

# Report on EN2 DTC Ensemble Task 2015: Testing of Stochastic Physics for use in NARRE

## **Motivation:**

With growing evidence that initial-condition uncertainties are not sufficient to entirely explain forecast uncertainty, the role of model uncertainty is receiving increasing attention. For instance, most global and regional ensemble systems are under-dispersive, producing unreliable and over confident ensemble forecasts. Current plans for a North-American/CONUS ensemble address model-related uncertainty by using a combination of multi-dycore and multi-physics, where the physics suites are those used in current operational systems and variations of these well-tested suites. However, an ensemble system based on multiple physics suites comes at the price of needing to maintain multiple physics packages as well as concerns related to individual member bias and climatology. To minimize ensemble system maintenance and better represent random (not individual member) forecast error, the ultimate goal is to move towards representing model uncertainty through a stochastic physics approach. Before moving forward with such a transition, the value added by applying a stochastic physics approach must be demonstrated.

## **Experiment Design:**

To test a number of different stochastic physics approaches versus multi-physics, the regional Rapid Refresh (RAP) model was chosen (operational domain in 2015 shown in Figure 1). Stochastic vs. multi-physics simulations were run using a data set consisting of seven 2013 warm season days. The seven days included May 23 and 29, June 7, 14, 20, 28 and July 5. The simulations were initialized at both 00Z and 12Z and run out 24 hours. In these experiment, we used a "cold start" approach to model initialization. In other words, there was no cycling or a rapid refresh component included. For initialization and lateral boundary conditions, the first eight members of the NCEP Global Ensemble Forecasting System (GEFS) were used.

Several different ensemble experiments were conducted to compare performance between the multi-physics ensemble and the ensembles created by employing stochastic approaches. The list of experiments is presented in Table 1. The first multi-physics (control) ensemble includes variations in convective and PBL treatments. The second ensemble, SPP, comprises stochastic parameter perturbations within the GF and MYNN PBL schemes. Ensemble three combines the SPP approach with SKEB. Similarly, ensemble four combines SPP with SPPT. Finally, ensemble five includes all three stochastic approaches combined together (SPP, SKEB, and SPPT).

Configurations of the multi-physics and the stochastic parameter perturbation ensemble members are presented in Table 2. The first four members

(Control0 to Control3) of the multi-physics ensemble have different convective treatments while the MYNN PBL scheme have the same. The other four multi-physics members (Control4 to Control7) include different PBL schemes and a constant convective treatment with the GF scheme. All of the stochastic physics ensemble members have the same operational RAP configuration except the first four members (Stoch0 to Stoch3) including a stochastically perturbed GF scheme, and the last four members (Stoch4 to Stoch7) included a stochastically perturbed MYNN PBL scheme. The perturbed parameters were chosen based on suggestions by the respective parameterization experts and developers. Information on specific parameter perturbations can be found in Table 3.

## **Results:**

While all simulations took place over the RAP operational North American domain, verification of results was limited to the continental United States (CONUS). An example of this domain is shown in Figure 2. Both precipitation and surface/upper-air verification was performed. Precipitation verification indicated that all ensembles were slightly under-dispersive, yet the stochastic method ensembles resulted in very similar rank histograms to the multi-physics ensemble (Fig. 3). The SPP ensemble (spp) performed the best in all categories of frequency bias, with all ensembles having a high frequency bias for light precipitation, and a low frequency bias for heavy precipitation (Fig. 4). Brier scores were better for all ensembles and for all precipitation thresholds during nighttime hours (Fig 5).

For the other variables analyzed, such as 2-m temperature, 10-m wind, 850 mb temperature, 250 mb wind and 500 mb height, RMSE was found to be similar and generally increased in time for all ensembles (Fig. 6). Spread also increased with time for each ensemble, but varied widely between ensembles, with SPP containing the least spread. When the spp ensemble was combined with other stochastic methods (SKEB and SPPT), a statistically significant increase in spread was observed. Therefore, spread/error ratios were consistently larger (closer to 1) for ensembles that combined different stochastic approaches (spp skeb, spp sppt and spp skeb sppt) for all variables and all lead times (Fig. 7). For most of the lead times and all variables, the best spread/error ratios were produced for the ensemble that combined all stochastic approaches together (spp skeb sppt).

While the stochastic parameter perturbation approach showed some promise, especially for precipitation simulations, its general performance is hampered by limited spread. To determine whether this scheme adds value in combination with other spread-generating approaches, we assessed the performance of the spp sppt experiment against that of using only SPPT (sppt). Figure 8 shows the percent difference of spp sppt over sppt for a number of verification metrics. The bars in Fig. 8 are plotted only for variables, lead times and statistics for which the difference was non-zero. Analysis shows that the use of SPP generally has a positive impact. This benefit is most pronounced by increasing spread, which consequently improves spread/error ratio. The impact on CRPS was overall positive, with the exception of upper air variables in 12 UTC model runs.

## **Conclusions:**

A new stochastic parameter perturbation (SPP) scheme was developed and compared alone and in combination with other stochastic methods (SKEB, and SPPT) against a multi-physics, baseline configuration. The SPP scheme introduces temporally and spatially varying perturbations to key parameters in the GF convection and MYNN PBL parameterizations. Twenty-four hour RAP ensemble simulations were verified for seven days during the summer of 2013, using GEFS data for 00 UTC and 12 UTC initial times and lateral boundary conditions. The most important findings are summarized below:

- Alone, the parameter perturbations of SPP introduce insufficient spread. However, when combined with SKEB and/or SPPT the spread is as large, and for some instances, even larger than for a multi-physics ensemble. Overall the ensemble mean error is changed only slightly.
- An ensemble created by combining three stochastic approaches (SPP, SKEB and SPPT; spp skeb sppt) generally outperformed the multi-physics, control ensemble for most of the examined variables, most of the evaluated lead times, and most of the employed statistics. SKEB made a larger impact on spread associated with upper level wind and geopotential heights, while SPPT had a larger impact on spread for near-surface temperature.
- Combining SPP with SPPT had generally a positive impact, on the order of a 2-10% improvement over an ensemble using SPPT alone. This is an important finding since SPP represents uncertainty at its source, in a physically consistent way, without introducing systematic member biases.

Results show that parameter perturbations alone do not generate sufficient spread to remedy the under-dispersion in short-term ensemble forecasts, and that a combination of several stochastic schemes can outperform any single scheme. This result implies that a synthesis of different approaches is best suited to capture model error in its full complexity. This finding may potentially have a large impact on the design of next-generation high-resolution regional, as well as global, operational ensembles. The single-physics stochastic approach clearly provides a viable alternative to using the costly and theoretically problematic multi-physics approach in operations.

**Tables:**

<b>Number</b>	<b>Ensemble</b>
1	Multi-physics (Control)
2	Stochastic parameter perturbation for GF and MYNN (stoch)
3	Stochastic parameter perturbation and SKEB (stoch_skeb)
4	Stochastic parameter perturbation and SPPT (stoch sppt)
5	Stochastic parameter perturbation and SKEB and SPPT (stoch skeb sppt)

Table 1. List of ensembles created for comparison between multi-physics and stochastic approaches.

<b>Multi-Physics Members</b>	<b>Convective Parameterization</b>	<b>PBL Parameterization</b>	<b>LSM Parameterization</b>
Control0	OSAS	MYNN	RUC
Control1	BMJ	MYNN	RUC
Control2	GF	MYNN	RUC
Control3	NSAS	MYNN	RUC
Control4	GF	MYJ	RUC
Control5	GF	YSU	RUC
Control6	GF	BOULAC	RUC
Control7	GF	MYNN	RUC
<b>Stochastic Members</b>	<b>Convective Parameterization</b>	<b>PBL Parameterization</b>	<b>LSM Parameterization</b>
Stoch0	GF-perturbed	MYNN	RUC
Stoch1	GF-perturbed	MYNN	RUC
Stoch2	GF-perturbed	MYNN	RUC
Stoch3	GF-perturbed	MYNN	RUC
Stoch4	GF	MYNN-perturbed	RUC
Stoch5	GF	MYNN-perturbed	RUC
Stoch6	GF	MYNN-perturbed	RUC
Stoch7	GF	MYNN-perturbed	RUC

Table 2. List of members for multi-physics and stochastic approaches.

<b>Perturbations to Grell-Freitas Convection Scheme</b>		
<b>Perturbed Parameter</b>		<b>Name</b>
Closure		xf_ens
<b>Namelist Parameter</b>	<b>Value (in SI)</b>	<b>Value (other than SI)</b>
rand_perturb2	1	
gridpt_stddev_rand_pert2	0.3	
stddev_cutoff_rand_pert2	3.0	
lengthscale_rand_pert2	150000 m	150 km
timescale_rand_pert2	21600 s	6h
<b>Perturbations to Planetary Boundary Layer Scheme</b>		
<b>Perturbed Parameter</b>		<b>Name</b>
Turbulent Mixing Length		el
Subgrid Cloud Fraction (correlated with EL)		cldfra_bl
Thermal and Moisture Roughness Length (anti-correlated and twice the amplitude)		CZIL
<b>Namelist Parameter</b>	<b>Value (in SI)</b>	<b>Value (other than SI)</b>
rand_perturb3	1	
gridpt_stddev_rand_pert3	0.15	
stddev_cutoff_rand_pert3	2.0	
lengthscale_rand_pert3	700000 m	700 km
timescale_rand_pert3	21600 s	6h

Table 3: Summary of namelist parameter settings for stochastic perturbation patterns.

**Figures:**

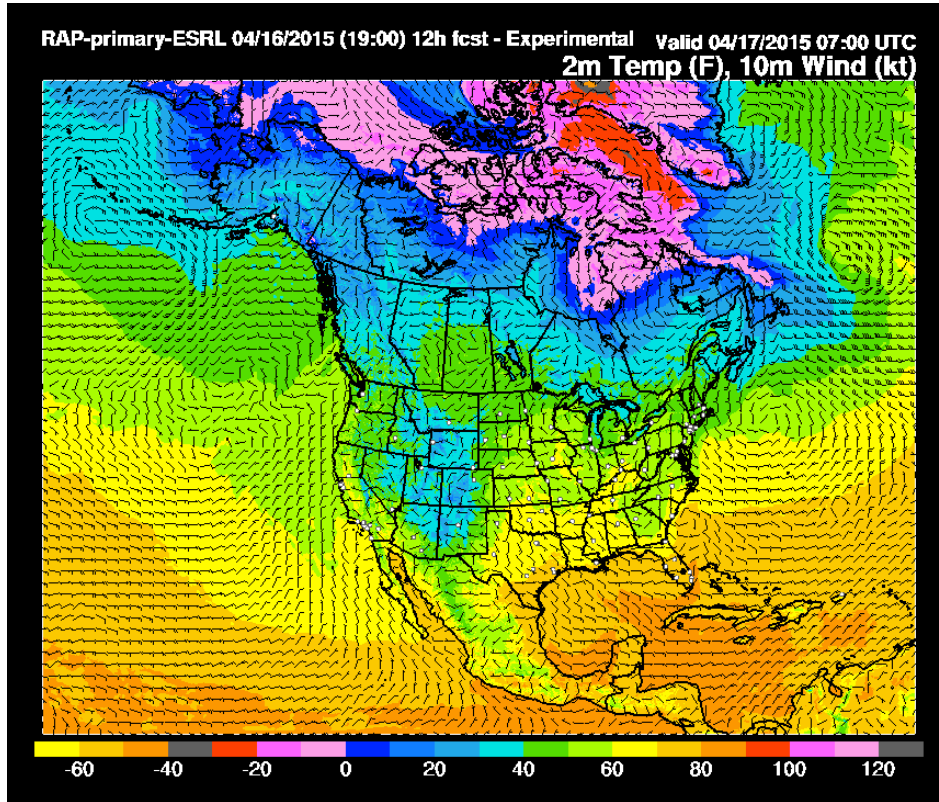


Figure 1. Example of the RAP operational domain.

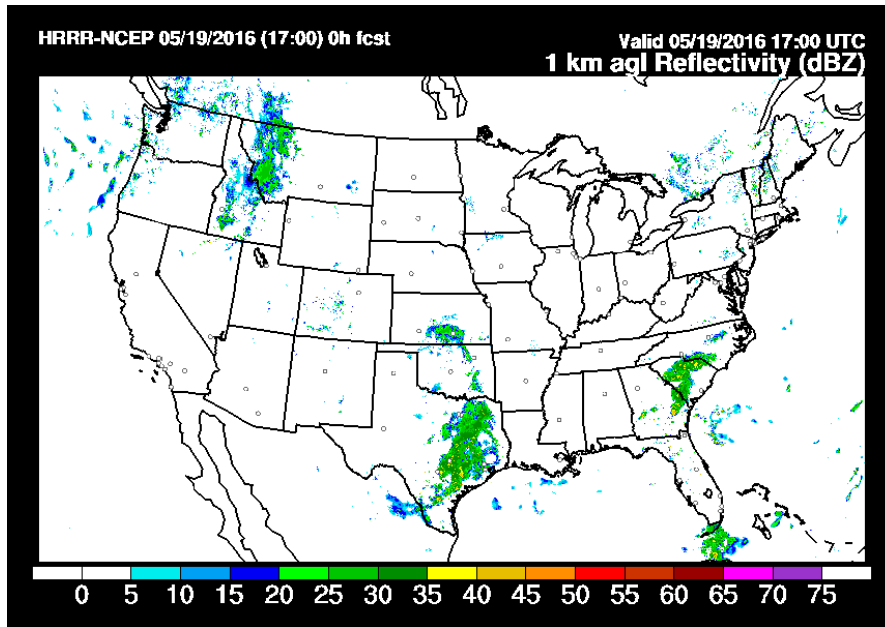


Figure 2. Example of the HRRR operational domain.

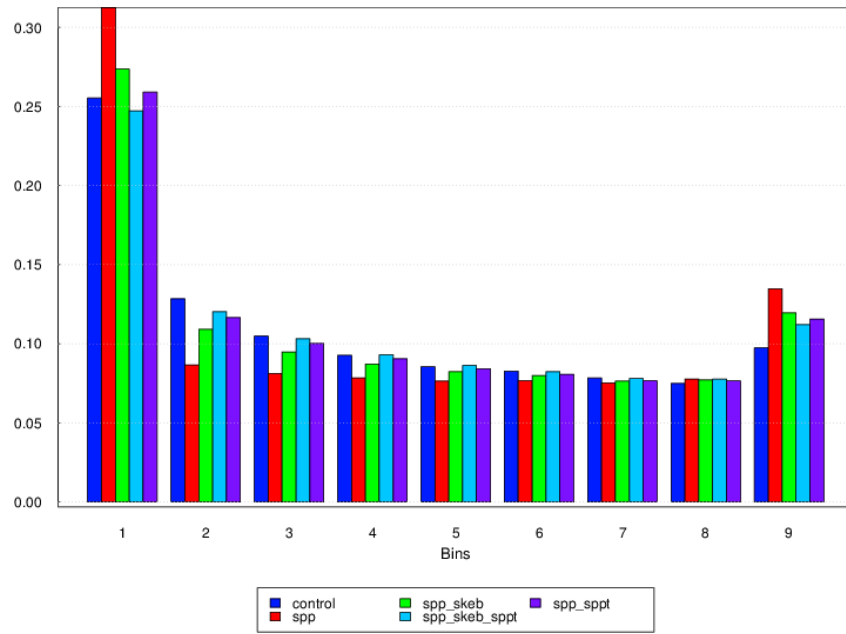


Figure 3. Rank histogram for 3-hourly precipitation for 00Z initialization and for control (red), stochastic (lime), stoch\_skeb (green), stoch\_sppt (dark blue), and stoch\_skeb\_sppt (purple) ensembles.

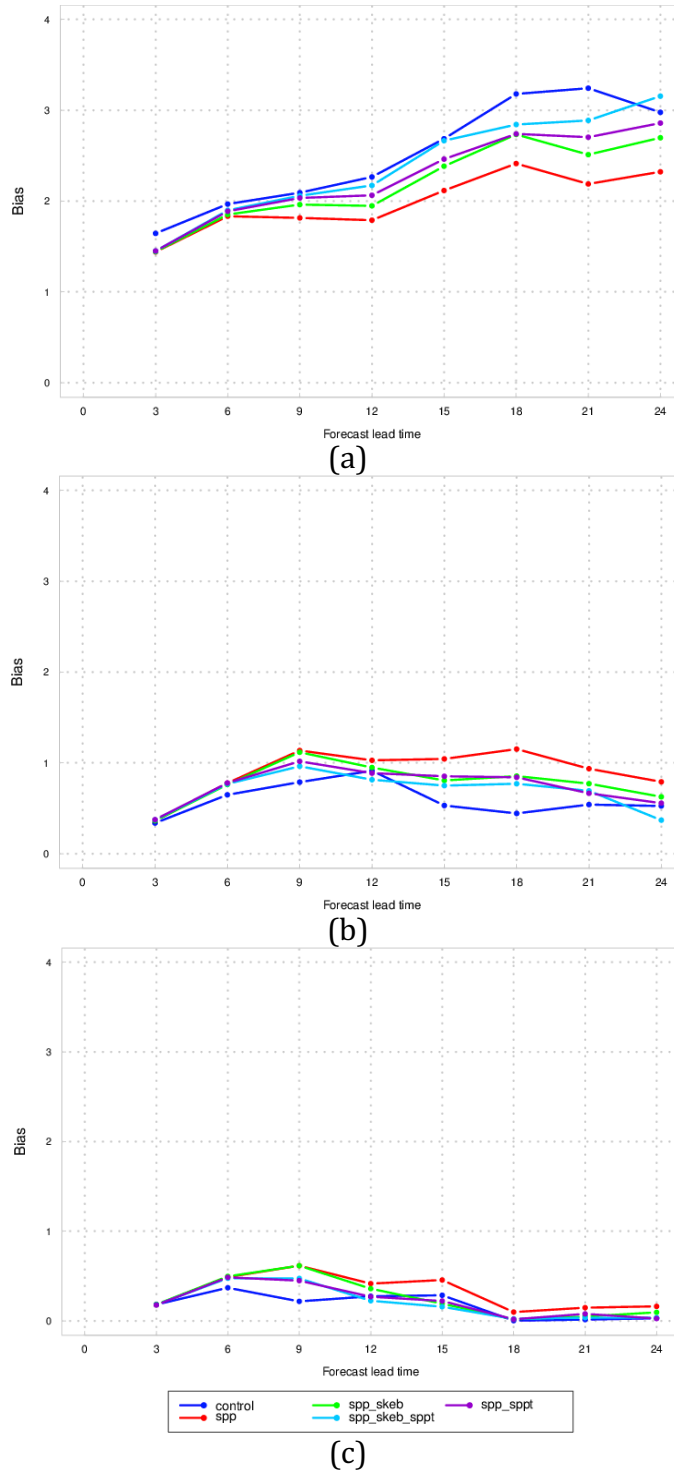
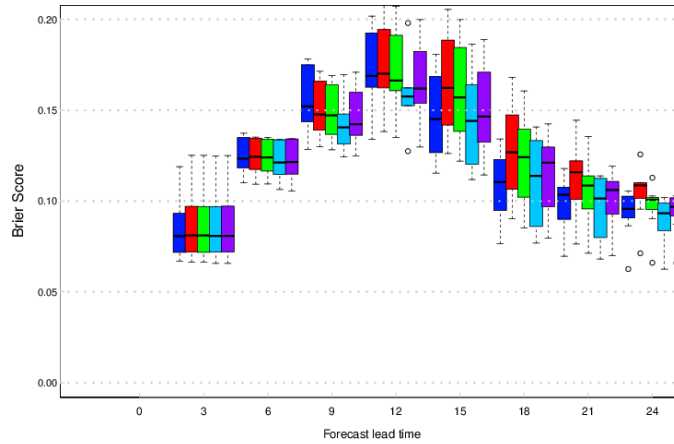
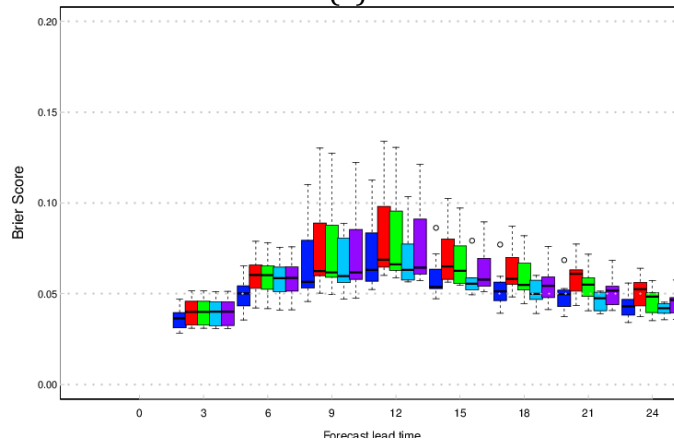


Figure 4: Frequency bias for 00Z initialization and for three precipitation thresholds a) 0.254 mm, b) 6.35 mm, and c) 12.7 mm.

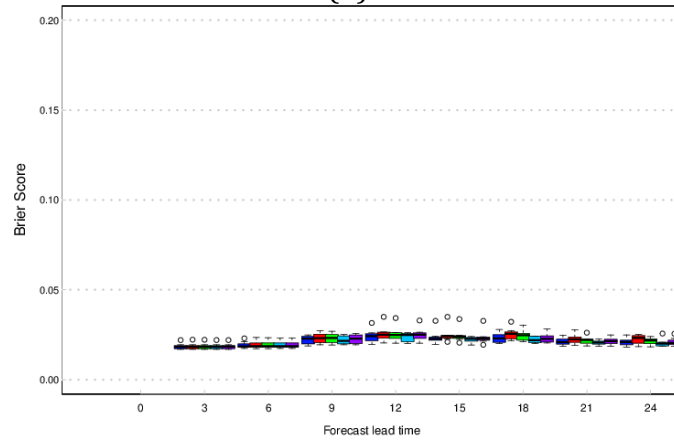




(a)



(b)



(c)

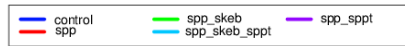


Figure 5. Brier score for 12Z initialization and for three precipitation thresholds: a) 0.254 mm, b) 2.54 mm, and c) 12.7 mm.

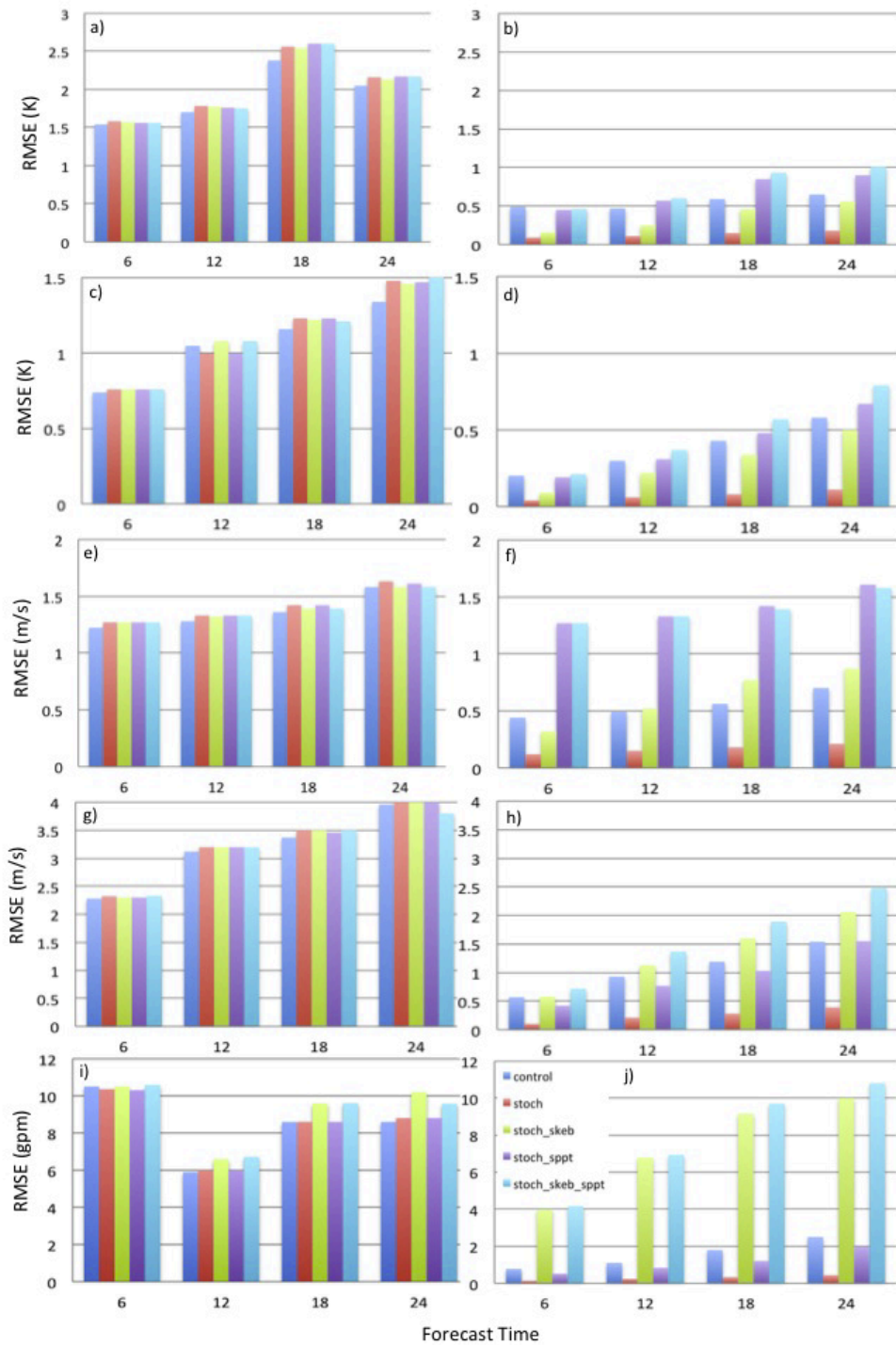


Figure 6: RMSE (left column) and corresponding spread (right column) for a) and b) 2-m temperature, c) and d) 850-hPa temperature, e) and f) 10-m zonal wind, g) and h) 250-hPa wind, i) and j) and 500-hPa geopotential height for simulations initialized at 00 UTC.

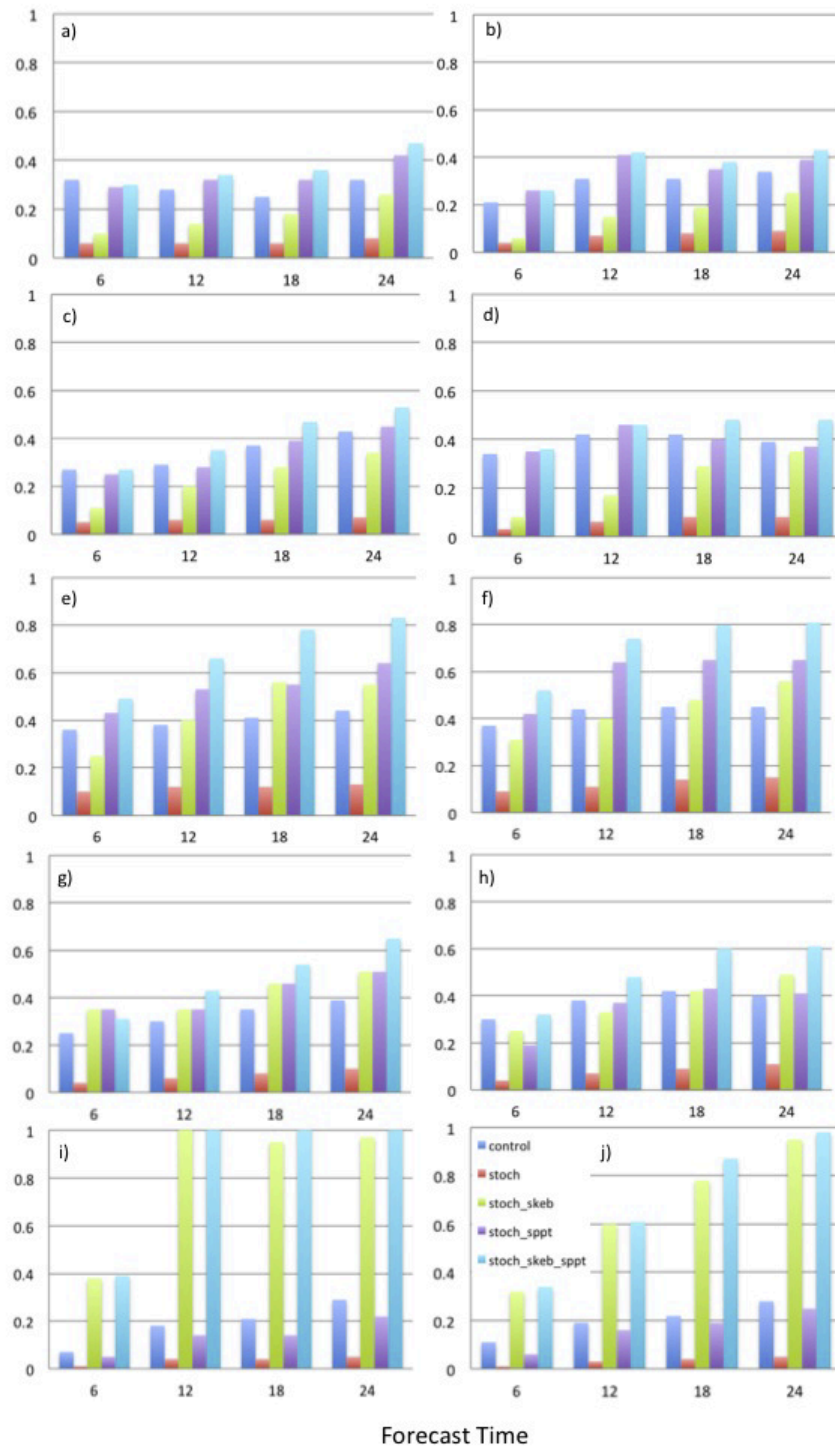


Figure 7: Spread/Error ratio for a) and b) 2-m temperature, c) and d) 850-hPa temperature, e) and f) 10-m zonal wind component, g) and h) 250-hPa zonal wind component, and i) and j) 500-hPa geopotential height, respectively for 00 UTC initialization (left column) and 12 UTC initialization (right column).

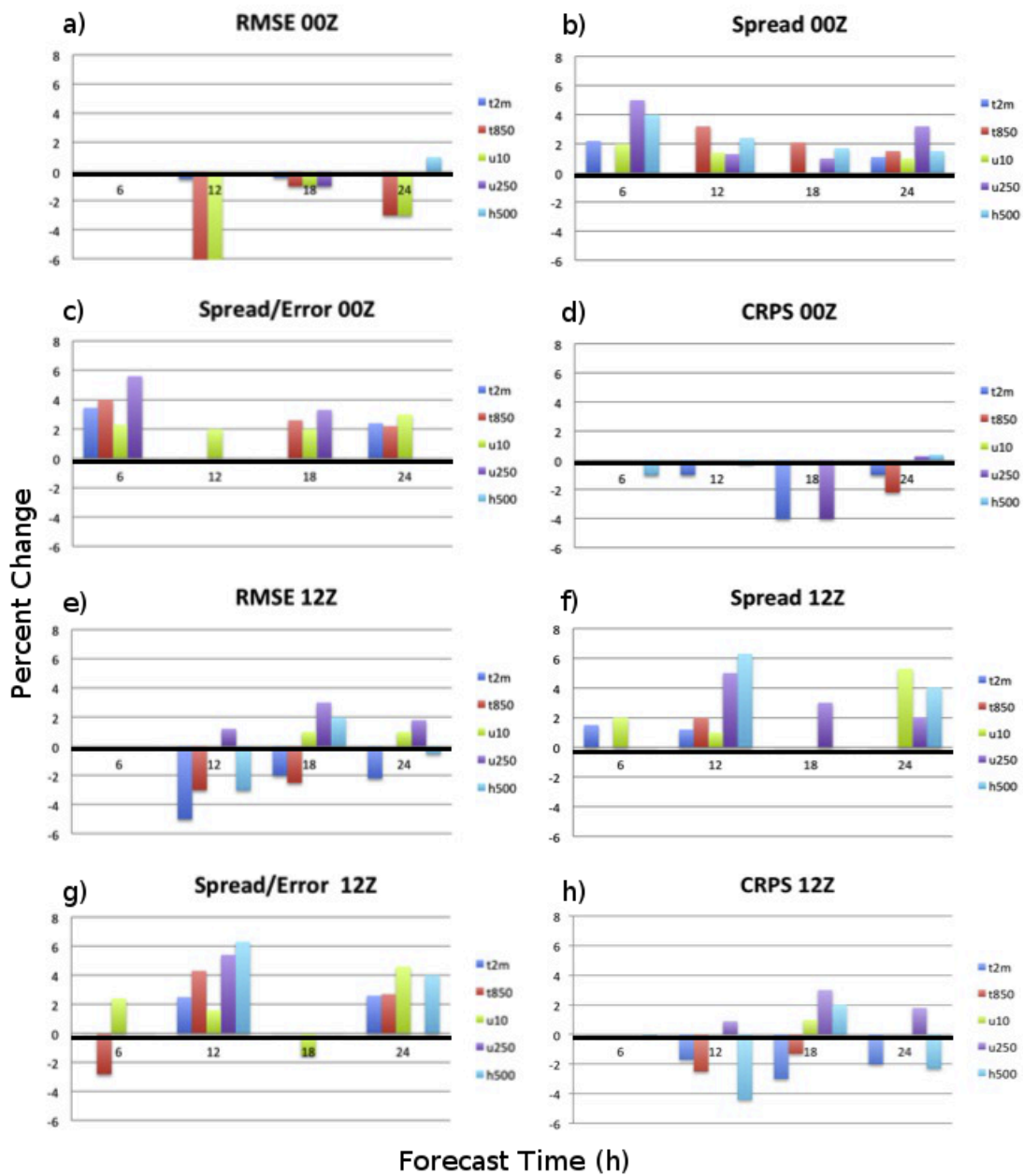


Figure 8: Percentage change of the SPP over spp sppt ensemble in a) RMSE, b) spread, c) spread/error ratio, and d) CRPS for 00 UTC initialization. Similarly, percentage change of the SPP over spp sppt ensemble in e) RMSE, f) spread, g) spread/error ratio and h) CRPS for runs initialized at 12 UTC.