

Hybrid Ensemble Prediction System: a New Ensembling Approach

Jun Du*

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

1 ABSTRACT

A new ensembling concept, Hybrid Ensembling, was proposed and tested at NCEP recently. This method essentially applies the spread or uncertainty information from the coarse ensemble prediction system (EPS) to the forecast with the most “detail” and least overall error to form a more robust new ensemble – hybrid ensemble. It has solid theoretical foundation to almost guarantee an improvement over the original ensemble. The concept, test results with the NCEP ensemble systems as well as its potential impact on future ensemble research and operation was discussed. This method could be easily and widely adopted by all major operational centers to have a potential jump in the performance of their current ensemble systems.

2 Introduction

Currently, at all major operational weather prediction centers in the world such as NCEP and ECMWF, ensemble prediction systems (EPS) and high-resolution single model runs are treated as two separate forecasting systems. This imposes at least the following two dilemma to managers and scientists.

Although an EPS usually runs with a reduced resolution model, it certainly competes limited computer resources with single higher-resolution model runs. It’s quite challenging for managers to justify how much resources should be spent on each system between higher-resolution model runs and lower-resolution ensemble runs in their decision making. This competition is likely to continue in future.

Secondary, scientists and forecasters view those two forecasting systems as competitors too in regarding their utility of how well predicting day-to-day weather. Obviously, both sides won’t always win in reality since higher-resolution model runs are usually more detailed and more accurate in deterministic measurement but lack of uncertainty information, while ensemble runs provide uncertainty information but are hampered in detail and accuracy by reduced model resolution and somewhat less accurate individual members.

Therefore, it will be very ideal if one can find a solid scientific approach to incorporate those two systems into one unified forecasting system which retains both detailed, higher accuracy and uncertainty information. Under the umbrella of such a unified forecasting system, ensemble runs and high-resolution runs are no longer competing each other but benefiting and compensating each other. It is such a new approach called **Hybrid Ensembling** that is proposed and tested in this paper.

* *Corresponding author address:* Dr. Jun Du, Environmental Modeling Center/NCEP, 5200 Auth Road, Camp Springs, MD 20746, USA; e-mail: <Jun.Du@noaa.gov>. This paper is from the preprint (p4.2) of the 50th NWP Conference, June 14-17, 2004, University of Maryland, College Park, Maryland, USA.

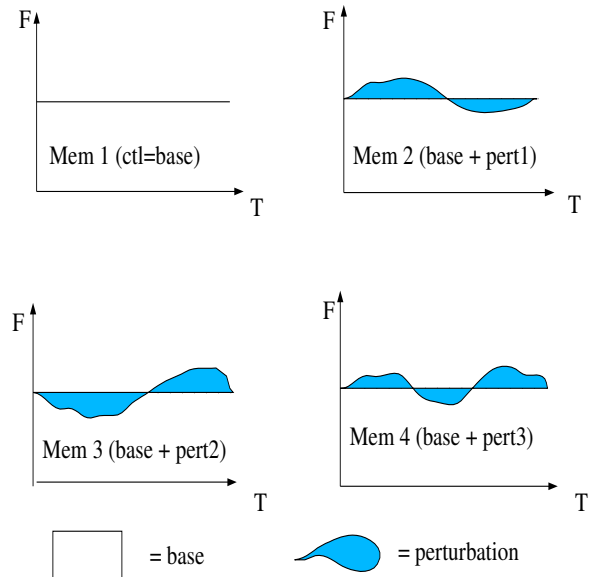


Figure 1: A demonstration from an ensemble with 4 members – control run plus 3 perturbed members: each ensemble member forecast can be decomposed into *base* and *perturbation* two components.

3 Concept

Figure 1 demonstrates that each member in an ensemble can be decomposed into two parts: *base* and *perturbation*. *i.e.*,

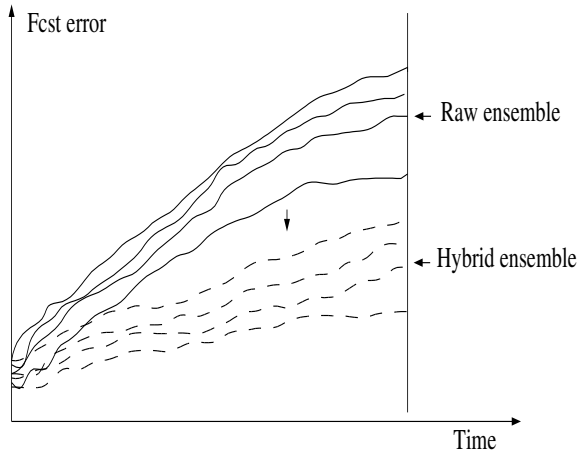
$$\text{member forecast} = \text{base} + \text{perturbation}$$

Therefore, if either part *base* or *perturbation* can be improved, this EPS will be improved as a whole.

For all current ensemble systems in the world, both *base* and *perturbation* in a member forecast come as one value (not decomposed) from a very same version of a model (same resolution, same dynamics and same physics) or are produced by a very same procedure. Since at an operational center, a high-resolution single run is usually available when EPS forecasts are available, we could use this high-resolution run which is the best deterministic forecast you can get to replace the *base* part but retain the *perturbation* part from the lower-resolution EPS to form a new **hybrid EPS**. It should be expected that the performance of this hybrid EPS will be improved steadily and even dramatically since a high-resolution model run is usually more accurate and detailed in structure than its lower-resolution ensemble runs especially for short range forecasts such as 1-3 days (Fig. 2).

For longer range forecasts such as beyond a week, a high-resolution forecast might not be superior to its lower-resolution forecast. Under such circumstance, this approach could be easily generalized by using any best single forecast available (not only limiting to a high-resolution run) as a replacement to this *base* portion.

Two approaches might be applied to the formation of a hybrid ensemble: one is by just adding the *perturbation* to



$F_{old} = B_{old} + pert$ & $F_{new} = B_{new} + pert$
 since $B_{new} > B_{old}$ (more accurate)
 hence $F_{new} > F_{old}$ (more accurate)

Figure 2: Hybrid ensemble should perform better than raw ensemble due to improved forecast direction which leads to less forecast error.

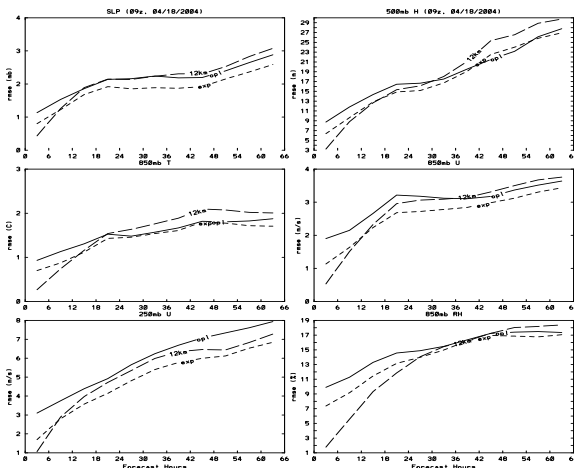


Figure 3: rms forecast errors of six selected fields for the forecast initialized at 09z, April 18, 2004. Curve labeled with “opl” is the NCEP operational SREF mean, “exp” is the hybrid SREF mean, and “12km” is the operational 12km Eta single forecast.

the *base* – “**one-side approach**” and another is by symmetrically adding and subtracting the *perturbation* to and from the *base* – “**two-side approach**”. In the “one-side approach” the high-resolution run or the best available single forecast (*i.e.* the new *base*) acts as a control member in the new hybrid ensemble, while in the “two-side approach” this new *base* acts not only as a control member but also as new ensemble mean. The advantage of “one-side approach” is that resulted ensemble mean is usually more accurate than the base (control) member, while ensemble spread and ensemble size remain the same. The advantage of “two-side approach” is that hybrid-ensemble size will be doubled and spread is increased (might be good in terms of ensemble coverage if the EPS is still under dispersive, but could hurt resolution of probabilistic forecast too), while ensemble mean forecast accuracy remains the same as the base (control) forecast.

4 Test

This hybrid EPS concept was tested with the NCEP operational SREF system (Du and Tracton 2001). The NCEP SREF system has 15 members with 3 subcomponents: 5 from 48km Eta-BMJ, 5 from 48km Eta-KF and 5 from 48km RSM. The IC perturbations are generated by the breeding method (Toth and Kalnay, 1993). For more details, readers can be referred to the following web site: <http://www.emc.ncep.noaa.gov/mmb/SREF/SREF.html>.

To form a hybrid ensemble, the *base* part was replaced by the operational 12km Eta for the 10 48km-Eta members and by the T254 GFS run for the 5 48km-RSM members, respectively, while the *perturbations* were retained from the 48km SREF system: *i.e.* differences between perturbed Eta members and Eta control member for the 10 Eta runs and differences between perturbed RSM members and RSM control member for the 5 RSM members. Thus, a new 15-member hybrid ensemble was formed using the “one-side approach” *i.e.* adding the *perturbations* to the *base*. Verification was done on the grid-212 domain covering the entire U.S. (40km, 185x129 gridpoints) and presented in the next two subsections.

4.1 Ensemble Mean

Figure 3 shows that the rms error reduction to ensemble mean forecast achieved by this hybrid ensembling approach is quite remarkable for all six selected variables in the 09z, April 18, 2004 case. The wind fields improve the most. Similar big improvements were also seen if measured by correlation coefficient (Fig. 4). Note that since in the new hybrid ensemble derived from “one-side approach”, 12km Eta run is the control member but not the center of the ensemble cloud, the new ensemble mean (center) could now easily outperform this high-resolution run. It is generally true for this case shown by both Figs. 3-4 except for the earlier time of low level moisture field. This exception might be attributed to negative impact from GFS forecast in predicting moisture field (to be confirmed).

4.2 Ensemble Spread

Since the hybrid ensemble mean is more accurate as shown in the last subsection, solutions of each member in the hybrid ensemble are presumably more correct too in general. Therefore, the chance for analysis or truth falling out of

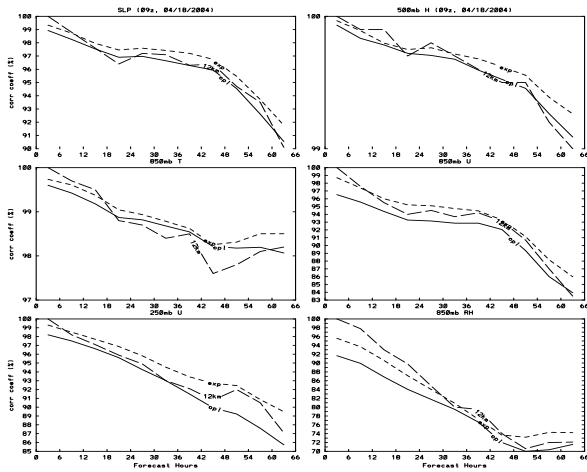


Figure 4: Same as Fig. 3 but measured by correlation coefficient.

ensemble range enveloped by all members, defined as “outlier”, should be reduced in the hybrid ensemble comparing to the raw ensemble. This has been confirmed by the results shown in Figs. 5-7. Figures. 5-6 show the ranked histograms or Talagrand distributions – where the two end-bins are outliers – for the raw ensemble and the hybrid ensemble, respectively. The sums of two end-bins or outliers from Figs. 5-6 were then plotted in Fig. 7. Note that the theoretically expected chance to be out of ensemble range due to limited ensemble size – $2 \times [1/(15+1)] \times 100 = 12.5$ – has been subtracted in Fig. 7.

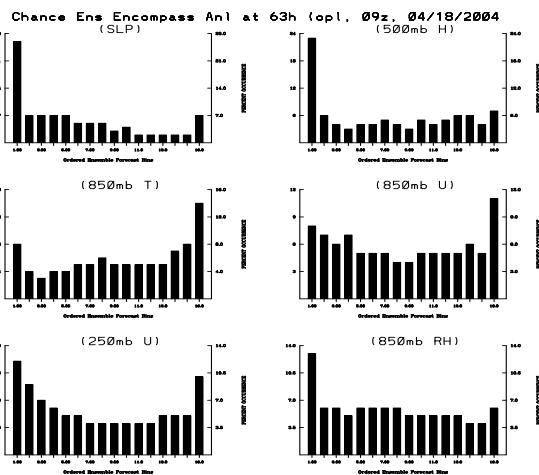


Figure 5: Ranked Histograms of 63-hr forecasts of six variables based on the raw operational NCEP SREF 15 members, initialized from 09z, April 18, 2004.

Since ensemble spread is a measure of atmospheric uncertainty level, ideally speaking ensemble spread and rms error of ensemble mean should be equal to each other. Currently, almost all EPSs in the world including the NCEP

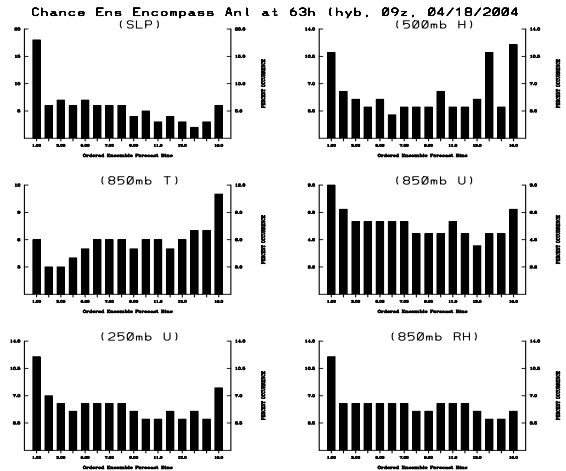


Figure 6: Same as Fig. 5 but for the hybrid ensemble. Also note that the vertical scale is different from Fig. 5.

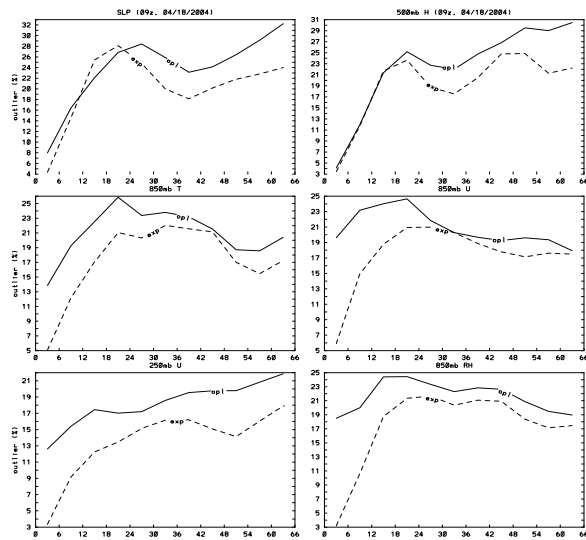


Figure 7: Outliers (percentage): sums of the two end-bins in Fig. 5 and 6 minus the expected value (see content). Solid curve (labeled “opl”) is for the raw SREF and dashed curve (labeled “exp”) is for the hybrid SREF.

SREF are under dispersive *i.e.* spread is substantially smaller than ensemble mean error (Fig. 8). Given this under-dispersive nature, the improvement in relationship between spread and error through hybrid ensembling should be obvious due to large error reduction in ensemble mean forecast (Fig. 3). Figures 8-9 do indicate that ensemble spread is much closer to rms error of ensemble mean in the hybrid ensemble (Fig. 9) than that in the raw ensemble (Fig. 8). Although majority of variables are still under dispersive, the sea-level pressure shows almost near-perfect spread-error relation in the hybrid ensemble (Fig. 9).

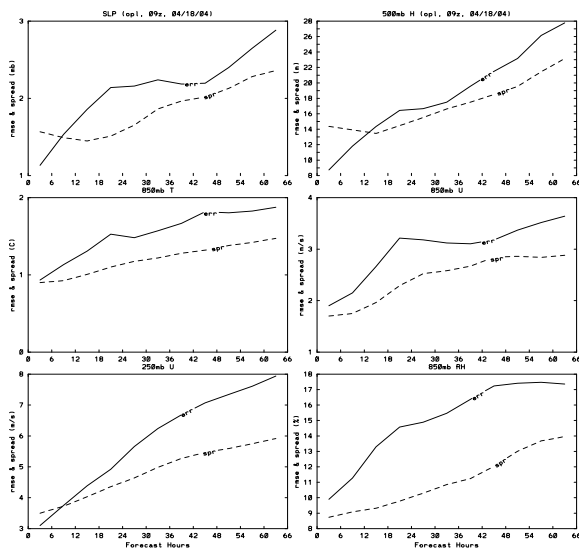


Figure 8: Spread (dash curve labeled “spr”) and rms error of ensemble mean forecast (solid curve labeled “err”) of six variables based on the raw operational NCEP SREF 15 members, initialized from 09z, April 18, 2004.

Although no verification was done for probabilistic forecasts when this paper was written, performance of probabilistic forecasts should be expected to improve too through this hybrid ensembling method given the large improvements in both forecast accuracy and ensemble spread.

5 Outlook

A new ensembling approach, Hybrid Ensembling, was proposed by combining a best single forecast (such as a high-resolution run) and a lower-resolution EPS. The uncertainty information extracted from the lower-resolution EPS is added to the best single forecast to form a new hybrid EPS. The method has solid theoretical foundation to almost guarantee an improvement over the original EPS and was tested for a case with very promising results. Therefore, this method could be easily and widely adopted by all major operational centers to have a potential jump in the performance of their current ensemble systems.

Under this hybrid EPS framework, the aim of an EPS will now be more focused *i.e.* to provide uncertainty information as reliable as possible (*e.g.* better spread-skill

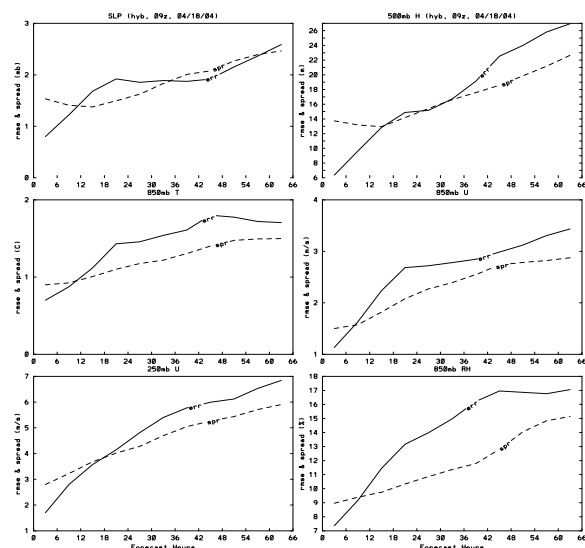


Figure 9: Same as Fig. 8 but for the hybrid SREF.

relationship or better representation of spread to real atmospheric predictability or better resolution and reliability of a probabilistic forecast etc.) rather than to improve ensemble mean performance which will automatically be taken care of by the new ensemble base from a best single forecast. Also, questions like “can an ensemble mean outperform a higher-resolution forecast” should be no longer asked in future research since a “yes” answer is certain by hybrid ensembling. This simplified goal (but might be a more difficult one) could potentially have a major impact on future ensemble research and practice: **the requirement for an EPS might be higher in size (large number of members) but less in model resolution.** This could be a very good news to mesoscale ensemble effort.

Since small-scale and storm-scale details (such as convective modes) becomes increasingly important in mesoscale prediction such as in the NCEP Storm Prediction Center’s Spring Program. Very-high resolution model will certainly play an unreplaceably important role in this regard. However, due to always limited computer resources, one cannot run all ensemble members with the highest possible model resolution. Therefore, a future EPS especially mesoscale EPS could be constructed like a Constellation in structure (Fig. 10): there are several (say 3-5) very high resolution runs (stars) to provide uncertainty information on storm scale, then many planets (lower-resolution ensemble members) go around a star to provide variations of larger-scale as well as of storm’s location and intensity. Within each star system, a sub-hybrid ensemble could be formed. Finally, by combining all those sub-hybrid ensembles together, uncertainty information of both small scale and larger-scale could be provided at the same time.

Because hybrid ensembling could double ensemble size if the “two-side approach” is used, this could potentially have a big contribution to ensembling too. However, if this “doubling size” has positive or negative impact on ensemble performance, it needs to be investigated.

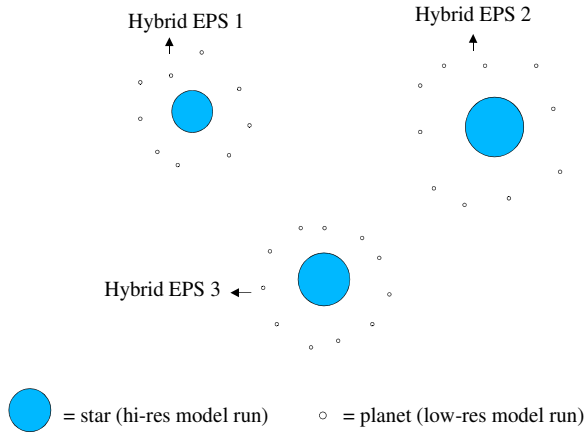


Figure 10: Constellation-like ensembling system: a couple of very high-resolution runs with various models or model versions (*i.e.* a mini-ensemble of stars) plus a large number of lower-resolution runs around its parent model (*i.e.* sub-hybrid EPSs).

Since *base* and *perturbation* come from two separate forecasting systems (such as hi-res and low-res models) in a hybrid ensemble, whether they (flow and its spread) are dynamically matched to each other such as with respect to storm locations is a natural question to ask especially when extreme nonlinearity occurs. This issue certainly needs to be investigated carefully. To avoid this, it is suggested that hybrid ensembling is better to be based on a same model to have more consistency in physics and dynamics.

References

- Du, J., and M. S. Tracton, 2001: Implementation of a real-time short-range ensemble forecasting system at NCEP: an update. Preprints, *9th Conference on Mesoscale Processes*, Ft. Lauderdale, Florida, Amer. Meteor. Soc., 355-356
- Toth, Z., and Kalnay, E., 1993: Ensemble Forecasting at the NMC: The generation of perturbations. *Bull. Amer. Meteorol. Soc.*, 74, 2317-2330.