9.5 AN ANALYSIS OF CCFP FORECAST PERFORMANCE FOR THE 2005 CONVECTIVE SEASON

Michael P. Kay*
Cooperative Institute for Research in Environmental Sciences (CIRES) University of Colorado/NOAA Research-Earth System Research Laboratory
Boulder, Colorado

Jennifer L. Mahoney
NOAA Research-Earth System Research Laboratory
Boulder, Colorado

Joan E. Hart
Cooperative Institute for Research in Environmental Sciences (CIRES) University of Colorado/NOAA Research-Earth System Research Laboratory
Boulder, Colorado

1. INTRODUCTION

The Collaborative Convective Forecast Product (CCFP) continues to evolve as the primary tool through which the meteorological community and the commercial aviation industry have come together to create a common forecast to address the impact of convective weather on the national airspace at strategic time scales (2 to 6 hours). Since its inception in 1998, the product has been treated as a prototype that is modified annually through continuous and posterior feedback.

The intended purpose of the CCFP is stated as being contingent on the view of either the consumer or the producer (WAWG 2005). For the consumer, the forecast is to be used for strategic planning of air traffic flow during the en route phase of flight (WAWG 2005). For the producers (forecasters), the aim of the forecast is to accurately represent convection that is most significant for managing the flow of air traffic at strategic time frames.

The difference in views of the CCFP between producer and consumer lead to difficulty in determining appropriate verification techniques to satisfy all interested parties. From a consumer's view, one should evaluate the forecast with regards to traffic flow management decisions made with respect to each forecast and the resulting success and failure of those decisions. From a producer's perspective, the forecast should be evaluated with regards to the weather (i.e., convection), and the definition of the forecast, and not in terms of how it is potentially being used or interpreted. While both views are important, the challenge of performing systematic analyses of the relation of CCFP to the performance of the national airspace system is a very complex problem not easily solved. By starting with an assessment from the forecaster's point of view, where the forecast is clearly defined (see Section 2a.), one may begin to understand and improve the primary input to a larger decision support system that is used in strategic decision making for commercial aviation. Mahoney et al. (2004) have combined aspects of both the producer and consumer's views to provide verification results for the 2004 CCFP.

In this paper Section 2 will describe the CCFP in more detail to support the verification from a producer's point of view, the datasets used to verify the CCFP, and the verification methodology. Section 3 will describe the results of the verification of key components of the CCFP. Section 4 will present conclusions and discussion concerning the results found from the analyses.

2. DATA AND METHODOLOGY

a. CCFP

CCFP is defined as a set of one or more polygons (areas) that meet the following minimum criteria:

1. The areal extent of the polygon is at least 3,000 sq. mi.
2. Composite reflectivity of at least 40 dBZ is expected to cover at least 25% of the area, and

**Corresponding author address*: Michael P. Kay,NOAA/OAR/ESRL, R/GSD5, 325 Broadway, Boulder, CO 80305. email: mike.kay@noaa.gov

3. Echo tops (defined as the height of the 18.5 dBZ reflectivity value), at least 25 kft MSL, are expected to cover at least 25% of the area, and
4. Forecaster confidence in criteria 1 through 3 being met equals, or exceeds, 25%.

In addition to the so-called CCFP minimum criteria listed above, there are several additional attributes associated with each forecast polygon. Forecast coverage is defined as the expected coverage of convection within the polygon. Values include Low (24-49%), Medium (50-74), or High (75-100%) coverage. Forecaster confidence is specified by one of two values: Low (25-49%) or High (50-100%). Maximum echo tops within each polygon may take on one of three values: 25-31-, 31-37-, or greater than 37 kft MSL. Movement of each area of convection and a growth rate for convection within the polygon are the final two attributes.

b. National Convective Weather Detection Product

Convective coverage is assessed using the National Convective Weather Detection Product (NCWD) (Meganhardt et al. 2000). The data have a native spatial resolution of 4 km and an approximate temporal resolution of 5 minutes. Unless specified otherwise the nearest NCWD file within 10 minutes (+/- 5 min.) of the valid time of a CCFP forecast is used to verify the forecast. The NCWD data are thresholded at VIP level 3 to approximately match the requirements for CCFP issuance of 40 dBZ.

c. UNISYS national radar mosaic

The echo top dataset is the UNISYS national radar mosaic created from the national WSR-88D network operated by the National Weather Service. This data has 4 km horizontal resolution and is available on the same grid as the NCWD. The vertical resolution is 5,000 ft with intervals extending to 70,000 ft. Data are time matched to CCFP valid times for verification.

The approach taken by this study for the verification of CCFP is to assess the forecast quality at the individual polygon level as opposed to the forecast level which represents the collection of all polygons that share issuance and valid times. Past studies such as Mahoney et al. (2004) have focused on forecast level results treating CCFP forecast polygon sets in a dichotomous manner and deriving overall statistics such as probability of detection, false alarm rate, and bias. In verifying the CCFP on a polygon by polygon basis, each polygon's attributes are assessed individually; something that cannot be done when aggregating all polygons for a given forecast. The directly verifiable attributes are coverage and echo tops. Forecaster confidence can be evaluated but not verified because a subjective belief is neither correct nor incorrect. Growth rate and movement are not verifiable attributes in the current definition of CCFP because the forecast does not define the time periods over which movement or growth are to be computed.

CCFP forecasts are studied for the period 1 March 2005 through 4 October 2005. There were 6334 unique forecasts analyzed, representing the collection of 2230 issuance times with their 2-, 4-, and 6-h forecast components. The remaining 356 forecasts did not contain any forecast polygons and therefore are not addressed by this study. Within the CCFP forecasts, only the forecast areas specifically defined as areas are verified in this study; line-type forecast areas are not considered.

CCFP polygons with centroids west of -105 longitude are not verified in this study owing to the limitations of the radar coverage within the western United States associated with sparse coverage of the network in addition to terrain effects (Howard et al. 1997). Additionally, CCFP polygons that do not lie wholly within the the continental U.S. are also discarded to eliminate cases where observed coverage computations would be effected by incomplete radar coverage at the edges of the network. Data for these polygons are used in discussions of climatological aspects of the forecasts, but are not used for verification.

3. RESULTS

For all analyses, CCFP polygons are gridded onto the 4 km grid used by the NCWD and UNISYS radar mosaic products in order to derive statistics unless otherwise noted.

a. Coverage

Coverage of CCFP is the most readily viewed attribute since most users rely on displays combining forecast areas with radar and other spatial data that is available in their forecasting environments. The frequency at which CCFP polygons meet minimum coverage (25% or greater) when verified using NCWD data at its native 4 km resolution is shown in Table 1. Significant overforecasting is evident in all categories. High coverage forecasts are the best indicator that a forecast polygon will meet minimum coverage requirements. Forecasts of High coverage represent only a very small fraction of the observed number of CCFP polygons. More information on CCFP forecast issuances and behavior can be found in Seseske et al. (2006).

Table 1. Percent of polygons issued for each coverage category that had at least 25% coverage using 4 km data. Values in parentheses indicate number of polygons issued.

Coverage	Verified
Low	0.4% (13569)
Medium	4.9% (1884)
High	16.5% (109)

Given the small number of polygons meeting minimum coverage requirements another set of computations were performed using NCWD observation regridded onto a 40 km grid. For the regridding process, if any 4 km grid box contained a VIP 3 or higher value the bounding 40 km grid box was marked as having convection. The results, shown in Table 2, indicate a substantial increase in the number of polygons meeting minimum coverage, especially for Medium and High coverage issuances.

Table 2. As in Table 1 except for 40 km verification data.

Coverage	Verified
Low	31.9% (13569)
Medium	66.4% (1884)
High	83.5% (109)

To further understand the nature of observed coverage values, Fig. 1 shows the distribution of observed coverages for each forecast category. To aid the comparison with values currently computed by the Real-Time Verification System (RTVS; Mahoney et al. 2002) the distribution of coverages using NCWD data regridded to 40 km with the addition of a +/- 10 minute buffer around the CCFP valid time is shown as well. If a grid box contains convection (40 dBZ or higher) for any observation file (typically 3 or 4 are available) in that time window it is marked as having convection. The time window provides an additional relaxation on the observations that acts to give a forecast credit for event occurrence near a valid time rather than exactly at a valid time.

Changing the verifying resolution from 4 km, to 40 km, and then to 40 km with temporal smoothing significantly increases the percentage of polygons that meet minimum coverage requirements. The increased coverages do not however cause observed coverages to correspond to predicted coverages for the various forecast categories. Median coverage values for all combinations except Low Coverage, 40 km time-windowed NCWD, fall below the expected coverage ranges and in most cases, significantly below them. For the 40 km NCWD data, the percentages of Low-, Medium-, and High-coverage polygons verifying with the correct observed coverage categories were 19%, 28%, and 5%, respectively.

b. Confidence

CCFP polygons are issued only when forecaster confidence meets or exceeds 25% that the coverage and echo top requirements will be satisfied. Without considering echo tops, Table 3 presents the percent of CCFP polygons that met or exceeded 25% coverage irrespective of the coverage attribute assigned to each polygon.

Table 3. Percent of polygons issued that had at least 25% coverage of NCWD data at 4 and 40 km resolution for each confidence level. Values in parentheses indicate number of polygons issued.

Confidence	4 km	40 km
Low	0.4% (9122)	27.7%
High	2.0% (6447)	48.7%

As noted in the coverage discussion, the spatial resolution chosen for computing coverage dominates the results but does not change the underlying signal; at both resolutions the polygons are significantly more likely to meet minimum coverage criteria when the forecasters have high confidence. This can be further broken down to

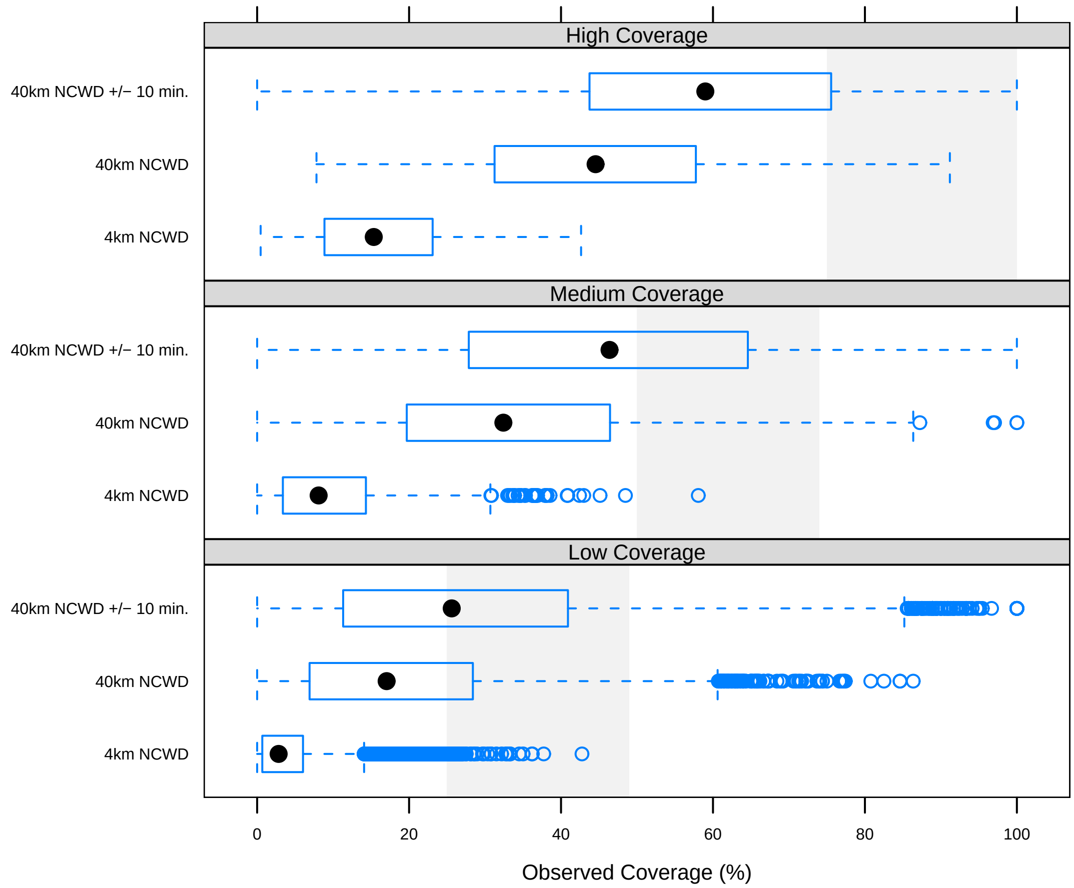

Fig. 1. Box and whisker diagram of observed coverage percentages for 4km NCWD, 40 km NCWD with no temporal smoothing, and 40 km NCWD with 20 min. time window approaches for Low, Medium, and High coverage CCFP polygons. Gray zone within each panel signifies expected range for observational values based upon forecast category.

consider the difference between 2- and 6-h forecasts (Table 4).

Table 4. As in Table 3 except for 2- and 6-h forecasts.

Confidence	2-h	6-h
Low	35.0% (2918)	23.0% (3112)
High	60.0% (2901)	35.0% (1524)

Low confidence forecasts are issued with increasing frequency from 2- to 6-h while high confidence forecasts are primarily used at 2-h. When treating polygons as either correct or incorrect and confidence in terms of the probability ranges defined for each, the forecasts can be seen as reasonably reliable except for high confidence 6-h forecasts, which are significantly overforecast.

c. Echo tops

Echo top information is used in CCFP to describe both the minimum criteria (25% or more of the area should be covered with echo tops at least 25 kft MSL at valid time) and the altitude of maximum echo tops expected in the area at the valid time. Each of these uses will be discussed below. The forecast and observed coverages of CCFP areas with echo tops at least 25 kft MSL when verified at 4 km resolution are shown in Table 5.

Table 5. Percent of polygons having at least 25% coverage of echo tops as defined by CCFP using 4 km radar data. Values in parentheses indicate number of polygons issued.

Echo Top Category	Verified
> 37 kft	4.8% (8633)
31-37 kft	2.1% (120)
25 – 31 kft	0.6% (8)

Most CCFP polygons fail to meet the CCFP minimum threshold of 25% coverage for any of the forecast categories. The distribution of forecast values indicates that forecasters clearly believe that in most convective situations that deep convection will exceed 37 kft MSL.

To assess the impact of observation resolution on echo top distribution, the echo tops were regridded to 40 km using a similar strategy to the one used to bring NCWD data from 4 km to 40 km (Table 5). The maximum echo top value within the set of 4 km grid boxes comprising each 40 km grid box was assigned as the value for the 40 km box. The results show significant improvement with over 40% of all polygons exceeding 25% coverage.

Table 6. As in Table 5 except using 40 km NCWD data.

Echo Top Category	Verified
> 37 kft	62.9% (8633)
31-37 kft	47.3% (120)
25 – 31 kft	42.2% (8)

The conditional probabilities of echo top observations given forecasts are presented in Table 7. These values represent the distributions of the four possible observation categories for each of the three forecast categories.

Table 7. Conditional probability of echo top observation category given forecast category using the 4 km echo top data.

		Observed			
		> 37 kft	31-37 kft	25-31 kft	< 25 kft
Fcst	> 37 kft	0.41	0.10	0.20	0.30
	31-37 kft	0.27	0.10	0.27	0.36
	25-31 kft	0.15	0.10	0.32	0.44

Several interesting aspects of the data can be seen in Table 7. First, a very small percentage of the observations fall within the 31-37 kft range. Of greater interest is how often the forecast category does not agree with the observed category. Observed values are usually significantly below the forecast value. While the WSR-88D echo top algorithm is known to underestimate storm tops in some situations (Howard et al. 1997) the discrepancy here is much larger than is typically observed purely due to radar limitations. The primary reason for the differences appears to be that the climatological distribution of observed echo top values is considerably lower than one might expect and therefore the lower height values are much more probable. The probability density function of observed echo tops from the month of July 2005 is shown in Fig. 2 where it is clear that values between 15 kft and 25 kft MSL are most probable. The most likely level, 15 kft, is below the CCFP minimum-defined echo top height.

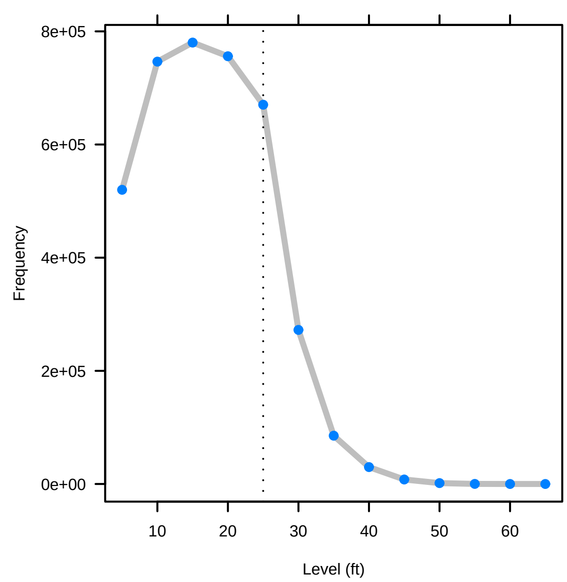

Fig. 2. Distribution of observed echo tops within CCFP polygons during July 2005. Dashed line indicates CCFP threshold height of 25 kft.

d. CCFP minimum criteria

The results have thus far focused on the verification and evaluation of the primary attributes of the CCFP. Attention is now turned to the evaluation of polygons in terms of whether or not all parts of the CCFP minimum requirements are met. Results for the number of polygons meeting these requirements are presented at both 4- and

40 km resolution along with counts for the coverage and echo top components in Table 8.

Table 8. Counts of number of polygons meeting minimum requirements for echo top, coverage, and the combination of echo top and coverage. Total number of possible polygons is 13562.

Resolution	Combined	Coverage	Echo Tops
4 km	27	165	541
40 km	1254	5670	2686

Using the current datasets, even the most liberal comparison leads to minimum requirements being met for under ten percent of all polygons issued with less than one percent verifying at the native 4 km resolution. Of great interest is the discrepancy between the coverage, echo top, and combined counts. In a large number of cases there is a lack of coincidence between NCWD data thresholded at VIP level 3 (approximately 40 dBZ) and WSR-88D echo tops that exceed 25 kft MSL, as illustrated by the difference between the combined and coverage columns in Table 8. If there was no difference the two columns would be identical owing to the combined field representing the intersection of the two datasets.

Additionally, the results suggest that CCFP polygons are more often covered with anvil material than precipitation cores. The echo top data tends to be a much smoother, more coherent field than the NCWD. This is supported by the fact that the regridding process is most beneficial to fields which are scattered such as higher reflectivity precipitation cores. The ratio of the 40 km counts to the 4 km counts can be used to indicate an approximate benefit of being regridded to a coarser resolution. This ratio of the 40 km counts to the 4 km counts is 5.0 for echo tops and 34.4 for NCWD data.

4. DISCUSSION

The results presented in Section 3 suggest that the forecast definition was created without consideration for the climatological aspects of convection that form the basis for the forecast itself. For instance, when high coverage (75-100% of an area) was forecast the average verifying coverage was 16.5%. This may be due to the fact that the forecasts are defined to be valid at a snapshot in time rather than valid over a certain time window. Convective timescales are such that there is considerable variability in most situations at any given time. It is possible that forecasters are instead producing a time-integrated areal forecast. This is supported by the significantly improved verification results when temporal smoothing of areal coverage was used. Given that high coverage CCFP forecasts are usually linked with significant impacts to air traffic flow, it is understandable that one would predict large coverage values that correlate with them. An adjustment to the coverage values that are linked with the existing categorical labels that are based upon climatological values of convective coverage within CCFP polygons would allow users to continue to utilize CCFP for decision making in its current form. Such an adjustment would provide a more sound definition for CCFP that better matches observed data.

For future CCFP definitions a clarification of the time period over which the movement and growth and decay attributes are valid would be beneficial for allowing a complete assessment of all CCFP attributes.

Future work will focus on understanding the differences between the WSI echo top mosaic and other higher-resolution fields such as the Corridor Integrated Weather System (CIWS) echo tops field and an echo top field derived from the National Mosaic and Multisensor Quantitative Precipitation Estimation Project (NMQ) dataset and the impact that the choice of echo top observations have on the echo top verification of CCFP. The lack of correlation between NCWD forecasts of strong convection and low echo top values observed in this study will be addressed as well.

5. ACKNOWLEDGMENTS

Thank you to Sean Madine and Nick Matheson for numerous stimulating conversations related to this work.

6. REFERENCES

Howard, K.W., J.J. Gourley, R.A. Maddox, 1997: Uncertainties in WSR-88D measurements and their impacts on monitoring life cycles. *Wea. Forecasting*, **12**, 166-174.

Mahoney, J.L., J. K. Henderson, B.G. Brown, J.E.

Hart, A. Loughe, C. Fischer, and B. Sigren, 2002: The Real-Time Verification System (RTVS) and its Application to Aviation Weather Forecast. Preprints, *10th Conference on Aviation, Range, and Aerospace Meteorology*, 13-16 May, Portland, OR, Amer. Met. Soc.

Mahoney, J.L., S. Seseske, J.E. Hart, and M. Kay, 2004: Defining observation fields for verification of spatial forecasts of convection: Part 2. Preprints, *11th Conference on Aviation, Range and Aerospace Meteorology*, Hyannis, MA, Amer. Met. Soc.

Meganhardt, D., C.K. Mueller, N. Rehak, and G. Cunning, 2000: Evaluation of the National Convective Weather Product. Preprints, *9th Conference on Aviation and Range Meteorology*, Orlando, FL, Amer. Met. Soc., 171-176.

Seseske, S., J.E. Hart, 2006: Verification of the CCFP: Creating Operationally-Relevant Statistics. Preprints, *12th Conference on Aviation and Range Meteorology*, Atlanta, GA, Amer. Met. Soc., this volume.

Weather Applications Working Group, 2005: CCFP User Needs Document, K. Mullen and F. Hayoz, Ed., 37 pp. Available from http://cdm.metronaviation.com/Workgroups/WxApps/CCFPUserNeeds05(V8.71)01-27-05.doc.