Naval Research Laboratory

Stennis Space Center, MS 39529-5004

NRL/MR/7320--17-9722

Global Ocean Forecast System 3.1 Validation Testing

E. Joseph Metzger
Robert W. Helber
Patrick J. Hogan
Pamela G. Posey
Prasad G. Thoppil
Tamara L. Townsend
Alan J. Wallcraft

Ocean Dynamics and Prediction Branch
Oceanography Division

Ole Martin Smedstad
Deborah S. Franklin

Vencore Services and Solutions, Inc.
Stennis Space Center, Mississippi

Luis Zamudio-Lopez

Florida State University
Tallahassee, Florida

Michael W. Phelps

Jacobs Technology Inc.
Herndon, Virginia

May 4, 2017

Approved for public release; distribution is unlimited.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing this collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number. **PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.**

1. REPORT DATE *(DD-MM-YYYY)*
04-05-2017

2. REPORT TYPE
Memorandum Report

3. DATES COVERED *(From - To)*

4. TITLE AND SUBTITLE

Global Ocean Forecast System 3.1 Validation Test

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER
0603207N

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER
73-5094-27-5

6. AUTHOR(S)

E. Joseph Metzger, Robert W. Helber, Patrick J. Hogan, Pamela G. Posey, Prasad G. Thoppil, Tamara L. Townsend, Alan J. Wallcraft, Ole Martin Smedstad,[1] Deborah S. Franklin,[1] Luis Zamudio-Lopez,[2] and Michael W. Phelps[3]

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Naval Research Laboratory
Oceanography Division
Stennis Space Center, MS 39529-5004

8. PERFORMING ORGANIZATION REPORT NUMBER

NRL/MR/7320--17-9722

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)

Office of Naval Research
One Liberty Center
875 North Randolph Street, Suite 1425
Arlington, VA 22203-1995

10. SPONSOR / MONITOR'S ACRONYM(S)

ONR

11. SPONSOR / MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT

Approved for public release; distribution is unlimited.

13. SUPPLEMENTARY NOTES

[1]Vencore Services and Solutions, Inc., Services and Solutions Group, Stennis Space Center, MS
[2]Florida State University, Tallahassee, FL
[3]Jacobs Technology Inc., Herndon, VA

14. ABSTRACT

The Global Ocean Forecast System (GOFS) 3.1 is comprised of the 1/12° HYbrid Coordinate Ocean Model that is two-way coupled to the Community Ice CodE in a daily update cycle with the Navy Coupled Ocean Data Assimilation. Additionally it uses Improved Synthetic Ocean Profiles to project surface information downward into the water column. GOFS nowcasts/forecasts the ocean's "weather", which includes the three-dimensional ocean temperature, salinity and current structure, the surface mixed layer, the location of mesoscale features, and ice concentration, thickness and drift in both hemispheres. It is scheduled to replace GOFS 3.0 for the ocean and the Arctic Cap Nowcast/Forecast System (ACNFS) for sea ice, both of which are the existing operational systems at the Naval Oceanographic Office. This report describes the system comparisons against unassimilated observations at both the nowcast time and as a function of forecast length. An ocean scorecard provides the relative performance of GOFS 3.1 vs. GOFS 3.0. Overall, GOFS 3.1 is performing equal to or better than both GOFS 3.0 and ACNFS, and is recommended to become the next global operational system.

15. SUBJECT TERMS

HYCOM ISOP
CICE Global ocean nowcast/forecast system
NCODA Model validation

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
					E. Joseph Metzger
a. REPORT	b. ABSTRACT	c. THIS PAGE	Unclassified Unlimited	59	19b. TELEPHONE NUMBER *(include area code)*
Unclassified Unlimited	Unclassified Unlimited	Unclassified Unlimited			(228) 688-4762

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std. Z39.18

Contents

1.0 INTRODUCTION **1**

2.0 GOFS 3.1 CONFIGURATION CHANGES **2**

2.1 HYCOM changes 2

2.2 Ice model changes 4

2.3 NCODA changes 4

2.3.1 ISOP modifications to improve the performance of GOFS 3.1 5

2.3.2 Ice assimilation 8

2.4 GOFS runstream configuration changes 9

2.5 Atmospheric forcing differences 11

3.0 VALIDATION TESTING RESULTS **12**

3.1 Validation hindcasts and forecasts 12

3.2 Ocean validation 12

3.2.1 Temperature/salinity vs. depth error analysis 13

3.2.2 Acoustical proxy error analyses 19

3.2.3 Upper ocean velocity validation 26

3.3 Ice validation 28

3.3.1 Ice edge location – Northern Hemisphere 30

3.3.2 Ice edge location – Southern Hemisphere 38

3.3.3 Ice thickness 39

3.3.4 Ice drift 42

4.0 SUMMARY, SCORE CARDS AND RECOMMENDATIONS **46**

5.0 ACKNOWLEDGMENTS **50**

6.0 REFERENCES **52**

7.0 TABLE OF ACRONYMS **56**

1.0 INTRODUCTION

This Validation Test Report (VTR) documents the continuing development, validation, verification and transition of the Global Ocean Forecast System (GOFS) for operational use at the Naval Oceanographic Office (NAVOCEANO) and run on the Navy DoD Supercomputing Resource Center (DSRC) computers. GOFS (version) 3.0 is based on the 1/12º global HYbrid Coordinate Ocean Model (HYCOM) and the Navy Coupled Ocean Data Assimilation (NCODA) and was previously documented by Metzger et al. (2008, 2010, 2014). It is the existing US Navy's operational global ocean nowcast/forecast system. GOFS 3.1 is scheduled to replace GOFS 3.0 pending its successful validation by NRL and OPerational TESTing (OPTEST) by NAVOCEANO and National Ice Center (NIC).

In September 2014 the original GOFS 3.1 VTR (Metzger et al., 2015) was delivered to NAVOCEANO. The conclusion of that report was that, in the net, GOFS 3.1 outperformed GOFS 3.0 for all ocean analyses and outperformed the Arctic Cap Nowcast/Forecast System (ACNFS) for the majority of the ice analyses. Thus the recommendation of the Validation Test Panel (VTP) was that GOFS 3.1 should enter OPTEST for the ocean (NAVOCEANO) and for ice (NIC).

NAVOCEANO and NIC both proceeded with the OPTEST over the course of the next year. However, in the summer of 2015 an issue associated with the assimilation of erroneously processed satellite sea surface temperature (SST) data in the Arctic caused significant delays until rectified in the real time systems. In September 2015, NIC gave verbal approval that GOFS 3.1 had passed the OPTEST with regard to ice validation. NAVOCEANO was in the process of finalizing their OPTEST when it came to NRL's attention that upper ocean velocity and energy levels in GOFS 3.1 were significantly reduced compared to GOFS 3.0, despite ocean velocity metrics in the VTR that indicated otherwise (see Metzger et al., 2015 section 3.2.4). The problem was significant enough to warrant a delay and GOFS 3.1 was pulled from OPTEST while NRL began work to rectify the problem. As a side note, GOFS 3.1 continued to run in real time at the Navy DSRC under NAVOCEANO control in order to supply Northern and

Manuscript approved February 16, 2017.

Southern Hemisphere ice products to NIC as these were not adversely affected by the aforementioned problem.

This VTR will detail the issues associated with the reduced upper ocean velocity and the changes made to NCODA/ISOP to correct the deficiencies. A new year long GOFS 3.1 hindcast and series of 5-day forecasts have been integrated and validation metrics similar to Metzger et al. (2015) are documented within. The upper ocean velocity metrics (Section 3.2.3) have been modified to provide a more accurate comparison of GOFS upper ocean velocity to independent drifting buoys. Many of the other metrics are the same as in the original GOFS 3.1 VTR. Ocean metrics will be compared between GOFS 3.1 and GOFS 3.0, while ice metrics will be compared between GOFS 3.1 and ACNFS, the US Navy's operational ice prediction system.

2.0 GOFS 3.1 CONFIGURATION CHANGES

Section 2 of Metzger et al. (2015) described new capabilities and improvements in GOFS 3.1 compared to GOFS 3.0/ACNFS with regard to the ocean/ice model components, data assimilation and the runstream configuration. Much of that same information is repeated here in order for this document to be self-contained and more recently added system modifications are also presented.

2.1 HYCOM changes

The ocean model software in GOFS 3.1 is HYCOM version 2.2.99DH vs. HYCOM version 2.2.19 in GOFS 3.0. In addition, the model configuration has changed significantly (some, but not all, of the changes rely on the newer HYCOM version). The models are on the same 4500 x 3298 tripole global grid (~9 km equatorial resolution, ~7 km mid-latitude resolution and ~3.5 km resolution at the North Pole), but the coastline and bathymetry are improved in GOFS 3.1. In GOFS 3.0, the model coastline is the 10 m isobath. In GOFS 3.1 it is the 0.1m isobath, i.e. the actual coastline, but with depths shallower than 5 m set to 5 m and depths greater than 6500 m scaled by 0.2 (maximum resulting depth is 7200 m). GOFS 3.0 uses a 7-term polynomial equation of state and has 32 layers in the vertical with at least the top 4 layers acting as z-levels (all layers become z-levels at high latitudes). GOFS 3.1 instead uses a 17-term rational function that approximates the Thermodynamic Equation of Seawater (TEOS-10) (Jackett et al.,

2006) and has 41 layers in the vertical. Most of the original 32 layers are retained, with the additional 9 layers all located near the surface such that at least the top 14 layers are always sigma-z levels so that water shallower than 84 m is always in fixed depth coordinates, and mixed layers are typically better resolved. In GOFS 3.0, isopycnal layers often upwelled onto the shelf which meant that only a few layers spanned the water column there, which was problematic for resolving the velocity profile. In GOFS 3.1 upwelled water masses have the correct temperature (T) and salinity (S) but are not represented by isopycnal layers over the shelf, thus allowing more degrees of freedom for the velocity profile. The "sigma-z" setup has also been modified such that the top layer is 1 m thick everywhere (GOFS 3.0, with conventional sigma coordinates, has a top layer that gets thinner as the total depth shallows), as this simplifies layer 1 temperature's association with SST. The surface forcing has also changed: GOFS 3.0 reads in wind stress that has been calculated off-line using a bulk parameterization based on the atmospheric model's 10 m winds and SST, whereas GOFS 3.1 reads in 10 m winds and calculates the wind stress using HYCOM SST and also taking into account HYCOM surface currents. The bulk parameterizations have been updated to use atmospheric surface pressure (vs. constant surface pressure) as part of the wind stress and heat flux coefficient computations. The maximum wind speed for the wind stress and heat flux bulk parameterizations is 34 m/s in GOFS 3.1, based on Sraj et al. (2013). GOFS 3.1 also inputs downward (vs. net) surface radiation (shortwave and longwave) fluxes. Ocean turbidity is based on a monthly chlorophyll climatology in GOFS 3.1 (Joliff et al., 2014) vs. photosynthetically available radiation in GOFS 3.0. GOFS 3.0 is initialized from the Generalized Digital Environment Model (GDEM) 3 (Carnes, 2009) and uses the Polar Science Center Hydrographic Climatology 3.0 (Steele et al., 2001) 10 m salinity as its sea surface salinity (SSS) climatology. GOFS 3.1 was initialized from GDEM 4 (Carnes et al., 2010) and uses GDEM 4.2 (improved coastal regions) surface salinity as its SSS climatology. Relaxation to SSS climatology is weaker in GOFS 3.1 and is turned off completely when the ocean model SSS is more than 0.5 psu away from climatology (which helps near rivers and in regions where advection dominates the SSS tendency).

2.2 Ice model changes

GOFS 3.0 contains a thermodynamic "energy-loan" sea ice model which is a standard sub-model within the HYCOM source code. It allows ice to grow/melt in response to changes in temperature and heat fluxes, but ice is not advected by the wind or ocean currents. In GOFS 3.1, HYCOM is fully two-way coupled to the Community Ice CodE (CICE) (Hunke and Lipscomb, 2008), the latter developed at Los Alamos National Laboratories. CICE has more sophisticated physics and improvements over earlier sea ice models that includes multiple ice thickness layers, multiple snow layers and new ice ridging parameterizations. The ocean and sea ice models share the same horizontal grid and are coupled via the Earth System Modeling Framework (Hill et al., 2004) with fields exchanged every hour. The CICE software is nearly the same as that employed by the ACNFS, that was validated by Posey et al. (2010) and Hebert et al. (2015). The main difference is that the ocean stress imparted on the ice is based on 3 m average ocean currents in GOFS 3.1, but on 10 m average ocean currents in ACNFS, and this required doubling the ice-ocean drag coefficient in GOFS 3.1. With regard to ice validation in this report, GOFS 3.1 will be compared against ACNFS rather than GOFS 3.0. Since GOFS 3.1 is a global system it also produces ice nowcasts and forecasts for the Southern Hemisphere.

2.3 NCODA changes

NCODA is able to represent the 3D ocean mesoscale by creating synthetic profile data that projects surface information downward into the water column for assimilation into the forecast model. In GOFS 3.0, NCODA uses the Modular Ocean Data Analysis System (MODAS) (Fox et al., 2002) to infer subsurface ocean properties from the relatively abundant real-time satellite-based observations of sea surface height (SSH) anomaly and SST. This method of synthetic temperature and salinity profile generation is based on historical relationships between surface and subsurface ocean observations. In GOFS 3.1, MODAS is replaced with the Improved Synthetic Ocean Profile (ISOP) system (Helber et al., 2013), which became available for use in NCODA (Townsend et al. 2015) after GOFS 3.0 was transitioned to operations. The ISOP system employs a more versatile 1D variational approach to utilize historical vertical covariances that not only preserves the relationship between the surface and subsurface

ocean properties but also the vertical gradients of temperature and salinity, which were weaknesses in MODAS. Assimilation of synthetic profiles with more accurate vertical structure results in more accurate prediction of the ocean vertical structure which in turn results in more accurate acoustic propagation/transmission loss predictions.

2.3.1 ISOP modifications to improve the performance of GOFS 3.1

The symptom of weak upper ocean velocity and energy levels in the original GOFS 3.1 was related to the quality of the ISOP synthetics being assimilated into HYCOM. The ISOP synthetic profiles failed to represent the true SSH variability of the ocean. As a result, upper ocean velocities were too weak and the eddy kinetic energy of the mesoscale variability was too low. It has since been determined that both the SSH anomalies observed from satellite altimeters and their associated error levels were incorrectly applied within the original ISOP algorithm. The problems amounted to a double accounting of the SSH anomaly errors and an overly strict culling of large observed SSH anomalies. To correct these problems, changes to the ISOP system in order of importance include: (1) reducing the input SSH anomaly error estimate for computing the synthetics, (2) revising and improving the databases used for constraining the SSH anomalies and errors, and (3) increasing the allowable range of input SSH anomalies. Making changes to ISOP in these three areas restored the skill of GOFS 3.1 to surpass GOFS 3.0 for all metrics including those for upper ocean velocity and eddy kinetic energy. These changes are described in greater detail in the following paragraphs.

A key aspect of data assimilation systems is the representation of errors. In the original ISOP system, the error in satellite observed SSH anomalies was applied when making the synthetic profiles. In the context of the GOFS cycling forecast system this became a double accounting of errors because the SSH anomaly errors are also used to estimate the output error level of the ISOP synthetics themselves. The solution is to reduce the SSH anomaly error for making ISOP synthetics but not when estimating the error of the synthetic. This is the most important change for correcting ISOP in GOFS 3.1.

Three databases files are used to modify the SSH anomaly error levels and remove large SSH anomalies. In the new version of ISOP, the file that increases SSH anomaly errors to account for steric

versus non-steric SSH anomaly errors is set to zero between 48°S and 48°N. The slope correction file is also updated to better account for the non-steric SSH anomaly contribution. The slope correction now depends only on latitude, based on HYCOM's total SSH anomaly after incremental insertion. In addition, the file used for removing large SSH anomalies has been updated and is now based on the latest SSH anomaly standard deviation estimate from a GOFS 3.0-like Ocean Reanalysis over the years 2003-2012. The standard practice of removing SSH anomaly values that are too far from the average was also too strict in the original version of ISOP. In the revised version, larger values of SSH anomalies are allowed. In total, the changes required to restore the skill of ISOP within NCODA amounted to three new database files and the aforementioned software code changes.

Figure 1 illustrates upper ocean (15 m) eddy kinetic energy (EKE) in GOFS 3.1 and GOFS 3.0 compared to one derived from independent drifting buoys with a drogue at 15 m depth (Lumpkin and Johnson, 2013). Drifters that lost their drogue have been excluded from the comparison. Here the GOFS results are based on one year of output while the drifter data spans a much longer time period, 1979-2012. GOFS 3.1 agrees more closely with the drifter data, especially in the low EKE regions west of the Americas and west of southern Africa. Overall GOFS 3.0 is too energetic at this depth as will be noted in the velocity comparisons below. The basin-wide average EKE in GOFS 3.1 has nearly doubled compared to the original GOFS 3.1 (not shown) and is also more consistent with the drifter derived data. Figure 2 shows cumulative percent and a histogram of 15 m EKE variability between 60°S-60°N that also clearly indicates GOFS 3.1 more closely matches the drifter derived EKE.

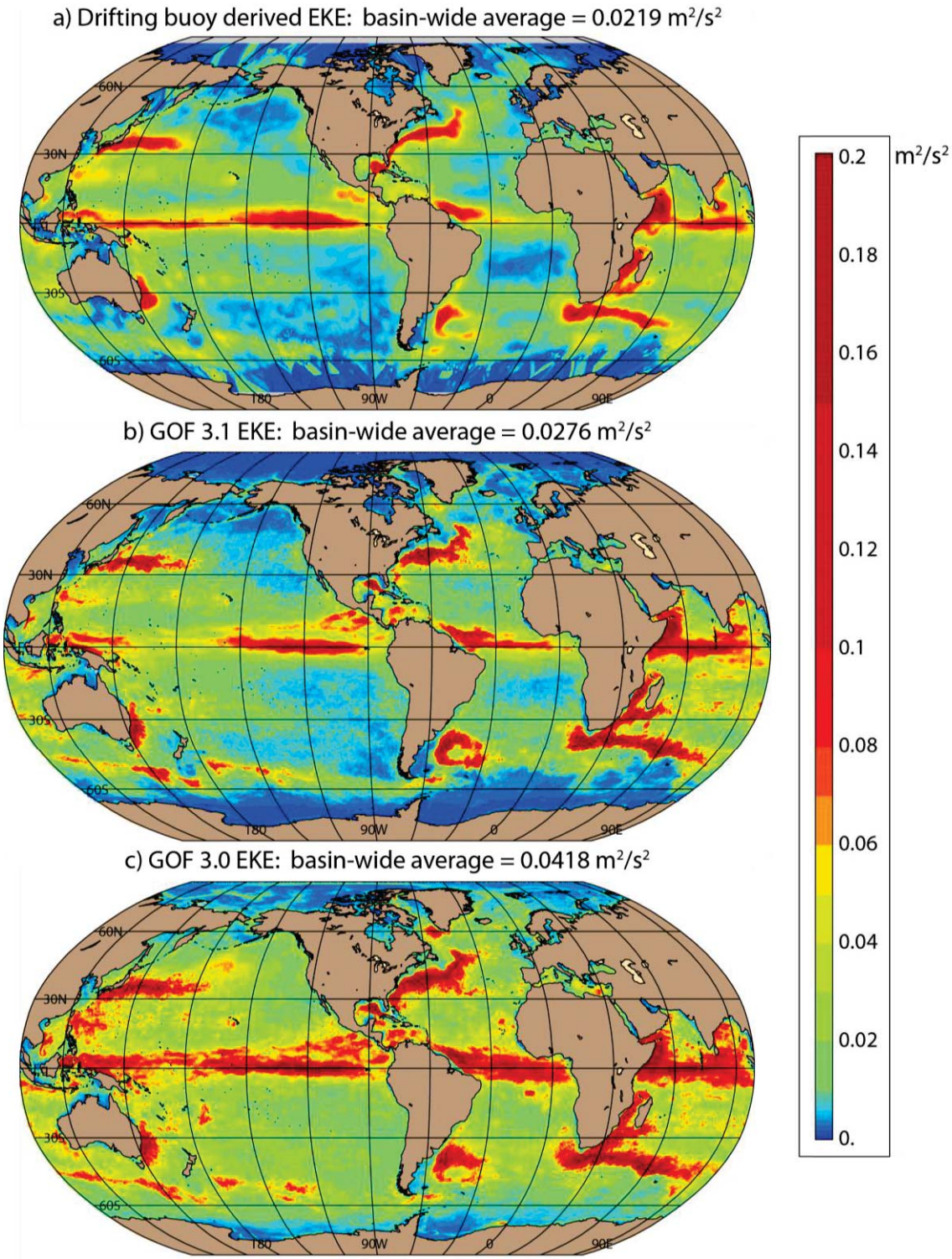

Figure 1: Eddy kinetic energy (EKE, m^2/s^2) at 15 m depth a) derived from drifting buoys, b) from GOFS 3.1, and c) from GOFS 3.0. The drifter derived data spans 1979-2012 while the GOFS EKE covers the hindcast period (7/2014 – 6/2015.) The basin wide average is noted above each panel.

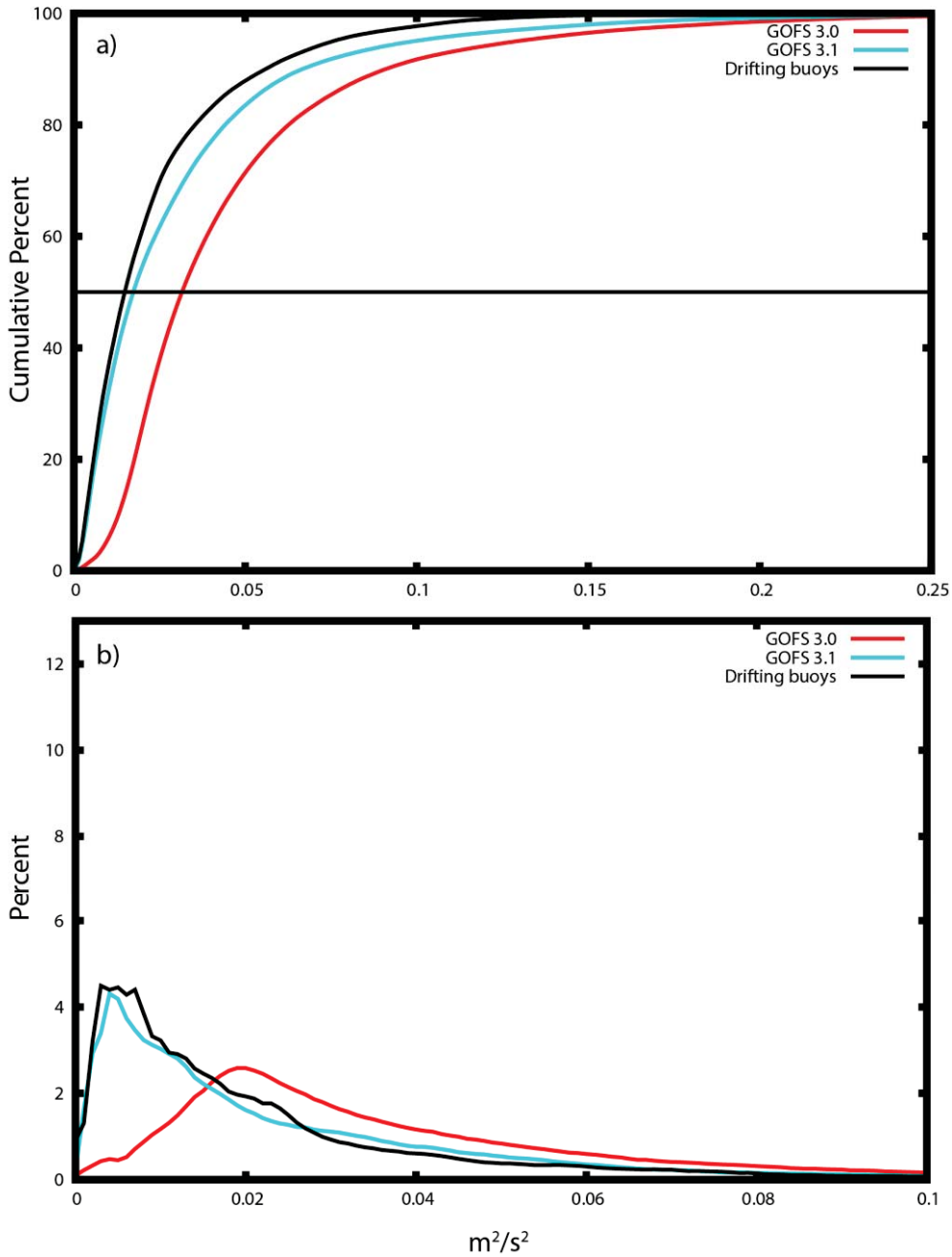

Figure 2: (top) Cumulative percent and (bottom) percent of EKE (m^2/s^2) in 0.005 bands. The curves are based on the panels in Figure 1 at 15 m depth and computed between 60°S-60°N. The black curve is from the drifting buoy data, the cyan curve from GOFS 3.1 and the red curve from GOFS 3.0.

2.3.2 Ice assimilation

The assimilation of ice concentration is also treated differently between GOFS 3.1 and ACNFS. In both systems, NCODA produces a daily ice concentration analysis based on yesterday's 24-hour CICE

forecast as the background and derived ice concentration products from 1) the Special Sensor Microwave Imager/Sounder (SSMIS) (~25 km resolution), 2) the Advanced Microwave Scanning Radiometer 2 (AMSR2) (~12.5 km resolution) and the 4 km gridded Interactive Multisensor Snow and Ice Mapping System (IMS) sea ice mask (see Posey et al. (2015) for specific details). The NCODA ice concentration analysis field is then directly inserted into CICE at 09Z in GOFS 3.1 and at 18Z in ACNFS. ACNFS assimilates the NCODA ice concentration analysis only near the ice edge and leaves the model field untouched away from the ice edge, while GOFS constrains its concentration in the interior to be within 10% of the NCODA analysis. The data sources noted above are used for the entirety of the GOFS 3.1 hindcast but were implemented in operational ACNFS on 3 February 2015. Prior to this date ACNFS only assimilated SSMIS ice concentration observations. Posey et al. (2015) documented a large improvement in ice edge error using the blended data sources, thus ACNFS ice edge location skill is degraded for approximately the first seven months of the hindcast period.

2.4 GOFS runstream configuration changes

Schematics of the GOFS 3.0 and 3.1 runstreams are shown in Figure 3. In GOFS 3.0 the single daily update cycle starts with the NCODA analysis at 18Z using the 24-hour HYCOM forecast as a first guess. Then, HYCOM is integrated for 24 model hours with the NCODA incremental analysis update (Bloom et al., 1996) applied to the ocean model over the first six hours (18Z → 00Z). Thus at 00Z HYCOM has fully assimilated all the observational data. Every day the system goes back 102 hours from the nowcast time because of late-arriving satellite altimeter data. The NCODA analysis and HYCOM hindcast cycle repeats itself five times up to the nowcast time (t=0), and HYCOM continues to run in (non-assimilative) forecast mode out to t=168 hours (one week).

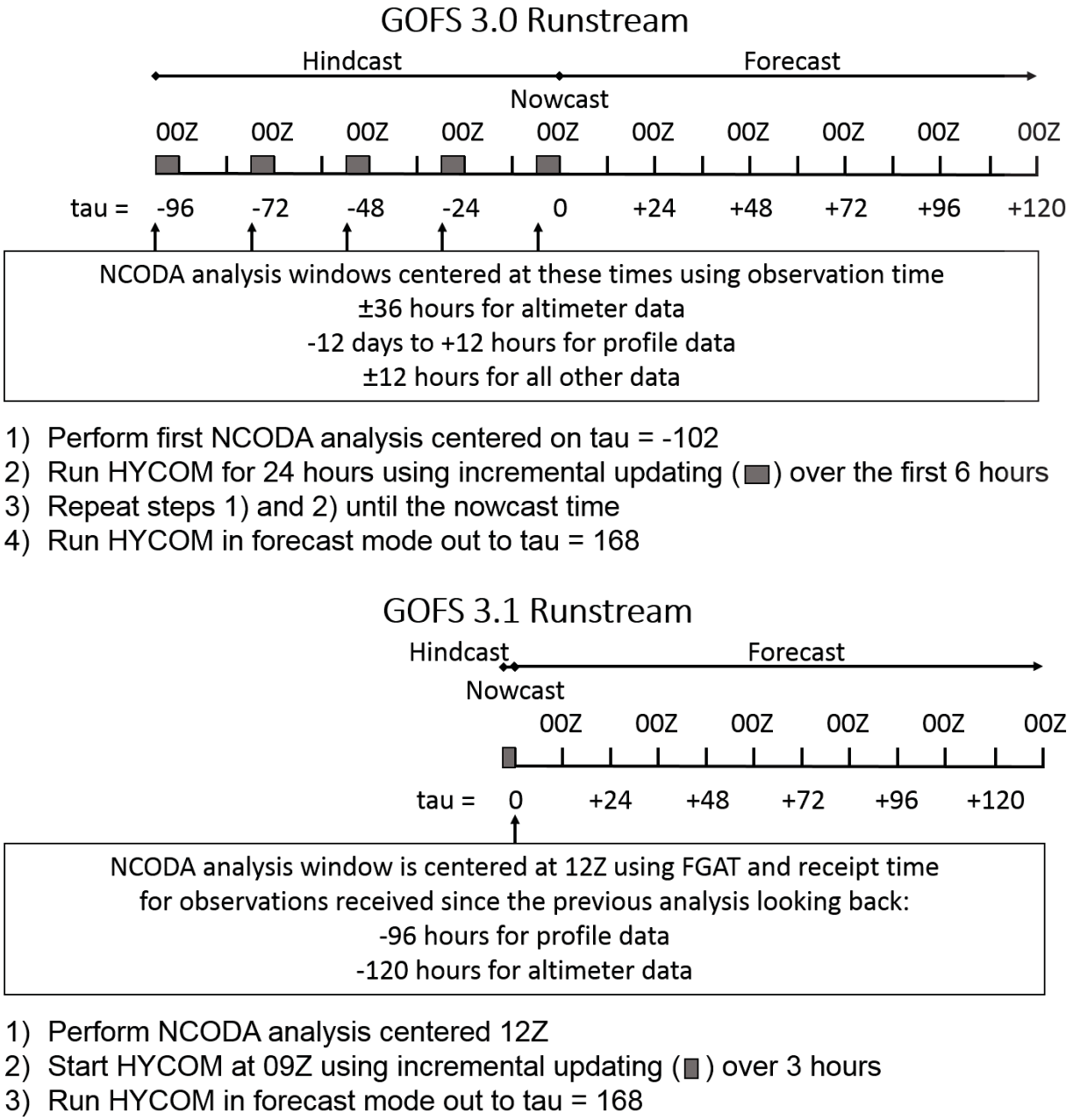

Figure 3: The GOFS 3.0 (top) and GOFS 3.1 (bottom) runstream configurations. The forecast length is only shown to five days (120 hours) but actually extends to seven days (168 hours).

In GOFS 3.1, the NCODA analysis is performed at 12Z (rather than 18Z) and the need for a daily four-day hindcast is mitigated by using the First Guess at Appropriate Time (FGAT) capability within NCODA. Comparing observations to model background fields valid at different times is a source of error

for the NCODA analysis. This error is reduced by comparing observations against time-dependent background fields using the FGAT technique. (Cummings and Smedstad, 2014) Thus, late arriving observations can be assimilated without the need to do a complete NCODA analysis N days back from the nowcast time and this new methodology greatly reduces the daily run-time of GOFS 3.1 compared to GOFS 3.0. HYCOM runs forward starting at 09Z with the incremental analysis update window between 09Z → 12Z, i.e. starting before the analysis time. Thus at 12Z HYCOM has fully assimilated all the observational data and the nowcast time and NCODA analysis time are synchronized. The total run-time on a Cray XC40 from the beginning of the hindcast to the end of the forecast takes 6.5 hours on 713 cores for GOFS 3.0, but only 3.8 hours on 900 cores for GOFS 3.1 despite the addition of nine layers to HYCOM and the inclusion of the more dynamically sophisticated CICE model.

2.5 Atmospheric forcing differences

When declared operational in March 2013, GOFS 3.0 used atmospheric forcing from the Navy Operational Global Atmospheric Prediction System (NOGAPS). NRL-Monterey has since developed and transitioned to operations at Fleet Numerical Meteorology and Oceanography Center (FNMOC) the NAVy Global Environmental Model (NAVGEM) (Hogan et al., 2014) that replaced NOGAPS. Initial comparisons of NOGAPS and NAVGEM surface fields indicated large differences in some variables such that GOFS 3.0 and ACNFS could not simply switch to NAVGEM with the expectation that the ocean/ice model response would be unchanged. Thus, a series of wind speed and heat flux calibrations were required for consistency between atmospheric forcing products (Metzger et al., 2013). For the hindcast period, GOFS 3.0 and ACNFS used NAVGEM 1.2 forcing, whereas GOFS 3.1 used NAVGEM 1.4 forcing. Because of the rapid NAVGEM upgrade schedule, it is difficult to ensure consistency in the atmospheric forcing between the operational systems and the system being transitioned. However, similar wind speed, surface radiation and net heat flux calibrations were applied to both, so the different atmospheric forcing in GOFS 3.0/ACNFS and GOFS 3.1 did not bias the ocean and ice validation metrics within.

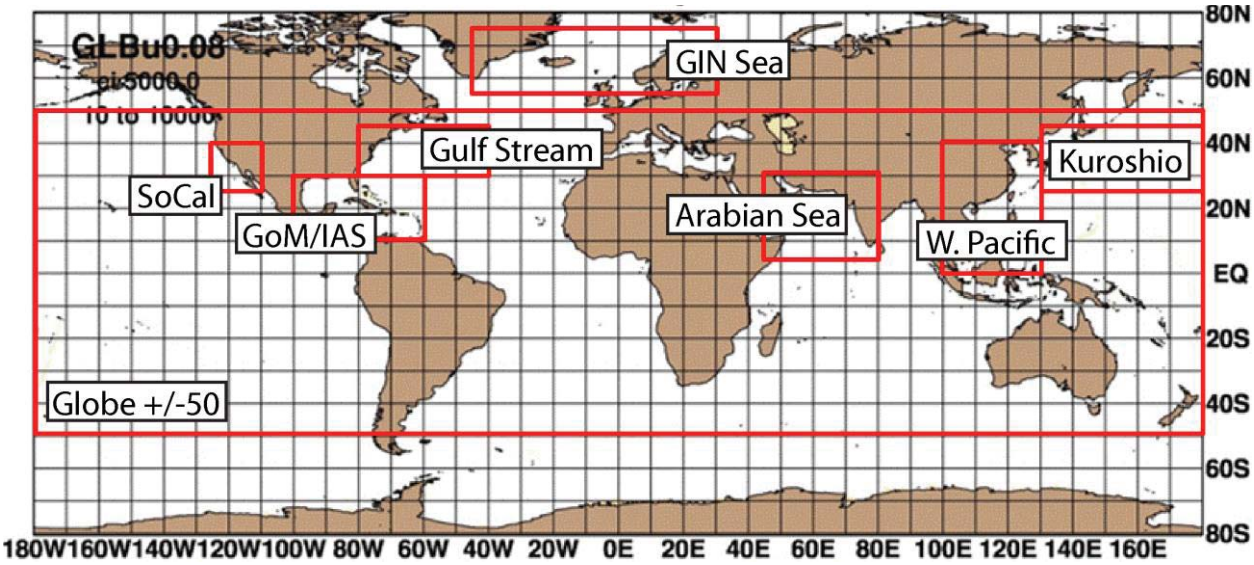

Figure 4: The ocean analysis regions used in this report: Globe: 180°W-180°E, 50°S-50°N; West Pacific marginal seas: 100-130°E, 0-40°N; Kuroshio Extension: 130°E-180°, 25-45°N; Arabian Sea/Persian Gulf/Gulf of Aden: 44-80°E, 5-30°N; Gulf Stream: 80-40°W, 30-45°N; Southern California (SoCal): 125-110°W, 25-40°N; Greenland/Iceland/Norwegian (GIN) Seas: 45°W-30°E, 55-75°N; and Gulf of Mexico (GoM)/Intra-Americas Sea (IAS): 100-60°W, 10-30°N.

3.0 VALIDATION TESTING RESULTS

3.1 Validation hindcasts and forecasts

A year-long GOFS 3.1 hindcast was integrated to emulate a real-time operational system, assimilating all available observational data. (In NRL nomenclature this is GLBb0.08-56.3.) It cycled forward one day at a time over the time period 1 July 2014 – 30 June 2015. In addition, a series of 5-day forecasts were made to examine short-term forecast skill (GLBb0.08-56.4). These forecasts were initialized from the GOFS 3.1 hindcast and forced with **forecast quality** NAVGEM 1.4 forcing. The frequency of the 5-day forecasts was dictated by the availability of NAVGEM forecast output that was typically every third day for a total of 10 forecasts each month. No ocean/ice data were assimilated into the forecast, thus all observational data used for validation are independent.

3.2 Ocean validation

The ocean error analyses used for validation in this report are similar to those used for GOFS 3.0 (Metzger et al., 2010), namely temperature/salinity vs. depth and acoustical proxies against unassimilated observational profile data at both the nowcast time out through a short-term (5-day) forecast. Additional

error analyses include predictions of surface layer trapping of acoustical frequencies, validation of upper ocean currents against drifting buoys and a comparison of SSH variability along satellite altimeter tracks. The analysis regions (Figure 4) have been chosen by the Validation Test Panel (VTP) to be both Navy relevant and dynamically diverse.

3.2.1 Temperature/salinity vs. depth error analysis

A temperature/salinity vs. depth error analysis within the top 500 m of the water column is performed using unassimilated profiles. For a given observation, both GOFS 3.0 and GOFS 3.1 are sampled at the nearest model grid point. In the vertical, both the observations and model output are linearly remapped to a common set of depth levels: 0, 2, 4, 6, 8, 10, 12, 15, 20, 25, 30, 35, 40, 45, 50, 60, 70, 80, 90, 100, 125, 150, 200, 250, 300, 350, 400, and 500 m. Model-data differences that exceed three standard deviations are excluded, i.e. the 99% confidence interval is used, and this is the reason why the observation counts differ between the two systems. Figure 5 illustrates how profile data are sampled relative to each system. Recall that for GOFS 3.0, the NCODA analysis is performed at 18Z and uses profiles forward in time to 06Z, whereas in GOFS 3.1 the NCODA analysis is at 12Z and uses profiles forward in time to 00Z. In order to make sure only unassimilated data are used in the analysis, profiles between 06Z → 00Z+1 day are compared against the 00Z HYCOM archive file for GOFS 3.0 (Figure 5a top panel) and profiles between 00Z → 00Z+1 day are compared against the 00Z HYCOM archive file for GOFS 3.1 (Figure 5a bottom panel). Note that for GOFS 3.0, the first 00Z time point after the NCODA analysis is the nowcast time, but for GOFS 3.1 it is actually a 12 hour forecast. When sampling the forecasts (Figure 5b), profiles ±1.5 hours on either side of the 3-hourly HYCOM archive files centered ±12 hours over 12Z are used.

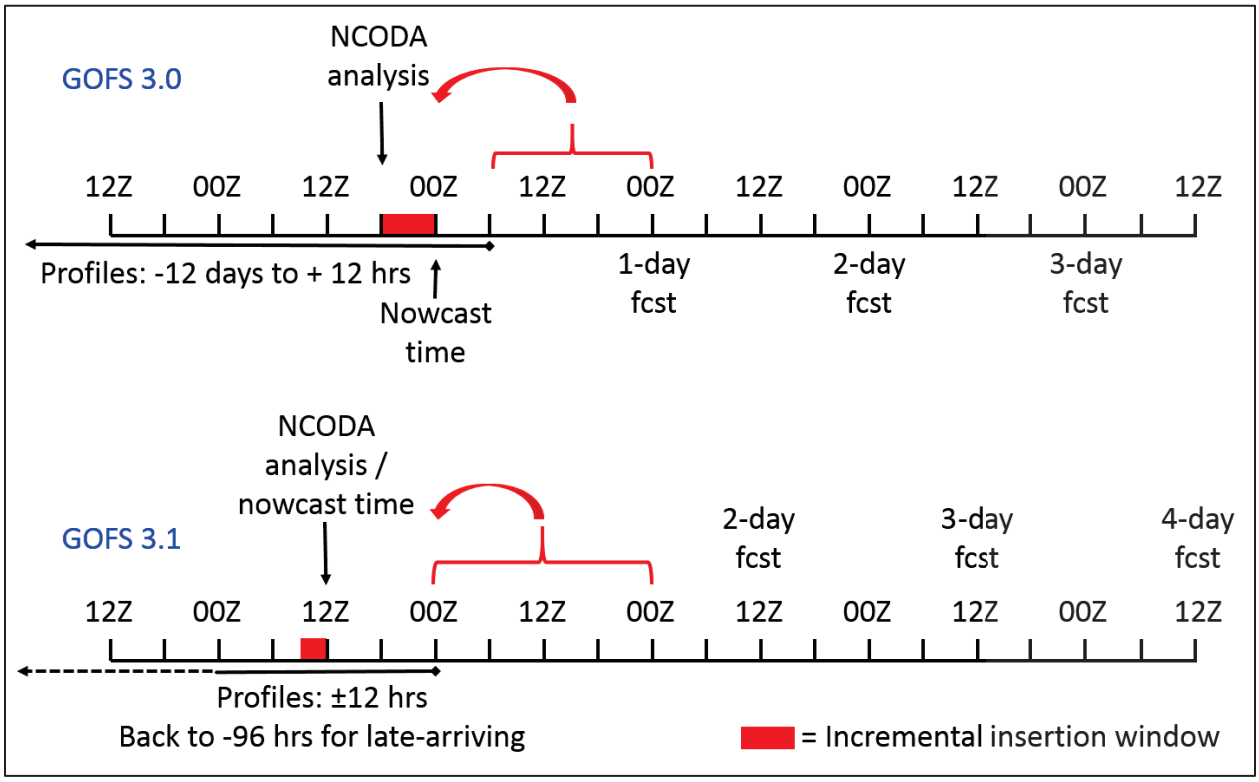

Figure 5a: Methodology for sampling profile data from HYCOM archive files at the nowcast time.

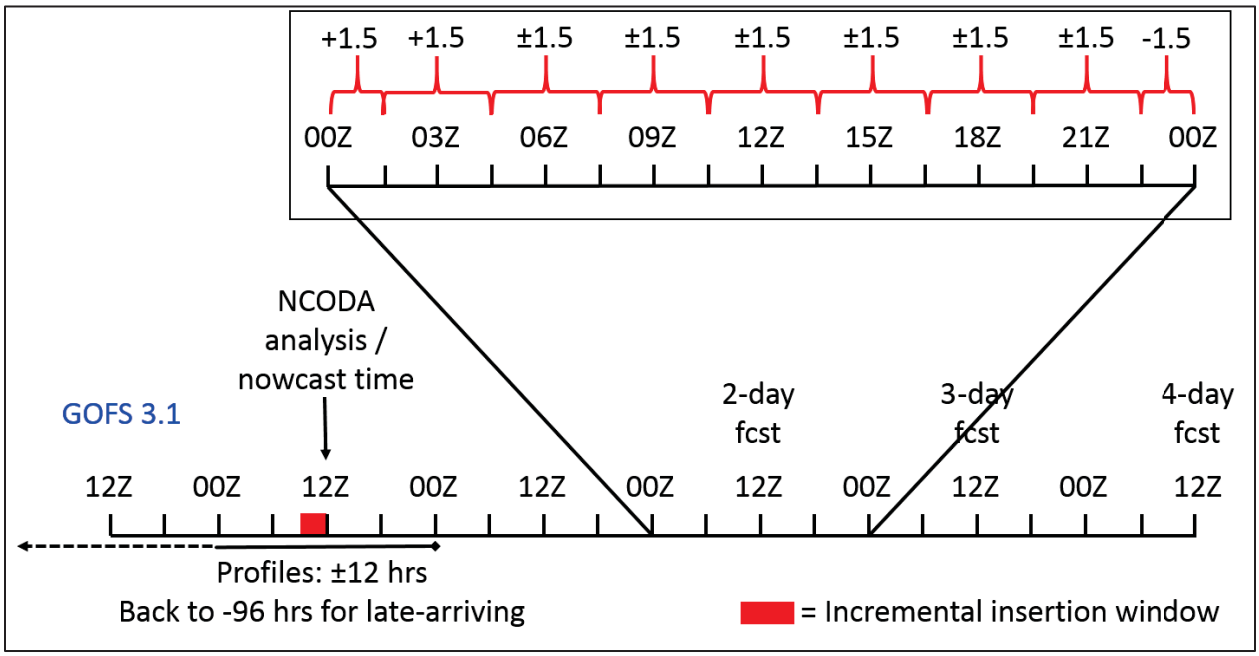

Figure 5b: As Fig. 3a except for sampling HYCOM at the forecast time. Here the 2-day forecast is shown.

Figure 6 shows mean error (ME) (bias) and root mean square error (RMSE) for temperature in the eight analysis regions (Figure 4) at the "nowcast" time. In general, the ME is small (<0.25°C) in both systems and typically well within the NAVOCEANO-defined tolerances for the upper ocean (0.5°C). Over the globe (±50°) there is a small overall cool bias in both systems. With regard to the individual analysis regions, the ME results are mixed with one system performing slightly better than the other depending upon the region that is selected. However, with regard to RMSE, GOFS 3.1 clearly outperforms GOFS 3.0 in all regions and at nearly every depth level. As expected, the highest RMSE is seen at the depth level of the thermocline.

A salinity vs. depth error analysis is shown in Figure 7. Recall from section 2. 1 that both systems relax to climatological salinity at the sea surface but this is weaker in GOFS 3.1. Similar to temperature, the two systems have comparable ME that is small (<0.1 psu) whereas again GOFS 3.1 generally has lower RMSE than GOFS 3.0.

Upper ocean temperature and salinity error analyses as a function of short-term (5-day) forecast length are shown in Figures 8 and 9, respectively. The black curve is the 1-day forecast, the red curve is the 3-day forecast and the cyan curve is the 5-day forecast. The change in ME vs. forecast length is typically largest near the surface, but the degradation in skill as the forecast length increases is still small nonetheless. The error growth with forecast length is more obvious for RMSE, and averaged over the top 500 m, the global domain error between the nowcast time and the 5-day forecast grows by 16% (0.51 to 0.59 °C) and 7% (0.15 to 0.16 psu) for temperature and salinity, respectively. The error growth is less consistent for the regional domains and there are times when the 5-day forecast error is smaller than the 1-day forecast error, although, the difference is generally small in such cases. The largest spread in ME is noted within the top 150 m for the Southern California region for T and for S the RMSE spread is highest for this region and depth range.

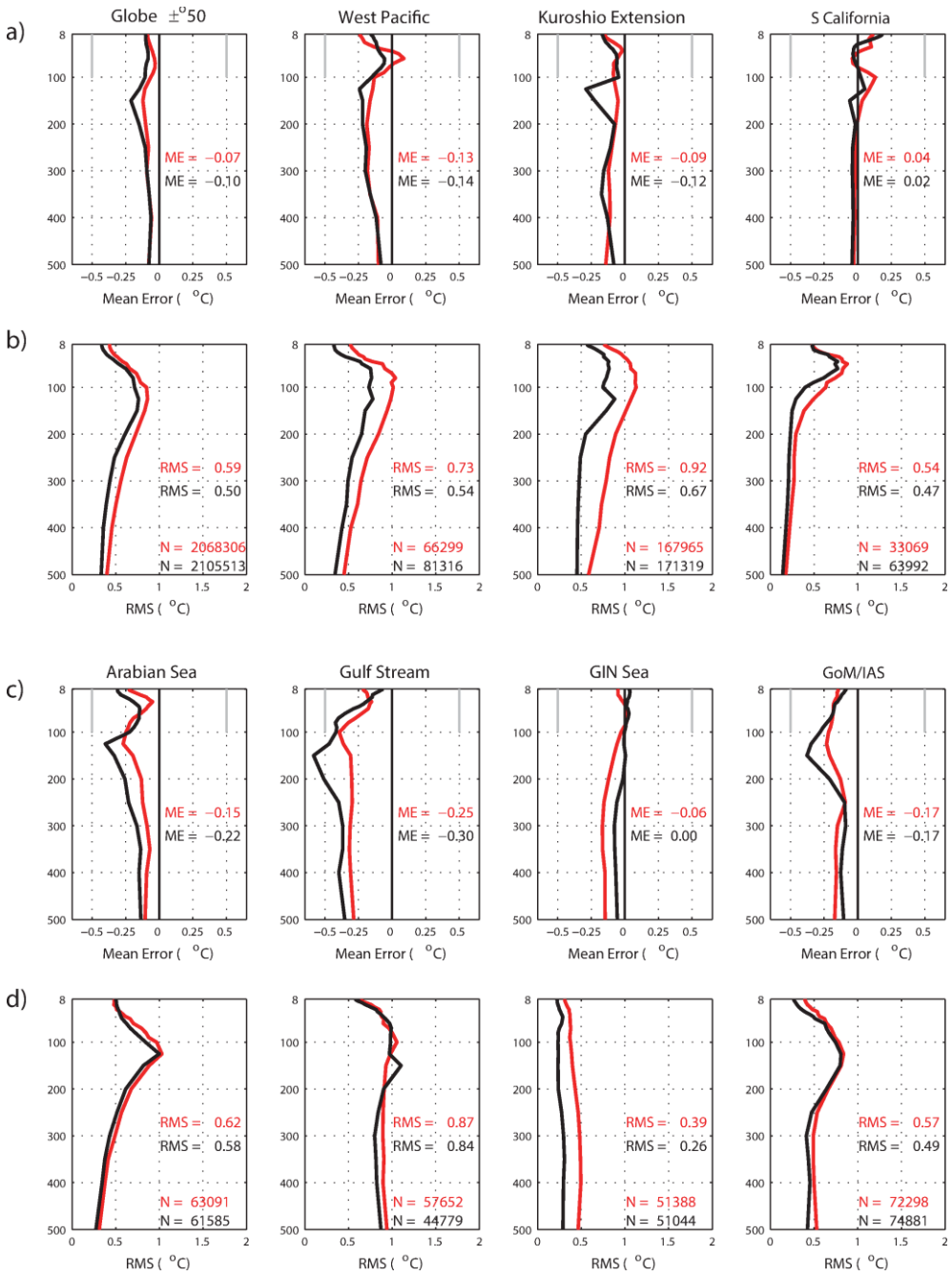

Figure 6: Temperature (°C) vs. depth error analysis in the upper 500 m against unassimilated profile observations at the "nowcast" time for the eight regions defined in Figure 4 spanning the hindcast period July 2014 – June 2015. Rows a) and c) are mean error (ME) while rows b) and d) are root mean square error (RMSE). The red curves are from GOFS 3.0 and the black curves are GOFS 3.1. The numeric values for ME/RMSE are averages from 8-500 m. N = xxxxxx represents the number of observations used in each region, e.g. if a single profile samples 18 times within this depth range, that number is added to the others to define the total. N/20 gives the approximate number of profiles in each region. The minimum depth is 8 m because that is often the last depth sampled by Argo profiles. The gray lines in the ME plots are the tolerances set by NAVOCEANO for the temperature bias in the GOFS 3.0 OPTEST.

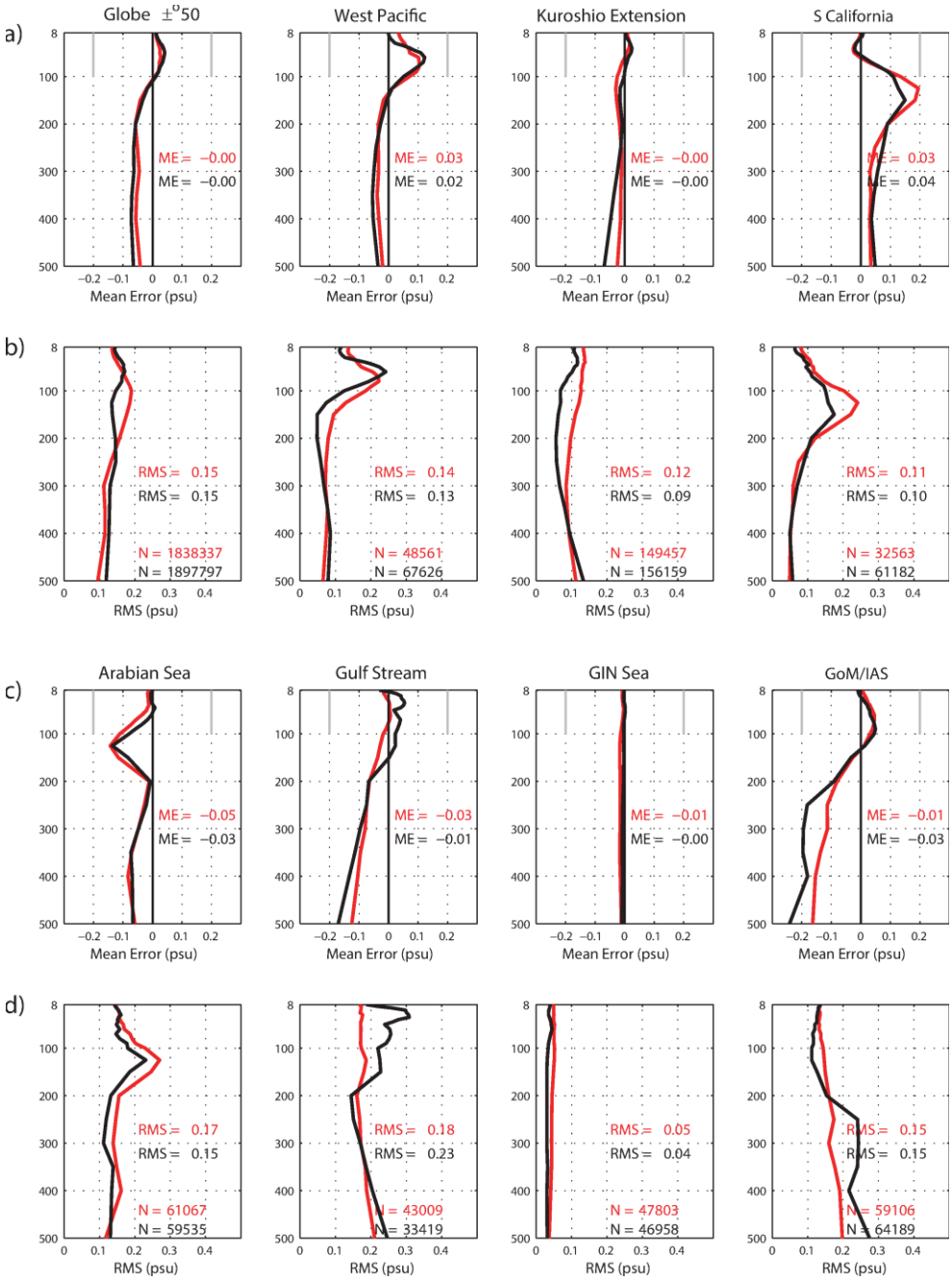

Figure 7: As in Figure 6 except for salinity (psu) vs. depth error analysis in the upper 500 m against unassimilated profile observations.

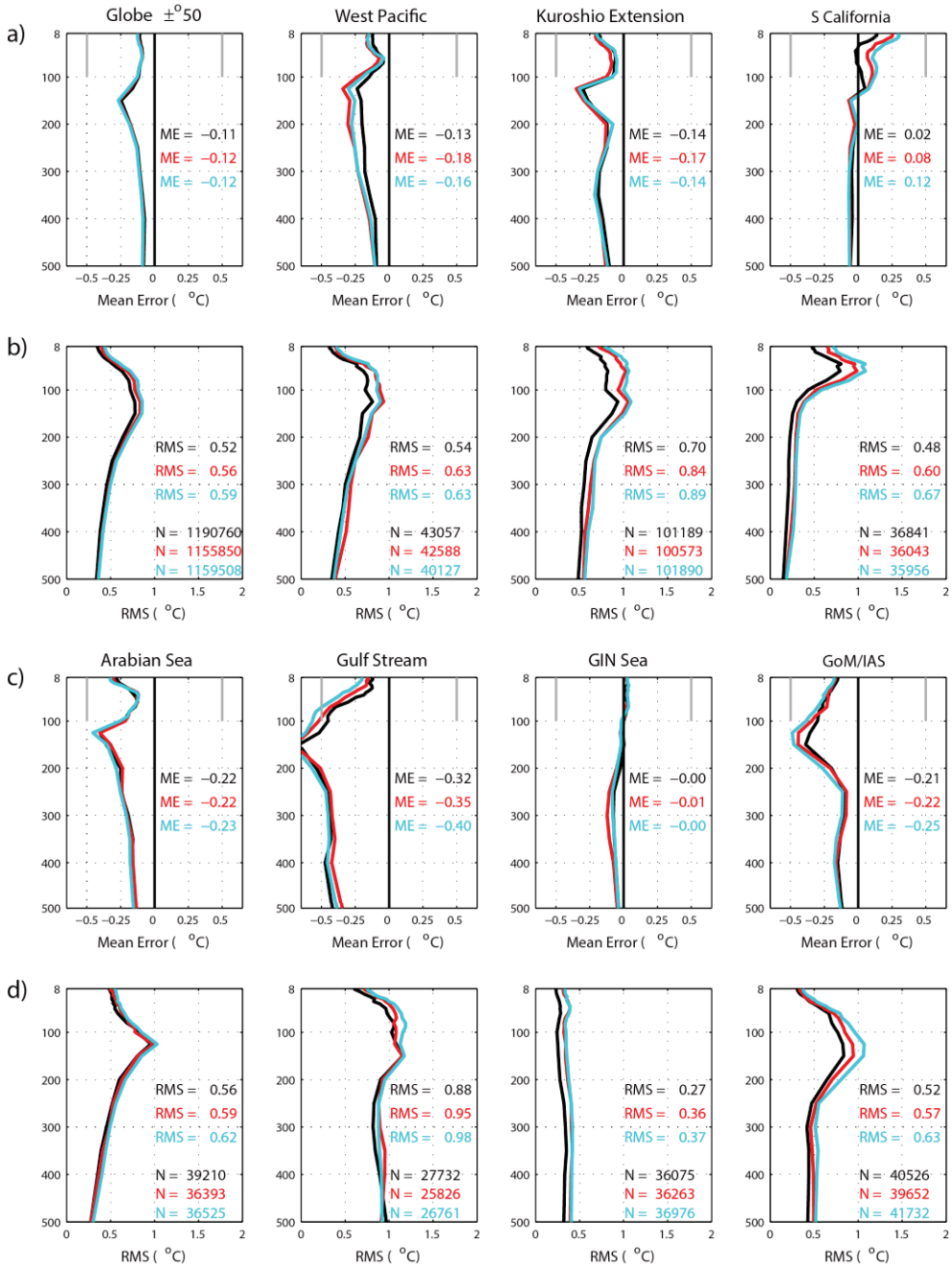

Figure 8: Temperature (°C) vs. depth error analysis in the upper 500 m against unassimilated profile observations for the eight analysis regions for the 5-day forecasts initialized from the hindcast period July 2014 – June 2015. Rows a) and c) are mean error while rows b) and d) are root mean square error. All curves are from GOFS 3.1 with black for the 1-day forecast, red for the 3-day forecast, and cyan for the 5-day forecast. The numeric values for ME/RMSE are averages from 8-500 m. N = xxxxxx represents the number of observations used in each region, e.g. if a single profile samples 18 times within this depth range, that number is added to the others to define the total. N/20 gives the approximate number of profiles in each region. The GOFS 3.1 forecasts used forecast quality NAVGEM 1.4 forcing.

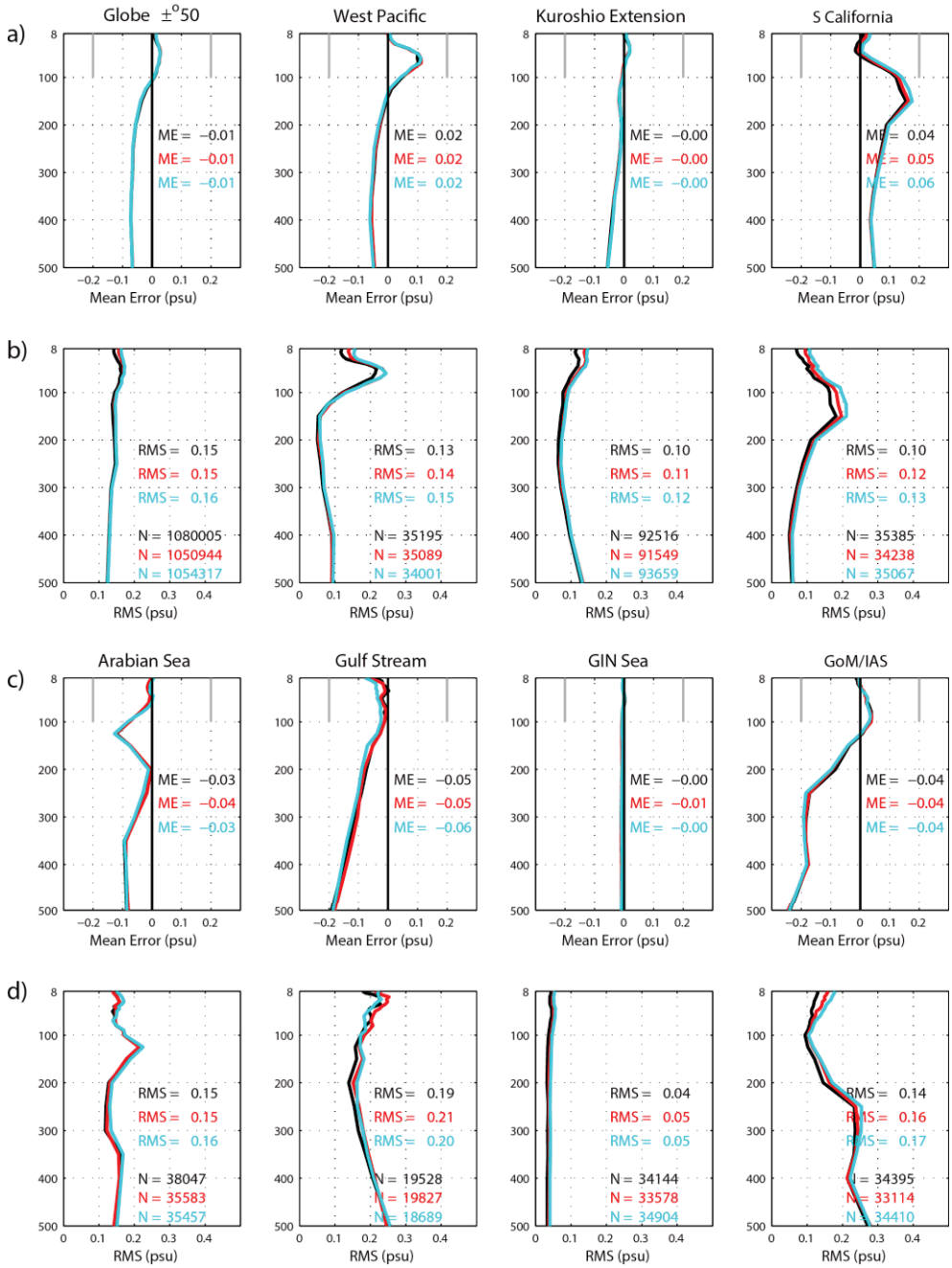

Figure 9: As in Figure 8 except for salinity (psu) vs. depth error analysis in the upper 500 m against unassimilated profile observations.

3.2.2 Acoustical proxy error analyses

Accurate knowledge of the underwater acoustical environment can lead to significant tactical advantages during naval operations. The 3D T and S structure and the upper ocean mixed layer largely

determine the sound speed profile that characterizes the acoustical ducts within the water column. Thus, an ocean nowcast/forecast system must be able to accurately simulate mixed layer depth (MLD), sonic layer depth (SLD), and below layer gradient (BLG). Here MLD is defined by a positive density difference of 0.15 kg/m^3 between the surface and a given depth if both T and S profiles are available, or a negative temperature difference of 0.5°C between the surface and a given depth if only T profiles are available. SLD is the vertical distance from the surface to the depth of the sound speed maximum, often but not always at the base of the mixed layer. Lastly, BLG is defined as the sound speed rate of change with depth per 100 feet in the first 300 feet below the SLD or below the surface if the SLD is absent. These quantities are derived from Naval Oceanographic Office Reference Publication 33 (RP33, 1992), with the exception that the sound speed equation by Chen and Millero (1977) and later correction by Millero and Li (1994) is used rather than that by Wilson (1960).

Histograms of ME and RMSE for MLD, SLD and BLG are shown in Figure 10. For the global domain ±50°, both systems have a shallow bias in MLD and SLD, although it is small (< 2 m) and closer to zero in GOFS 3.1. For the other subdomains, lower ME is about equally divided between GOFS 3.0 and 3.1. In all cases for MLD and SLD, there is lower RMSE in GOFS 3.1. Examining BLG error, the bias is similar over the majority of the analysis regions (including the global domain ±50°), but again GOFS 3.1 has lower RMSE for most analysis regions. The error growth as a function of short-term forecast length depicted in Figure 11 shows a reasonable error growth rate for most subdomains. Unsurprisingly, MLD ME and RMSE are high in the GIN Sea region because of weak stratification.

The spatial distribution of the acoustical proxy error is shown in Figures 12-14. Here median bias error (MdBE) is used and defined as median (model) minus median (observation). For MLD and SLD, GOFS 3.1 has lower error across most of the Pacific Ocean, but slightly higher errors in the Gulf Stream region. BLG is similar between the two but the basin-wide RMSE is lower in GOFS 3.1.

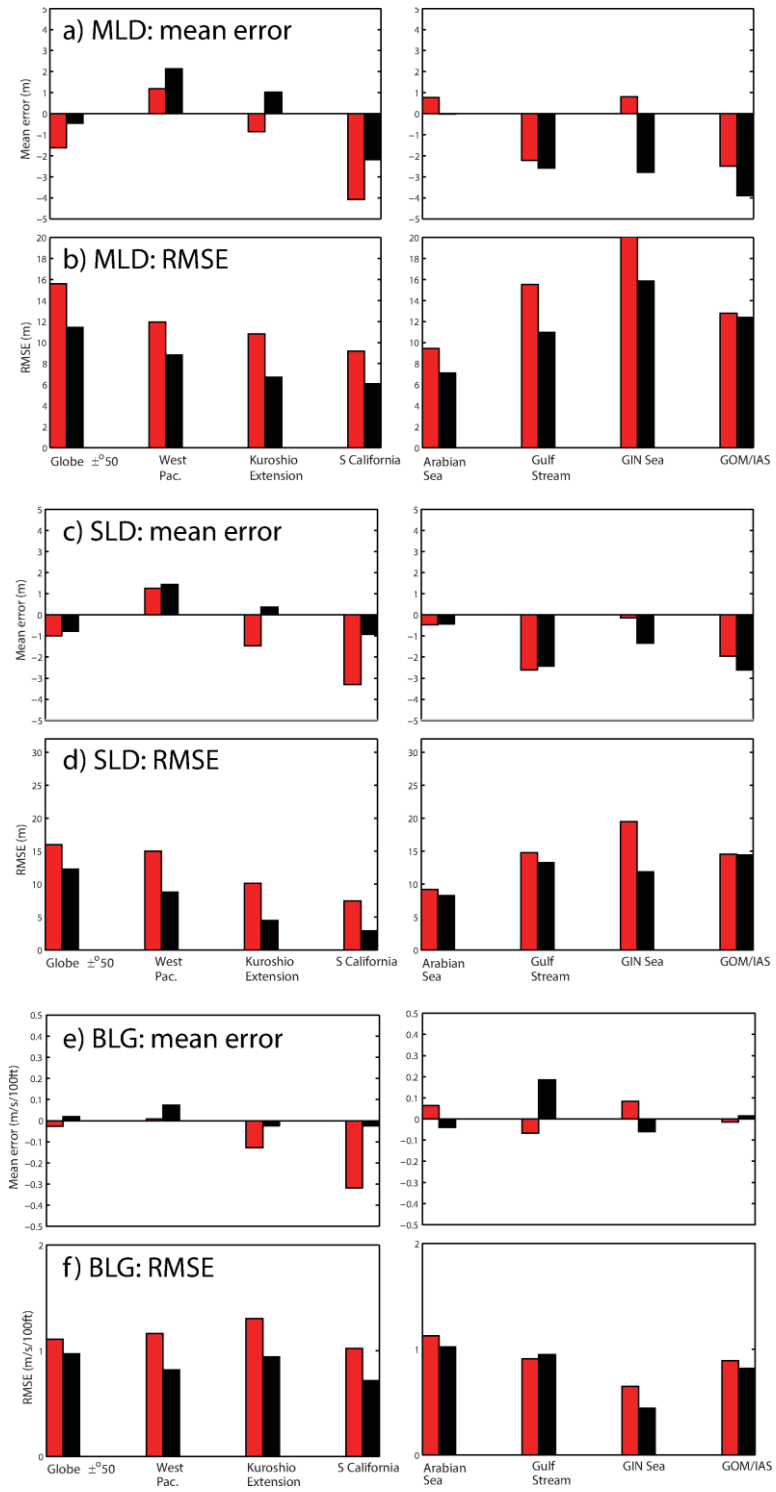

Figure 10: Histograms of mixed layer depth (a, b), sonic layer depth (c, d), and below layer gradient (e, f) mean error (a, c and e) and root mean square error (b, d and f) against unassimilated profile observations for GOFS 3.1 (black) and GOFS 3.0 (red) at the "nowcast" time for the eight analysis regions spanning the hindcast period July 2014 – June 2015. Units for MLD and SLD are meters and units of BLG are m/s/100 ft.

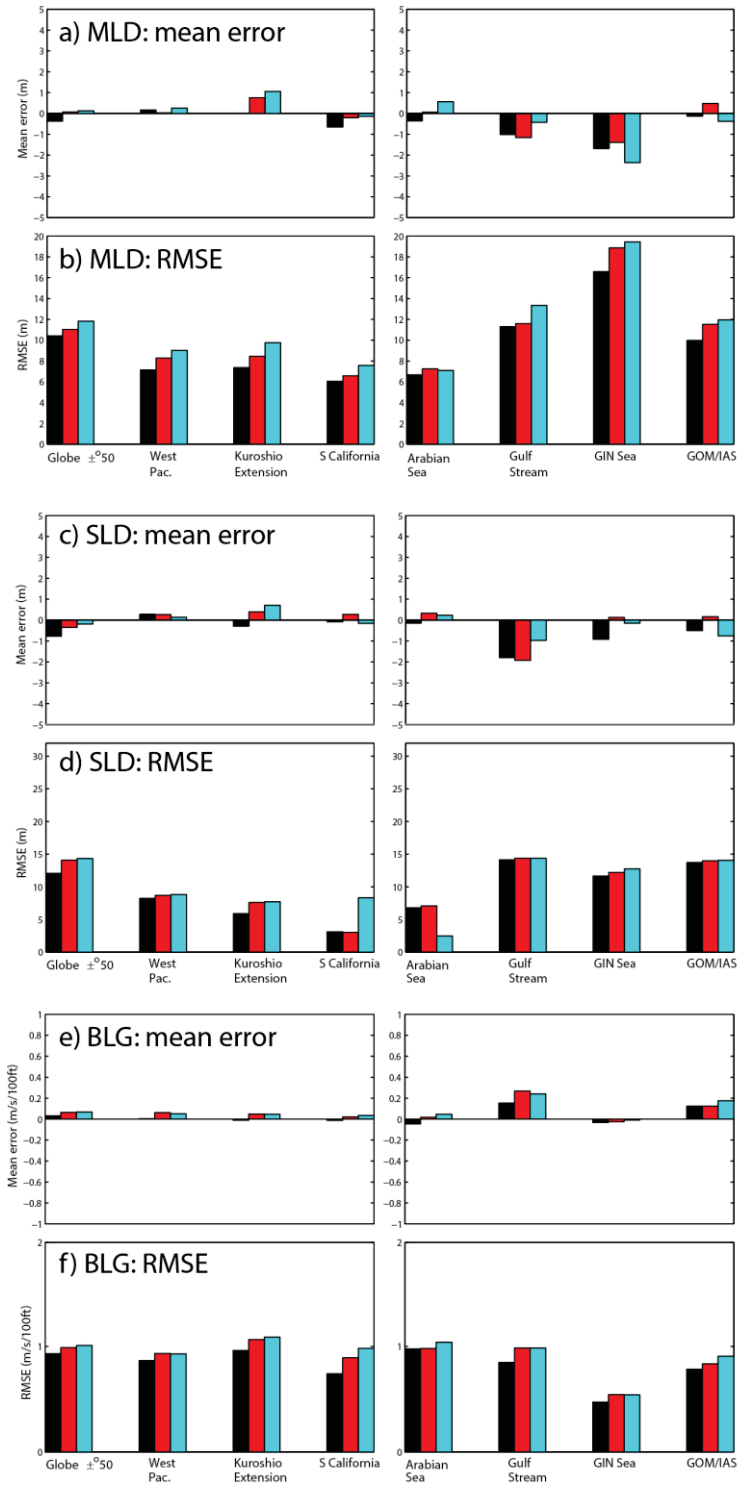

Figure 11: As in Figure 10 except for the GOFS 3.1 1-day (black), 3-day (red) and 5-day forecasts initialized from the hindcast period July 2014 – June 2015.

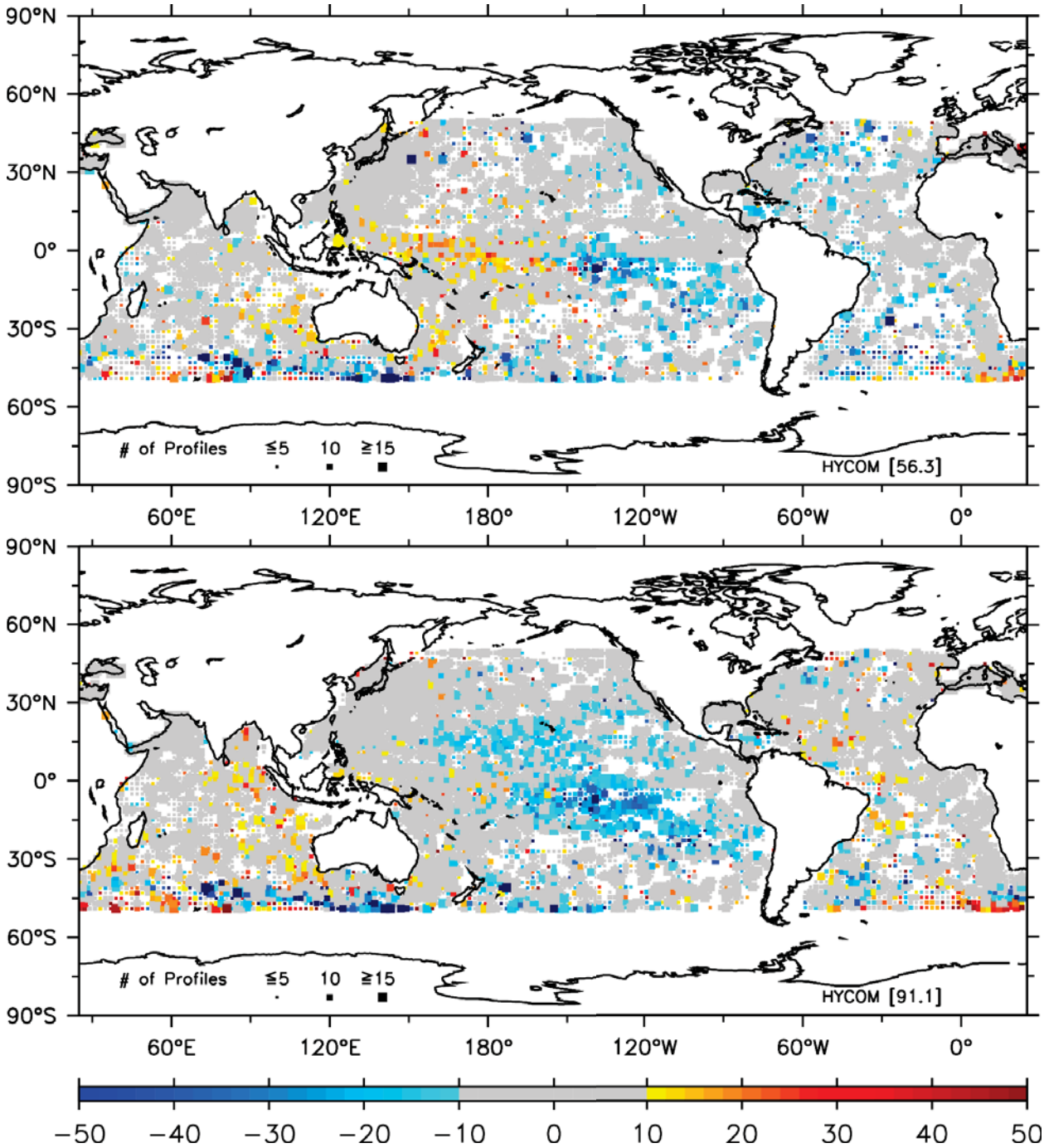

Figure 12: MLD median bias error (MdBE) (m) for GOFS 3.1 (top) and GOFS 3.0 (bottom) against unassimilated profiles over the hindcast period July 2014 – June 2015. Red values indicate the nowcast MLD is deeper than observed while blue values indicate it is shallower. The data are averaged over 2° bins and the number of profiles within each bin is indicated by the size of each individual square as denoted by the legend in Antarctica. The MdBE and RMSE (respectively) for GOFS 3.1 is -0.1 m and 30.0 m, whereas it is -0.8 m and 34.5 m for GOFS 3.0 over this domain.

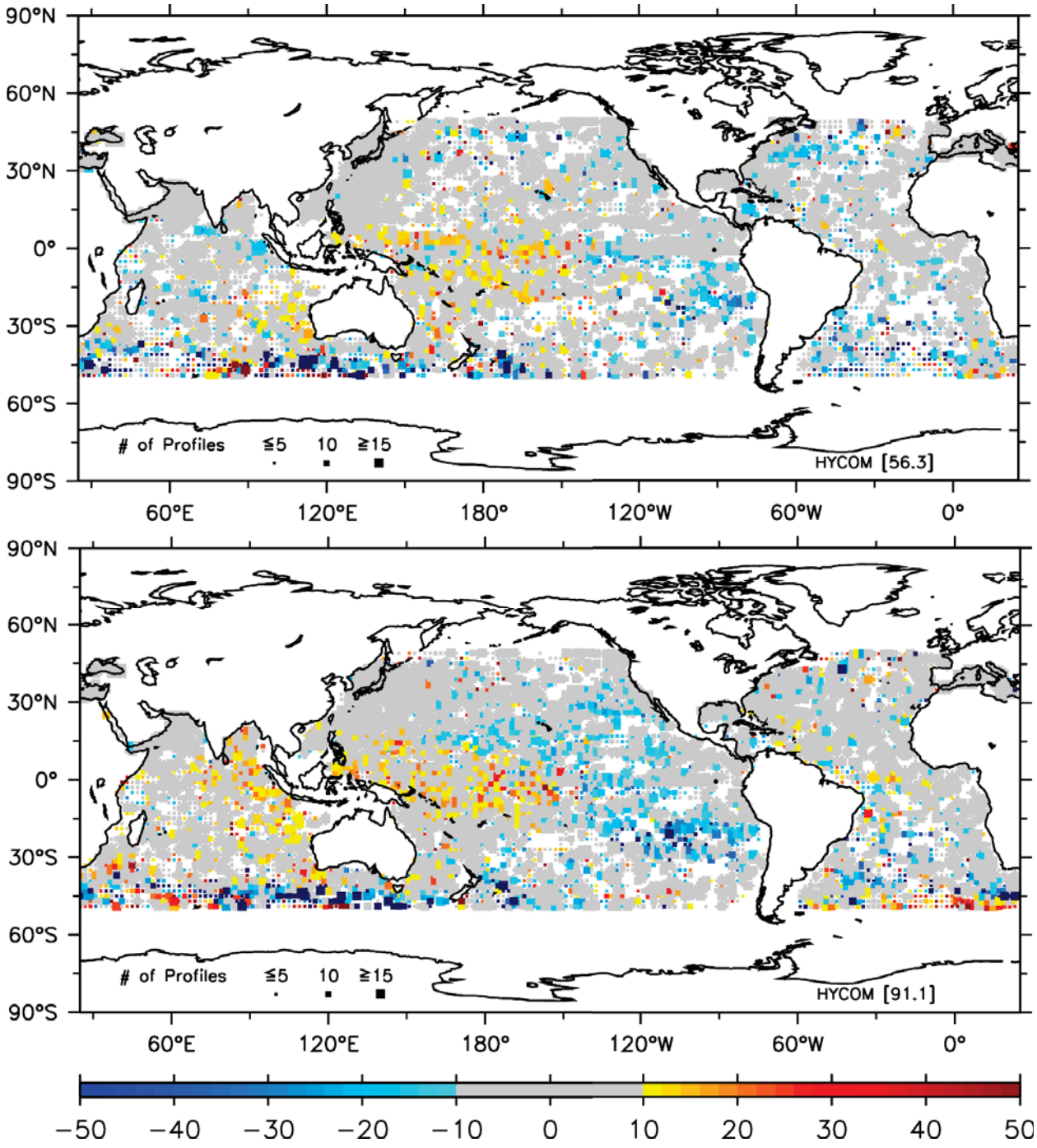

Figure 13: As in Figure 12 except for SLD. The MdBE and RMSE (respectively) for GOFS 3.1 is 0 m and 51.4 m, whereas it is 0 m and 51.4 m for GOFS 3.0 over this domain.

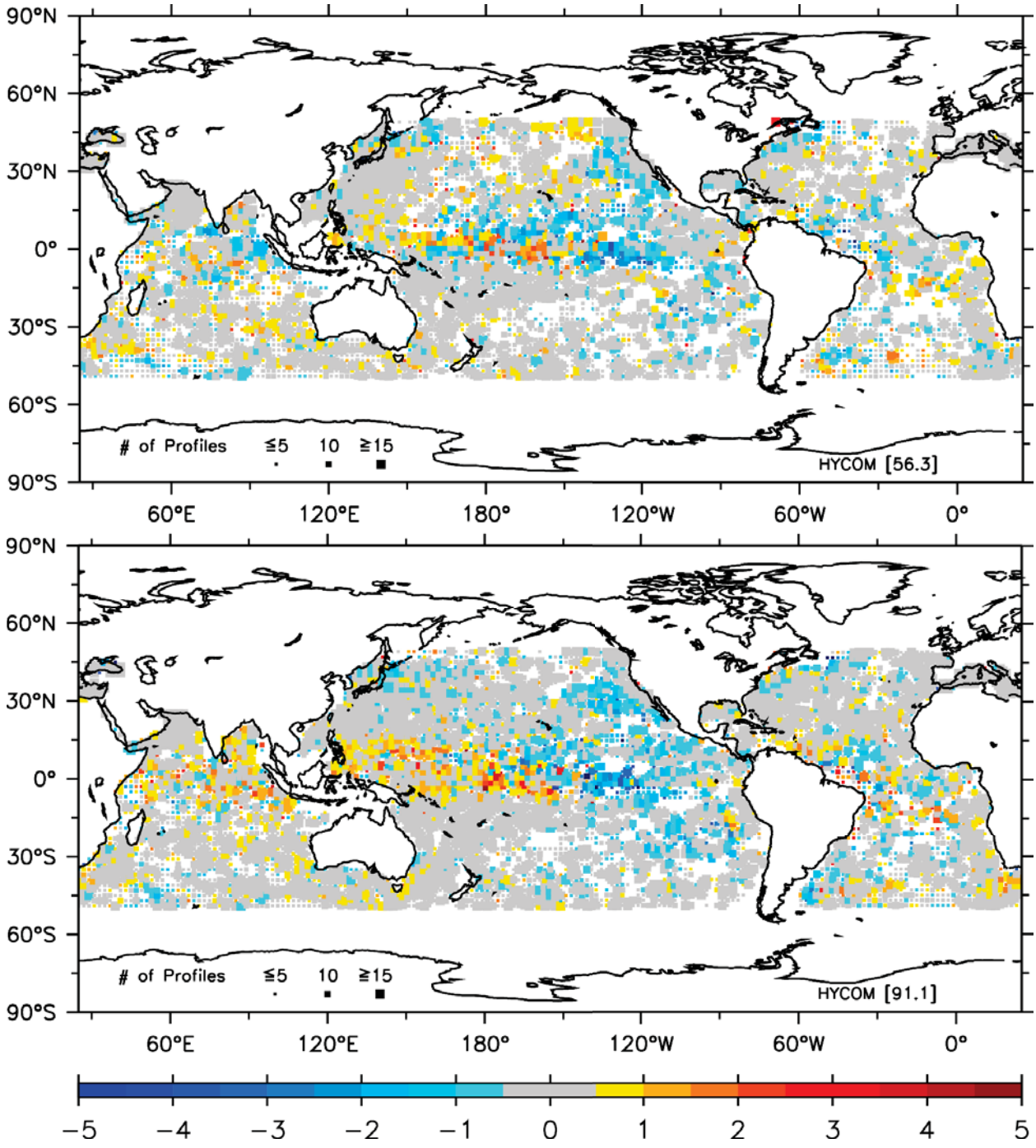

Figure 14: As in Figure 12 except for BLG. The MdBE and RMSE (respectively) for GOFS 3.1 is 0.0 m/s/100 ft. and 1.5 m/s/100 ft., whereas it is 0.0 m/s/100 ft. and 1.7 m/s/100 ft. for GOFS 3.0 over this domain.

3.2.3 Upper ocean velocity validation

Lastly for the ocean, GOFS 3.1 and 3.0 upper ocean velocity is compared against drifting buoys obtained from the National Oceanic and Atmospheric Administration Global Drifter Program. Observed SST from these drifters may be assimilated into both systems (if it is reported in a timely fashion and appears on Global Telecommunications System), but the velocities are not and thus they are an independent validation data set for this error analysis. The drifters have a drogue at 15 m and instantaneous GOFS velocity output are extracted at this depth. In Metzger et al. (2015), separation distance was used as a metric for upper ocean velocity. Here the error analyses are changed to ME and RMSE of 15 m speed, and vector correlation (Crosby et al., 1993) to provide directional error. When these new metrics were applied to the previously transitioned GOFS 3.1 output (with sub-optimal ISOP), the slower upper ocean velocities relative to GOFS 3.0 were evident. Thus, we are confident they provide a more accurate assessment of near surface currents than separation distance.

Figure 15 shows speed ME, RMSE and vector correlation as a function of forecast length for the 8 analysis regions. For the majority of the regions (including global ±50°), the speed bias in GOFS 3.1 is closer to zero, while GOFS 3.0 shows a fast bias nearly everywhere. For all of the analysis regions, the RMSE of speed is lower and directional vector correlation is higher in GOFS 3.1 than in GOFS 3.0. The degradation in skill as a function of forecast length is smaller for speed than it is for direction.

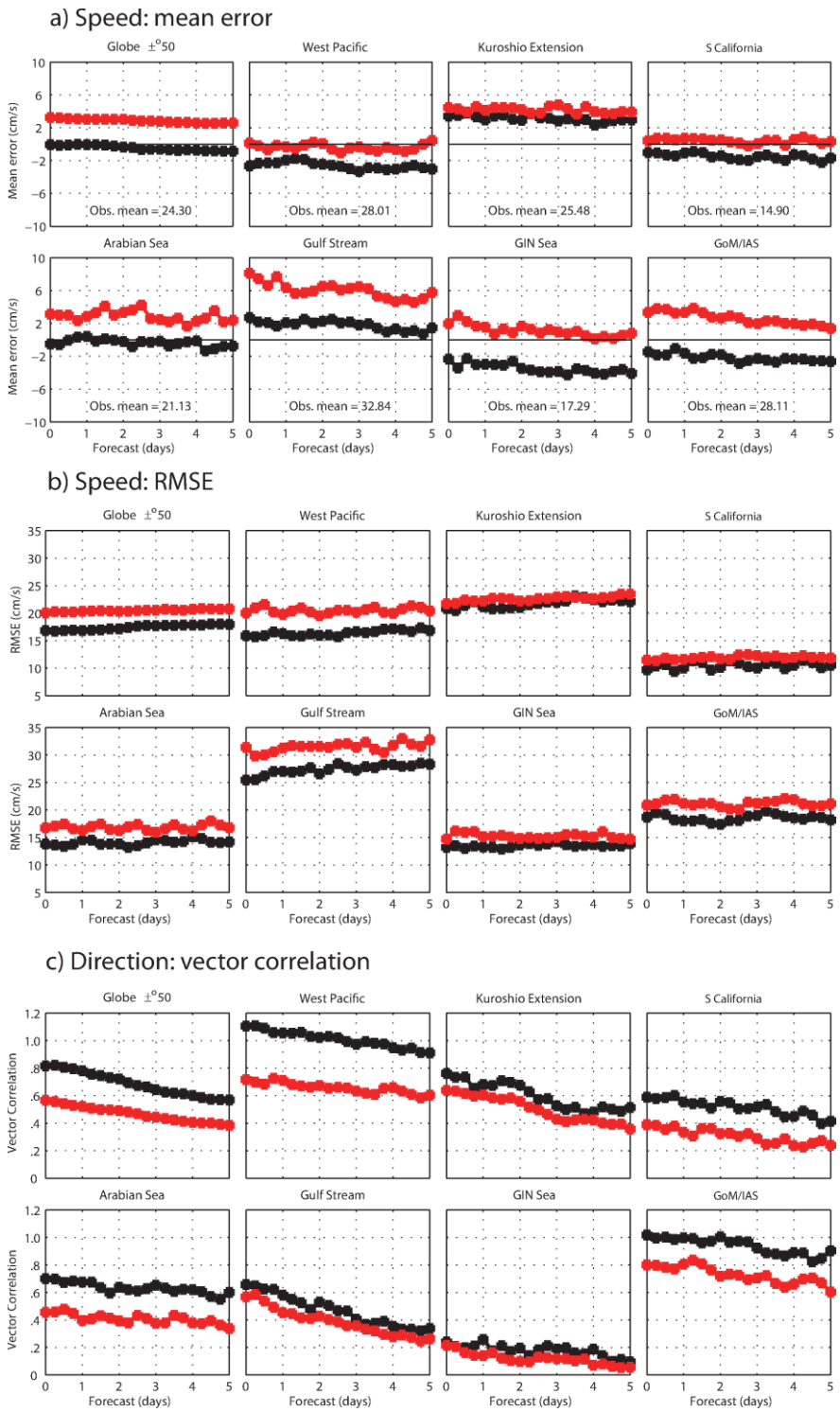

Figure 15: Speed mean error (cm/s) (a) and RMSE (cm/s) (b), and direction vector correlation (c) of unassimilated drifting buoys at 15 m depth compared against GOFS 3.1 (black) and GOFS 3.0 (red) as a function of forecast length for the eight analysis regions. The mean observed drifting buoy speed for each region is noted in the top eight subpanels. The vector correlation is based on Crosby et al. (1993) with 0 indicating no correlation and 2 indicating perfect correlation. This analysis spans the hindcast period July 2014 – June 2015.

3.3 Ice validation

As mentioned previously, validation of GOFS 3.1 ice is performed relative to operational ACNFS for the Northern Hemisphere. GOFS 3.1 also provides nowcasts and forecasts of Southern Hemisphere sea ice and while there is not another US Navy operational system to compare against, the ice edge error analyses are presented here for completeness. Six regions in each hemisphere are chosen that correspond to the regional seas in the polar latitudes (Figure 16).

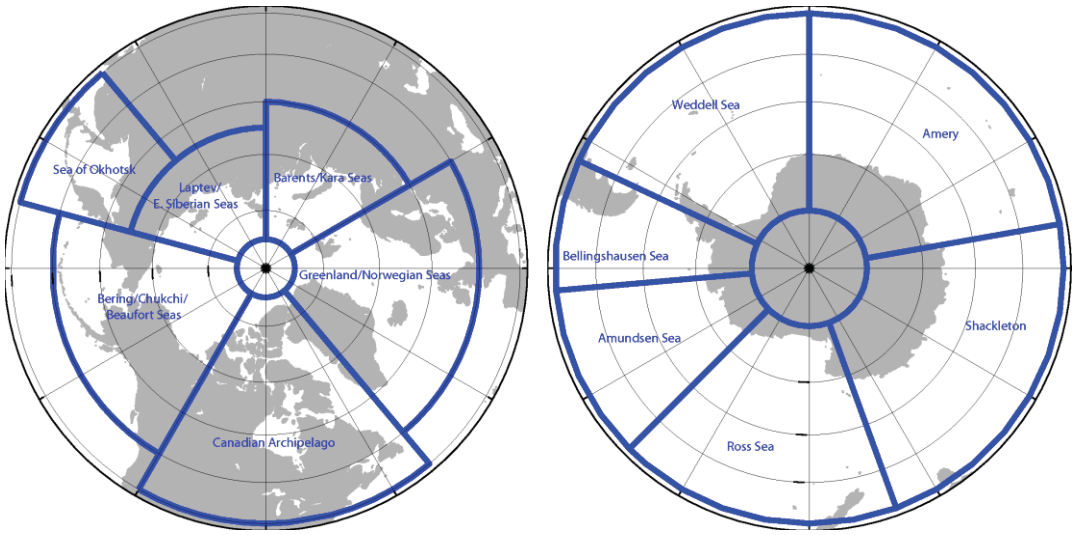

Figure 16: Arctic (left) and Antarctic (right) ice validation analysis regions defined by the regional seas in each hemisphere.

Ice extent in GOFS 3.1 and ACNFS is largely controlled by the assimilation of satellite ice concentration data. Because the two systems assimilate the same data streams (after 3 February 2015), there is little difference in ice extent, but the same cannot be said for ice thickness (and ice concentration to a lesser degree) which can be attributed to the changes in ice assimilation described in section 2.3.2. Figure 17 shows ice concentration, ice thickness and their differences between ACNFS and GOFS 3.1. The largest concentration differences are seen in the marginal ice zone where the assimilation is also different. The largest thickness differences are in the Beaufort Sea, with GOFS 3.1 depicting thinner ice than ACNFS. This is also generally true for the vast majority of the Arctic. These thickness differences may be attributed to the initial conditions that started each assimilative hindcast. ACNFS was initialized

with a much thicker ice state than GOFS 3.1, and has maintained the thick ice over many years of integration.

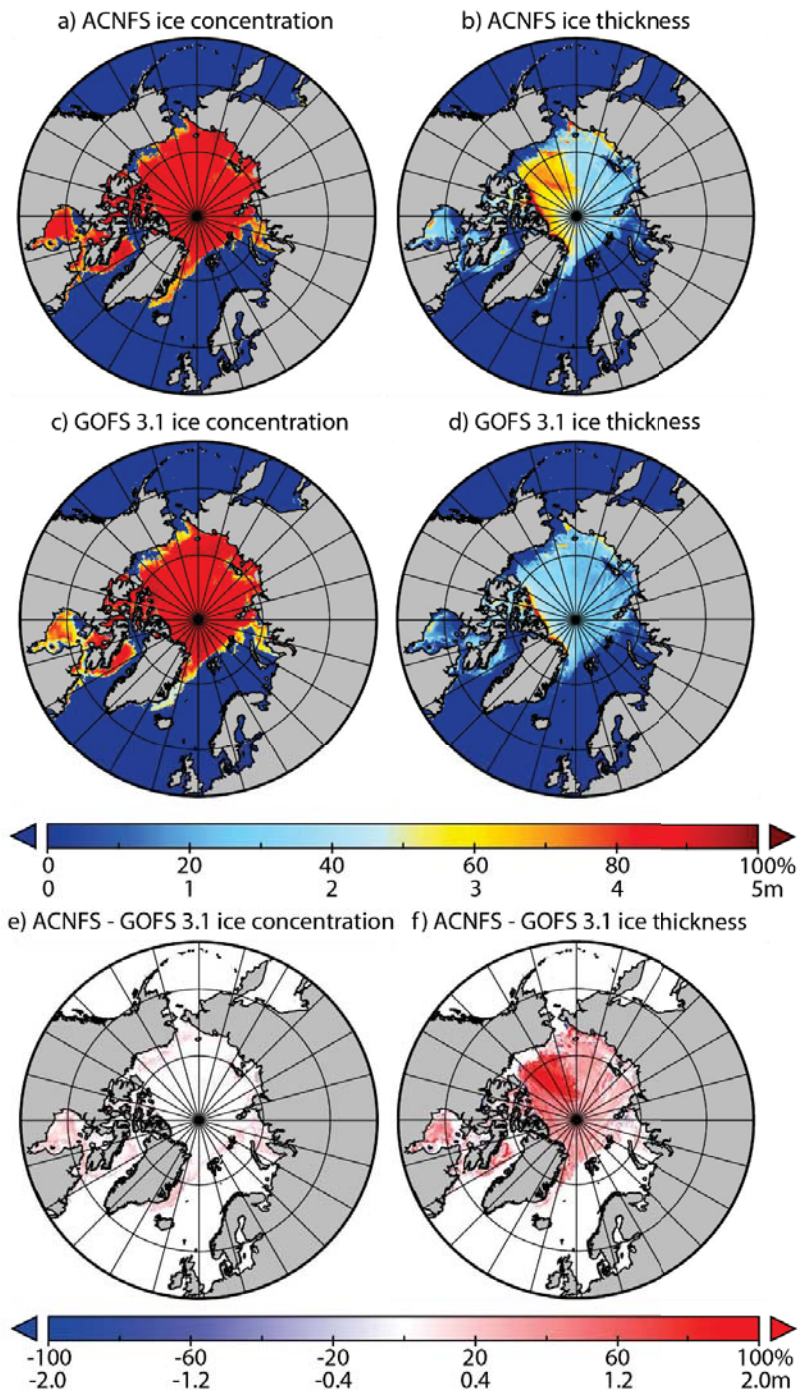

Figure 17: Arctic Ocean (a,c) ice concentration (%) and (b,d) ice thickness (m) for 20 June 2016 00Z from ACNFS (a,b) and GOFS 3.1 (c,d). Difference panels (ACNFS minus GOFS 3.1) are shown in e) for ice concentration and in f) for ice thickness. The GOFS 3.1 (ACNFS) panels are twelve (six) hours after the NCODA ice analysis has been directly inserted into CICE.

3.3.1 Ice edge location – Northern Hemisphere

Maritime operations in polar latitudes rely on an accurate prediction of the ice edge. The NIC produces a daily ice edge analysis for both hemispheres based on a variety of satellite sources (visible images, infrared, scatterometer, synthetic aperture radar (SAR), and passive microwave data) and define the ice edge as areas of <10% sea ice concentration ('Open Water'). Here it is used as an independent observation type and compared against GOFS 3.1 and ACNFS (Figure 18). Model ice edge locations are defined as those grid points that exceed a certain threshold value for ice concentration and that also have a neighboring point that falls below that value. In this case a threshold of 5% was used to determine the model ice edge. The distances between each NIC observed point and the nearest model-derived ice edge location were then calculated, from which a daily mean was computed for each model day. The impact of AMSR2 and IMS assimilation in ACNFS (Posey et al., 2015) is easily seen at the beginning of February 2015. In the last five months of the hindcast, 24-hour forecast ice edge error between the two systems is more similar, but GOFS 3.1 shows lower error (Table 1).

Region	GOFS 3.1	ACNFS	% change
Pan Arctic	22.2	25.2	12
GIN Seas	20.6	26.7	23
Barents/Kara Seas	24.2	26.3	8
Sea of Okhotsk	18.9	21.9	14
Bering/Chukchi/Beaufort Seas	21.1	25.5	17
Canadian Archipelago	23.8	24.0	1
"Wins"	6	0	

Table 1: Mean 24-hour forecast ice edge location error (km) of GOFS 3.1 and ACNFS compared against the NIC observed ice edge for the validation period February – June 2015 for the six analysis regions of the Arctic Ocean. Lower error is highlighted with green. During this time frame, both GOFS 3.1 and ACNFS use the same ice assimilation data sources.

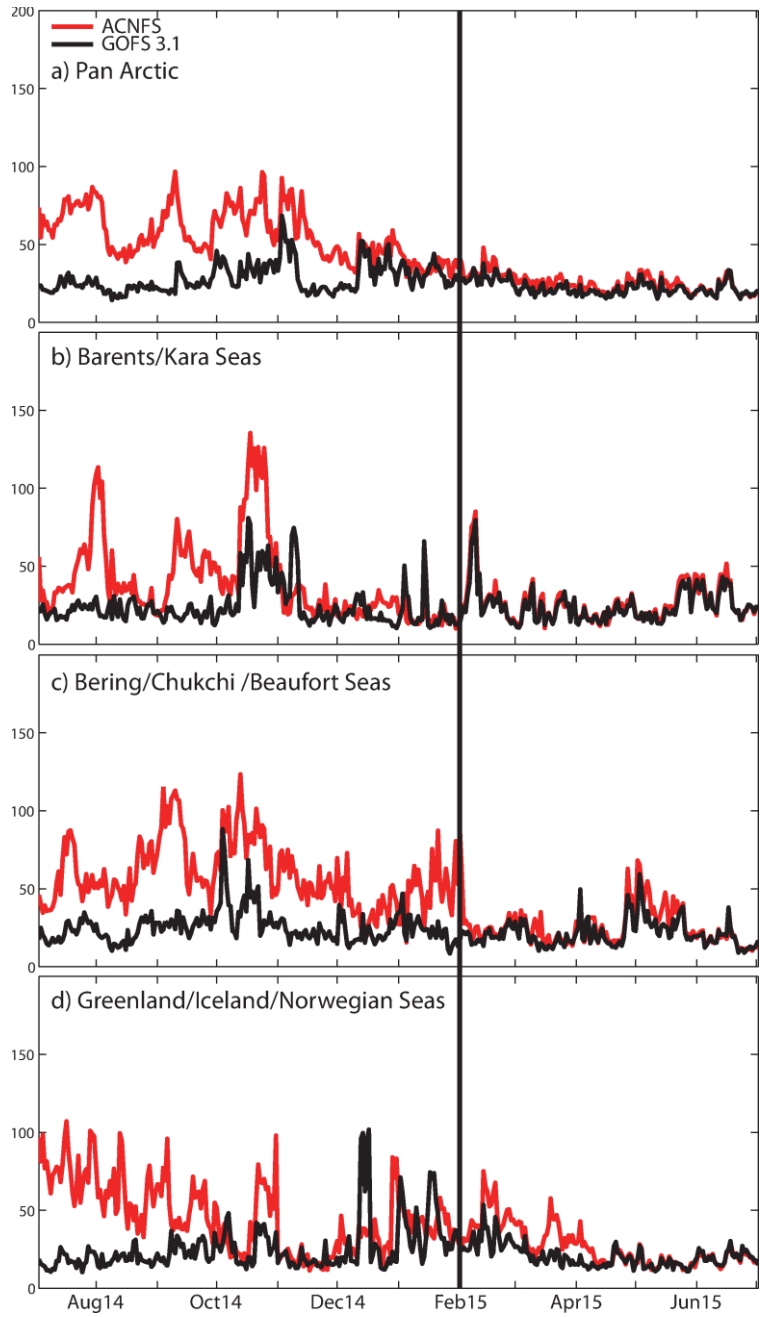

Figure 18: Daily 24-hour forecast ice edge location error (km) versus time for GOFS 3.1 (black) and ACNFS (red) ice edge (defined as the 5% ice concentration) against the independent ice edge analysis from the NIC over the validation period July 2014 – June 2015 for four Arctic Ocean regions: a) the Pan Arctic Ocean, b) the Barents/Kara Seas, c) the Bering/Chukchi/Beaufort Seas and d) the Greenland/Iceland/Norwegian Seas. All comparisons are performed at 00Z. The NCODA ice analysis is performed at 12Z (18Z) in GOFS 3.1 (ACNFS) and directly inserted into CICE, and so this represents a 12 (6) hour forecast. The black vertical line indicates when ACNFS began assimilating the blended AMSR2 and IMS products. The GOFS 3.1 output is from the daily output of GLBb0.08-56.3.

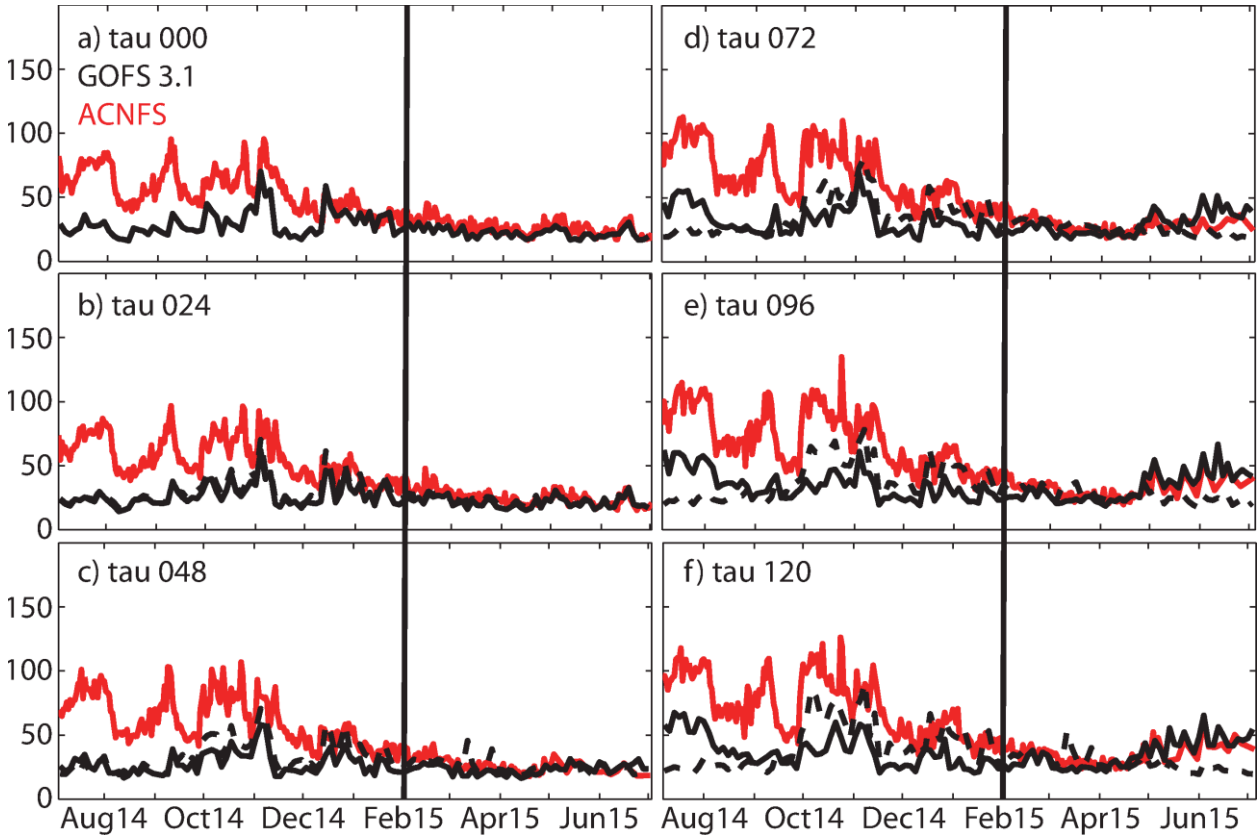

Figure 19: Ice edge location error (km) versus time for the 0 through 5-day forecast for GOFS 3.1 (black) and ACNFS (red) ice edge against the independent ice edge analysis from the NIC for the Pan Arctic Ocean. In addition, the dashed black line is a "forecast" of GOFS 3.1 persistence, i.e. same nowcast field is used throughout. The black vertical line indicates when ACNFS began assimilating the blended AMSR2 and IMS products. The GOFS 3.1 output is from the GLBb0.08-56.4 forecasts (black solid lines) initialized from the GLBb0.08-56.3 hindcast (black dashed lines).

Ice edge location error as a function of forecast length (out to 5 days) for the Pan Arctic region is shown in Figure 19, along with persistence of the GOFS 3.1 nowcast ice state. Results for all the analysis regions are shown in Table 2 for the period of common ice assimilation in GOFS 3.1 and ACNFS. Note that the tau 024 values of Table 2 are not identical to Table 1 because forecasts are only made every third day (GLBb0.08-56.4), while Table 1 results are for every day of the hindcast (GLBb0.08-56.3). Performance varies regionally with GOFS 3.1 producing lower forecast ice edge error at all times for the GIN Seas, but ACNFS is better at most times for the Canadian Archipelago region. Overall, GOFS 3.1 has more "wins" from the nowcast through 2-day forecasts, whereas ACNFS has lower error at longer

forecast times. Note also that for some regions GOFS 3.1 exhibits a rather large jump in ice edge error between tau 048 and tau 072, but the cause for this has not yet been determined.

Pan Arctic Region			**Barents/Kara Seas**		
Tau	GOFS 3.1	ACNFS	Tau	GOFS 3.1	ACNFS
000	22.2	26.7	000	28.3	34.6
024	21.4	25.2	024	23.9	27.5
048	23.3	26.4	048	24.1	26.3
072	29.2	28.2	072	24.5	25.7
096	33.4	31.2	096	27.1	26.7
120	35.3	33.4	120	30.1	28.8
"wins"	3	3	"wins"	4	2
Bering/Chukchi/Beaufort Seas			**Greenland/Iceland/Norwegian Seas**		
Tau	GOFS 3.1	ACNFS	Tau	GOFS 3.1	ACNFS
000	22.2	27.5	000	22.1	28.8
024	21.0	25.5	024	20.2	26.7
048	22.8	25.8	048	20.7	29.7
072	28.9	27.6	072	24.0	30.4
096	32.4	31.0	096	27.1	33.4
120	34.5	33.0	120	28.3	34.6
"wins"	3	3	"wins"	6	0
Sea of Okhotsk			**Canadian Archipelago**		
Tau	GOFS 3.1	ACNFS	Tau	GOFS 3.1	ACNFS
000	20.4	23.2	000	23.4	25.1
024	19.3	21.9	024	24.1	24.0
048	24.3	23.5	048	26.6	24.6
072	34.5	26.7	072	34.7	27.9
096	38.5	29.7	096	41.3	32.2
120	40.8	33.2	120	42.2	34.4
"wins"	2	4	"wins"	1	5

Table 2: GOFS 3.1 and ACNFS ice edge error (km) as a function of forecast length compared against the observed NIC ice edge for the validation period February – June 2015 for the six analysis regions of the Arctic Ocean. Lower errors/higher "wins" are highlighted with green. During this time frame, both GOFS 3.1 and ACNFS use the same ice assimilation data sources.

Figure 20 shows the GOFS 3.1 and ACNFS ice concentration difference from a 5-day forecast minus the verifying analysis for 3 June 2015 during boreal summer. The red areas indicate less forecast ice, and blue indicates more forecast ice. For this date, it is clear that GOFS 3.1 has more ice melt than a 5-day ACNFS forecast, hence the lower ACNFS ice edge error at the longer forecast times.

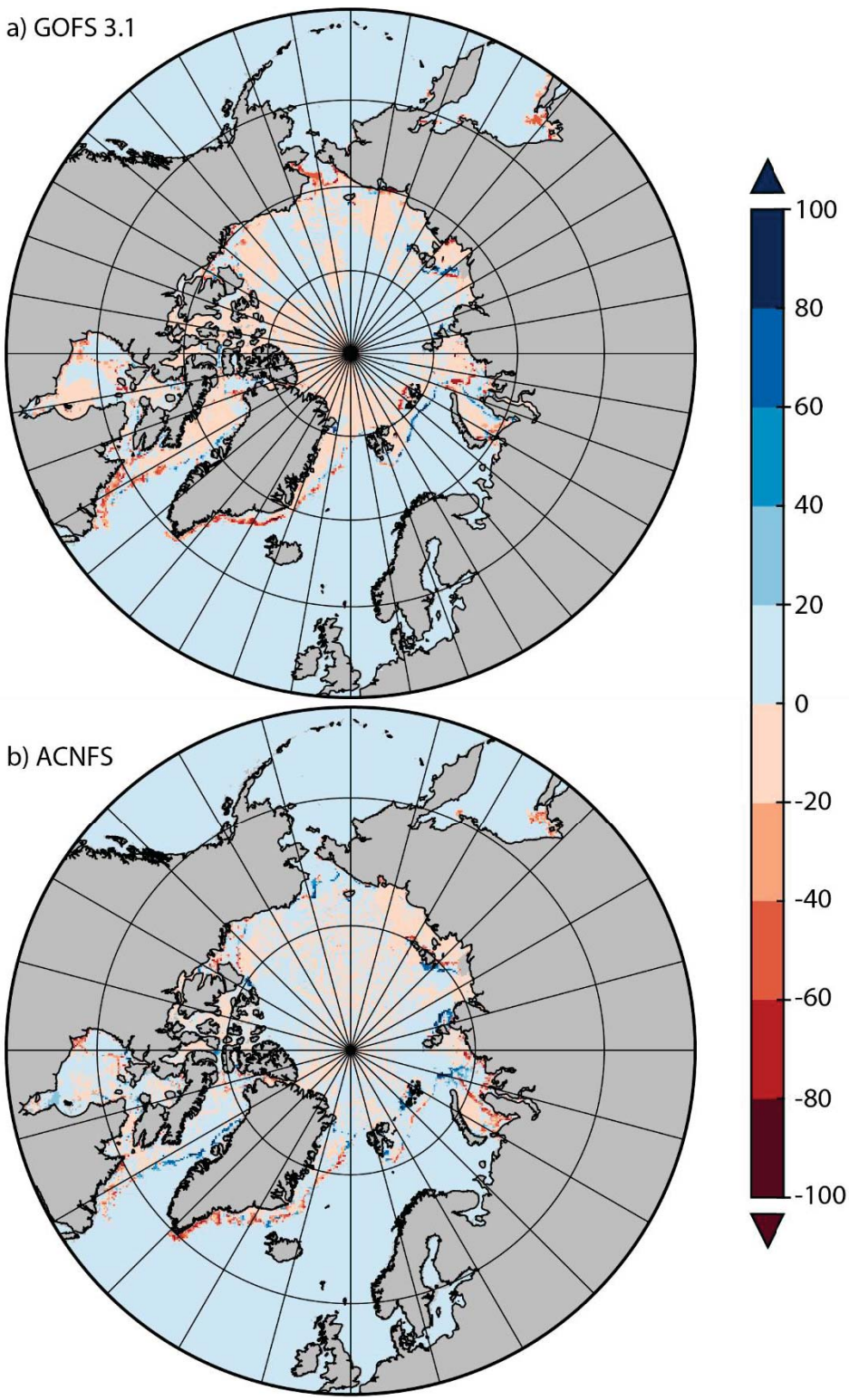

Figure 20: a) GOFS 3.1 and b) ACNFS ice concentration difference (%) of the 5-day forecast minus the verifying analysis on 3 June 2015. Red indicates less forecast ice (i.e. more melt) and blue indicates more forecast ice (i.e. more freeze-up).

Figure 19 also provides insight of GOFS 3.1 ice edge forecasts against the nowcast GOFS 3.1 ice edge persisted for the duration of the 5-day forecast. At long forecast times (Figure 19f), GOFS 3.1 forecasts are better than persistence during the freeze season, but less skillful than persistence during the melt season. Again, this is due to excessive ice melt as GOFS 3.1 forecast length increases. Table 3 compares GOFS 3.1 ice edge error as a function of forecast length versus GOFS 3.1 persistence. GOFS 3.1 forecasts have lower error than GOFS 3.1 persistence out through tau 072, but beyond that GOFS 3.1 persistence provides better forecasts. Table 4 shows GOFS 3.1 ice edge error as a function of forecast length versus ACNFS persistence and again demonstrates that GOFS 3.1 exhibits similar skill out to the tau 072 forecast length.

Pan Arctic Region			Barents/Kara Seas		
Tau	GOFS 3.1	GOFS 3.1 Persistence	Tau	GOFS 3.1	GOFS 3.1 Persistence
000	26.6	26.6	000	24.5	24.5
024	24.8	27.5	024	23.6	26.1
048	26.2	29.9	048	24.7	28.8
072	31.3	31.9	072	27.0	31.2
096	34.5	33.6	096	29.2	32.9
120	36.7	36.0	120	31.4	35.0
Bering/Chukchi/Beaufort Seas			**Greenland/Iceland/Norwegian Seas**		
Tau	GOFS 3.1	GOFS 3.1 Persistence	Tau	GOFS 3.1	GOFS 3.1 Persistence
000	25.1	25.1	000	25.4	25.4
024	23.4	25.8	024	23.5	26.6
048	24.9	28.0	048	22.9	28.3
072	30.5	30.9	072	26.5	29.2
096	33.2	32.8	096	30.4	31.3
120	35.1	35.2	120	32.1	32.6
Laptev/East Siberian Seas			**Sea of Okhotsk**		
Tau	GOFS 3.1	GOFS 3.1 Persistence	Tau	GOFS 3.1	GOFS 3.1 Persistence
000	29.1	29.1	000	25.3	25.3
024	28.1	30.2	024	23.5	25.8
048	29.1	32.6	048	28.0	28.1
072	35.0	34.6	072	36.3	31.2
096	38.4	38.3	096	39.1	32.8
120	42.0	42.0	120	41.4	34.1
Canadian Archipelago			**Total "wins"**		
Tau	GOFS 3.1	GOFS 3.1 Persistence	Tau	GOFS 3.1	GOFS 3.1 Persistence
000	30.1	30.1	000	–	–
024	28.5	30.8	024	7	0
048	30.0	33.0	048	7	0
072	36.9	33.6	072	4	3
096	40.8	35.2	096	2	5
120	42.9	38.1	120	2	5

Table 3: GOFS 3.1 and GOFS 3.1 persistence ice edge error (km) as a function of forecast length compared against the observed NIC ice edge for the validation period July 2014 – June 2015 for the seven analysis regions of the Arctic Ocean. The lower right quadrant also summarizes the total number of "wins". Lower errors/higher "wins" are highlighted with green.

Pan Arctic Region			Barents/Kara Seas		
Tau	GOFS 3.1	ACNFS Persistence	Tau	GOFS 3.1	ACNFS Persistence
000	22.2	26.7	000	23.8	27.5
024	21.4	27.6	024	24.1	29.2
048	23.3	28.6	048	24.5	31.0
072	29.2	29.7	072	27.1	32.6
096	33.3	30.6	096	30.1	33.7
120	35.3	31.7	120	32.1	34.6
Bering/Chukchi/Beaufort Seas			**Greenland/Iceland/Norwegian Seas**		
Tau	GOFS 3.1	ACNFS Persistence	Tau	GOFS 3.1	ACNFS Persistence
000	22.2	27.5	000	22.1	28.8
024	21.0	27.8	024	20.2	29.8
048	22.8	28.8	048	20.7	30.8
072	28.9	30.0	072	24.0	31.7
096	32.4	31.2	096	27.1	32.4
120	34.4	33.1	120	28.3	32.6
Laptev/East Siberian Seas			**Sea of Okhotsk**		
Tau	GOFS 3.1	ACNFS Persistence	Tau	GOFS 3.1	ACNFS Persistence
000	–	–	000	20.4	23.2
024	–	–	024	19.3	23.7
048	–	–	048	24.3	24.4
072	–	–	072	34.5	25.1
096	–	–	096	38.6	25.7
120	–	–	120	40.8	26.6
Canadian Archipelago			**Total "wins"**		
Tau	GOFS 3.1	ACNFS Persistence	Tau	GOFS 3.1	ACNFS Persistence
000	23.4	25.1	000	6	0
024	24.1	25.8	024	6	0
048	26.6	26.6	048	5	0
072	34.7	27.6	072	4	2
096	41.2	28.5	096	2	4
120	42.2	29.4	120	2	4

Table 4: GOFS 3.1 and ACNFS persistence ice edge error (km) as a function of forecast length compared against the observed NIC ice edge for the validation period February – June 2015 for the seven analysis regions of the Arctic Ocean. The Laptev/East Siberian Sea section is blank because of insufficient data during this time period. The lower right quadrant also summarizes the total number of "wins". Lower errors/higher "wins" are highlighted with green.

3.3.2 Ice edge location – Southern Hemisphere

Ice edge location error for the Southern Hemisphere in GOFS 3.1 is shown for selected regions in Figure 21 and statistics for all regions in Table 5. Southern Hemisphere errors are two to three times higher than the Northern Hemisphere (Table 1). Animations of ice concentration demonstrated that GOFS 3.1 typically under predicts ice extent compared to the NIC ice edge analysis.

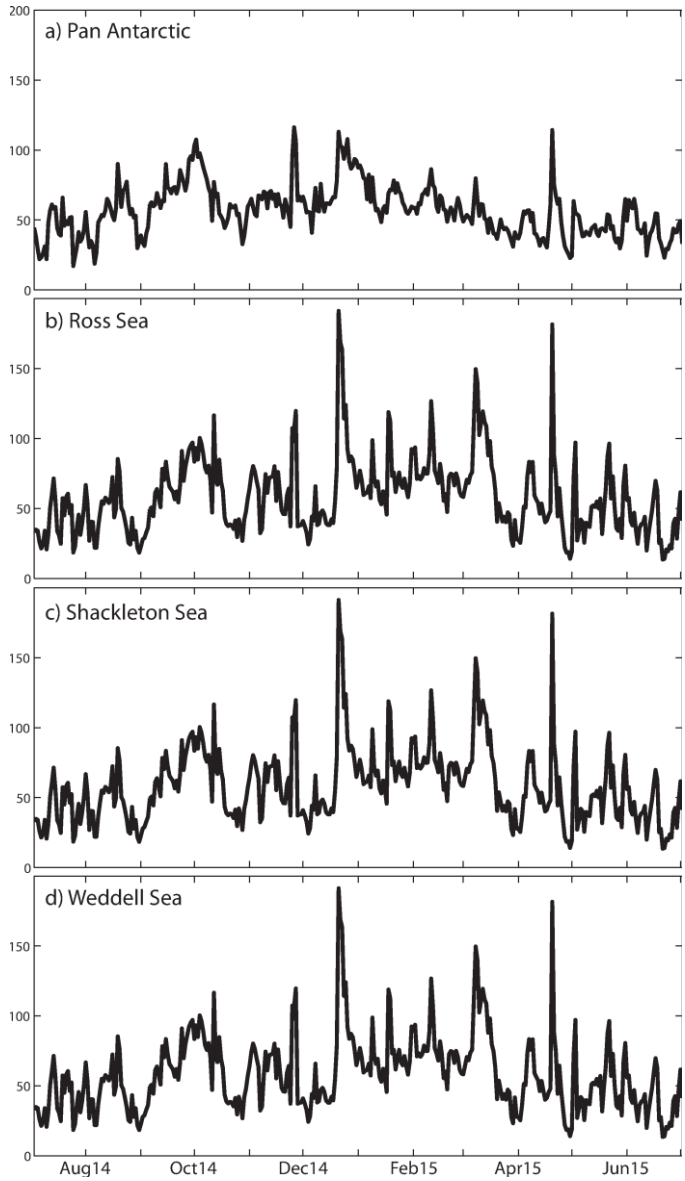

Figure 21: As in Figure 18 except for four Antarctic Ocean regions: a) the Pan Antarctic Ocean, b) the Ross Sea, c) the Shackleton Sea and d) the Weddell Sea.

Region	GOFS 3.1
Pan Antarctic	56.6
Amery Sea	59.9
Amundsen Sea	61.7
Bellingshausen Sea	48.8
Ross Sea	58.4
Shackleton Sea	53.1
Weddell Sea	62.1

Table 5: Mean 24-hour forecast ice edge location error (km) of GOFS 3.1 compared against the NIC observed ice edge for the validation period July 2014 – June 2015 for the seven analysis regions of the Antarctic Ocean.

3.3.3 Ice thickness

The NASA Operation IceBridge mission (Kurtz et al., 2013) collects airborne remote sensing measurements to bridge the gap between NASA's Ice, Cloud and land Elevation Satellite (ICESat) mission and the upcoming ICESat-2 mission. Although limited in spatial extent, short in length and only containing a small number of flights (Figure 22), IceBridge data is used in validating GOFS 3.1 and ACNFS ice thickness. It is serendipitous that some flights cover the region where the ice thickness differences are largest between the two forecast systems (Figure 17f).

The IceBridge flight data have very high temporal resolution. The data were filtered via a running mean that smoothed the time series sufficiently. Several time filters were examined and a six-minute filter produced accurate results. This equates to approximately 45 km of flight path. GOFS 3.1 and ACNFS output were interpolated to these flight paths for the following error analysis.

Figure 23 illustrates the ice thickness comparison for two sample flights while Table 6 presents various statistics for all flights. The IceBridge data exhibit significantly more spatial variability than either forecast system. In general, GOFS 3.1 ice thickness is closer to the observed flight data in the regions of the Beaufort Sea and Canadian Archipelago, whereas ACNFS has lower error for the region north of Greenland. This is clearly evident in Figure 23a for the flight that starts north of Greenland and ends north of Alaska. At the beginning of this flight ACNFS thickness compares favorably against the IceBridge data but by the end of the flight GOFS 3.1 thickness is in better agreement. Table 6 also

indicates that ice thickness is generally higher than observed in the Beaufort Sea and Canadian Archipelago whereas it is too thin north of Greenland. The better performance of GOFS 3.1 in the Beaufort Sea and Canadian Archipelago is important because these are regions of significant commercial activity (oil exploration, fishing, etc.).

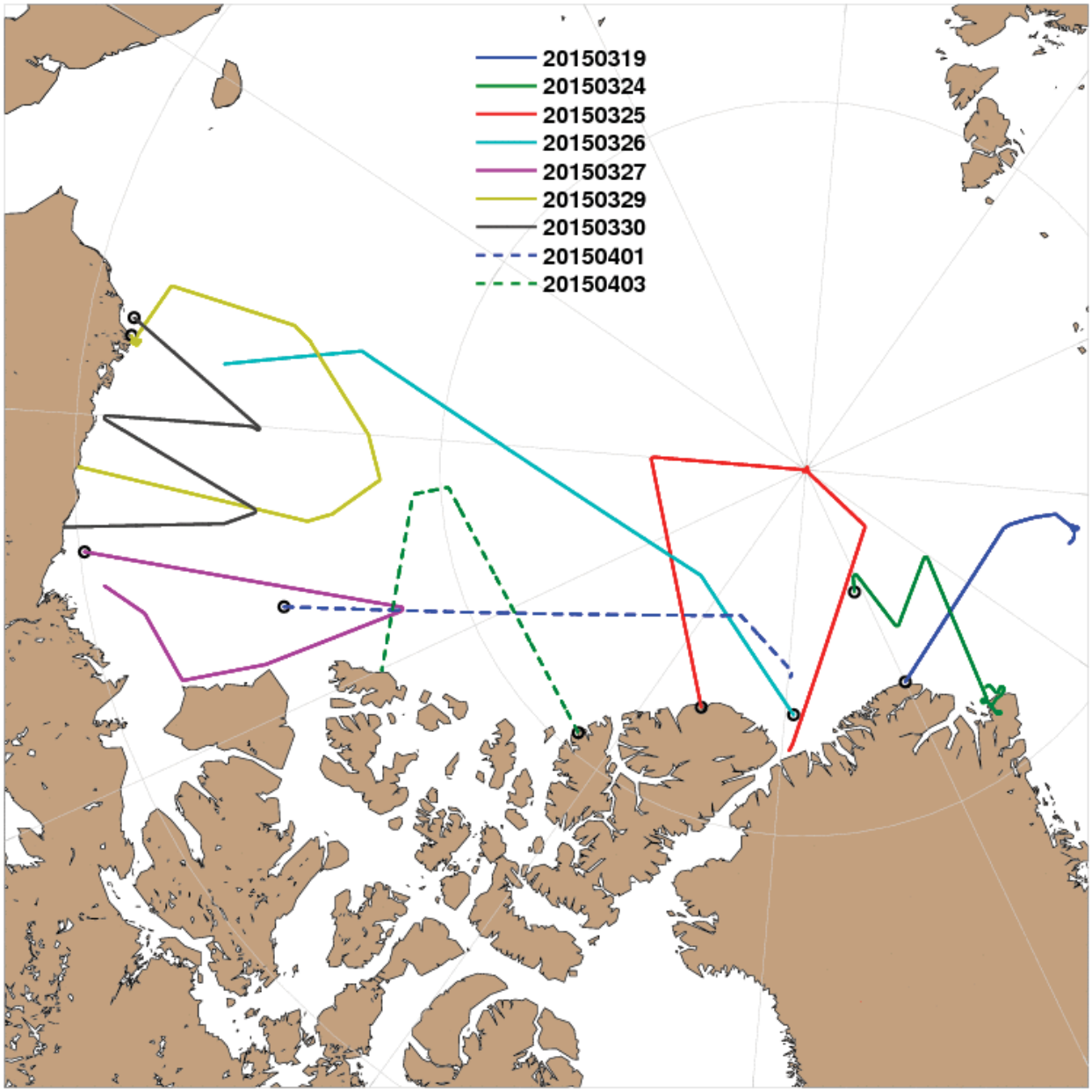

Figure 22: IceBridge flight paths for the spring of 2015. Flight data for the 26 March 2015 (cyan) and 30 March 2015 (black) paths are depicted in Figure 23. The black circle indicates the starting point of each flight.

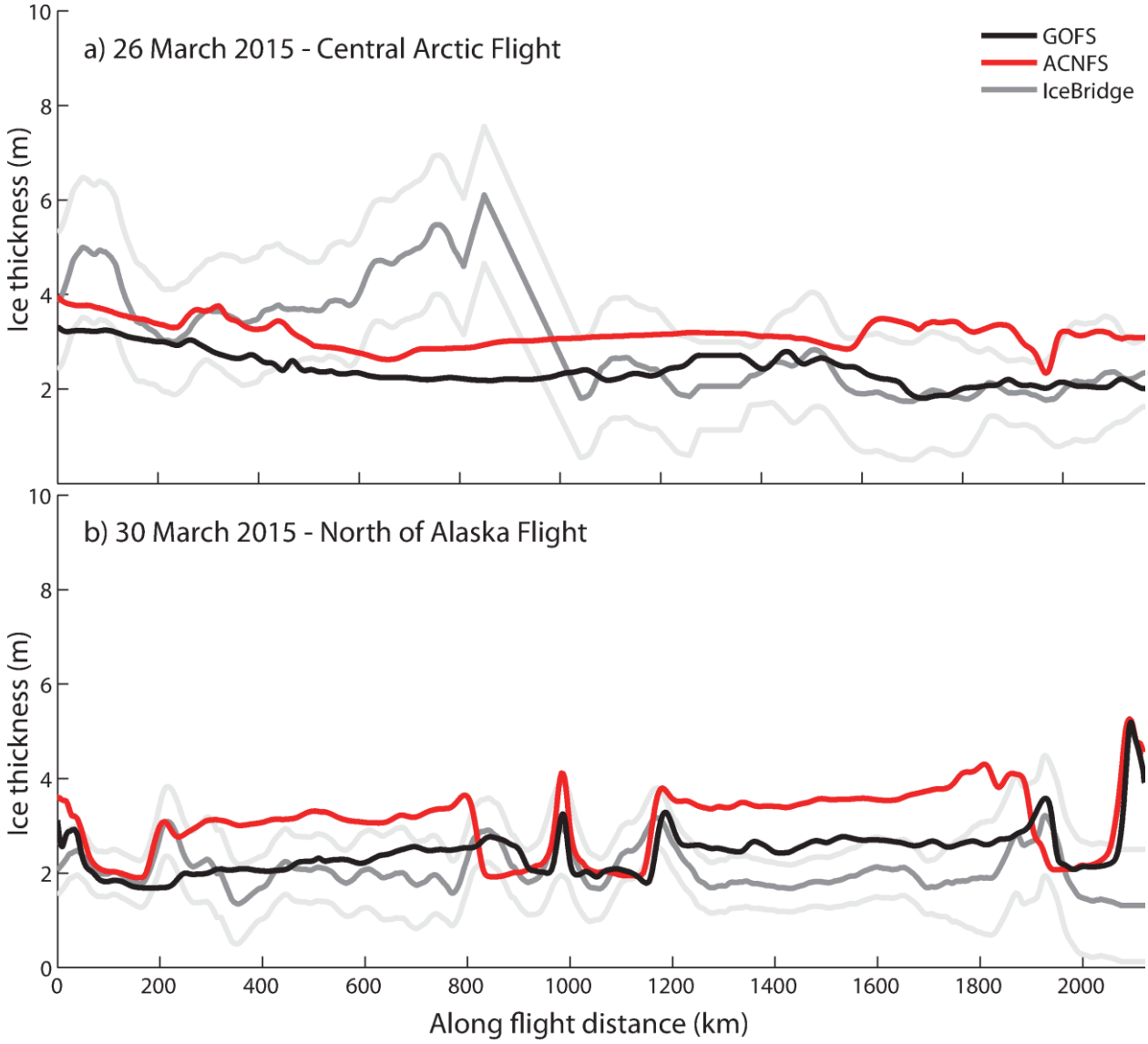

Figure 23: "Quick-look" IceBridge (dark gray line with uncertainty estimate as light gray lines) ice thickness (m) versus GOFS 3.1 (black) and ACNFS (red) for a) the 26 March 2015 flight that begins at northern Greenland and ends north of Alaska, and b) the 30 March 2015 flight that begins near Barrow, Alaska and zig-zags into the Beaufort Sea and back toward the Alaskan coast.

Flight	Mean Error		Absolute Mean Error		RMS Difference	
	GOFS 3.1	ACNFS	GOFS 3.1	ACNFS	GOFS 3.1	ACNFS
20150319	-0.45	-0.33	0.64	0.61	1.48	1.43
20150324	-0.39	-0.26	0.98	1.01	1.12	1.24
20150325	-1.16	-0.82	1.18	0.87	1.68	1.46
20150326	-0.59	0.17	0.77	0.97	1.15	1.16
20150327	0.42	0.64	0.88	1.14	1.24	1.34
20150329	0.37	1.11	0.59	1.19	0.65	1.40
20150330	0.31	0.97	0.51	1.10	0.60	1.26
20150401	0.27	0.72	0.54	0.82	0.63	0.97
20150403	-0.81	-0.14	0.99	0.70	1.20	0.84
"wins"	4	5	6	3	6	3

Table 6: Statistics comparing "quick-look" IceBridge ice thickness data against GOFS 3.1 and ACNFS sampled along the same flight paths and on the same dates. Units are meters. Lower errors/total "wins" are highlighted with green.

3.3.4 Ice drift

Ice drift from the two forecast systems is compared against drifting buoys from the International Arctic Buoy Program (IABP; http://iabp.apl.washington.edu/index.html). A total of 254 drifting buoys (Figure 24) exist over the July 2014 – June 2015 hindcast period, and the entire period is used in this ice drift analysis. From the daily latitude/longitude pairs of the ice-bound drifters, observed ice drift components are derived using the haversine formula to determine the x- and y-direction distances travelled each day. Results are then converted from km/day to m/s. Model ice velocity components are interpolated via cubic splines to the observed positions and GOFS 3.1 or ACNFS drift is derived from the mean ice velocity at tau 000 and tau 024.

Similar to the ocean drift error analysis, we computed ice speed mean error and RMSE, and ice direction vector correlation. Figure 25 shows the monthly variation of these quantities over the Pan Arctic region for the 24-hour forecasts, while Table 7 provides the regional statistics. Relative to ACNFS, GOFS

3.1 has lower speed error for the majority of the regions and higher directional correlations for all the regions.

Northern Hemisphere ice drift is also categorized by ice concentration to evaluate system performance within both the pack ice and the marginal ice zone (Tables 8 and 9). However, the analysis is only performed for the Pan-Arctic domain since the number of ice drifters gets very small or non-existent in the smaller subregions. Table 8 shows observed ice drift speed along with the corresponding GOFS 3.1 or ACNFS speed and the bias. Note that except for the very lowest ice concentration category, both GOFS 3.1 and ACNFS have ice drift that is too fast. The exception is the very lowest category in which the observations are significantly faster than either system. Table 9 compares similar statistics as Table 7. In general, GOFS 3.1 has lower error at the higher ice concentration categories, whereas ACNFS is better at the lower concentrations.

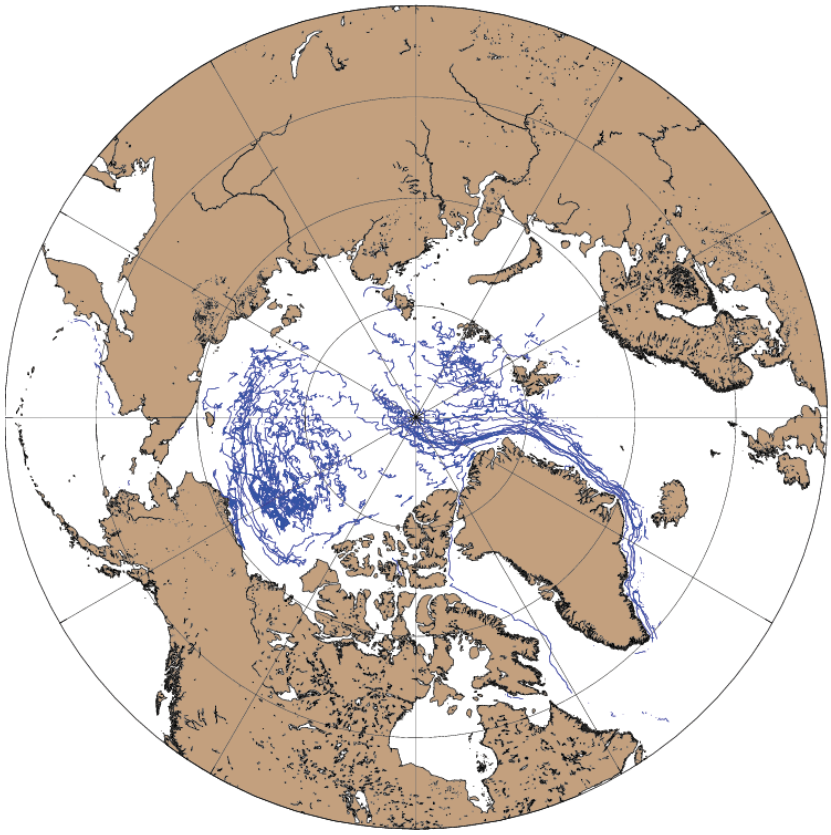

Figure 24: A "spaghetti" plot of International Arctic Buoy Program ice-bound drifting buoy tracks over the July 2014 – June 2015 hindcast time frame. A total of 254 buoys are used in the ice drift analysis.

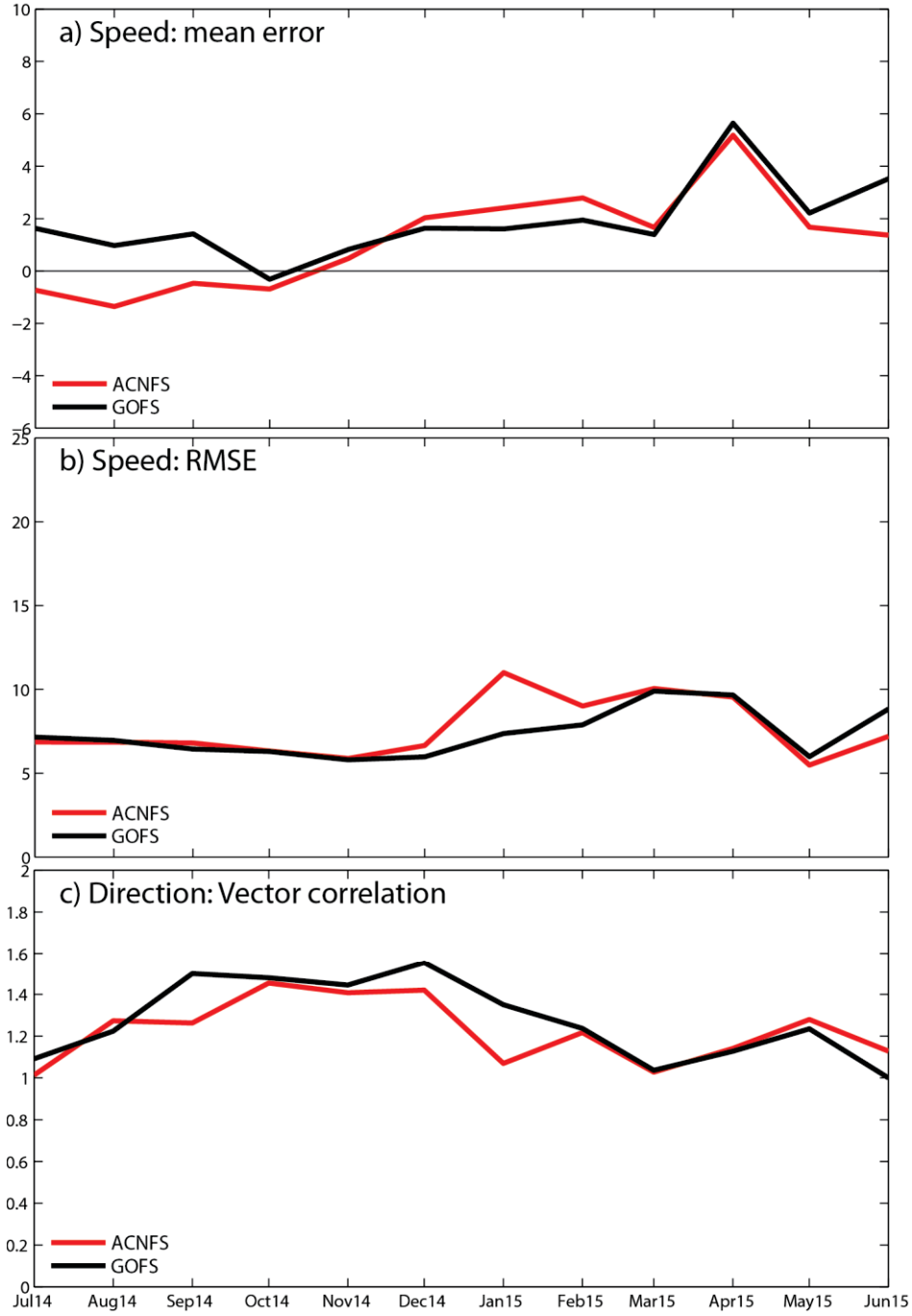

Figure 25: Monthly a) speed mean error (cm/s), b) RMSE (cm/s), and c) direction vector correlation of unassimilated IABP ice-bound drifting buoys for the Pan Arctic region compared against GOFS 3.1 (black) and ACNFS (red) at all ice concentrations. The vector correlation is based on Crosby et al. (1993) with 0 indicating no correlation and 2 indicating perfect correlation. This analysis spans the hindcast period July 2014 – June 2015.

Region	Absolute ME		RMSE		Vector Correlation	
	GOFS 3.1	ACNFS	GOFS 3.1	ACNFS	GOFS 3.1	ACNFS
Pan Arctic	4.8	4.8	7.3	7.7	1.29	1.21
GIN Seas	7.0	7.6	11.1	13.1	1.03	0.92
Barents/Kara Seas	3.6	3.8	5.3	5.6	1.58	1.35
Laptev/E. Siberian Seas	3.7	3.9	5.3	5.5	1.43	1.36
Bering/Chukchi/Beaufort	4.7	4.5	7.0	6.8	1.26	1.23
Canadian Archipelago	6.1	5.5	8.9	9.0	0.89	0.86
Central Arctic	3.8	4.0	5.2	5.5	1.57	1.47
"wins"	4	2	6	1	7	0

Table 7: Ice drift statistics of speed absolute mean error and RMSE (both in cm/s), and direction vector correlation for IABP ice-bound drifting buoys versus GOFS 3.1 and ACNFS for the 24-hour forecast over the July 2014 – June 2015 hindcast period at all ice concentrations. The vector correlation is based on Crosby et al. (1993) with 0 indicating no correlation and 2 indicating perfect correlation. Lower errors/total "wins" are highlighted with green.

Ice concentration range (number of buoys)	Buoy speed	GOFS 3.1		ACNFS	
		Speed	Bias	Speed	Bias
0.80 < icecon ≤ 1.00 (204)	9.7	11.7	2.0	11.0	1.3
0.60 < icecon ≤ 0.80 (117)	13.1	13.9	0.8	10.9	-2.2
0.30 < icecon ≤ 0.60 (82)	15.0	16.5	1.5	17.3	2.3
0.10 < icecon ≤ 0.30 (20)	23.1	27.7	4.6	26.5	3.4
0.00 < icecon ≤ 0.10 (38)	28.9	17.6	-11.3	22.5	-6.4

Table 8: Ice drift speed mean error (cm/s) categorized by ice concentration for IABP ice-bound drifting buoys over the Pan-Arctic region versus GOFS 3.1 and ACNFS for the 24-hour forecast over the July 2014 – June 2015 hindcast period.

Ice concentration range (number of buoys)	Absolute ME		RMSE		Vector Correlation		"Wins"
	GOFS 3.1	ACNFS	GOFS 3.1	ACNFS	GOFS 3.1	ACNFS	GOFS/ ACNFS
0.80 < icecon ≤ 1.00 (204)	4.2	4.2	6.1	6.3	1.47	1.38	2 / 0
0.60 < icecon ≤ 0.80 (117)	5.4	5.6	7.4	7.6	1.20	1.11	3 / 0
0.30 < icecon ≤ 0.60 (82)	8.3	7.7	11.7	12.3	0.90	0.86	2 / 1
0.10 < icecon ≤ 0.30 (20)	16.4	14.4	20.3	19.7	0.64	0.64	0 / 2
0.00 < icecon ≤ 0.10 (38)	18.3	17.6	24.6	26.0	0.52	0.64	1 / 2

Table 9: Ice drift statistics of speed absolute mean error and RMSE (both in cm/s), and direction vector correlation categorized by ice concentration for IABP ice-bound drifting buoys over the Pan-Arctic domain versus GOFS 3.1 and ACNFS for the 24-hour forecast over the July 2014 – June 2015 hindcast period. The vector correlation is based on Crosby et al. (1993) with 0 indicating no correlation and 2 indicating perfect correlation. "Wins" are highlighted with green.

4.0 SUMMARY, SCORE CARDS AND RECOMMENDATIONS

Since the transition of GOFS 3.0 to NAVOCEANO, research and development at NRL have led to improvements in the global system with respect to the ocean/ice models and the data assimilation. The new GOFS 3.1 represents the existing state-of-the-art Navy prediction system. Major differences between it and GOFS 3.0/ACNFS follow. With respect to HYCOM, GOFS 3.1 incorporates improved coastlines and bathymetry, uses a 17-term equation of state (vs. 9-term), adds nine near surface layers to better resolve the mixed layer, uses an improved ocean turbidity scheme and an improved surface salinity relaxation scheme. The CICE sub-model used within GOFS 3.1 and ACNFS is the same, but the ice assimilation scheme is different. ACNFS utilizes the satellite observations near the edge only and largely leaves the model field untouched away from the edge, whereas GOFS 3.1 also trusts the satellite observations near the edge, but away from it uses the NCODA ice analysis ±10%. Ocean data assimilation changes include the use of receipt time for observations and the FGAT technique. This has reduced the length of the daily hindcast to 1 day (vs. 4 days) and significantly reduced the total daily runtime of the system. Lastly, GOFS 3.1 employs newly improved ISOP for the projection of surface

information downward into the water column (vs. MODAS) that has been shown to provide more accurate vertical ocean structure.

Error analyses similar to past GOFS VTRs are performed. For the ocean these include temperature/salinity vs. depth, acoustical proxies, surface layer trapping of acoustic frequencies, and upper ocean velocities. For ice validation, error analyses of ice edge location, ice thickness and ice drift are performed.

To quantify the performance of GOFS 3.1 and GOFS 3.0, an ocean score card developed by Zamudio et al. (2015) is employed. To track the system's performance of each different ocean metric, errors are translated into scores via the following. Let *Error_model(i)* be the model error for the individual ocean metrics *(i)*, and model is either GOFS 3.1 or GOFS 3.0. And, let

Maximum_error(i) = largest of *Error_model(i),*

Score_model(i) = 1 – (*Error_model(i) / Maximum_error(i))*

Total_score_model = mean (*Score_model(i))*

be the maximum error over the hindcast period for each individual region, the scores of the models per each individual metric, and the total score per each model, respectively. A model score of 1 indicates perfect skill relative to the observations, whereas a model score of 0 indicates no skill. The difference between the GOFS 3.1 and 3.0 scores shows the relative performance to each other and allows a determination as to which system is more skillful. Mean absolute error (MAE) and RMSE are the metrics that contribute to the score for T vs. depth, S vs. depth, MLD, SLD, and BLG. The drifting buoy score uses these same metrics and additionally the inverse of vector correlation. Lastly, the false / true ratio is the metric to compute the score for the cutoff frequency metric.

Figure 26 shows the monthly scores for the global analysis region along with the number of observations. Panel a) shows *Total_score_model* while the other panels show *Score_model(i).* For all metrics with the exception of S vs. depth, GOFS 3.1 shows higher skill than GOFS 3.0. This is especially true for the drifting buoy analysis. Figure 27 shows *Total_score_model* for all the analysis regions. Again, GOFS 3.1 shows higher skill for all regions with the exception of a few months in the Gulf Stream and

GoM/IAS analysis regions. Note that in Figure 27, because each region is normalized by the maximum error of that specific region, it is not possible to compare the relative skill of one analysis region with another. That can be done by normalizing by the maximum error over **all** regions (not shown), but that can give the false impression of high skill if one region happens to be particularly unskillful.

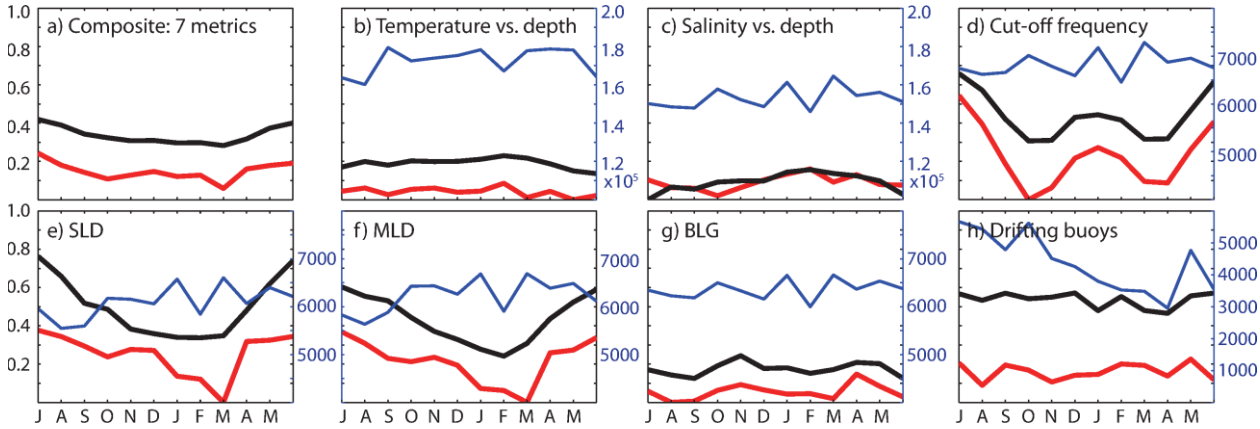

Figure 26: Monthly scores of the global (50°S-50°N) analysis region for a) *Total_score_model* (the composite of all seven metrics), and *Score_model(i)* with b) T vs. depth, c) S vs. depth, d) cut-off frequency, e) SLD, f) MLD, g) BLG and h) drifting buoys. A score of 1 indicates perfect skill, whereas a score of 0 indicates no skill. The black line is GOFS 3.1, the red line is GOFS 3.0 and the blue line (and right side y-axis) is the number of observations used in the error analysis in panels b), c) and h), while it is the number of profiles in panels d), e), f) and g).

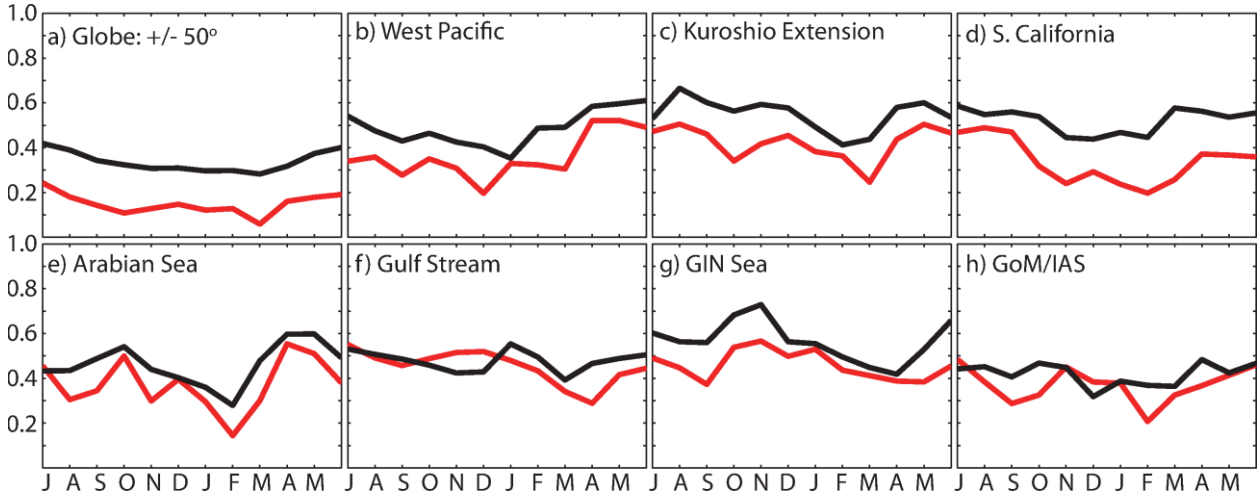

Figure 27: Monthly *Total_score_model* for the analysis regions: a) global (50°S-50°N), b) West Pacific, c) Kuroshio Extension, d) Southern California, e) Arabian Sea, f) Gulf Stream, g) GIN Sea and h) GoM/IAS. A score of 1 indicates perfect skill, while a score of 0 indicates no skill. The black line is GOFS 3.1 and the red line is GOFS 3.0. Because each region is normalized by the maximum error of that specific region, these do not compare the relative skill of one analysis region with another.

An ice score card similar to the ocean score card has not yet been refined, so instead we count the number of "wins" for the various metrics that are summarized in Table 10. With regard to ice edge location error, GOFS 3.1 is the superior system from the nowcast time to forecast lengths out to tau 048, while ACNFS provides better forecasts beyond that. With regard to ice thickness and ice drift, GOFS 3.1 outperforms ACNFS. GOFS 3.1 ice edge forecast performance was also compared against "forecasts" of GOFS 3.1 and ACNFS persisted ice edge (Table 11). Here GOFS 3.1 has lower ice edge error from the nowcast time out to tau 072, but beyond that persistence is the better forecast. Thus, in the net, GOFS 3.1 outperforms ACNFS at the nowcast and to forecasts out to tau 048. Beyond this time window, its overall skill diminishes, although still remains high for the GIN Seas out through tau 120.

NH ice edge location error "wins"		
Tau	GOFS 3.1	ACNFS
000	6	0
024	5	1
048	4	2
072	2	4
096	1	5
120	1	5
NH ice thickness error "wins"		
Metric	GOFS 3.1	ACNFS
ME	4	5
MAE	6	3
RMSE	6	3
NH ice edge drift error "wins"		
Metric	GOFS 3.1	ACNFS
MAE	4	2
RMSE	6	1
VC	7	0

Table 10: "Wins" of the various ice metrics for GOFS 3.1 and ACNFS defined in section 3.3. ME = mean error, MAE = mean absolute error, RMSE = root mean square error and VC = vector correlation.

NH ice edge location error "wins"		
Tau	GOFS 3.1	GOFS 3.1 persistence
000	–	–
024	7	0
048	7	0
072	4	3
096	2	5
120	2	5
NH ice edge location error "wins"		
Tau	GOFS 3.1	ACNFS persistence
000	6	0
024	6	0
048	5	0
072	4	2
096	2	4
120	2	4

Table 11: "Wins" of the NH ice edge location error for GOFS 3.1 compared against GOFS 3.1 persistence and ACNFS persistence "forecasts".

Overall, this report has determined that GOFS 3.1 is performing better than GOFS 3.0 with respect to the ocean and GOFS 3.1 is performing better than ACNFS with respect to sea ice. It is recommended that GOFS 3.1 move to OPTEST at NAVOCEANO (ocean) and NIC (ice). Pending success at the OPTEST level, it can then replace both existing operational systems, GOFS 3.0 and ACNFS.

5.0 ACKNOWLEDGMENTS

This work was funded as part of the NRL 6.4 Large Scale Prediction and 6.4 Ocean Data Assimilation projects, managed by the Office of Naval Research under program element 0603207N. The numerical hindcasts and forecasts were performed on the Navy DSRC Cray XC40 computer (Gordon) at Stennis Space Center, Mississippi using grants of computer time from the DoD High Performance Computing Modernization Program. Thanks are extended to Pete Spence for the acoustical trapping software. The Validation Test Panel consisted of: Mark Cobb (NAVOCEANO), Sean Helfrich (National

Ice Center), Steve Morey (Florida State University), Pat Hogan, Pam Posey and Joe Metzger (NRL). This is NRL contribution NRL/MR/7320--17-9722, which is approved for public release and distribution is unlimited.

6.0 REFERENCES

Bloom, S.C., L.L. Takacs, A.M. da Silva and D. Ledvina. 1996. Data assimilation using incremental analysis updates. *Mon. Wea. Rev*. **124**:1256–1271.

Carnes, M.R., 2009. Description and Evaluation of GDEM-V 3.0. *NRL Memo. Report*, NRL/MR/7330--09-9165. (Available at http://www7320.nrlssc.navy.mil/pubs.php.)

Carnes, M.R., R.W. Helber, C.N. Barron and J.M. Dastugue, 2010. Validation Test Report for GDEM4. *NRL Memo. Report*, NRL/MR/7330--10-9271.

Chen, C.T. and F.J. Millero, 1977: Speed of sound in seawater at high pressures. *J. Acoust. Soc. Am*., **62**, 1129-1135.

Crosby, D.S., L.C. Breaker and W.H. Gemmill, 1993: A proposed definition for vector correlation in geophysics: Theory and application. *J. Atm. Ocean. Tech*., **10**, 355-367.

Cummings, J.A. and O.M. Smedstad, 2014. Ocean data impact in global HYCOM. *J. Atmos. Ocean. Technol.,* **31**, 1771-1791, DOI: 10.1175/JTECH-D-14-00011.1.

Fox, D.N., W.J. Teague, C.N. Barron, M.R. Carnes and C.M. Lee, 2002: The Modular Ocean Data Analysis System (MODAS). *J. Atmos. Ocean Technol*., **19**, 240-252.

Hebert, D.A., R.A. Allard, E.J. Metzger, P.G. Posey, R.H. Preller, A.J. Wallcraft, M.W. Phelps and O.M Smedstad, 2015: Short-term sea ice forecasting: An assessment of ice concentration and ice drift forecasts using the U.S. Navy's Arctic Cap Nowcast/Forecast System (ACNFS). *J. Geophys. Res.- Oceans*, doi:10.1002/2015JC011283.

Helber, R.W., C.N. Barron, M. Gunduz, P.L. Spence and R.A. Zingarelli, 2010: Acoustic metric assessment for UUV observation system simulation experiments. *U.S. Navy Journal of Underwater Acoustics,* **60**(1), 101-123.

Helber, R.W., T.L. Townsend, C.N. Barron, J.M. Dastugue and M.R. Carnes, 2013: Validation Test Report for the Improved Synthetic Ocean Profile (ISOP) System, Part I: Synthetic Profile Methods and Algorithm. *NRL Memo. Report*, NRL/MR/7320—13-9364.

Hill, C., C. DeLuca, V. Balaji, M. Suarez, and A. da Silva, 2004: The Architecture of the Earth System Modeling Framework. *Computing in Science and Engineering*, **6**:18-28.

Hogan, T.F., M. Liu, J.A. Ridout, M.S. Peng, T.R. Whitcomb, B.C. Ruston, C.A. Reynolds, S.D. Eckermann, J.R. Moskaitis, N.L. Baker, J.P McCormack, K.C. Viner, J.G. McLay, M.K. Flatau, L. Xu, C. Chen, and S.W. Chang, 2014: The Navy Global Environmental Model. *Oceanography*, Vol. 27, No. 3, 116-125, http://dx.doi.org/10.5670/oceanog.2014.66.

Hunke, E.C. and W. Lipscomb, 2008. CICE: The Los Alamos sea ice model, documentation and software user's manual, version 4.0. *Tech. Rep. LA-CC-06-012 Los Alamos National Laboratory*, Los Alamos, NM. (http://climate.lanl.gov/models/cice/index.htm).

Jackett, D.R., T.J. McDougall, R. Feistel, D.G. Wright and S.M. Griffies, 2006: Algorithms for Density, Potential Temperature, Conservative Temperature, and the Freezing Temperature of Seawater. *J. Atmos. Oceanic Technol*., **23**, 1709–1728, doi: http://dx.doi.org/10.1175/JTECH1946.1.

Joliff, J.K., C.N. Barron, A.J. Wallcraft, E.J. Metzger, J.F. Shriver and J. Wesson, 2014: An improved solar radiation scheme for global HYCOM with updated global ocean color climatology. *NRL Memo. Rpt.,* NRL/MR/7330-2014-9528. (Submitted.)

Kurtz, N.T., S.L. Farrell, M. Studinger, N. Galin, J.P. Harbeck, R. Lindsay, V.D. Onana, B. Panzer and J.G. Sonntag, 2013: Sea ice thickness, freeboard, and snow depth products from Operation IceBridge airborne data. *The Cryosphere*, **7**, 1035-1056, doi:10.5194/tc-7-1035-2013.

Lumpkin, R. and G.C. Johnson, 2013: Global Ocean Surface Velocities from Drifters: Mean, Variance, ENSO Response, and Seasonal Cycle. *J. Geophys. Res.-Oceans*, 118, 2992-3006, doi:10.1002/jgrc.20210 .

Metzger, E.J., P.G. Posey, P.G. Thoppil, T.L. Townsend, A.J. Wallcraft, O.M. Smedstad, D.S. Franklin, L. Zamudio and M.W. Phelps, 2015: Validation Test Report for the Global Ocean Forecast System V3.1 - 1/12º HYCOM/NCODA/CICE/ISOP. *NRL Memo. Report*, NRL/MR/7320--15-9579. (Distribution limited to US government agencies and their contractors.)

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,* Vol. 27. No. 3, 32-43, http://dx.doi.org/10.5670/oceanog.2014.66.

Metzger, E.J., O.M. Smedstad, P.G. Thoppil, H.E. Hurlburt, D.S. Franklin, G. Peggion, J.F. Shriver T.L. Townsend and A.J. Wallcraft, 2010: Validation Test Report for the Global Ocean Forecast System V3.0 - 1/12º HYCOM/NCODA: Phase II. *NRL Memo. Report*, NRL/MR/7320--10-9236. (Available at http://www7320.nrlssc.navy.mil/pubs.php.)

Metzger, E.J., O.M. Smedstad, P.G. Thoppil, H.E. Hurlburt, A.J. Wallcraft, D.S. Franklin, J.F. Shriver and L.F. Smedstad, 2008: Validation Test Report for the Global Ocean Prediction System V3.0 - 1/12º HYCOM/NCODA: Phase I. *NRL Memo. Report*, NRL/MR/7320--08-9148. (Available at http://www7320.nrlssc.navy.mil/pubs.php.)

Metzger, E.J., A.J. Wallcraft, P.G. Posey, O.M. Smedstad and D.S. Franklin, 2013: The switchover from NOGAPS to NAVGEM 1.1 atmospheric forcing in GOFS and ACNFS. *NRL Memo. Report*, NRL/MR/7320--13-9486. (Available at http://www7320.nrlssc.navy.mil/pubs.php.)

Millero, F.J. and X. Li, 1994: Comments on "On equations for the speed of sound in seawater". *J. Acoust. Soc. Am.*, **95**(5), 2757-2759.

Posey, P.G., E.J. Metzger, A.J. Wallcraft, D.A. Hebert, R.A. Allard, O.M Smedstad, M.W. Phelps, F. Fetterer, J.S. Stewart, W.N. Meier and S.R. Helfrich, 2015: Improving Arctic sea ice edge forecasts by assimilating high horizontal resolution sea ice concentration data into the US Navy's ice forecast systems. *The Cryosphere*, doi: 10.4194/tc-9-1735-2015.

Posey, P.G., E.J. Metzger, A.J. Wallcraft, R.H. Preller, O.M. Smedstad and M.W. Phelps, 2010: Validation of the 1/12º Arctic Cap Nowcast/Forecast System (ACNFS). *NRL Memo. Report*, NRL/MR/7320--10-9287. (Available at http://www7320.nrlssc.navy.mil/pubs.php.)

RP33: Fleet Oceanographic and Acoustic Reference Manual, 1992: *Naval Oceanographic Office Reference Publication 33*, Naval Oceanographic Office, Stennis Space Center, MS.

Smith, S., J.A. Cummings, C. Rowley, P. Chu, J. Shriver, R. Helber, P. Spence, S. Carroll, O.M. Smedstad and B. Lunde, 2011: Validation Test Report for the Navy Coupled Ocean Data Assimilation 3D Variational Analysis (NCODA-VAR) System, Version 3.43. *NRL Memo. Report*, NRL/MR/7320--11-9363.

Sraj, I., and Coauthors, 2013: Bayesian Inference of Drag Parameters Using AXBT Data from Typhoon Fanapi. *Mon. Wea. Rev*., 141, 2347–2367, doi: http://dx.doi.org/10.1175/MWR-D-12-00228.1.

Steele, M., R. Morley and W. Ermold, 2001: PHC: A global ocean hydrography with a high quality Arctic Ocean. *J. Climate*, **14**, 2079-2087.

Townsend, T.L., C.N. Barron and R.W. Helber, 2015: Ocean prediction with Improved Synthetic Ocean Profiles (ISOP). *2015 NRL Review,* NRL/PU/7320--2015-2812, pp. 105-112.

Wilson, W.D., 1960: Equation for the speed of sound in sea water. *J. Acoust. Soc. Amer*., **32,** 1357.

Zamudio, L., P. Spence and G.A. Jacobs, 2015: Verification and validation report for Navy operational regional forecast models. Part II: Evaluation. *NRL internal report,* 30 April 2015.

7.0 TABLE OF ACRONYMS

3DVAR	3-Dimensional VARiational analysis
ACNFS	Arctic Cap Nowcast/Forecast System
AMSR2	Advanced Microwave Scanning Radiometer 2
BLG	Below Layer Gradient
CICE	Community Ice CodE
COF	CutOff frequency
DoD	Department of Defense
DSRC	DoD Supercomputing Resource Center
EKE	Eddy Kinetic Energy
FGAT	First Guess at Appropriate Time
FNMOC	Fleet Numerical Meteorology and Oceanography Center
GDEM	Generalized Digital Environment Model
GOFS	Global Ocean Forecast System
HYCOM	HYbrid Coordinate Ocean Model
IABP	International Arctic Buoy Program
ISOP	Improved Synthetic Ocean Profiles
MdBE	Median Bias Error
ME	Mean Error
MLD	Mixed Layer Depth
MODAS	Modular Ocean Data Analysis System
MVOI	Multi-Variate Optimum Interpolation
NAVGEM	NAVy Global Environmental Model
NAVOCEANO	Naval Oceanographic Office
NIC	National Ice Center
NCODA	Navy Coupled Ocean Data Assimilation
NOGAPS	Navy Operational Global Atmospheric Prediction System
NRL	Naval Research Laboratory
OPTEST	Operational Test
RMSE	Root Mean Square Error
S	Salinity
SAR	Synthetic Aperture Radar
SLD	Sonic Layer Depth
SSH	Sea Surface Height
SSMIS	Special Sensor Microwave Imager/Sounder
SSS	Sea Surface Salinity
SST	Sea Surface Temperature
T	Temperature
TEOS	Thermodynamic Equation of Seawater
VTP	Validation Test Panel
VTR	Validation Test Report