Report on EN6 DTC Ensemble Task 2014: Preliminary Configuration of North American Rapid Refresh Ensemble (NARRE)

Motivation

As an expansion of computing resources for operations at EMC is becoming available through the support of the Sandy Supplemental Program, plans on how to best utilize these resources to address important forecasting questions have been made. One of the items on the EMC roadmap is an option to create an extension of the existing Short Range Ensemble Forecasting (SREF) system by adding the rapid refresh component. The idea is to have SREF continuing to run on 6-hourly cycles out to 84 forecast hours and NARRE will be a subset of 6 to 8 SREF members updated hourly and running out to 18-24 forecast hours. Having these members as a subset of SREF means that the model uncertainty, at least at the beginning, is going to be addressed by use of two dynamic cores ARW (RAP) and NAM (NMMB) and variations in physics. According to the EMC roadmap, this new system, consisting of 6 members, is expected to be implemented in operations during 2017 for improved aviation and probabilistic forecasts for other short-range applications. The work involves a very close collaboration between GSD, EMC and DTC staff. To further facilitate collaboration between the groups, common NARRE software versioning and revision control (SVN) will be in place by the end of the next DTC performance period, which is April 2016. The SVN development branch, to be set up by GSD, will contain all NARRE related code, including both data assimilation and modeling code.

Initial work included development of preliminary configurations that were tested in retrospective experiments. Initialization will ultimately be hourly ("Rapid"), but for the initial testing, the DTC employed less frequent updates. The forecast length used in testing was 24 hrs. The final configuration will depend on computing resources dedicated to this task and discussions with EMC colleagues on pre-NARRE design. NARRE will be configured with 13-km horizontal grid spacing and 60 vertical levels. For the initial testing we used Rapid Refresh (RAP) operational domain, which is somewhat smaller than the domain the system will use in operations.

Experiment Design

As previously mentioned, NARRE consists of two dynamic cores, ARW and NMMB. The first task was to set up the domain on the same rotated lat/lon grid. The integration domain is the same as the RAP operational domain (Fig. 1), with horizontal grid spacing of 13 km. In operations, the system will run over the larger NAM operational domain. The decision was made to perform experiments using eight members, four coming from RAP and four from NAM models. In addition to employing two different dynamic cores, model uncertainty is addressed by variations in physics. The experiment was designed to assess how physics diversity affects the ensemble performance. At the time, not many options for different physics packages were available for the NMMB members, so the decision was made to keep the operational physics suite for all NMMB members and perturb only initial and boundary conditions (Table 1). For this purpose, the control member was initialized with the Global Forecasting System (GFS) and the remaining three with three members of the Global Ensemble Forecasting System (GEFS). On the other hand, for ARW members, more physics variations were available, so changes were made in convection, surface physics, PBL physics, and microphysics. The majority of changes were based on combinations of the two, RAP and NAM, operational physics suites. In addition to the control being the RAP configuration, we added eight more members with variations in physics. The same initial and boundary conditions used for the NMMB members were used for the ARW members. Experiment testing, regarding the impact of the physics, focused on a five-day period: May 26-31, 2013.

Three members were drawn out of the eight experimental RAP configurations and combined with the control RAP member and four NAM members for eight-member ensemble performance evaluation. This approach resulted in the evaluation of fifty-six different perturbations. For the purpose of assessing the ensemble performance when using variations in physics, the ensemble verification system developed by EMC was employed. To evaluate the fifty-six perturbations of physics, we focused on changes in a few key statistics: RMSE of the ensemble mean, the Spread/Error ratio and the Continuous Rank Probability Skill Score (CRPSS). On average, the difference between the best and the worst ensemble configuration ended up being 8% for the RMSE, 10% for Spread/Error ratio, and 5% in CRPSS. The configuration with the highest scores was selected and the chosen members are highlighted in blue in Table 1.

Additional performance measures were evaluated, such as reliability diagrams and rank histograms. Figure 2 illustrates the reliability for 700-mb relative humidity and 850-mb temperature for three different lead times. It can be seen that for these fields reliability was generally good for all times, especially in the case of relative humidity. The 700-mb relative humidity rank histogram indicates some under dispersion, which is not too surprising for an ensemble with such a small number of members (Fig. 3). For the spread/skill measure, the ratio between the ensemble mean RMSE and the spread was evaluated. The ratio results for different lead times and various variables (T2m, T850, u10, v10, U500, V500 and RH700) are presented in Figure 4a. The values indicated fairly good spread skill correlation, with values relatively close to one for all variables. Similar analyses were performed when the same ensemble configuration was used for simulations of the cold season period (Figure 4b). The idea was to confirm that the same configuration would produce similar results for different seasons. Results are presented in Fig. 4b, and it can be seen that for the cold season, the spread-skill relation was slightly improved as compared to the warm season experiment. Based on these results, we proceeded with this configuration for real-time runs.

Since November 15, 2014, the preliminary NARRE configuration has been running in real time. In order to participate in HMT-WPC Winter Weather Experiment (WWE), NARRE had to run out to at least 48 hours. In the real-time mode, the rapid refresh component was employed only for ARW members (the rapid refresh infrastructure for NMMB members was not available at the time). ARW members were cycled hourly but 48-hour forecasts from all members were made only at 00Z and 12Z.

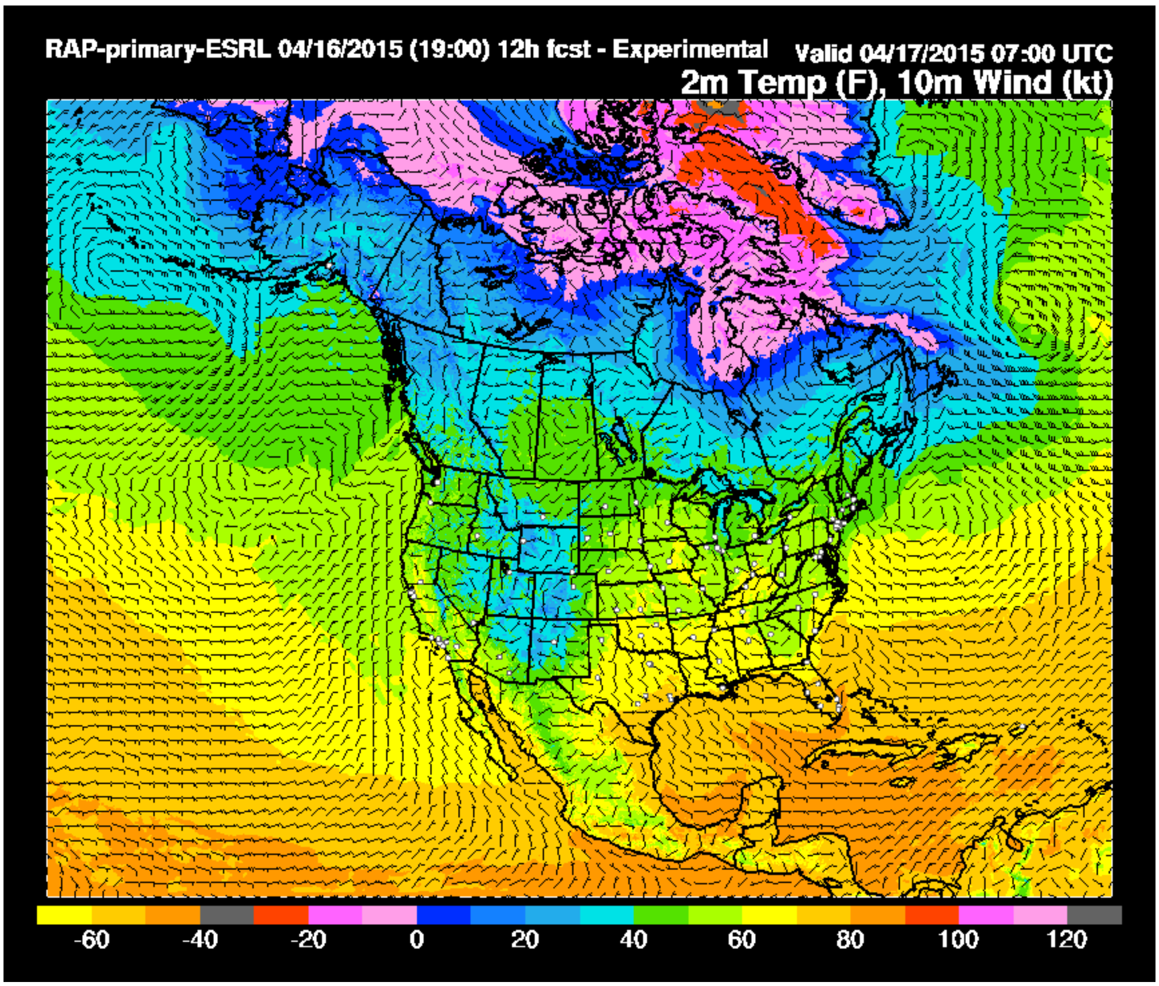

Figure 1. Example of the RAP operational domain

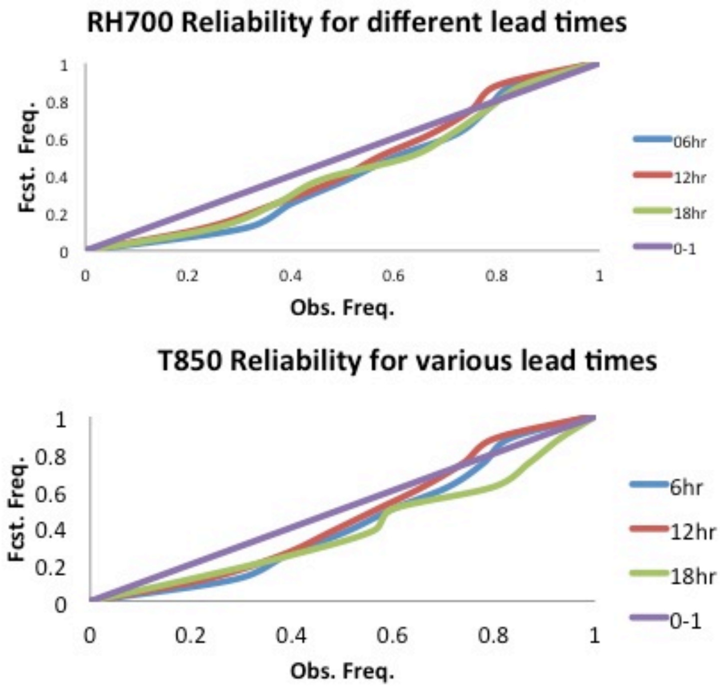

Figure 2. Reliability diagrams for 700-mb relative humidity and 850-mb temperature for different lead times and warm season period of interest.

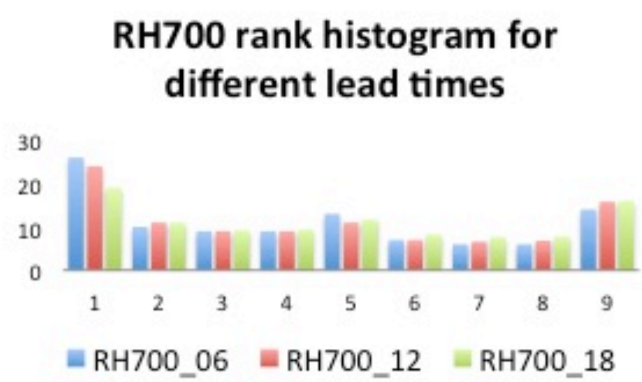

Figure 3. Rank Histogram for 700-mb relative humidity for different lead times and the warm season period of interest.

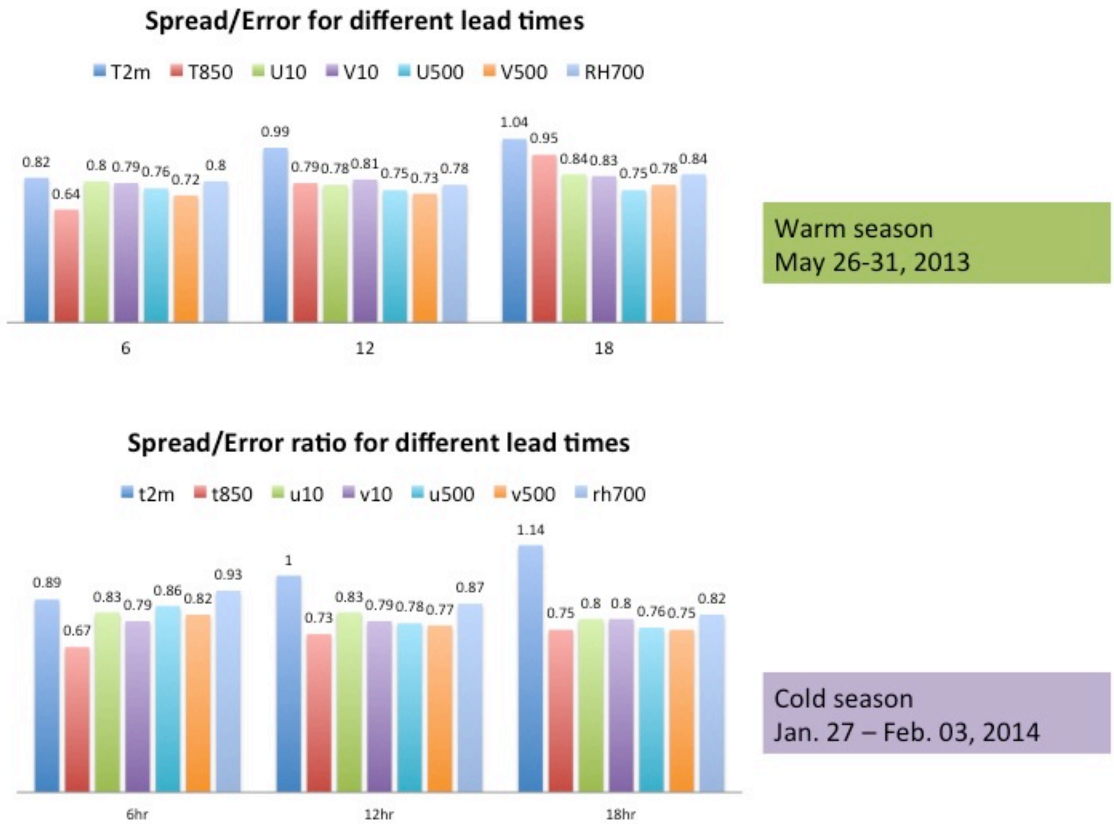

Figure 4. Spread/Error ratio for 2-m temperature, 850-mb temperature, 10-m wind, 500-mb wind and 700-mb relative humidity for different lead times and two periods of interest.

	MP	Sfclay	Sfcphy	PBL	CU	IC/LBs
rap ctl	Thompson	MYNN	RUC	MYNN	GF	GFS
rap1	Thompson	MO-MYJ	RUC	MYJ	BMJ	GEP01
rap2	Ferrier	MO-YSU	RUC	YSU	BMJ	GEP02
rap3	Ferrier	MO-MYJ	RUC	MYJ	BMJ	GEP03
rap4	Ferrier	MO-MYJ	NOAH	MYJ	BMJ	GEP01
rap5	Ferrier	MYNN	RUC	MYNN	GF	GEP02
rap6	Ferrier	MYNN	RUC	MYNN	BMJ	GEP03
rap7	Thompson	MO-YSU	RUC	YSU	BMJ	GEP04
rap8	Ferrier	MO-YSU	RUC	YSU	GF	GEP01
nmmb ctl	Ferrier	MYJ	NOAH	MYJ	BMJ	GFS
nmmb1	Ferrier	MYJ	NOAH	MYJ	BMJ	GEP01
nmmb2	Ferrier	MYJ	NOAH	MYJ	BMJ	GEP02
nmmb3	Ferrier	MYJ	NOAH	MYJ	BMJ	GEP03

Table 1. List of members tested. The green color indicates NMMB members and blue indicates ARW members selected based on the results of the experiment.