Application of Method for Object-Based Diagnostic Evaluation to GFS, GFSX, and ECMWF Precipitation Verification

Tracey Dorian, Fanglin Yang
(Email: tracey.dorian@noaa.gov)
IMSG - NOAA/NCEP/EMC

Introduction:

Forecast verification is a key step to improving model performance. Traditional verification measures offer forecast quality assessment scores, but do not provide diagnostic information about why forecast skill was high or low and traditional verification approaches penalize forecasts twice for missing observed precipitation and for giving a false alarm (Gilleland 2009). Additionally, measures-based verification measures "tend to favor smoother forecast fields of coarser-resolution models" and therefore do not provide useful information on the benefits of high resolution models (Wolff 2014). Object-based precipitation verification, unlike traditional skill scores, offers spatial information about how close the precipitation forecasts are to the observations in location, size, orientation, and intensity. The Method for Object-Based Diagnostic Evaluation (MODE), developed by the National Center for Atmospheric Research (NCAR) Developmental Testbed Center (DTC), is an example of a verification tool that provides statistics about the spatial differences between forecast objects and observation objects. Precipitation objects are defined in both forecasts and observations based on two parameters, the accumulation threshold and a smoothing radius (Davis 2006). Examples of diagnostic information that can be obtained from MODE include centroid distance, angle difference, area ratio, percentile intensity ratio, and more.

Methods:

MODE precipitation verification was applied to three global numerical weather prediction models – the ECMWF, GFS, and GFSX (Parallel GFS) – to the 2014 seasons. Multiple thresholds combined with a fixed smoothing radius were used to compare model forecasts to observations. The 00Z forecasted 24-hour precipitation accumulations ending on the 36, 60, 84, 108, 132, 156, and 180 forecast lead times from the models were compared to Climatologically Calibrated Precipitation Analysis (CCPA) observations. The Median of the Maximum Interest with respect to observation objects (MMIO) is an output summary statistic generated for each MODE run. Larger values of MMIO indicate higher interest values which imply better matches across forecast and observation fields.

Results:

For all four seasons in 2014, the ECMWF generally has the largest MMIO values during most forecast hours. The Parallel GFS has higher MMIO values compared to the T574 GFS. In the summer, the MMIO values are lowest compared to other seasons, and there seems to be a larger drop in MMIO values with forecast lead time. Also in the summer, the MMIO values drop quicker as threshold increases compared to winter, spring, and fall. All three models consistently overestimate the extreme light rain intensities, or the 10^{th} and 25^{th} percentile intensities, averaged over all forecast hours. The median light rain intensities in the GFSX, however, do appear to be closest to observations during all four seasons. For median intensities, or 50^{th} percentile intensities, all three models do quite well throughout the entire year. The GFSX is farthest from observations in summer where it underestimates the intensities. For the extreme high intensities, or the 75th and 90th percentile intensities, the three global models generally underestimate intensities and are farthest from observations in late winter/early spring and in the summer. The GFSX intensities, once again, are usually closest to observations in 2014. The variation in model bias from day to day is much larger for the 75th and 90th percentile intensities compared to the lower percentile intensities. Similarly, the variation in median percentile intensities between forecast lead times is larger for heavy rain. The GFSX is closest to observations in total object count, especially in summer where the GFS and ECMWF largely underestimate the number of objects - the same is true for fall 2014. For winter and spring of 2014, the GFSX overestimates total object count. The GFS appears closest to observations in total object count in the spring. Finally, for all seasons in 2014 the European model and GFS median object areas are larger than observations during most forecast lead times. The Parallel GFS is closest to observations for the median object area, but also tends to forecast larger object areas compared to observations during all seasons.

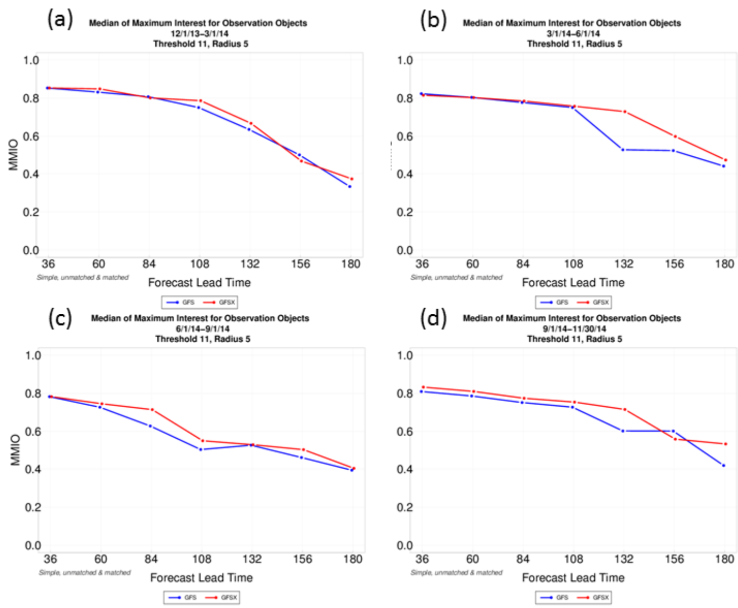

Figure 1: The Median of Maximum Interest for observation objects versus forecast lead time for accumulation threshold 11mm and convolution radius of 5 grid squares for (a) winter 2013-2014, (b) spring 2014, (c) summer 2014, and (d) fall 2014. T575 GFS is in blue, GFSX T1534 is in red. Forecasts are for 24-h accumulated precipitations ending on x-axis forecast lead times.

References:

Davis, C. et al. (2006): Object-Based Verification of Precipitation Forecasts. Part I: Methodology and Application to Mesoscale Rain Areas. *Monthly Weather Review.*,134, 1772-1784. doi: http://dx.doi.org/10.1175/MWR3145.1

Gilleland, E. et al. (2009): Intercomparison of Spatial Forecast Verification Methods. *Weather and Forecasting.*, 24., 1416-1430. doi: http://dx.doi.org/10.1175/2009WAF2222269.1

Wolff, J.K. et al. (2014): Beyond the Basics: Evaluating Model-Based Precipitation Forecasts Using Traditional, Spatial, and Object-Based Methods., *Weather and Forecasting.*, 29., 1451-1472. doi: http://dx.doi.org/10.1175/WAF-D-13-00135.1