IMPACT OF MODEL ERROR AND IMPERFECT INITIAL CONDITION PERTURBATIONS ON ENSEMBLE-BASED PROBABILISTIC FORECASTS: UNPREDICTABLE SPOTS

Jun Du*

SAIC at Environmental Modeling Center/NCEP/NOAA, Washington DC, U.S.A.

1 Motivation

Given existing of intrinsic uncertainties in both initial condition (IC) and model (physics and dynamics), ensemble approach is the way to possibly give a full picture of future state of atmosphere. The ultimate goal of ensemble forecasting is to reliably estimate the time-evolution of probabilistic density function (PDF) of meteorological fields. However, due to model error and imperfect IC perturbations used in real-world ensemble prediction systems (EPS), it is believed that this task is extremely difficult if not impossible. Currently, little is known about the impact of model error or imperfect IC perturbations on the evolution of ensemble-based PDF in operational numerical weather prediction (NWP) models.

In this study, using the NCEP short-range ensemble forecasting (SREF) system (Du et. al. 2004, http://wwwt.emc.ncep.noaa.gov/mmb/SREF/SREF.html), the following three issues will be discussed: (1) given a near-perfect EPS, how well PDF can be predicted? (2) how to identify "bad" PDF forecast regions [defining as *"unpredictable spots"* (see the context below)] so that a special post-processing might be applied to these regions? and (3) what is the relative importance between model error and imperfect IC perturbations over PDF evolution? Since there is no way to know "true" PDFs in real atmosphere, two "perfect model" experiments were conducted to study these issues.

2 Results

Table 1 describes the design of Experiment I. Since it's reasonable to assume that the difference between current NWP models (analyses) and real atmospheric system (state) is much bigger than that between any two "good" operational NWP models (analyses), the EPS of Eta.KF could represent a near *"perfect"* system, while the "RSM.SAS" a very *"good"* system with respect to the *"true"* system (Eta.BMJ). Given such ideal EPSs, how well PDF could be predicted?

For simplicity as well as its importance, only 12h accumulated precipitation (12h-apcp) field was investigated as an example. For both *"good"* and *"perfect"* EPSs, their general performances (domain-averaged scores), as expected, are reasonably good in all aspects including ensemble mean (Fig. 1a-d), spread (Fig. 1e) and probability distribution (Fig. 1f) verifying against the *"truth"*.

Although the domain-wise performance is reasonably good, what is spatial variation of the performance? Figure 2 shows the percentage distribution of grid points over six RPS score categories (Table 2) from the near *"perfect"* EPS. Although majority of grid points (about 90 %) have good probability forecast (category 1 and 2), about 1 % of grid points have extremely bad probability forecast (categories 4 and 5) averaged over 20 cases during August and September 2003. For the *"good"* EPS, the number of grid points having "bad" probability forecast increased to about 2 % with a small number of grid points even entering category 6! (not shown). Further study shows that over these "bad-PDF" regions, not only probability forecast is bad, but ensemble spread and mean forecasts are all bad (Fig. 3). These "bad PDF" spots can be defined as **"unpredictable spots"** since even a near-perfect EPS cannot predict it well over those spots.

One can imagine that almost any of current statistical post-processing methods (bias, spread, PDF corrections) to an EPS is based on general performance (statistics) of past forecasts, therefore, won't help to correct a future forecast over those "unpredictable spots" where calibration is really needed the most. Since "unpredictable spots" changes in location and time from case to case, **a location and time dependent post-processing method is strongly desired!**

A big question is that is it possible and how to identify those "unpredictable spots" *in prior*?. Is it possible that those "unpredictable spots" are associated with highly unpredictable regions where ensemble spread is large? Unfortunately, the answer is NO. Figure 4 shows that "unpredictable spots" are not closely correlated to ensemble spread at all. Therefore, it's very difficult if not impossible to locate those spots.

Table 3 is the design of experiment II. 9 cases from August 2003 were investigated. Results are shown in Fig. 5 (model error only) and Fig 6 (imperfect IC perturbation only). Figure 5 tells us that **as long as model error exists (always the case in real world), it is almost certain that there are some spots which cannot be predicted even given a perfect IC perturbation distribution!** Figure 6 implies that only if given a perfect model plus a very realistic IC perturbations (Eta.BMJx), good probability forecast at all locations becomes a possibility although it's still not a perfect forecast (majority enter category 2 and a small number of points still enter category 3 which is not a good but useful forecast). This result vividly illustrates that how tough the task is to predict PDF correctly based on ensemble in operational environment!

* *Corresponding author address:* Dr. Jun Du, Environmental Modeling Center/NCEP, 5200 Auth Road, Camp Springs, MD 20746, USA; e-mail <Jun.Du@noaa.gov>

References

Du, J., J. McQueen, G. DiMego, T. Black, H, Juang, E. Rogers, B. Ferrier, B. Zhou, Z. Toth and M. S, Tracton, 2004: The NOAA/NWS/NCEP short-range ensemble forecast (SREF) system: evaluation of an initial condition vs multi-model physics ensemble approach. Preprints (CD), 16th Conference on Numerical Weather Prediction, Seattle, Washington, Amer. Meteor. Soc.

EPS	membership	model/physics	IC/perturbations	representativeness
Eta.BMJ	5	Eta with BMJ convective scheme	EDAS/bred from Eta.BMJ	*"truth"*
RSM.SAS	5	RSM with SAS convective scheme	GDAS/bred from RSM.SAS	very *"good"* system
Eta.KF	5	Eta with KF convective scheme	EDAS/bred from Eta.KF	near *"perfect"* system

Table 1: Design of Experiment I: how well PDF can be predicted with a near perfect EPS?

category	1	2	3	4	5	6
RPS value	0	(0,1)	[1,2)	[2,3)	[3,4)	4
meaning	perfect	good	useful	bad	worse	worst

Table 2: Definition of RPS score category. For a probability distribution over 5 MECE categories, the perfect RPS score is 0.0 and the worst is 4.0 (completely opposite).

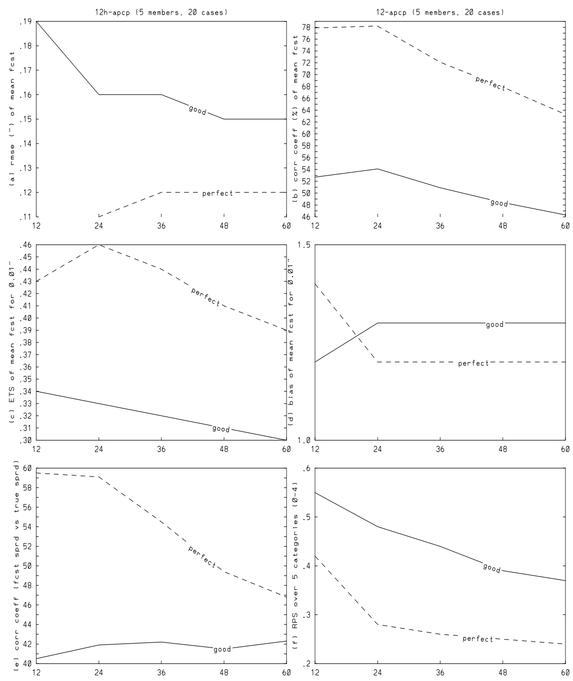

Figure 1: Scores of very *"good"* and near *"perfect"* EPS against the *"truth"*. The verification was performed over a 40km US Continental Domain (185x129=23865 grid points) averaged over 20 cases during August and September 2003. (a)-(d) for ensemble mean (rms error, corr coeff, Equitable Threat Score and bias, respectively), (e) corr coeff between fcst spread and true spread; and (f) Ranked Probability Score of probability distribution over 5 MECE categories (0.01", 0.25", 0.5", 1").

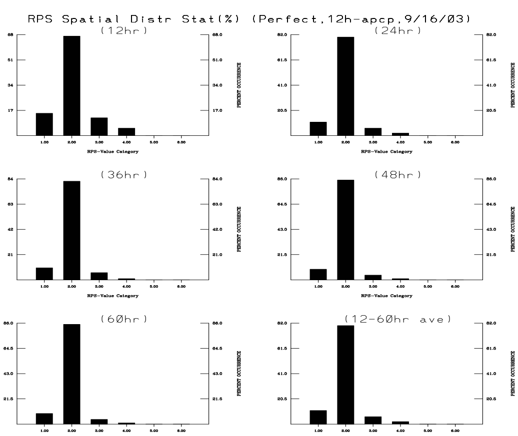

Figure 2: Percentage distribution of grid points over six RPS score categories for 12h, 24h, 36h, 48h and 60h probability forecasts as well as 12-60h average. It is from the near *perfect* EPS using 09z, Sept. 16, 2003 case as an illustration.

EPS	membership	model/physics	IC/perturbations	representativeness
Eta.BMJ	5	Eta with BMJ scheme	EDAS/bred from Eta.BMJ	truth
Eta.BMJx	5	Eta with BMJ scheme	EDAS/bred from Eta.KF	slight diff IC pert/no model error
Eta.KF	5	Eta with KF scheme	EDAS/bred from Eta.BMJ	same IC pert/slight model error

Table 3: Design of Experiment II: what is the relative importance between model error and imperfect IC perturbations over PDF evolution?

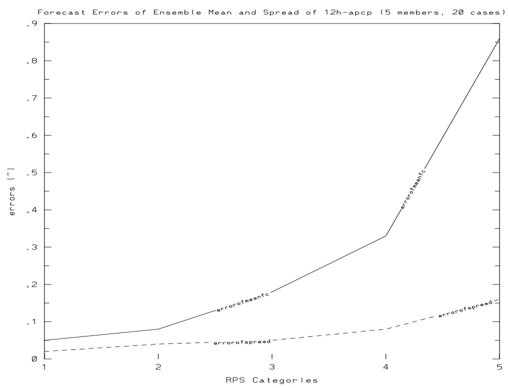

Figure 3: Increase of forecast errors with decrease of probability forecast accuracy (from category 1 to 5). Solid curve is the absolute difference of ensemble mean between *"good"* EPS and the *"true"* EPS; while the dash curve the absolute difference of ensemble spread between the two EPSs. Result is averaged over 20 cases.

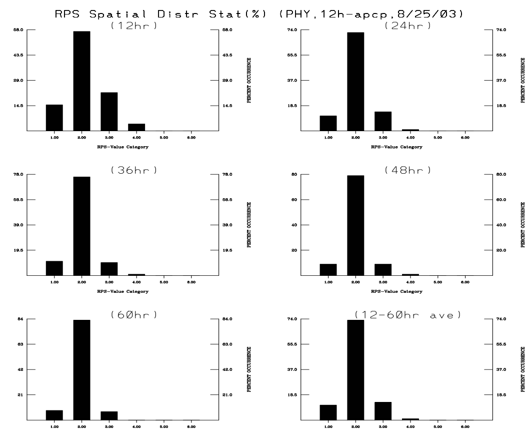

Figure 5: Same as Fig. 2 but for "perfect IC perturbation but slight model error" scenario (Eta.KF) using 21z, Aug. 25, 2003 case as an example.

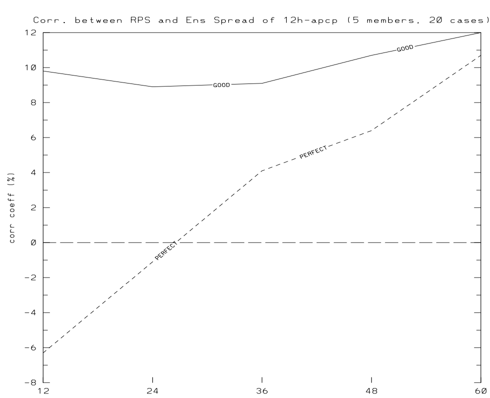

Figure 4: Spatial correlation between RPS score and ensemble spread for near *"perfect"* (dash) and *"good"* (solid) EPS, averaged over 20 cases during August and September 2003.

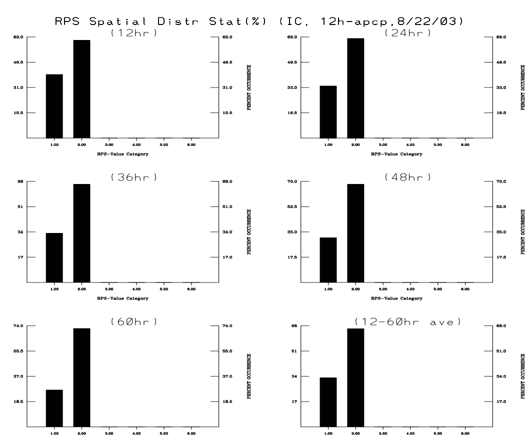

Figure 6: Same as Fig. 2 but for "perfect model but slightly imperfect IC pertuebation" scenario (Eta.BMJx) using 21z, Aug. 22, 2003 case as an example.