Ensembles: A Critical Review

Zoltan Toth, Jie Feng, Jing Zhang, and Malaquias Pena

Global Systems Division, NOAA/OAR/ESRL

Soon after the emergence of numerical weather prediction, forecasts in chaotic systems like the atmosphere or the coupled Earth system were found to develop unavoidable and growing errors. Under such conditions, a single forecast can be rather misleading as its variance comprises of a priori indistinguishable components: one that both matches with, and another that differs from future reality. To mitigate the situation, the idea of ensembles, or a multitude of equally likely forecast scenarios was born. The practical implementation of the widely used concept (1) doubles the estimated error in the analysis (initial perturbations), and then (2) further compounds forecast errors with additional stochastic perturbations. The mean of the "perturbed" forecasts filters out unpredictable features that become misaligned across members, lowering the error compared with the unperturbed (control) forecast, an effect deemed desirable.

The intentionally inserted initial and forecast errors, however, irrevocably compromise predictive information: The error in all perturbed forecasts is significantly higher than that in the control. Centering the perturbed analyses around the control shields the ensemble mean from immediate degradation. Deceivingly, it also makes any single quantity of reality appear as a random draw from the ensemble. In the multidimensional space of atmospheric dynamics, however, the ensemble cloud does not encompass, but is further displaced from reality. A single dynamical prediction started from the best initial condition provides the most forecast information. Therefore we advocate for the dedication of NWP resources to the generation of such an unperturbed forecast. In turn, statistically processing can be used to filter out unpredictable noise, assess error characteristics, and provide probabilistic statements associated with the best forecast.

Unlike in forecast applications where the odds of finding a global solution superior to the control remains the same, astronomically low value, in data assimilation, more ensemble members guarantee ever enhanced information in the analyzed state. So for the application of dynamically generated ensembles in data assimilation - yes, in forecasting - no.