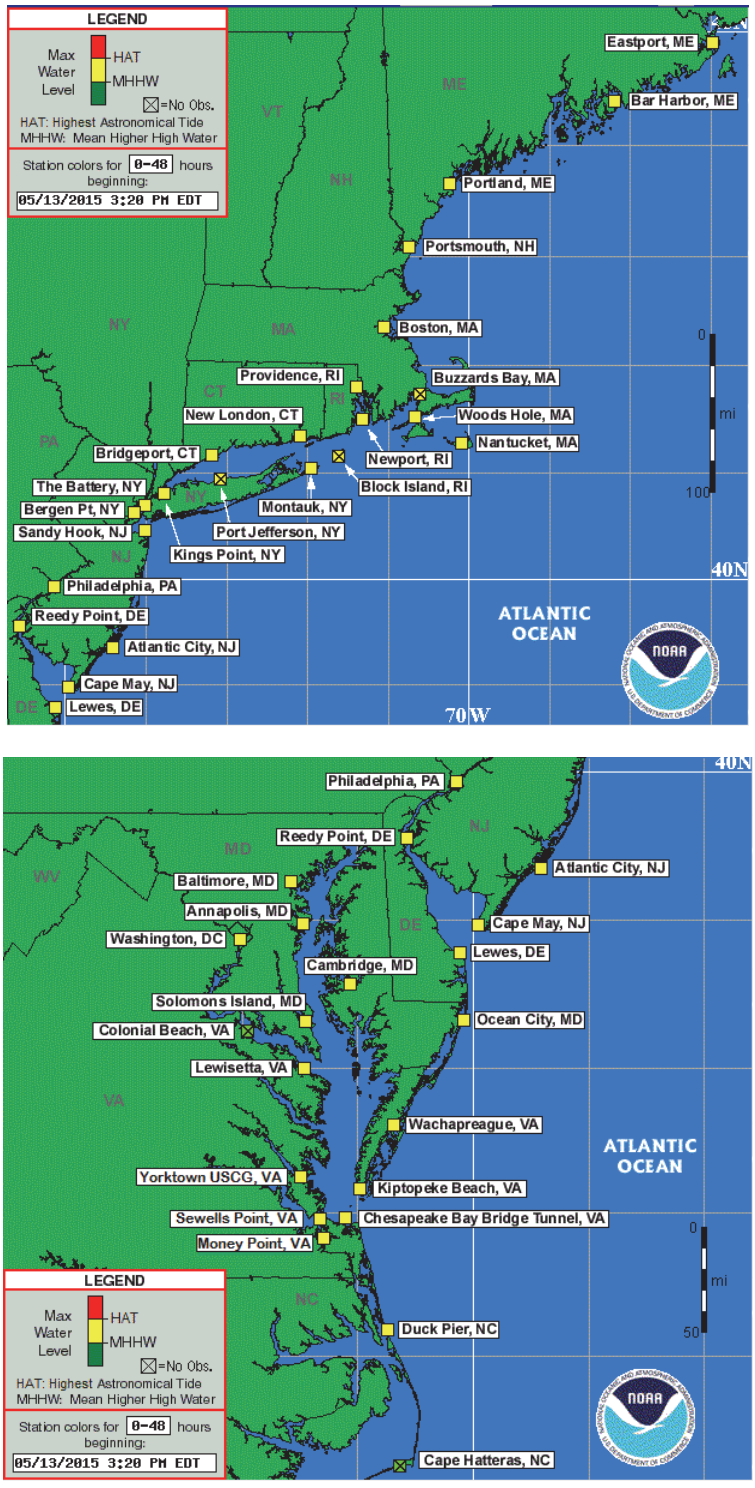


Figure 1.4. Station locations of the ETSS point forecast guidance along the U.S. east coast and Gulf of Mexico. This figure is a screen shot from http://www.nws.noaa.gov/mdl/etsurge/index.php?page=map&region=ne&datum=msl&list=&map=0-48&type=&stn=

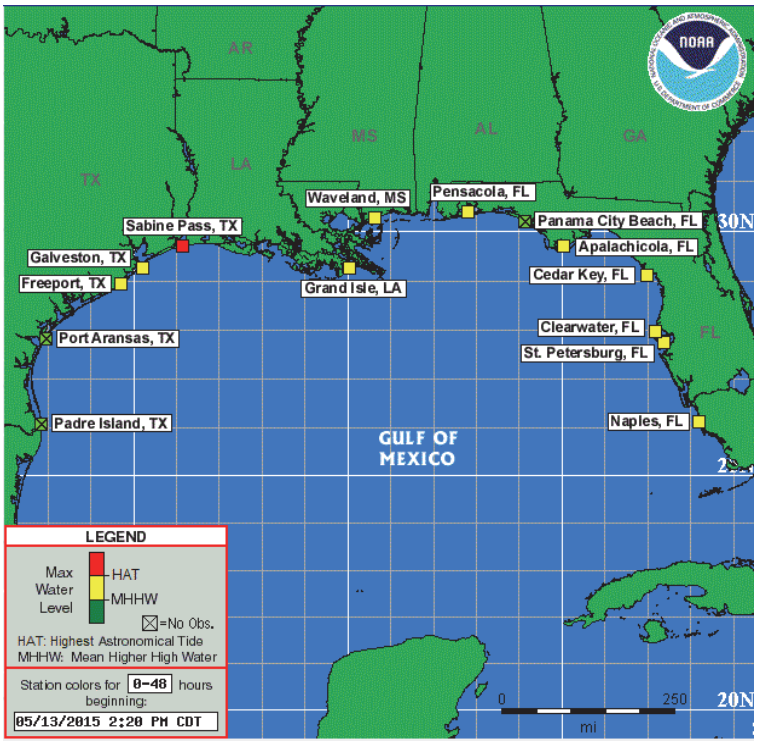


Figure 1.4. (Continued)

1.2.3. ESTOFS

The Coast Survey Development Laboratory (CSDL) of the NOS created, in collaboration with the Environmental Modeling Center (EMC) of NCEP, the ESTOFS for the Western North Atlantic basin. The hydrodynamic model employed by the ESTOFS is the Advanced CIRculation (ADCIRC) finite element model (Luettich et al. 1992; Luettich and Westerink 2004). The ADCIRC model is a two dimensional 2-D barotropic model which has several beneficial features making it suitable for this system. It has been demonstrated to be effective at predicting tidal circulation and storm surge propagation in complex coastal systems. It makes use of an unstructured grid which can readily and accurately represent irregular shorelines including barrier islands, rivers, and waterways.

The ESTOFS was implemented operationally by NCEP to provide forecasts of surge with tides, astronomic tides, and sub-tidal water levels throughout the domain. The system generates four forecast cycles per day and each cycle extends to 180 hours. It is forced with tides from the ADCIRC tidal database, and by wind and air pressure from the GFS fields. The system output is disseminated in the form of either point or gridded format. We made use of the point format output for our study.

Figure 1.5. The ESTOFS horizontal grid. This grid plot is obtained from http://polar.ncep.noaa.gov/global/about/.

1.2. Observed Data

We used observed water level data and sea-surface temperature data collected at water level and physical oceanography observation stations of the Center for Operational Oceanographic Products and Services (CO-OPS) of NOAA. We used SST data from NOAA's National Data Buoy Center (NDBC) buoy measurements. We used SST and SSS data from the 2009 version of the World Ocean Atlas (WOA09) database developed by NODC. The following sections describe in detail the data sources and retrievals.

1.2.1. CO-OPS Water Level

Figures 1.6 (a), (b), and (c) display maps of CO-OPS water level stations at which we evaluated the performance of G-RTOFS, ETSS, and ESTOFS forecasts, respectively. Since G-RTOFS and ETSS both focus on forecasts in the open ocean and coastal areas, we purposefully selected only offshore stations for the model-data comparison and excluded those located in estuaries or embayments.

With regard to observed data, we used quality-controlled hourly water level observations from 47 CO-OPS water level stations for comparison with ETSS (Table A.1). We used 52 CO-OPS water level stations for comparison with G-RTOFS (Table A.2), and we used 54 CO-OPS water level stations for comparison with ESTOFS (Table A.3). We low-pass filtered the total water level time series to remove the tidal components, obtaining a sub-tidal water level time series. We then compared the sub-tidal water level time series with the G-RTOFS, ETSS, and ESTOFS results.

(a)

Figure 1.6. Maps of CO-OPS water level stations used for evaluating performances of (a) ETSS, (b) G-RTOFS, and (c) ESTOFS subtidal water level forecasts.

(b)

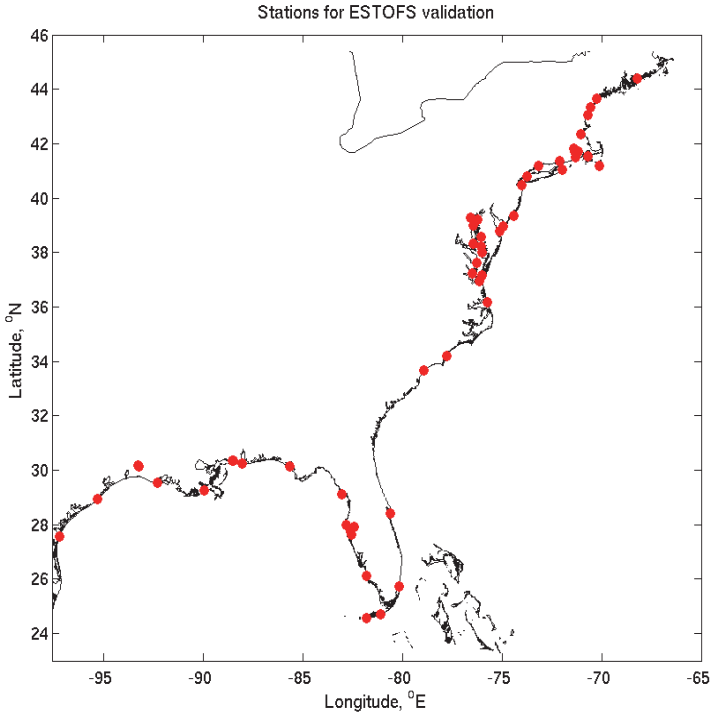


Figure 1.6. (Continued)

1.2.2. CO-OPS SST

We used hourly SST measurements collected at 45 CO-OPS physical oceanography observation stations (Figure 1.7) to evaluate the performance of G-RTOFS SST forecast. Table B.1 lists the station identification (ID) numbers, names, and their geographical locations in longitude and latitude. We downloaded hourly SST time series data from the CO-OPS online database.

1.2.3. NDBC SST

We used SST measurements from 78 NDBC buoys along the U.S. east coast and the Gulf of Mexico coast. Table C.1 lists the buoy IDs, station names, and station locations by longitude and latitude. Figure 1.8 displays a map of these stations. In general, the NDBC stations are located further offshore than are the CO-OPS stations. Hence, the NDBC data are more representative of offshore surface temperature conditions. We downloaded the hourly SST time series data from http://www.ndbc.noaa.gov/data/realtime2.

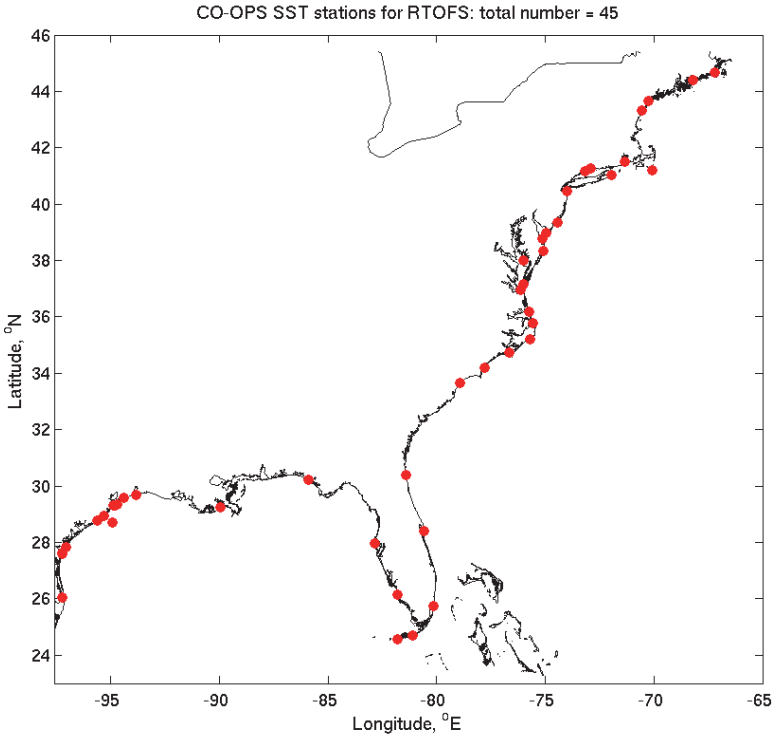


Figure 1.7. The map of the CO-OPS physical oceanography observation stations selected for evaluating the performance of the G-RTOFS SST forecast.

G-RTOFS stations:

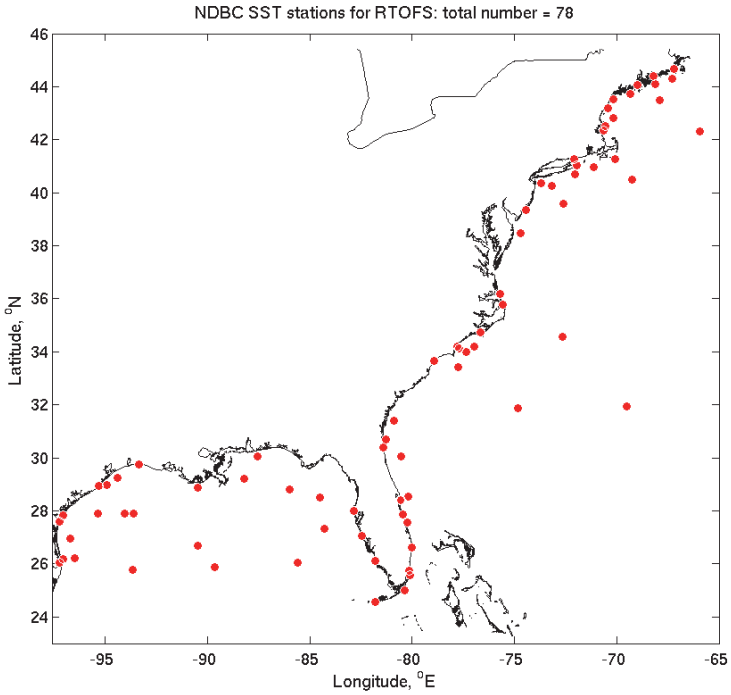


Figure 1.8. The map of NDBC buoy stations selected for evaluating the performance of G-RTOFS SST forecast.

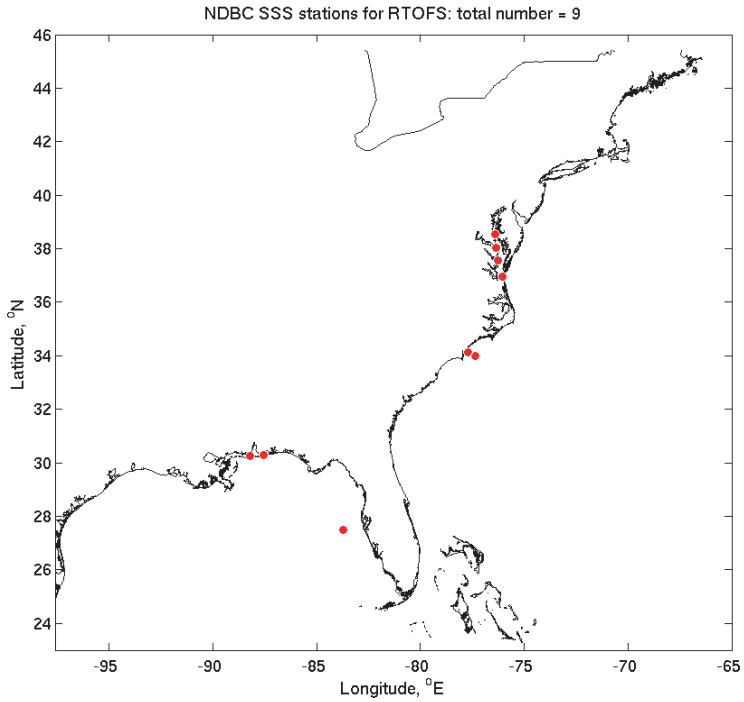


Figure 1.9. The map of NDBC buoy stations selected for evaluating the performance of G-RTOFS SSS forecast.

1.2.4. WOA09 SST and SSS

The World Ocean Atlas 2009 (WOA09) is a set of objectively analyzed (1° grid) climatological fields of in situ temperature, salinity, dissolved oxygen, apparent oxygen utilization (AOU), percent oxygen saturation, phosphate, silicate, and nitrate at standard depth levels for annual, seasonal, and monthly compositing periods for the World Ocean. It also includes associated statistical fields of observed oceanographic profile data interpolated to standard depth levels on both 1° and 5° grids. We downloaded the data set contained in a NetCDF file from the NDBC Web site: http://www.nodc.noaa.gov/OC5/WOA09F/pr_woa09f.html.

In the WOA09 database, quantities defined on each grid node represent a composite of observations collected in nearby areas. In the present study, we treat each grid node as an in situ observing station. We used SSS values from 9 coastal grid points (Table D.1) in WOA09 to assess the agreement of the G-RTOFS forecast to climatology (Figure 1.9).

2. METHODS OF MODEL SKILL ASSESSMENT

We evaluated the performance of G-RTOFS forecast guidance on subtidal water level (SSH), sea-surface temperature (SST), and sea-surface salinity (SSS) as well as the performance of the ETSS (SSH) forecast guidance. We assessed the model performance from two perspectives: forecast cycle-based (FC) method and forecast hour-based (FH) method. The former evaluates the performance over the forecast cycle as a whole, whereas the latter evaluates the performance at individual hours within the forecast cycle.

G-RTOFS generates four forecast cycles, one every six hours (00z, 06z, 12z, 18z), per day. Each cycle produces forecast guidance at 3 hour intervals from hour 00 to hour 144 or hour 192 (Section 1.2). Since all four cycles utilize the same model configuration (grid) and forcing data of the same quality, it is reasonable to believe that the model will perform equally well across all four cycles. For this reason, we focus on evaluating the model performance of the 00 UTC cycle. In the present study, we focus on the forecast through 144 hours. Hence, every forecast cycle produces a 49 point time series of SSH, SST, and SSS corresponding to hours 00, 03, 06, …, 144.

The ETSS water level forecast is generated once a day at 00 UTC. This daily forecast cycle produces an hourly forecast from hour 01 through hour 96. So each forecast cycle produces a 96 point time series of SSH corresponding to hours 01, 02, 03, …, 96.

ESTOFS generates four forecast cycles at hours 00z, 06z, 12z, and 18z each day. Each cycle produces hourly forecast guidance from hour 01 through hour 180. So each forecast cycle produces a 180 point time series of SSH corresponding to hours 01, 02, 03, … 180.

For each variable (SSH, SST, and SSS) at a given station, we first compile two coincident time series. The first is from the model output and the second is from the observed data. Data points in the two time series correspond to the same model forecast guidance hour. For each pair of values, model and observed, we subtract the observed value from the model value to get a time series representing the model-data difference. We then apply the FC and FH methods to this new time series to assess the model performance.

Sections 2.2 and 2.3 describe details of the FC and FH methods, respectively. Since the methods used for SSH, SST, and SSS are the same, in the following sections we will designate the variable names only if necessary.

2.1. Model and Data Time Series

To compare the forecast with observations, we first identify the model grid cells located closest to the observed data stations. We create the model-data difference time series by subtracting each observed data value from the corresponding model value. Since we utilize one forecast cycle, the forecast cycle generated at hour 00 UTC, we end up with one time series for each day. For G-RTOFS, each time series consists of 49 data points corresponding to the three-hourly forecast guidance from hours 00 through 144 UTC. For ETSS, each time series consists of 96 data points corresponding to the hourly forecast guidance from hours 01 through 96. For ESTOFS, each time series consists of 180 data points corresponding to the hourly forecast guidance from hours 01

through 180. In a given month, this results in a 30 (for April) or 31 (for January, July, and October) point time-series which forms the base-line data set for all data analysis.

Figures 2.1-2.6 display sample model and data time series. Figure 2.1 depicts the ETSS and CO-OPS subtidal water levels. Figure 2.2 depicts the G-RTOFS and CO-OPS subtidal water levels. Figure 2.3 depicts the ESTOFS and CO-OPS subtidal water levels. Figure 2.4 depicts the G-RTOFS and CO-OPS SST. Figure 2.5 depicts the G-RTOFS and NDBC buoy SST. Finally, Figures 2.6 and 2.7 display the G-RTOFS and WOA09 SST and SSS, respectively.

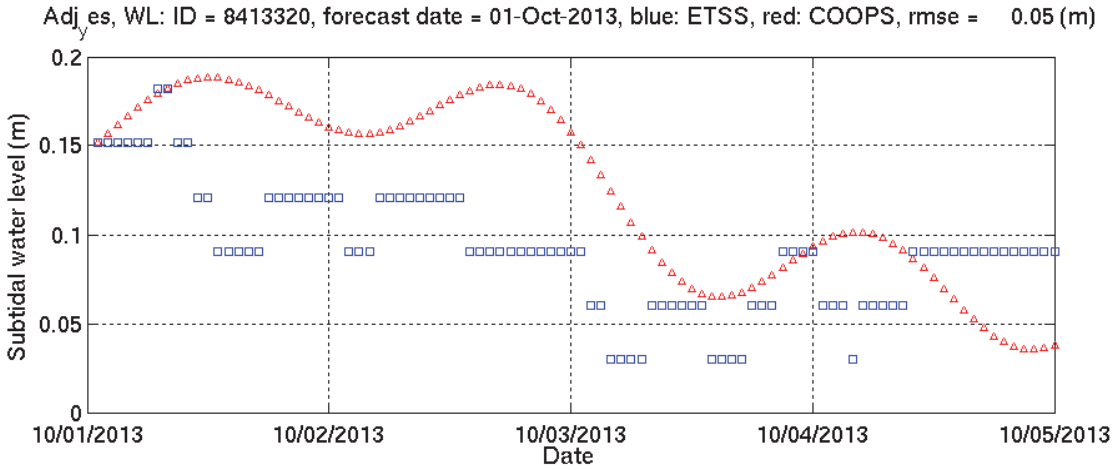


Figure 2.1. Subtidal water level time series of ETSS (blue) and CO-OPS (red) at CO-OPS water level station 8413320 during the ETSS forecast cycle on 10/01/2013.

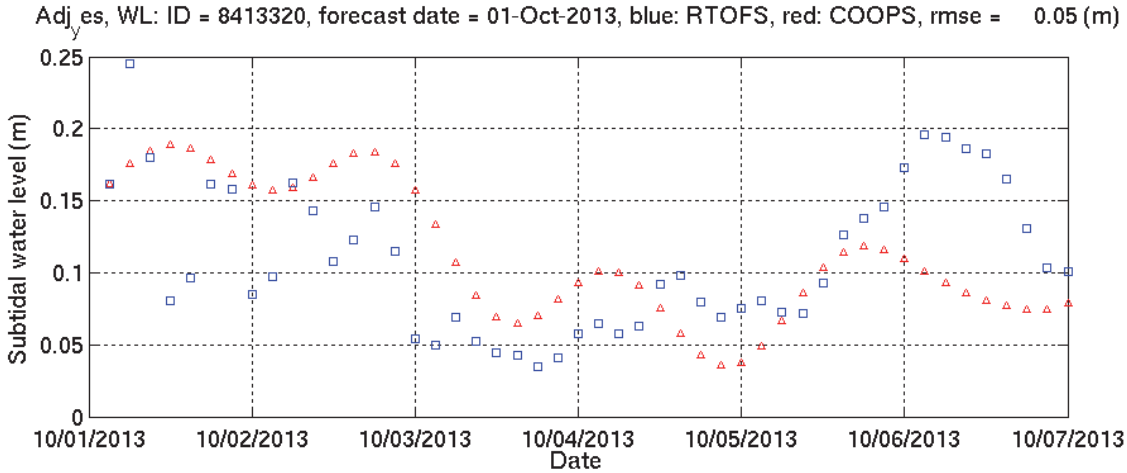


Figure 2.2. Subtidal water level time series of G-RTOFS (blue) and CO-OPS (red) at water level station 8413320 during the G-RTOFS forecast cycle on 10/01/2013.

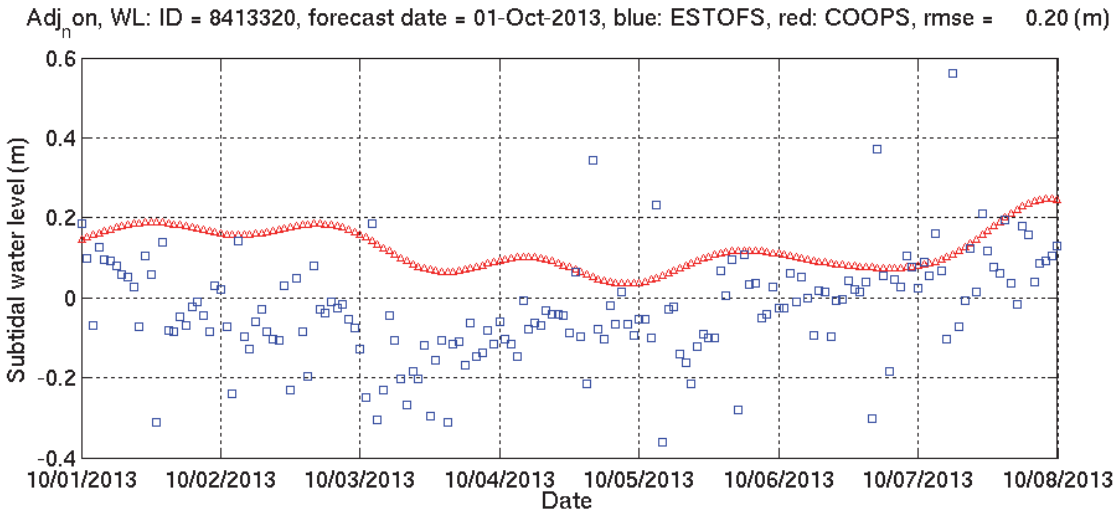


Figure 2.3. Subtidal water level time series of ESTOFS (blue) and CO-OPS (red) at water level station 8413320 during the ESTOFS forecast cycle on 10/01/2013.

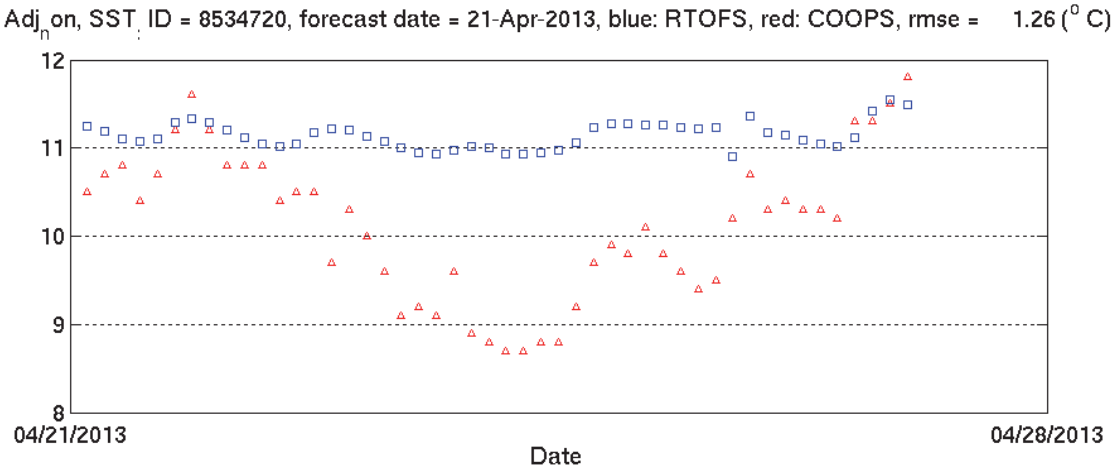


Figure 2.4. Sea-surface temperature time series of G-RTOFS (blue) and CO-OPS (red) at the CO-OPS station 9415020 during the G-RTOFS forecast cycle 00z UTC 4/21/2013.

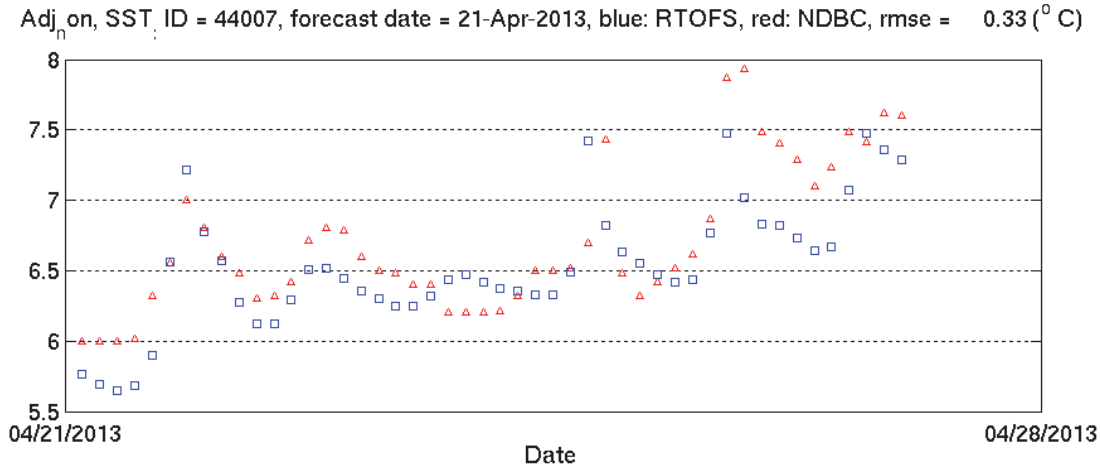


Figure 2.5. Sea-surface temperature time series of G-RTOFS (blue) and NDBC buoy 46247 (red) during the G-RTOFS forecast cycle 00z UTC 4/21/2013.

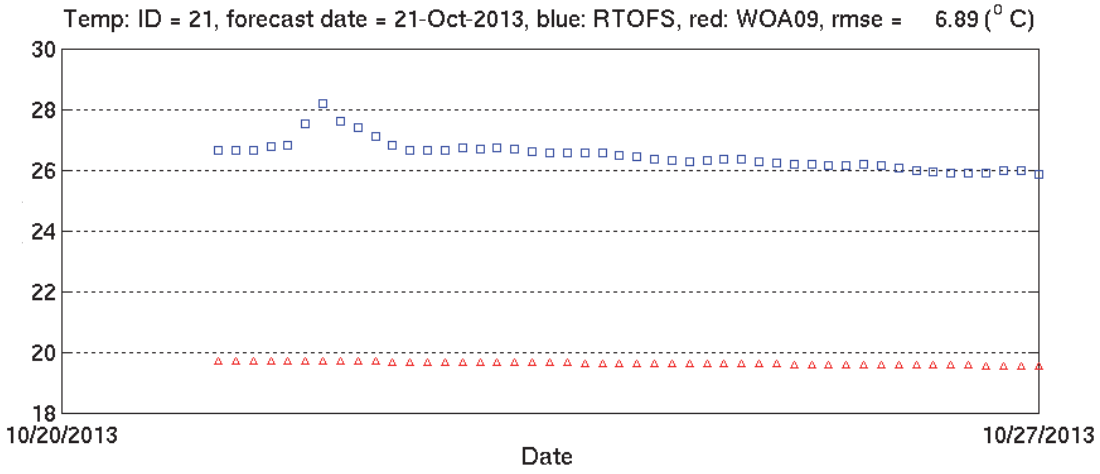


Figure 2.6. Sea-surface temperature time series of G-RTOFS (blue) and WOA09 (at NDBC SST station 21) (red) (Table C.1 in Appendix C) during the G-RTOFS forecast cycle on 10/21/2013.

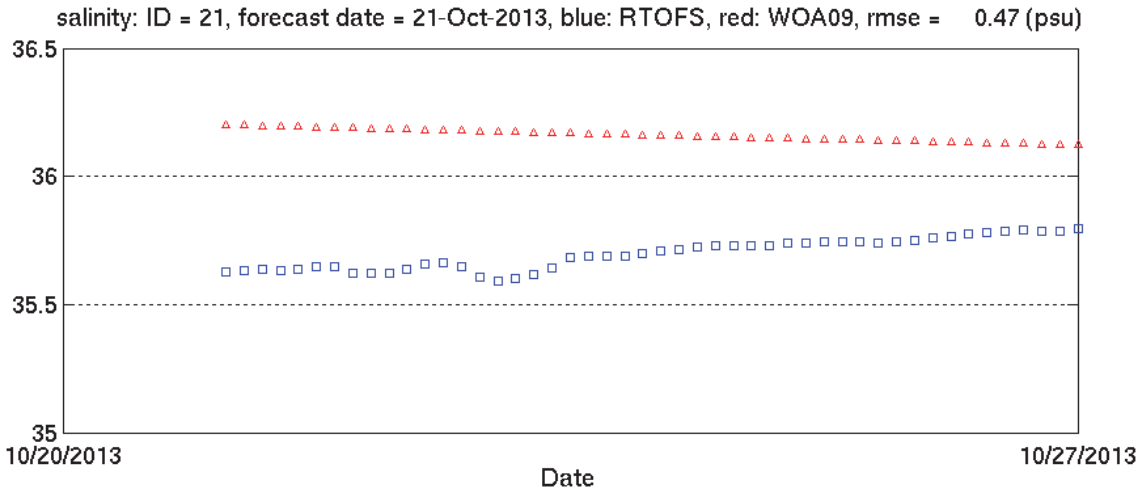


Figure 2.7. Sea-surface salinity time series of G-RTOFS (blue) and WOA09 Station 21 (red) (Table D.1 in Appendix D) during the G-RTOFS forecast cycle on 10/21/2013.

2.2. Forecast cycle (FC) based method

The Forecast cycle based method evaluates the model performance across the entire forecast cycle. The cycle length for G-RTOFS is 144 hours, for ETSS it is 96 hours, and for ESTOFS it is 180 hours. From each model-data difference time series (Section 2.1), we calculate the mean, standard deviation (STD), and the root-mean-square error (RMSE).

One model-data time series is created for each forecast cycle (Section 2.1) at each station. Since one forecast cycle is utilized per day, either 30 or 31 model-data difference time series are created for each month. For each of these time series, we calculate the daily mean, STD, and RMSE. Hence, we obtain 30 or 31 values for each statistical parameter for each month. Taking the average of these values, for each parameter, gives us a monthly mean value of bias, STD, and RMSE for that station. To obtain the monthly mean values which appear in Table 3.1, we then average over all stations.

2.3. Forecast hour (FH) based Method

The Forecast hour based method evaluates the model performance at a given forecast hour across multiple forecast cycles. This method reveals how the model skill evolves with the increasing of the forecast guidance hour.

Figure 2.8 illustrates how this method organizes the model and observed data to estimate the model error statistics. As described in Section 2.1, we compile either 30 or 31 model-data pairs for each given month, with each pair corresponding to a time series from one forecast cycle. This essentially forms a two-dimensional array; one dimension is the forecast guidance hour and the other is the daily forecast cycle. The FH method focuses on contrasting model-data discrepancies at each forecast guidance hour. We therefore subset the 2-D array into multiple 1-D time series according

to the forecast guidance hour, i.e., each 1-D series consists of data points corresponding to the same forecast hour, but from multiple forecast cycles. The number of pairs of data points in each time series is either 30 or 31 depending on the number of days (i.e., forecast cycles) in the month under consideration.

The FH based method was created to investigate the evolution of model skill with the increasing of forecast hour. To characterize the trend of variation, we offset the model time series SSH by the amount of the first data point minus the first model point. In essence, we adjust the initial forecast point to match the corresponding observed data point. We then apply this same offset to all remaining points in the forecast cycle.

Figures 2.9- 12 display the forecast hour derived time series. Specifically, Figure 2.9 depicts the G-RTOFS and CO-OPS subtidal water levels. Figure 2.10 depicts the G-RTOFS and CO-OPS SST. Figure 2.11 depicts the ETSS and CO-OPS SSH. Figure 2.12 displays the ESTOFS and CO-OPS SSH.

Following the same approach as was taken with the FC method, for each forecast hour, we calculate the mean, STD, and RMSE of model-data differences at each station. We then average each quantity over all stations to derive station-averaged model statistics.

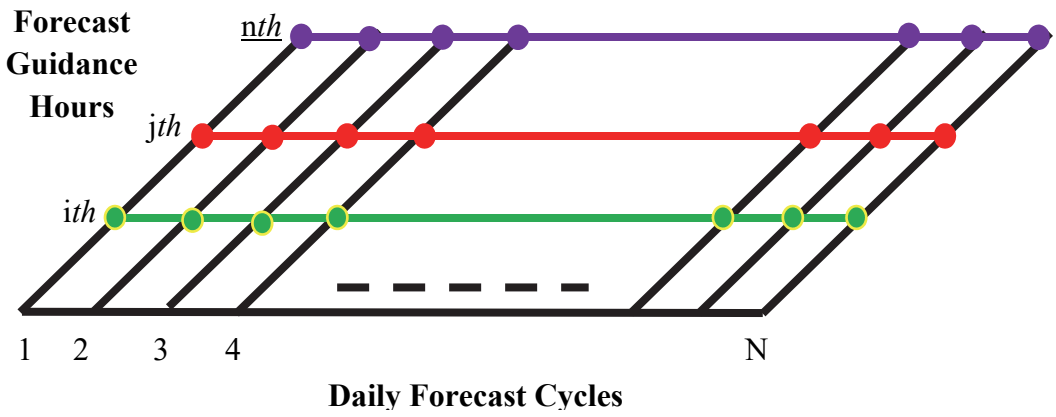


Figure 2.8 – Data structure of the two-dimensional model-data time series array.

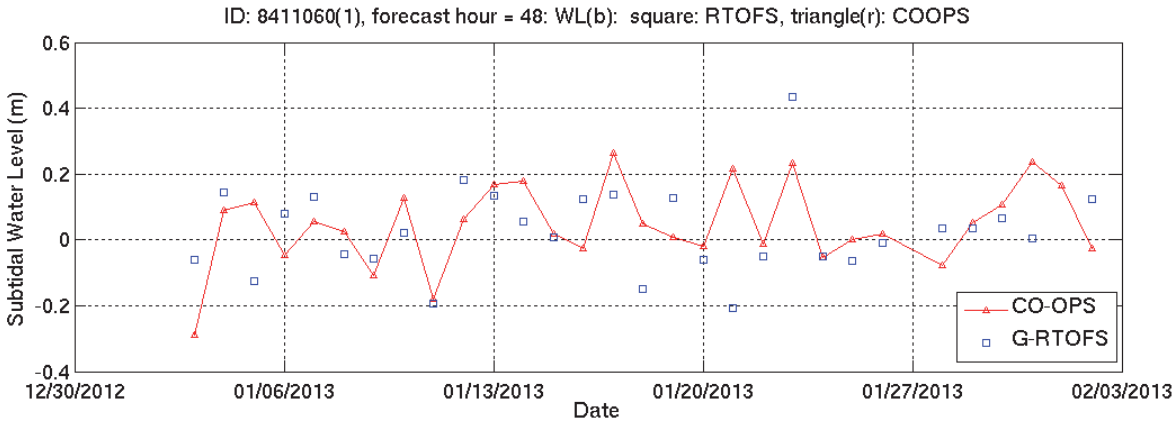


Figure 2.9. Forecast hour based daily sub-tidal water level time series of G-RTOFS (blue) and CO-OPS observations (red) from 12/30/2012 to 02/03/2013.

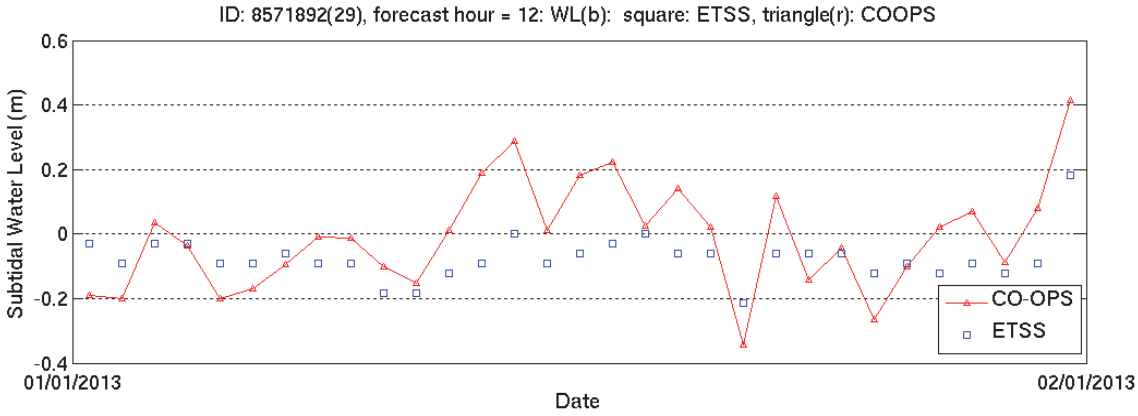


Figure 2.10. Forecast hour based daily SSH time series of ETSS (blue) and CO-OPS observations (red) from 9/28/2012 to 11/12/2012.

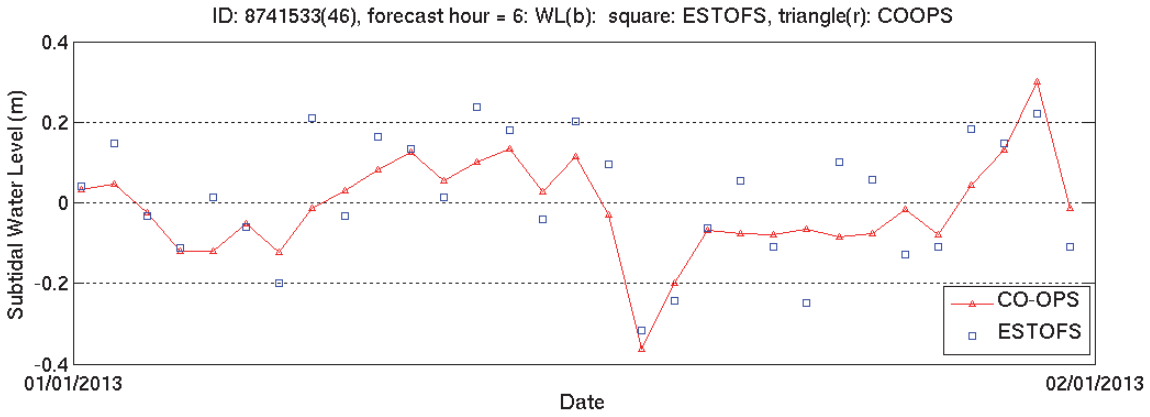


Figure 2.11. Forecast hour based daily SSH time series of ESTOFS (blue) and CO-OPS observations (red) from 1/01/2013 to 2/01/2013.

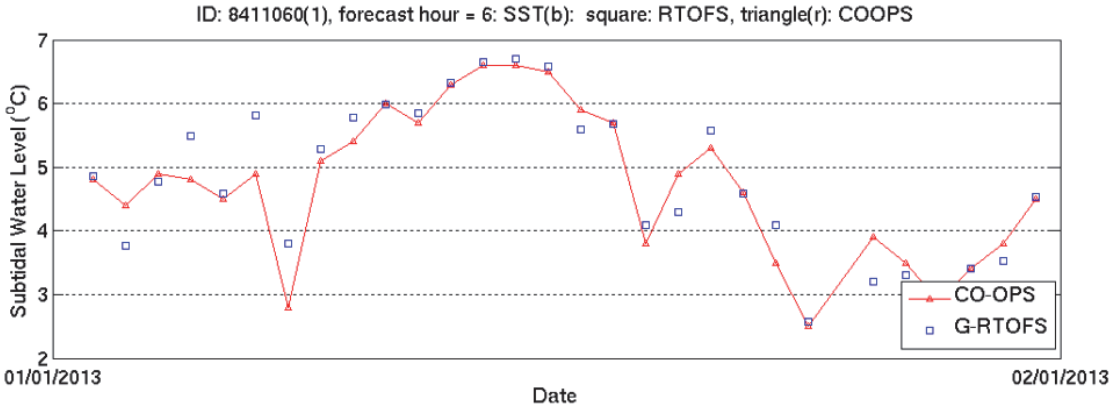


Figure 2.12. Forecast hour based daily SST time series of G-RTOFS (blue) and CO-OPS observations (red) from 9/28/2012 to 11/12/2012.

3. PERFORMANCE OF WATER LEVEL FORECASTS

In this chapter, we evaluate the performance of G-RTOFS, ETSS, and ESTOFS for SSH. We apply both the forecast cycle (FC) method and the forecast hour (FH) method in the analysis.

3.1. G-RTOFS Forecast

In this session, we compare the G-RTOFS water level forecasts with measurements at 52 CO-OPS water level stations. Table A.1 lists the station IDs, names, and geographical locations in longitude and latitude.

3.1.1. FC Based Assessment

Figures 3.1(a)-(d) display model RMSE maps for January, April, July, and October of 2013. The RMSE values are greatest during April, and are relatively smaller for the remaining months. The RMSE ranges between 30 and 70 cm in January 2013, 35 and 80 cm in April 2013, 25 and 75 cm in July 2013, and between 30 and 65 cm in October 2013. G-RTOFS performed worst in the spring. As can be seen in Table 3.1, April 2013 has a station averaged RMSE of 55.6 cm. In addition, April 2013 has the largest number of stations with RMSE values which approach 80 cm.

The maximum RMSE by season is 70 cm, 80 cm, 75 cm, and 65 cm which occur in January 2013, April 2013, July 2013, and October 2013, respectively. The RMSE demonstrates a distinct spatial disparity. It varies in a 40-cm range in January, a 45 cm range in April, a 50-cm range during July, and a 35-cm range in October. Throughout the seasons of the year, the Gulf of Mexico stations appear to outperform the stations of the mid-Atlantic and of the northeast coast.

Figures 3.2(a)-(d) display the mean bias and STD of model errors at 52 CO-OPS water level stations in January 2013, April 2013, July 2013, and October 2013. There is a very large negative bias throughout all four seasons. This negative bias appears largest in spring, with a station averaged value of -53.6 cm (Table 3.1), and is relatively smaller for the remaining months. This large negative bias throughout the region, especially in the Gulf of Mexico and along the southern U.S. Atlantic coast, demonstrates the datum problem with the G-RTOFS model/system. This datum issue stems from the fact that G-RTOFS is a global model and that the model domain is the entire ocean surface of the world. In all four seasons, there is one station (Lewisetta, 8635750) in the Chesapeake Bay with a large positive bias ranging between 20 and 35 cm. This one station is a statistical outlier and the elevated values of mean difference should not be considered valid.

The range of STD is moderate, from around 5 to 11 cm, with the largest deviations occurring in January (10.2 cm). Standard deviation values appear to be smallest during July (5.6 cm), and they are small during October as well (7.4 cm). Typically, January is subject to more stormy weather while July tends to more placid. The breaks in plot continuity are due to observed data gaps occurring at several stations during these months.

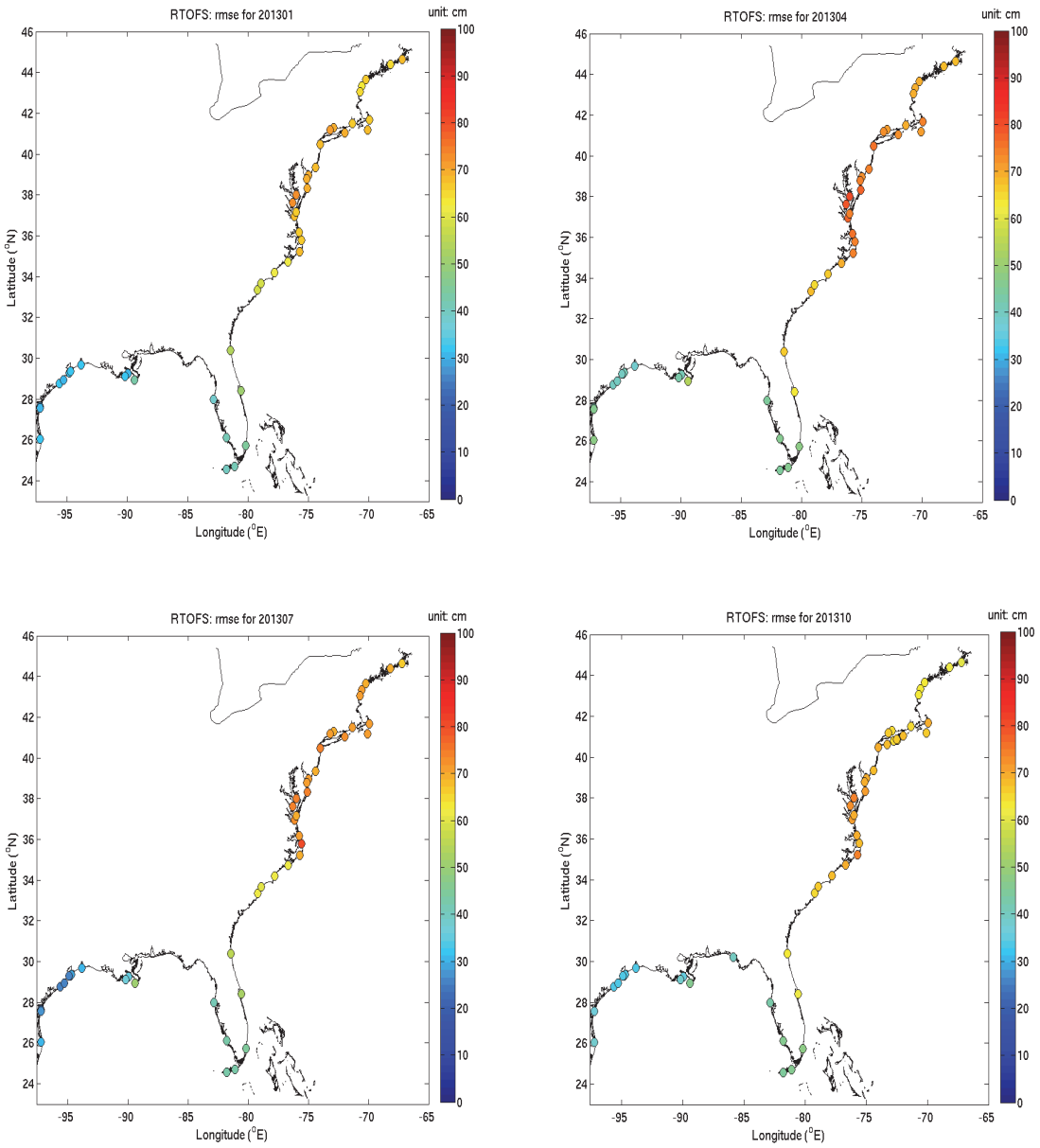
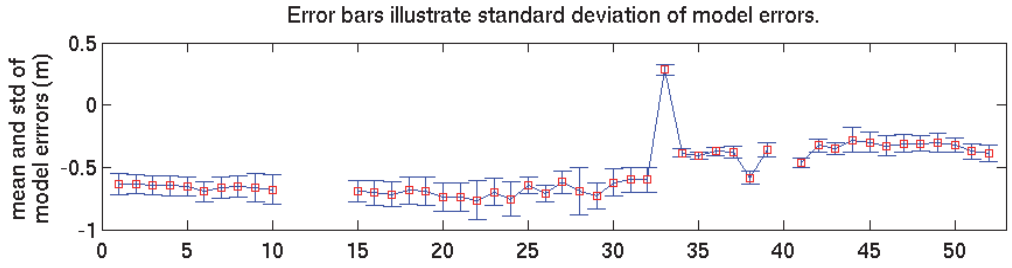


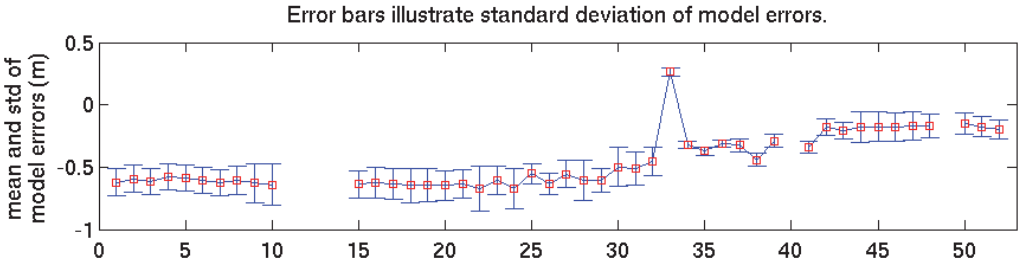
Figure 3.1. Color coded RMSE maps of G-RTOFS water level forecasts in (a) January 2013, (b) April 2013, (c) July 2013, and (d) October 2013 at 52 CO-OPS water level stations (Table A.1).

01/2013



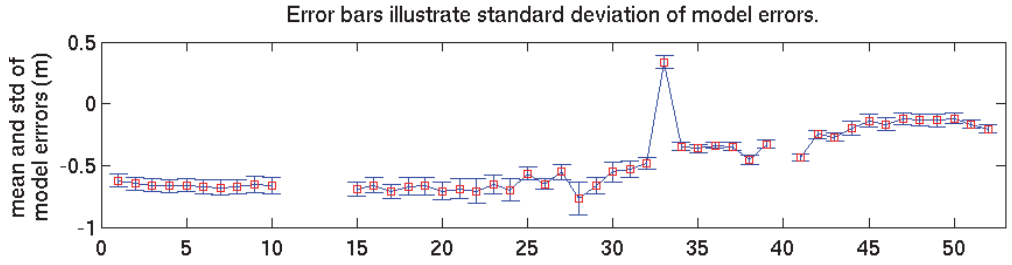
(a)

04/2013



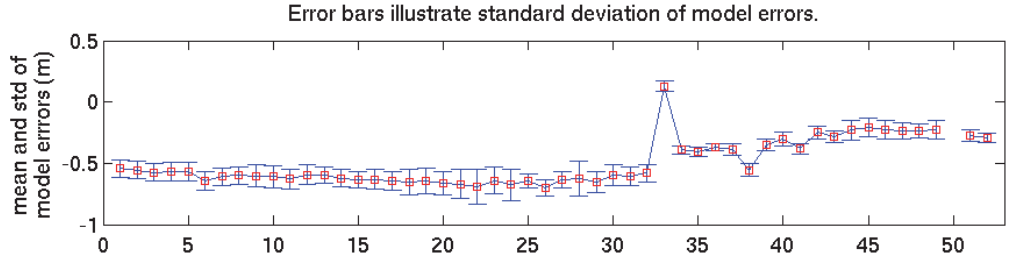
(b)

07/2013



(c)

10/2013



(d)

Figure 3.2. Mean (red squares) and standard deviation (blue error bars) of the G-RTOFS water level errors for 2013 at 52 CO-OPS water level stations from: (a) January 2013, (b) April 2013, (c) July 2013, and (d) October 2013. The horizontal axis shows the station numbers (Table A.1).

3.1.2. Forecast Hour Based Method

In this section, we apply the FH method to investigate the G-RTOFS performance for SSH at individual hours within the forecast cycle.

Station-Averaged RMSE Figure 3.3 displays station-averaged RMSE at each forecast hour in (a) January 2013, (b) April 2013, (c) July 2013, and (d) October 2013. In all four months, the RMSE increases rapidly as the forecast hour extends and then, upon reaching a transition point, the rate of deterioration decreases. When that transition hour occurs varies by month.

The ranking of months with respect to G-RTOFS RMSE (from best to worst, high RMSE value) performance is as follows: July, October, April, and then January. The corresponding RMSE was around 12 cm, 14 cm, 16 cm, and 18 cm, respectively. The transition point occurred around hours 80, 40, 55, and 25, respectively.

RMSE of Individual Stations Above, we discussed station-averaged model performance. The calculations reflect the overall model skill across the entire U.S. east coast and Gulf of Mexico regions. We now assess the model skill at the individual stations. Figures 3.4 – 3.7 display station maps of water level RMSE for (a) January 2013, (b) April 2013, (c) July 2013, and (d) October 2013. Each figure consists of six plots which correspond to forecast hours 6, 24, 48, 72, 96, and 144.

In general, model results demonstrate favorable agreement with observations within the first six hours of a forecast cycle. In all four months, RMSE values were typically less than 12 cm at hour 6. From there, the model performance deteriorates markedly by forecast hour 24.

As the forecast hour increases, the skill decreases. However, the degree of deterioration varies by month. The G-RTOFS performance was least favorable in January with many stations at 25 cm or above by hour 144. January was followed by April, then October. G-RTOFS performance was most favorable in July with only a couple of stations going over 20 cm RMSE by hour 144.

The G-RTOFS forecasts demonstrate spatial variability. In general, the forecasts demonstrate better skill at the stations located on the Gulf of Mexico (especially the southern Florida stations), the east coast south of Virginia, and at stations located to the northeast of Long Island Sound.

It is worth noting that the model skill demonstrates the least spatial and temporal variability in the summer season; see Figure 3.13 for the July 2013 result. The RMSE is uniform across the region at hour 6 and there is only slight variability by hour 24. There is one station in the mid-Atlantic region (Oregon Inlet Marina, NC, 8652587) with a problematically high RMSE at all hours throughout the forecast cycle.

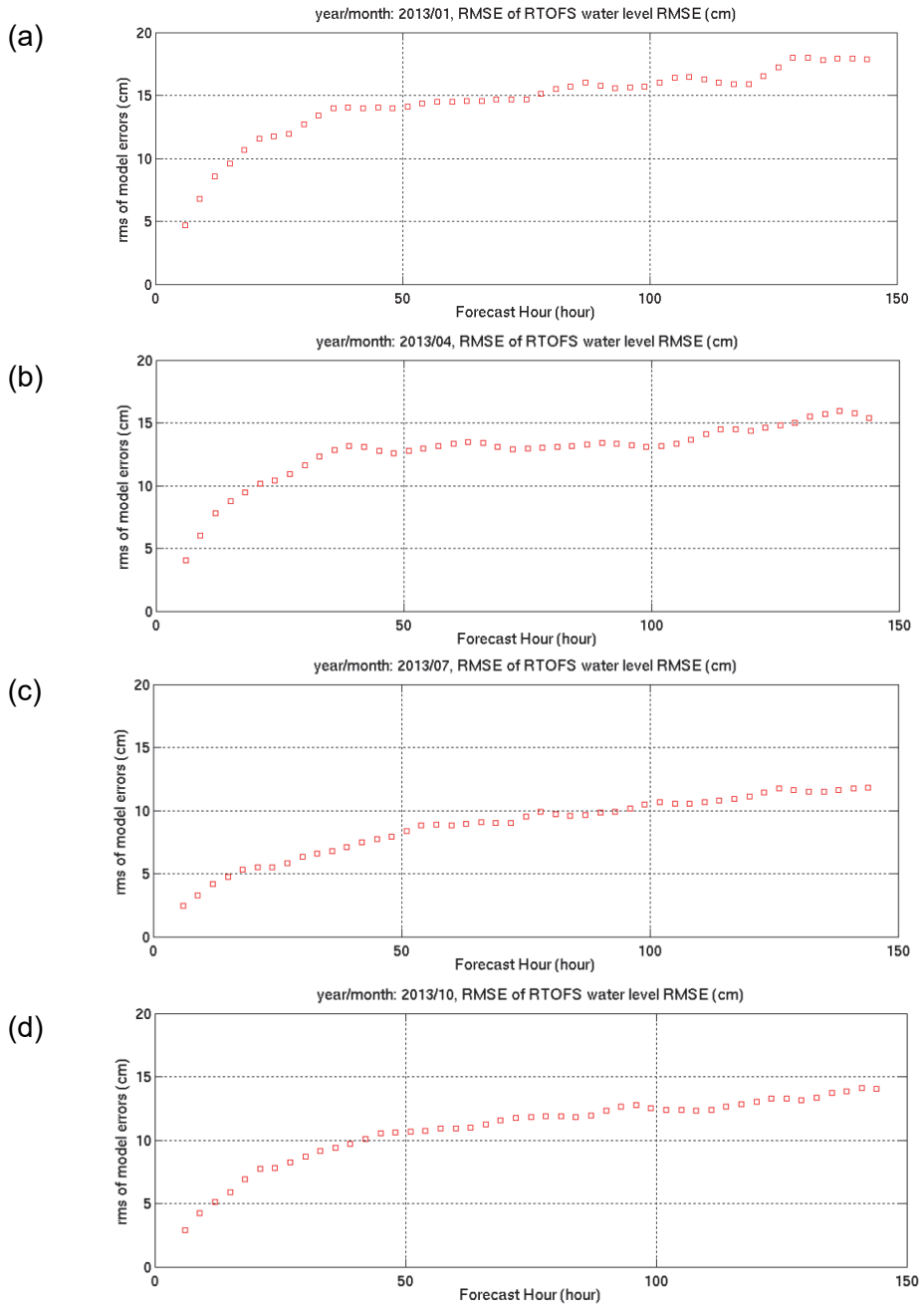


Figure 3.3. Station-averaged root mean squared errors of G-RTOFS water level forecasts in (a) January 2013, (b) April 2013, (c) July 2013, and (d) October 2013.

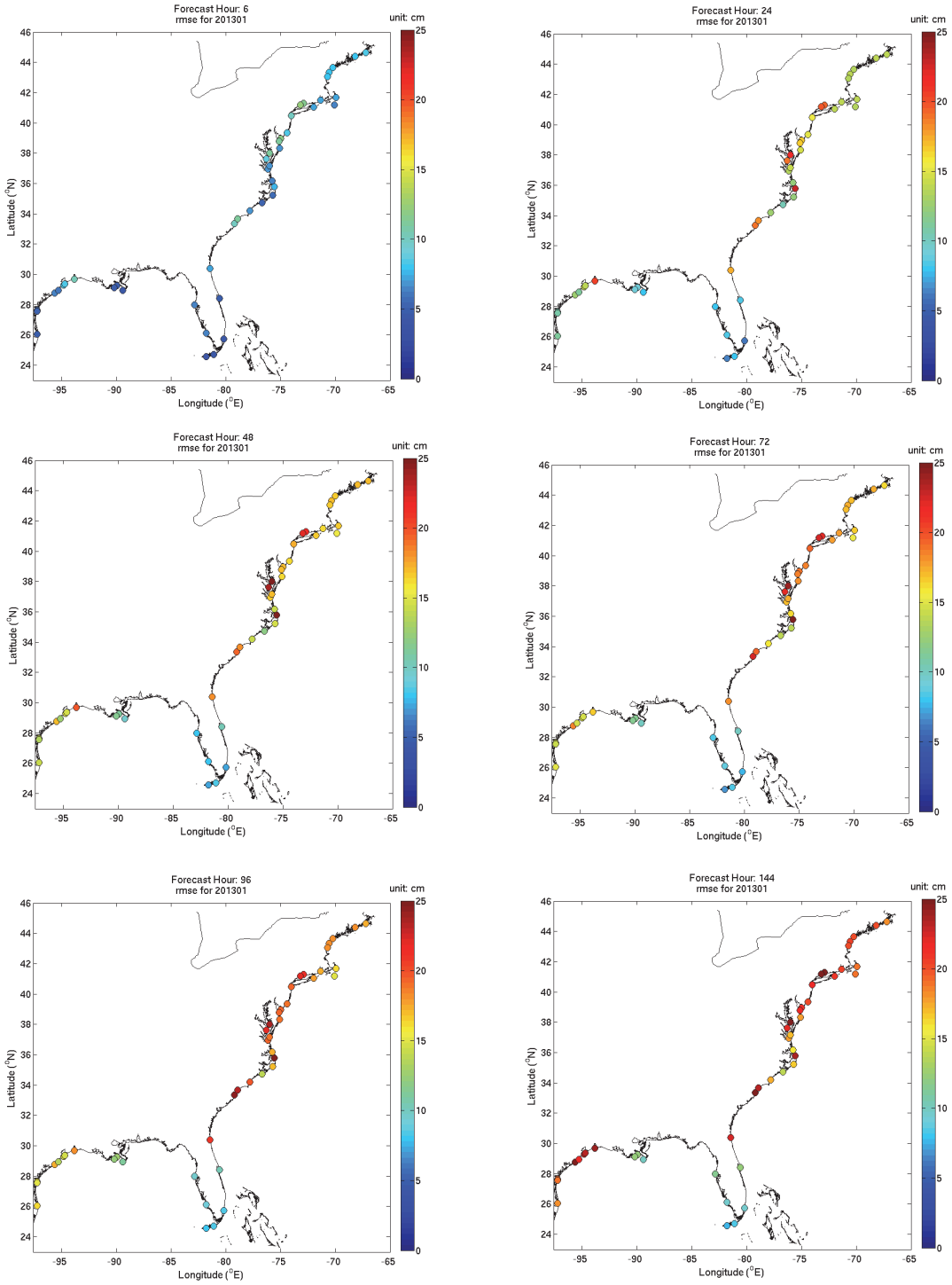


Figure 3.4. Color coded RMSE maps of G-RTOFS water level forecasts in January 2013. The six plots correspond to forecast guidance hours (a) 6, (b) 24, (c) 48, (d) 72 (e) 96 and (f) 144, respectively.

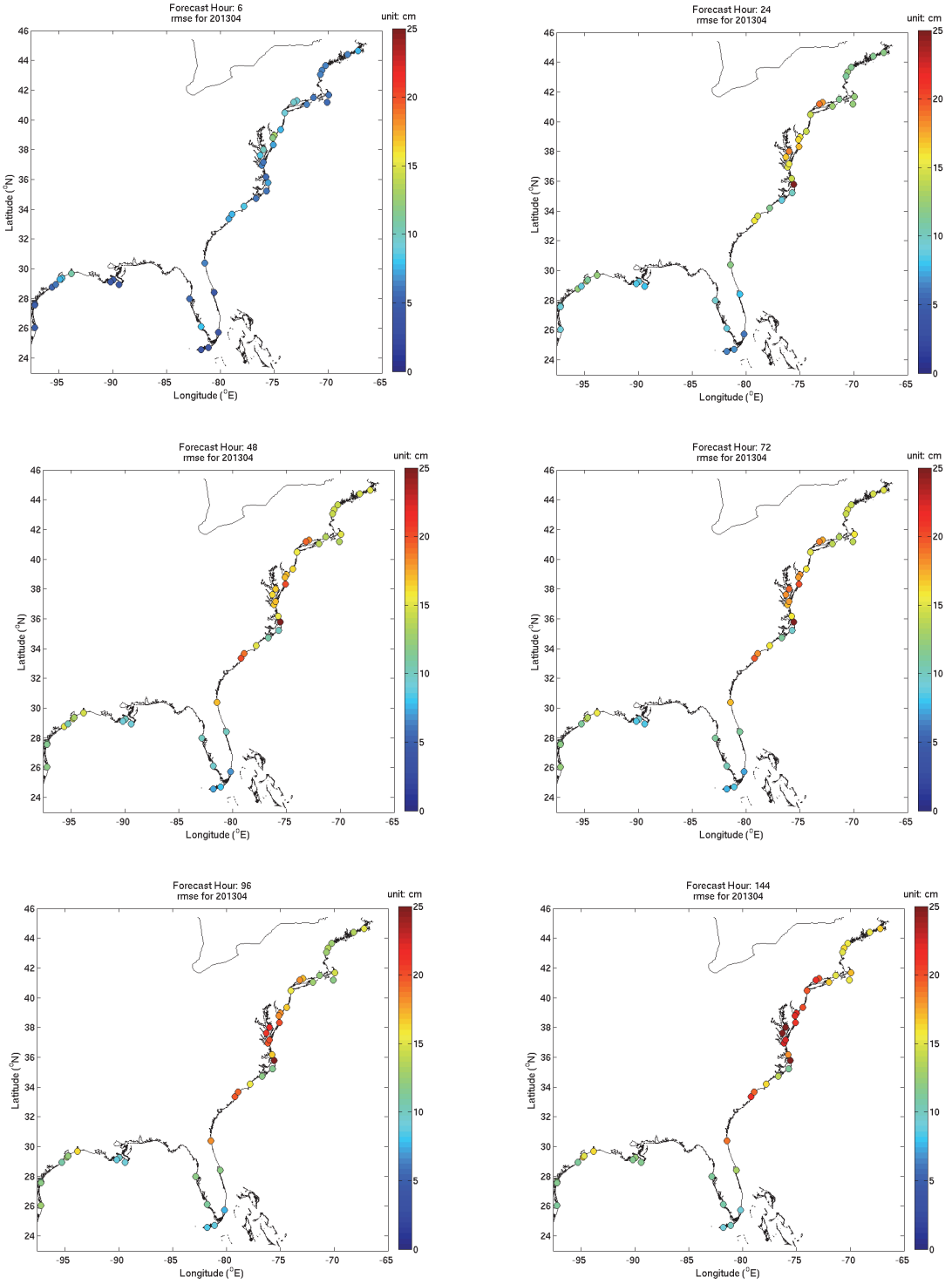


Figure 3.5. Color coded RMSE maps of G-RTOFS water level forecasts in April 2013. The six plots correspond to forecast guidance hours (a) 6, (b) 24, (c) 48, (d) 72, (e) 96, and (f) 144, respectively.

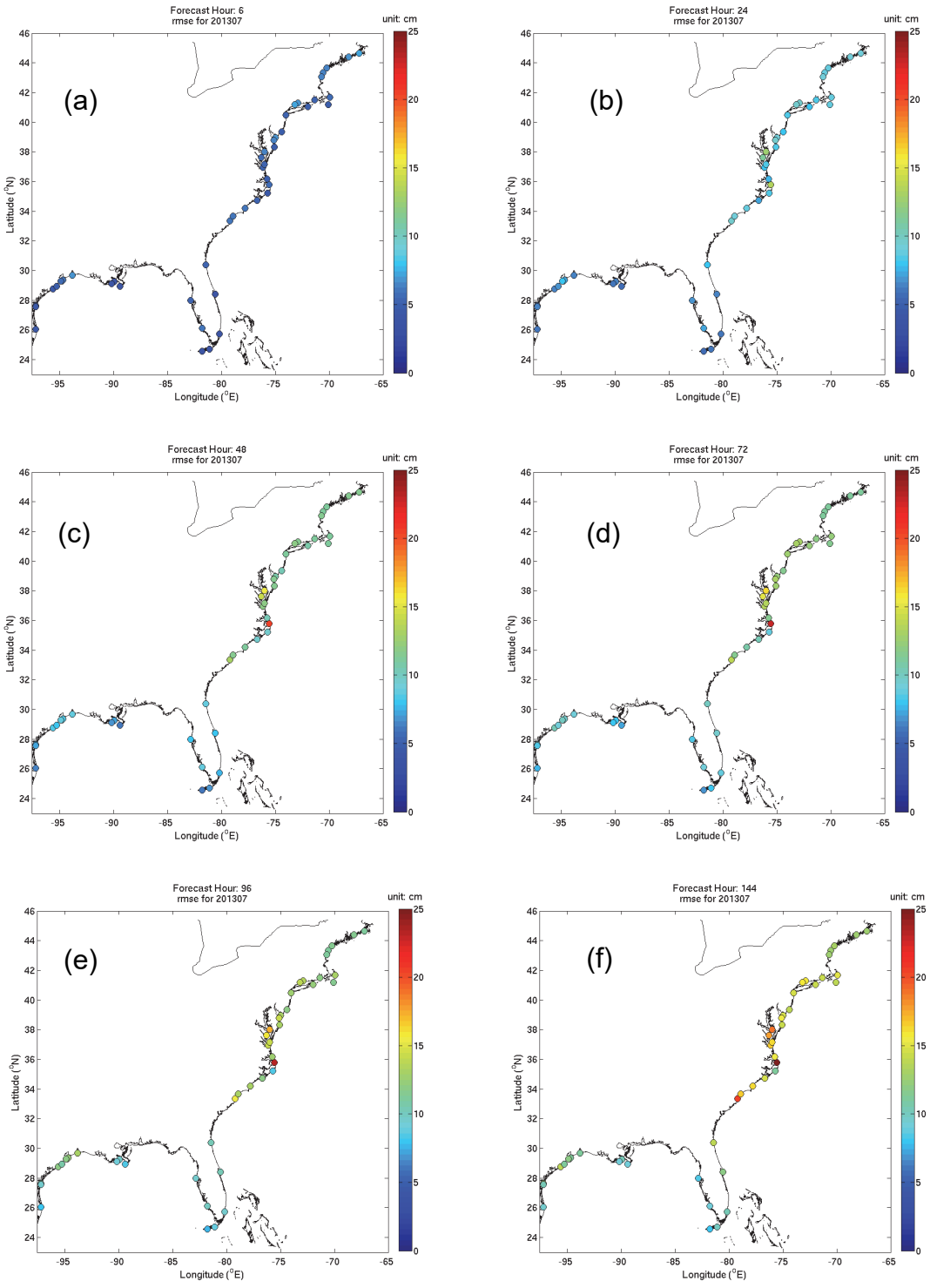


Figure 3.6. Color coded RMSE maps of G-RTOFS water level forecasts in July 2013. The six plots correspond to forecast guidance hours (a) 6, (b) 24, (c) 48, (d) 72, (e) 96, and (f) 144, respectively.

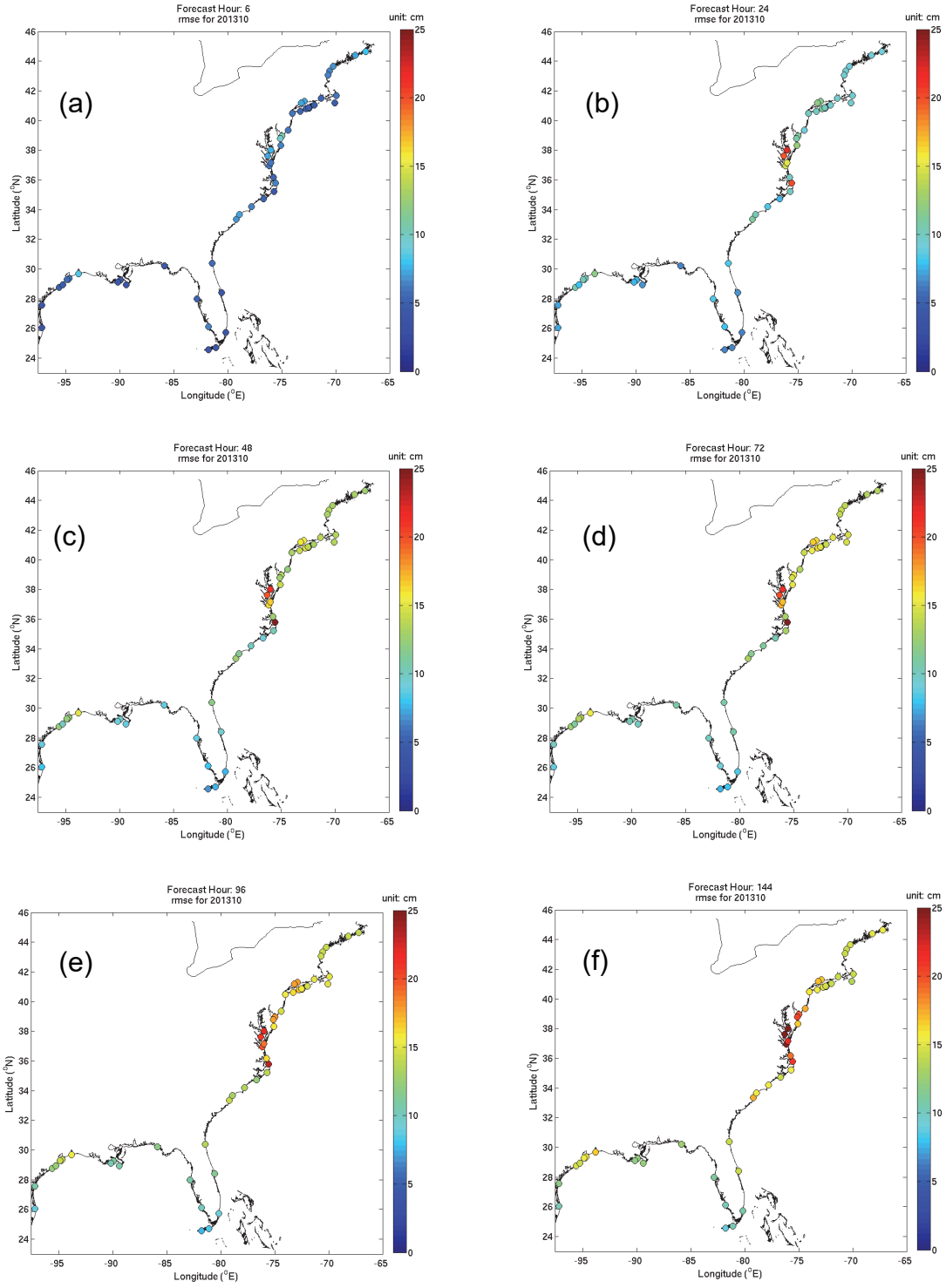


Figure 3.7. Color coded RMSE maps of G-RTOFS water level forecasts in October 2013. The six plots correspond to forecast guidance hours (a) 6, (b) 24, (c) 48, (d) 72, (e) 96, and (f) 144, respectively.

3.2. ETSS Forecast

In the following, we will compare ETSS point guidance output with measurements from 47 CO-OPS water level stations. Table A.2 lists the station meta data including their identification (ID) numbers, names, and geographical locations in longitude and latitude.

3.2.1. FC Based Assessment

Figures 3.8(a)-(d) display maps of the model root-mean-square error (RMSE) in January, April, July, and October of 2013. The RMSE typically ranges between 10 and 36 cm with variations in different months. It is around 10 to 22 cm in January, 10 to 24 cm in April, 12 to 29 cm in July, and somewhat larger around 18 to 36 cm in October. In this instance, ETSS performs better in winter and in the spring than it does in the summer and in fall.

The RMSE was close to a uniform value of 10 to 22 cm across the region in January with just slightly greater RMSE across the mid-Atlantic region in April. Performance deteriorated across the region in July, most notably across the mid-Atlantic and the eastern Gulf of Mexico. As shown in Table 3.1, the station averaged RMSE value for July was 16.9 cm. Performance deteriorated further in October, most notably across the southern Atlantic coast. The station averaged RMSE was 22.2 cm and RMSE values reached a high in the mid-Atlantic region of around 36 cm.

Figures 3.9(a)-(d) display the mean bias and standard deviation (STD) of model error at the 47 water level stations in January, April, July, and October of 2013. In all four months, the station averaged bias is below 0 suggesting that the model under predicts the water level. The negative bias is fairly modest in winter and spring at around -8 to -9 cm. This negative bias increases to around -21 by October, however. In January and in April, the STD was moderate, around 6 to 7 cm. For the remaining months, the STD is relatively smaller.

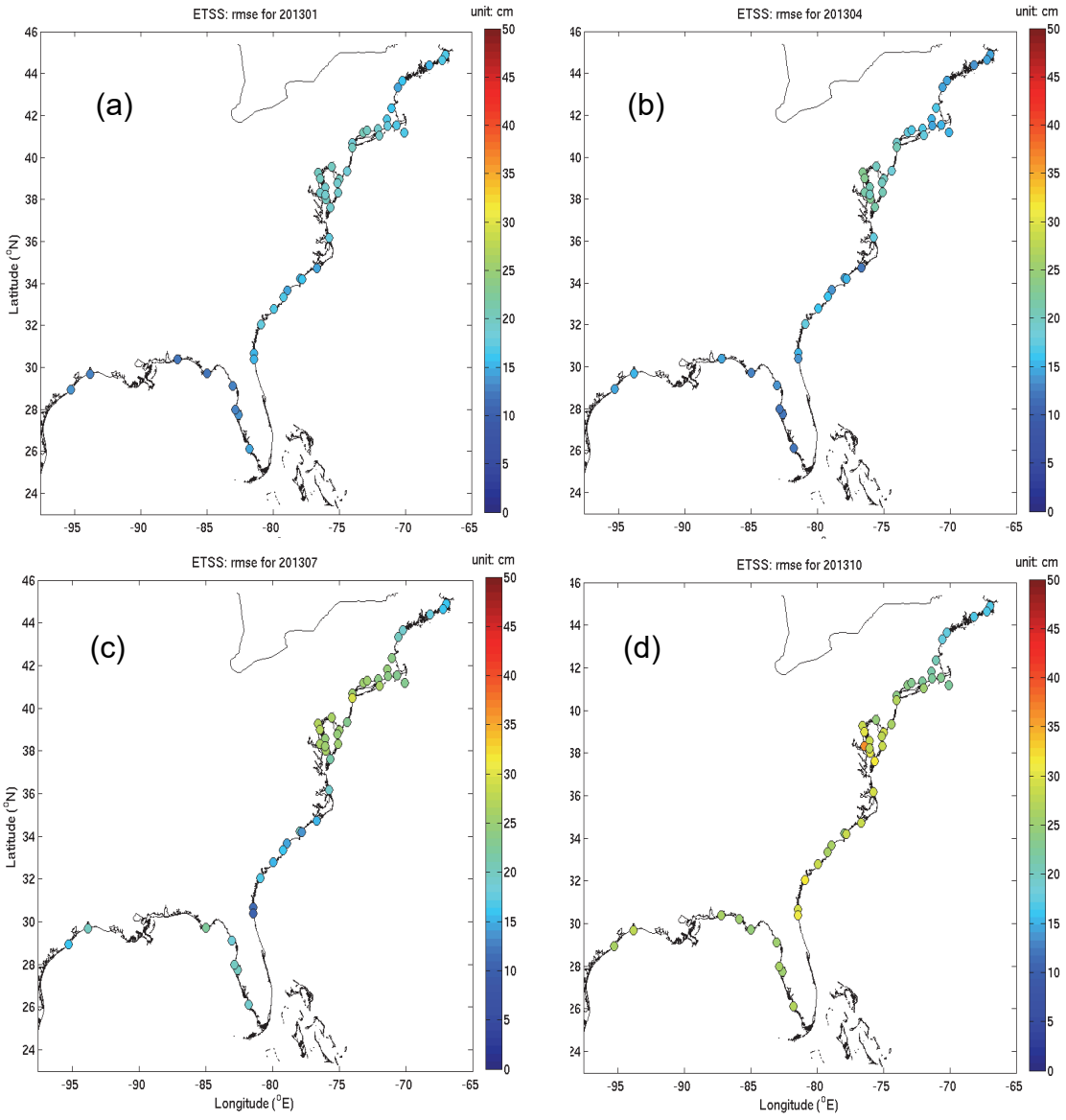
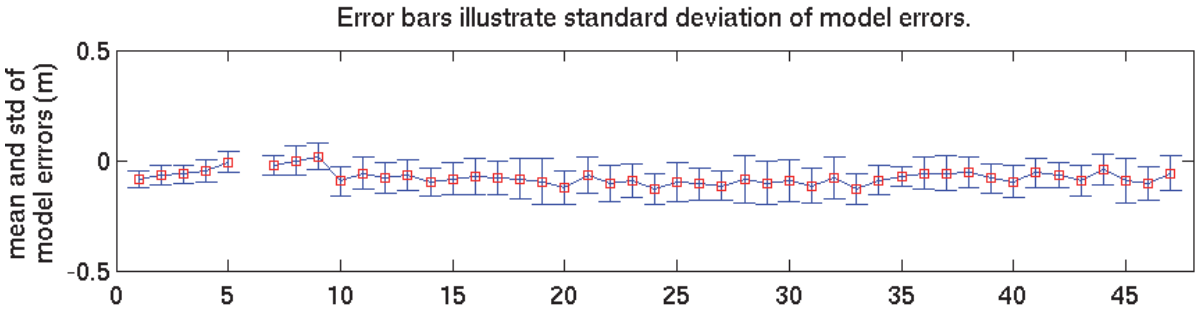


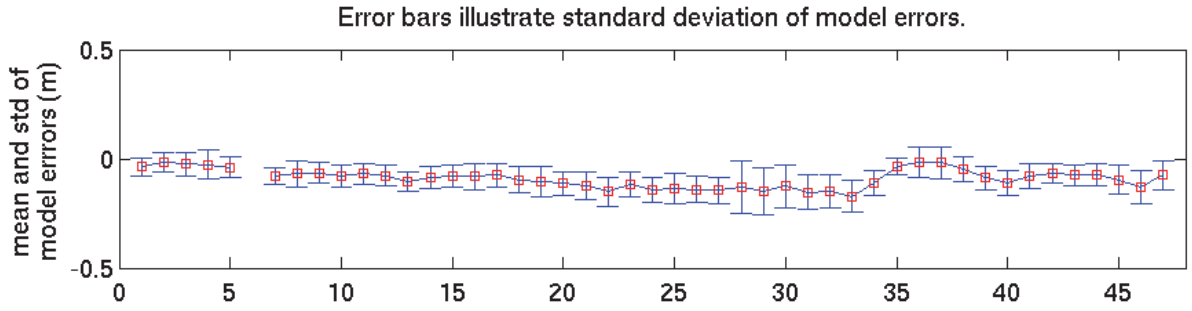
Figure 3.8. Color coded RMSE maps of the ETSS water level forecast in (a) January 2013, (b) April 2013, (c) July 2013 and (d) October 2013 at 47 CO-OPS water level stations (Table A.2).

01/2013



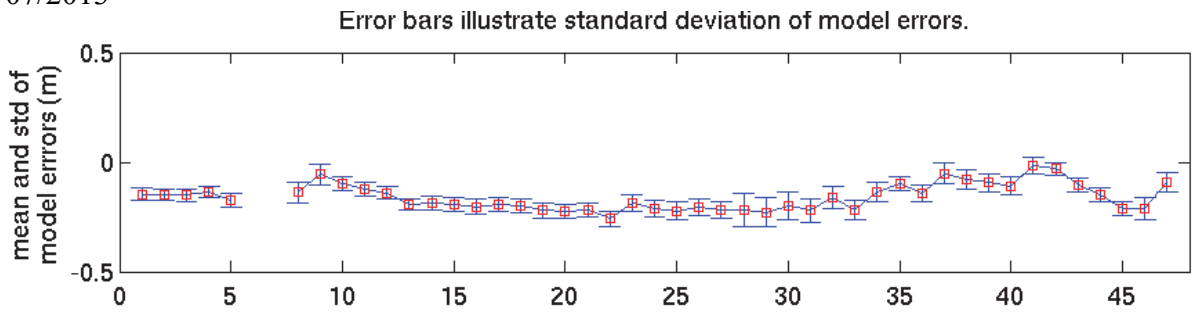
(a)

04/2013



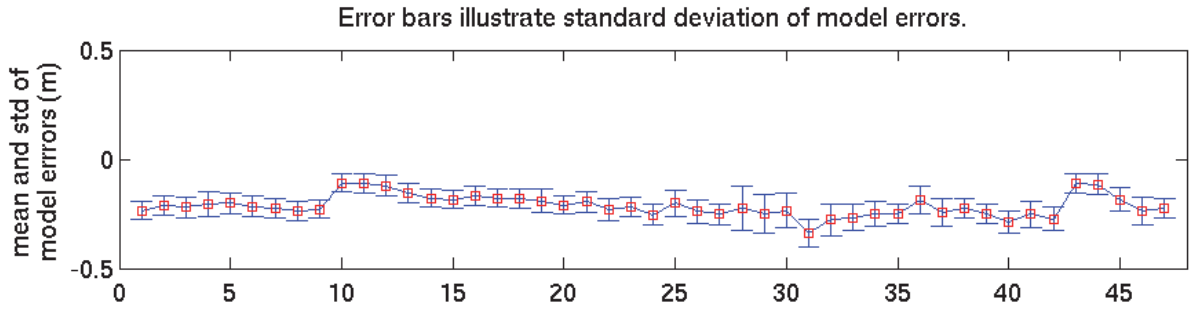
(b)

07/2013



(c)

10/2013



(d)

Figure 3.9. Means (red squares) and standard deviations (blue error bars) of the ETSS water level forecast errors in (a) January 2013, (b) April 2013, (c) July 2013, and (d) October 2013 at 47 CO-OPS water level stations. The horizontal axis shows the station numbers (Table A.2).

3.2.2. FH Method

In this section, we apply the FH method to evaluate the model performance for SSH at individual hours within a forecast cycle.

Station-Averaged RMSE Figure 3.10 displays the station averaged RMSE of ETSS water levels at each forecast hour in (a) January 2013, (b) April 2013, (c) July 2013, and (d) October 2013. As we would anticipate, model skill gradually degrades with the increasing forecast hour; the RMSE increased from 1 to 2 cm at hour 2 to 8 to 14 cm at hour 96.

The ranking of model performance by month from best to worst is July 2013, followed by October, April and January of 2013. For instance, at a time about the middle of a forecast cycle, the RMSE in July, October, April, and January was about 8, 9, 11, and 12 cm, respectively.

RMSE at Individual Stations In the previous passage, we discussed the model skill averaged over all stations. Now, we examine the skill at the individual stations. Figures 3.11 – 3.14 display station maps of water level RMSE for (a) January 2013, (b) April 2013, (c) July 2013, and (d) October 2013, respectively. The four plots in each figure correspond to forecast hours 12, 24, 48, and 96.

In general, the model demonstrates evenly satiable skills at hour 12 across all stations and seasons with RMSE generally less than 15 cm.

At each station, model skill gradually degrades with increasing forecast hour. Model skill is good within the first 12 hours of the forecast cycle. For instance, for all four months, the RMSE ranges between 6 and 15 cm at hour 12. The plots for all four months clearly demonstrate the rate of RMSE deterioration is fastest from hours 12 through 24 of the forecast cycle. The RMSE increases to as high as 20 cm at hour 24 at mid-Atlantic stations during January and April. The RMSE gradually reaches its peak between hours 48 and 96.

However, the degradation varies by month. It was most severe in January and April, less severe in July, and at an immediate level in October.

The model performance varies by latitude. The RMSE deterioration is most marked at the mid-latitude stations around 38 – 40^{o} and in the region of New York Harbor and Long Island Sound at around 40 – 42^{o} (see Table A.1 for the station meta data). The more problematic stations are in the Chesapeake Bay and Delaware Bay regions, and in the New York Harbor and the Long Island Sound region. The spatial disparity is for the most part evident by hour 24 in the forecast cycle.

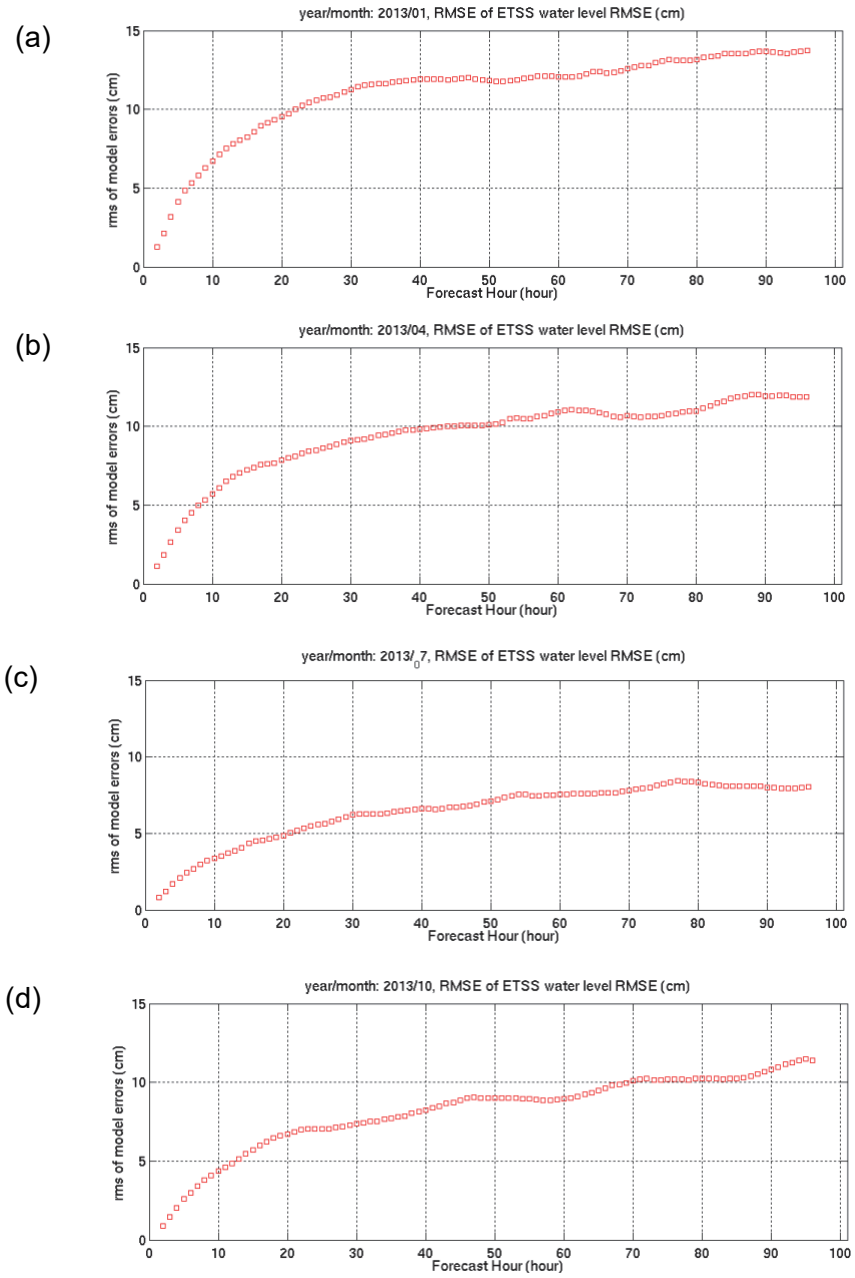


Figure 3.10. The RMSE of the ETSS water level forecast at each forecast hour in (a) January 2013, (b) April 2013, (c) July 2013, and (d) October 2013.

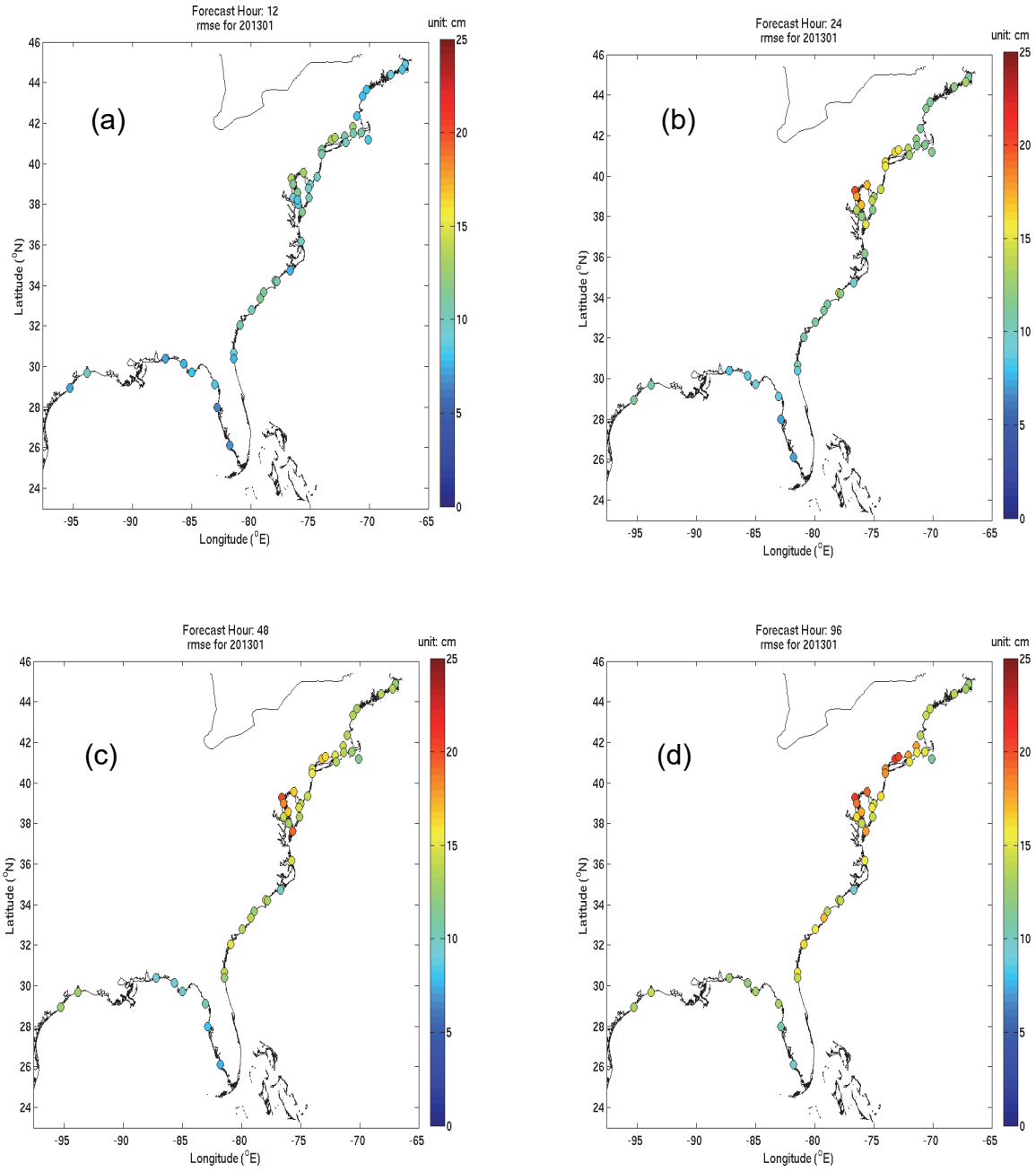


Figure 3.11 Color coded RMSE maps of the ETSS water level forecast in January 2013. The four plots correspond to forecasts at hours (a) 12, (b) 24, (c) 48, and (d) 96, respectively.

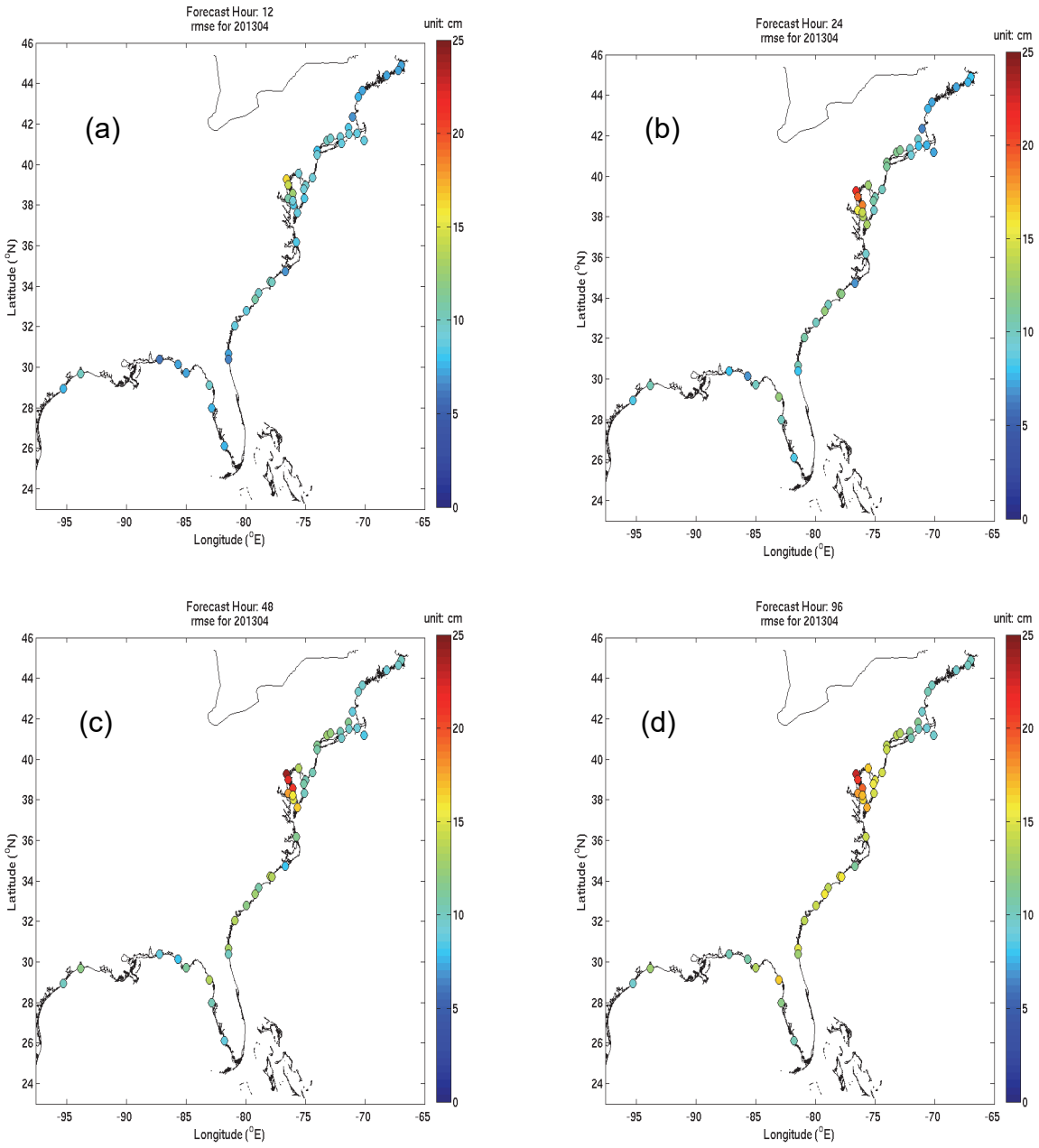


Figure 3.12. Color coded RMSE maps of the ETSS water level forecast in April 2013. The four plots correspond to forecasts at hours (a) 12, (b) 24, (c) 48, and (d) 96, respectively.

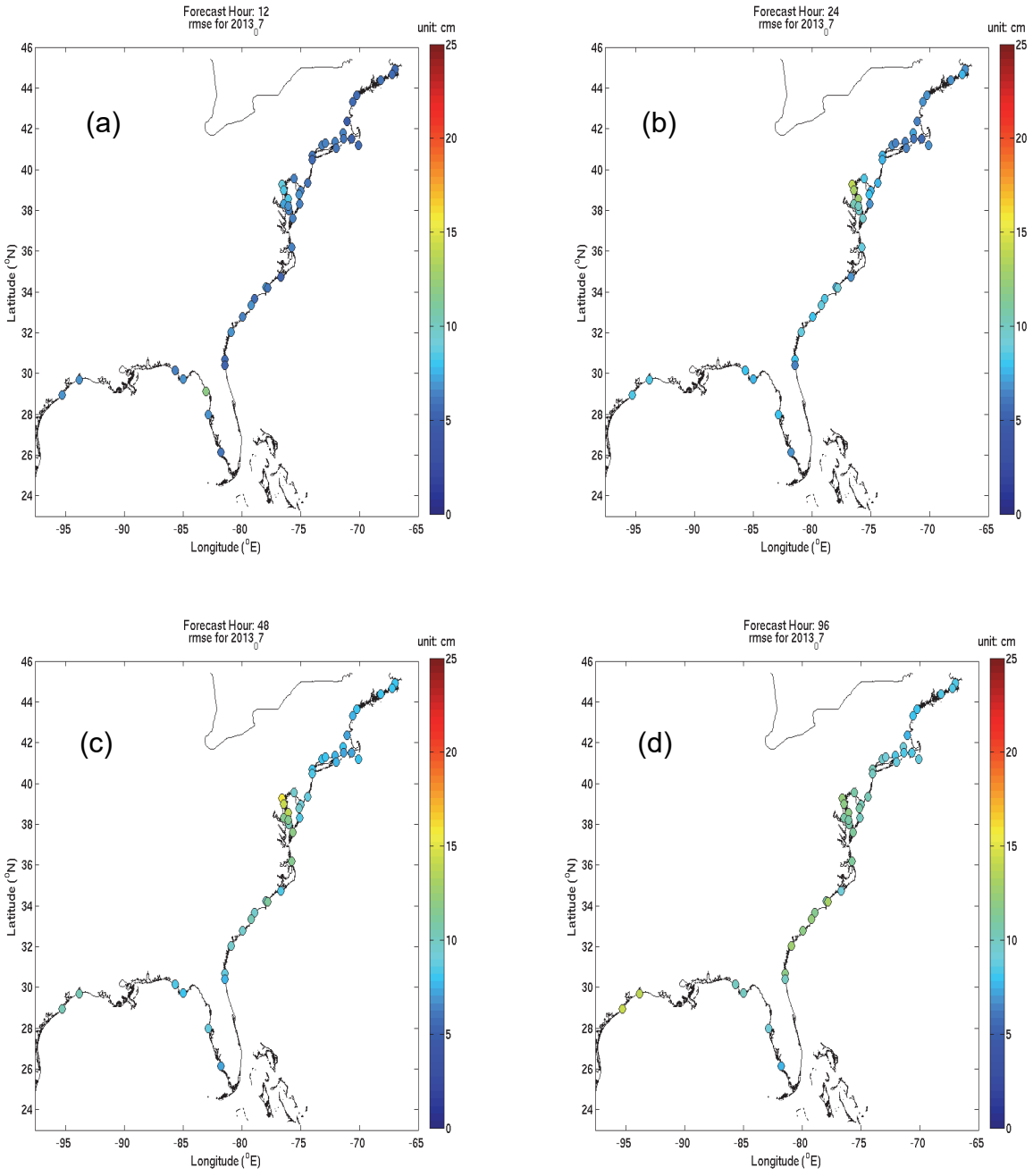


Figure 3.13. Color coded RMSE maps of the ETSS water level forecast in July 2013. The four plots correspond to forecasts at hours (a) 12, (b) 24, (c) 48, and (d) 96.

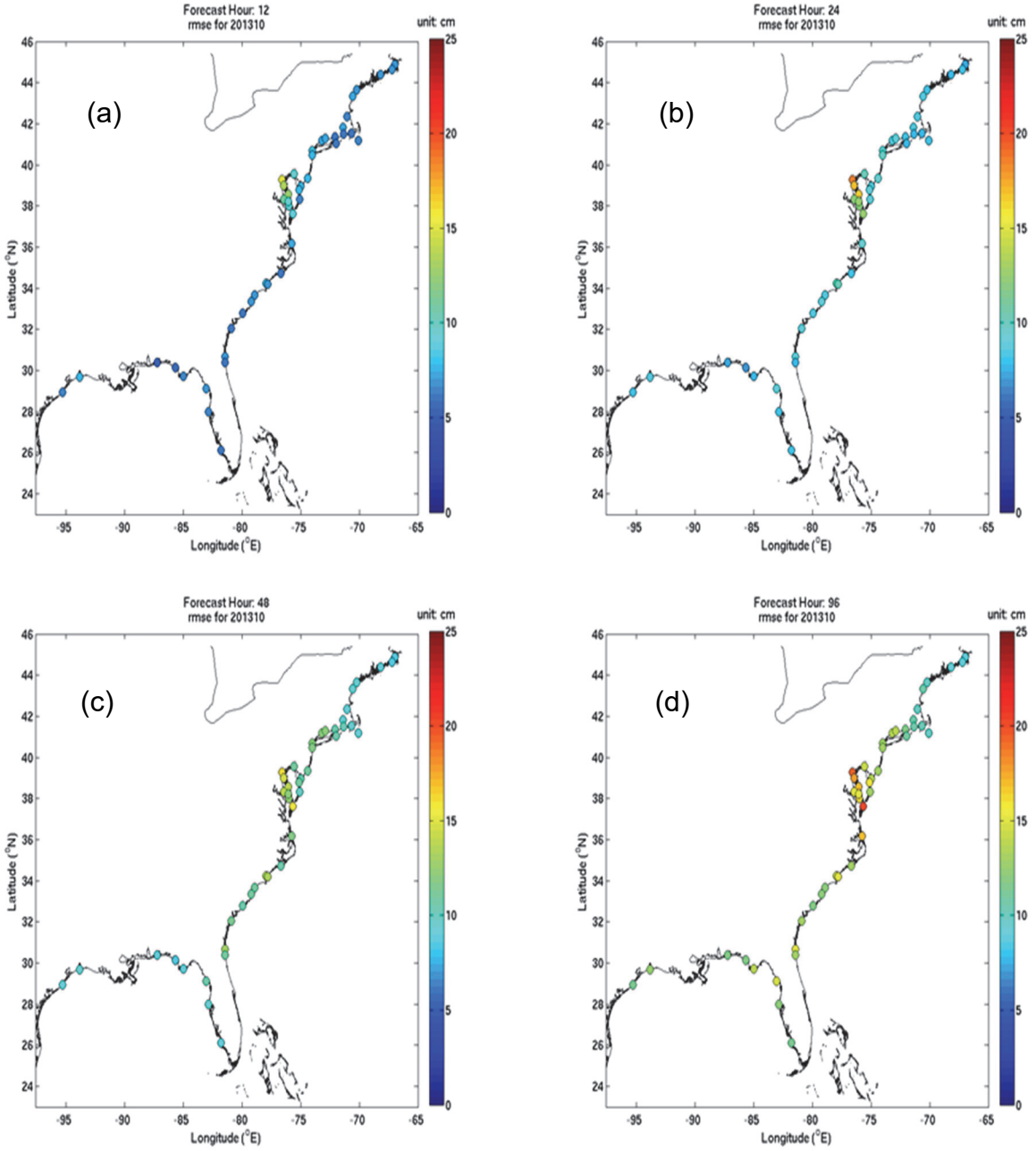


Figure 3.14. Color coded RMSE maps of the ETSS water level forecast in October 2013. The four plots correspond to forecasts at hours (a) 12, (b) 24, (c) 48, and (d) 96.

3.3. ESTOFS Forecast

We now compare the ESTOFS water level forecasts with measurements at 54 CO-OPS water level stations. Table A.3 lists the station IDs, names, and geographical locations in longitude and latitude.

3.3.1. FC Based Assessment

Figures 3.15(a)-(d) display model RMSE maps for January, April, July, and October of 2013. The RMSE typically ranges between 16 and 30 cm in January 2013, 12 and 28 cm in April 2013, 15 and 28 cm in July 2013, and between 18 and 30 cm in October 2013. The distinction from season to season is not as clear cut with ESTOFS as it is with ETSS and G-RTOFS.

The maximum RMSE by season, starting with January, is 30 cm, 28 cm, 28 cm, and 30 cm. The largest station averaged value of RMSE, 21.7 cm, occurred in October. There appears to be measureable spatial disparity for all months except April 2013. For January, July, and October of 2013, the RMSE at the mid-Atlantic stations of the Chesapeake Bay, the Delaware Bay, and Long Island Sound are noticeably greater. The spatial variability is of interest in October. Along the U.S. east coast, there are four distinct regions of RMSE values: the southern east coast up to VA., the mid-Atlantic region, the New York Harbor/Long Island Sound region as far north as Boston (8443970), and the northeast (Gulf of Maine) region.

Figures 3.16(a)-(d) display the mean bias and STD of model errors at the 54 CO-OPS water level stations for January 2013, April 2013, July 2013, and October 2013. For the majority of stations, the mean bias is below 0 for all four seasons indicating that ESTOFS tends to under-predict subtidal water levels throughout the year. The bias does seem to vary by season. The station averaged bias is as small as -3.3 cm during April and as large as -19.1 cm during October.

The magnitude of standard deviation is relatively larger, almost by a factor of 2, for ESTOFS than it is for ETSS. STD values appear to be largest in January at about 5 and 20 cm, with a station averaged deviation of 12.5 cm. The standard deviation decreases slightly to about 5 to 12 cm during April. STD is smallest during July with a station averaged value of 6.2 cm. Typically, the spring and fall are subject to more stormy weather while July tends to more placid.

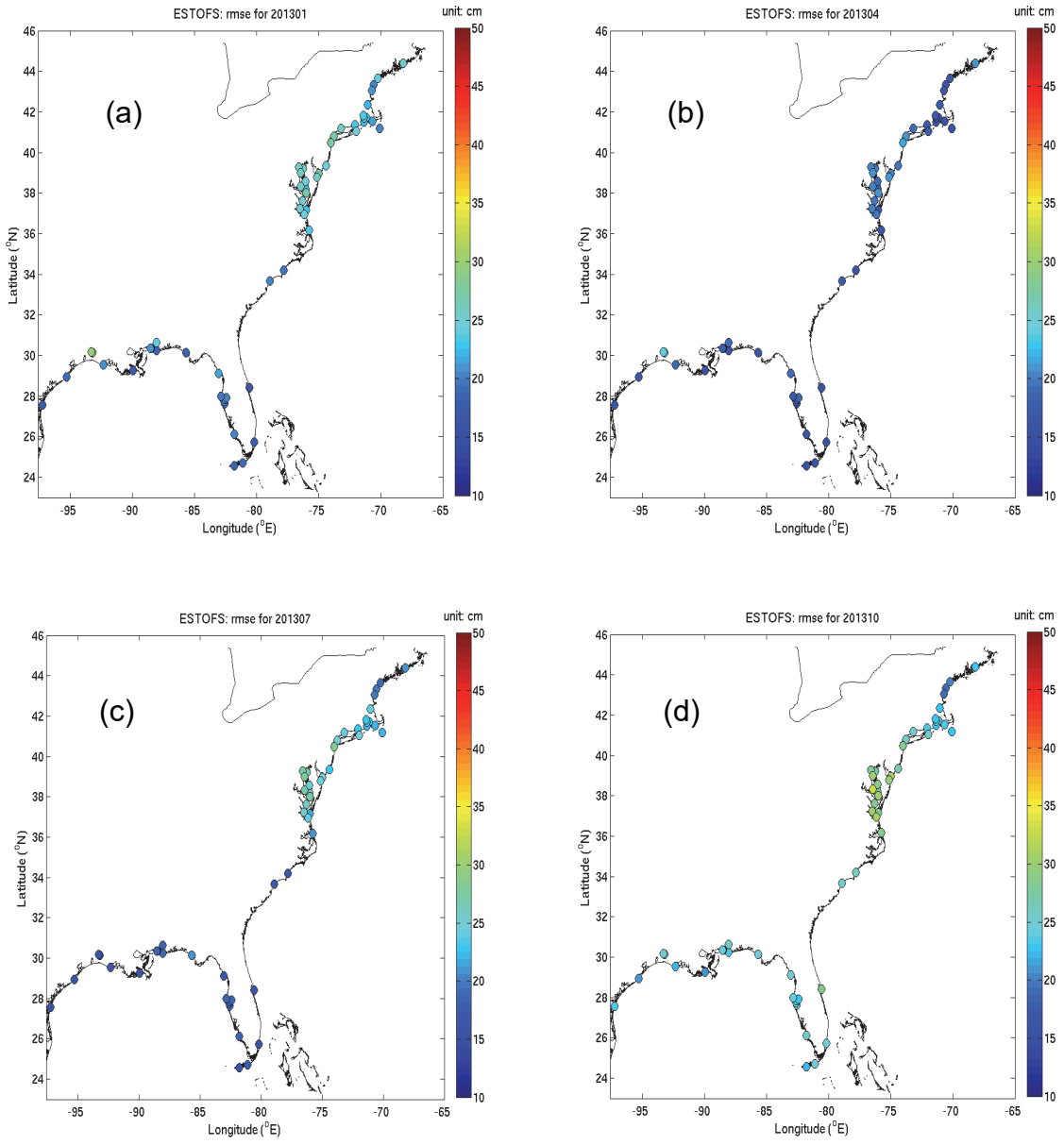
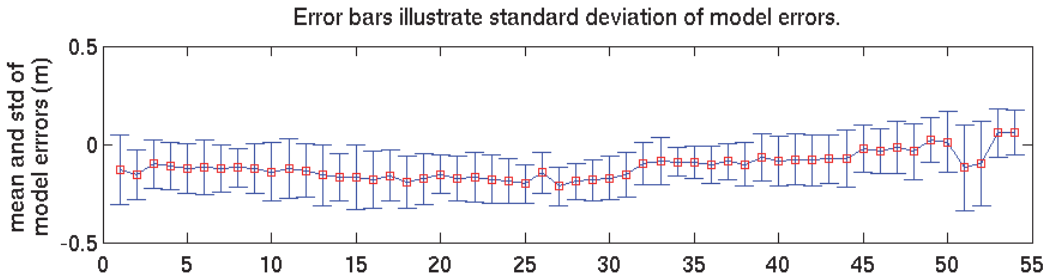


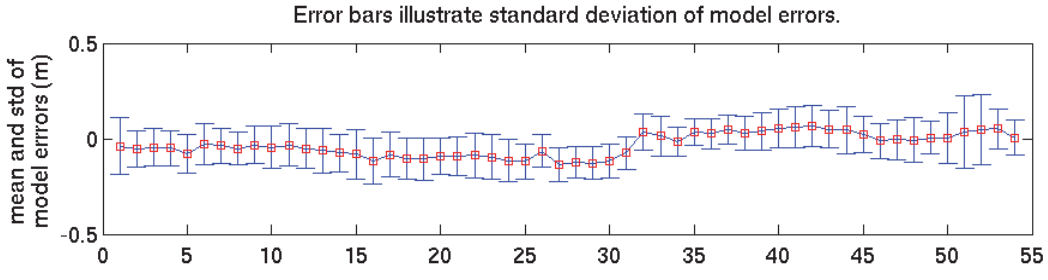
Figure 3.15. Color coded RMSE maps of ESTOFS water level forecasts in (a) January 2013, (b) April 2013, (c) July 2013, and (d) October 2013 at 54 CO-OPS water level stations (Table A.3).

01/2013



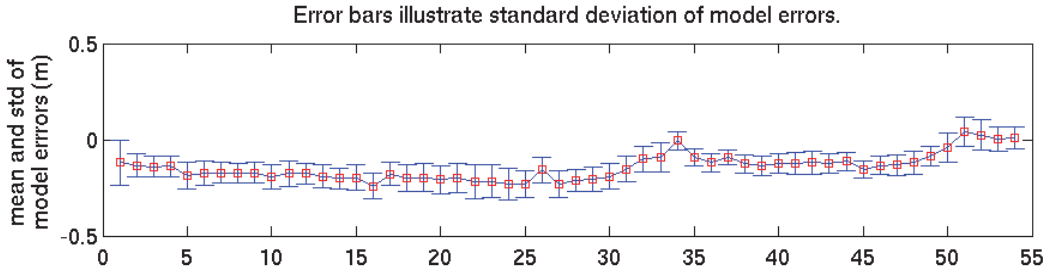
(a)

04/2013



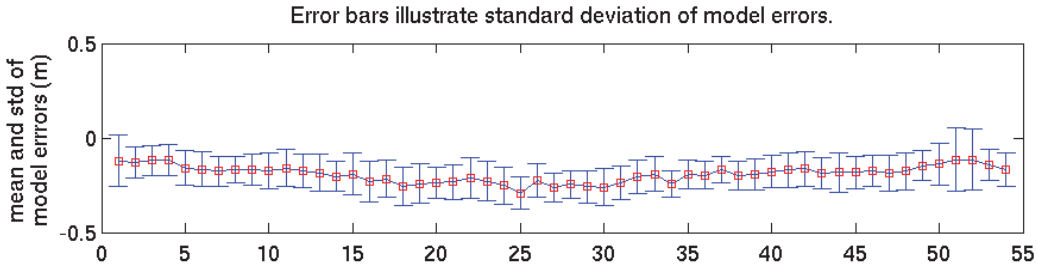
(b)

07/2013



(c)

10/20113



(d)

Figure 3.16. Mean (red squares) and standard deviation (blue error bars) of the ESTOFS water level errors for 2013 at 54 CO-OPS water level stations from: (a) January 2013, (b) April 2013, (c) July 2013, and (d) October 2013. The horizontal axis shows the station numbers (Table A.3).

3.3.2. FH Method

In this section, we apply the FH method to evaluate the model performance for SSH at individual hours within a forecast cycle.

Station-Averaged RMSE Figure 3.17 displays the station averaged RMSE of ESTOFS water levels at each forecast hour in (a) January 2013, (b) April 2013, (c) July 2013, and (d) October 2013. As we would anticipate, model skill degrades with increasing forecast hour. The RMSE deteriorated in all four months, though the deterioration was most spectacular during January and April where the RMSE increased to 22 cm and 18 cm, respectively. The ranking of model performance by month from best to worst was July, followed by October, April, and January of 2013. The ESTOFS water level forecasts do not appear to asymptotically approach a limiting value.

The RMSE plots display a small sinusoidal oscillation with a period of about 12 hours. The oscillation is due to an alias from the low sampling rate, daily in this case, of the underlying time series.

RMSE at Individual Stations In the previous passage, we discussed the model skill averaged over all stations. Now, we will examine the skill at each individual station. Figures 3.18 – 3.21 display station maps of water level RMSE for (a) January 2013, (b) April 2013, (c) July 2013, and (d) October 2013, respectively. The four plots in each figure correspond to forecast hours 6, 24, 48, 96, 144, and 180.

The plots for all four months demonstrate that the rate of RMSE deterioration is much more gradual for ESTOFS than for either ETSS or G-RTOFS. There is very little deterioration in model skill during January even by hour 180. Two stations on the southern east coast and two stations in the Gulf of Mexico show measurable deterioration to about 25 cm. April is the month showing the greatest deterioration in model skill with forecast hour. Several stations in the Gulf of Mexico have an RMSE above 30 cm. Skill deterioration subsides somewhat in July, and further still in October.

The model performance varies by region and by latitude. The RMSE deterioration is most marked at the mid-latitude east coast stations around 38 – 40^{o} and in the region of New York Harbor and Long Island Sound at around 40 – 42^{o} (see Table A.3 for the station meta data). The more problematic stations are located in the Chesapeake Bay: Cambridge (8571892), Tolchester Beach (8573364), Baltimore (8574680), Annapolis (8575512), Solomons Island (8577330), Kiptopeake (8632200), and Lewisetta (8635750), in the Delaware Bay region: Lewes (8557380), Cape May (8536110), and in the New York Harbor and Long Island Sound region: Sandy Hook (8531680), Kings Point (8516945), and Bridgeport (8467150). There are several stations along the gulf coast which appear to be problematic as well.

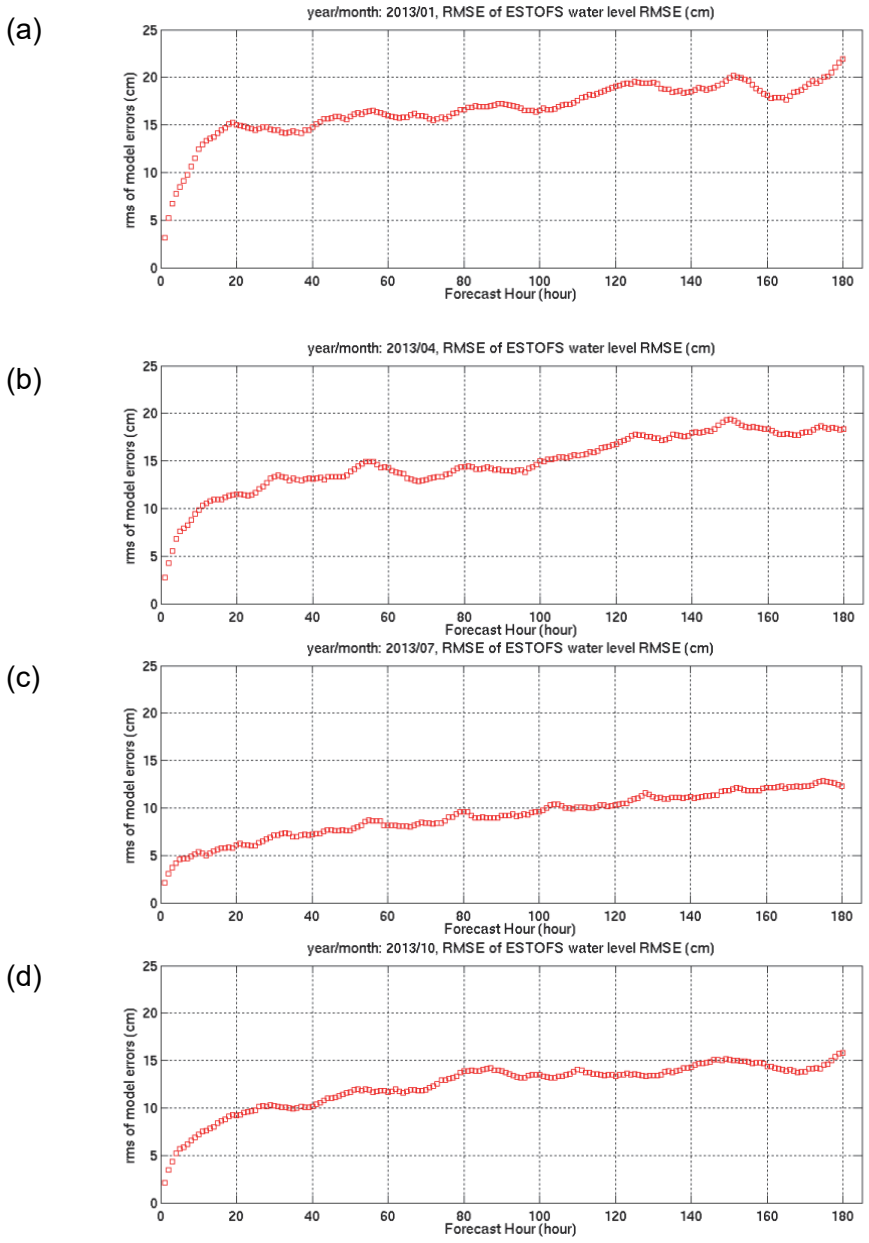


Figure 3.17. Station-averaged root mean squared errors of ESTOFS water level forecasts in (a) January 2013, (b) April 2013, (c) July 2013, and (d) October 2013.

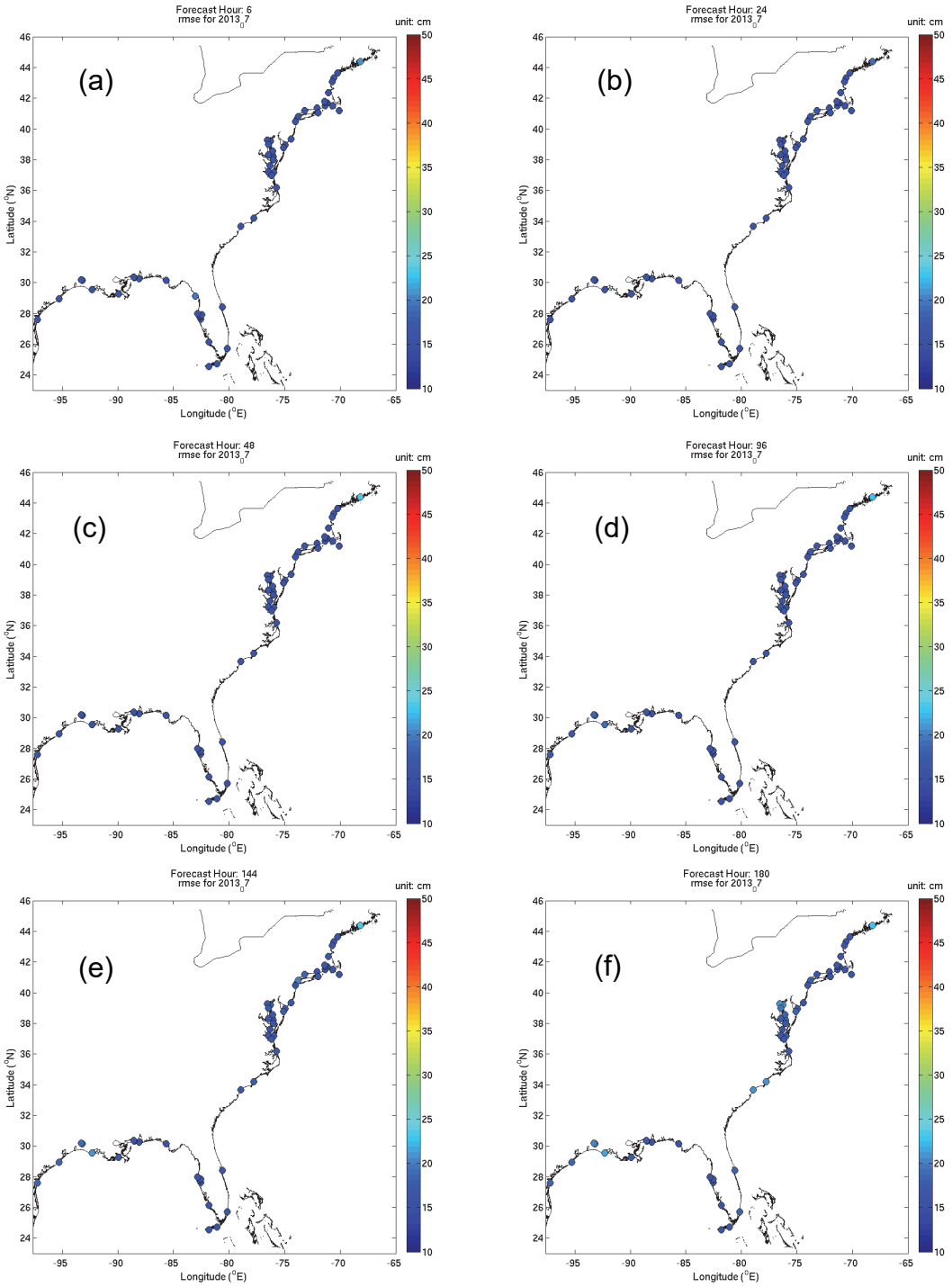


Figure 3.18. Color coded RMSE maps of ESTOFS water level forecasts from January 2013. The six plots correspond to forecast guidance hours (a) 6, (b) 24, (c) 48, (d) 96, (e) 144, and (f) 180, respectively.

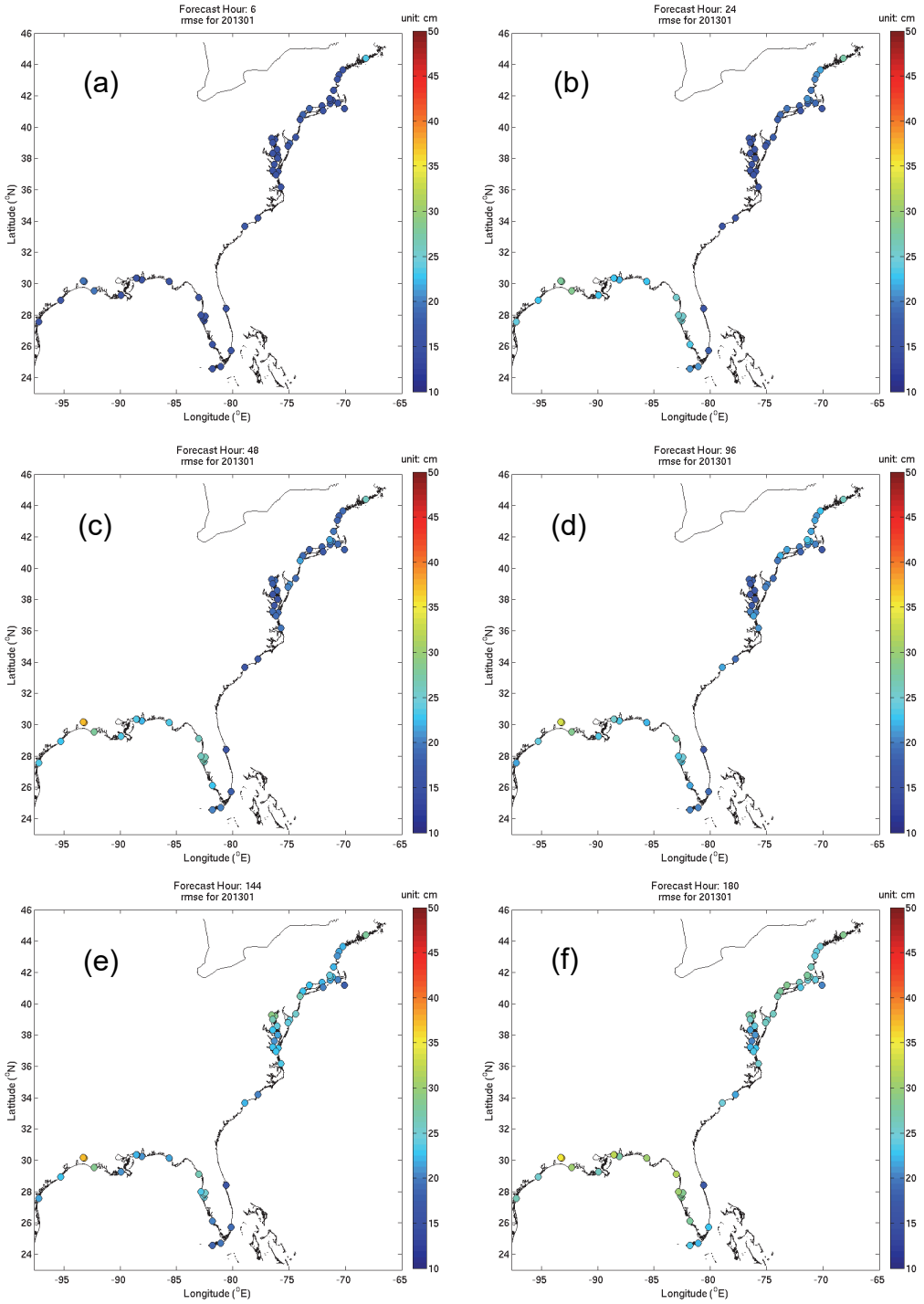


Figure 3.19. Color coded RMSE maps of ESTOFS water level forecasts from April 2013. The six plots correspond to forecast guidance hours (a) 6, (b) 24, (c) 48, (d) 96, (e) 144, and (f) 180, respectively.

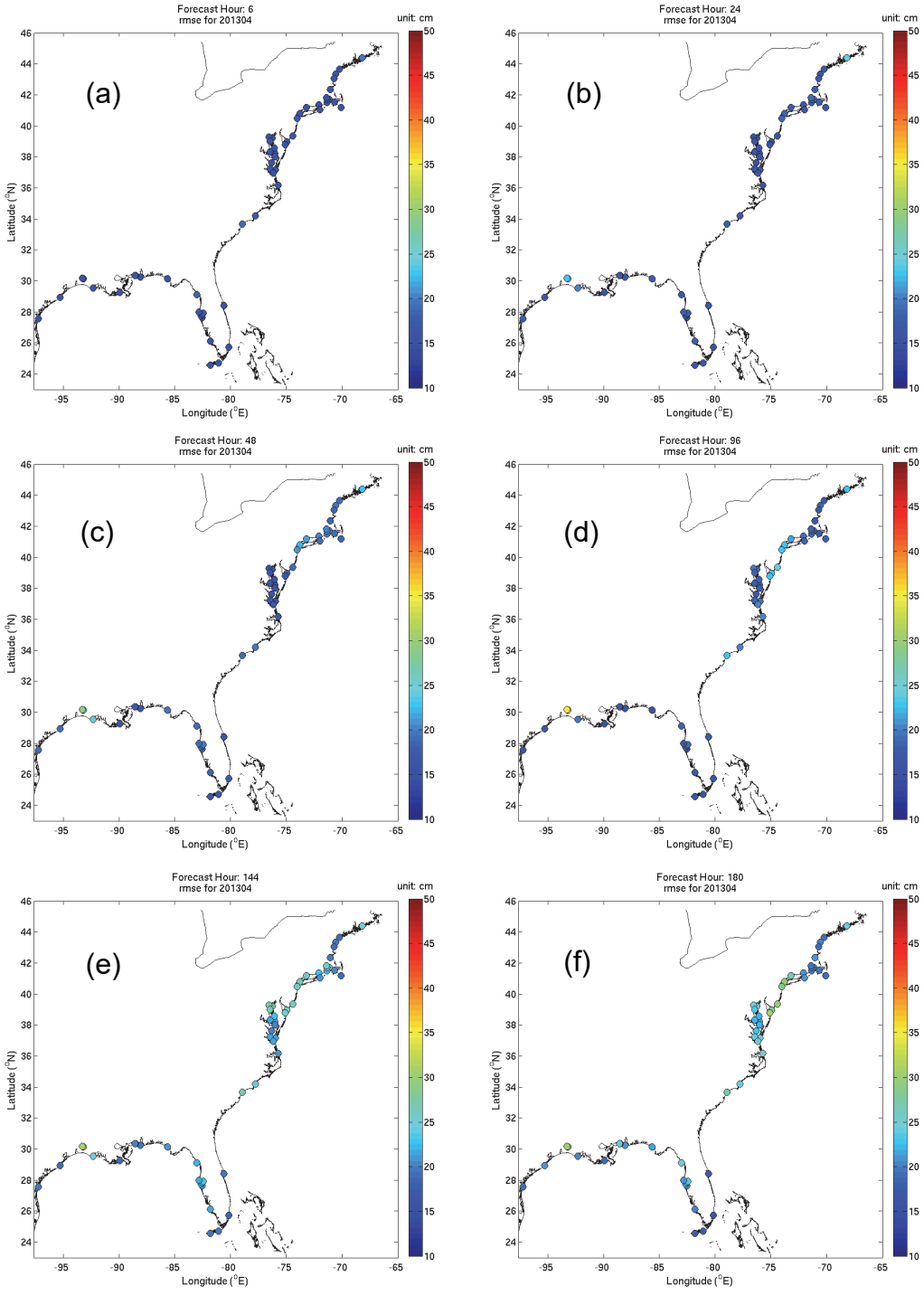


Figure 3.20. Color coded RMSE maps of ESTOFS water level forecasts from July 2013. The six plots correspond to forecast guidance hours (a) 6, (b) 24, (c) 48, (d) 96, (e) 144, and (f) 180, respectively.

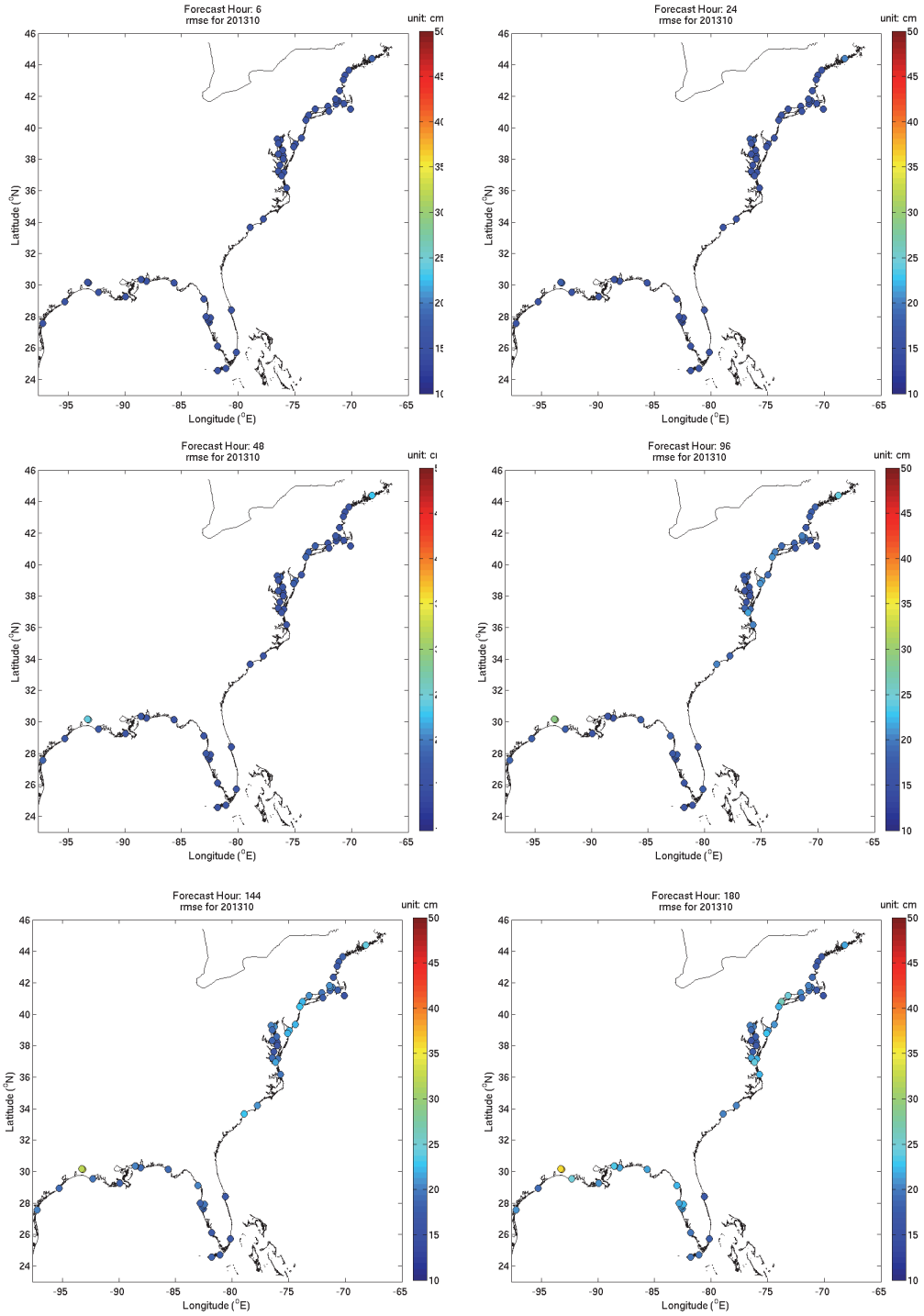


Figure 3.21. Color coded RMSE maps of ESTOFS water level forecasts from October 2013. The six plots correspond to forecast guidance hours (a) 6, (b) 24, (c) 48, (d) 96, (e) 144, and (f) 180, respectively.

3.4. Summary

We assessed the performance of the G-RTOFS, ETSS, and the ESTOFS SSH forecast guidance by comparing model results with observations at 52, 47, and 54 CO-OPS water level stations, respectively (Tables A.1, A.2, and A.3). We applied both the forecast-cycle (FC) based method and the forecast-hour (FH) based method in the analysis (Section 2).

Table 3.1 lists the station averaged bias, standard deviation, and root-mean-squared error for the winter, spring, summer, and fall, estimated using the FC based method. G-RTOFS displays a very large negative bias throughout the year with station averaged values of -45.3 cm in the winter, -53.6 cm in the spring, -47.6 cm in the summer, and -49.7 cm in the fall. The size of the negative bias of G-RTOFS indicates a datum issue within the model/system. The negative bias of ETSS starts small but grows more negative through the year. The station averaged bias is -7.6 cm in the winter and -9.1 cm in the spring. It jumps to -16.0 cm in the summer, then to -21.3 cm in the fall. The values of negative ESTOFS bias are much more modest and vary somewhat through the year with a range of -20 to 0 cm in the winter, -5 to 3 cm in the spring, -20 to 0 cm in the summer, and -25 to 0 cm in the fall.

The standard deviation of G-RTOFS ranges between 5 and 20 cm. The standard deviation of ETSS is the smallest, ranging between 1 and 9 cm. The STD of ESTOFS is somewhat larger, ranging between 5 and 20 cm over the four seasons.

The RMSE for G-RTOFS ($RMSE_{G\text{-}RTOFS}$) peaks during the spring with values at many stations reaching 70 cm or greater and a station averaged value of 55.6 cm. RMSE values drop a bit in summer to a station average of 49.5 cm, and then rise just a bit in fall to an average of 50.9 cm. The RMSE of ETSS is moderately low in both January and in April with station averaged values of 11.9 cm and 11.7 cm, respectively. The RMSE increases in summer to a station averaged value of 16.9 cm. The RMSE peaks during October with values ranging between 20 and 38 cm and a station averaged value of 22.2 cm. The RMSE of ESTOFS is more modest than that of G-RTOFS throughout the year. The RMSE never goes much above 30 cm with seasonal highs of 30, 28, 28, and 30 corresponding to the seasons of winter, spring, summer, and fall, respectively. Overall, the RMSE for ETSS and ESTOFS are comparable, while the $RMSE_{G\text{-}RTOFS}$ is greater by a factor of two or three than that of the other two.

Table 3.1 Forecast cycle-based statistics of the G-RTOFS, ETSS, and ESTOFS SSH errors.

	SSH of G-RTOFS (144 hour Forecast Cycle)			**SSH of ETSS (96 hour Forecast Cycle)**			**SSH of ESTOFS (180 hour Forecast Cycle)**		
	bias (cm)	std (cm)	rmse (cm)	bias (cm)	std (cm)	rmse (cm)	bias (cm)	std (cm)	rmse (cm)
Winter (Jan. 2013)	-45.3	10.2	47.9	-7.6	7.2	11.9	-11.3	12.5	18.5
Spring (April. 2013)	-53.6	8.4	55.6	-9.1	6.1	11.7	-3.3	10.4	13.2
Summer (July. 2013)	-47.6	5.6	49.5	-16.0	3.9	16.9	-14.4	6.2	16.7
Fall (Oct. 2013)	-49.7	7.4	50.9	-21.3	5.2	22.2	-19.1	9.3	21.7
Mean	-49.1	7.9	51.0	-13.5	5.6	15.7	-12.0	9.6	17.5

Table 3.2 Forecast hour-based statistics of the G-RTOFS, ETSSS, and ESTOFS SSH errors.

	SSH of G-RTOFS		**SSH of ETSS**		**SSH of ESTOFS**	
	rmse (cm) (144 hour cycle)	rmse (cm) (48 hours)	rmse (cm) (96 hours)	rmse (cm) (48 hours)	rmse (cm) (180 hours)	rmse (cm) (48 hours)
Winter (Jan. 2013)	14.2	10.5	11.1	9.4	16.6	13.3
Spring (April. 2013)	12.5	9.6	9.5	7.8	14.8	11.2
Summer (July. 2013)	8.6	5.3	6.4	5.0	9.2	6.1
Fall (Oct. 2013)	10.7	7.1	8.2	6.4	12.3	8.8
Mean	11.5	8.1	8.8	7.2	13.2	9.9

4. PERFORMANCE OF G-RTOFS FOR SEA-SURFACE TEMPERATURE

In this chapter, we discuss the performance of the G-RTOFS sea-surface temperature (SST) forecast.

4.1. Compared with CO-OPS Observations

We compared the G-RTOFS SST forecast with measurements from 45 CO-OPS physical oceanography observation stations (Figure 1.6). Table B.1 lists the station meta data including IDs, names, and geographical locations in longitude and latitude. In the following, we describe the results derived from both the forecast cycle (FC) and forecast hour (FH) based methods.

4.1.1. Forecast Cycle-Based Method

Figures 4.1(a)-(d) display RMSE maps for January 2013, April 2013, July 2013, and October 2013. The RMSE displays somewhat greater magnitude and spatial variability in summer than during the other seasons. The RMSE is no higher than 6 ºC in January, no higher than 5 ºC in April, and ranges from about 2 to 7 ºC in July. It goes no higher than 4 ºC in October. July had the highest station averaged mean RMSE at 2.05 ºC.

Figures 4.2(a)-(d) display the model bias and standard deviation (STD) of model errors at the 45 CO-OPS stations in January 2013, April 2013, July 2013, and October 2013. Generally, the average model-data difference during July ranges between 5 ºC on the positive side and -8 ºC on the negative end (though the -8 ºC could be an outlier). The magnitude of bias is around -2 to 5 ºC in January and between -1 and 2 ºC in October. July has the largest magnitudes of bias with values ranging between -8 ºC and 5 ºC. However, the station averaged bias for July is a rather modest 0.39 ºC.

The corresponding values of STD display similar seasonal variability, with typical values of 1ºC or less in January and April. There is a slightly greater range of dispersion in July and October with some values of STD reaching 2 ºC. The station averaged STD is 0.99 for July compared with 0.72 for January. Model values are less scattered about the mean in winter and spring than in summer than in fall.

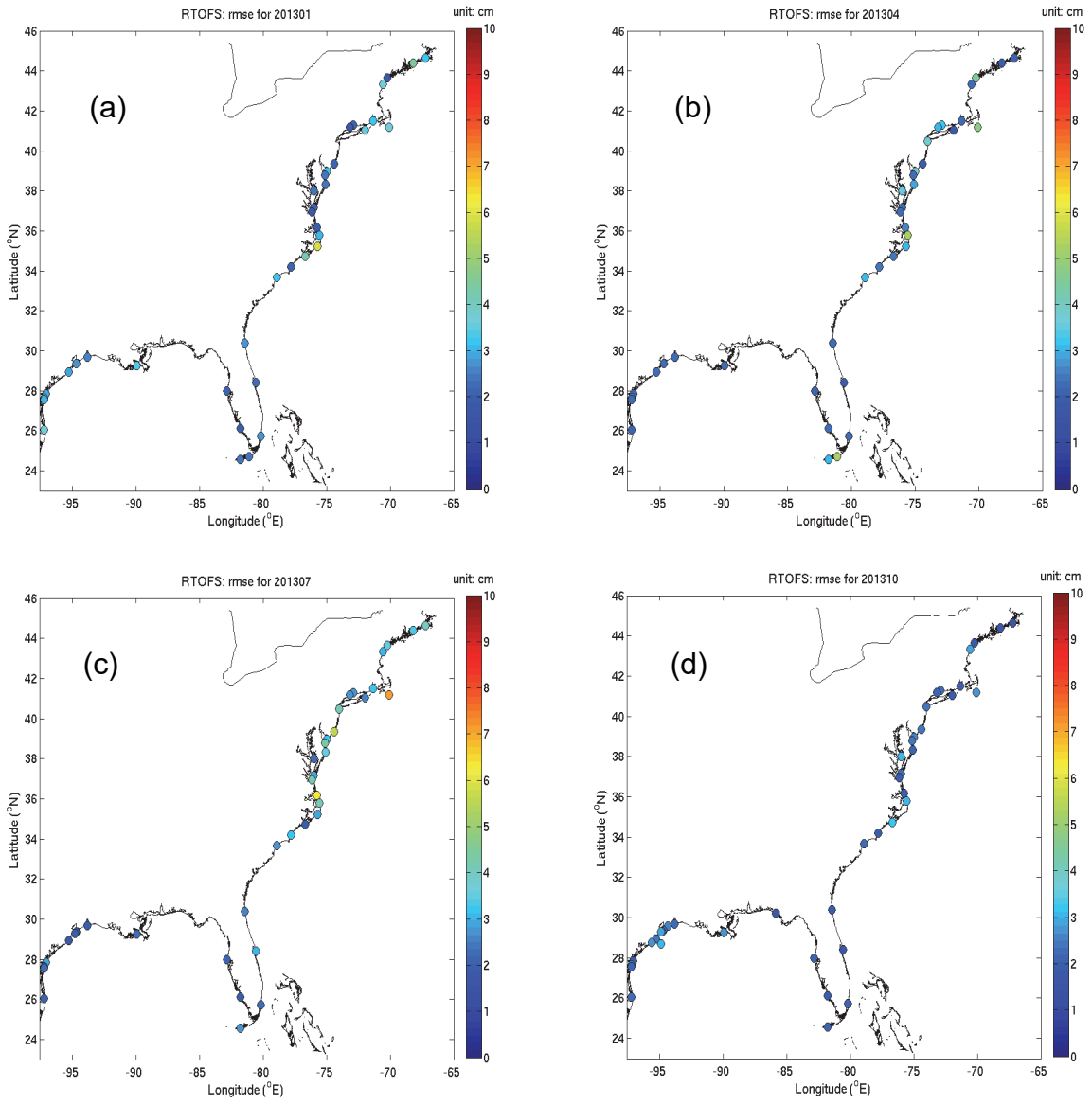
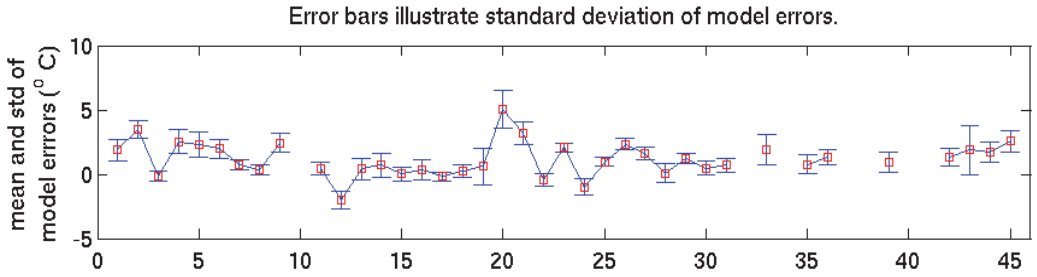


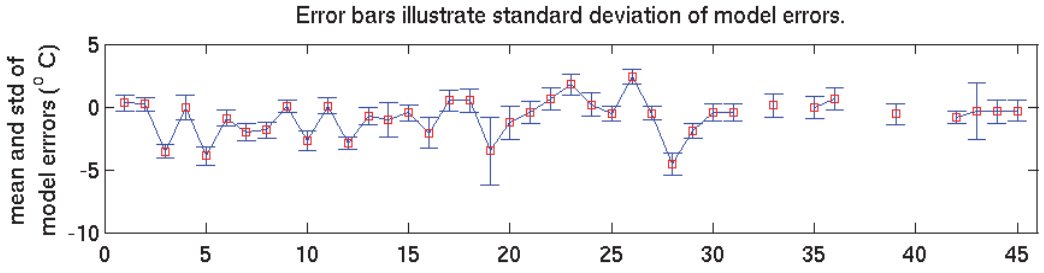
Figure 4.1. Color coded RMSE maps of the G-RTOFS SST forecast in (a) January 2013, (b) April 2013, (c) July 2013, and (d) October 2013 at 45 CO-OPS physical oceanography observation stations (Table B.1).

01/2013



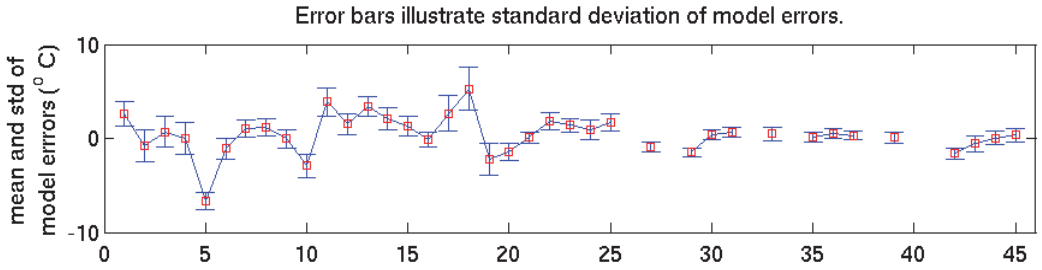
(a)

04/2013



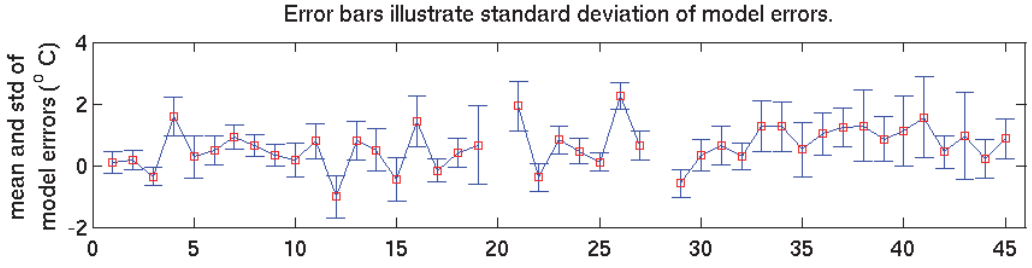
(b)

07/2013



(c)

10/2013



(d)

Figure 4.2. Means (red squares) and standard deviations (blue error bars) of the G-RTOFS SST errors in (a) January 2013, (b) April 2013, (c) July 2013, and (d) October 2013 at 45 CO-OPS physical oceanography observation stations (Table B.1).

4.1.2. Forecast Hour (FH) Based Method

In this section, we investigate the model performance at individual hours within the forecast cycle. We first describe the results in terms of station-averaged RMSE and then investigate the RMSE at all individual stations.

Station-averaged RMSE Figure 4.3 displays station averaged RMSE by forecast hour in (a) January 2013, (b) April 2013, (c) July 2013, and (d) October 2013. During the months of January, April, and October, the RMSE follows a roughly sinusoidal pattern. The sinusoidal pattern is very clear in October, starting at about 1.1 ºC near hour 0 and ending near 1.5 º C, with an amplitude of about 0.1 to 0.2 º C.

RMSE at Individual Stations We now examine the skill at the individual stations. Figures 4.4-4.7 display RMSE maps for all stations in (a) January 2013, (b) April 2013, (c) July 2013, and October 2013, respectively. Six plots in each figure correspond to forecast hours 6, 24, 48, 72, 96, and 144, respectively.

For the months of January, April, and July, the distinguishing characteristic is the spatial variation of the RMSE. During January, two stations on the Carolina coast, as well as a station to the northeast of the Gulf of Maine, appear to be problematic. During April, the skill at the Carolina stations appears to be improved, but the problematic station appears to be one station to the north on the coast of Virginia. In July, the RMSE at hour six ranges from about 1ºC to around 5ºC (or above). The model skill does not appear to degrade that quickly by hour 144. The rate at which model skill degrades appears to be much slower for temperature than for surface water levels. The RMSE during October appears to be quite a bit smaller than for the other months, never going much over 2.5 ºC.

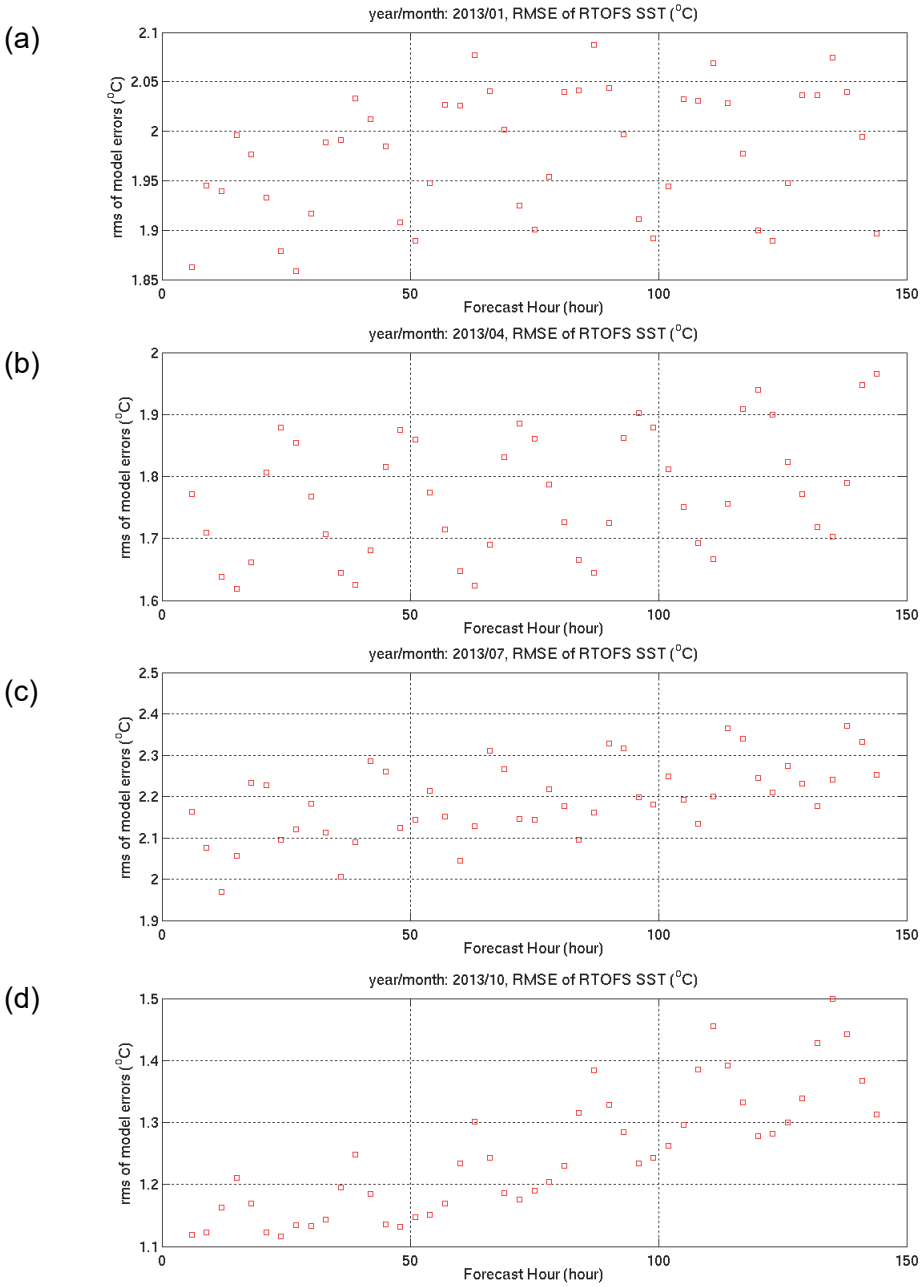


Figure 4.3. Root mean squared errors (RMSE) of the G-RTFOS SST forecast at forecast hours 1-144 in (a) January 2013, (b) April 2013, (c) July 2013, and (d) October 2013.

January

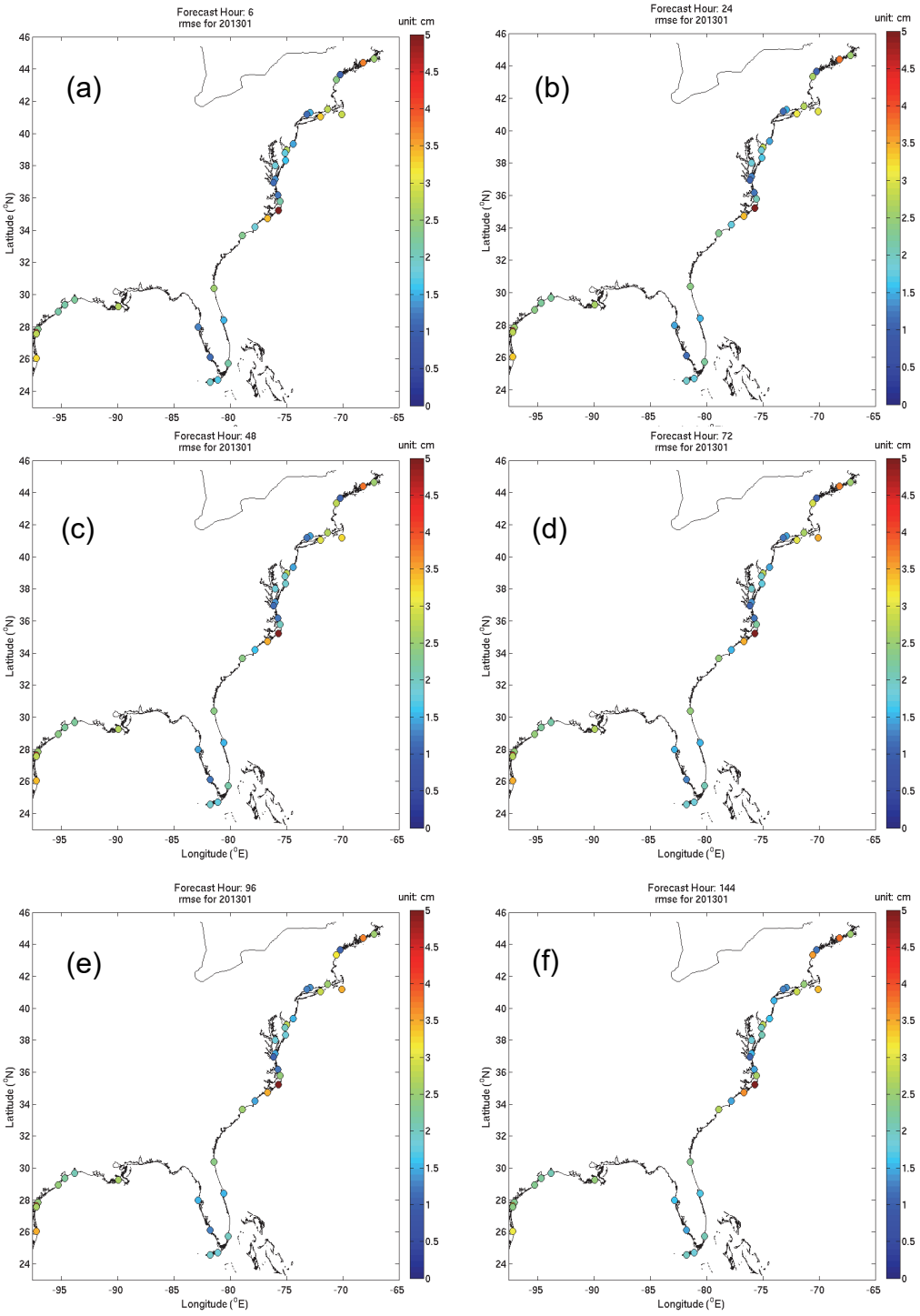


Figure 4.4. Color coded RMSE maps of the G-RTOFS SST forecast in January 2013. The six plots correspond to forecasts at hours (a) 6, (b) 24, (c) 48, (d) 72, (e) 96, and (f) 144.

April

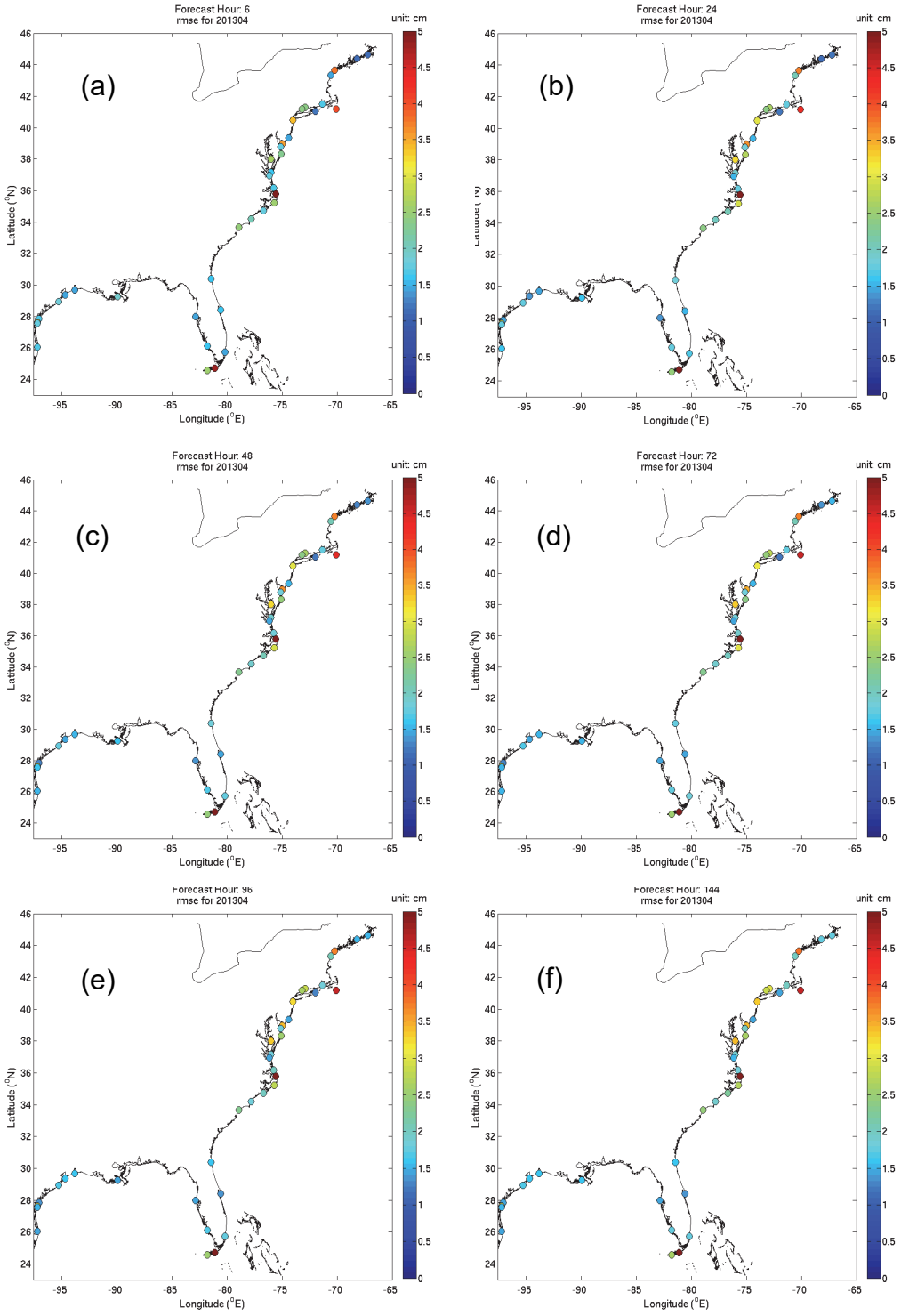


Figure 4.5. Color coded RMSE maps of the G-RTOFS SST forecast in April 2013. The six plots correspond to forecasts at hours (a) 6, (b) 24, (c) 48, (d) 72, (e) 96, and (f) 144.

July

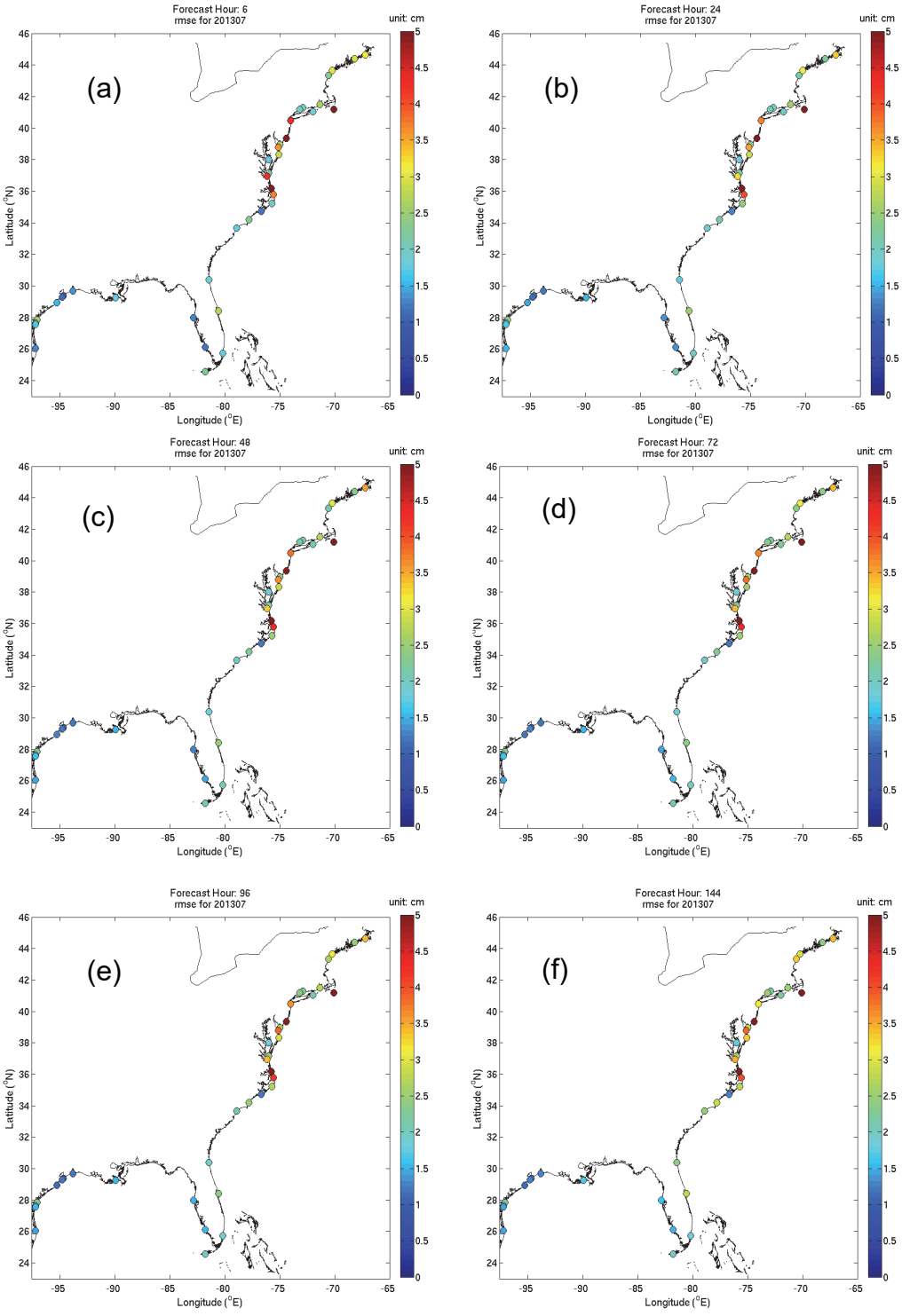


Figure 4.6. Color coded RMSE maps of the G-RTOFS SST forecast in July 2013. The six plots correspond to forecasts at hours (a) 6, (b) 24, (c) 48, (d) 72, (e) 96, and (f) 144.

October

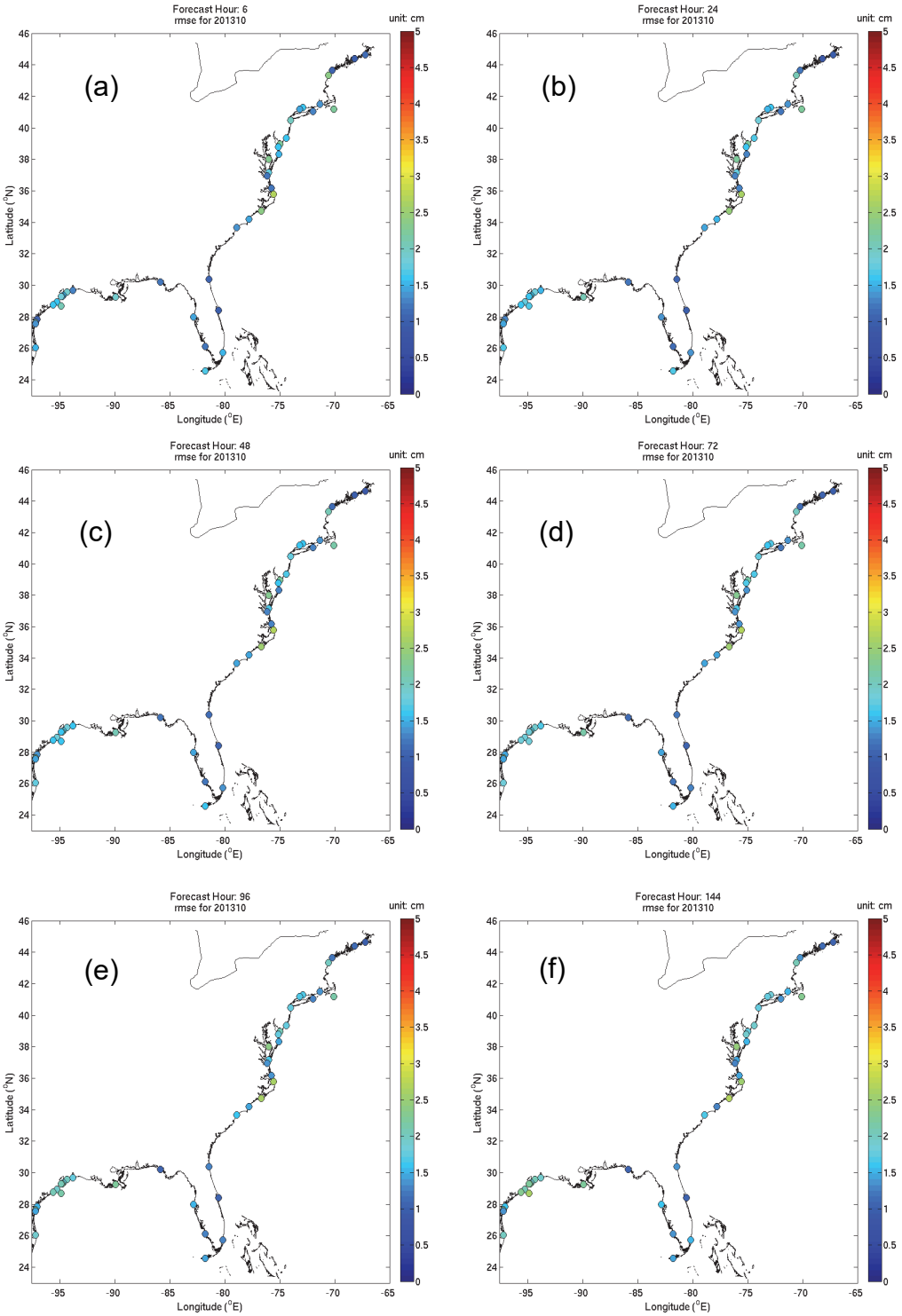


Figure 4.7. Color coded RMSE maps of the G-RTOFS SST forecast in October 2013. The six plots correspond to forecasts at hours (a) 6, (b) 24, (c) 48, (d) 72, (e) 96, and (f) 144.

4.2. Compared with NDBC Buoy Observations

We compared the G-RTOFS SST forecast with measurements at 78 NDBC buoys. These NDBC buoys are depicted in Figure 1.7. Table C.1 lists the station IDs, names, and geographical locations in longitude and latitude. In the following, we describe the results using both the forecast cycle (FC) and the forecast hour (FH) based methods.

4.2.1. Forecast Cycle Based Method

Figures 4.8(a)-(d) display RMSE maps for January, April, July, and October of 2013, respectively. All four months show a similar range of amplitude and spatial variability of RMSE. The RMSE was around 1 to 4 ºC for January and around 1 to 5 ºC for April. The RMSE was highest during July at around 1 to 5.5 ºC with a station averaged RMSE of 1.53 ºC. The RMSE subsided in October to about 1 to 4 ºC and a station averaged RMSE of 0.73 ºC.

Figures 4.9 (a)-(d) show the bias and standard deviation of model errors at 78 NDBC stations in January, April, July, and October of 2013. The typical model-data difference ranged between -4 and 4 ºC with some variation by month. The station averaged model bias was highest in January at 0.64 ºC and lowest in April at -0.14 ºC.

The corresponding values of STD display a large variation in size. The majority of values in January are quite small, but there are a few stations with rather large values of 2 or 3 ºC. This trend continues throughout the year and could be due to outliers in the observed data. The standard deviations in April are quite small at the majority of stations. The range of standard deviation values is greater in July, with the largest values being in the range of 2 ºC or more. The station averaged STD for July is 0.79 ºC. Standard deviation values are the smallest of any month in October with a station averaged value of 0.39 ºC.

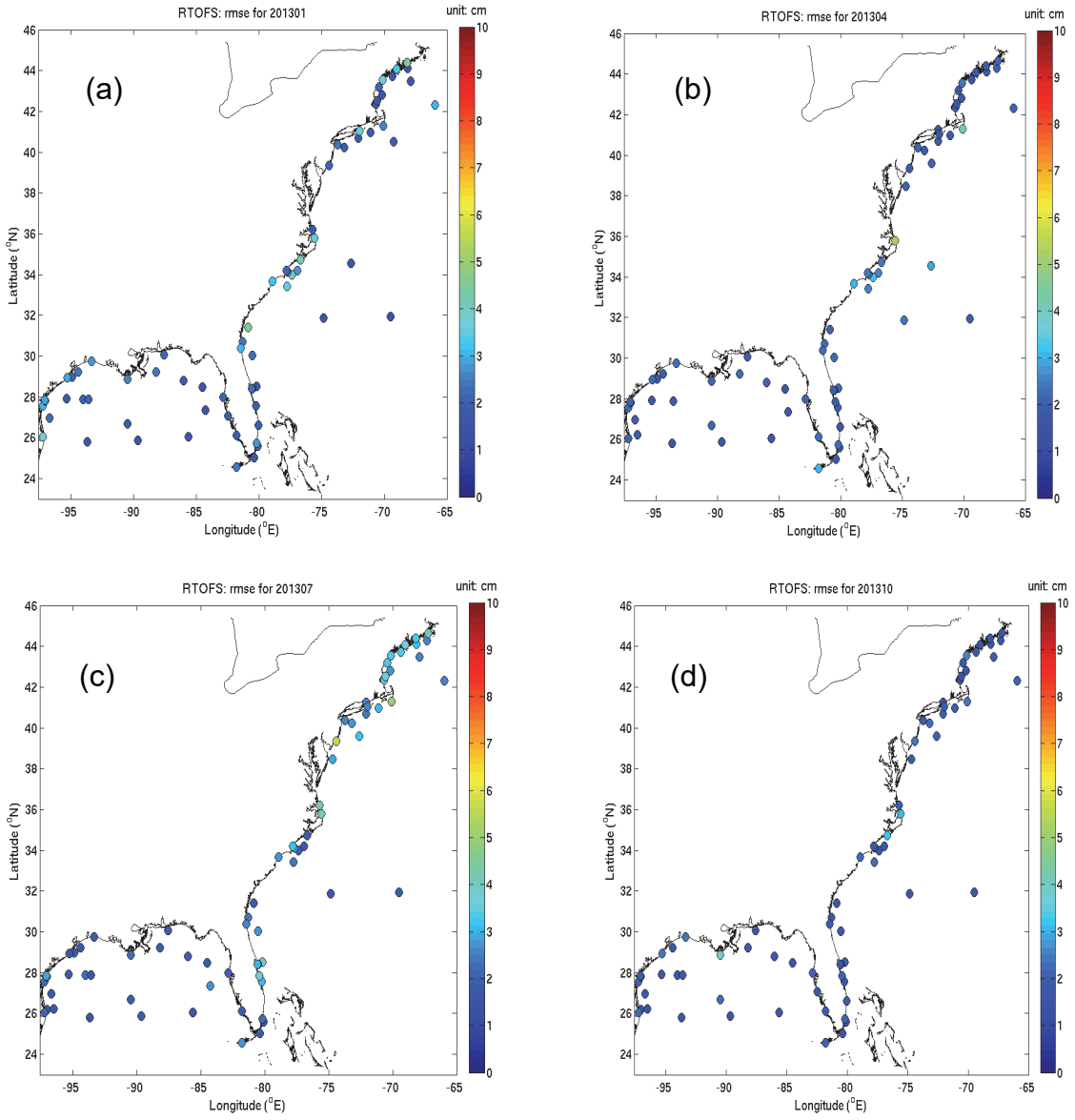
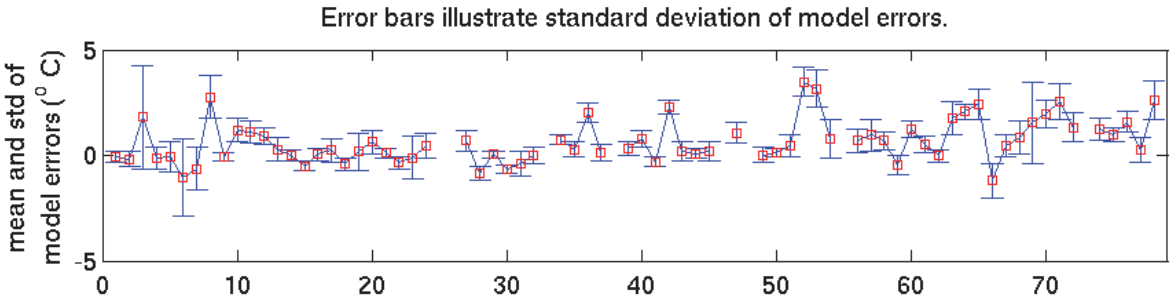


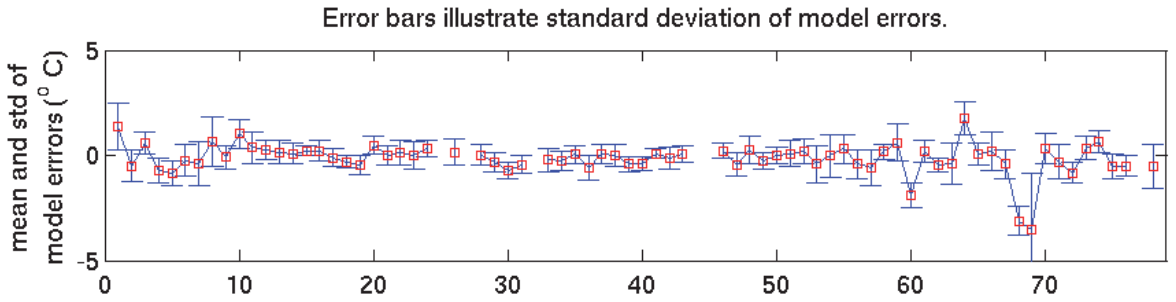
Figure 4.8. Color coded RMSE maps of the G-RTOFS SST forecast in (a) January 2013, (b) April 2013, (c) July 2013, and (d) October 2013 at 78 NDBC stations (Table C.1).

01/2013



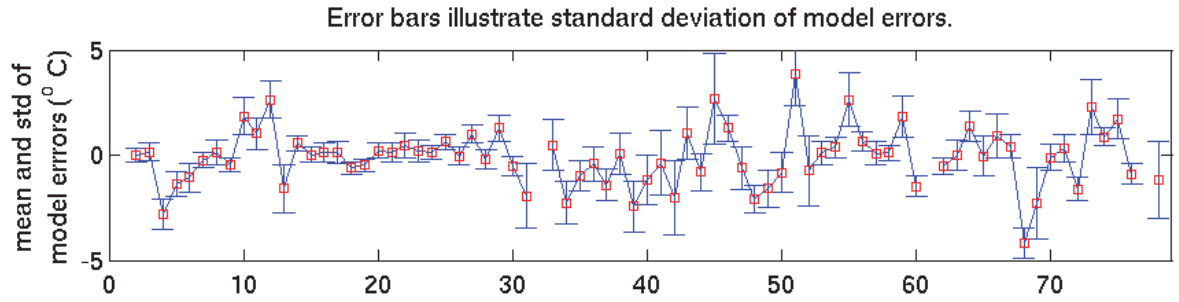
(a)

04/2013



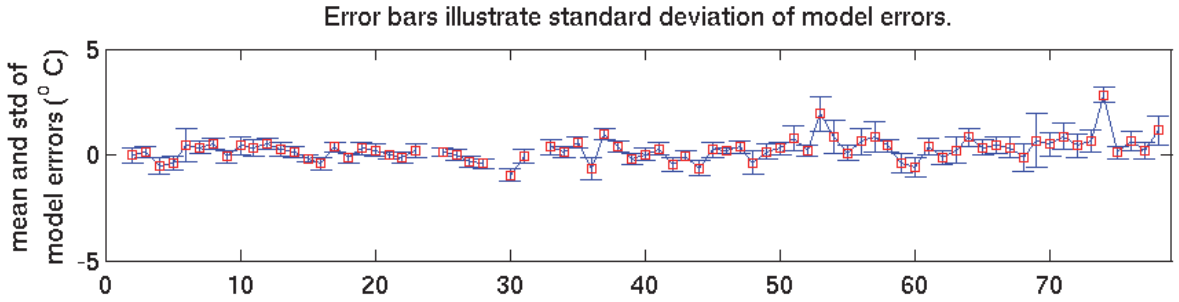
(b)

07/2013



(c)

10/2013



(d)

Figure 4.9. Means (red squares) and standard deviations (blue error bars) of the G-RTOFS SST errors in (a) January 2013, (b) April 2013, (c) July 2013, and (d) October 2013 at 78 NDBC stations (Table C.1).

4.2.2. Forecast Hour Based Method

In this section, we discuss the evolution of the model performance with increasing forecast hour. We first describe the results in terms of station-averaged RMSE and then investigate RMSE at the individual stations.

Station-averaged RMSE Figure 4.10 displays the RMSE averaged over all 78 stations at each forecast hour in (a) January 2013, (b) April 2013, (c) July 2013, and October 2013. In all four months, the RMSE increases gradually. The increase is, however, on the order of 0.1 to 0.2 ^{o}C. The sinusoidal oscillation is clear in all four months as well. The amplitude of the oscillation ranges between 0.05 and 0.15 ^{o}C.

RMSE at Individual Stations We now examine skill at the individual stations. Figures 4.11 – 4.14 display RMSE maps for all stations in (a) January 2013, (b) April 2013, (c) July 2013, and (d) October 2013. Six plots display snap shots of RMSE maps at forecast hours 6, 24, 48, 72, 96, and 144.

The RMSE maps for individual forecast hours reveal consistent results across all four months as the RMSE values increase gradually from hour 0 to hour 144 (Figure 4.10). Across all four months, the deterioration of model skill with forecast hour is not that marked. The feature which appears to stand out most is the spatial variability. There is a cluster of stations off the Carolina coast which appear to have an RMSE of 1 to 2^{o} C higher than for the remainder of the stations, but somewhat less so for October. There is a northeastern station which appears to be problematic as well, but only for January. In both April and July, there is a station off the coast of North Carolina which appears to be problematic.

Model skill was highest in October. Skill was moderate for April and July, and least for January.

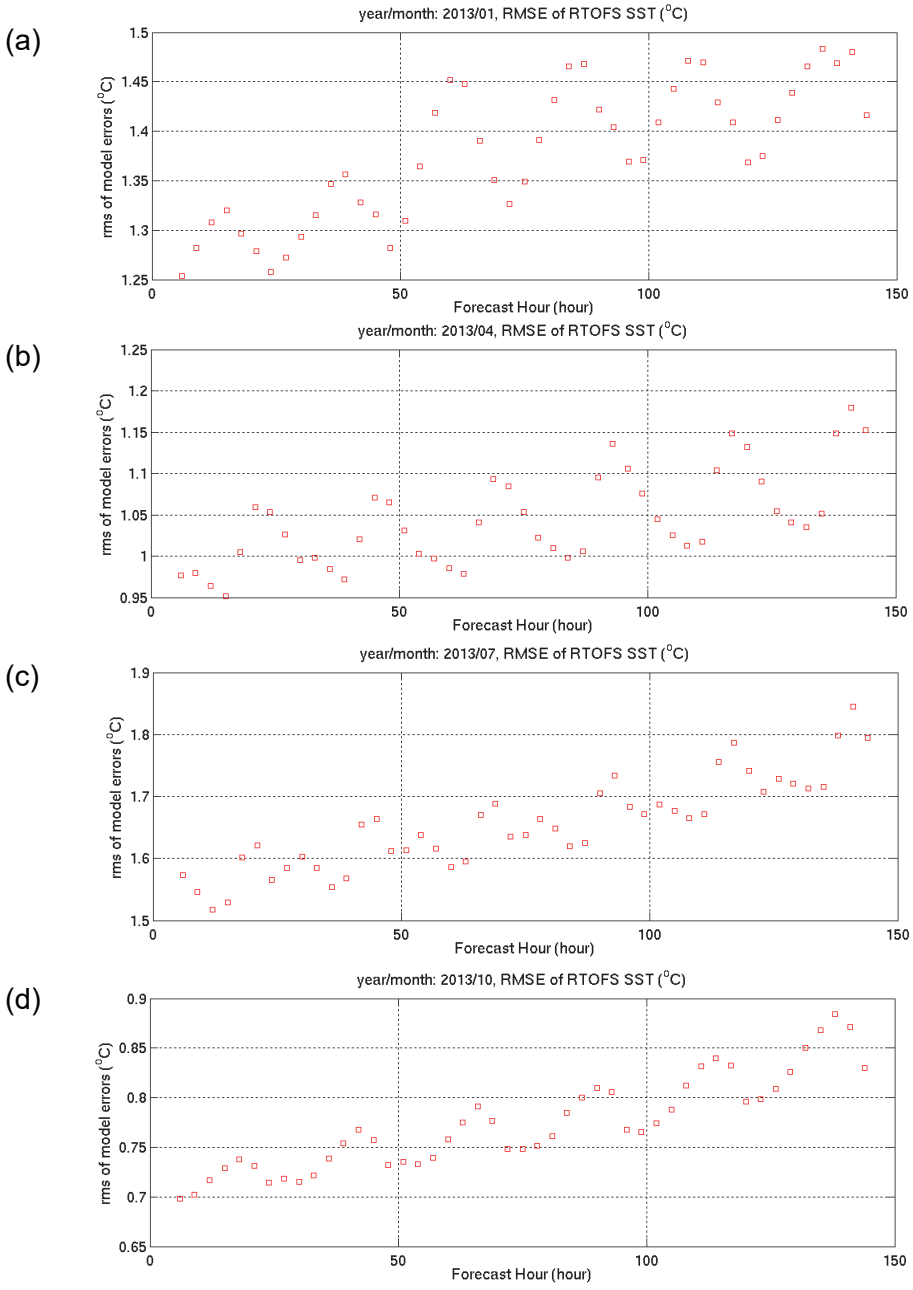


Figure 4.10. Root mean squared errors (RMSE) of the G-RTOFS SST forecast at forecast hours 1-144 in (a) January 2013, (b) April 2013, (c) July 2013, and (d) October 2013.

January

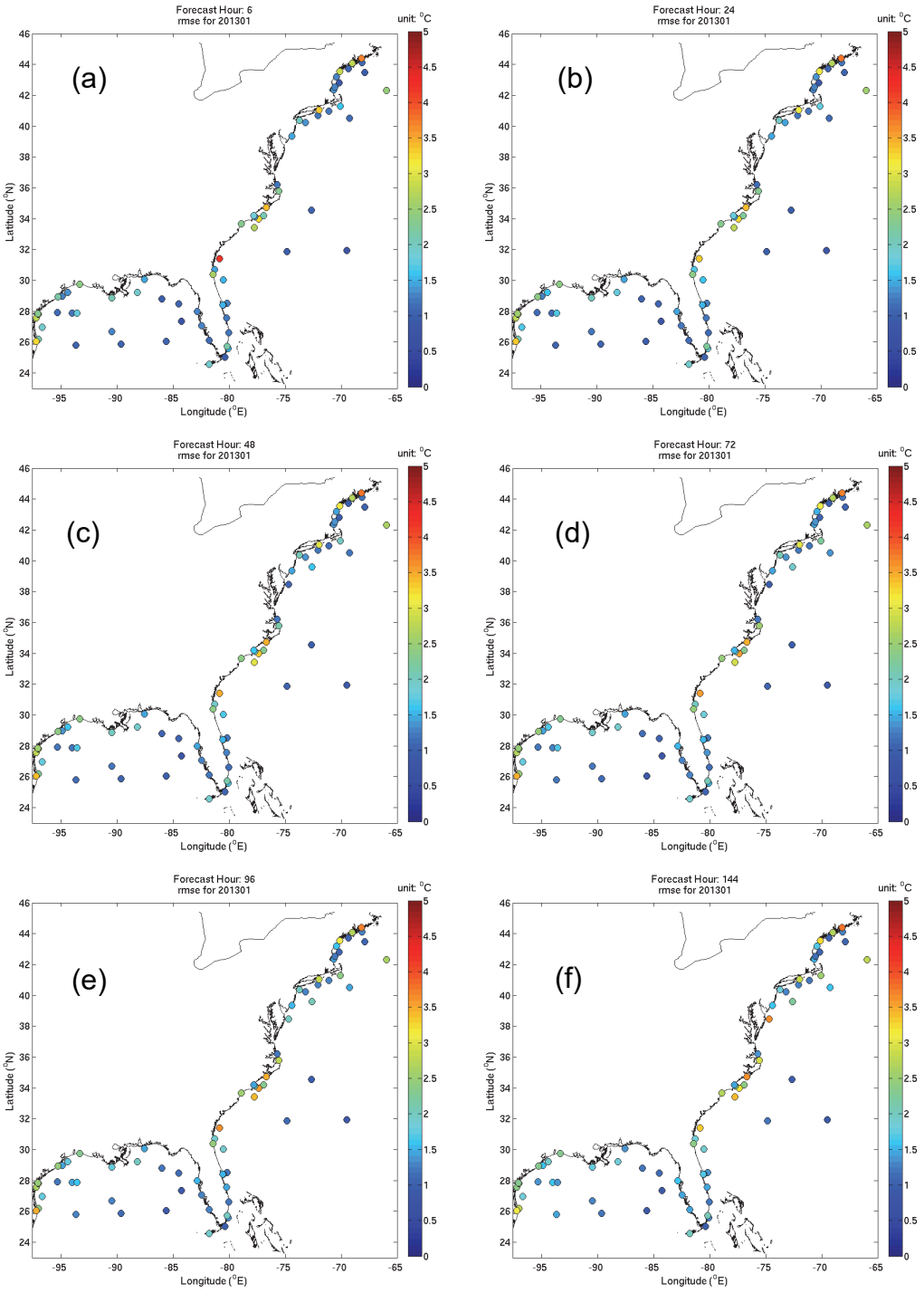


Figure 4.11. Color coded RMSE maps of the G-RTOFS SST forecast in January 2013. The model-data comparison was made against SST measured at 78 NDBC buoys (Table C.1). The six plots correspond to forecasts at hours (a) 6, (b) 24, (c) 48, (d) 72, (e) 96, and (f) 144.

April

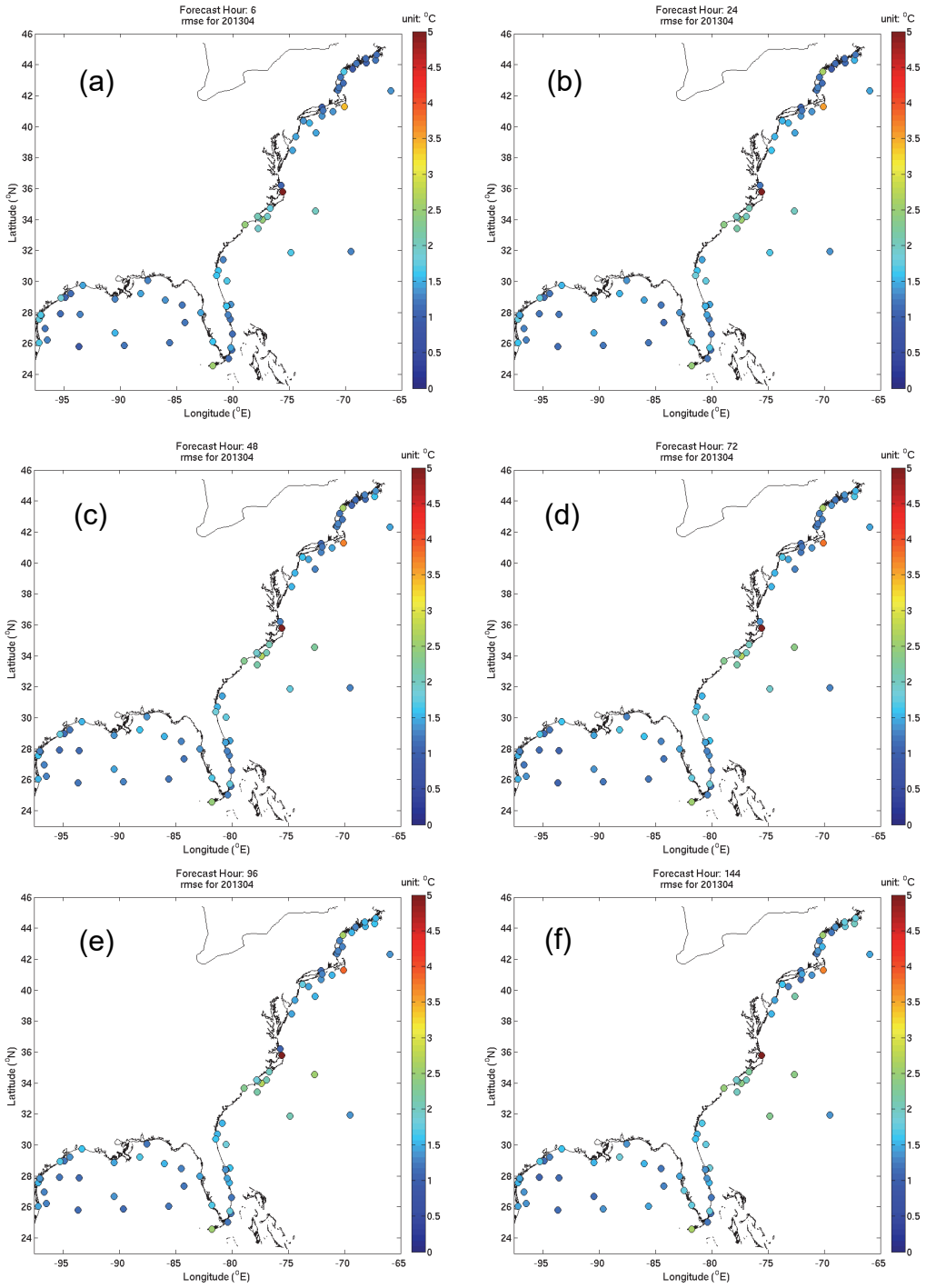


Figure 4.12. Color coded RMSE maps of the G-RTOFS SST forecast in April 2013. The model-data comparison was made against SST measured at 78 NDBC buoys (Table C.1). The six plots correspond to forecasts at hours (a) 6, (b) 24, (c) 48, (d) 72, (e) 96, and (f) 144.

July

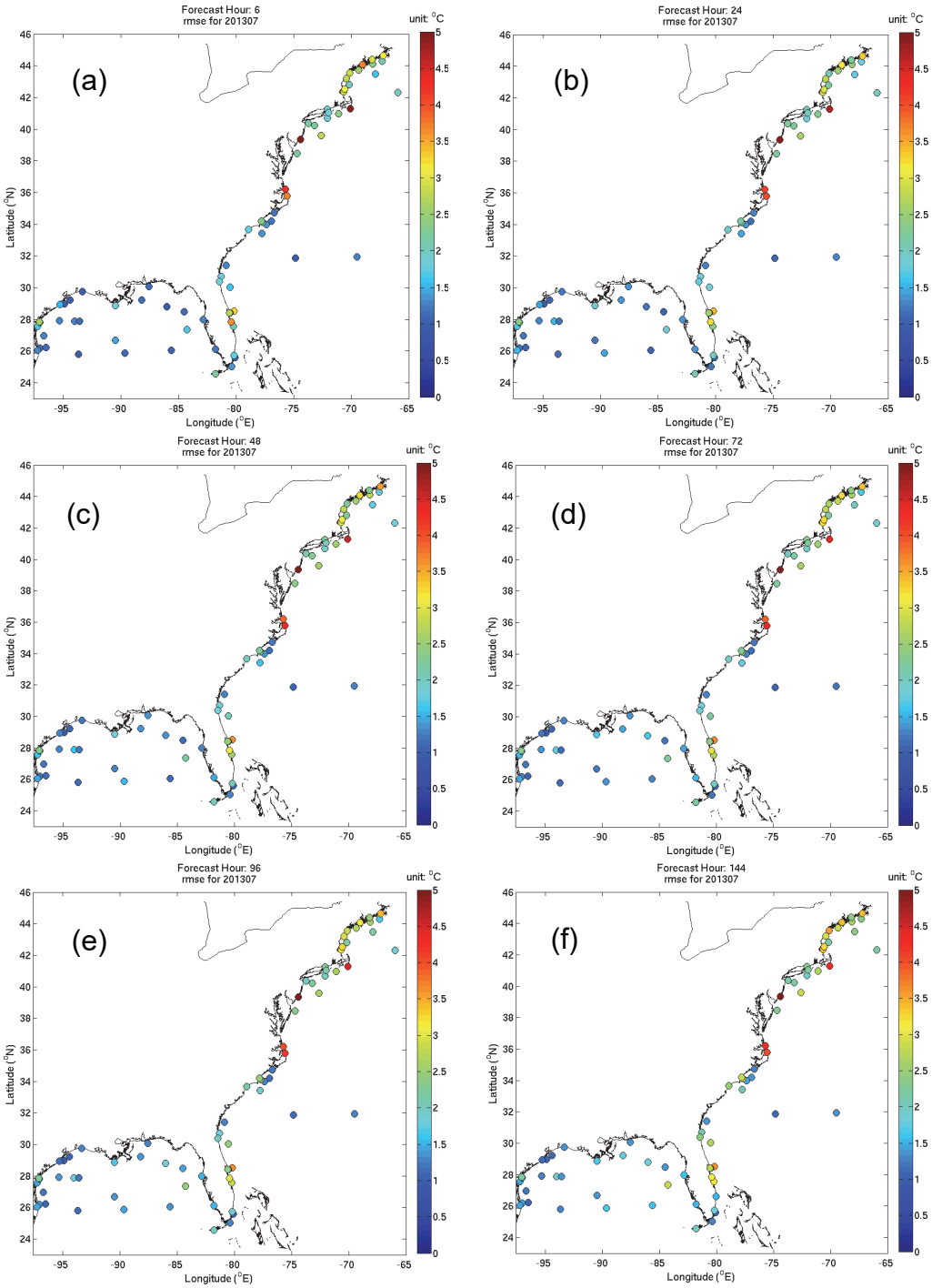


Figure 4.13. Color coded RMSE maps of the G-RTOFS SST forecast in July 2013. The model-data comparison was made against SST measured at 78 NDBC buoys (Table C.1). The six plots correspond to forecasts at hours (a) 6, (b) 24, (c) 48, (d) 72, (e) 96, and (f) 144.

October

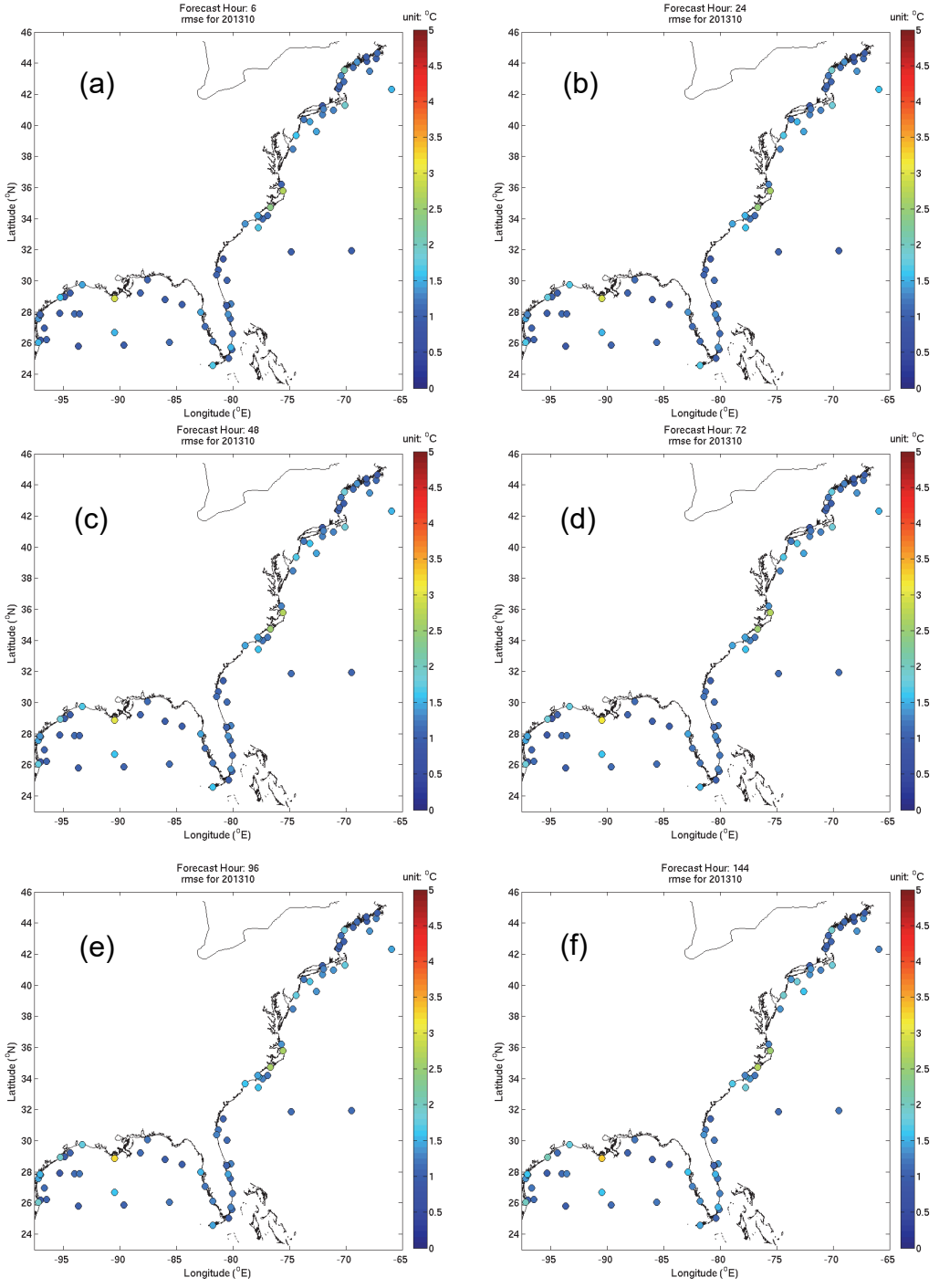


Figure 4.14. Color coded RMSE maps of the G-RTOFS SST forecast in October 2013. The model-data comparison was made against SST measured at 78 NDBC buoys (Table C.1). The six plots correspond to forecasts at hours (a) 6, (b) 24, (c) 48, (d) 72, (e) 96, and (f) 144.

4.3. Compared with WOA09 Data

In the following, we compare the G-RTOFS with the WOA09 SST using only the forecast cycle (FC) based method. Since WOA09 represents climatological monthly means of oceanic quantities, any comparison with this database reflects model skill in a climatological sense.

4.3.1. Forecast Cycle (FC) Based Method

Figures 4.15(a)-(d) display maps of the model root-mean-square error for January, April, July, and October of 2013. The model exhibits much greater RMSE in July 2013 than in the non-summer months. October is the next highest month. The RMSE was typically less than 6 ºC in January and April. In July, the RMSE rose to higher than 10 ºC and almost as high in October. The mean RMSE was 8.88 ºC for July and 6.35 ºC for October.

The RMSE displayed spatial variability across all four months. It displayed a large spatial variability in July 2013. The range of RMSE values was almost as great for October. The range of RMSE values was much less, and the RMSE values were quite a bit less as well for January and April of 2013. In January, there appear to be several outlier stations with high RMSE values around 8 ºC, but generally, the RMSE does not go above 6ºC. The mean RMSE was 2.63 ºC for January. In a climatological sense, the G-RTOFS had much better skill in the winter and spring than in the summer and fall.

Figures 4.16(a)-(d) display the mean and standard deviation of model errors at 72 WOA09 data points. Table C.1 lists the geographical locations in longitude and latitude for these data points. The mean model-data difference ranges from -10 to 5 ºC in January, and from -5 to 4 ºC in April. In both cases, the mean bias is fairly close to 0: -0.80 ºC for January and 0.26 ºC for April. In July, there is a large change as bias values range from 0 to over 10 ºC, a dramatic shift towards a positive bias. The mean bias for July is 8.76 º C. This positive bias is again evident in October as all station bias values are over 0 ºC. Overall, the range of model bias is much greater when the G-RTOFS results are compared with WOA09 than when compared with either the CO-OPS observations or with the NDBC observations.

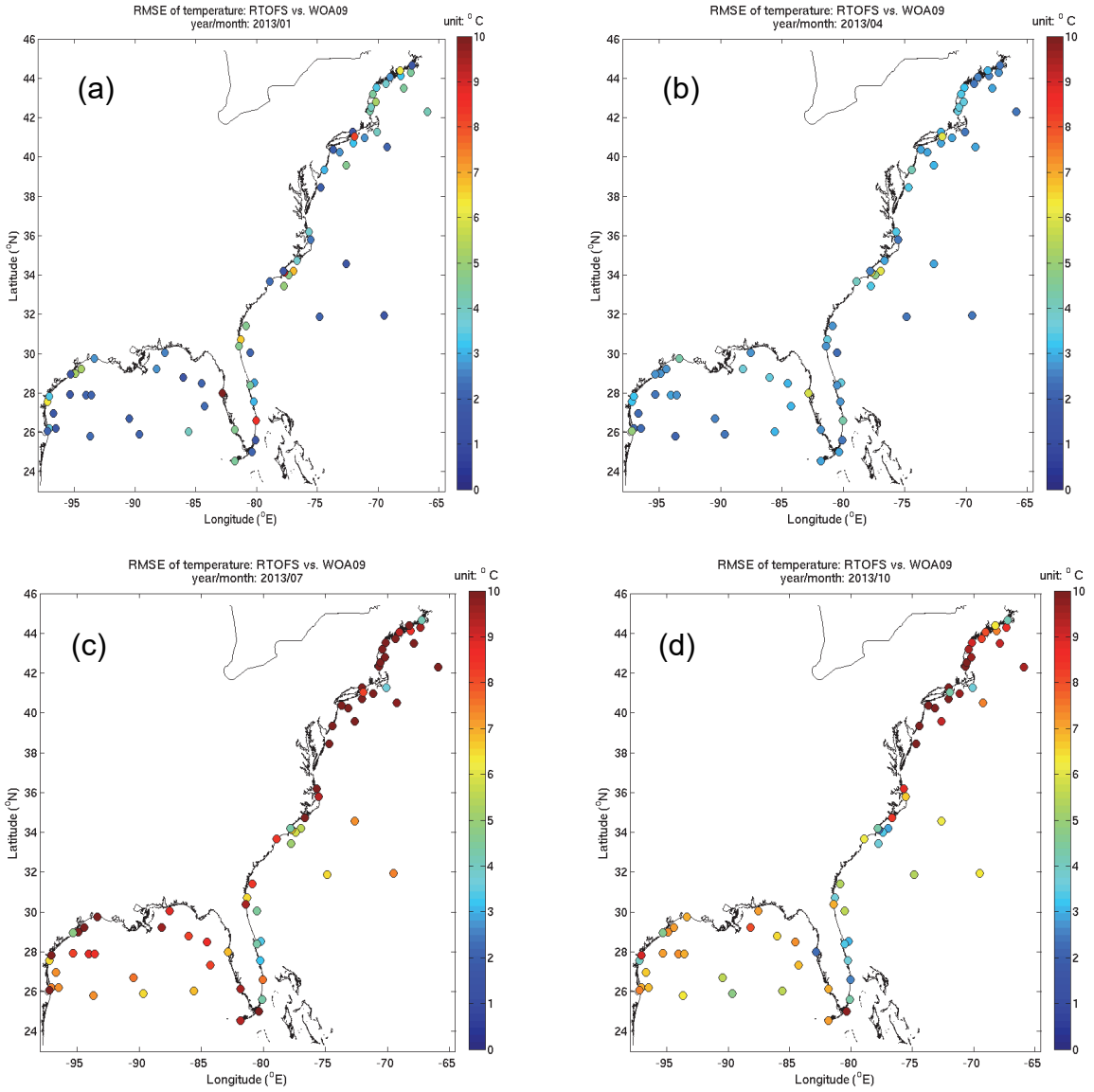


Figure 4.15. Color coded RMSE maps of the G-RTOFS SST forecast in (a) January 2013, (b) April 2013, (c) July 2013, and (d) October 2013 at 72 WOA09 data grid points (Table C.1).

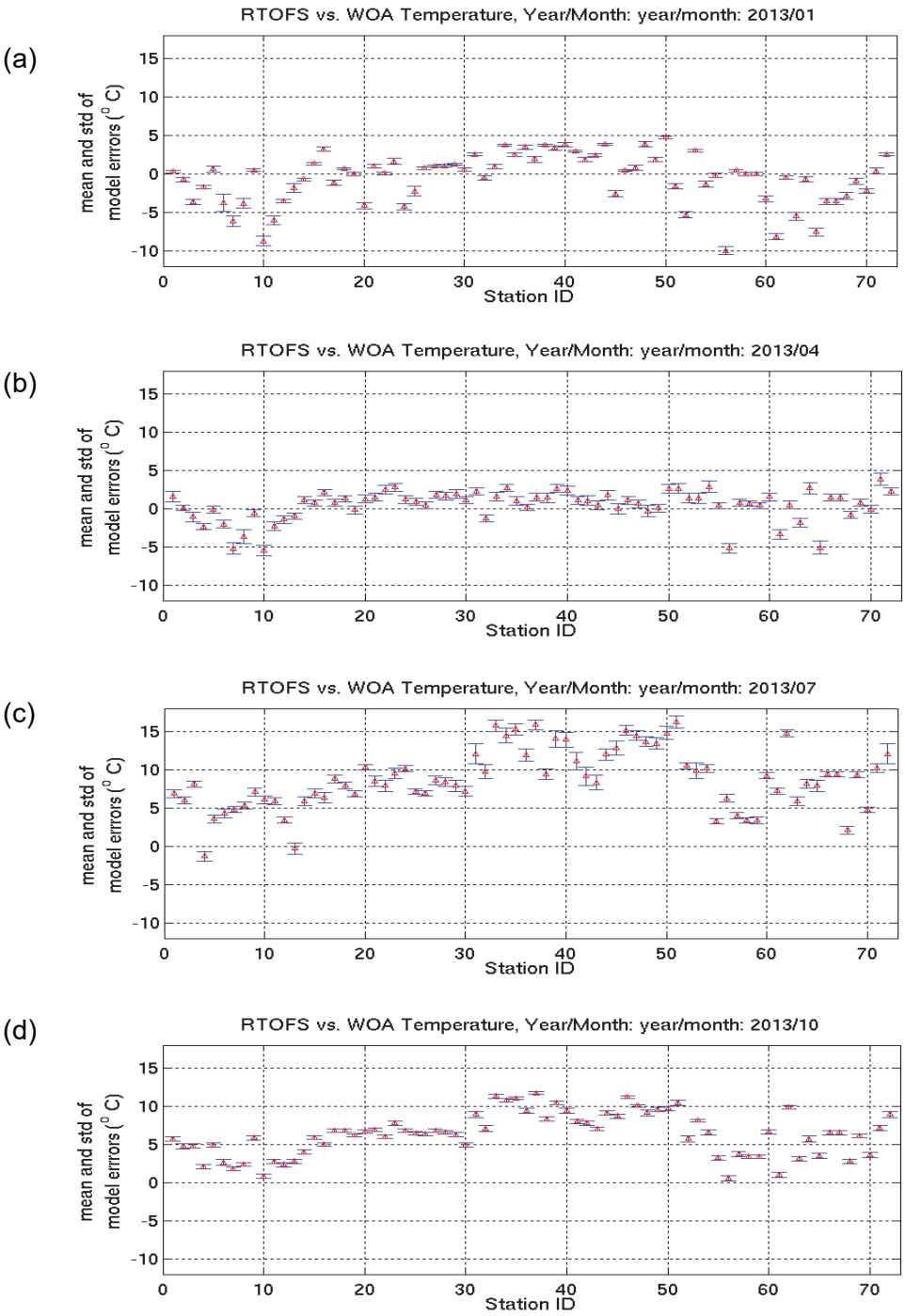


Figure 4.16. Means (red squares) and standard deviations (blue error bars) of the G-RTOFS SST errors in (a) January 2013, (b) April 2013, (c) July 2013, and (d) October 2013 at 72 WOA09 data grid points. The horizontal axis shows the station numbers (Table C.1).

4.4. Summary

We evaluated the performance of the G-RTOFS SST forecast guidance by comparing the model results with observed data from CO-OPS physical oceanography observation stations, NDBC buoys, and the WOA09 database. Overall, the RMSE from the WOA09 comparison reveals the largest magnitudes and seasonal variability of model error.

Table 4.1 lists the station averaged bias, standard deviation, and root-mean-squared error for the winter, spring, summer, and fall estimated using the FC based method. The bias of G-RTOFS with respect to the CO-OPS physical oceanography observation stations is variable throughout the year, though it does tend positive. It varies between -2 and 5 ºC in the winter, and between -5 and 3 ºC in the spring. The bias reaches its most negative station averaged value during the spring at -0.79 ºC. The bias varies between -8 and 5 ºC in the summer, with one station (number 5) having the anomalous value of -8 ºC. The range of bias values is more limited in the fall with the average value at 0.62 ºC. G-RTOFS vs. the NDBC buoy data demonstrates similar variability with regard to model bias. The model bias reaches almost 5 ºC in the winter, and overall leans slightly positive throughout the year. The model bias with respect to WOA09 locations reaches extreme values of greater than 15 ºC at a number of stations during the summer, and as high as 12 ºC during the fall. The model bias reaches negative values of less than -10 ºC during the winter. These extreme examples of model bias require further investigation.

The standard deviations of the G-RTOFS results vs. the CO-OPS physical oceanography observation stations range between 1 and 3 ºC, with the larger deviations occurring in the spring and summer. The standard deviations of the G-RTOFS results vs. the NDBC buoy data tend to be small in both winter and spring, but there are a few outliers reaching perhaps 3 ºC. In the summer, the standard deviations grow larger with a station average value of 0.79 ºC. The standard deviations of the G-RTOFS results vs. the WOA09 climatology are quite small throughout the year, never much larger than 1 ºC.

The RMSE of G-RTOFS vs. the CO-OPS physical oceanography observation stations displays somewhat greater magnitude and spatial variability in summer than during the other seasons. The RMSE is no higher than 6 ºC in January, and ranges from about 1.5 to 7 ºC in July. The station averaged value of RMSE is 2.05 ºC in July. The RMSE of G-RTOFS vs. the NDBC buoy stations was highest during July as well, with RMSE values at around 1.5 to 5.5 ºC. The station averaged $RMSE_{G\text{-}RTOFS}$ was 1.53 º C. The RMSE of G-RTOFS vs. the WOA09 climatology displayed spatial variability across all four months. The largest spatial variability occurred in July 2013; many stations along the Atlantic coast and in the Gulf of Mexico had an RMSE approaching 10ºC or more. The range of RMSE values was almost as great for October. In January, the RMSE generally does not go above 6ºC. Overall, when compared with observed data, the RMSE values seem quite large for G-RTOFS vs. the WOA09 climatology.

Model skill does not appear to degrade rapidly as the forecast hour projects into the future. This result differs from the case of the G-RTOFS SSH forecast, where model skill gradually degrades with increasing forecast hour (Section 3.2).

Table 4.1 Forecast cycle-based statistics of the G-RTOFS SST errors.

Data sources	**CO-OPS**			**NDBC**			**WOA09**		
	bias (ºC)	std (ºC)	rmse (ºC)	bias (ºC)	std (ºC)	rmse (ºC)	bias (ºC)	std (ºC)	rmse (ºC)
Winter (Jan. 2013)	1.18	0.72	1.80	0.64	0.57	1.25	-0.80	0.32	2.63
Spring (Apr 2013)	-0.79	0.86	1.70	-0.14	0.62	0.98	0.26	0.54	2.12
Summer (July 2013)	0.39	0.99	2.05	-0.10	0.79	1.53	8.76	0.62	8.88
Fall (Oct 2013)	0.62	0.64	1.18	0.24	0.39	0.73	6.33	0.28	6.35
Mean	0.35	0.80	1.68	0.16	0.59	1.12	3.64	0.44	5.00

5. PERFORMANCE OF G-RTOFS FOR SEA-SURFACE SALINITY

In this chapter, we discuss the performance of the G-RTOFS sea-surface salinity (SSS) forecast. We compare the model results with both in situ observations and the WOA09 database.

5.1 Compared with NDBC Observations

We compared the G-RTOFS SSS forecast with measurements from nine NDBC buoys (Figure 1.8). Table D.1 in Appendix D lists the station meta data including IDs, names, and geographical locations in longitude and latitude. In the following, we describe the results derived from both the forecast cycle (FC) and forecast hour (FH) based methods.

5.1.1. Forecast Cycle (FC) Based Method

Figures 5.1(a)-(d) display RMSE maps for January, April, July, and October of 2013, respectively. All four months show a similar range of amplitude and spatial variability of RMSE. The RMSE was around 1 to 10 psu for January and around 1 to 10 psu again in April. The RMSE was highest during July with all but three stations near 10 psu. The RMSE subsided a bit in October especially in the Gulf where the RMSE dropped to around 6 to 7 psu. The station near the entrance to the Chesapeake Bay dropped to about 6 psu as well. The RMSE is huge in the mid-Atlantic region of the Chesapeake and Delaware Bays. Here, the RMSE is 10 psu or greater year round.

Figures 5.2(a)-(d) show the mean and standard deviation of model errors at nine NDBC stations in January, April, July, and October of 2013. The typical mean model-data difference ranges between 0 and 20 psu with some variation by month. The large model error occurred at stations 4, 5, and 6 where the model error is near 20 psu year round. The model mean is biased lower than the mean SSS of the observed data. The observed data locations are inside of the Chesapeake and Delaware Bays. Within the body of an estuarine system, the salinity will be lower. The locations of the grid cells do not precisely match the physical locations. The model calculates a value of salinity appropriate for a location outside of the estuarine, a more ocean like value of salinity.

The corresponding values of STD display little variation in size. Through all seasons, the STD is very small, except at station 9 where the STD grows to perhaps 5 psu.

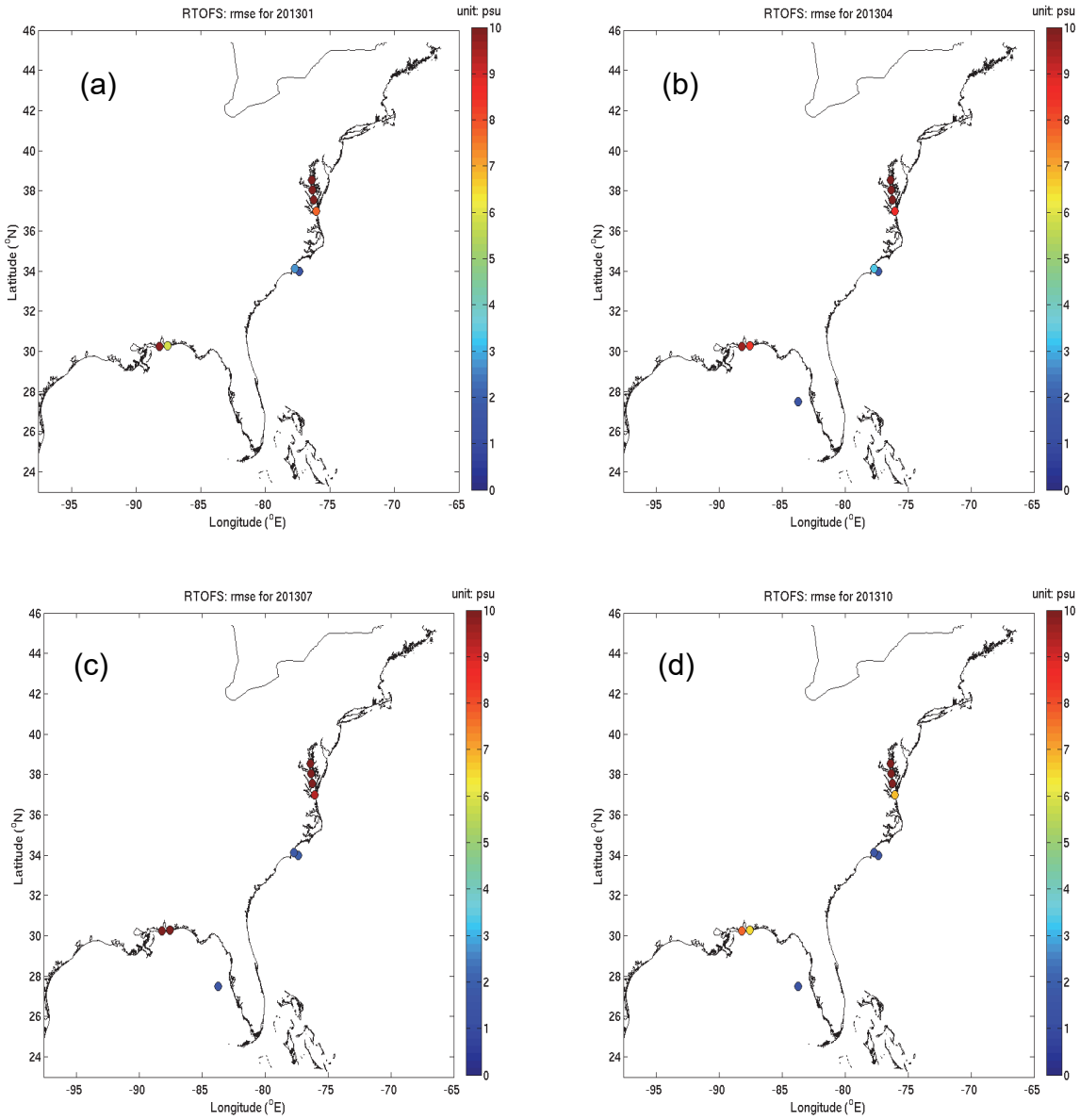
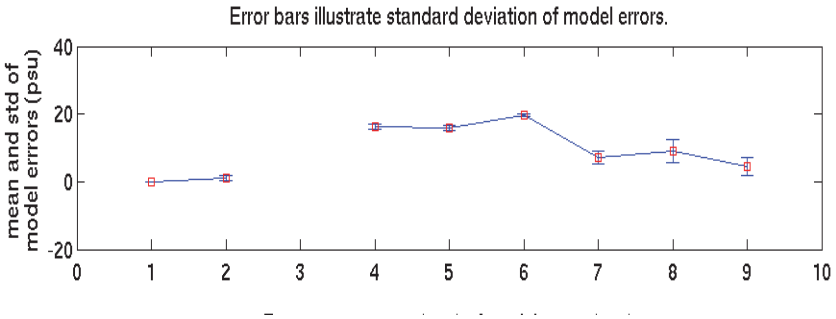


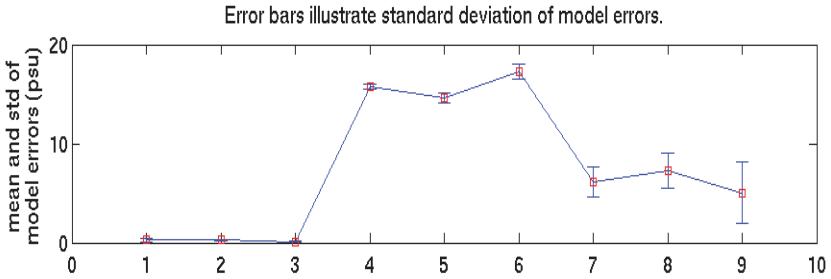
Figure 5.1. Color coded RMSE maps of the G-RTOFS SSS forecast in (a) January 2013, (b) April, 2013, (c) July 2013, and (d) October 2013 at 9 NDBC stations (Table D.1).

01/2013



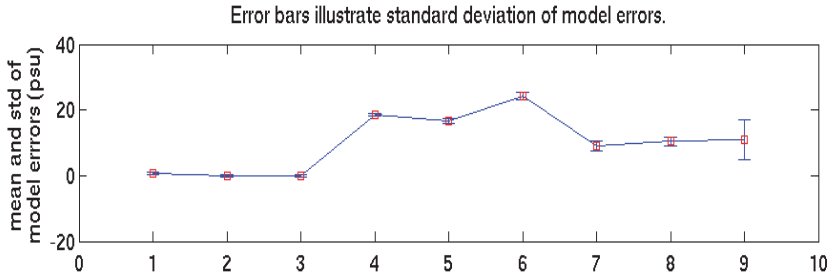
(a)

04/2013



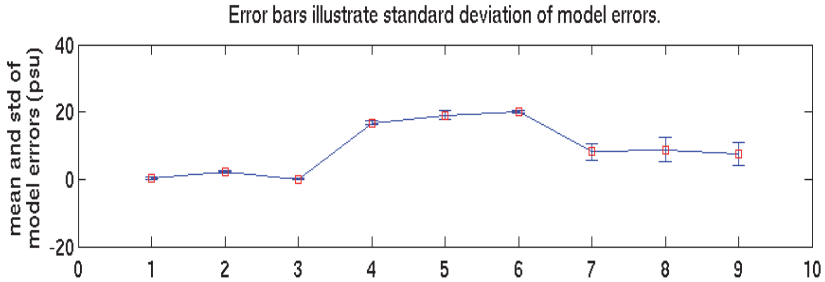
(b)

07/2013



(c)

10/2013



(d)

Figure 5.2. Means (red squares) and standard deviations (blue error bars) of the G-RTOFS SSS errors in (a) January 2013, (b) April 2013, (c) July 2013, and (d) October 2013 at 9 NDBC stations. The horizontal axis shows the station numbers (Table D.1).

5.1.2. Forecast Hour (FH) Based Method

In this section, we discuss the evolution of the model performance with increasing forecast hour. We first describe the results in terms of station-averaged RMSE and then investigate RMSE at the individual stations. For the reason described in the beginning paragraph of Section 4.3, we will focus on assessing the model skill in terms of the station-averaged RMSE.

Figures 5.3(a)-(d) display RMSE averaged over 9 NDBC stations at all 144 forecast hours in (a) January 2013, (b) April 2013, (c) July 2013, and (d) October 2013. Remarkably, RMSE remained at a nearly constant level, or in the case of April, dropped slightly through the forecast cycle. The RMSE averaged around 9.5 psu in January, 9.4 psu in April, 10.5 psu in July, and 7.5 psu in October 2013.

For the months of April, July, and October, these plots reveal the familiar sinusoidal, diurnal period, signal. The plot from January 2013 lacks this feature, however. The plotted points depict more of a spread pattern.

Station-averaged RMSE Figure 5.3 displays the RMSE averaged over all nine stations at each forecast hour in (a) January 2013, (b) April 2013, (c) July 2013, and October 2013. In all four months, the RMSE either remains constant, or counter intuitively, decreases slightly. The RMSE from January 2013, July 2013, as well as from October 2013 remains largely constant. The RMSE from April 2013 seems to decrease slightly with increasing forecast hour. The sinusoidal oscillation is clear in three of the four months as well. The amplitude of the oscillation is around 0.5 in all months except July, where it is closer to 1.0 psu.

RMSE at Individual Stations To examine skill at the individual stations, we examine figures 5.4 – 5.7 which display RMSE maps for all stations in (a) January 2013, (b) April 2013, (c) July 2013, and (d) October 2013. Six plots display snap shots of RMSE maps at forecast hours 6, 24, 48, 72, 96, and 144.

The RMSE maps for all four seasons reveal consistent results. The deterioration of model skill with forecast hour is not that marked. The feature which stands out most is the spatial variability. There is a cluster of stations in the mid-Atlantic region (Chesapeake and Delaware Bays) which have high RMSE values and two stations along the Gulf coast with high RMSE values. The RMSE values for psu of 5 or greater probably require further investigation.

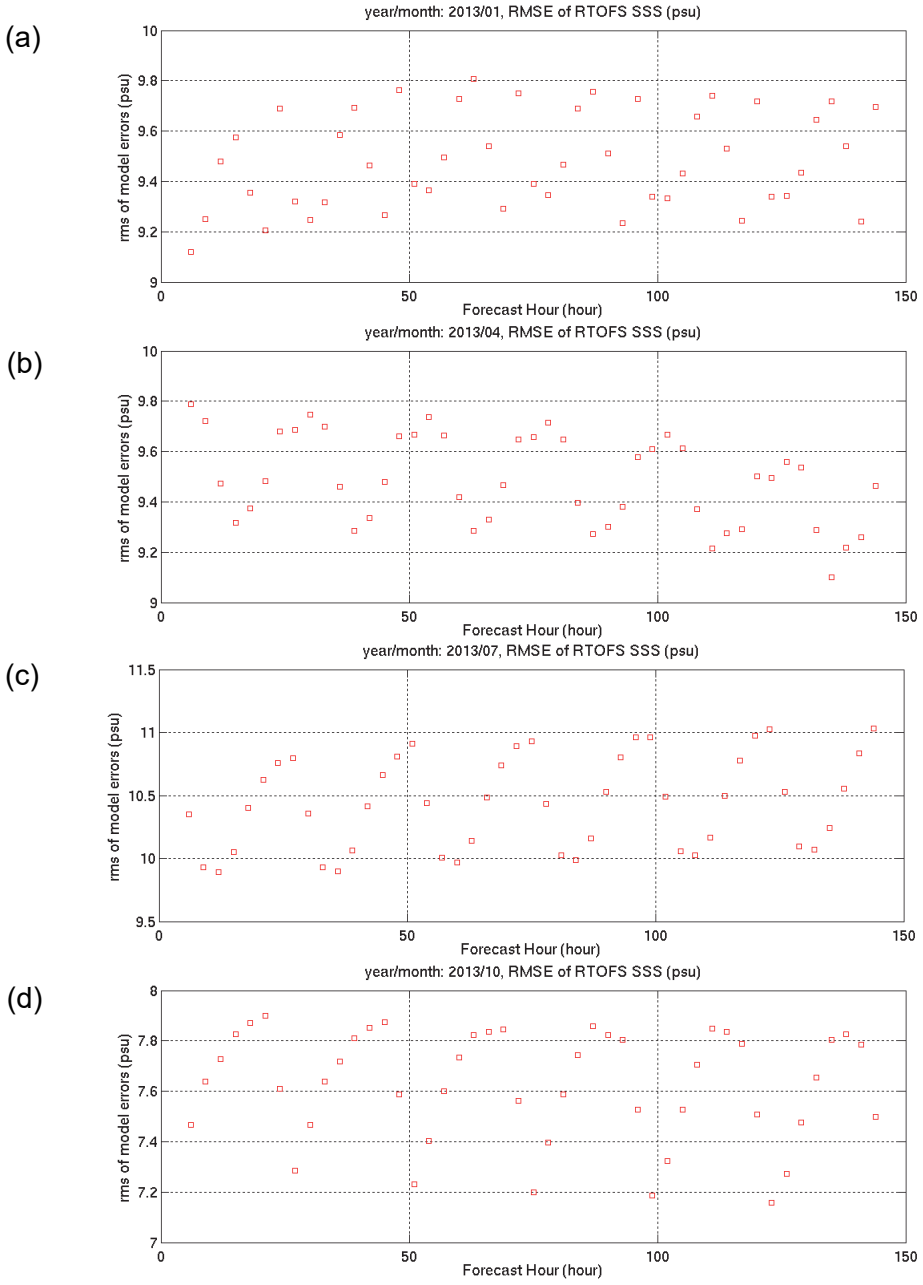


Figure 5.3. Root mean squared errors (RMSE) of the G-RTOFS SSS forecast at forecast hours 1-144 in (a) January 2013, (b)April 2013, (c) July 2013, and (d) October 2013.

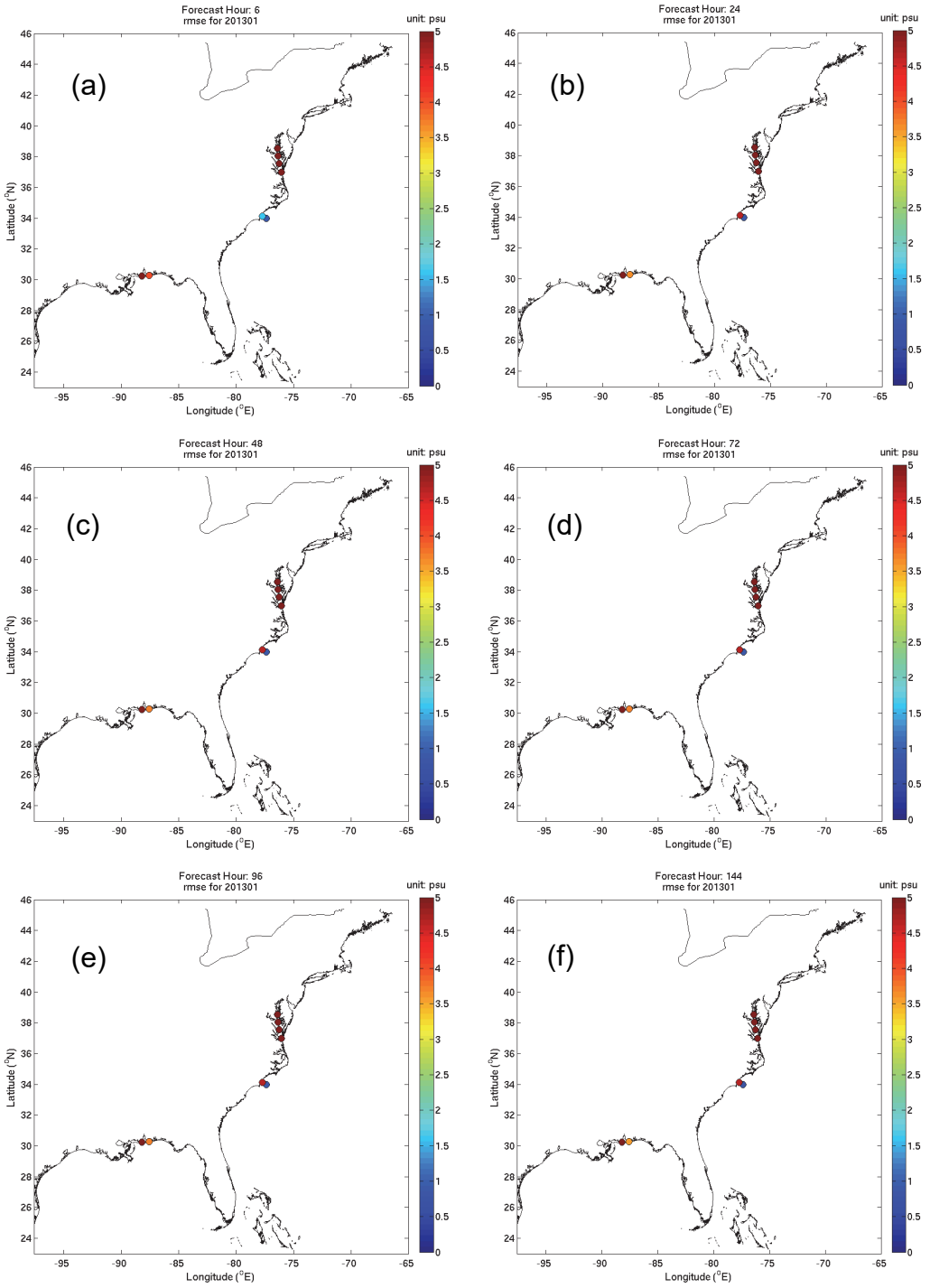


Figure 5.4. Color coded RMSE maps of the G-RTOFS SSS forecast in January 2013. The model-data comparison was made against SSS measured at 9 NDBC buoys (Table D.1). The six plots correspond to forecasts at hours (a) 6, (b) 24, (c) 48, (d) 72, (e) 96, and (f) 144.

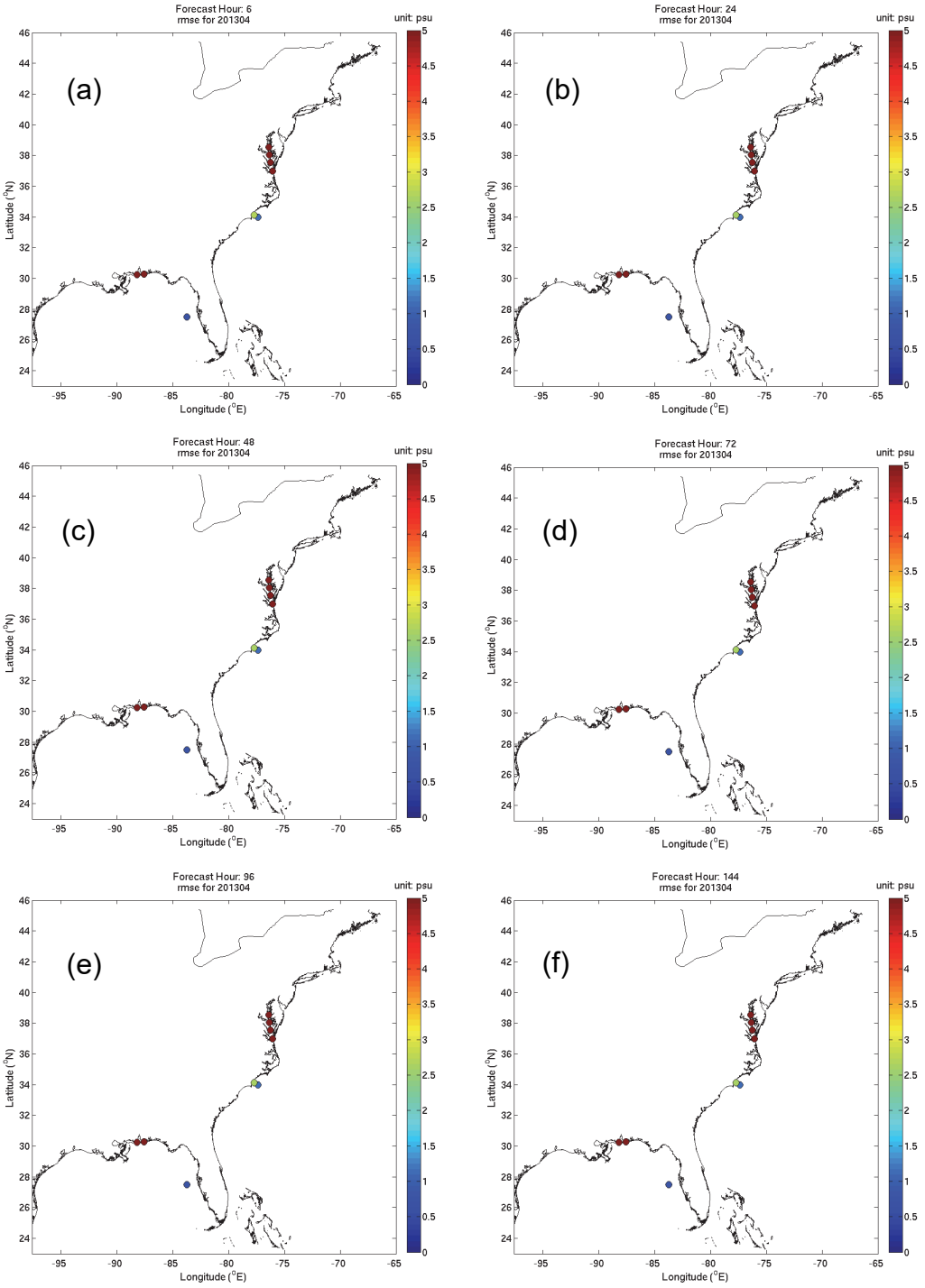


Figure 5.5. Color coded RMSE maps of the G-RTOFS SSS forecast in April 2013. The model-data comparison was made against SSS measured at 9 NDBC buoys (Table D.1). The six plots correspond to forecasts at hours (a) 6, (b) 24, (c) 48, (d) 72, (e) 96, and (f) 144.

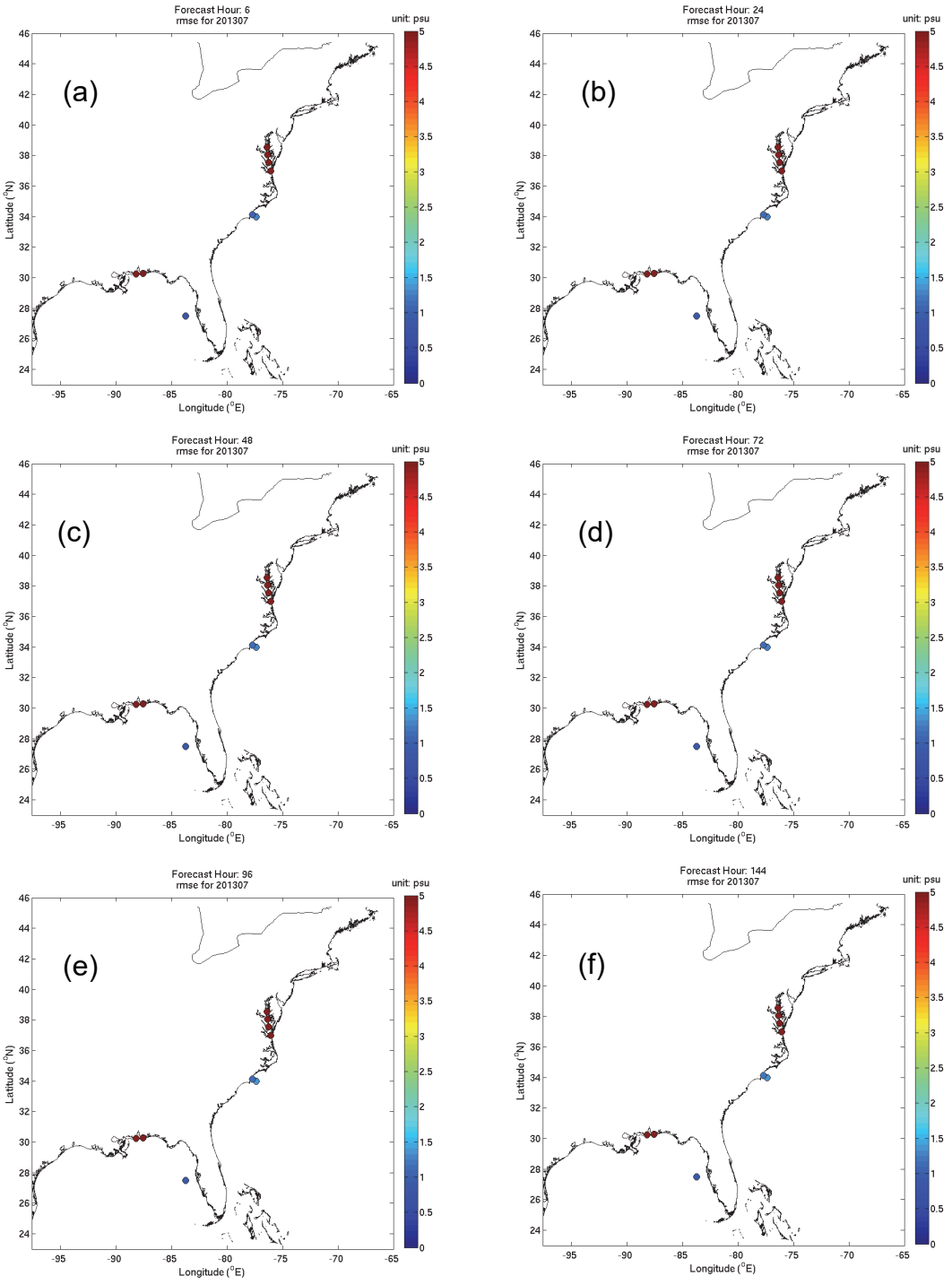


Figure 5.6. Color coded RMSE maps of the G-RTOFS SSS forecast in July 2013. The model-data comparison was made against SSS measured at 9 NDBC buoys (Table D.1). The six plots correspond to forecasts at hours (a) 6, (b) 24, (c) 48, (d) 72, (e) 96, and (f) 144.

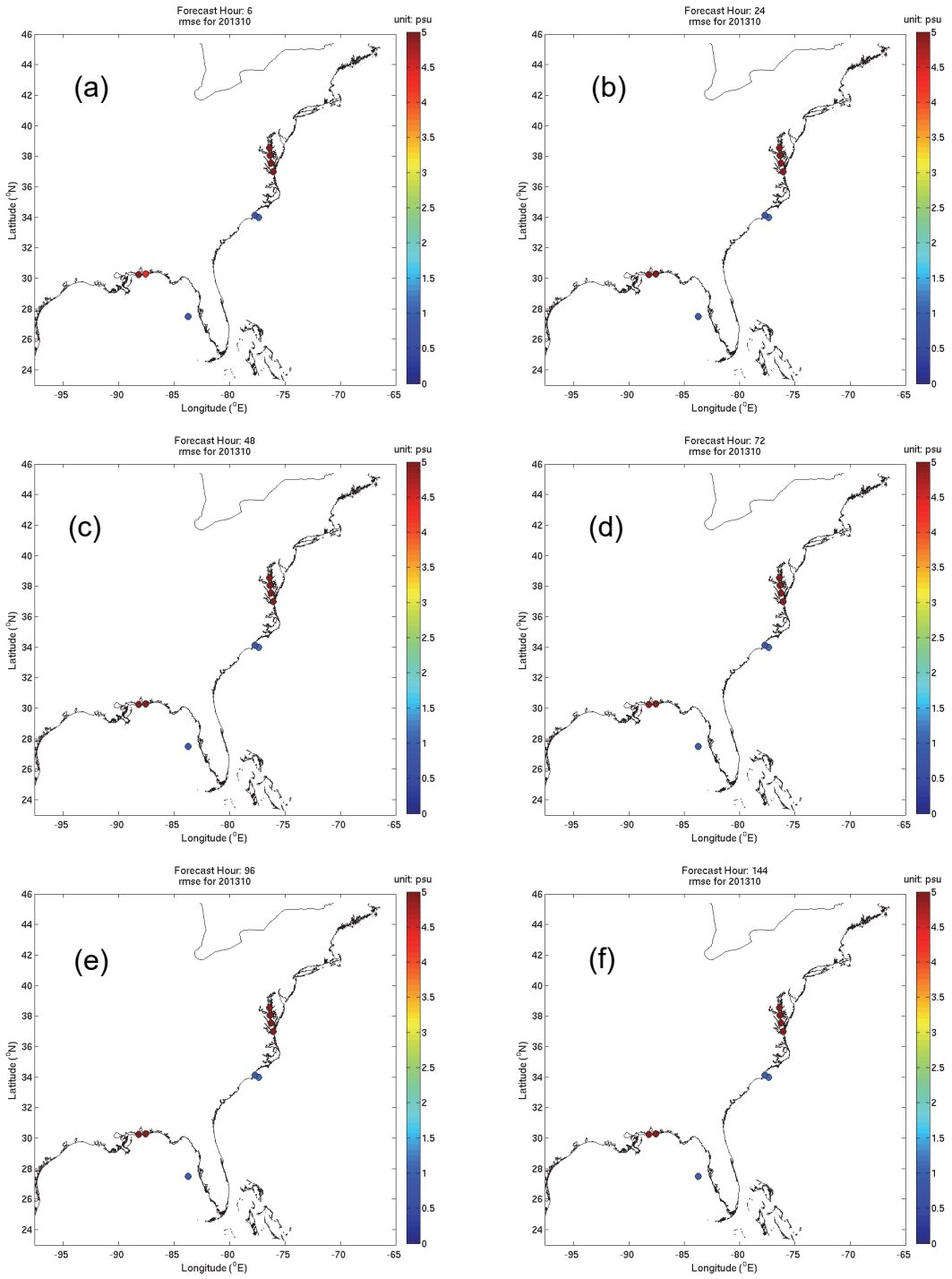


Figure 5.7. Color coded RMSE maps of the G-RTOFS SSS forecast in October 2013. The model-data comparison was made against SSS measured at 9 NDBC buoys (Table D.1). The six plots correspond to forecasts at hours (a) 6, (b) 24, (c) 48, (d) 72, (e) 96, and (f) 144.

5.2. Compared with WOA09 Data

Here, we use the climatological monthly mean SSS in WOA09. Hence, any comparisons with this database reflect the model skill in a climatological sense. Since we are comparing the model results with climatological data sets, it would not make sense scientifically to address details of the forecast hour based model performance. As a result, we will investigate only the FC based method.

Figures 5.8(a)-(d) show the maps of the model root-mean-square error in January, April, July, and October of 2013. The RMSE was, in general, remarkably steady both spatially and across the four seasons. The RMSE was generally under 4 psu and at many stations in the 1 to 2 psu range.

Figures 5.9(a)-(d) display the mean and standard deviation of model errors at 72 WOA09 stations. During July, the model means are remarkably close to 0 at most stations with five stations having outlier values. The greatest amount of deviation between stations occurs in January where there are about three clear cut outliers, and the non-outliers having around a 1.5 psu range of deviation. The results for April and October are similar to the results for July except for a slightly greater range of deviation amongst the non-outliers.

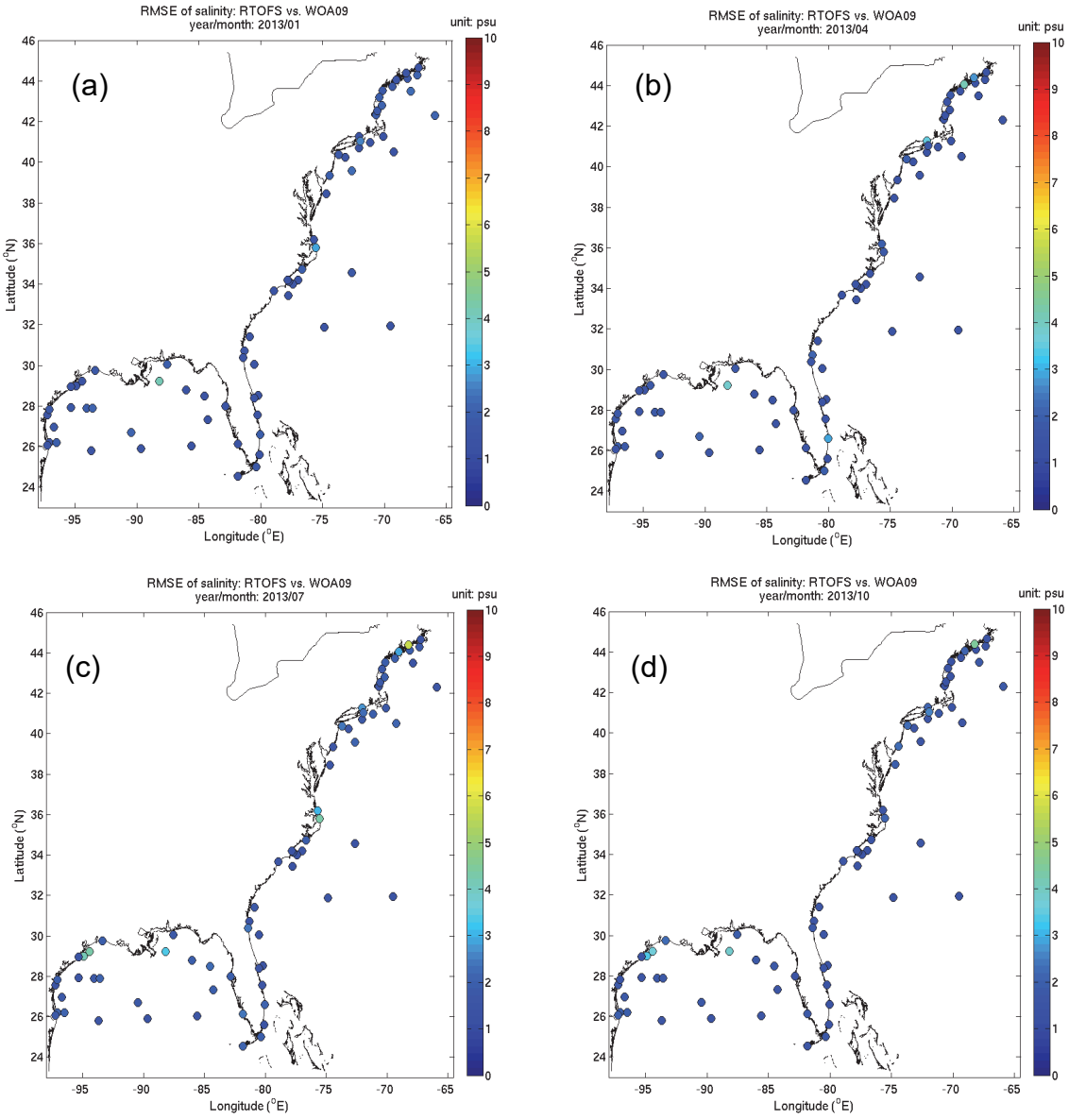
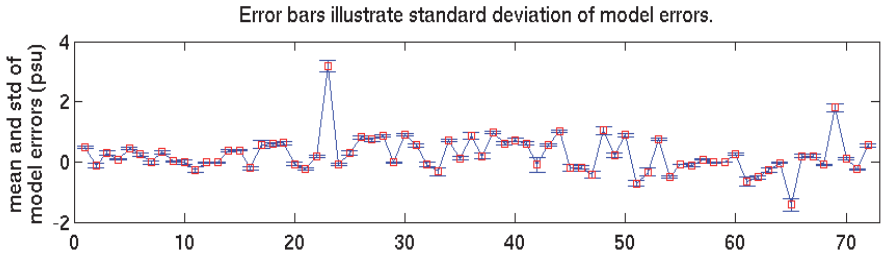


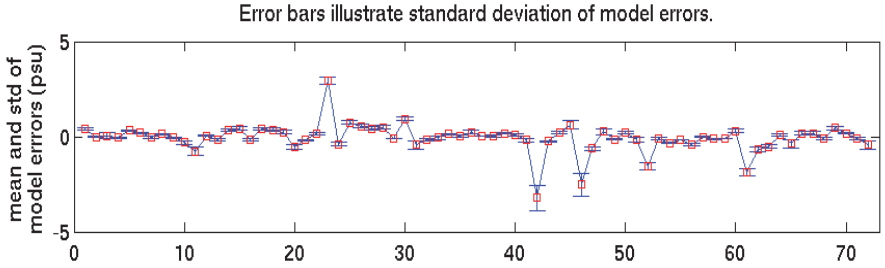
Figure 5.8. Color coded RMSE maps of the G-RTOFS SSS forecast in (a) January 2013, (b) April 2013, (c) July 2013, and (d) October 2013 at 72 WOA09 data grid points (refer to Appendix C for station locations).

01/2013



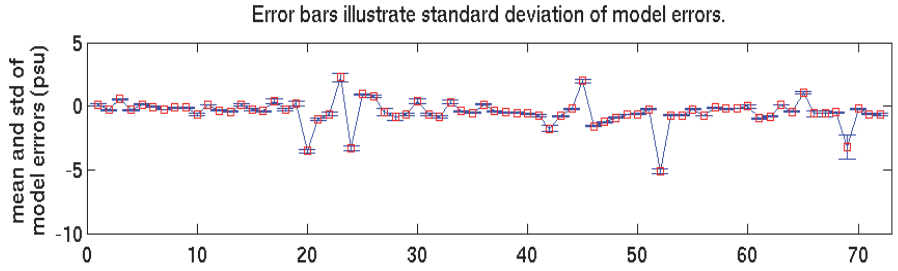
(a)

04/2013



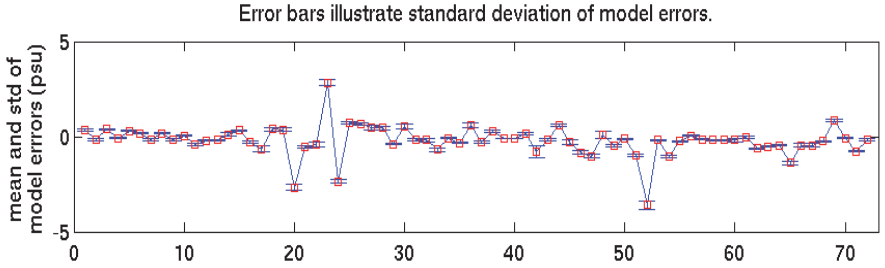
(b)

07/2013



(c)

10/2013



(d)

Figure 5.9. Means (red squares) and standard deviations (blue error bars) of the G-RTOFS SSS errors in (a) January 2013, (b) April 2013, (c) July 2013, and (d) October 2013 at 72 WOA09 data grid points. The horizontal axis shows the station numbers (refer to Appendix C for station locations).

5.3. Summary

The SSS forecast guidance of G-RTOFS demonstrates poor agreement with data from the nine NDBC buoy stations. The station bias ranges between 0 and 20 psu. The large RMSE values occur at stations 4,5, and 6. These stations are located within the estuaries of the mid-Atlantic. The station averaged RMSE is highest during the summer at about 10.5 psu and lowest during the fall at about 7.5 psu.

The SSS forecast guidance of G-RTOFS demonstrates a satisfactory agreement with the WOA09 database. The station bias ranges between -4 and 4 psu through all four seasons, including outlier values. The STD is uniformly small.

6. SUMMARY

The Operational Nowcast/Forecast Systems created by the NOS of NOAA use sea-surface height, temperature, and salinity outputs from the Global Real-Time Ocean Forecast System, as well as sea-surface height from the ETSS and from ESTOFS when G-RTOFS SSH is not available. These forecast outputs are used in preparing open ocean boundary forcing conditions. To support future development of the NOS OFS in eastern U.S. coastal waters, and in the Gulf of Mexico, we assessed the performance of SSH forecasts from the G-RTOFS, ETSS, and ESTOFS and the forecasts of sea-surface temperature and sea-surface salinity from the G-RTOFS.

This report summarizes the procedures utilized and the results which were obtained. Chapter 1 describes background information concerning the set up and operation of the G-RTOFS, the ETSS, and the ESTOFS. The chapter also explains how the observed data (water level, SST, and SSS) used in model-data comparisons is obtained and how it is processed. Chapter 2 explains technical details of the two methods used to assess model performance: the forecast cycle (FC) based method and the forecast hour (FH) based method. Chapters 3-5 discuss the model performance of SSH (from G-RTOFS, ETSS, and ESTOFS), SST (from G-RTOFS), and SSS (from G-RTOFS) forecast guidance, respectively.

Our intention is to investigate both the temporal (seasonal) variability and the spatial variability of model performance. We chose four monthly periods: January, April, July, and October of 2013. These four months align roughly with the seasons of winter, spring, summer, and fall.

We assessed the model performance from two perspectives. The first was to assess the performance across the entire forecast cycle over multiple forecast cycles. The second was to assess performance with respect to a single forecast hour, over multiple forecast cycles. These two perspectives correspond to the methods described in Chapter 2: the forecast cycle (FC) based method and the forecast hour (FH) based method. The FC based method calculates bias and root-mean-squared error (RMSE) across each forecast cycle. The FH based method assesses how the model performance evolves as a forecast cycle progresses through the hours of the forecast cycle. This method calculates bias and RMSE of model results at each forecast hour over multiple forecast cycles.

G-RTOFS displays a very large negative water level bias year round reaching a magnitude of 75 cm at one station during the spring. The station averaged bias value for G-RTOFS is highest during the spring at -53.6 cm. The year round negative bias is indicative of the datum issue that G-RTOFS has being a global model. In contrast, the model bias from ETSS is moderate year round. The station averaged bias is -7.6 cm in the winter and -9.1 cm in the fall. That bias value increases to -16.0 cm in the summer, and increases again to -21.3 cm in the fall. The range of bias values that ESTOFS displays through the seasons is of similar magnitude to that of ETSS, and is also negative.

The RMSE of ETSS ($RMSE_{ETSS}$) appears to be moderate year round. The $RMSE_{ETSS}$ ranges from around 10 to 21 cm in the winter, and from around 18 to 36 cm in the fall. The one station with a value of 36 cm appears to be an outlier. The station averaged value of RMSE is 22.2 cm. The RMSE of G-RTOFS is of an order of magnitude more than double that of ETSS. $RMSE_{G\text{-}RTOFS}$ ranges from approximately 25 to 75 cm during July and from around 30 to 65 cm during October. The station averaged RMSE value for G-RTOFS during July is 49.5 cm compared with a value of

16.9 cm for ETSS. The ESTOFS RMSE is comparable to that of ETSS; it ranges from around 16 to 30 cm in the winter and around 18 to 30 cm in the fall.

With respect to the forecast hour, the skill of all three models degrades as a forecast progresses into the later hours of the cycle. $RMSE_{ETSS}$ increases monotonically from less than 1 cm at the beginning of a forecast cycle to around 8 to 14 cm near the end of the cycle. The RMSE of the G-RTOFS SSH forecast increases monotonically from less than 2 cm at the start of a forecast cycle, to around 12 to 18 cm at the end of the cycle. Model skill exhibits seasonal variability. The maximum $RMSE_{ETSS}$ are 14 cm, 12 cm, 8 cm, and 11.5 cm in the winter, spring, summer, and fall, respectively. The maximum $RMSE_{G\text{-}RTOFS}$ are 18 cm, 15 cm, 12 cm, and 14 cm which correspond to the winter, spring, summer, and fall.

The performance of G-RTOFS SST forecast guidance demonstrates some seasonal variability. The RMSE of G-RTOFS vs. the CO-OPS physical oceanography observation stations displays somewhat greater magnitude and spatial variability in summer than during the other seasons. The RMSE is no higher than 6 ºC in January, and ranges from about 1.5 to 7 ºC in July. The RMSE of the G-RTOFS vs. NDBC buoy stations was highest during July as well, with RMSE values at around 1.5 to 5.5 ºC. The RMSE of G-RTOFS vs. the WOA09 climatology displayed spatial variability across all four months. The largest spatial variability occurred in July 2013 where many stations along the Atlantic coast and in the Gulf of Mexico had an RMSE approaching 10ºC or more. The range of RMSE values was almost as great for October. In January, the RMSE generally does not go above 6ºC. Overall, when compared with observed data, the RMSE values seem rather large for G-RTOFS surface temperature no matter which data source is used for comparison. The performance is more satisfactory in the winter and spring than in the summer and fall. The agreement between model results and WOA09 is very poor, especially during summer and fall.

The magnitude of bias of G-RTOFS SST values vs. observed is highest when compared with values from the WOA09 database. The model bias reaches extreme values of greater than 15 ºC at a number of stations during the summer, and as high as 12 ºC during the fall. The bias reaches negative values of less than -10 ºC during the winter. These extreme examples of model bias require further investigation. Aside from the WOA09 comparison, the only anomalous values for bias are from the summer comparison with CO-OPS data which resulted in a bias of -8 to 5 ºC.

Model skill does not appear to degrade rapidly as the forecast hour projects into the later hours of the cycle as it did in the case of the G-RTOFS SSH forecast, where model skill steadily degrades with increasing forecast hour.

The G-RTOFS SSS forecast guidance fares poorly when compared with data from the nine NDBC buoy stations. The station bias ranges between 0 and 20 psu. The large RMSE values occur at the mid-Atlantic stations, numbers 4, 5, and 6. These stations are located within the estuaries of the Chesapeake and Delaware Bays. The station averaged RMSE is highest during the summer at about 10.5 psu and lowest during the fall at about 7.5 psu.

The G-RTOFS SSS forecast guidance fares relatively better when compared with the WOA09 database, however. The station bias ranges between -4 and 4 psu through all four seasons, including outlier values. The STD is uniformly small. The RMSE is generally within the 1 to 3 psu range.

ACKNOWEDGEMENTS

We are very grateful to Drs. Frank Aikman and Edward Myers, Chiefs of the Coastal Marine Modeling Branch, Coast Survey Development Laboratory (CSDL) for providing the project guidance and critical resources. We would like to express our sincere appreciation to Dr. Kurt Hess at CSDL for his insightful comments and suggestions to improve this document. We would also like to thank Dr. Avichal Mehra of the National Weather Service's Environmental Modeling Center for providing guidance on how to retrieve G-RTOFS forecast data archived in the NCEP's High Performance Storage System (HPSS). We wouldn't have been able to complete this project without their gracious help and encouragement.

REFERENCES

Global Real-Time Ocean Forecast System (G-RTOFS), 2011. hppt://polar.ncep.noaa.gov/global.

Metzger, E.J., O.M. Smedstad, P.G. Thoppil, 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. US Navy operational global ocean and Arctic ice prediction systems. *Oceanography* 27(3):32–43.

NOAA/NCEP's Extra Tropic Storm Surge system (ETSS), 1992. http://www.opc.ncep.noaa.gov/et_surge/et_surge_info.shtml

Appendix A. CO-OPS Water Level Stations Used for the Water Level Skill Assessment

Table A.1. CO-OPS water level stations used in G-RTOFS water level skill assessment

Station	ID	Station Name	Longitude (°E)	Latitude (°N)
1	8411060	Cutler Farris Wharf, Maine	-67.210	44.657
2	8413320	Bar Harbor, Maine	-68.205	44.392
3	8418150	Portland, Maine	-70.247	43.657
4	8419317	Wells, Maine	-70.563	43.320
5	8423898	Fort Pt., NH	-70.712	43.072
6	8447435	Chatham, MA	-69.950	41.688
7	8449130	Nantucket Island, MA	-70.097	41.202
8	8452660	Newport RI	-71.327	41.505
9	8465705	New Haven, CT	-72.908	41.283
10	8467150	Bridgeport, CT	-73.182	41.173
11	8515186	Fire Island, NY	-73.260	40.627
12	8513388	Mariches Coast Guard Station, NY	-72.750	40.787
13	8512354	Shinnecack Inlet, NY	-72.480	40.837
14	8512451	Ponquogue Point, NY	-72.503	40.850
15	8510560	Montauk, NY	-71.960	41.048
16	8534720	Atlantic City, NJ	-74.418	39.355
17	8531680	Sandy Hook, NJ	-74.008	40.467
18	8536110	Cape May, NJ	-74.960	38.968
19	8557380	Lewes, DE	-75.120	38.782
20	8570283	Ocean City Inlet, Md	-75.092	38.328
21	8638863	Chesapeake Bay Bridge Tunnel, VA	-76.147	36.967
22	8635750	Lewisetta, VA	-76.013	37.995
23	8632200	Kiptopeake, VA	-75.988	37.165
24	8636580	Windmill Point, VA	-76.290	37.615
25	8656483	Beaufort, NC	-76.670	34.720
26	8654467	USCG Hatteras, NC	-75.703	35.208
27	8658163	Wrightsville Beach, NC	-77.787	34.213
28	8652587	Oregon Inlet Marina, NC	-75.548	35.795
29	8651370	Duck, NC	-75.747	36.183
30	8662245	Oyster Landing (North Inlet Estuary), SC	-79.187	33.352
31	8661070	Springmaid Pier, SC	-78.918	33.655
32	8720218	Mayport (Bar Pilots Dock), FL	-81.430	30.397
33	8722670	Lake Worth Pier, FL	-80.033	21.612
34	8723970	Vaca Key, FL	-81.105	24.712
35	8723214	Virginia Key, FL	-80.162	25.730
36	8724580	Key West, FL	-81.807	24.555
37	8725110	Naples, FL	-81.807	26.132
38	8721604	Trident Pier, FL	-80.592	28.415
39	8726724	Clearwater Beach, FL	-82.832	27.978
40	8729210	Panama City Beach, FL	-85.878	30.213
41	8760922	Pilots East, SW Pass, LA	-89.407	28.932
42	8761724	Grand Isle, LA	-89.957	29.263

43	8762075	Port Fourchon, LA	-90.198	29.113
44	8770822	Texas Point, Sabine Pass, TX	-93.842	29.688
45	8771450	Galveston Pier 21, TX	-94.793	29.310
46	8771341	Galveston Bay Entrance, North Jetty, TX	-94.723	29.357
47	8772985	Sargent, TX	-95.617	28.772
48	8772447	USCG Freeport, TX	-95.302	28.943
49	8771486	Galveston Railroad Bridge, TX	-94.897	29.302
50	8775792	Packery Channel, TX	-97.237	27.633
51	8775870	Bob Hall Pier, Corpus Christi, TX	-97.217	27.580
52	8779770	Port Isabel,TX	-97.215	26.060

Table A.2. Water level stations used in ETSS water level skill assessment

Station	ID	Station Name	Longitude (°E)	Latitude (°N)
1	8725110	Naples, FL	-81.807	26.132
2	8726520	St Petersburg, FL	-82.627	27.760
3	8726724	Clearwater Beach, FL	-82.832	27.978
4	8727520	Cedar Key, FL	-83.032	29.135
5	8728690	Apalachicola, FL	-84.982	29.727
6	8729210	Panama City Beach, FL	-85.878	30.213
7	8729840	Pensacola, FL	-87.210	30.403
8	8770822	Texas Point, Sabine Pass, TX	-93.842	29.688
9	8772447	USCG Freeport, TX	-95.302	28.943
10	8410140	Eastport, ME	-66.982	44.903
11	8413320	Bar Harbor, ME	-68.205	44.392
12	8418150	Portland, ME	-70.247	43.657
13	8443970	Boston, MA	-71.053	42.353
14	8449130	Nantucket Island, MA	-70.097	41.202
15	8447930	Woods Hole, MA	-70.672	41.523
16	8454000	Providence, RI	-71.400	41.807
17	8452660	Newport RI	-71.327	41.505
18	8461490	New London, CT	-72.090	41.360
19	8467150	Bridgeport, CT	-73.182	41.173
20	8510560	Montauk, NY	-71.960	41.048
21	8518750	The Battery, NY	-74.013	40.700
22	8531680	Sandy Hook, NJ	-74.008	40.467
23	8534720	Atlantic City, NJ	-74.418	39.355
24	8536110	Cape May, NJ	-74.960	38.968
25	8551910	Reedy Point, DE	-75.573	39.558
26	8557380	Lewes, DE	-75.120	38.782
27	8570283	Ocean City Inlet, MD	-75.092	38.328
28	8574680	Baltimore, MD	-76.578	39.267
29	8575512	Annapolis, MD	-76.460	38.983
30	8571892	Cambridge, MD	-76.068	38.573
31	8577330	Solomons Island, MD	-76.452	38.317
32	8631044	Wachapreague, VA	-75.685	37.607
33	8635750	Lewisetta, VA	-76.013	37.995
34	8651370	Duck, NC	-75.747	36.183
35	8656483	Beaufort, NC	-76.670	34.720
36	8658120	Wilmington, NC	-77.953	34.227
37	8658163	Wrightsville Beach, NC	-77.787	34.213
38	8661070	Springmaid Pier, SC	-78.918	33.655
39	8665530	Charleston, SC	-79.925	32.782
40	8670870	Fort Pulaski, GA	-80.902	32.033
41	8720030	Fernandina Beach, FL	-81.465	30.672
42	8720218	Mayport (Bar Pilots Dock), FL	-81.430	30.397
43	8411060	Cutler Farris Wharf, ME	-67.210	44.657
44	8419317	Wells, ME	-70.563	43.320
45	8465705	New Haven, CT	-72.908	41.283

46	8571421	Bishops Head, MD	-76.030	38.220
47	8662245	Oyster Landing (North Inlet Estuary), SC	-79.187	33.352

Table A.3. CO-OPS water level stations used in the ESTOFS water level skill assessment

Station	ID	Station Name	Longitude (°E)	Latitude (°N)
1	8413320	Bar Harbor, ME	-68.205	44.392
2	8418150	Portland, ME	-70.247	43.657
3	8419317	Wells, ME	-70.563	43.320
4	8423898	Fort Point, NH	-70.712	43.072
5	8443970	Boston, MA	-71.053	42.353
6	8447386	Fall River, MA	-71.163	41.703
7	8447930	Woods Hole, MA	-70.672	41.523
8	8449130	Nantucket Island, MA	-70.097	41.202
9	8452660	Newport, RI	-71.327	41.505
10	8452944	Conimicut Light, RI	-71.343	41.717
11	8454000	Providence, RI	-71.400	41.807
12	8461490	New London, CT	-72.090	41.360
13	8467150	Bridgeport, CT	-73.182	41.173
14	8510560	Montauk, NY	-71.960	41.048
15	8516945	Kings Point, NY	-73.763	40.810
16	8531680	Sandy Hook, NJ	-74.008	40.467
17	8534720	Atlantic City, NJ	-74.418	39.355
18	8536110	Cape May, NJ	-74.960	38.968
19	8557380	Lewes, DE	-75.120	38.782
20	8571421	Bishops Head, MD	-76.030	38.220
21	8571892	Cambridge, MD	-76.068	38.573
22	8573364	Tolchester Beach, MD	-76.245	39.213
23	8574680	Baltimore, MD	-76.578	39.267
24	8575512	Annapolis, MD	-76.460	38.983
25	8577330	Solomons Island, MD	-76.452	38.317
26	8632200	Kiptopeke, VA	-75.988	37.165
27	8635750	Lewisetta, VA	-76.013	37.995
28	8636580	Windmill Point, VA	-76.290	37.615
29	8637689	Yorktown USCG Training Center, VA	-76.478	37.227
30	8638863	Chesapeake Bay Bridge Tunnel, VA	-76.147	36.967
31	8651370	Duck, NC	-75.747	36.183
32	8658163	Wrightsville Beach, NC	-77.787	34.213
33	8661070	Springmaid Pier, SC	-78.918	33.655
34	8721604	Trident Pier, FL	-80.592	28.415
35	8723214	Virginia Key, FL	-80.162	25.730
36	8723970	Vaca Key, FL	-81.105	24.712
37	8724580	Key West, FL	-81.807	24.555
38	8725110	Naples, FL	-81.807	26.132
39	8726384	Port Manatee, FL	-82.562	27.638
40	8726520	St Petersburg, FL	-82.627	27.760
41	8726607	Old Port Tampa, FL	-82.552	27.857
42	8726667	Mckay Bay Entrance, FL	-82.425	27.913
43	8726724	Clearwater Beach, FL	-82.832	27.978
44	8727520	Cedar Key, FL	-83.032	29.135

45	8729108	Panama City, FL	-85.667	30.152
46	8735180	Dauphin Island, AL	-88.075	30.250
47	8741041	Dock E, Port of Pascagoula, MS	-88.505	30.347
48	8741533	Pascagoula NOAA Lab, MS	-88.562	30.367
49	8761724	Grand Isle, LA	-89.957	29.263
50	8766072	Freshwater Canal Locks, LA	-92.305	29.555
51	8767816	Lake Charles, LA	-93.222	30.149
52	8767961	Bulk Terminal, LA	-93.300	30.190
53	8772447	USCG Freeport, TX	-95.302	28.943
54	8775870	Corpus Christi, TX	-97.217	27.580

Appendix B. CO-OPS Physical Oceanography Observation Stations Used for the G-RTOFS SST Skill Assessment

Table B.1. CO-OPS stations used for the G-RTOFS SST skill assessment

Station	ID	Station Name	Longitude (ºE)	Latitude (ºN)
1	8411060	Cutler Farris Wharf, Maine	-67.210	44.657
2	8413320	Bar Harbor, Maine	-68.205	44.392
3	8418150	Portland, Maine	-70.247	43.657
4	8419317	Wells, Maine	-70.563	43.320
5	8449130	Nantucket Island, MA	-70.097	41.202
6	8452660	Newport RI	-71.327	41.505
7	8465705	New Haven, CT	-72.908	41.283
8	8467150	Bridgeport, CT	-73.182	41.173
9	8510560	Montauk, NY	-71.960	41.048
10	8531680	Sandy Hook, NJ	-74.008	40.467
11	8534720	Atlantic City, NJ	-74.418	39.355
12	8536110	Cape May, NJ	-74.960	38.968
13	8557380	Lewes, DE	-75.120	38.782
14	8570283	Ocean City Inlet, MD	-75.092	38.328
15	8632200	Kiptopeake, VA	-75.988	37.165
16	8635750	Lewisetta, VA	-76.013	37.995
17	8638863	Chesapeake Bay Bridge Tunnel, VA	-76.147	36.967
18	8651370	Duck, NC	-75.747	36.183
19	8652587	Oregon Inlet Marina, NC	-75.548	35.795
20	8654467	USCG Hatteras, NC	-75.703	35.208
21	8656483	Beaufort, NC	-76.670	34.720
22	8658163	Wrightsville Beach, NC	-77.787	34.213
23	8661070	Springmaid Pier, SC	-78.918	33.655
24	8720218	Mayport (Bar Pilots Dock), FL	-81.430	30.397
25	8721604	Trident Pier, FL	-80.592	28.415
26	8722670	Lake Worth Pier, FL	-80.033	21.612
27	8723214	Virginia Key, FL	-80.162	25.730
28	8723970	Vaca Key, FL	-81.105	24.712
29	8724580	Key West, FL	-81.807	24.555
30	8725110	Naples, FL	-81.807	26.132
31	8726724	Clearwater Beach, FL	-82.832	27.978
32	8729210	Panama City Beach, FL	-85.878	30.213
33	8761724	Grand Isle, LA	-89.957	29.263
34	8770808	High Island, TX	-94.390	29.593
35	8770822	Texas Point, Sabine Pass, TX	-93.842	29.688
36	8771341	Galveston Bay Entrance, North Jetty, TX	-94.723	29.357
37	8771450	Galveston Pier 21, TX	-94.793	29.310
38	8771486	Galveston Railroad Bridge, TX	-94.897	29.302
39	8772447	USCG Freeport, TX	-95.302	28.943
40	8772985	Sargent, TX	-95.617	28.772
41	8773146	East Matagorda, TX	-94.913	28.710
42	8775237	Port Aransas, TX	-97.073	27.838

43	8775792	Packery Channel, TX	-97.237	27.633
44	8775870	Bob Hall Pier, Corpus Christi, TX	-97.217	27.580
45	8779770	Port Isabel, TX	-97.215	26.060

Appendix C. NDBC Buoy Stations Used for the G-RTOFS SST Skill Assessment

Table C.1. NDBC buoy stations used for the G-RTOFS SST skill assessment

Station	ID	Station Name	Longitude (°E)	Latitude (°N)
1	41001	East Hatteras, NC	-72.631	34.561
2	41002	South Hatteras	-74.835	31.862
3	41008	Grays Reef	-80.868	31.400
4	41009	Canaveral	-80.184	28.523
5	41012	St Augustine, FL	-80.534	30.042
6	41013	Frying Pan Shoals, NC	-77.743	33.436
7	41036	Onslow Bay, NC	-76.949	34.207
8	41037	Wrightsville Beach, NC	-77.363	33.988
9	41048	West Bermuda	-69.497	31.95
10	41110	Masonboro Inlet, NC	-77.709	34.141
11	41112	Offshore Fernadina Beach, FL	-81.292	30.709
12	41113	Cape Canaveral Nearshore, FL	-80.530	28.400
13	41114	Fort Pierce, FL	-80.225	27.551
14	42001	Mid Gulf 180 nm South of SW Pass, LA	-89.658	25.888
15	42002	West Gulf-207nm East of Brownsville, TX	-93.666	25.790
16	42003	East Gulf	-85.612	26.044
17	42012	44 NM SE of Mobile, AL	-87.555	30.065
18	42019	60 nm South of Freeport, TX	-95.353	27.913
19	42020	60 nm SSE of Corpus Christi, TX	-96.694	26.968
20	42035	22 nm East of Galveston, TX	-94.413	29.232
21	42036	112 NM WNW of Tampa, FL	-84.517	28.500
22	42039	115 NM ESE of Pensacola, FL	-86.006	28.794
23	42040	64 NM South of Dauphin Island, AL	-88.207	29.212
24	42043	GA-252 TABS B	-94.919	28.982
25	42044	PS-1126 TABS J	-97.051	26.191
26	42045	PI-745 TABS K	-96.500	26.217
27	42046	HI-A595 TABS N	-94.037	27.890
28	42047	HI-A389 TABS V	-93.597	27.897
29	42099	Offshore St. Petersburg, FL	-84.277	27.342
30	42360	Walker Ridge 249	-90.460	26.700
31	44007	12 NM SE of Portland, ME	-70.144	43.531
32	44008	Nantucket	-69.247	40.502
33	44009	Delaware Bay	-74.703	38.461
34	44013	Boston, ME	-70.651	42.346
35	44017	Montauk Pt, NY	-72.048	40.694
36	44024	Northeast Channel	-65.927	42.312
37	44025	30 miles South of Islip, NY	-73.167	40.250
38	44027	Jonesport, ME	-67.307	44.287
39	44029	Bouy A01, MA	-70.570	42.520
40	44030	Bouy A01, Ma-Western Maine Shelf	-70.418	43.183
41	44032	Buoy E01 Central Maine Shelf	-69.358	43.715
42	44033	Buoy F01-West Penobscot Bay	-69.000	44.060
43	44034	Buoy Io1 Eastern Maine Shelf	-68.110	44.110

44	44037	Buoy M01 Jordan Basin	-67.883	43.484
45	44056	Duck, NC	-75.714	36.200
46	44060	Eastern Long Island Sound, CT	-72.067	41.263
47	44065	New York Harbor, NY	-73.703	40.369
48	44066	Texas Towers #4, NJ	-72.600	39.584
49	44097	Block Island, RI	-71.117	40.981
50	44098	Jeffrey's Ledge, NH	-70.169	42.801
51	ACYN4	Atlantic City, NJ	-74.418	39.355
52	ATGM1-8413320	Bar Harbor, ME	-68.205	44.392
53	BFTN7	Beaufort, NC	-76.670	34.720
54	CAPL1	Calcasieu, LA	-93.343	29.768
55	CFWM1-8411060	Cutler Farris Wharf, ME	-67.210	44.657
56	CWBF1	Clearwater Beach, FL	-82.832	27.997
57	FCGT2	USCG Freeport, TX	-95.303	28.943
58	FWYF1	Fowey Rock, FL	-80.097	25.591
59	JMPN7	Wrightsville Beach, NC	-77.795	34.210
60	KYWF1	Key West, FL	-81.808	24.553
61	LKWF1	Lake Worth, FL	-80.033	26.612
62	MLRF1	Molassas Reef, FL	-80.376	25.012
63	MQTT2	Corpus Christi, TX	-97.217	27.580
64	MROS1	Springfield Pier, SC	-78.918	33.655
65	MTKN6	Montauk, NY	-71.960	41.048
66	MYPF1	Mayport, FL	-81.430	30.397
67	NPSF1	Naples, FL	-81.807	26.130
68	NTKM3	Nantucket Island, ME	-70.097	41.285
69	ORIN7	Oregon Inlet Marsh, NC	-75.548	35.795
70	PTAT2	Port Aransas, TX	-97.050	27.828
71	PTIT2	Port Isabelle, TX	-97.215	26.060
72	RTAT2	Port Aransas, TX	-97.073	27.840
73	SIPF1	Sabastian Inlet State Park, FL	-80.445	27.862
74	SPLL1	South Timbalier Block, LA	-90.483	28.867
75	TRDF1	Trident Pier, FL	-80.593	28.415
76	VAKF1	Virginia Key, FL	-80.162	25.732
77	VENF1	Venice, FL	-82.450	27.070
78	WELM1-8419317	Wells Maine	-70.144	43.531

APPENDIX D. NDBC Stations Used for the G-RTOFS SSS Skill Assessment

Table D.1. NDBC stations used in G-RTOFS SSS skill assessment

Station	NOS Number	Station Name	Longitude (°E)	Latitude (°N)
1	41037	ILM3 – 27 miles SE of Wrightsville Beach, NC	-77.363	33.988
2	41038	5 miles SE of Wrightsville Beach, NC	-77.721	34.142
3	RTAT2	Port Aransas, TX	-97.073	27.840
4	SIPF1	Sabastian Inlet State Park, FL	-80.445	27.862
5	SPLL1	South Timbalier Block, LA	-90.483	28.867
6	TRDF1	Trident Pier, FL	-80.593	28.415
7	VAKF1	Virginia Key, FL	-80.162	25.732
8	VENF1	Venice, FL	-82.450	27.070
9	WELM1-8419317	Wells Maine	-70.144	43.531