3.8 **SKILL OF THE AVIATION WEATHER CENTER'S COLLABORATIVE CONVECTIVE FORECAST PRODUCT (CCFP)**

Jonathan W. Slemmer*
NOAA/NWS/NCEP/Aviation Weather Center, Kansas City, Missouri

1. INTRODUCTION

The Collaborative Convective Forecast Product (CCFP) was initiated in 1999 at the Aviation Weather Center (AWC) to meet the need of the Federal Aviation Administration (FAA) of getting better convective forecasts for managing air traffic (Fahey, et al. 2004). Over the years, CCFP has evolved through user feedback and analytical feedback on forecaster precision with the intent to improve the forecast usefulness and accuracy (Torbert, et al. 2004).

The purpose of this study is to introduce and evaluate three new methodologies for assessing CCFP skill using Convective SIGMETs (CSIG). The objective is to provide users and forecasters with insight into the CCFP forecasting process and provide data usable for improving the forecasts.

In general, outstanding problems in the assessment of convective forecast skill are:

a) A measure of convection that matches the convective forecast criteria.
b) The use of appropriate assessment methodologies.

In previous CCFP skill studies, National Convective Weather Detection (NCWD) data has been used as the observed convection in different ways (Kay, et al. 2006, Mahoney, et al. 2004, and Seseske, et al. 2006). In this study, CCFP skill is assessed using CSIGs as the observed convection. Although the criteria used for generating CCFP and CSIG areas are different, they have enough similarities to find value in comparing them. As a practical matter, these forecast criteria cannot be distinguished very well by a forecaster. Various CCFP skill statistics can also be found on the Real Time Verification System RTVS web-site (http://www-ad.fsl.noaa.gov/stage2fvb/rtvs/) including using CSIG to determine CCFP skill which was developed by NOAA's Earth System Research Laboratory (ESRL) with funds provided by the FAA Aviation Weather Research Program.

CCFP is a graphical representation of expected significant convective occurrence at 2-, 4-, and 6-hours after issuance time that regularly occur at 2 hour intervals from March through October except for one late night forecast. Significant convection for the purposes of CCFP forecast areas is defined as a polygon of at least 3000 square miles that contains:

a) A coverage of at least 25% with echoes of at least 40 dBZ composite reflectivity.
b) A coverage of at least 25% with echo tops of FL250, or greater.
c) A confidence of at least 25%.

For further information on CCFP refer to: http://aviationweather.gov/products/ccfp/docs/pdd-ccfp.pdf

CSIGs are a graphical representation of current significant convection whose purpose is to warn aviation interests of convective hazards every hour based on the following criteria:

a) Severe thunderstorm(s) and embedded thunderstorm(s) occurring for more than 30 minutes of the valid period regardless of the size of the area.
b) A line of thunderstorms.
c) An area of active thunderstorms (VIP level of 4 or greater and/or having significant satellite signatures and affecting at least 40 percent of

Figure 1. High volume air traffic area considered for CCFP skill evaluation in red defined as an area bounded by four VOR points from 150NW ORD to 225ENE JFK to 225ESE ATL to 150SW DFW.

* *Corresponding author address*: Jonathan W. Slemmer, NOAA/National Weather Service, National Center for Environmental Prediction, Aviation Weather Center, 7220 NW 101st Terrace, Room 118, Kansas City, MO 64153-2371; email: Jonathan.Slemmer@noaa.gov

the area outlined; the national lightning network is also used to determine the level of activity) affecting at least 3,000 square miles.

Forecasters analyze current significant convective hazards through a combination of radar, lightning, and satellite data.

For further information on CSIG refer to: http://www.nws.noaa.gov/directives/sym/pd01008011curr.pdf

The three new methodologies for assessing CCFP skill using CSIG that will be introduced in the paper are:

a) Interpertating the CCFP diurnal skill scores (POD, FAR, CSI, BIAS) for a subjectively determined high volume air traffic area (Fig. 1) from evaluation of each grid point except for individual CCFP areas which were evaluated for the CONUS.
b) Assessing “forecaster work” in cases where CCFP and CSIG areas did not occur (“null cases”) using the Rapid Update Cycle (RUC) model K-Index analyses in the high volume air traffic area (Fig. 1) from evaluation of each grid point.
c) Evaluating convective jet airway segment (JAS) impacts defined as an area 20 nm either side of a portion of a jet airway from a major airport to the next major intersection. The entire JAS will be considered to have been impacted by either a forecast and/or observed convection if any portion of the JAS is affected.

2. METHODOLOGY

ASCII text files of CCFP products for the summers (June through August) of 2005 and 2006 are mapped to a grid of 8 x 8 km for each 2- and 6-hour forecast valid time. The 4-hour forecast was not analyzed due to time and processing constraints and 4-hour skill scores may have added little additional value to this study. Similarly, ASCII text files of CSIG products are interpolated to the same grid for each CCFP forecast valid time.

The CCFP has six possible area combinations (low confidence/low coverage, low confidence/ medium coverage, low confidence/high coverage, high confidence/low coverage, high confidence/medium coverage, and high confidence/high coverage) and solid lines. The first two previously mentioned methodologies in assessing CCFP skill (diurnal verification and assessing “forecaster work” in “null cases”) will treat all CCFP areas equally, whereas the last one (JAS impacts) will assess four of these seven possible CCFP objects into three sub-groupings which are low coverage/low confidence, low coverage/high confidence, and medium or high coverage/high confidence and ignore the others since their frequency of occurrence is rare (less than 1 percent of all issuances).

Diurnal CCFP skill scores were generated for both the 2- and 6-hour CCFP forecasts by assessing each grid point within the high volume air traffic area for each hour that the CCFP forecasts were valid during the summers of 2005 and 2006.

While verification schemes typically analyze the value of the forecasting process when something has occurred (forecasted and/or observed), little has been done to analyze this value when nothing has been forecasted and observed (“null cases”). However, not all “null cases” involve significant “forecaster work”. Instances of significant “forecaster work” occurs when the forecaster takes time to analyze areas that might have significant enough convective potential to require the issuance of a CCFP area. Insignificant “forecaster work” occurs when a forecaster is quickly able to exclude areas that have little or no potential for significant convection meeting CCFP issuance criteria. In order to be able to distinguish between insignificant and significant “forecaster work,” the RUC initial analyses K-Index were compared to the “null cases.” The K-Index incorporates two of the three most important convective ingredients (moisture and instability but not lift) and is frequently used by forecasters to assess convective potential, and thus was considered to be a good single parameter to indicate the potential for significant convective development.

RUC K-Index initial analyses are output to gridded files. Since the CCFP and CSIG gridded files are a higher resolution, the CCFP and CSIG grid points that are the closest to the nearest RUC gridded points are interpolated to the RUC grid. From these gridded files, all grid points over the high traffic area that do not have a CCFP and/or CSIG area are assessed for the potential of significant convective development using the K-Index value in order to evaluate “forecaster work".

A method was developed for assessing the potential impact to a jet airway into major airports where commercial air traffic volume can be high. JASs that connected from near one of four major airports (Chicago, New York, Atlanta, Dallas/Fort Worth) to other major intersections were analyzed for containing a CSIG and/or CCFP area(s) determined from the 8 x 8 km gridded files.

Since any observed or forecasted convection along a JAS may have an impact on a entire JAS; a forecast will be considered correct when a CSIG and CCFP area occurred along the JAS, a forecast will be considered missed when a CSIG area occurred but a CCFP area did not occur along the JAS, and a forecast will be a false alarm when a CCFP area occurred but a CSIG area did not occur along the JAS. The highest CCFP confidence and coverage sub-grouping (low confidence/low coverage, high confidence/low coverage, and high confidence/medium or high coverage) will verify as having been detected when a CSIG also occurs anywhere along the route.

3. DIURNAL CCFP SKILL

The typical convective forecast day generally begins in the morning with the forecaster trying to determine where and when diurnally driven convection will develop. This can be very difficult and is reflected in the relatively high frequency of missed forecasts (Figs. 2a-b) around mid-day which is reflected in the lower CSI score (Fig. 2c).

Once most of the convection has developed or is beginning to develop by the afternoon, the forecaster knows where convection is most likely to occur over the next several hours. This results in a decrease of the missed forecasts and an increase of the correct forecasts which also corresponds to the peak of the CSI and POD scores (Figs. 2a, 2c). However, the forecaster tendency is to not diminish the convection quickly enough by the evening resulting in a higher BIAS and FAR (Figs. 2b, 2d).

As would be expected, the 2-hour forecast has greater overall skill than the 6-hourforecasts. Overall CCFP skill for 2005 as compared to 2006 was very similar with only a slight improvement in 2006 (not shown).

Besides the previous skill scores that were evaluated at each grid point, individual CCFP areas were also evaluated in the CONUS for the minimum areal coverage threshold of 25 percent. Figure 3 shows the percent frequency of when a CCFP area had at least 25 percent CSIG coverage. The lowest percentage of verifying areas occurs in the later morning while the highest is the late afternoon. Overall, 60 and 43 percent of 2-hourand 6-hourCCFP area forecasts meet minimum CSIG coverage of 25 percent or greater,

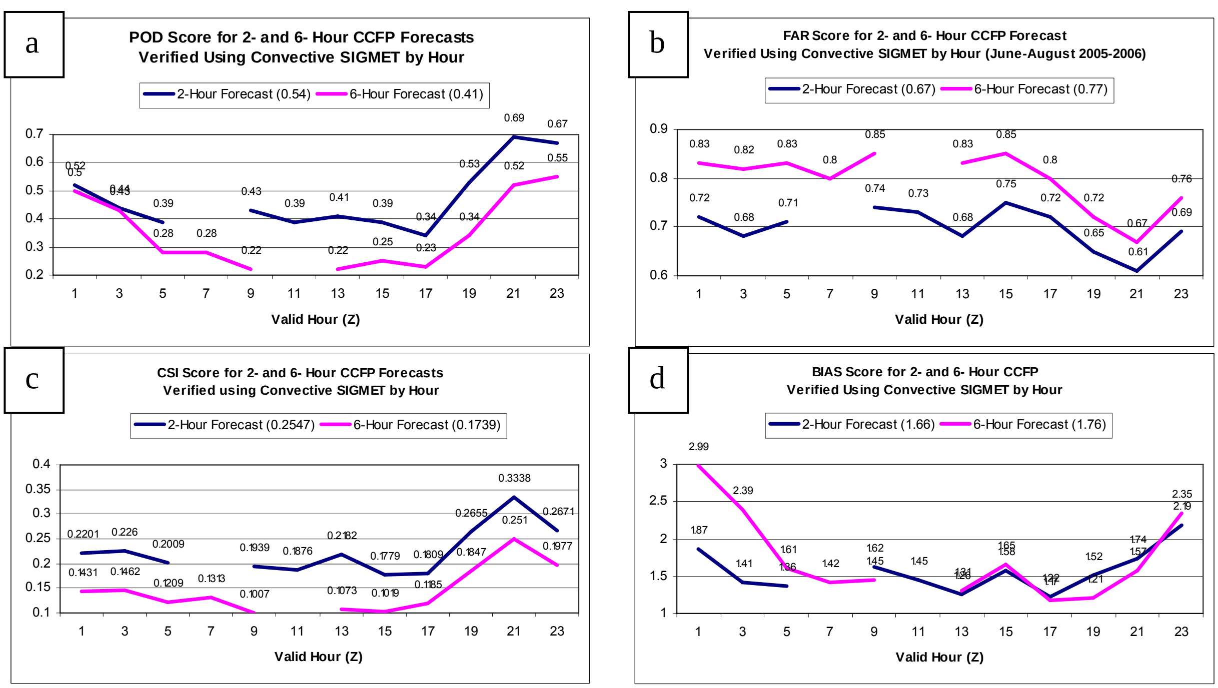

Figure 2. June through August 2005-2006 CONUS CCFP skill scores: a) POD b) FAR c) CSI d) BIAS.

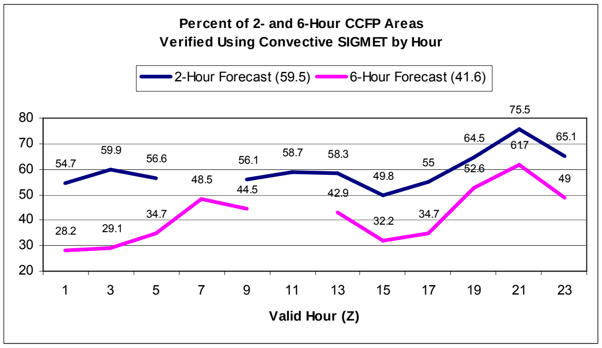

Figure 3. Percent of CCFP areas verifying by hour with at least 25 percent CSIG coverage for June through August 2005-2006 in the CONUS.

respectively.

4. ASSESSING FORECASTER WORK IN NULL CASES

The literature suggests 50 percent or greater areal coverage of thunderstorms is possible when KI ≥ 25 for east of the Rockies. To better assess whether a KI ≥ 25 was a good value to use for assessing the potential for the development of significant convection, different K-Index thresholds were also compared to the times when CSIGs were issued to ascertain what value best correlated with significant convection.

KI ≥ 25 correlated with forecasted and observed convection when CSIG and CCFP areas overlapped or HITs around 95 percent of the time for each 2- and 6-hourCCFP valid hour (Fig. 4a-b). Incorrect forecasts or FARs (forecasted but not observed) and MISSes (observed and not forecasted) were generally slightly less at about 90 to 95 percent of the time (Fig. 4a-b).

KI for other values were also tested to see whether they were a better KI threshold value compared to a value of 25 for determining the potential for significant convective development. KI threshold values of 20 and 30 were tested for the 2- and 6-hourCCFP forecasts and CSIG correlations occurred around 95 and from about 65 to 75 percent of the time each hour, respectively (not shown). Any additional benefit from KI ≥ 20 was minimal and if used could led to additional "forecaster work" that mostly might be unnecessary. KI ≥ 30 resulted in a significant correlation decrease with CSIG occurrence which could result in excluding areas from forecaster consideration that had significant convective potential. This confirmed that KI ≥ 25 was the best choice for evaluating the "forecaster work" in "null cases" and will be used in this study.

The 2- and 6-hourCCFP forecast valid for each hour were assessed for each grid point for the high volume air traffic area defined in Figure 1 and were classified into three categories: a) forecasted and/or observed convection, b) no forecasted and no observed convection ("null cases") and KI ≥ 25, and c) no forecasted and no observed convection ("null cases") and KI < 25.

In general, approximately half of all "null cases" had KI ≥ 25 during less convectively active times and decreased to less than half the time during the more convective times in the afternoon (Fig. 5a). Cases where convection was forecasted and/or observed varied from about 20 percent in the late afternoon to 3 to 5 percent in the early morning (Fig. 5b). Generally, "forecaster work" when "null cases" are combined with forecasted and/or observed convection may be significant more than half the time for the high volume air traffic area considered using this method.

Figures 5c-d shows the hourly percentage distribution of the cases when convection was forecasted and/or observed. As an example at 21Z, 20 percent of the time convection was observed and/or forecasted (Fig. 5b). Of this 20

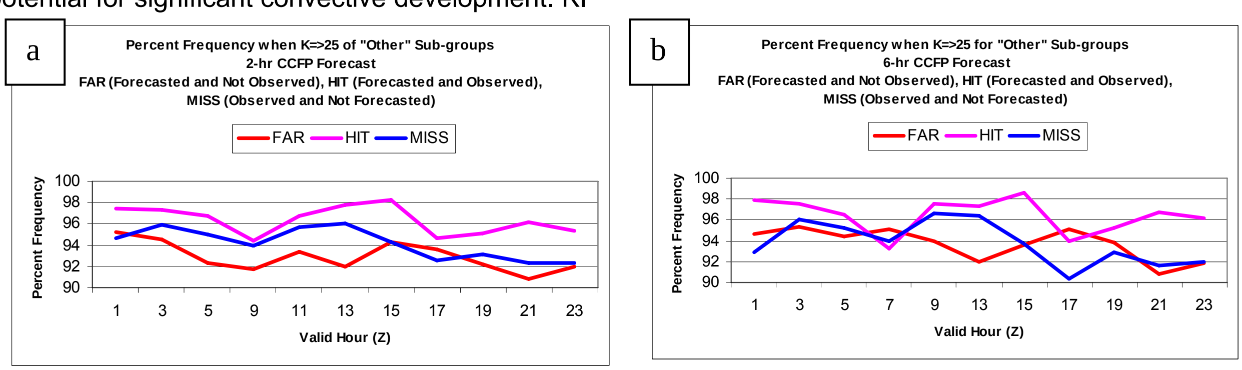

Figure 4. June through August 2005-2006 High Traffic Volume CCFP skill scores of percent correlation of correct, incorrect, and missed CCFP forecasts with CSIGs: a) 2-hourb) 6-hr.

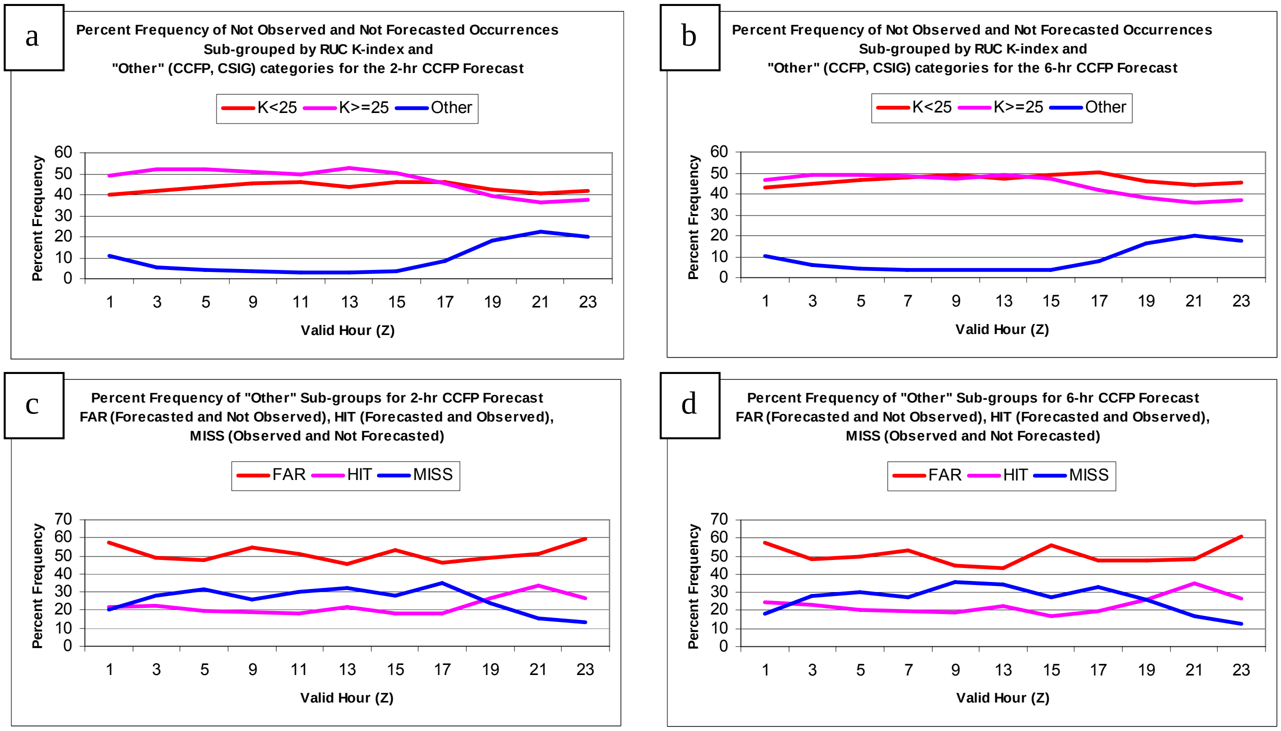

Figure 5. June through August 2005-2006 High Traffic Volume CCFP skill scores: a) Percent Distribution of "Null Cases" when $KI \geq 25$ and $KI < 25$ and "Non-Null Cases" *for 2-hourforecast b) same as previous description for 6-hourforecast c) Percent distribution of "Non-Null Cases" into correct forecasts, missed forecasts, and incorrect forecasts for 2-hourforecast d) same as previous description for 6-hourforecast.*

percent, figure 5d shows that 48 percent of the time was an incorrect forecast (FAR), 27 percent of the time were correct forecasts (HIT), and 17 percent of the time were missed forecasts (MISS). The peak in the percentage of correct forecasts occurs in the late afternoon with a low around mid-night. The peak in the percentage of missed forecasts occurs at mid-day with a low in the early evening while the incorrect forecasts high and low occur at the exact opposite times.

5. JAS IMPACTS

Figure 6 shows the conceptual model of the JAS CCFP verification scheme. Figure 6a shows an example of a correctly forecasted low coverage/low confidence CCFP area since portions of both the CCFP and CSIG areas occur within the JAS. Figure 6b shows an example of an incorrectly forecasted low coverage/low confidence CCFP area since a portion of the CCFP area occurs within the JAS but the CSIG area does not. Figure 6c shows an example of a missed low coverage/high confidence CCFP forecast as a portion of the CSIG area occurs within the JAS but the CCFP area does not. Figure 6d shows an example of a correctly forecasted medium coverage/high confidence CCFP area since portions of both the CSIG and CCFP area occur within the JAS.

JASs that connected from near one of four major airports (Chicago, New York, Atlanta, Dallas/Fort Worth) to the nearest other major intersection were analyzed to see whether they contained a CSIG and/or CCFP area(s). Percent frequencies of correctly forecasted CCFP areas for the three coverage and confidence sub-groupings earlier discussed and also the frequency of time when a CSIG occurred with no CCFP area occurring along the JAS were put into tables for each airport for the 2- and 6-hour CCFP forecasts.

Figure 7 shows the statistics for all JASs into the ATL (Atlanta) VOR[1] which is located very close to

[1] VHF Omnirange. A ground-based navigation aid transmitting very high frequency (VHF) navigation signals 360° in azimuth, on radials oriented from magnetic north. The VOR periodically identifies itself by Morse Code and may have an additional voice identification feature. Voice features can be used by ATC or FSS for transmitting information to pilots.

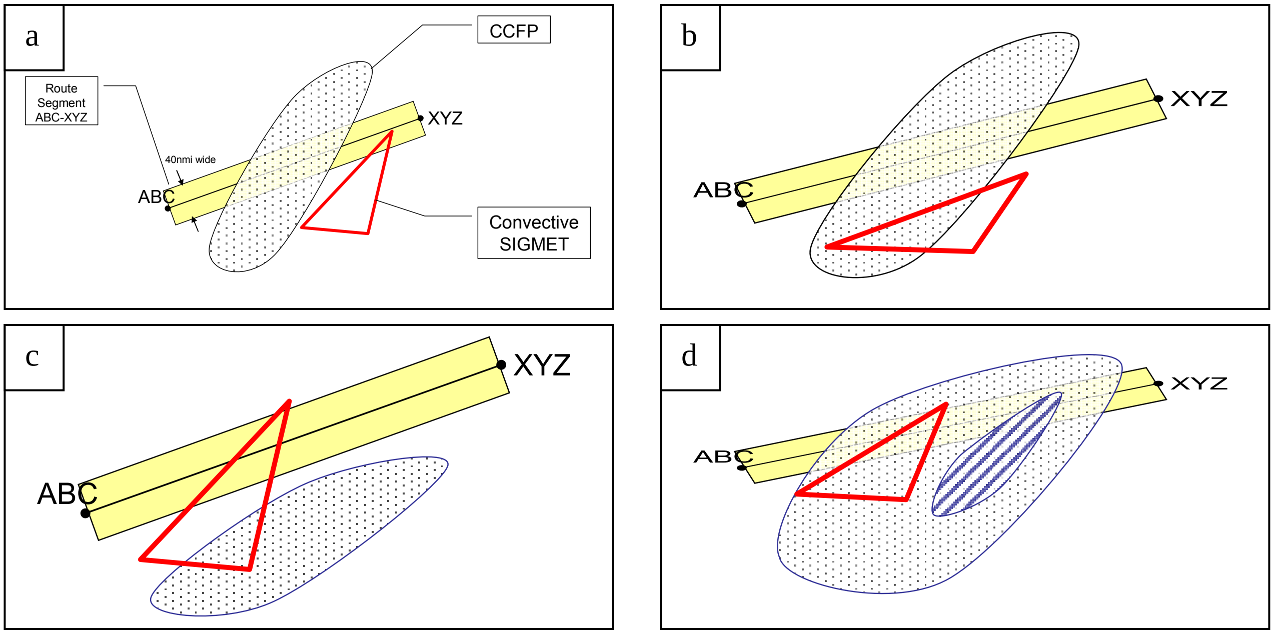

Figure 6. Conceptual Model of the JAS CCFP verification scheme: a) low coverage/low confidence CCFP area was correctly forecasted b) low confidence/low coverage CCFP area was incorrectly forecasted c) low confidence/high coverage CCFP area was a missed forecast d) medium confidence/high coverage CCFP area was correctly forecasted.

the Atlanta International Airport. In general, as the coverage and confidence increases the frequency at which the CCFP area verifies increases. Although the overall 2-hour CCFP forecasts verify with a higher frequency than the 6-hour CCFP forecasts and the percent frequency of missed convection where a CSIG area was occurring but not a CCFP area decreased slightly, the forecast accuracy improvement of the 2- versus 6-hourforecast is surprisingly small using the JAS method.

Figure 8 shows another way to look at the JAS verification by combining all JASs for an airport for each CCFP forecast hour. The general trend of these statistics suggest the typical peak in convection in the afternoon has the best verification of CCFP areas while mid-day has the highest frequency of misses. Verification of the JAS is the worst at either end (mid-morning and late-evening). Once again surprisingly, the 2-hourCCFP forecast skill does not seem much better than the 6-hourCCFP forecast.

a

VOR and Route Number	Freq CSIG/CCFP Occurence Low Conf/Low Cov	Freq CSIG/CCFP Occurence High Conf/Low Cov	Freq CSIG/CCFP Occurence High Conf/Med+ Cov	Percent CSIG with No CCFP (Missed)
VUZ (14)	52	71	91	7
SPA (14)	51	72	83	6
SZW (43)	55	79	93	7
VXV (43)	50	65	85	7
MCN (45)	47	69	81	6
BNA (45)	49	66	86	8
IRQ (52)	47	73	89	7
OTK (89)	52	75	94	6
IIU (89)	52	66	96	8
CTY (91)	55	80	97	8
MEI (239)	55	80	97	8
Average	52	73	91	7

b

VOR and Route Number	Freq CSIG/CCFP Occurence Low Conf/Low Cov	Freq CSIG/CCFP Occurence High Conf/Low Cov	Freq CSIG/CCFP Occurence High Conf/Med+ Cov	Percent CSIG with No CCFP (Missed)
VUZ (14)	51	71	90	8
SPA (14)	47	63	76	8
SZW (43)	57	77	93	8
VXV (43)	47	64	82	8
MCN (45)	50	59	75	7
BNA (45)	52	59	96	9
IRQ (52)	49	68	88	8
OTK (89)	54	80	81	6
IIU (89)	51	65	88	10
CTY (91)	60	81	91	10
MEI (239)	57	79	90	10
Average	53	71	87	8

Figure 7. Percent frequencies of correctly forecasted CCFP areas for the three confidence and coverage categories and also the frequency of time when CSIG(s) occurred with no CCFP area(s) occurring along the JAS for the ATL (Atlanta) VORTAC for June through August 2005-2006: a) 2-hourCCFP forecast b) 6-hourCCFP forecast.

a

	13z	15z	17z	19z	21z	23z	01z	03z
Freq CSIG/CCFP Occurence Low Conf/Low Cov	**28**	**44**	**57**	**70**	**63**	**51**	**40**	**46**
Freq CSIG /CCFP Occurence High Conf/Low Cov	**65**	**73**	**59**	**78**	**82**	**66**	**75**	**58**
Freq CSIG /CCFP Occurence High Conf/Med+ Cov	**100**	**67**	**74**	**91**	**100**	**84**	**93**	**95**
Percent CSIG with No CCFP (Missed)	**3**	**4**	**18**	**12**	**8**	**6**	**10**	**8**

b

	13z	15z	17z	19z	21z	23z	01z	03z
Freq CSIG Occurence Low Conf/Low Cov	**32**	**38**	**53**	**68**	**73**	**53**	**38**	**33**
Freq CSIG Occurence High Conf/Low Cov	**26**	**57**	**59**	**78**	**83**	**70**	**64**	**47**
Freq CSIG Occurence High Conf/Med+ Cov	**100**	**100**	**60**	**100**	**100**	**79**	**89**	**80**
Percent CSIG with No CCFP (Missed)	**3**	**5**	**18**	**19**	**12**	**9**	**6**	**8**

Figure 8. Percent frequencies of correctly forecasted CCFP areas for the three coverage and confidence categories and also the frequency of time when CSIG(s) occurred with no CCFP area(s) occurring along all JAS for each CCFP forecast hour except for the late night and early morning for the ATL (Atlanta) VORTAC for June through August 2005-2006: a) 2-hourCCFP forecast b) 6-hourCCFP forecast.

6. DISCUSSION

The purpose of this study was to introduce three new methodologies for assessing CCFP skill that may provide utility and were not found in prior literature searches. Previous CCFP skill studies have typically used NCWD data as the observed convection but for the purposes of this study CSIG was used. It is hoped these methodologies will provide insight into the CCFP forecast process and illustrate different ways that CCFP skill can be evaluated regardless of whether these new assessment methods are deemed appropriate.

The diurnal CCFP skill results help to explain the typical daily convective forecast cycle. These results may provide forecasters with information on how to improve their skill, such as CCFP areas generally need to be diminished more quickly during the evening in order to reduce the high false alarm rate at that time.

Most verification results in previous studies are based upon the phenomenon either being observed and/or forecast which helps to assess forecaster performance and workload. Little has been done to assess “forecaster work” in cases when the phenomenon is neither observed or forecast which also contributes to forecaster performance and workload. An attempt was made to quantify this difficult to measure component called “forecaster work” by trying to measure convective potential where convection was neither observed or forecasted (“null cases”). Convective potential was assessed by using the RUC K-Index initial analysis field for cases when neither a CSIG or CCFP areas were in effect. When convective potential was high in “null cases”, some forecaster work may have been necessary. When convective potential was low, little if any forecaster work may have been necessary in “null cases,” Using this assessment method for the summers of 2005 and 2006, some “forecaster work” may have been necessary almost half the time in “null cases”.

Segments of jet airways (JAS) going into four major airports were assessed for convective impacts for the most common types of CCFP coverage/confidence combinations. Almost all of the previously discussed methodologies assessed CCFP skill for each grid point, but this method assesses CCFP skill for the entire JAS, since a convective impact on any portion of a JAS may affect air traffic. Generally, as the coverage/confidence increased the CCFP forecast skill also increased but surprisingly very little improvement was seen with the 2-hour CCFP forecast versus the 6-hour CCFP forecast. Diurnal JAS CCFP skill trends were in general qualitative agreement with the diurnal CCFP skill earlier discussed.

Potential future work could include a further analysis of these diagnostic tools if they are found of value and a comparison of the advantages and/or disadvantages of using CSIG as observed convection versus other types such as NCWD.

7. ACKNOWLEDGEMENTS

The author wishes to thank Jack May of the AWC for the conceptual model of the JAS verification scheme. The author also wishes to thank Bruce Entwistle, Steve Silberberg, and Andy Fischer of the AWC; Dave Rodenhuis of the University of Victoria; and Jay Lawrimore of the National Climatic Data Center for reviewing this manuscript.

8. REFERENCES

Fahey, T. and D. Rodenhuis 2004: Continual Evolution of CCFP – User Needs For Extended Range Prediction. *11th Conference on Aviation, Range, and Aerospace Meteorology,* Hyannis, MA, Amer. Met. Soc.

Kay, M. P., J. Mahoney, and J. Hart. 2006: An Analysis of CCFP Forecast Performance for the 2005 Convective Season. *12th Conference on Aviation, Range, and Aerospace Meteorology,* Atlanta, GA, 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.

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

Torbert, D. and D. Rodenhuis 2004: Operational Feedback Reports to Providers of Aviation's Collaborative Convective Forecast Product. *11th Conference on Aviation, Range, and Aerospace Meteorology,* Hyannis, MA, Amer. Met. Soc.