

## RUC MODEL-BASED CONVECTIVE PROBABILITY FORECASTS

Stephen S. Weygandt and Stanley G. Benjamin

NOAA Research - Forecast Systems Laboratory, Boulder, CO

### 1. INTRODUCTION

The low predictability of warm season convective storms remains one of the outstanding challenges to the safety and efficiency of aviation travel. In particular, two unique aspects of thunderstorms combine to make them perhaps the most significant weather-related challenge confronting the aviation community: 1) the rapidity with which they develop and 2) the extreme hazard they pose to aircraft. In contrast to large winter storms, which have a reasonable degree of predictability at a 1-2 day lead time, inherent uncertainties in convective forecasting severely limit the usefulness of explicit thunderstorm predictions beyond about 2 h.

In spite of these thunderstorm prediction difficulties, the increasing utilization of the National Air Space has led to a growing need for longer lead-time (2-6 h) thunderstorm likelihood information. This information is needed as guidance to aviation meteorologists and traffic flow managers as they work together to make strategic aircraft routing decisions to optimize air traffic relative to developing thunderstorm clusters. In addition to the obvious safety benefits from improved long lead-time thunderstorm likelihood guidance, significant improvements in air travel efficiency are likely from such information.

To address this need, NOAA Forecast Systems Laboratory has developed a convective probability forecast product based on the Rapid Update Cycle (RUC) model (Benjamin et al 2004a,b). Known as the RUC Convective Probability Forecast (RCPF), this product was first tested in a real-time mode during the summer of 2003, with verification of 2-, 4-, and 6-h forecasts performed within the Real-Time Verification System (RTVS, Mahoney et al 2002). D2 status within the Aviation Weather Technology Transfer (AWTT, Knapp et al. 2002) process

---

\* *Corresponding author address:* Stephen S. Weygandt, NOAA/FSL, R/FS1, 325 Broadway, Boulder, CO 80305, Stephen.Weygandt@noaa.gov

was obtained during the spring of 2004. A series of improvements have since been made to the product and a statistical comparison of the 2003 and 2004 versions is ongoing.

In this paper, we first discuss the rationale for a probabilistic forecast of convection, then describe the techniques employed to create this forecast from the RUC model output. Next, results from the summer 2003 season are summarized, followed by a description of the 2004 accomplishments. We conclude with a discussion of the plans for further development and operational implementation of this product.

### 2. PROBABILISTIC CONVECTIVE FORECASTS

The inherent low predictability of atmospheric convection severely limits the accuracy of deterministic thunderstorm forecasts. For any weather phenomenon, the time limit for predictability in a numerical model (defined as the point at which the forecast becomes practically useless because model errors overwhelm the true solution) is related to the life cycle of the modeled phenomena. Because convective systems are composed of individual convective updrafts, which have a life cycle on the order of an hour, their predictability limit is much shorter than that for winter storms, which have a life cycle of several days.

From an aviation standpoint, lines and clusters of thunderstorms pose a more significant hazard than individual convective elements, because they are less easily avoided and can restrict large areas of otherwise congested air-space. Fortunately, many years of operational model evaluation have shown that prediction skill for these larger convective systems decreases more gradually with time than for individual convective elements, and numerical models retain at least some degree of skill for many hours. At long lead-time, numerical models typically capture large-scale characteristics of convection, but have significant phase errors in specific convective details. As such, these predictions are best used to provide guidance on general areas of thunderstorm

likelihood, as opposed specific point forecasts of thunderstorm occurrence. Over the past several years, attempts to utilize long lead-time deterministic model thunderstorm forecasts have met with rather limited success.

An alternative to deterministic model forecasts of convection is to utilize various model fields as inputs to an ensemble approach that yields probabilistic thunderstorm likelihood information. We have utilized a simple form of this ensemble approach in the creation of the RCPF. Inherent in this approach is a coarsening of the deterministic model inputs to a larger-scale more indicative of actual thunderstorm prediction skill. Determination of the optimal coarsening (indicated by a characteristic length-scale) as a function of forecast lead-time and other factors is a significant challenge in the generation of optimal probabilistic forecasts.

It is important to emphasize that the creation of convective probability forecasts from deterministic model output should be viewed as complimentary to the task of improving deterministic thunderstorm predictions. While ensemble techniques may yield probabilistic forecasts that provide the most useful information to the user, it must be remembered that the quality of these probability forecasts depends critically on the skill of the underlying deterministic forecasts.

### 3. METHODOLOGY

The research hypothesis guiding this probabilistic thunderstorm forecasting effort is that at long lead-time (> 2h) a forecast of convection at a specific point from a single deterministic model run is less likely to be correct from a statistical standpoint than a well-chosen ensemble of model predictors. Thus, at long lead-time the best use of high-resolution model forecasts is to form an ensemble for the purpose of generating probabilistic forecasts of thunderstorm likelihood. Such an ensemble may be created in a number of ways, including: 1) multi-model ensembles, 2) initial and boundary condition ensembles, 3) model physics ensembles, 4) time-lagged ensembles, and 5) grid-point ensembles. Our initial research approach (tested in real-time during summer 2003) has been to aggregate the model convective information to larger time and space scales through the use of a gridpoint ensemble approach.

The construction of a gridpoint ensemble is very straightforward as illustrated in Fig. 1. Given a distribution of model predicted convective

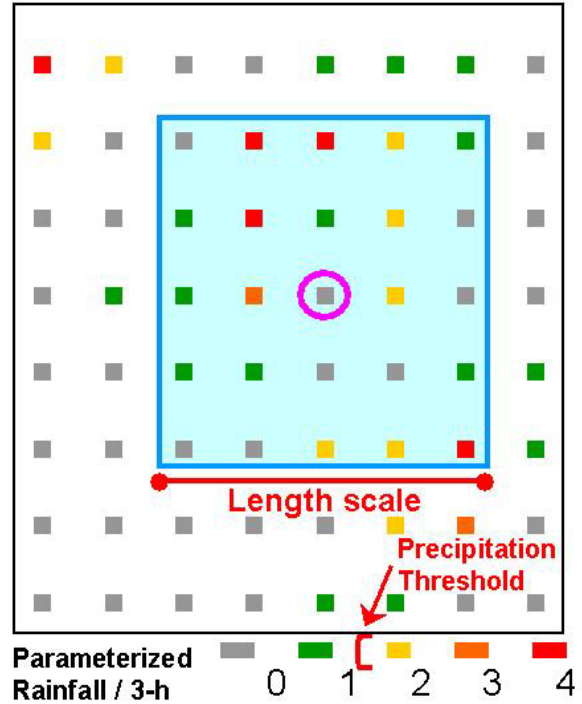


Fig. 1. Schematic diagram depicting the application of the gridpoint ensemble technique to a hypothetical model convective precipitation field. The technique requires specification of two parameters, a length-scale and a convective precipitation threshold, as indicated in red (5 gridpoints and 1 mm in this example). For a given model gridpoint (indicated in lavender) the probability of convection is calculated as the fraction of gridpoints within the box that have convective precipitation exceeding the threshold. For this hypothetical case, the probability is calculated as 10/25 or 40%.

(parameterized) precipitation values, one must specify two parameters: 1) a characteristic length scale (box size) and 2) a model convective precipitation rate threshold indicative of aviation constraining convection. Using these two parameters, it is easy to calculate the probability of convective precipitation in excess of the threshold (a proxy for the probability of aviation constraining convection) within a given distance of each model gridpoint. Note that this probability of convection within a given distance is equivalent to a fractional coverage of convection within the box. Application of this simple algorithm over an entire model domain leads to a probability of convection field as shown in Fig. 2. These plots are from a 5-h RUC forecast of afternoon convection from a challenging summer 2003 case that resulted in significant air-traffic issues.

The present gridpoint ensemble algorithm is highly dependent on the choice of the two

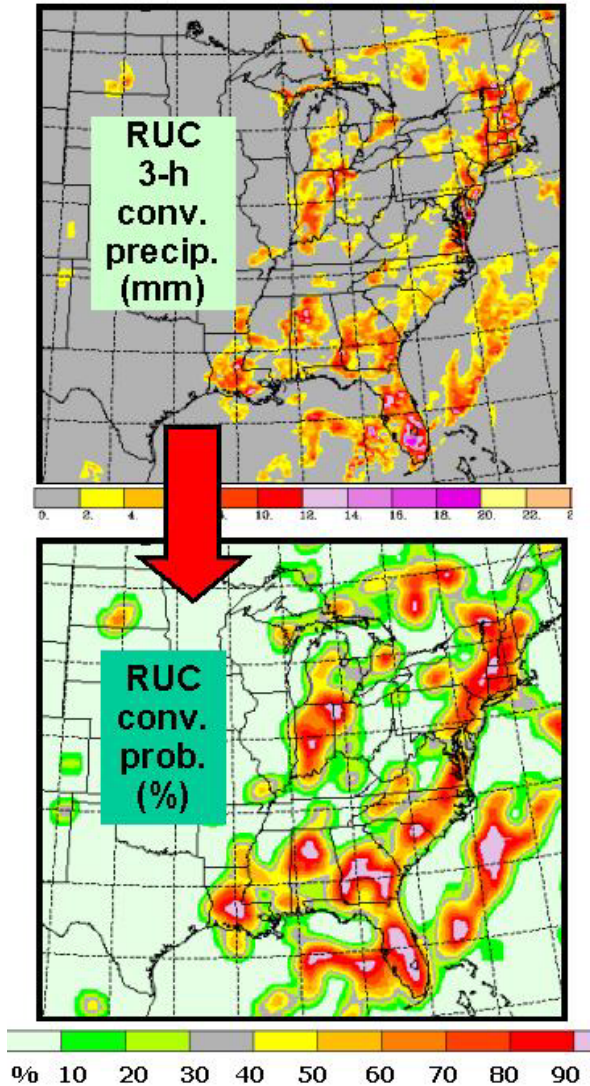


Fig. 2. Graphical depiction of creation of the RUC convective probability forecast (RCPF) from the model convective precipitation field for a sample 5-h forecast valid 1900 UTC, 4 Aug. 2003. a) Raw convective (sub-gridscale) 3-h accumulated precipitation (mm). b) Convective probability forecast obtained via the gridpoint ensemble, using a box size of 7 points and a 3-h convective precipitation threshold of 2 mm.

parameters (convective precipitation threshold and length scale). We have tested a number of threshold parameter combinations to find the optimal values. Fig. 3 shows the resultant convective probability fields for four combinations of these two parameters. Clearly, increasing the length-scale parameter spreads the probabilities over a larger area and decreases the value of the maxima. Increasing the convective precipitation threshold parameter decreases the probabilities and eliminates small areas of low probability. To

facilitate the optimization of the two parameters, a verification procedure has been implemented whereby the RCPF is thresholded at specific probability levels to create categorical (yes/no) forecasts of convection. These forecasts are then verified against the National Convective Weather Diagnostic product (NCWD) defined on a 40-km grid. The NCWD (Mueller et al. 1998) is a product developed at NCAR that depicts ongoing convection by combining national radar mosaic information with lightning data. By verifying a suite of categorical forecasts produced using different probability thresholds, the impact of variations in the two parameters (precipitation threshold and length scale) on the quality of the resultant probability forecast can be examined in a number of ways. One simple approach is to compare standard categorical skill scores for different parameter combinations over a range of probability thresholds. Fig. 4 shows an example of such an exercise, a sample plot showing the critical success index (CSI) as a function of the bias for the four threshold/length-scale combinations depicted in Fig. 3. For each threshold/length-scale combination, the individual points defining the respective curve are obtained by computing the CSI and bias for categorical forecasts created by thresholding probability forecast at a series of probabilities (5% to 95% at 5% intervals). As expected, the CSI is maximized for a bias value somewhat in excess of 1.0 for all experiments. Overall, the curves are fairly similar; however, differences occur in the important bias range of 0.75 to 1.5. In that range, increasing the length-scale and the precipitation threshold clearly leads to forecast improvement.

Next we examine ability of the gridpoint ensemble technique to add skill over a forecast based directly on the model's deterministic convective precipitation forecast. This can be evaluated by comparing relative operating characteristic (ROC) curves created by thresholding probabilistic and deterministic forecasts at a variety of values. Fig. 5 shows the results of this assessment, sample ROC curves relating the probability of detection (POD) to the probability of false detection (POFD) over a series of probability and precipitation thresholds. Comparison of the ROC curves indicates a superiority for the probability forecast over the deterministic forecast over a wide range of POD values.

An inherent limitation of the gridpoint ensemble approach is that the non-zero probability can only extend out from the edge of the model predicted convective precipitation a distance equal to the length-scale. Moreover, this approach is

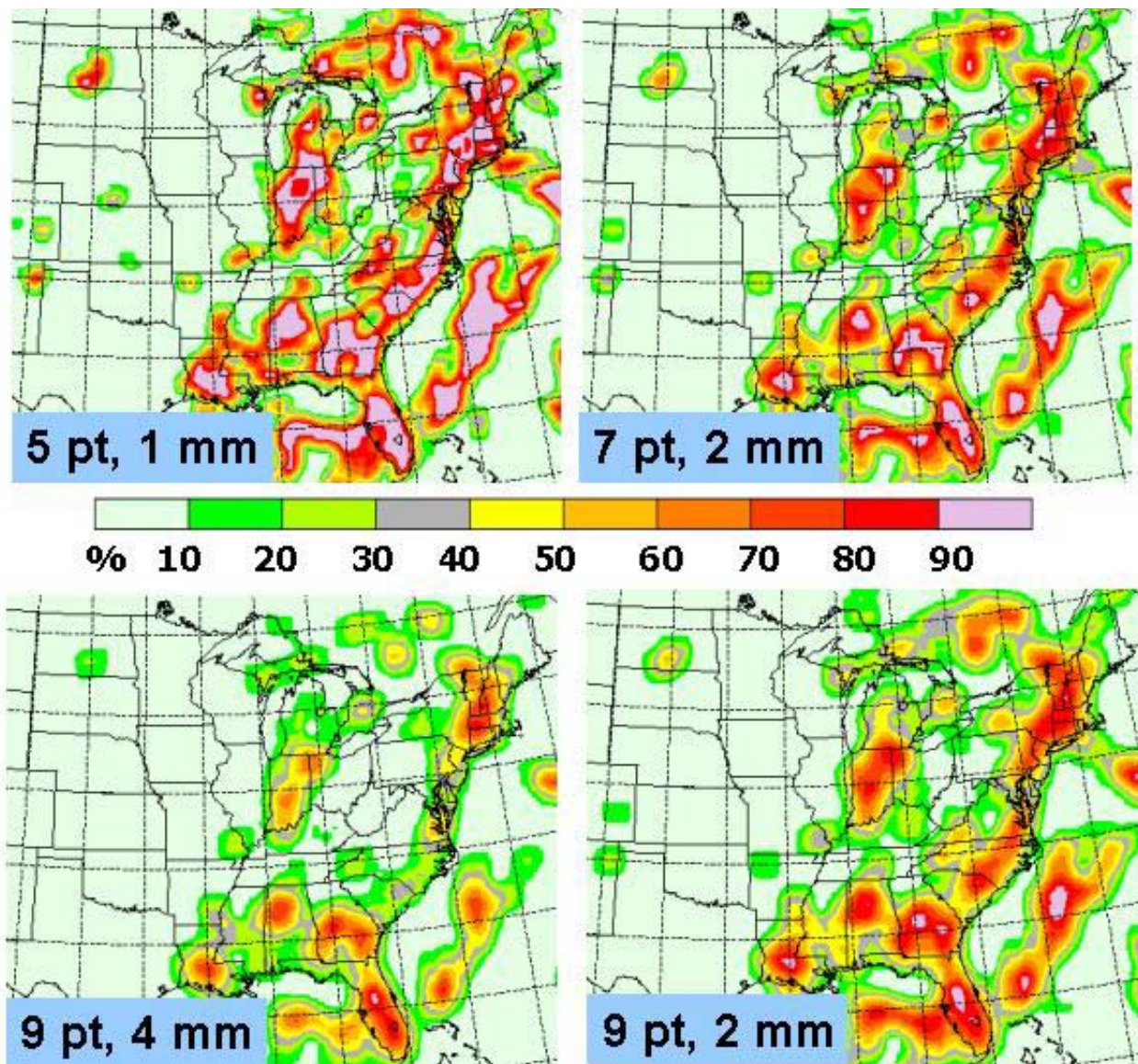


Fig. 3. Comparison of convective probability forecasts obtained via the gridpoint ensemble technique using four different box-size, precipitation threshold parameters. a) 5 points, 1 mm; b) 7 points, 2 mm; c) 9 points, 2 mm; d) 9 points, 4mm. All data from 5-h forecast valid 1900 UTC, 4 Aug. 2003.

constrained by the details of a single model run. To address these shortcomings, additional ensemble information is being added as the product matures. This includes information from time-lagged ensembles, multiple-model ensembles, and ensemble closure information from the cumulus parameterization. In addition to the basic convective probability field, a number of complementary fields are available or planned. The first among these is the RUC potential convective cloud top field (sample shown in Fig. 6). This field is determined by the level of convective overshoot, as computed from the model sounding. Tests by by

MIT/Lincoln Laboratory have shown that this field quite accurately predicts the radar-observed 18 dBZ echo top. Additional derived products that are envisioned as the RCPF is more fully developed include 1) a storm-motion field, 2) a probability change field, and 3) and an areal coverage fields.

#### 4. 2003 REALTIME TEST RESULTS

During the summer of 2003 a preliminary version of the RCPF was tested in real-time. Consistent with our vision of the RCPF as a

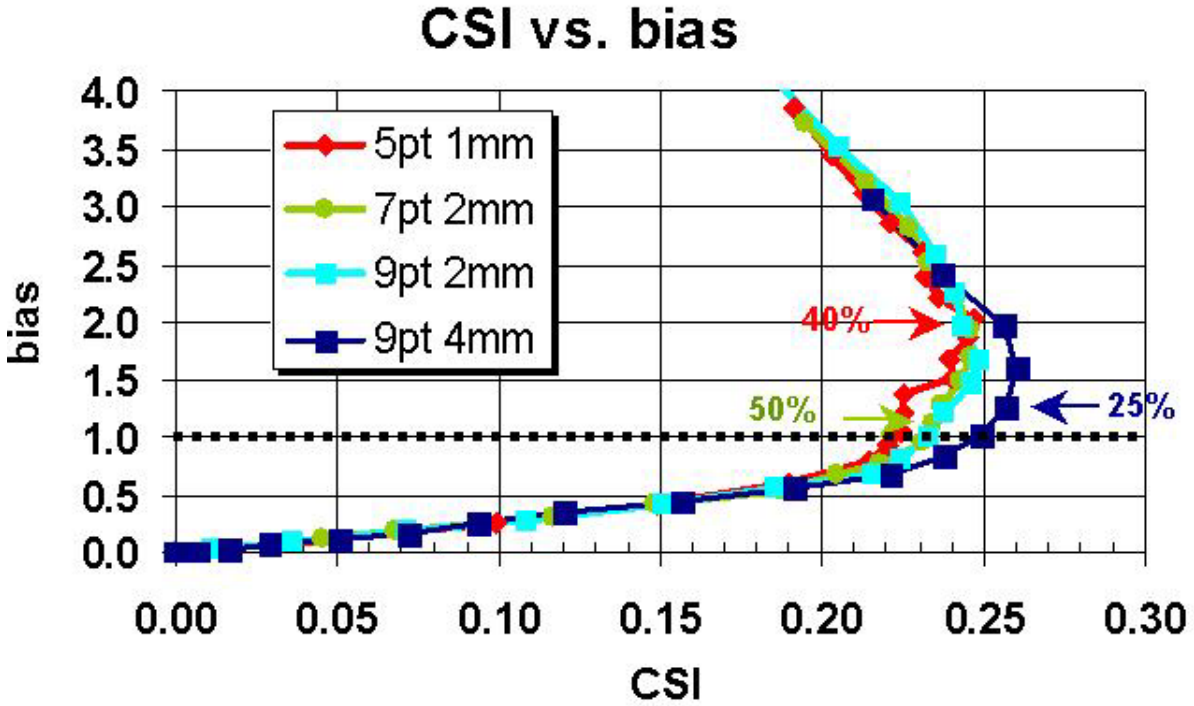


Fig. 4. Critical Success Index (CSI) as a function of bias (obtained by thresholding the RUC Probabilistic Convective Forecast, RCPF, at a series of probability thresholds) for the four sets of parameter combinations depicted in Fig. 3

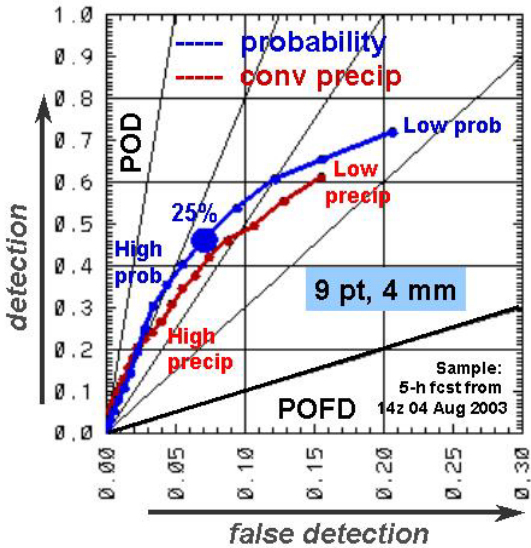


Fig. 5. Relative Operating Characteristic (ROC) curve comparison of the RUC Probabilistic Convective Forecast (RCPF) calculated with the 9 point, 4 mm threshold versus the model-predicted convective precipitation for a 5-h forecast valid 1900 UTC, 4 Aug. 2003. RUC curves obtained by thresholding the fields are various values (% for RCPF, mm for convective precipitation). Position of the RCPF curve above and to the left of the convective precipitation curve illustrates the superiority of the probability forecast.

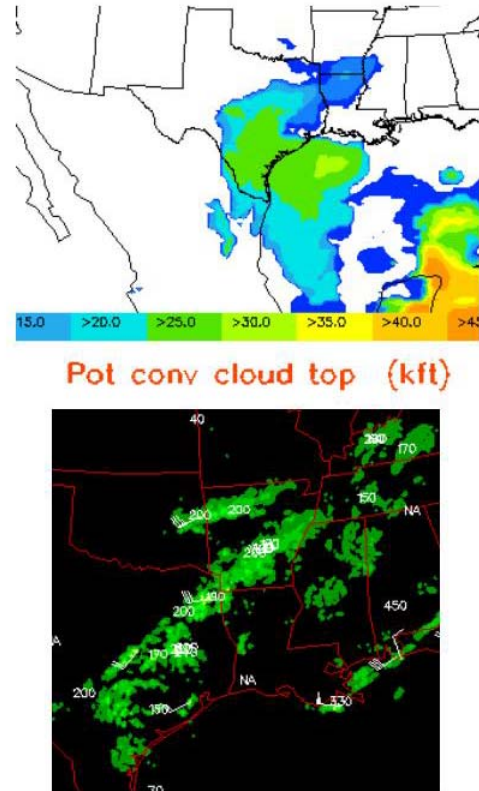


Fig. 6. a) Sample RUC potential convective storm-top and b) corresponding radar-derived echo tops from 1900 UTC Feb 9, 2004.

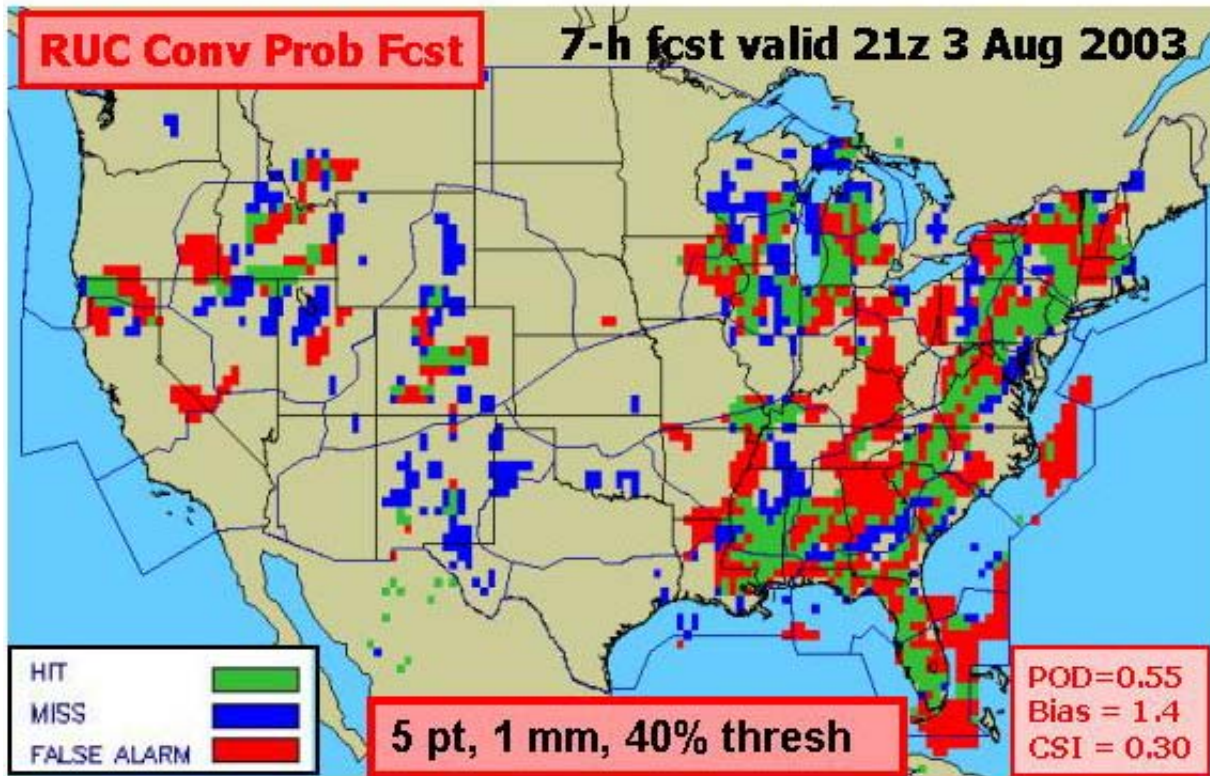


Fig. 7. Annotated sample RCPF verification from the Real-Time Verification System (RTVS) for 7-h fcst valid 2100 UTC 3 August 2003.

guidance tool for human-generated products such as the Collaborative Convective Forecast Product (CCFP, Sims et al. 2004), 2-, 4-, and 6-h forecasts were generated every 2-h with a product latency of about 45 min. The 2003 formulation utilized a 5-point spatial filter and 1 mm per 3-h accumulated convective precipitation fields threshold applied to model output from a backup version of the RUC run at FSL. This version of the RUC is run in a hardened quasi-operational setup yielding very high reliability, but the version lacks some of the latest improvements found in developmental versions of the RUC run at FSL. Quantitative verification of the RCPFs was completed within the RTVS and results were displayed via a web interface.

Fig. 7 shows an annotated sample product verification display from the RTVS web-page. As described above, a categorical forecast was derived by thresholding the RUC convective probability forecast, in this case at 40%. Within the display, green pixels indicate places where the RUC categorical product correctly forecast ongoing convection (as verified against the 40-km NCWD product). Blue pixels indicate places where the RUC product failed to predict convection that actually occurred, while red pixels indicate places

where the RUC erroneously predicted convection. Across the eastern half of the country, the RUC product predicts most regions of convection as reflected in the favorable POD and CSI scores. One area for concern is the high degree of small-scale detail in the RCPF, which may not represent the scale of actual forecast skill and may make the product more difficult to use for strategic guidance. This small-scale detail is a result of the small filter (5 point) used during the 2003 season.

Analysis of the RTVS verification statistics for a two-month period (July, August) from 2003 indicates that the RCPF provides useful strategic thunderstorm likelihood information. The average diurnal cycle of three standard skill scores (POD, bias, and CSI) is shown in Fig. 8 for 6-h RCPFs. Readily apparent in the plot is the enhanced skill for early morning forecasts valid in the mid-afternoon. This morning forecast of initial convective development is extremely important, as it can provide the basis for strategic re-routes of trans-continental plane flights, around which other shorter duration flights are planned. It should be noted that these scores are from the 2003 product, which utilized the 5 point / 1 mm threshold

## 2003 Version Categorical Verification for RCPF

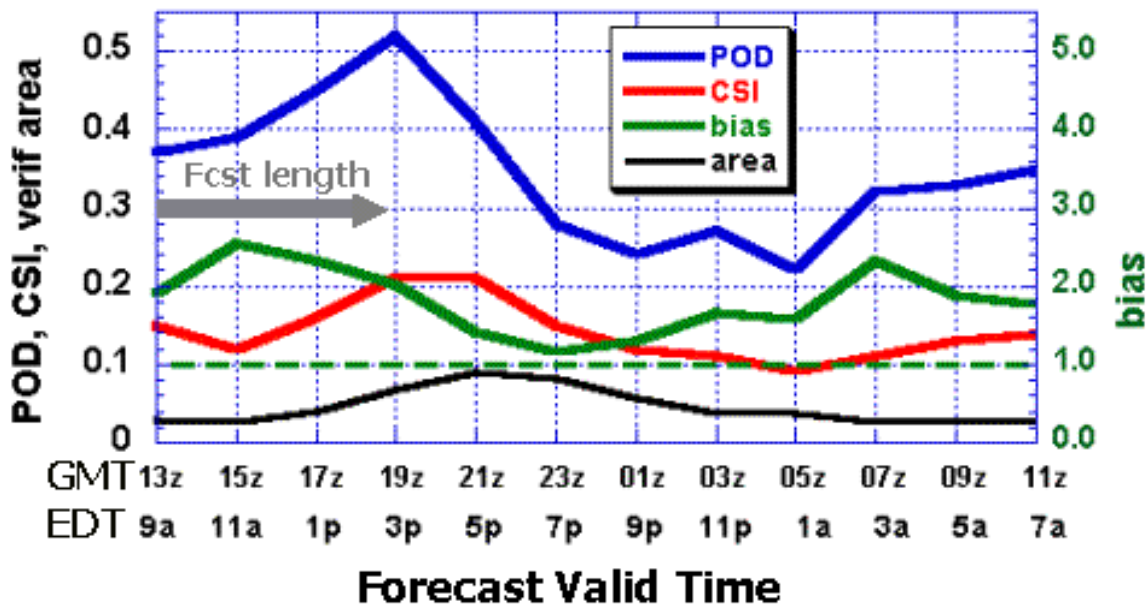


Fig. 8. Diurnal cycle of average skill scores (from July-Aug, 2003 period) for 6-h RUC convective probability forecasts (RCPFs). Shown are the probability of detection (POD), critical success index (CSI) and bias (bias), as well as the areal coverage of the verifying convection (area). The dashed line green line (bias = 1.0) is included for reference.

combination. As shown in Fig. 4, improved skill can be expected from the use a larger spatial filter. One shortcoming with the 2003 RCPF, as indicated by Fig. 8, is the rather high bias. During the late morning and early afternoon hours average bias values exceeded two.

Comparisons of the RCPF with the CCFP have been examined for the purposes of improving the RCPF and evaluating the suitability of the RCPF as a potential guidance product to be used by CCFP forecasters. It is important to note that CCFP forecast polygons must satisfy a number of constraints including storm top, areal coverage, forecaster confidence, and polygon size requirements. As such, direct comparison of the CCFP with the RCPF, a gridded convective probability forecast, must be interpreted with caution. Nevertheless, comparisons from the 2003 season (not shown) have indicated a general consistency between CCFP and RCPF products and highlighted some strengths of the RCPF and areas for improvement. Consistent with Fig 8., the RCPF is most useful for morning long range (6-h)

predictions of initial convective development. In contrast to the CCFP, 2-h RCPFs sometimes do not show significant improvement over 6-h forecasts valid at the same time. This is especially true for forecast of ongoing convection, and is consistent with known deficiencies in initializing active convective in numerical models. Current development work on radar reflectivity and radial velocity assimilation offers the potential for improvements in this area. A final issue noted for improvement in the 2003 RCPF was in the area of temporal consistency between subsequent forecasts. Improvements in this area can be expected from the use of time-lagged ensembles.

### 5. 2004 REALTIME TEST RESULTS

Based on the results from the initial real-time test of the RCPF conducted during the summer of 2003, a number of areas for improvement were noted. These included 1) reducing the high bias while maximizing skill scores, and 2) improving spatial coherency and temporal coherency. Following an extensive testing period during the

spring and early summer of 2004, a number of significant changes were made to the RCPF. These include: 1) use of a larger spatial filter, 2) use of time-lagged ensembles created from multiple output times and multiple RUC forecast cycles, 3) use of a diurnally varying precipitation threshold, 4) a special treatment of convection over the Western U.S., and 5) a 1-h reduction in the product latency (requiring the forecast to be an hour older). The larger filter combined with the time-lagged ensemble significantly improved spatial coherency and temporal consistency and greatly improved robustness by eliminating many false alarms. Because convection clusters over the Western U.S. tend to be smaller-scale and produce less precipitation than further east, one consequence of the larger spatial filter was a great reduction in the probability of convection across that region. To remedy this shortcoming, a smaller spatial filter and lower precipitation threshold were used west of 104° longitude.

Bias of the categorical forecast depends, of course, on the selected probability threshold and as shown in Fig. 8 exhibited a large diurnal variability in 2003. Through use of the larger spatial filter and

time-lagged ensemble, combined with a diurnally varying precipitation threshold, we have been able to achieve the goals established for 2004 (improved bias, improved spatial coherency, improved temporal consistency, reduced product latency) while maintaining a similar degree of skill as measured by the CSI. Fig. 9 shows a comparison of average CSI and bias scores for a one-week period (6-12 Aug. 2004) for the 2003 and 2004 versions of the RCPF. As can be seen, the excessive bias produced by the 2003 product is significantly reduced in the 2004 product, while the CSI scores are nearly the same. This is especially significant because the model input for the 2004 product is an hour older (longer forecast) to facilitate the hour reduction in the product latency.

As depicted in Fig. 10, a sample product comparison between 2003 and 2004, the 2004 product has a better spatial coherency and avoids some of the false alarms associated with the 2003 product. Fig. 11 illustrates the ability of the RCPF to accurately predict the development of new convection, in this case an active band of storms stretched across a very important air traffic corridor.

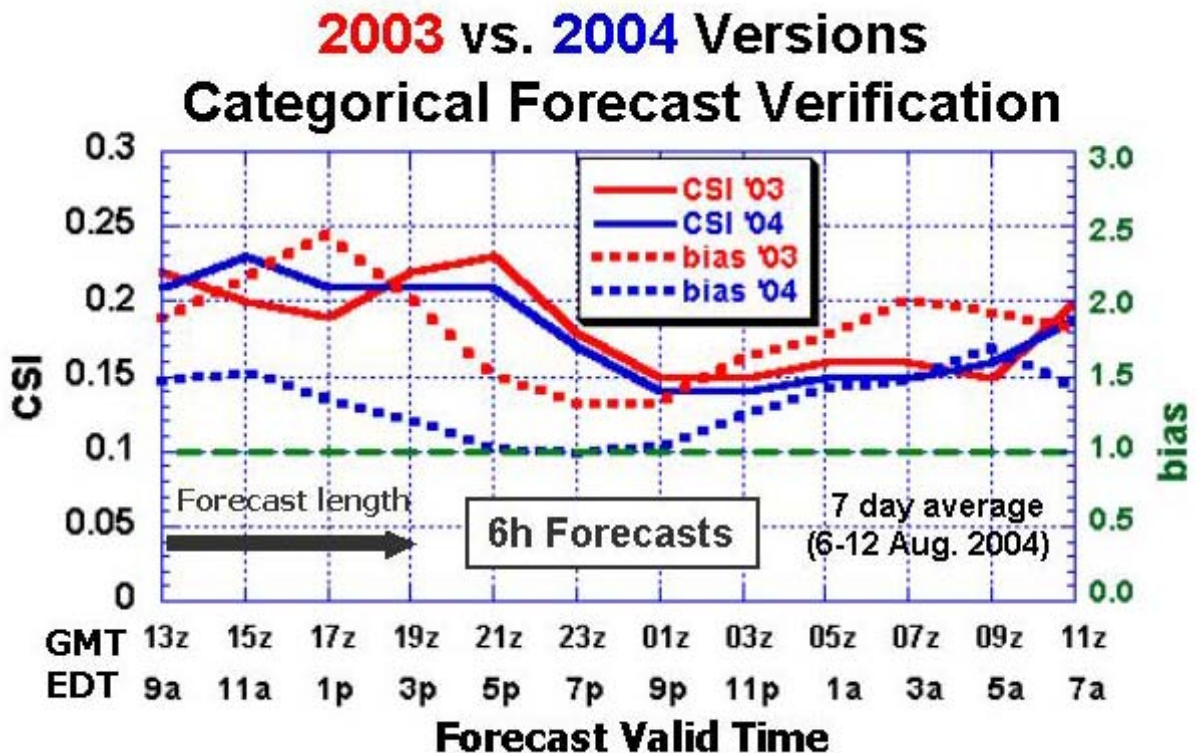


Fig. 9. Comparison of the diurnal cycle of average skill scores for the 2003 and 2004 versions of the RUC Convective Probability Forecast (RCPF). Averages are computed from a 7-day period (Aug 6-12) for 6-h RCPFs. Shown are the critical success index (CSI) and bias (bias), with a dashed line green line (bias = 1.0) included for reference.



## Sample comparison 2003 & 2004 products

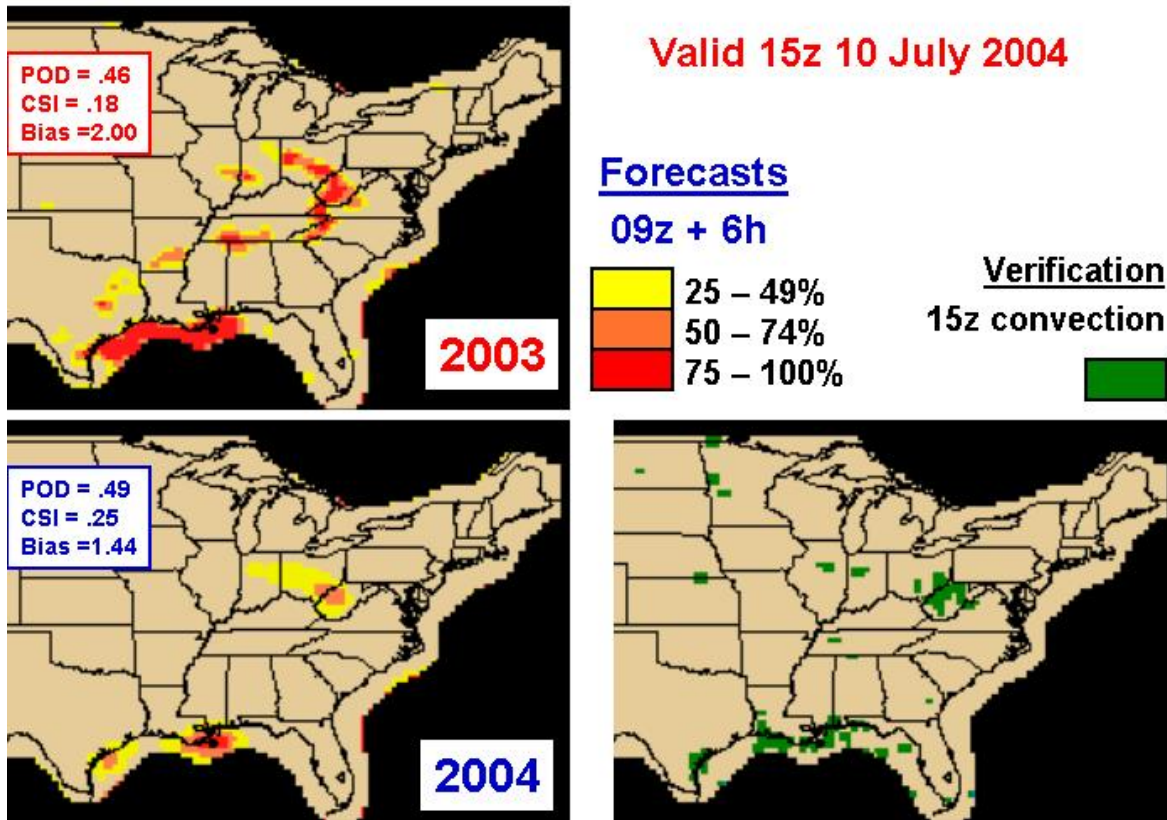


Fig. 10. Sample product comparison of 2003 and 2004 versions of the RUC Convective Probability Forecast for 6-h forecast valid 1500 UTC 10 July 2004. Shown also is the NCWD verification valid at 1500 UTC.

Accumulation of statistics for the 2004 real-time test is ongoing as are efforts to utilize other RUC model-predicted fields (lifted index, convective inhibition, horizontal convergence, etc.) in the construction of the RCPF. Additionally, real-time web-based display of the probability forecast as well as predicted storm tops will soon be available

### 6. SUMMARY AND FUTURE PLANS

Beginning in the spring of 2003 with an initial concept for a model-based probabilistic convective forecast, we have developed and refined the RUC Convective Probability Forecast (RCPF). Results from a real-time evaluation of the initial prototype conducted during the summer of 2003 indicated encouraging results and suggested a number of areas for improvement. Based on these results, D2 status within the AWTT process was obtained during the late spring of 2004 and a number of improvements to the product were implemented.

Preliminary results from a real-time comparison of the 2003 and 2004 products indicates that the goals established for the 2004 product have been satisfied. In particular, the excessive bias has been reduced, spatial coherency and temporal consistency has been improved, and product latency has been reduced without sacrificing performance as measured by the CSI. Real-time RCPFs are now available 15 minutes before the forecast hour, so that they could be examined before the issuance of the final CCFP product.

Work is proceeding on two fronts. First, we are working to facilitate the dissemination of RCPF output fields through a variety of pathways including real-time display on the RUC web-site and on the experimental Aviation Digital Data Services (ADDS, Sherretz et al 2002) web-site. Second, developmental work is ongoing or planned to enhance and improve the product. Enhancements include the display of additional fields, such as storm-top heights and storm motion.

# Sample 2004 RUC Convective Probability Forecast

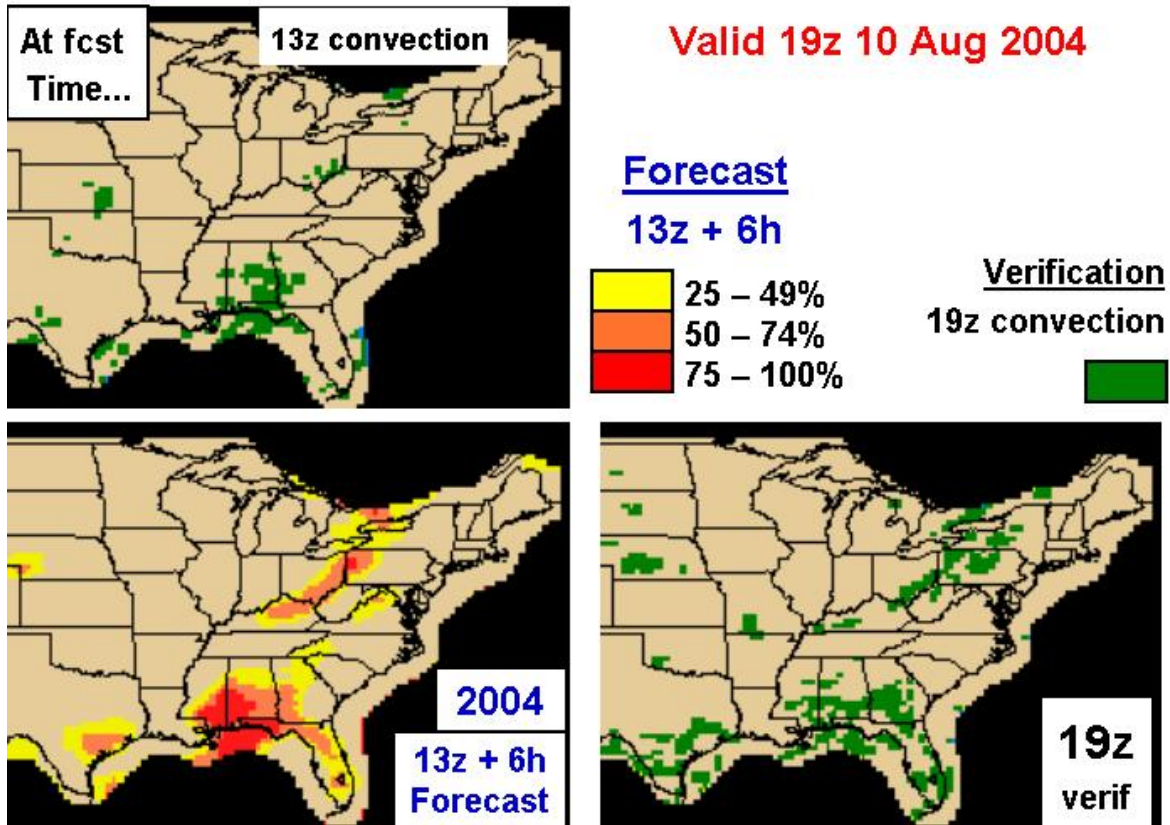


Fig. 11. Sample product from the 2004 version of the RUC Convective Probability Forecast for 6-h forecast valid at 1900 UTC 10 August 2004. Shown also is the NCWD verification field at the forecast issue time (1300 UTC) and the forecast valid (1900 UTC) illustrating the actual convective development.

Improvements include the use of ensemble closure information from the Grell-Devenyi (2002) cumulus parameterization scheme and other RUC fields in the generation of the probability forecast. Of course the RCPF will continue to improve from enhancements to the underlying RUC model forecasts. Relevant enhancements underway for the RUC include improvements to the cumulus parameterization, assimilation of radar data, more effective use of METAR information throughout the boundary layer in the RUC analysis (Benjamin et al. 2004c), and increased horizontal resolution (fully cycled 13-km RUC now in real-time testing, Benjamin et al 2004d). Plans for 2005 include the wider dissemination of the RCPF to interested users. Potential users could include the Aviation Weather Center (AWC), Central Weather Service Units, and the FAA Strategic Command Center.

The evolution path for this product is to merge it with the radar data-based short-range thunderstorm guidance products, such as the

National Convective Weather Forecast (Meganhardt et al. 2004), leading to the creation of a seamless 0-6 h thunderstorm likelihood product. As the product evolves, we envision a higher output frequency resulting in a more time continuous thunderstorm probability forecast. Forecast length extensions beyond 7-h are also possible.

## 7. ACKNOWLEDGMENTS

This research is in response to requirements and funding by the Federal Aviation Administration (FAA). The views expressed are those of the authors and do not necessarily represent the official policy or position of the FAA. Discussions with Cindy Mueller and others from NCAR, MIT/LL, AWC, and other organizations have been very helpful. Thanks also to Jennifer Mahoney and Joan Hart for their verification work. We acknowledge Tracy Smith and Nita Fullerton for careful scientific and technical reviews, respectively.

## 8. REFERENCES

- Benjamin, S.G., G. A. Grell, J.M. Brown, and T.G. Smirnova, 2004a: Mesoscale weather prediction with the RUC hybrid isentropic-terrain-following coordinate model. *Mon. Wea. Rev.* **132**, 473-494.
- Benjamin, S.G., D. Devenyi, S.S. Weygandt, K.J. Brundage, J.M. Brown, G.A.Grell, D. Kim, B.E.Schwartz, T.G. Smirnova, and T.L. Smith 2004b: An hourly assimilation-forecast cycle: The RUC. *Mon. Wea. Rev.* **132**, 495-518.
- Benjamin, S.G., S.S. Weygandt, D. Devenyi, J.M. Brown, G. Manikin, T.L. Smith, and T.G. Smirnova, 2004c: Improved moisture and PBL initialization in the RUC using METAR data. Preprints, *22nd Conf. on Severe Local Storms*, Hyannis, MA, Amer. Meteor. Soc., this volume.
- Benjamin, S.G., T.G. Smirnova, K.J. Brundage, S.S. Weygandt, T.L. Smith, B.E.Schwartz, D. Devenyi, J.M. Brown, and , G.A.Grell, 2004d: A 13-km RUC and beyond: Recent developments and future plans. Preprints, *11th Conf. on Aviation, Range, and Aerospace Meteorology*, Hyannis, MA, Amer. Meteor. Soc., this volume.
- Grell, G.A. and D. Devenyi, 2002: A generalized approach to parameterizing convection combining ensemble and data assimilation techniques. *Geophys. Res. Let.* **29**, 30-1-4.
- Knapp, D.I., R.J. Olson, F.R. Mosher, J.A. May, and S.R. Silberburg, 2002: Technology transfer at the aviation weather center: developing, testing, and implementing new forecast tools. Preprints, *10th Conf. on Aviation, Range, and Aerospace Meteorology*, Portland, OR, Amer. Meteor. Soc., 20-23.
- Mahoney, J.L., J.K. Henderson, B.G. Brown, J.E. Hart, A.F. Loughe, and C. Fischer, and B. Sigren, 2002: The real-time verification system (RTVS) and its application to aviation weather forecasts. Preprints, *10th Conf. on Aviation, Range, and Aerospace Meteorology*, Portland, OR, Amer. Meteor. Soc., 20-23.
- Meganhart, D., C. Mueller, S. Trier, D. Ahijevych, and N. Rehak, 2004: NCWF-2 probabilistic nowcasts. Preprints, *11th Conf. on Aviation, Range, and Aerospace Meteorology*, Hyannis, MA, Amer. Meteor. Soc., this volume.
- Mueller, D.I., C.B. Fidalgo, D.W. McCann, D. Meganhart, N. Rehak, and T. Carty, 1998: National Convective Weather Forecast Product. Preprints, *8th Conf. on Aviation, Range, and Aerospace Meteorology*, Dallas, TX, Amer. Meteor. Soc., 230-234.
- Sherretz, L., G. Pratt, G. Thompson, and J. Henderson, 2002: Implementing ADDS operationally at the NWS Aviation Weather Center. Preprints, *10th Conf. on Aviation, Range, and Aerospace Meteorology*, Portland, OR, Amer. Meteor. Soc., 81-84.
- Sims, D., D. Rhodenuis, and M. Szkil, 2004: Program management for the CCFP. Preprints, *11th Conf. on Aviation, Range, and Aerospace Meteorology*, Hyannis, MA, Amer. Meteor. Soc., this volume.