**Probabilistic Prediction of Tropical Cyclone Track, Intensity and Structure
with an Analog Ensemble
NOAA Hurricane Forecast Improvement Project
Semi-annual Progress Report
1 September 2017 – 28 February 2018**

Project Personnel

Principal Investigator: William E. Lewis, Space Science and Engineering Center (SSEC) / University of Wisconsin-Madison (UW).

Co-Principal Investigator: Luca Delle Monache, National Center for Atmospheric Research (NCAR)

Co-Principal Investigator: Christopher Rozoff, NCAR

Co-Principal Investigator: Stefano Alessandrini, NCAR

Introduction

This project seeks to apply an Analog Ensemble (AnEn) technique (Delle Monache et al. 2011, 2013) to the problem of tropical cyclone (TC) track, intensity and structure prediction. The AnEn technique is used here to create a naturally calibrated ensemble prediction of TC track, intensity and structure from a training dataset composed of deterministic Hurricane Weather Research and Forecasting (HWRF) model forecasts. In the AnEn, a set of analog forecasts is created by searching for archived HWRF forecasts that share key features in common with a current forecast from the same configuration of HWRF. The actual intensity, structure and track displacement observations associated with each forecast are then used to produce an ensemble forecast.

The general AnEn technique applied to HWRF appears ideally suited for TC forecasting for the following reasons:

- One can use a higher resolution model for an ensemble prediction (since only one real-time forecast is needed for the AnEn),
- There is no need for initial conditions and model perturbation strategies to generate an ensemble,
- The forecasts are intrinsically reliable and no post-processing is needed,
- The flow-dependent error characteristics can be determined, and
- The AnEn is ideal for TC forecasting given its ability to improve the prediction of rare events, which may enhance the skill of HWRF's rapid intensification (RI) forecasts.

Key Scientific Accomplishments

The AnEn depends on a large dataset of archived forecasts made with a numerical model having a frozen (i.e. fixed) configuration. We collected all 2017 HWRF pre-implementation test forecasts in the Atlantic and Eastern Pacific Oceans (hereafter known as the H217 test set). There are 858 Atlantic and 1630 Eastern Pacific forecasts in this set (covering the years 2014-2016) and for each basin we processed predictors that are relevant to track, intensity, and structure. As part of the 2017 HFIP Demonstration, we configured and tuned the AnEn for rapid intensity prediction at 24-, 48- and 72-hr lead times. The list of predictors chosen for this purpose is shown in Table 1.

Table 1. Optimal Analog Ensemble (AnEn) predictors for Rapid Intensification (RI) by forecast lead time and basin.

	Atlantic	Eastern Pacific
0-24 hr	$\triangle v_{max}$ (HWRF), Symmetry of low-level inflow ($r = 0$ -100km), IVCN $\triangle v_{max}$	$\triangle v_{max}$ (HWRF), Min SLP, IVCN $\triangle v_{max}$
0-48 hr	$\triangle v_{max}$ (HWRF), CAPE ($r = 200 - 600$ km), Latent Heat Flux ($r = 0 - 50$ km), IVCN $\triangle v_{max}$	$\triangle v_{max}$ (HWRF), Total condensate ($r = 0 - 100$ km), IVCN $\triangle v_{max}$
0-72 hr	$\triangle v_{max}$ (HWRF), Storm translation speed, Latent Heat Flux ($r = 0 - 50$ km), IVCN $\triangle v_{max}$	$\triangle v_{max}$ (HWRF), Inertial stability ($r = 0 - 100$ km), IVCN $\triangle v_{max}$

Evaluation of the results from the 2017 HFIP Demonstration was undertaken and these findings are summarized in a series of plots of Brier Skill Score (BSS) evaluated at three forecast lead

times / intensity change thresholds. These are 24 hr ($\Delta V_{max} \geq 30$ kt), 48 hr ($\Delta V_{max} \geq 55$ kt), and 72 hr ($\Delta V_{max} \geq 65$ kt). The climatological probability of rapid intensification for each of these forecast lead times / thresholds is 5.9%, 3.9% and 6.8%, respectively (as computed from NHC's best-track database for the years 1987-2016). A positive BSS represents fractional increase in skill above the climatological baseline, and, likewise, a negative BSS conveys a decrease in skill.

Figure 1 displays results for the Atlantic basin for a homogeneous sample of 5 probabilistic models (HWMN, COMN, DTOP, SHIPS and AnEn). All models provide skillful forecasts at 24 hr, but only the AnEn displays skill at all lead times, performing particularly well at the all-important 48-hr forecast lead time.
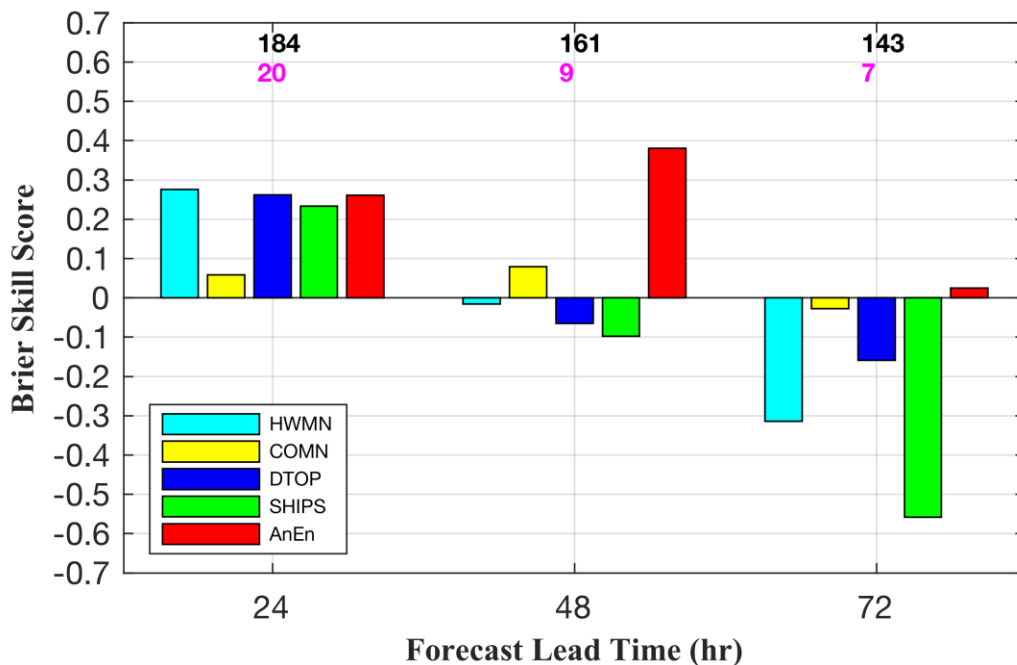

Figure 1. Brier Skill Scores for the 2017 Atlantic Hurricane season computed for the following lead times (rapid-intensification thresholds): 24 hr ($\Delta V_{max} \geq 30$ kt), 48 hr ($\Delta V_{max} \geq 55$ kt), and 72 hr ($\Delta V_{max} \geq 65$ kt). The models evaluated are HWMN (2017 HWRF ensemble), COMN (2017 COAMPS-TC ensemble), DTOP (Deterministic to Probabilistic Model), SHIPS (Statistical Hurricane Intensity Prediction Scheme), and AnEn (Analog Ensemble). Black (magenta) numbers at the top of the plot indicate the total number of forecasts (rapid intensification events) for a particular lead time. Results shown are for a 20-member analog ensemble.

During the analysis of the results it was noted that HWMN (the HWRF ensemble) and COMN (the Navy COAMPS-TC ensemble) had significantly fewer forecasts than the other models. Therefore, a separate comparison was made between the three remaining models (DTOP, SHIPS, and AnEn) whose collective homogeneous sample was thus significantly larger. These results are shown in figure 2. DTOP and AnEn are the superior performers, with skill in excess of 20% at each lead time, though the AnEn once again exhibits clear superiority at the 48-hr lead time.
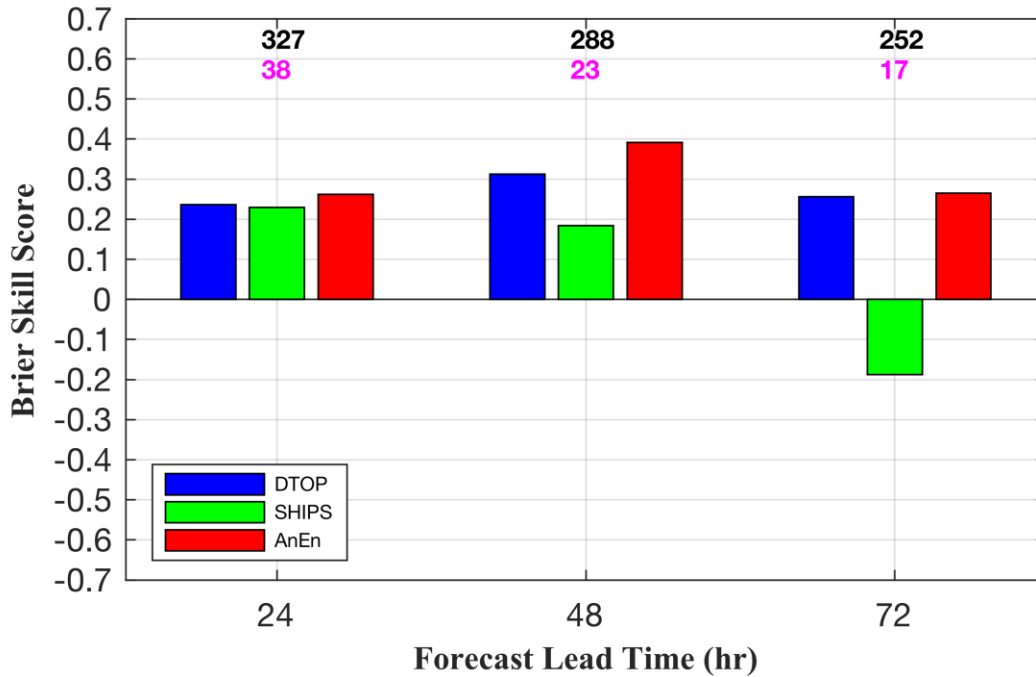
Figure 2. As in figure 1, but excluding HWMN and COMN.

There were significantly fewer forecasts in the eastern Pacific (the HFIP Demonstration period did not begin until 1 August, by which time nearly half the activity in the eastern Pacific had already occurred). Nevertheless, we evaluated model performance in this basin as well, choosing once again to exclude HWMN and COMN, as their inclusion in a homogenous sample would render the sample size too small to include any RI events.

Figure 3 shows that the Brier Skill Scores for each of the analyzed models in the eastern Pacific basin is generally smaller than its counterpart in the Atlantic. Whether this is associated with the particular predictability regime of the 2017 eastern Pacific season or is simply a consequence of the much smaller sample size is not yet known. However, what can be clearly noted is that, despite problems at 24 hr, the AnEn performs well at 48 and 72 hours, on par with DTOP and clearly superior to SHIPS.
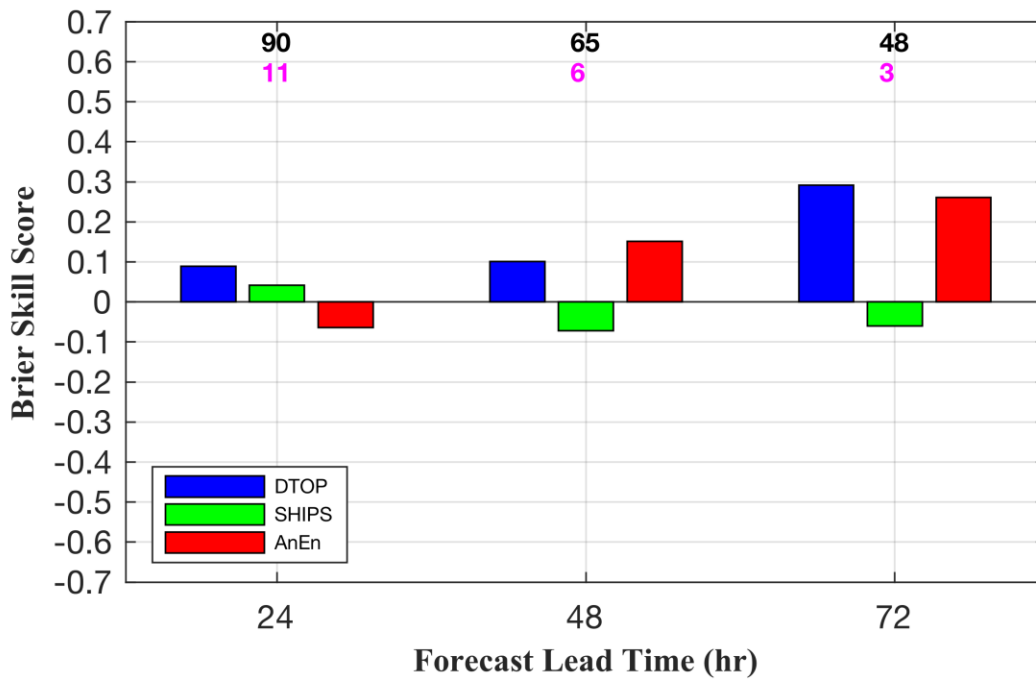
Figure 3. As in figure 2, but for the eastern Pacific basin.

In addition to the RI-version of the AnEn, significant progress has been made on the TC track-based AnEn system as well. Working with the H217 test set, the AnEn was configured to predict independent zonal and meridional displacements of the TC center at each forecast lead time. One potential complication of this approach is that, since separate ensembles are generated for each forecast lead time, the resultant tracks are not guaranteed to be smooth. In fact, initial tests produced tracks that frequently possessed numerous first-order discontinuities. However, the latest results use a Schaake Shuffle technique to couple the ensembles from one lead time to the next, producing tracks which are smooth and realistic.

Figure 4 shows mean displacement (i.e. position) errors for a comparison of the TC track-based AnEn and HWRF track forecasts. While it must be noted that the AnEn forecasts generally possess position errors which are of the same order (or slightly larger) than HWRF, the AnEn possesses excellent ensemble-spread / error characteristics (i.e. the ensemble spread at a given forecast lead time is a very good match for the RMSE). This represents a significant advance over dynamical ensembles, which generally are under-dispersive with respect to both intensity and track (and thus do not provide reliable estimates of model uncertainty).
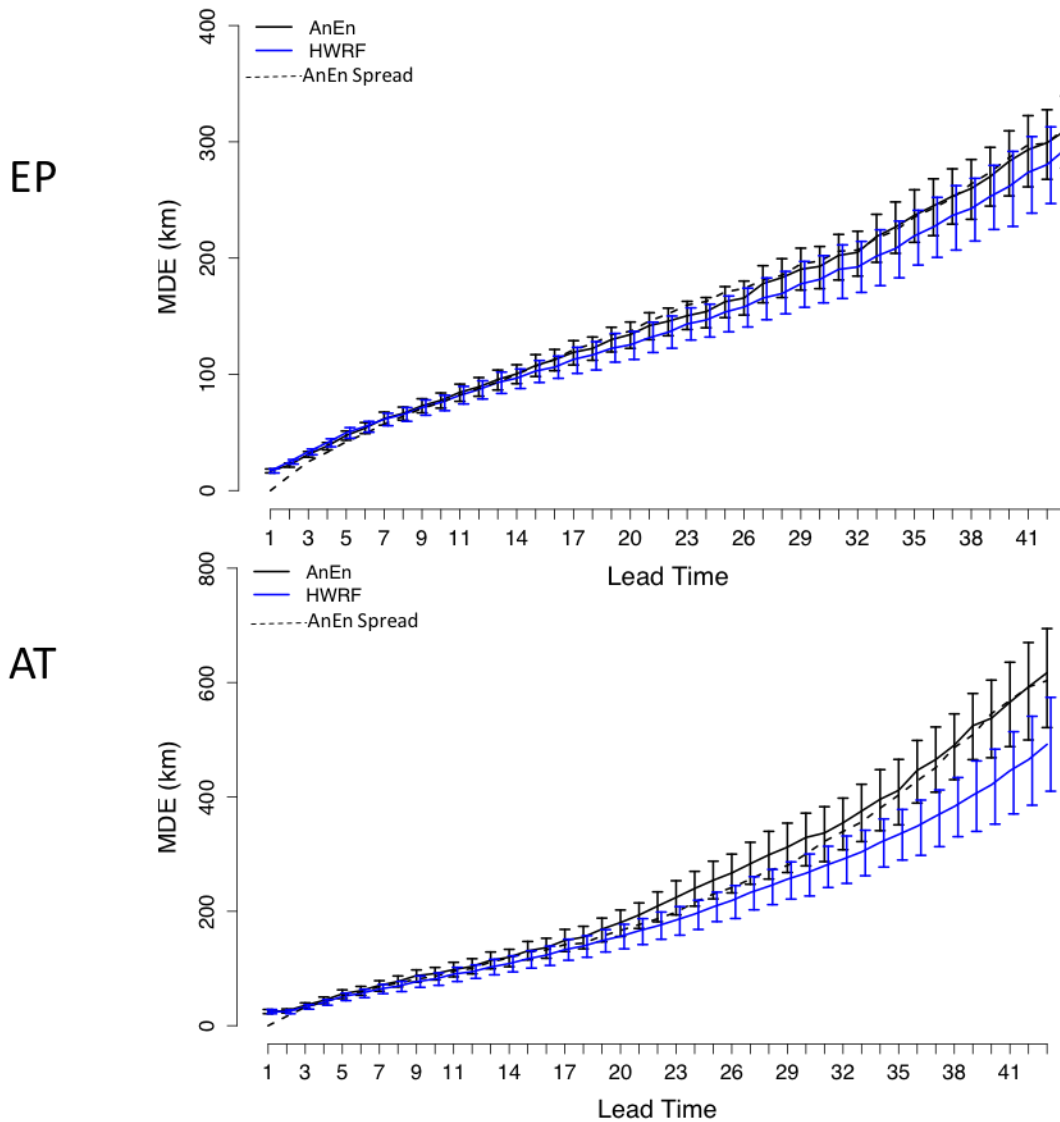
5

Figure 4. Mean displacement error (MDE, in km) and spread for TC-track based AnEn (black) and HWRF (blue) in the eastern Pacific and Atlantic basins. Results shown are for a 20-member analog ensemble.

Progress on testing, evaluation and/or verification of proposed improvements

We have completed all testing, evaluation and verification for the 2017 intensity-only / RI version of the AnEn. Testing and evaluation of the TC track-based and TC structure AnEn models continues using the H217 test set. All AnEn models will be tuned, tested and evaluated using the H218 test set as soon as it becomes available.

Interactions with NOAA/EMC, NHC and Developmental Test Center

We continue collaborating with the HFIP Ensemble Tiger Team led by Ryan Torn (SUNY Albany) and Mark DeMaria (NHC) and the DTC on comparing the AnEn's performance with other statistical and dynamical ensemble prediction systems. In particular, we will continue working with Ryan Torn to ensure that real-time AnEn output is made available to DTC (through which output will also be made available to NHC) this summer during the upcoming 2018 HFIP Demonstration.

Progress against milestones / schedule in Proposal

This project is slated to take place over a two-year period. The project timeline is shown below. Tasks 1 and 3 are complete, and tasks 2 and 4 (i.e. development of the track- and structure-based versions of the AnEn) will be completed within the next month. Tasks 5 and 6 are ongoing, as we will be tuning all AnEn systems for real-time operations in 2018 once the FY2018 HWRF (i.e. H218) pre-implementation forecast set is available.

Task	Activity
1	Develop HWRF-based predictors for all AnEn models (track, intensity, structure) from the HWRF retrospective test dataset (led by UW personnel) [Sep. 2016 – Aug. 2018]
2	Develop TC track-based AnEn system. [Sep. 2016 – Aug. 2017] (UW/NCAR)
3	Develop improved intensity and new intensity-change AnEn algorithms. (UW/NCAR) [Sep. 2016 – Aug. 2017]
4	Develop TC structure AnEn system. (UW/NCAR personnel) [Sep. 2017 – Aug. 2018]
5	Testing and Evaluation of AnEn systems, including coordination with NHC/HFIP Ensemble Tiger Team. (NCAR/UW personnel) [Sep. 2016 – Aug. 2018]
6	Real-time testing of AnEn systems during 2018 hurricane, delivery of code (UW/NCAR). [May 2018 – November 2018]

Previously unreported changes to the execution of the originally submitted proposal

None to report.

Outcomes that could be transitioned or offered to Operations

The 2017 HWRF intensity-only version of the AnEn was adapted for real-time rapid-intensification forecasts (one of NHC's highest priorities) during the 2017 HFIP demonstration. In this regard it displayed outstanding performance, providing probabilistic forecasts which were superior to the current state-of-the-art aids available to NHC in the Atlantic basin. These results have been compiled into a report for submission to *Weather and Forecasting*. The 2018 version of this code is currently in development and will be run in real time on NOAA's Jet supercomputer this summer; output will be communicated to forecasters at NHC (as well as other HFIP personnel and the general scientific community) via DTC and the AnEn website.

The 2018 track and structure versions of the AnEn are also nearing completion and will likewise be implemented in real-time on the Jet supercomputer during the HFIP demonstration later in 2018. Output will be made available via the AnEn website.

Budgetary issues

None to report.