DISTRIBUTION STATEMENT A. Approved for public release; distribution is unlimited.

ESPC WU-3 Coupled Global Ensemble Design

Justin McLay (MRY) / Clark Rowley (SSC)
Naval Research Laboratory
Monterey, CA 93943-5502
phone: (831) 656-4837 fax: (831) 656-4769
e-mail: justin.mclay@nrlmry.navy.mil
clark.rowley@nrlssc.navy.mil

Award Number: N0001415WX00994
Award Number: N0001415WX00995

LONG-TERM GOALS

The long-term goal of the project is to provide extended-range (weekly to monthly timescale) probabilistic environmental information to meet Navy and DoD operations and planning needs across the globe from undersea to upper atmosphere and from the Tropics to the Poles.

OBJECTIVES

The objective is the development of a fully coupled atmosphere/ocean/ice ensemble prediction system (EPS) that will provide daily probabilistic predictions out to 10 days and weekly probabilistic predictions out to 30 days. Initial operational capability is targeted for 2018.

APPROACH

It is recognized that for extended-range predictions that are beyond the predictability limit of many processes, the most important information that can be provided is the probability distribution function (PDF) of environmental conditions. To properly estimate this PDF, it will be necessary to combine the existing independent ensemble prediction capabilities for the atmosphere, ocean, and waves as well as implement new ensemble capabilities for ice and land systems. The ideal, long-term solution is to obtain the PDF as an extension of weakly or strongly coupled ensemble data assimilation techniques, such as those being developed under the ESPC WU-10 Coupled Data Assimilation subproject. However, there are a variety of other coupled ensemble generation schemes that present themselves as expedient, interim solutions. These include the lagged ensemble method, the ensemble transform (ET), conditioned random perturbations, and the "forecast anomaly" method of Oke et al. (2005). The approach favored here is to develop and test some of these interim solutions as backstops to the longer-term effort of generating ensembles via coupled ensemble data assimilation.

Key personnel: (NRL MRY): Justin McLay, Neil Barton, Carolyn Reynolds, Maria Flatau, (NRL SSC): Clark Rowley, Pat Hogan, Posey, Metzger, Mozheng Wei.

WORK COMPLETED FOR FY15

Milestone 1: Develop ensemble system with simple diurnal SST model

- Scorecard validations were completed for multiple forms of diurnal SST variation. The basic details of the validations are as follows:

  Technical details:

  - The diurnal models differed in the specifics of their latitudinal mask, cool skin layer, and warm-layer depth.
  - Multi-season, deterministic NAVGEM v1.3 integrations were performed, initialized every 12h for January 2014 and July 2014.
  - Integrations were made to T+336h.
  - The resolution of the integrations was T239L50, the same as the ensemble resolution.
  - The basic deterministic scorecards consisted of the RMSE and BIAS of 2m air temperature and 10m wind speed, validated against ECMWF analyses.

  Summary points:

  As with earlier tests of the diurnal SST variation, the latest forms of the diurnal model have substantial influence on tropical 2m air temperature and some lesser but still notable influence on 10m wind speed. Overall, the tropical 2m air temperature benefits from the inclusion of the diurnal model. This is particularly true in terms of reduction of tropical warm BIAS. The RMSE is also improved in certain cases. The 10m wind speed results are mixed, with benefit in some cases for RMSE but aggravated negative (slow) wind speed BIAS in other cases. Figures 1-2 show examples of the diurnal SST variation's impact on 2m air temperature BIAS. The latest diurnal models have fewer adverse effects on extratropical scores than earlier versions, thanks to the more restrictive latitudinal mask, and have similar favorable effects on tropical scores. This makes these latest versions the most promising versions tested to date.

  An evaluation of the effect of the diurnal SST variation on tropical cyclone (TC) track forecasts was also completed, based on deterministic integrations for the four month test period Aug-Nov 2014. The inclusion of diurnal SST variation produces very minor differences in TC track and intensity. On average, the differences over the 0-120h forecast interval are too small to be noteworthy. Specifically, the average difference in track error and intensity at T+120h is O(1 nm) and O(1 hPa), respectively.

  Since the diurnal SST variation yields a net positive deterministic scorecard and has neutral impact on TC scores, it is a viable candidate for operational transition.

  At the close of FY15, further testing of the diurnal model in the NAVGEM v1.3 ensemble is awaiting a reconfiguration of the ET ensemble generation scheme.*

  * Tests have found that the spatial distribution of variability of the NAVGEM v1.3 model is very different from that of v1.2. In particular, v1.3 supports much greater variability in the tropical upper atmosphere. In consultation with the NAVGEM modelers and NRL-DC upper-atmosphere researchers, the general consensus is that this added upper-atmosphere variability is desirable from a physical and dynamical standpoint. However, due to the nuances of the ET ensemble generation scheme, this added upper-atmosphere variability has been very disruptive to the quality of the forecast ensemble in the troposphere. To compensate, a significant reconfiguration of the ET for NAVGEM v1.3 is presently being undertaken.

Milestone 2: Implement persistent SST anomaly capability

- Scorecard validations were completed for the persistent anomaly capability. The experimental setup was similar to that used for the diurnal SST variation:

  Technical details:

  - Multi-season deterministic NAVGEM v1.3 integrations were carried out, initialized every 12h for January 2014 and July 2014.
  - Integrations were made to T+336h.
  - The resolution of the integrations was T239L50, the same as the ensemble resolution.
  - The time-varying climatological SST was obtained from the ECMWF ERA-40 reanalysis dataset.
  - The persistent-anomalies were spatially filtered, and the climatological SST was updated every 24h.
  - The scorecard was the same as that used for the diurnal SST variation (described above for Milestone 1).

  Summary points:

  As with the diurnal SST variation, indications are that the persistent-anomaly SST has substantial influence on 2m air temperature and some lesser but still notable influence on 10m wind speed. The persistent-anomaly SST has a particularly large impact on 2m air temperature bias (Figure 3). However, the scorecards reveal a major problem with the persistent-anomaly SST approach. Namely, its effects display a very clear flip-flop as a function of region and season. More specifically, it causes major benefit (detriment) to the tropical 2m airtmp bias and detriment (benefit) to the Extratropical 2m airtmp bias in summer (winter), based on using the ECMWF SST climatology and validating against ECMWF analyses. Figures 3 and 4 illustrate the seasonal flip-flop effect for the Tropics. It's not clear how to mitigate this problem. It's possible that it could be avoided if the NCEP SST climatology were to be used and the forecasts were to be validated against UKMET analyses or against buoys and ship obs. However, this remains to be seen.

  Owing to the flip-flop problem, work on the persistent SST anomaly capability has been tabled for the time being.

Milestone 3: Implement initial SST perturbation capability

- Scorecard validations were completed for multi-season ensemble runs with climatologically-derived SST variability in the uncoupled NAVGEM ensemble forecast system. The experimental details were as follows:

Technical details:

- SST initial perturbations were produced for each member of the ensemble for each forecast cycle.
- The perturbations were derived from historical anomalies with respect to climatology in the temporal neighborhood of the ensemble initialization date.
- The SST climatology was derived from 20-years of the ECMWF ERA-40 reanalysis dataset.
- The runs consisted of T239L50 20-member, 14d lead time NAVGEM v1.3 ensembles
- The SST variance had a pre-specified value of 0.50, implying an average SST perturbation magnitude of 0.7 degrees.

This climatologically-derived SST variability is a simple alternative to using SST variability produced by the ET analysis perturbation system, which is inadequate in several major respects. It also can serve as a useful baseline for other potential types of SST variability, such as SST variability from a lagged-ensemble method.

Summary points:

The ensembles with climatologically-derived SST variability display aggregate positive (favorable) scores relative to the control. The most consistent, major effect of the inclusion of the SST variability is to improve the ensemble variance. Figures 5 and 6 show two examples of the variance diagnostic for these ensemble runs. There are no major impacts of the SST variability on the RMSE or CRPS diagnostics, although there are a number of minor (mostly positive) impacts. Of the negative impacts, the most prominent is a degradation of 850 hPa air temperature bias. However, the negative impact on the 850 hPa air temperature bias is generally minor (i.e. less than 5% relative change in score).

These first results are promising and suggest that the initial SST perturbation capability will be a viable candidate for operational transition, pending the outcome of testing with the latest configuration of the NAVGEM v1.3 ensemble system.

At the close of FY15, further testing of the initial SST perturbations in the NAVGEM v1.3 ensemble is awaiting a reconfiguration of the ET ensemble generation scheme. See note (*) under Milestone 1 (above).

Milestone 4: Support global ocean ensemble cycling

- Implemented a perturbed-observation analysis capability in NCODA as an alternative ensemble generation technique for the global ocean/ice ensemble.

Technical details:

- The original NCODA perturbed observation method was performed by Sergei Frolov (NRLMRY).

- The NCODA perturbed obs analysis was modified to extend to all data types (including both real and synthetic profile observations) and for independent parallel analyses (the analyses can be run sequentially or in parallel with no communication required between members).
- Observation perturbations are based on the observation error estimates in NCODA.
- The updated code was initially exercised in an analysis-only cycle for the WestPac region.
- The code was ported to the DSRC Cray systems for use in the global HYCOM ensemble tests.

Summary points:

The observation and analysis perturbations are still being assessed, but subsurface temperature deviations appear consistent with forecast errors (Figure 9). The first implementation of a perturbation technique for the real profile data is to add simple uncorrelated random perturbations appropriate for instrument error, where some vertical correlation would more appropriately account for representation error. Figure 10 compares the perturbed temperature profiles for an in situ measurement and a synthetic profile generated using the Modular Ocean Data Assimilation System (MODAS). The synthetic profile perturbation is generated by perturbing the input surface state (sea surface height and temperature) parameters that are used to create the synthetic profile.

Milestone 5: HYCOM ensemble cycling

a) Modify scripts for retrospective ocean ensembles

b) Perform/test ocean retrospective ensembles

- Implemented and ran out a 20-member HYCOM ensemble on the NAVY DSRC.

Technical details:

- The first ensemble dataset was initialized for 01 Jan 2014 from the set of 01 Jan ocean model states from the HYCOM 20-year reanalysis.
- Each member cycled with an individual NCODA analysis of unperturbed observations from 01 Jan – 31 Mar 2014, then ran freely from 01 Apr – 30 Jun 2014.
- The second 10-member ensemble integration was repeated using the same setup for 01 Jul – 31 Dec 2014 to distinguish seasonal changes in variability from the impact of the DA and forecast.
- Software to calculate ensemble metrics for the global ocean ensemble is being developed, both for native HYCOM output and for post-processed z-level output in netCDF format (so-called NAVO netCDF).

Summary points:

This initial dataset is helping assess the timescale for the collapse of the ensemble spread with the unperturbed-observation DA, and the spread levels maintained during the extended forecast. The reanalysis ensemble will be extended with perturbed observations and forcing to assess those contributions to ensemble spread. As expected, the spread collapses rapidly from the interannual variability at the initial time when cycled with identical observations (Figure 11), and grows during the free-running forecast. Comparisons of the ensemble spread with the forecast error during the free-running forecast show the ensemble maintains skill past 30 d (anomaly correlations in excess of .6), and the ensemble spread increase tracks the growth of the RMS forecast error (Figure 12). The initial ensemble spread is low, consistent with the absence of analysis error in the ensemble generation. That will be addressed with the perturbed observation capability in NCODA.

c) Perform/test retrospective ensembles in coupled system with data components

Begins FY16

Milestone 6: Ensembles for Arctic prediction (NAVGEM/CICE/HYCOM)

- Ten-member time-lagged ensembles were completed for the Sea Ice Prediction Network (SIPN) for the June, July, and August SIPN reports.

  Technical details:

  SIPN is an international project to compare projections of September minimum sea ice extent. The runs were completed by starting the NAVGEM-HYCOM-CICE ESPC coupled system at using ten consecutive days at 12Z from the end of the previous month of the report. For example, the June report used 12Z initial conditions from May 20th 2015 to May 29th 2015. Initial conditions were provided from the respective data assimilation systems, i.e., NAVGEM's initial conditions were from NAVDAS-AR; and HYCOM and CICE's initial conditions were from NCODA.

  Summary points:

  Results from these lagged ensemble runs were consistent with other modeling centers.

- The Global Ocean Forecast System (GOFS 3.1) was run in forecast mode without data assimilation, initialized with May 1, June 1 and July 1 ice/ocean analyses, for ten simulations using the National Centers for Environmental Prediction (NCEP) Climate Forecast System Reanalysis (CFSR) atmospheric forcing fields from 2005-2014 to determine a seasonal projection of the September 2015 Arctic sea ice minimum.
-

  Technical details:
  Ice extent estimate was calculated using all grid cells with at least 15% ice concentration and averaged across all ensemble members. For the forecasts initialized from July 1, 2015, the sea ice extent estimate was 5.0 Mkm^2 with a range of 3.5-6.0 Mkm^2. The 2015 sea ice minimum was 4.41 Mkm^2 observed on September 11. The 2015 sea ice minimum is the fourth lowest extent recorded since satellites began measuring sea ice in 1979.

Additional tasks requested from the Sea Ice Outlook committee (and performed by NRL 7320) for the 2015 September minimum forecast efforts included: 1) determine the sea ice probability (SIP) map for September mean ice extent, 2) determine the first ice-free ordinal date map, 3) initialize GOFS 3.1 using May 1, 2015 University of Washington's PIOMAS ice thickness conditions and 4) initialize GOFS 3.1 using University of Washington's PIOMAS climatology ice thickness conditions.

Milestone 7: Ice cover and thickness perturbation based on perturbed observations and retrospective ensembles

Begins FY16

Milestone 8: Global ocean ensemble processing

Begins FY16

Milestone 9: Run and examine lagged ensembles

- Two tests of a time-lagged fully coupled ensemble were completed as part of an effort for the Sea Ice Prediction Network (SIPN).

  Technical details:

  - Each lagged ensemble consisted of 10 members.
  - Integrations were performed with the NAVGEM-HYCOM-CICE coupled system.
  - The integrations were initialized in May-July 2015 and were run to a lead time of either four or six months.
  - The resolution of the system components was NAVGEM (T359), CICE (1/12°), and HYCOM (1/12°).

  Summary points:

  These tests were valuable on multiple counts. They provided a first look at the variability and physical authenticity of the fully coupled system. In terms of variability, the lagged ensemble exhibits relatively little spread compared to the other coupled systems participating in SIPN (Figure 7). However, this is not all that surprising given the basic nature of the lagged method and the early days of the development process. In terms of physical authenticity, the ensemble-mean projection was for 4.5 million km2 of September minimum sea-ice extent. This is on the low end of dynamical models (Figure 7), but not unreasonable when compared with recent observed September minimums. The geographic pattern of sea ice extent is also plausible (Figure 8).

  These tests also highlighted some challenges in the output and storage of data, and enabled clock-timing tests to measure expected computational expense. The lessons learned from the tests prompted updates of the coupled-system run scripts to improve data management and to incorporate new output options. Also, two new coupled-system diagnostics (sea ice extent and

El Nino 3.4 temperature anomaly) were implemented to facilitate better interpretation of the test results.

- To complement the lagged ensemble effort, multiple additional baseline methods for generating ensemble perturbations in the fully coupled ensemble system were canvassed. They include variants of the ET method, conditioned random perturbations, and the "forecast anomaly" method of Oke et al. [2005]).

  Summary points:

  These additional methods are important as they are viable fallback options in the event that the lagged ensemble method proves unsuitable.

Milestone 10: Develop and implement coupled ensemble system infrastructure

Begins FY16

Milestone 11: Perform fully coupled retrospective ensembles

Begins FY16

Milestone 12: Develop and implement coupled ensemble forecast system

Begins FY17

Milestone 13: Ensemble evaluation and diagnostics

Begins FY16

RESULTS

Milestone 1: Develop ensemble system with simple diurnal SST model

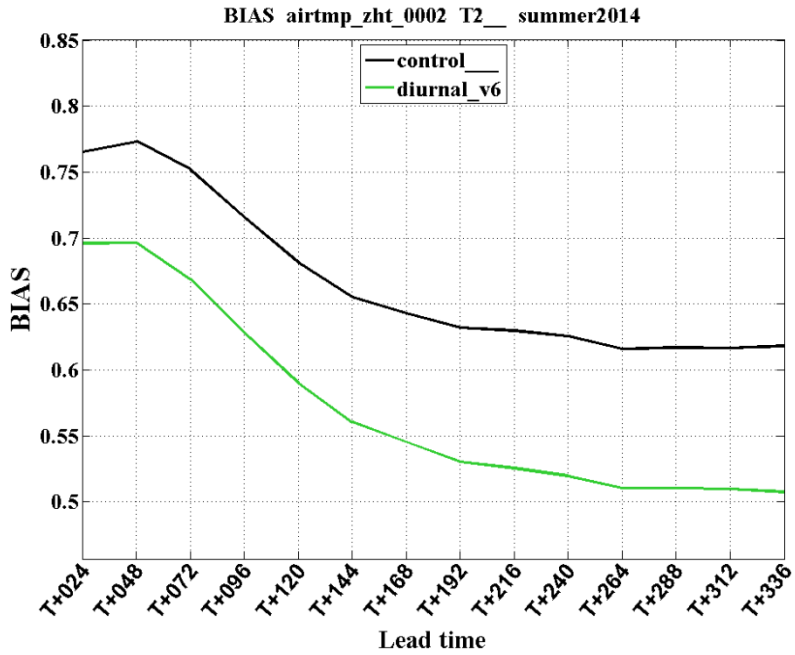

Figure 1 (above). BIAS of 2m air temperature for July 2014 for the Tropics [20S,20N] for the control run (black) and the diurnal SST model (green).

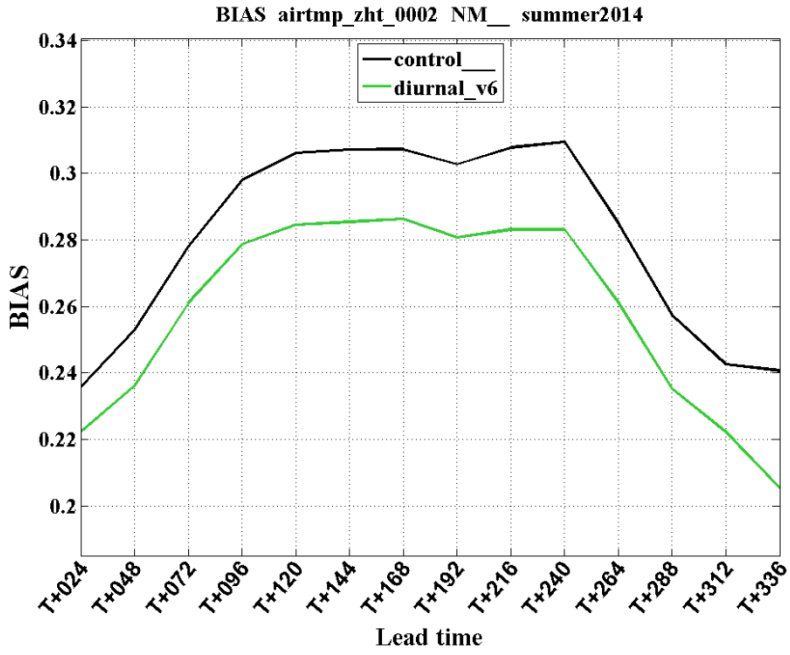

Figure 2 (above). BIAS of 2m air temperature for July 2014 for the NH Midlatitudes for the control run (black) and the diurnal SST model (green).

Milestone 2: Implement persistent SST anomaly capability

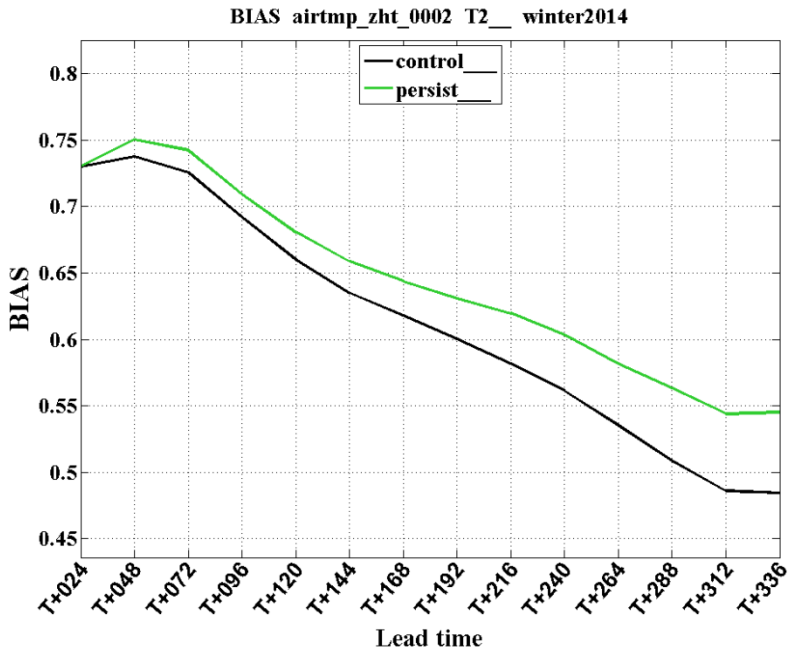

Figure 3 (above). Bias of 2m air temperature for January 2014 for the Tropics for the control run (black) and the persistent-anomaly SST capability (green).

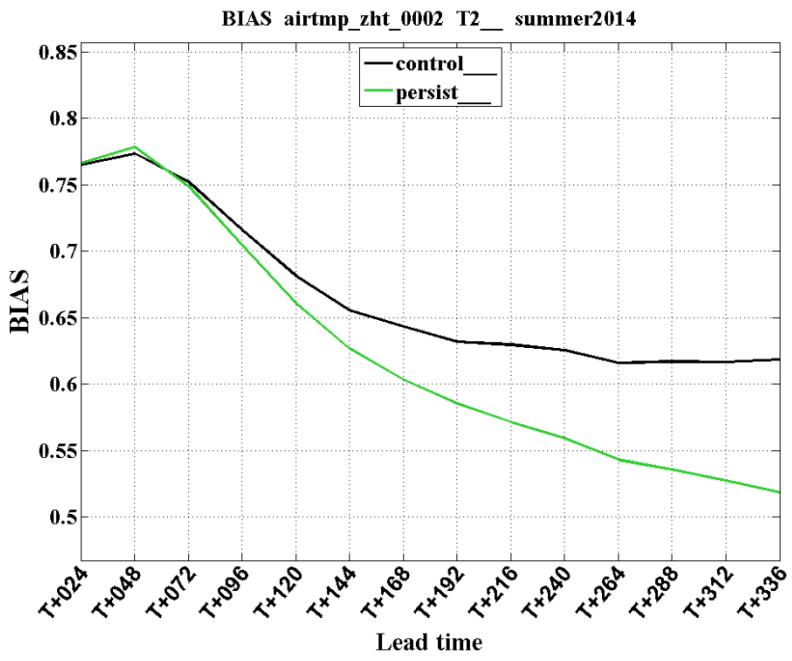

Figure 4 (above). Bias of 2m air temperature for July 2014 for the Tropics for the control run (black) and the persistent-anomaly SST capability (green).

Milestone 3: Implement initial SST perturbation capability

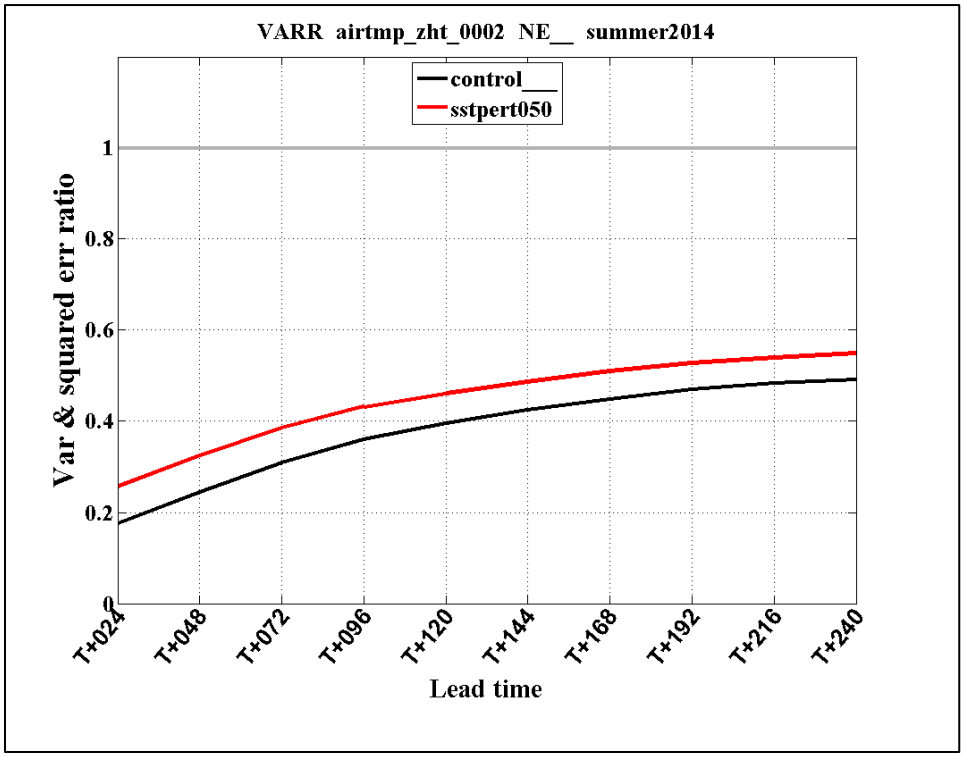

Figure 5 (above). Expected ratio of variance to squared-error for 2m air temperature for July 2014 for the NH Extratropics for the control run (black) and the initial SST perturbation capability (Red). Perfect forecasts have a ratio of one.

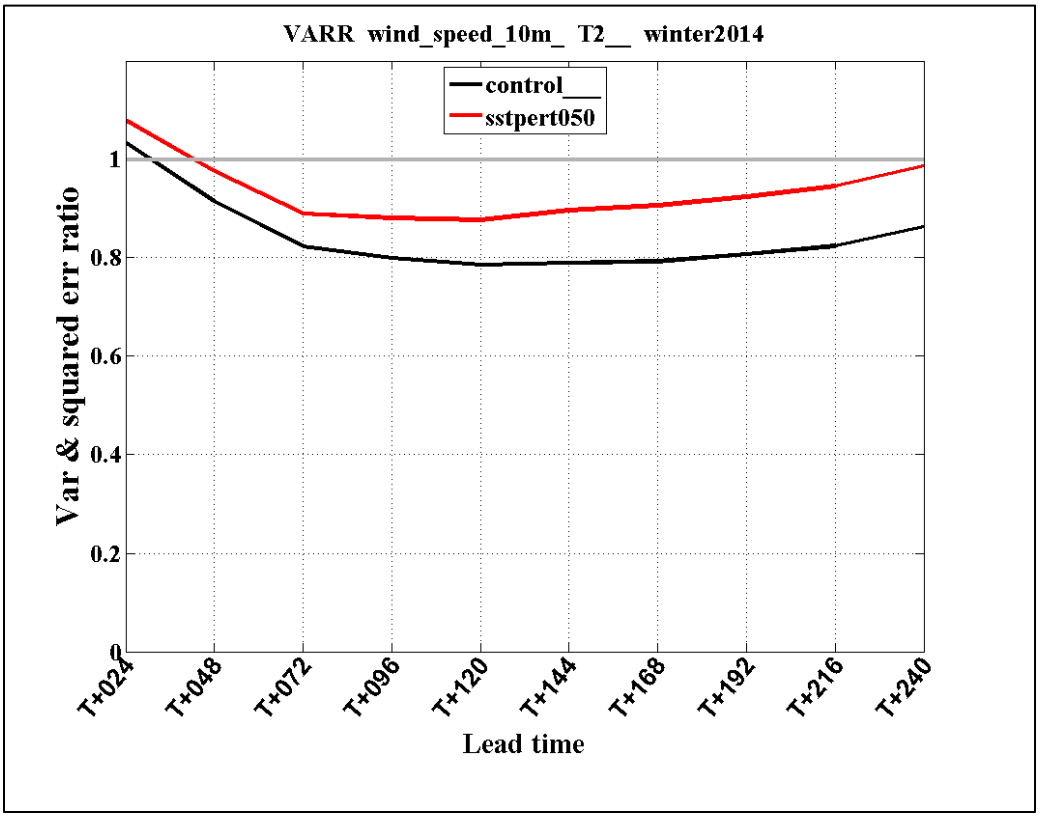

Figure 6 (above). Expected ratio of variance to squared-error for 10m wind speed for January 2014 for the Tropics (20°S to 20°N) for the control run (black) and the initial SST perturbation capability (Red). Perfect forecasts have a ratio of one.

Milestone 6: Ensembles for Arctic prediction (NAVGEM/CICE/HYCOM)
Milestone 9: Run and examine lagged ensembles

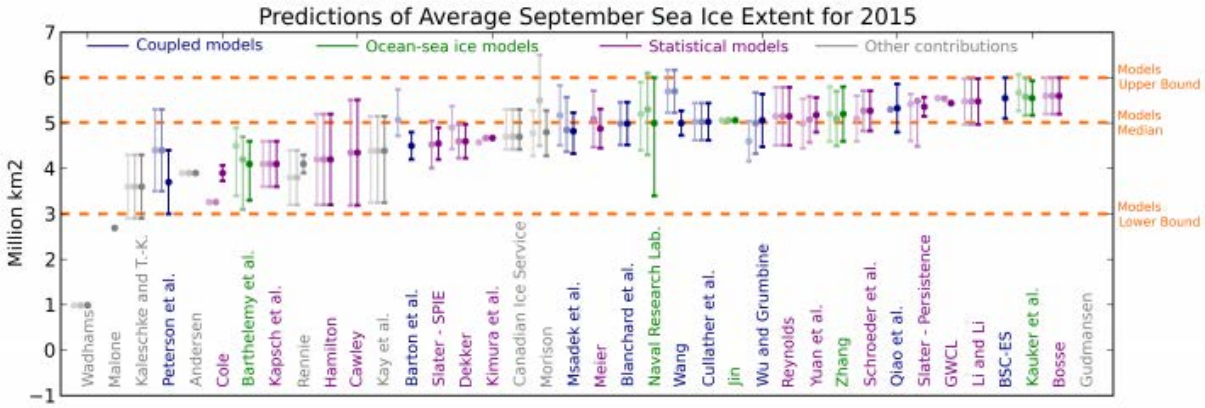

Figure 7 (above). Sea Ice Outlook predictions. Navy lagged ensemble is the entry labeled "Barton et al.", 13th from the left. The dots are the outlook estimates themselves and the intervals are the uncertainty ranges provided by the groups. The middle dashed horizontal line is the median of the August outlooks from dynamical modeling groups. The lower and upper dashed horizontal lines show, respectively, the lowest and highest bounds for all August modeling contributions. Figure courtesy of François Massonnet.

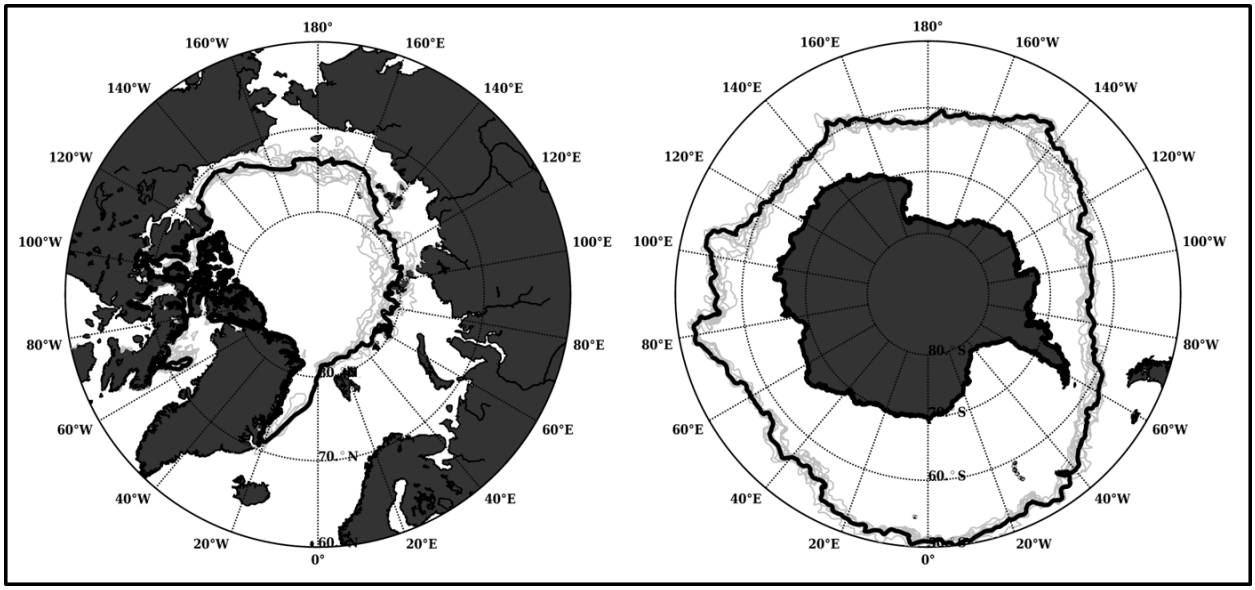

Figure 8 (above). September sea ice extent from the SIPN lagged ensemble experiment. Gray lines are the sea ice extent for each of the 10 ensemble members and the black line is the ensemble mean. Extent is defined by the 15% concentration threshold.

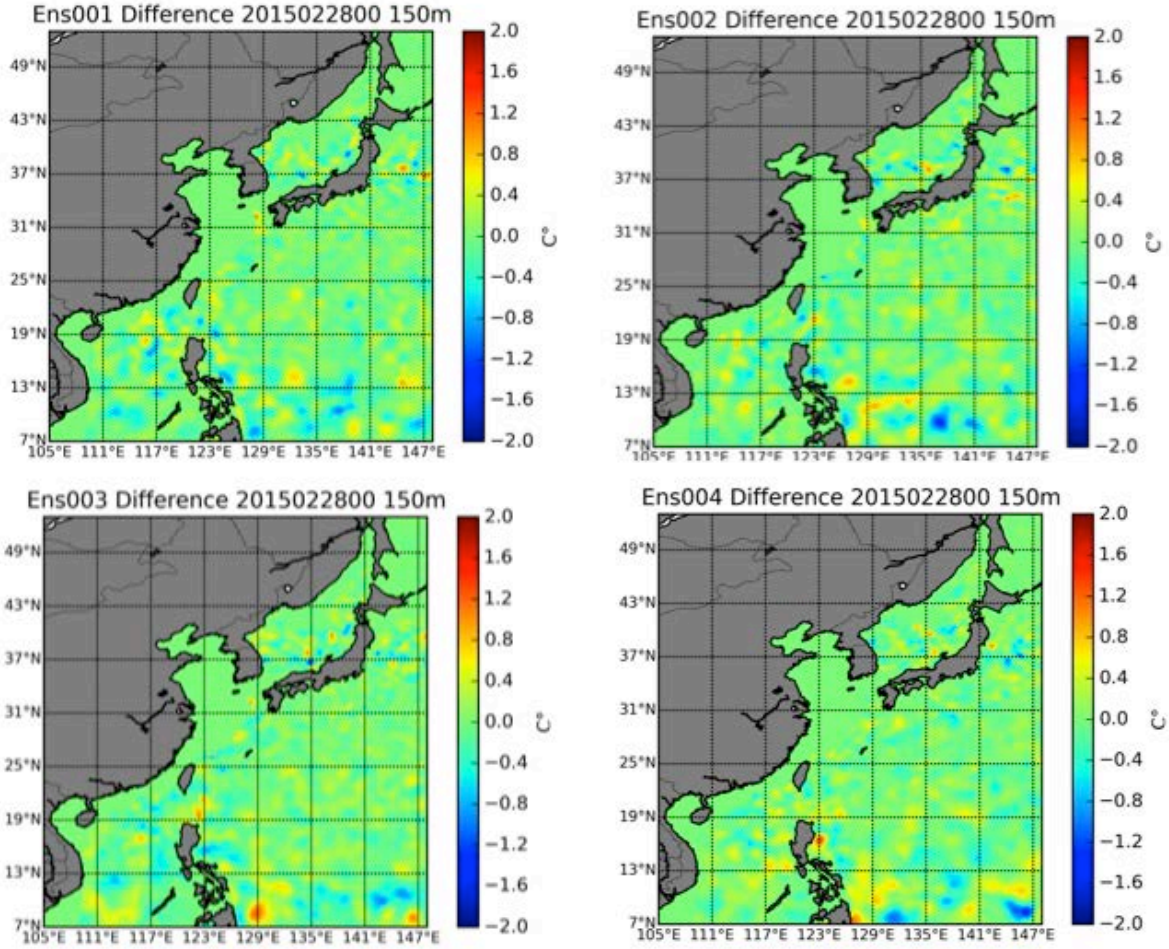

Figure 9. Sample NCODA ensemble perturbations from a 60-day analysis-only cycle of the perturbed-observation NCODA in the northwest Pacific, showing the 150 m temperature perturbations on 28 Feb 2015 for 4 members from the 16-member ensemble mean.

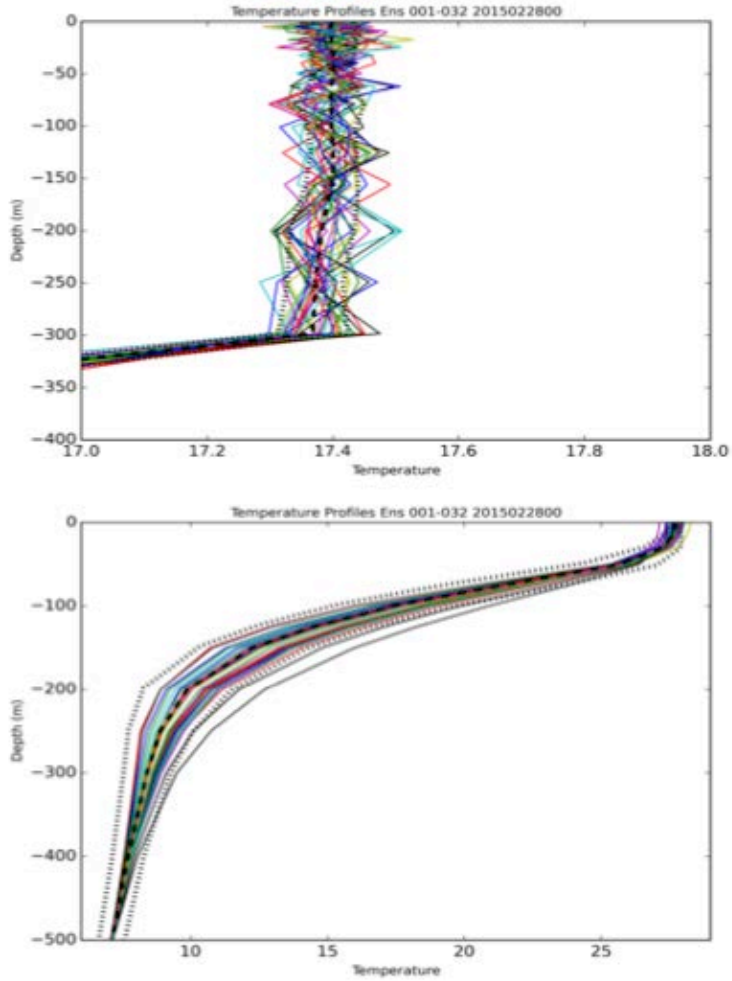

Figure 10. (upper panel) Sample perturbed temperature profile observations for an in situ profile using uncorrelated perturbations based on instrument error. (lower panel) Example perturbed temperature profile observations for a synthetic temperature profile using perturbed inputs to the synthetic profile algorithm (here, MODAS). In both the heavy dashed line shows the unperturbed profile.

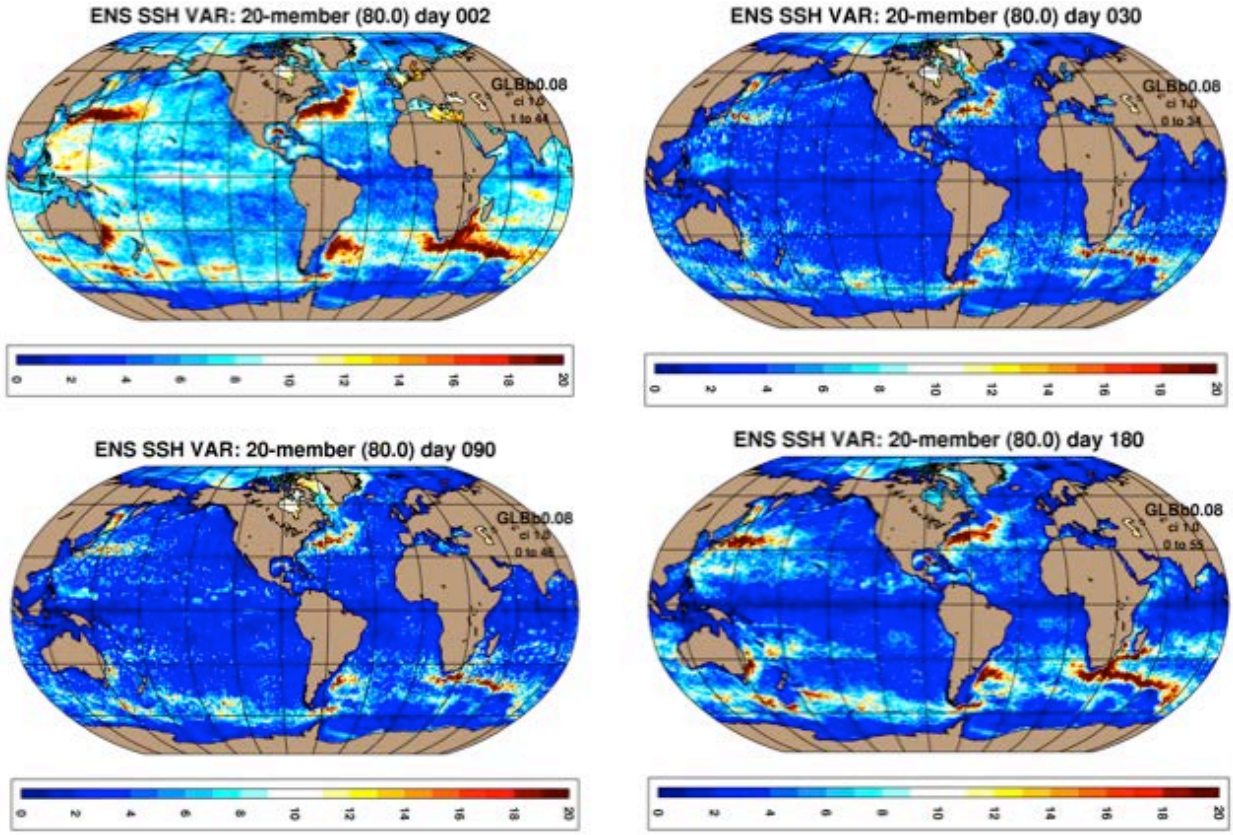

Figure 11. Sea surface height ensemble variance from the 20-member global HYCOM ensemble reanalysis, at days 2 (upper left), 30 (upper right), 90 (lower left), and 180 (lower right). The initial spread is the internannual variability from the 20-year reanslysis used to initialize the ensemble. The day 30 spread shows the collapse after one month of cycling with unperturbed observations. The growth from day 90 to day 180 represents the growth over the 3-month free run of the initial perturbations constrained by the observations.

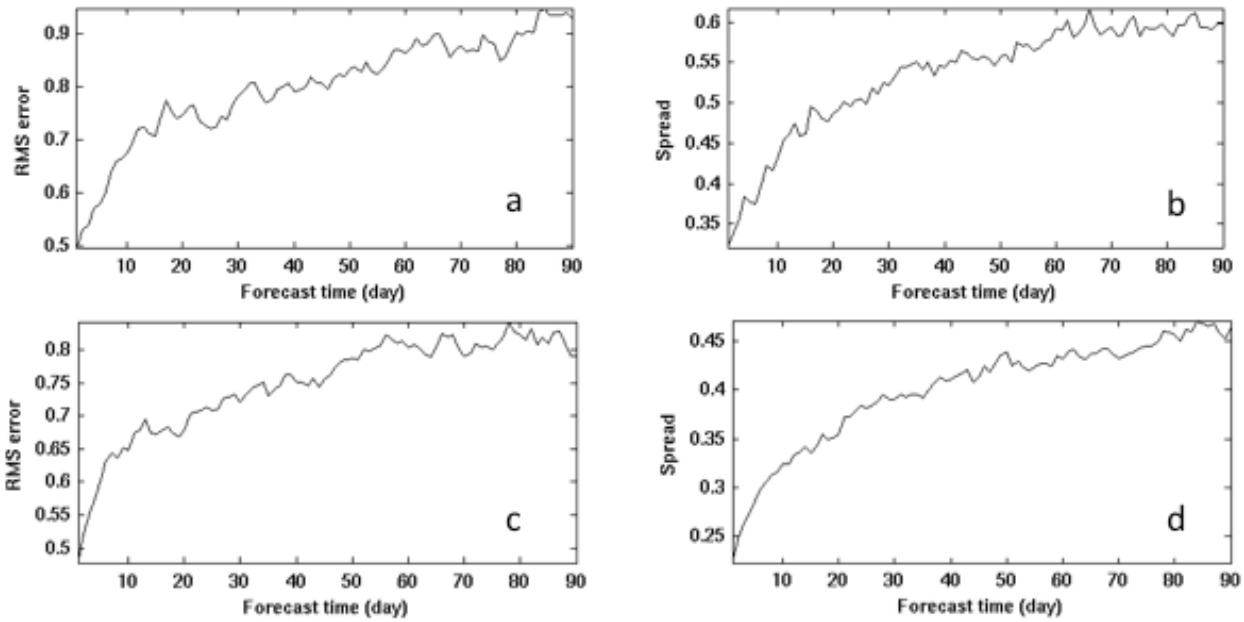

Figure 12. Comparison of daily temperature RMS error (a,c) and ensemble spread (b,d) of the free-running forecast of the 10-member global ensemble for 01 Oct – 31 Dec 2014, for the Atlantic (a,c) and Pacific (c,d) subdomains. The forecast was initialized at the end of the 01 Jul – 30 Sep 2014 analysis cycle using identical (unperturbed) observations.

IMPACT/APPLICATIONS

These experiments allow for an assessment of potential extended range probabilistic forecast utility (both in the atmosphere and in the ocean) for Navy-relevant metrics such as potential for high winds, extreme events, or tropical cyclones. Development of a skillful coupled ensemble forecast system will allow for DoD decision support on monthly time scales as well as improve capabilities reliant upon extended-range forecasts such as trans-ocean ship routing.

TRANSITIONS

The potential for extended-range forecasting demonstrated in this program will be transitioned to operations through existing and future 6.4 programs.

RELATED PROJECTS

Success of this project is dependent upon the success of the other ESPC components, particular the development of the coupled global prediction system (WU-1 ESPC), and the development of coupled data assimilation (WU-10 ESPC), which would provide consistent coupled initial perturbations for the ensemble system. Development of the atmospheric component of the coupled ensemble system will leverage work ongoing under the RTP project "Inclusion of Automated Environmental Uncertainty in Automated Ship Route Guidance" and under the 6.2 base-funded project "The Madden Julian Oscillation - Key to Coupled Extended-range Prediction".