Western Region Technical Attachment
No. 93-33
November 23, 1993

PREDICTABILITY AND ENSEMBLE FORECASTING

Lawrence B. Dunn - WSFO Salt Lake City

Recently, Dr. Jeffrey Anderson of GFDL/Princeton University spoke at Western Region Headquarters. Jeff worked at WR/SSD (Western Region/Scientific Services Division) for a number of summers while he was a student, and is now an active researcher in the field of predictability. The goal of research in this field essentially boils down to an attempt to determine when numerical models are likely to be correct and vice versa. Forecasters may not realize it, but they are also very active in this same field since, to a great extent, they must also determine when to believe the models, and when they should not be believed.

Predictability research has tended to focus on longer term forecasts (10 days and beyond). As we all have seen, every once in a while a 6-10 day forecast verifies with remarkable accuracy, although this is generally not the case. Obviously if one were able to identify the cases in which the 6-10 forecasts were correct, there would be great economic benefit. Research into this problem has taken many different approaches. Some have tried to identify certain regimes the models handle better than others, such as blocking regimes, or regimes that fit certain teleconnection patterns such as the PNA. Others have examined whether run- to-run consistency is an indication that the models are more likely to be correct. The results are somewhat mixed.

There has not been as much predictability research on short-range forecasts. Roebber (1990) used the NGM and the Canadian regional model to examine the relationship between run-to-run consistency and model accuracy for developing cyclones along the east coast of North America. He found the degree of model consistency was not well related to model skill. This somewhat surprising result flies in the face of conventional wisdom in which forecasters generally express greater confidence when the models are consistent.

Ensemble forecasting is another area of predictability research, and one that is beginning to have an impact on operations. This is the method of averaging different model runs together to see if the average has more skill and utility than individual model runs. This was the subject of Jeff's talk. The NMC is currently generating an ensemble with 14 members for the 6-10 day forecast period. The details of how they do this are described in Tracton and Kalnay (1993). The European Center is also generating ensemble forecasts.

One of the big issues in ensemble forecasting is: 'Which model runs should be included in the ensemble?' One obvious choice would be to create an ensemble from different models that all start from the same initial time. Another choice would be to create an ensemble from a single model, but from runs that start at different initial times (and data). This is often referred to as the 'Lagged-average method'. Lagged-average ensembles are inexpensive to produce because the old model runs have already taken place. Yet another method produces an ensemble made up of members from multiple runs of a single model that have each had different perturbations introduced into their initial conditions. This last method is aimed at identifying situations in which small changes in the initial conditions will have a large impact on the model solution. Of course, the method by which the perturbations is introduced is a topic of considerable debate. The NMC is using a combination of the last two methods to produce their ensemble forecast.

It is clear from the Tracton and Kalnay article that there is interest in pushing ensemble forecasting into the 1-5 day range. There are many issues to be resolved before this takes place, including:

- Is the evidence that run-to-run consistency may not be an indicator of model skill sufficient to argue against lagged-average ensembles?

- Does it make sense to create an ensemble from different models that each have different characteristic errors, and different data-cutoff times for their initial analyses?

- Should different members of the ensemble have different relative weights assigned to them or are all models assumed to be equal?

- What is the skill of the ensemble average relative to randomly selected individual model solutions, and what is it relative to the best individual solution?

- Can forecasters use knowledge of errors in initial conditions and characteristic errors of individual models to consistently select the best individual model solution in a given situation or would they be better off using an ensemble average?

- If a forecaster now uses the above knowledge to pick the best individual model solution, and then modifies it accordingly, how might a forecaster modify an ensemble solution, or is this meaningful?

The list could be quite long, and another list could be created discussing the best way to display ensemble forecasts. Tracton and Kalnay (1993) display a few possibilities.

Dr. Anderson encouraged field forecasters to become involved in this debate, and suggested the NMC is likely to welcome forecaster participation. Those interested in this topic should contact SSD.

A few other tidbits from Jeff's talk include the following:

- The ECMWF is still beating the MRF overall.

- The ECMWF has a later data cutoff time than the MRF.

- The ECMWF gives more weight to observations and less to the 1st guess than the MRF. The result is the ECMWF tends to flip/flop more than the MRF. Thus the ECMWF may be changing more from day-to-day than the MRF, making difficult to use, but if new data does result in a significantly different solution, the ECMWF will likely show this sooner than the MRF.

- Jeff said that research shows that when faced with the choice between run-to-run consistency and going with the model from the latest data, going with the latest data generally is best. Thus, in general, he felt model-to-model consistency from the latest data was better than sticking with a given solution because it was consistent with previous runs.

References:

Roebber, P. J., 1990: Variability in successive operational model forecasts of maritime cyclogenesis. *Wea. Forecasting*, **5**, 586-595.

Tracton, M. S., and E. Kalnay, 1993: Operational ensemble prediction at the National Meteorological Center: Practical aspects. *Wea. Forecasting*, **8**, 379-398.