A MULTI-MODEL DIAGNOSTIC STUDY OF MEDIUM-RANGE FORECAST SKILL DROPOUTS

Rolf H. Langland and Ryan N. Maue, NRL, Monterey, CA

1. Introduction

A continuing problem in numerical weather prediction is the occurrence of intermittent yet significant "dropouts" in forecast skill. Forecast dropouts are seen in the time series of 5-day anomaly correlation (AC) skill scores in which individual or a sequence of forecasts deviates significantly below the monthly average. The threshold for defining a dropout varies between models due to differences in their respective overall skill.

A primary goal of this research is to identify potential sources of forecast dropouts including, but not limited to: (a) the diagnosis of model physics or dynamics deficiencies, (b) initial condition error or uncertainty in the analysis, and (c) diagnosis of atmospheric flow regimes or patterns that can be described as less predictable. In the following discussion, we identify forecast dropouts from three current state-of-the-art deterministic forecast models, and provide evidence that particular atmospheric flow regimes and transitions are closely associated with below-normal forecast skill.

2. Data and methods

The data for this analysis consists of an archive of deterministic model forecasts out to at least 5-days made in real-time during the past several years. Table 1 describes details of the ECMWF, GFS, and NOGAPS modeling systems, time period of the archived forecasts, and download source. Each respective model has improved in terms of resolution and performance from 2004-present, and that aspect of forecast skill changes is clearly shown in our analysis. It is advantageous that these respective forecast centers maintain easily accessible archives of individual model runs. While we are examining the skill of several reanalysis products that utilize a homogeneous model, those results are not presented here.

The familiar definition of anomaly correlation coefficient (ACC; Equation 1) is used to quantify the spatial correlation between forecast and observed deviations from climatology (Miyakoda et al. 1972). Containing aspects of forecast error and bias, the ACC measures pattern similarity, which leads to higher scores during periods of amplified versus zonal flow. A subjective value of ACC=0.6 is described as the limit of synoptic-skill below which the forecast is less valuable.

$$ACC=\frac{\sum_{m=1}^{M} \mathrm{f}'_{\mathrm{m}} o'_m}{\left[\sum_{m=1}^{M}\left(\mathrm{f}'_{\mathrm{m}}\right)^2 \sum_{m=1}^{M}\left(\mathrm{o}'_{\mathrm{m}}\right)^2\right]^{1/2}} \quad \text{(Eq. 1)}$$

$$\mathrm{f}'_{\mathrm{m}} = \mathrm{f}_{\mathrm{m}} - c_m$$

$$o'_m = o_m - c_m$$

The forecast (*f*) and observed (*o*) 500hPa geopotential height are typically calculated over a given region (20°-80° N for the Northern Hemisphere) and compared to climatology (*c*) at each grid point (*m*). The climatology is calculated by generating 21-day centered means at each grid point and synoptic time from the new NCEP CFS Reanalysis (CFSRR) (Saha et al. 2010). Each model forecast is verified against its own analysis. Later in this discussion, twice-daily values of the Arctic Oscillation are calculated from 1000hPa geopotential height values obtained from the CFSRR using familiar EOF techniques (Thompson and Wallace 1998).

Model	Time Period	Grid/Fields	Source
ECMWF	Oct 2006-present	T799-T1279 ~26-40 km	ECMWF TIGGE
GFS	Feb 2004-present	T382-T574 0.5° -1.0°	NOMADS/NCDC
NOGAPS	Jan 2004 -present	T239-T319 0.5°x0.5°	NRL MONTEREY

Table 1: Data sources including modeling center, time period of forecast data archive, gridded output characteristics, and source.

3. Anomaly correlation coefficient

The anomaly correlation coefficient (ACC) time series is shown in Figure 1 for the Northern Hemisphere (NH) middle-latitudes from the ECMWF model during the period October 2006 – December 2010. The twice-daily individual *5-day* forecasts (green squares) show considerable variance from the 90-day centered average (blue line) with a noticeable negative skew (longer tail towards lower skill or dropouts). For the ECMWF model, the lowest 5% of the ACC distribution corresponds to a deviation of -0.06 from the 90-day centered average; however, that is dependent upon the season. It is clear that the winter ACC is considerably higher than the summer on average partly attributed to increased scores occurring during the typically highly amplified flow of NH winter.

A similar story is seen for GFS and NOGAPS in terms of seasonal skill changes. Figure 2 shows the dominance of ECMWF over the other two models in terms of ACC during the past 4-years, but with the gap tightening. The three model's skill scores are highly correlated on this seasonal time scale, but what about shorter time scales? The ACC is a noisy metric, with fluctuations on daily and synoptic time scales.

In order to glean information about potential flow regime connections, a 7-day centered average is applied to the ECMWF, GFS, and NOGAPS ACC time-series with their respective 90-day centered or seasonal averages subtracted (Figure 3). The resulting high-correlation between the individual deterministic forecasts is demonstrated by the nearly overlapping time-series. That is to say that the models, while forecasting the same atmosphere from their own data-assimilation generated initial conditions, tend to suffer and succeed at the same time. There are certainly many cases when one model outperforms another, yet the overall correspondence on synoptic time-scales is striking.

4. Multiple model dropout case-study and discussion

One of the most remarkable forecast dropouts occurred with the 00Z October 31, 2009 forecast cycle with verification on November 5 for a the 120-hour forecast. From the familiar multi-model ACC time series (Figure 4), all models fell considerably from the monthly average. The ECMWF model ACC was 0.68, one of the worst performances during the past 5-years. In these situations, it is necessary to diagnose the location of the 5-day forecast error and the forecast trajectory itself. This entails examining initial condition uncertainty – how well the model data assimilation is capturing the true state of the atmosphere (Langland et al. 2009) – and understanding how these errors grow during the forecast.

While an exhaustive case-study is not presented here, raw error maps of 500hPa geopotential height from the ECMWF model highlight the North Eastern Pacific as the area with considerable forecast error (Figure 5). The top panel is the 120-hour forecast map showing a cut-off low pressure system located just north of Hawaii, a so-called Kona Low development. However, the analysis at the verification time (November 5 00z) has a considerably different atmospheric state with a much stronger subtropical ridge and deeper Aleutian low. The raw errors are considerable of both signs indicative of a misplacement in phase and underestimate in magnitude of the large-scale NE Pacific synoptic systems.

There is considerable interest in determining the large-scale flow regimes that favor dropouts, and the Arctic Oscillation (AO) represents one such metric of the boreal winter circulation. During periods of the positive (negative) phase AO, there is preliminary evidence from Figure 6 that ECMWF AC skill is lower (higher) than average. Also, during transitions from one phase to another typically associated with weather extreme events, the reconfiguration of the large-scale circulation on a hemispheric scale may be associated with forecast dropouts at a significantly higher rate than during quiescent periods of AO phase. The influence of atmospheric blocking and cut-off low formation or Rossby-wave breaking is strongly connected to the modulation of the baroclinic storm tracks by tropical convection, and may teleconnect to remote regions leading to circumstances favorable for enhanced forecast errors.

5. References

Langland, R. H., R. N. Maue, and C. H. Bishop, 2008: Uncertainty in atmospheric temperature analyses. *Tellus*, **60A**, 598-603.

Miyakoda, K., Hembree, G. D., Strickler, R. F., Shulman, I., 1972: Cumulative results of extended forecast experiments. 1. Model performance for winter cases. *Mon. Wea. Rev.,* **100**, 836-855.

Saha, Suranjana, et. al., 2010: The NCEP Climate Forecast System Reanalysis. *Bull. Amer. Meteor. Soc.*, In Press (DOI: 10.1175/2010BAMS3001.1).

Thompson, David W. J.; John Michael Wallace (1998). "The Arctic oscillation signature in the wintertime geopotential height and temperature fields". *Geophys. Res. Lett.*, **25** (9): 1297–1300.

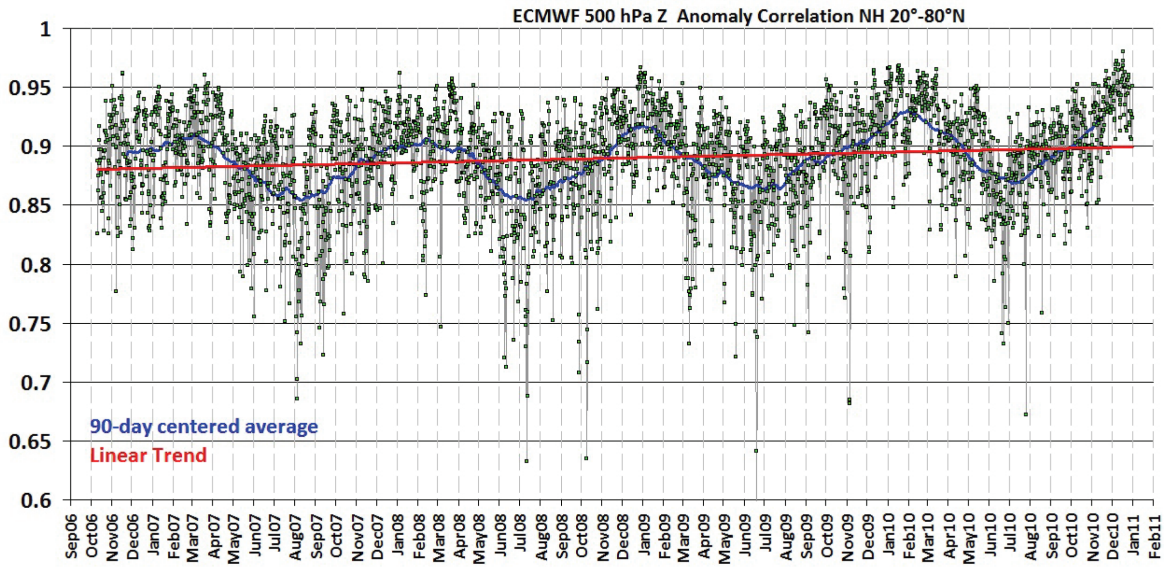

Figure 1: ECMWF Northern Hemisphere Anomaly Correlation: 500 hPa geopotential height from 20°-80°N. Individual 12-hourly forecasts are indicated by the green squares at verification time, with a 90-day centered average (blue line), and linear trend (red line).

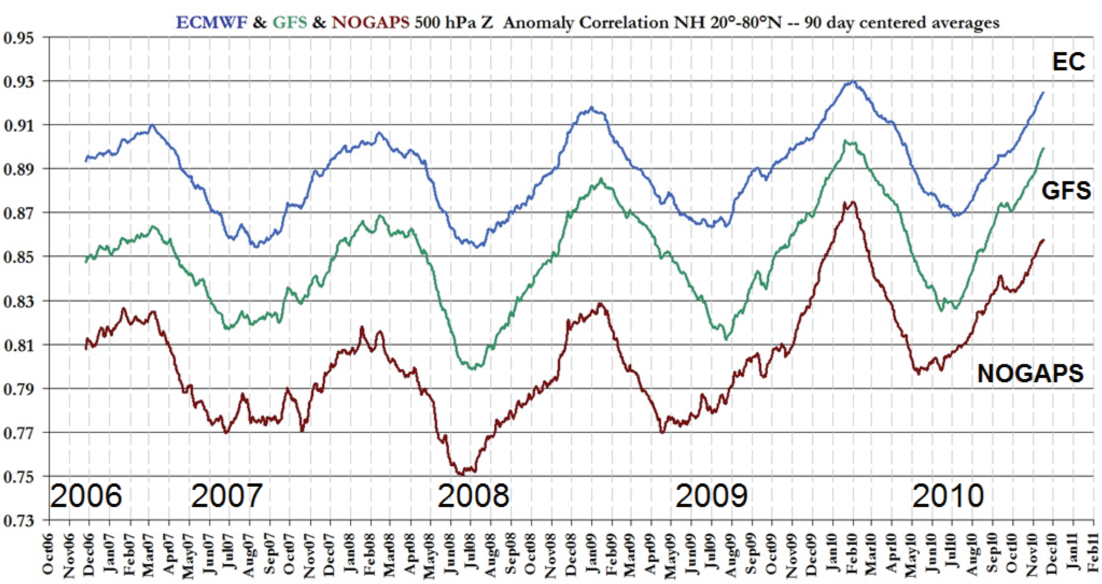

Figure 2: Anomaly correlation (AC) for the Northern Hemisphere (500-hPa geopotential height) for ECMWF, GFS, and NOGAPS deterministic forecasts. The twice-daily forecasts are verified against their own analyses, and a 90-day centered average is shown from 2006-2010.

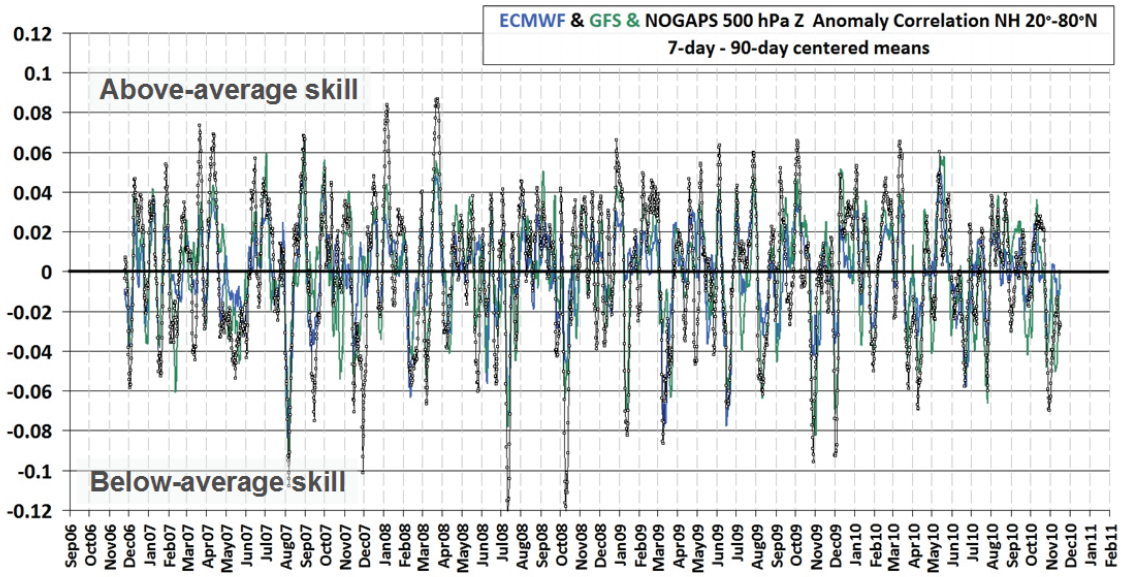

Figure 3: Same data as Figure 1, but a 7-day centered average minus the 90-day centered average in order to remove aspects of the seasonal cycle in the AC time series. Above (below) average skill is indicated by positive (negative) deviations of the 7-day average away from the 90-day average.

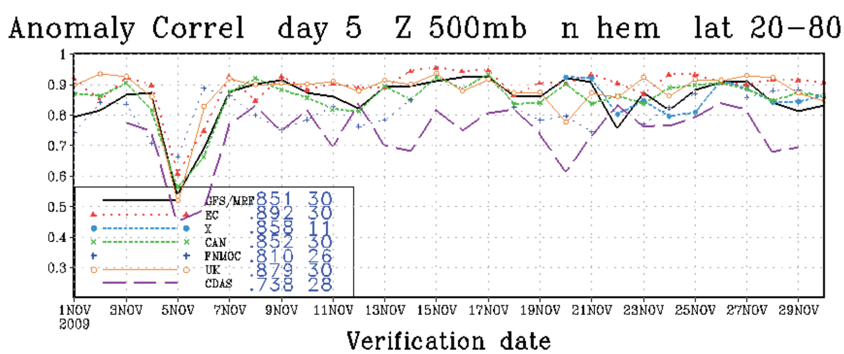

Figure 4: Multi-model anomaly correlation (500 hPa geopotential height) for the Northern Hemisphere at 00Z verification times during November 2009. Graphic source: http:// www.emc.ncep.noaa.gov/gmb/STATS/html/monarch.html

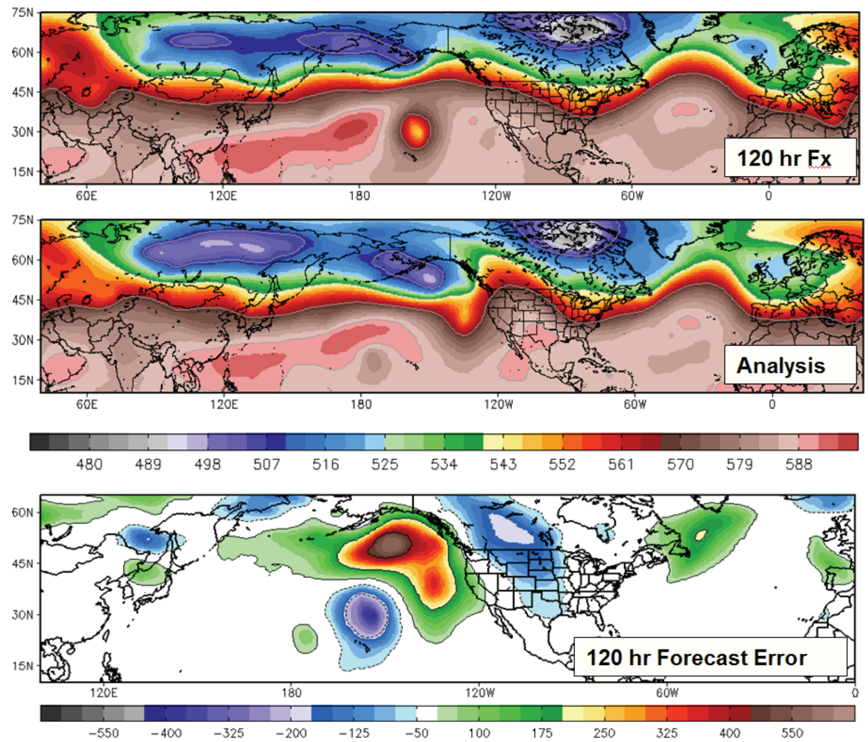

Figure 5: (top) ECMWF120-hour 500-hPa geopotential height (gpm) forecast verifying at 00Z on November 5, 2009, (middle) ECMWF analysis, and (bottom) the forecast error: forecast field minus analysis.

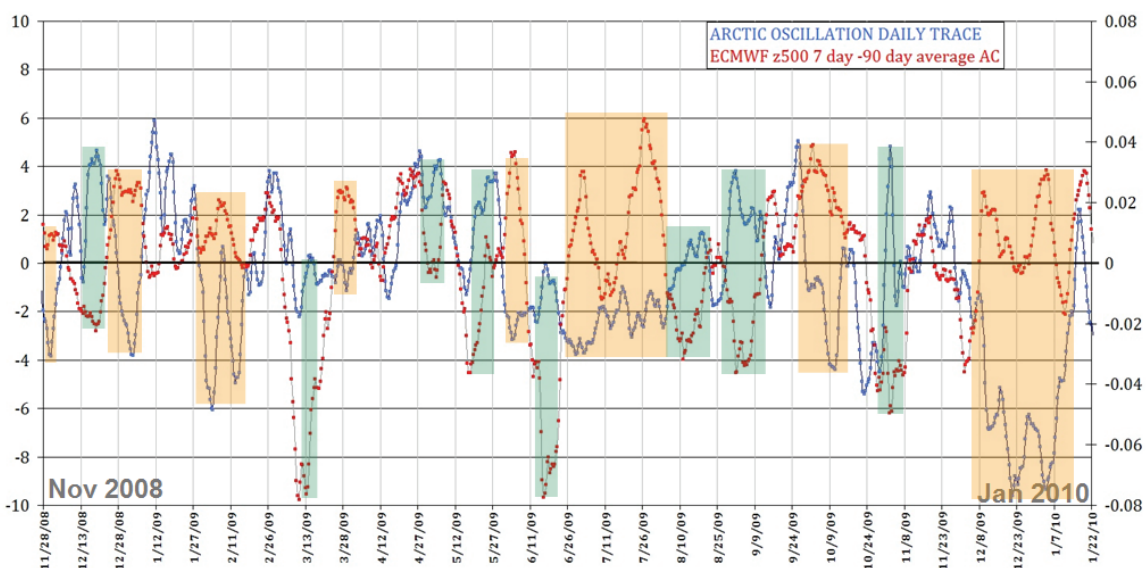

Figure 6: Daily time-series traces of the Arctic Oscillation (blue line, not normalized) and the ECMWF NH anomaly correlation (red line). Shaded tan (blue) regimes are indicative of periods of higher (lower) skill. The NH AC time series is a result of the differencing of two averaging procedures: 7 and 90 day centered averages.