1. INTRODUCTION

A number of research studies indicate that the recent history of convection contains a great deal of information on its future state out to 6 hours and beyond (e.g., Golding 2000). The scale of the convective feature often determines the length of time that it will persist and thus be predictable through extrapolation alone (Wilson et al. 1998). Additional work has been done in developing heuristic models of convection that can be used to predict the Lagrangian evolution (growth and decay) of convection (Hand 1996; Pierce et al. 2000; Megenchardt et al. 2004). The National Convective Weather Forecast (NCWF-2) combines extrapolation and a heuristic approach to produce 0-2 hr probabilistic forecasts of convection that are available in real-time for use by the aviation community. Mueller et al. (this conference) describe how this system has been extended out to 6 hours and discuss its performance. While these observation-based techniques perform well in the very short term (e.g., 0-3 hr), their skill decreases rapidly with lead time. On the other hand, despite efforts to assimilate a myriad of data streams (including radar), the skill of numerical weather prediction models continues to be lacking at lead times less than 3 hr. To mitigate poor skill of RUC in predicting convective precipitation, Weygandt and Benjamin (2004) developed a probabilistic approach that takes into account uncertainty inherent in the modeling system's ability to forecast the exact timing and location of convection. The new product called the RUC Convective Probability Forecast (RCPF) is being tested for operational use by the AWC.

The goal of this study is optimally blend these RCPF with 0-6 hr NCWF convective probability forecasts described by Muller et al. (this conference) to take advantage of the lead-time dependent relative skill of each method. Data from the spring and summer of 2005 are used to develop a summary of verification statistics for RUC-based and NCWF 0-6 hr probabilistic forecasts including CSI, POD, FAR, and Bias as a function of lead time. These performance parameters are then used to develop a weighting function that varies with lead time and time of day. The RUC-based and NCWF 0-6 hr probabilistic forecasts are then blended using this weighting function to produce a single merged probabilistic forecast. The new merged forecasts are then evaluated using a few unique cases from the spring/summer of 2005 (see J. Wilson's talk) with results being intercompared with the individual components of the system. Physical explanations for deficiencies and triumphs of the new system (to be called NCWF-6) are given.

2. COMPARISON OF MODEL AND OBS-BASED TECHNIQUES

Both NCWF and the RCPF forecast products are given as probabilities. The probability that convection will occur at a given point is determined using a spatial filter. The spatial filter differs in these two systems with NCWF using an elliptical filter following Wolfson et al. (1998) and RCPF using a square filter (Weygandt and Benjamin 2004). In both systems, a somewhat arbitrary threshold is used to determine the areas affected by convection strong enough to impact aviation. In NCWF the thresholded forecast variable is an extrapolated interest field (see Megenchardt et al. 2004) similar to VIL (a derived-product available from the WSR-88D radar network; Klazura and Imy 1993) while in the RCPF, the threshold forecast variable is rainfall rate. In the RCPF, a threshold of 1 or 2 mm hr-1 is used depending on time of day and longitude. The probability that convection will be present at a given location and time is determined by dividing the number of gridpoints where the threshold is exceeded within the filter area by the total number of gridpoints encompassed by the filter.

In NCWF the filter size is a function of lead time while in the RCPF system the filter size is fixed at 180 km. The NCWF filter size increases from 60 km at a leadtime of 1 hr to 180 km at a leadtime of 3 hrs, following Germann and Zawadski.
(2004). The filter size remains fixed between 3 and 6 h. Since the RUC has a grid mesh of 20 km, the filter encompasses a total of 81 grid points. While in the NCWF system, the filter is run on a 4 km grid, so that even the 60 km filter spans a much greater number of grid points (225) yielding better statistical information (e.g., distributions of convective intensity).

The probability forecasts from NCWF and RCPF for a single late spring afternoon (4 June 2005) are compared to illustrate their relative strengths and weaknesses (Figure 1). This case included a squall line initiation event that occurred in the Southern Great Plains and widely scattered cellular convection associated with large-scale instability characteristic of the SE US throughout much of the year. Storm cells initiated along a dryline/cold front around 2130 UTC and formed into a squall line by 0000 UTC. A second related area of convection to the north over northeast KS initiated around 1930 UTC and remained nearly stationary while growing in size. The widely-scatter storms in the southeast characterize the typical diurnal cycle of convection in areas of large-scale instability as is commonly present in this region of the country.

There are a number of features that standout in comparing RCPF and NCWF forecasts of convection (Figure 1). The RCPF tends to overestimate spatial coverages of convection resulting in very large biases, particularly in regions of large-scale instability such as that associated with the flow around the Bermuda high into the SE (not as evident on this day). The NCWF skill rapidly decreases with lead time (compare 2 and 4 h forecasts) because of convection initiation, dissipation, and difficulties associated with obtaining a motion vector for accelerating/turning storms. Even at lead times of only two hours, there is often initiation that cannot be forecasted by NCWF (as seen in Figure 1). In the case shown in Figure 1, the NCWF forecast misses the entire squall line at a lead time of 4 hours.

The strengths of the RCPF include its ability to predict regions and time-of initiation at lead times up to 8 hours and its ability to depict the nature (linear vs complex) of storms in the vicinity of synoptic-scale boundaries. The strengths of the NCWF are its ability to extrapolate existing storms - particularly larger convective elements that have lifetimes > 1 hr, its treatment of growth and dissipation at leadtimes < 2 hr (see Mueller talk), and the reduced number of false alarms particularly at shorter lead times due to the use of smaller filter sizes than is used in the RCPF and the need for convection to have been present in the past (i.e., incorporating persistence). Note that a longer lead times, dissipation and errors in storm motion result in increasing FAR and lowering CSI.

Both the RCPF and NCWF have significant weaknesses, however, these weaknesses can be overcome to some degree by combining the systems. Thus, it seems that a new operational algorithm could be developed that optimally combines these two disparate forecasts to produce 0-6 hour forecasts that are more accurate than either individual system could produce alone.

Figure 1. Probabilistic forecasts from the RCPF (gray shades) and NCWF (pink contours) systems during a convection initiation event that occurred over S. Great Plains on 4 June 2005 for forecast lead times of (a) 2 and (b) 4 hours issued at 2200 and 2000 UTC, respectively. The NCWF data are contoured at the 10% probability level. The RCPF probabilities are contoured at 25, 50 and 75%. WSI radar reflectivity data at forecast time (red contours depicting radar echoes > 35 dBZ) and valid time (reflectivity maps showing dBZ > 25). The valid time is 2400 UTC. “G”, “I”, “A” and “D” denote regions of storm growth, initiation, advection and dissipation, respectively.
3. METHODOLOGY FOR BLENDING

The new 0-6 hr NCWF product and RCPF product may be combined using a weighted average similar to the method of Golding (2000). Since the two different probability forecasts represent different scales and different rainfall intensities, the forecasts must be calibrated before they can be merged together. Calibration is performed by validating both the new NCWF and RCPF probabilistic forecast fields at a range of probability levels for each lead time. The validation is performed using coverage maps calculated from the observed radar reflectivity (thresholded at 35 dBZ) using a filter size that increases with time.

The temporal variability in statistical measures (e.g., Conditional Success Index (CSI = f(POD,FAR)) and bias) is evident in Figure 2. The verification scores are for the 24-hour period beginning 1500 UTC on 4 Jun 2005 for the RCPF and NCWF. This encompasses the time period discussed in detail in the previous section. Note the diurnal cycle in bias and CSI in both forecast systems. The CSI scores tend to be greater in NCWF forecasts for lead times of up to 4 hrs on this day. The RCPF CSI scores are at a maximum during the late afternoon between 00 and 03 UTC and at a minimum during the earlier morning hours (denoting problems with dissipation). The NCWF scores reach a maximum during the early evening hours (03-06 UT). This peak skill level has a lag with lead time (i.e., peak occurs later in the day for longer lead times). At leadtimes of 4 to 6 hours the NCWF forecast skill goes to zero between 20 and 01 UT owing to the period of initiation and growth which are not treated beyond leadtimes of 2 hrs.

The RUC CSI scores are only a weak function of probability threshold (i.e., the skill of 75% likelihood contour is just as good as the 25% likelihood contour in predicting 25% coverage). The scores are a much stronger function of the chosen probability level in the NCWF system. This difference gives an indication that NCWF forecasted probability levels more closely match
reality and can be used more reliably than those generated by the RCPF on this day.

Day-to-day variability in these results is large and depends on the synoptic regime and character/organization of the convection (not shown). This will be explored in more detail in the talk.

The statistics are evaluated for each set of forecasts to determine the probability level at which the CSI score is maximized and the bias is closest to unity (note RUC biases are much greater than NCWF – this needs to be accounted for somehow) in both methods at each lead time ranging from 1 to 6 hours. Then the NCWF and RCPF forecasts are combined using the probability levels that produced the best forecasts with weighting based on relative forecast skill. The weighting function may be determined dynamically using real time verification statistics (if possible given CPU constraints) or in a static mode based on climatological skill (more economical) – see talk.

4. REFERENCES


