8.3 AN EVALUATION OF THE PERFORMANCE OF THE CURRENT ICING POTENTIAL AT HIGH ALTITUDES

Michael B. Chapman*, Cory A. Wolff, Richard E. Bateman, and Ben C. Bernstein
National Center for Atmospheric Research

1. INTRODUCTION

The Current Icing Potential (CIP) is a diagnostic in-flight icing algorithm. Hourly diagnoses generated by CIP are based on a combination of observations along with 3-h forecasts from the 40-km Rapid Update Cycle (RUC) (Benjamin et al. 1999). Currently, CIP is operational at altitudes below 18 kft. Several recent formal and informal studies have been performed to verify CIP using observations of icing from pilot reports (PIREPs) and research aircraft (e.g. Brown et al. 1999; Bernstein et al. 2000). The results of these evaluations showed the efficiency of CIP in detecting icing conditions at altitudes below 18kft.

In an effort to examine the algorithms' performance at altitudes above 18 kft, a supplemental PIREP collection program was initiated in collaboration with Skywest Airlines from 13 – 25 August 2003. The purpose of this effort was to obtain a consistent observational data set that could be compared to and used in conjunction with regularly archived PIREPs over the same time period. These reports proved extremely valuable in assessing the accuracy of the regularly archived PIREPs, which have less certainty, especially with respect to the time and location of the report. An evaluation of CIP at high altitudes was also performed for the period 01 January – 31 March 2003 using regular PIREPs as observations in order to evaluate CIP's performance over several winter months. The study was designed to assess the overall performance of CIP during two different seasons at altitudes greater than 18 kft, and to examine in greater depth some specific cases that had varied verification results.

***Author contact information: Mike Chapman, NCAR, PO Box 3000, Boulder, CO 80307-3000
E-Mail:mchapman@ucar.edu

2. DATA SETS USED

2.1 CIP icing potential field

CIP is a physically-based, situational technique that produces an icing diagnosis by combining satellite, surface, radar, lightning and PIREP observations with fields from the 40-km RUC model (McDonough and Bernstein, 1999; Bernstein et al. 2004). The algorithm defines the icing field in terms of "potential," with floating point values from zero (no potential for icing) to 1.0 (icing very likely) (Fig. 1). Output is available at 1kft vertical increments.

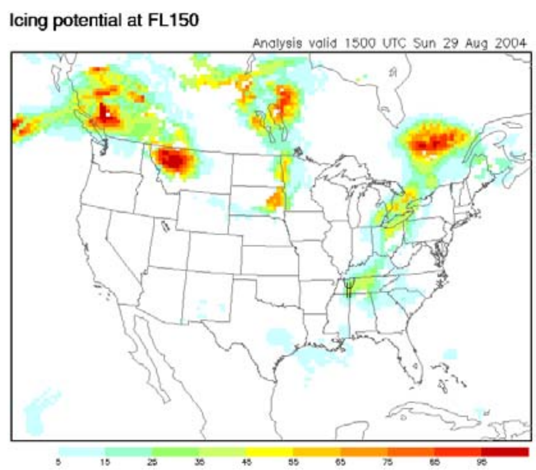

Fig 1. *Example of CIP from the Aviation Digital Data Service web page (http://adds.aviationweather.noaa.gov/)*

For the 13 -25 August 2003 Skywest verification 208 CIP files were compared to the PIREPs. A total of 1440 CIP files were used for the 01 January – 31 March 2003 verification.

2.2 Pilot reports (PIREPs)

PIREPs, which signify an observation of icing or lack thereof, are vital because they are the primary "ground truth" observations available to verify the presence or absence of icing at a specific location and time. However, standard PIREPs have several drawbacks as verification observations. For example, they typically do not provide information on a very fine spatial scale and they underreport the absence of icing. Also problematic is the small number of icing reports at altitudes greater than 18 kft. In order to compensate for these deficiencies, a supplemental PIREP collection program was initiated with the help of Skywest Airlines. The reports were taken specifically at altitudes above 18 kft regardless of icing conditions, with the time, location, and icing conditions recorded. In this study both standard and supplemental PIREPs were used in order to provide the maximum number of observations. Table 1 shows counts of the numbers of observations used in the verification for the 13 - 25 August 2003 and 01 January - 31 March 2003 time periods as well as the numbers of PIREPs available at lower altitudes (<18 kft)..

Table 1. PIREP counts for both (13 - 25 Aug 2003 and Winter 2003) time periods at altitudes above and below 18 kft including YES, NO and Moderate or greater (MOG).

PIREP type	*NO*	*YES*	*MOG*
Skywest	242	34	3
Standard(13-25 Aug 2003) <18 kft	2570	1829	102
Standard (13-25Aug 2003) >18 kft	34	219	61
Standard (Winter 2003) < 18 kft	12210	23566	7475
Standard (Winter 2003) > 18 kft	649	3279	782

3. VERIFICATION TECHNIQUE

Verification was accomplished by comparing the icing potential field to PIREPs of positive and negative icing at altitudes greater than 18,000 feet. The four grid points surrounding the PIREP and at 1-kft flight levels above and below it were examined. Since CIP incorporates information from PIREPs in the hour prior to the forecast time, this analysis only used PIREPs from the hour following the CIP valid time.

The methods utilized in the evaluation of CIP are based on standard techniques of forecast verification. They are described in greater detail in Brown et al. (1997). The icing forecast verification methodology treats icing forecasts and observations as Yes/No values. Brown et al. (1999) outlines how this method is extended to verify continuous, rather than binary fields. Icing diagnoses produced by CIP can be converted into a set of Yes/No values by applying a variety of thresholds. For example, applying a threshold of 0.20 to CIP diagnoses would lead to a Yes value for all grid points with an icing potential greater than or equal to 0.20 while each grid point with a value less than 0.20 would be assigned a No value. The verification methods are based on a two-by-two contingency table (Table 2). Each cell in this table contains a count of the number of times a particular forecast/observation pair occurred.

Table 2. Contingency table for YES/NO forecasts. Elements in cells are counts of forecast-observation pairs.

Forecast	*Observation*		*Total*
	YES	*NO*	
YES	YY	YN	YY+YN
NO	NY	NN	NY+NN
Total	YY+NY	YN+NN	YY+YN+ NN+NY

PODy and PODn are the primary verification statistics that are included in this evaluation. They are estimates of the proportions of Yes and No observations that are correctly diagnosed. Together, PODy and PODn measure the ability of the forecasts to discriminate between (or correctly categorize) Yes and No icing observations. PODy and PODn can be combined into an overall measure of this discrimination capability, the True Skill Statistic (TSS), also known as the Hanssen-Kuiper's discrimination statistic (e.g., Wilks 1995). Table 3 provides definitions and descriptions of these statistics.

Table 3. Verification Statistics used for the evaluation of CIP.

Statistic	*Definition*	*Description*
PODy	YY/(YY+NY)	Probability of detection of YES observations
PODn	NN/(NN+YN)	Probability of detection of NO observations
TSS	PODy+PODn-1	True Skill Statistic

The relationship between PODy and 1-PODn for different thresholds is the basis for the verification approach known as "Signal Detection Theory" (SDT). This relationship can be represented for a given algorithm by the curve joining the (1-PODN, PODy) points for different thresholds. The resulting curve is known as the "Relative Operating Characteristic" (ROC) curve in SDT. When PODy is plotted on the y-axis, the closer a given curve comes to the upper left corner, the better the forecast. The area under the curve is a measure of overall forecast skill and provides another measure that can be compared among forecast products. This measure is not dependent on the threshold used. A forecast with zero skill would have an ROC area of 0.5 (e.g., Mason 1982).

4. RESULTS

4.1 January 01 – March 31 2003

Figure 2 is an ROC diagram comparing the performance of the CIP at altitudes greater than 18 kft (CIP-WinterHigh) with the performance of the CIP at all levels (CIP-Winter), for the period 1 January to 31 March 2003. The plot shows the relatively good capability of CIP-Winter to discriminate between YES and NO observations. The results for CIP-WinterHigh are not as good, with an area under the curve of 0.59 for CIP-WinterHigh versus 0.76 for CIP-Winter.

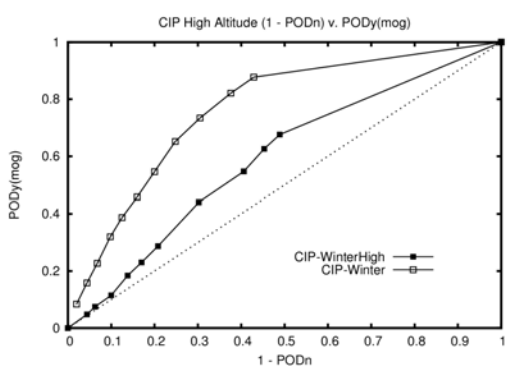

Figure 2. *ROC diagram of CIP (01 Jan – 31 Mar 2003) for all levels (CIP-Winter) and greater than 18,000 feet (CIP-WinterHigh).*

Table 4 compares the numerical verification results for the two winter evaluations. Based on varying thresholds the results in Table 4 show the limited ability of CIP-WinterHigh to capture YES(MOG) observations at even the lowest threshold (i.e., PODy(MOG) = 0.68 at a threshold of 0.01). Analysis of the TSS for both data sets shows the difference in skill from the CIP-Winter (TSS = 0.22 at a threshold of 0.7) compared to CIP-WinterHigh (TSS = 0.19 at a threshold of 0.01).

Table 4. PODy(MOG) and PODn statistics for CIP-WinterHigh and CIP-Winter.

Thresh -old	***CIP-WinterHigh***			***CIP-Winter***		
	PODy	**PODn**	**TSS**	**PODy**	**PODn**	**TSS**
0.01	**0.68**	0.51	**0.19**	0.88	0.57	0.45
0.1	0.63	0.55	0.18	0.82	0.62	0.44
0.2	0.55	0.59	0.14	0.73	0.70	0.43
0.3	0.44	0.70	0.14	0.65	0.75	0.40
0.4	.0.29	0.79	0.08	0.55	0.80	0.35
0.5	0.23	0.83	0.06	0.46	0.84	0.30
0.6	0.18	0.86	0.04	0.39	0.88	0.27
0.7	0.12	0.90	0.02	0.32	0.90	**0.22**
0.8	0.08	0.94	0.02	0.23	0.93	0.16
0.9	0.05	0.96	0.01	0.16	0.96	0.12

Table 5 shows a comparison of CIP-Winter-High PODy(MOG) statistics stratified by 3,000-ft layers to help identify specifically which layers are contributing most negatively to the overall verification results. The results show the greatest skill occurring in the 24-27 kft layer (TSS = 0.12 at a threshold = 0.30) followed by the 21-24 kft layer (TSS = 0.11 at a threshold = 0.02). The poorest results occurred on the 27-30 kft layer (TSS = 0.04 at the threshold = 0.01).

Table 5. TSS for CIP-WinterHigh in 3,000 foot layers.

Threshold	***18-21 kft***	***21-24 kft***	***24-27 kft***	***27-30 kft***
0.01	0.07	0.12	0.15	**0.04**
0.1	0.06	0.12	0.16	0.00
0.2	0.05	**0.11**	0.11	0.00
0.3	0.08	0.07	**0.12**	0.00
0.4	0.05	0.09	0.06	0.00
0.5	0.03	0.07	0.05	0.00
0.6	0.00	0.06	0.04	0.00
0.7	0.00	0.00	0.04	0.00
0.8	0.00	0.00	0.04	0.00
0.9	0.00	0.00	0.04	0.00

4.2 August 13 - 25 2003

The ROC diagram in Fig. 3 compares the CIP performance at high altitudes for the 13 – 25 August 2003 (CIP-August) period to the corresponding performance for winter 2003 (CIP-WinterHigh). This figure indicated that CIP performance at high altitudes was notably better during the August period than during the winter period, in terms of discriminating between Yes(MOG) and No observations of icing. Measurements of the areas under the ROC curves indicate CIP- August had considerably more skill than CIP-WinterHigh (0.76 vs. 0.59).

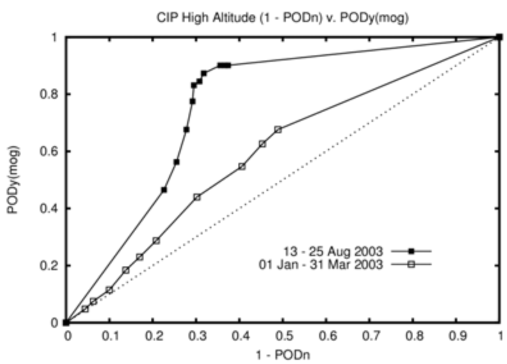

Figure 3*. ROC diagrams for CIP for two different time periods (13 – 25 Aug 2003 and 01 Jan - 31 Mar 200)3 for altitudes > 18 kft.*

Table 6 compares the statistics for the two time periods. The results in this table clearly demonstrate the better skill of CIP to capture the YES(MOG) PIREPs at high altitudes during the summer than during the winter. In particular, during August the CIP has large PODy(MOG) at high thresholds [PODy(MOG) > 0.78 and TSS > 0.49 at thresholds up to 0.6] whereas during winter large PODy(MOG) values are achieved only for much lower CIP thresholds [PODy(MOG) = 0.68 and TSS = 0.19 for a threshold of 0.01].

Table 6. PODy(MOG) and PODn statistics for CIP-August and CIP-WinterHigh.

Threshold	***CIP-August***			***CIP-WinterHigh***		
	PODy	PODn	TSS	PODy	PODn	TSS
0.01	0.90	0.63	0.53	**0.68**	0.51	**0.19**
0.1	0.90	0.63	0.53	0.63	0.55	0.18
0.2	0.90	0.64	0.54	0.55	0.59	0.14
0.3	0.87	0.68	0.55	0.44	0.70	0.14
0.4	0.85	0.69	0.54	0.29	0.79	0.08
0.5	0.83	0.71	0.54	0.23	0.83	0.06
0.6	**0.78**	0.71	**0.49**	0.18	0.86	0.04
0.7	0.68	0.72	0.40	0.12	0.90	0.02
0.8	0.56	0.75	0.31	0.08	0.94	0.02
0.9	0.47	0.77	0.24	0.05	0.96	0.01

Of the 105 observations of MOG icing between 24-30 kft during winter 2003, 61 were not associated with a CIP diagnosis of any icing potential. Figure 4 shows locations of the 61 missed observations of MOG icing. To gain insight into why these PIREPs were missed, a case study was undertaken. An additional case of summertime high altitude icing was also investigated.

Figure 4. *Map of missed MOG icing observations for CIP-WinterHigh over 24-30 kft layer.*

5. CASE STUDIES

Two cases were selected for further study: one from the winter 2003 time period from Fig. 4 where the verification results were poor above 24 kft, and a second case from August 2003 where CIP was successful in diagnosing icing at altitudes greater than 24kft. The case studies include an examination of the meteorological conditions (satellite, soundings, surface, and upper-air synoptic charts) as well as the RUC model temperature and relative humidity forecasts that were available for CIP to use.

5.1 March 12th 2003, 16Z

Two PIREPs near Seattle, WA, one at 24 kft and the other at 28 kft, were recorded for this case at 1640Z and 1651Z. Each PIREP reported moderate icing. The 12Z sounding from Quillayute, WA (KUIL; Fig. 5) on 12 March 2003 shows that deep saturated conditions were present up to very cold temperatures (T<-65°C). These conditions indicate that the ice process most likely was quite active, which resulted in at least partial glaciation of the clouds. The area circled in red on Fig. 5 indicates the approximate altitude range of interest for this case (24 – 28kft). The observed temperatures for this layer were very low (-34°C to -40°C).

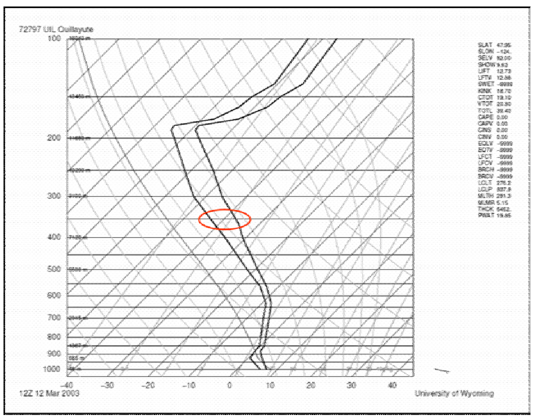

Figure 5. Skew-t diagram of the 12Z 13 March 2003 KUIL sounding

The corresponding temperatures (-31°C to -40°C; Fig. 6a) and relative humidities (RH; 75 – 68%; Fig. 6b) from the RUC model for this layer match the KUIL sounding quite well. Note also that CIP showed a fairly deep layer of icing potential (6-20kft.; Fig 6C).

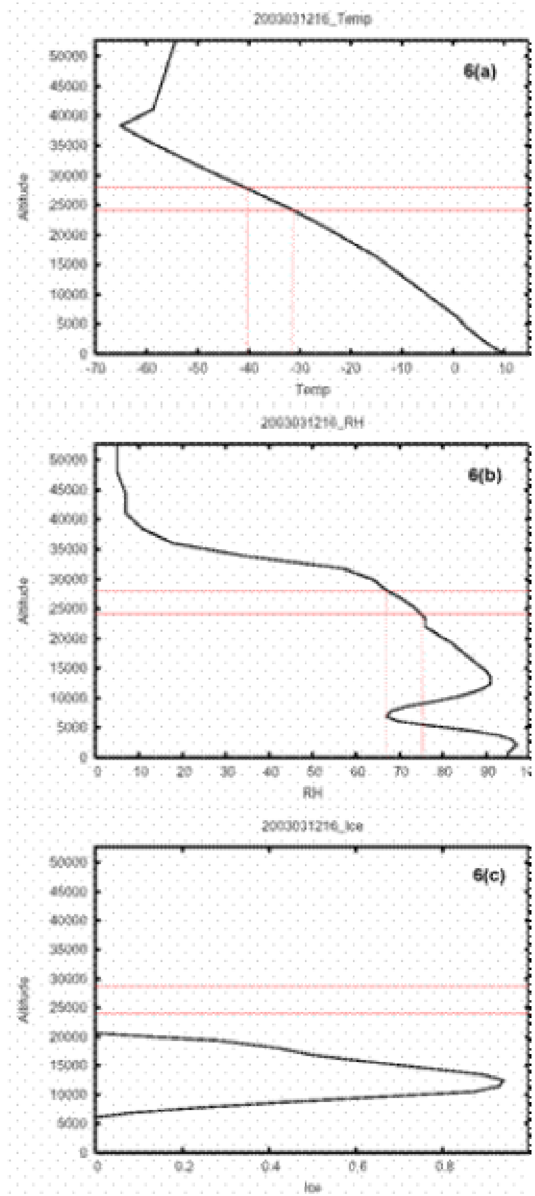

Figure 6. Plot of altitude v. (a) RUC temperature, (b) RUC RH, and (c) CIP Icing Potential for 12 March 2003 near MOG Icing PIREPs valid for 16Z. .

An infrared satellite image from 1545Z on 13 March 2003 shows cold cloud tops (Fig. 7), which are consistent with both the observed and forecast soundings. The 1545Z image was used because of its timing relative to the PIREPs and the 16Z CIP.

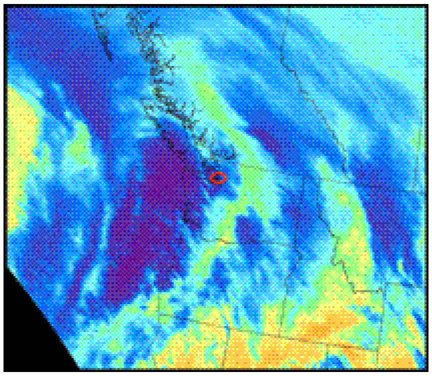

Figure 7. Infrared satellite image for 12 March 2003 valid at 1545 Z. PIREP locations were located in the red oval.

The Seattle radar (KATX; Fig. 8) was examined and showed several areas of deep stratiform precipitation over the area as well. The 1646Z reflectivity was used because of its proximity in time to the two PIREPs (1640Z and 1651Z).

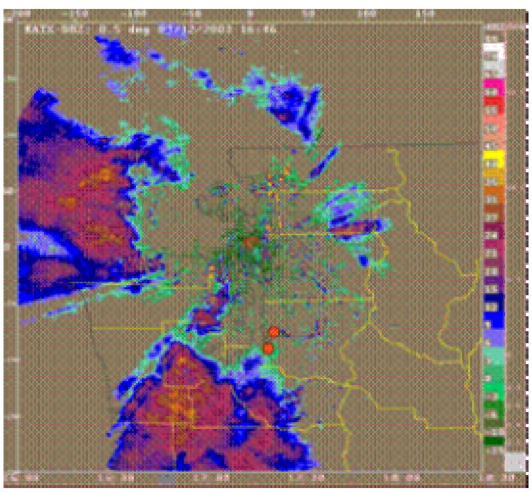

Figure 8. 12 March 2003, 1646Z reflectivity image for KATX. PIREP locations are symbolized with red dots.

As stated earlier, icing is rare at T<-25°C, and especially so as temperature approaches -40°C. The CIP temperature interest map (Fig. 9) reflects this tendency, with interest = 0 at T<-25°C. Because the operational version of CIP running at the time of this case used the "Old" interest map the interest value was 0 for even higher temperatures (T<-22°C). This interest mapping led to the CIP diagnosis of zero icing potential for the two high altitude PIREPs near Seattle on this day. Note that if the new temperature interest mapping had been used, the PIREPs still would have been missed by CIP.

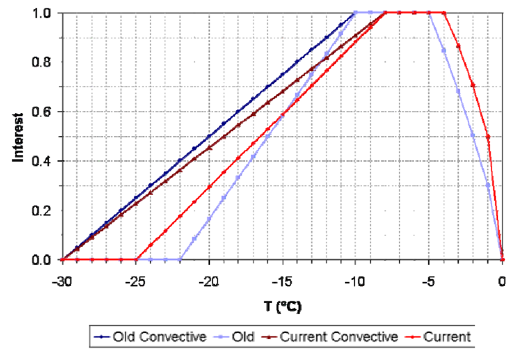

Figure 9. Old and new temperature interest maps used in CIP for diagnosis of the potential for icing. The convective map is used in the presence of lightning observations.

5.2 August 18th 2003, 00Z

For this case, a moderate icing PIREP was recorded over Eastern Colorado at 27 kft. The 0Z, 18 August 2003 sounding taken at Denver, CO (KDNR; Fig. 10) shows a temperature of approximately -22°C and high humidities at the level of the PIREP.

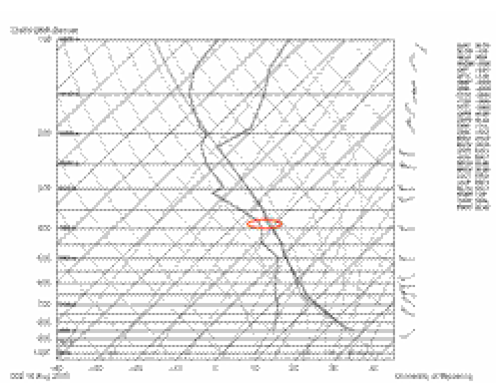

Figure 10. Skew-t diagram of the 0Z, 18 August 2003 KDNR sounding.

The RUC model temperature value at the location and level of the PIREP is similar to the observed value at that level (-23.3°C; Fig. 11a) but the RUC RH value is a bit low (39%; Fig. 11b) compared to the actual sounding value.

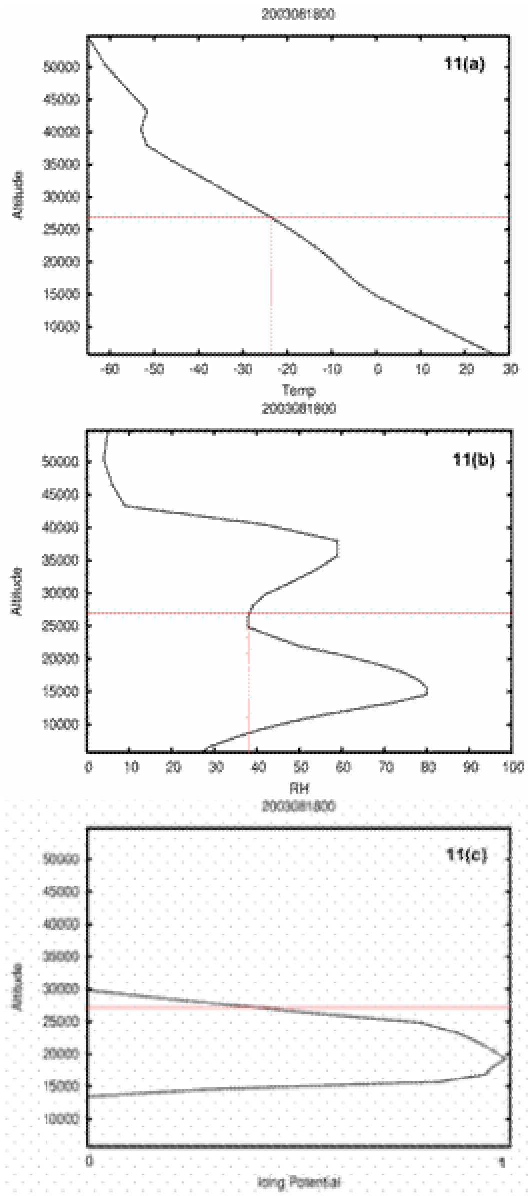

Figure 11. Plot of Altitude v. (a) RUC Temperature, (b), RUC RH, and (c) CIP Icing Potential for 18 August 2003 near MOG icing PIREPs.

Lightning was also detected within 49 km of the PIREP location, which suggests that the clouds may have been convective. The temperature interest map for CIP changes when convection is present (Fig. 7; red line), which allows icing to be diagnosed at temperatures down to -30°C. This is evident in figure 11c where icing potential is diagnosed all the way up to 30 kft. Therefore CIP was successful in capturing the MOG icing PIREP at 27 kft.

Reflectivity observations from the radar at the Denver, CO – Front Range Airport (KFTG; Fig. 20) indicate that widespread small-scale convective precipitation was present in the region.

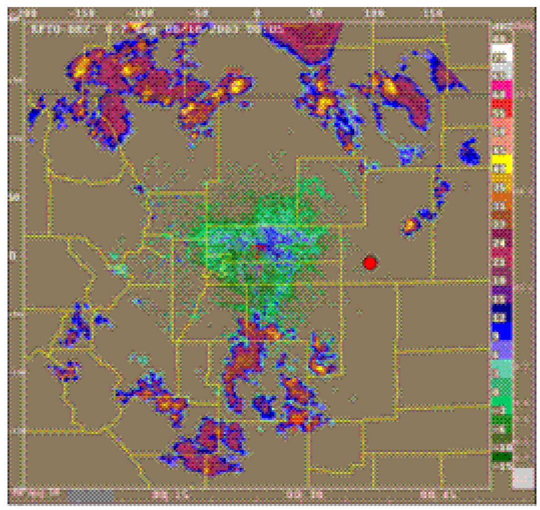

Figure 12. 0005Z reflectivity on 18 Aug 2003 from (KFTG) radar. The location of the PIREP is marked with a red dot.

6. CONCLUSIONS

6.1 Verification results

The statistical verification results indicate that the CIP performed well at levels above 18 kft during the 13 – 25 August 2003 time period, based on both the standard PIREPs and the Skywest Airlines supplemental PIREPs. The Skywest observations provided valuable information, particularly regarding no-icing conditions. One possible explanation for the success of the CIP during this season may be attributed to the warmer temperatures at higher altitudes during that season (Bernstein et al., 2002). Another possible explanation may be that more of the icing events during the summer are convective, which makes it possible to diagnose icing conditions at lower temperatures.

The CIP's performance during the 01 January – 31 March 2003 time period was not as good as in the summer period. When compared to CIP performance for all levels during this period, the results for high altitudes indicated notably less skill. One possible reason for this poor performance may be related to the form of the CIP temperature interest maps, which only provide a possibility for icing down to -22°C. Another contributing factor may be the lack of pilot reports at the higher altitudes (>24 kft), where only 15% of the observations of MOG icing were reported in the 18 – 30kft layer. CIP diagnosed no icing conditions for 61 of the 105 MOG icing reports between 24 and 30 kft.

6.2 Case studies

The case study from 12 March 2003 is an example of a time when MOG icing observations were present but no icing potential was diagnosed. It was apparent that clouds were present but that the layers were colder than CIP would consider having icing conditions in the absence of deep convection, which was not present. The temperatures were extremely cold in this case (T<-32°C); thus, the introduction of the new temperature interest maps would not have helped in this case. However, diagnosis of icing conditions at these cold temperatures would lead to gross over-warning.

For the 18 August 2003 case study, CIP was able to capture the MOG icing reports at high altitudes. The clouds in this case were convective, which led CIP to use its convective temperature interest map. This mapping allows icing to be diagnosed at colder temperatures, which in turn made it possible to capture the MOG icing observations in this case.

7. FUTURE WORK

In the future, a verification study will be undertaken using the same observations with the latest version of CIP, which uses the new interest maps and the 20-km RUC. This evaluation will provide a valuable comparison to the results from this study. As part of this new evaluation, the case studies will be repeated using the new version of CIP. In addition, more cases that are meteorologically and regionally similar will provide more confidence in our assessment of the CIP's performance at high altitudes.

8. ACKNOWLEDGEMENTS

The authors would like to thank the pilots of Skywest Airlines for providing a valuable and high quality data set on very short notice. Thanks also to Anne Holmes for her programming support as well as Patrick Boylan for his support.

This research is in response to requirements and funding by the Federal Aviation Administration. The views expressed are those of the authors and do not necessarily represent the official policy and position of the U.S. Government.

9. REFERENCES

Benjamin, S.J., J.M. Brown, K.J. Brundage, D. Kim, B. Schwartz, T. Smirnova, and T.L. Smith, 1999: Aviation forecasts from the RUC-2. *Preprints, 8th Conference on Aviation, Range, and Aerospace Meteorology*, Dallas, TX. 10-15 January, American Meteorological Society (Boston), 486-490.

Bernstein, B.C., F. McDonough, M.K. Politovich, and B.G. Brown, 2000: A research aircraft verification of the Integrated Icing Diagnostic Algorithm (IIDA). *Preprints, 9th Conf. on Aviation, Range and Aerospace Meteorology*, 11-15 Sept., Orlando, FL, 280-285.

Bernstein, B.C., and F. McDonough, 2002: An Inferred Icing Climatology – Part II: Applying a version of IIDA to 14-years of coincident soundings and surface observations. *Preprints, 10th Conf. on Aviation, Range and Aerospace Meteorology*, Portland, OR, 13-16 May, American Meteorological Society (Boston) J21-24.

Bernstein, B.C., F. McDonough, M.K. Politovich, B.G. Brown, T.P. Ratvasky, D.R. Miller, C.A. Wolff, and G. Cunning, 2004: Current Icing Potential (CIP): Algorithm Description and Comparison with Aircraft Observations. Conditionally accepted to Journal of Applied Meteorology.

Brown, B.G., G. Thompson, R.T. Bruintjes, R. Bullock, and T. Kane, 1997: Intercomparison of in-flight icing algorithms. Part II: Statistical Verification Results. Weather and Forecasting, 12, 890-914.

Brown, B.G., T.L. Kane, R. Bullock, and M.K. Politovich, 1999: Evidence of improvements in the quality of in-flight icing algorithms. *Preprints, 8th Conf on Aviation, Range and Aerospace Meteorology,* Dallas, TX, 10-15 January, American Meteorological Society (Boston), 48-52.

Mason, I., 1982: A model for assessment of weather forecasts. *Australian Meteorological Magazine*, **30**, 291-303.

McDonough, F. and B.C. Bernstein, 1999: Combining satellite, radar, and surface observations with model data to create a better aircraft icing diagnosis. *Preprints, 8th Conf. on Aviation, Range and Aerospace Meteorology, 10-15 Jan., Dallas TX. 467-471*.

Wilks, D.S., 1995: *Statistical Methods in the* Atmospheric *Sciences*. Academic Press, 467 pp.