10.2 COMPARISON OF IN-SITU, MODEL AND GROUND BASED IN-FLIGHT ICING SEVERITY

Christopher J. Johnston [a]*, David J. Serke [a], Daniel R. Adriaansen [a], Andrew L. Reehorst [b], Marcia K. Politovich [a], Cory A. Wolff [a] and Frank McDonough [a]

[a] NCAR- National Center for Atmospheric Research
Research Applications Laboratory
Boulder, Colorado

[b] NASA Glenn Research Center
Cleveland, Ohio

1. INTRODUCTION

In-flight icing is a significant hazard for the aviation industry. It occurs when supercooled liquid water (SLW) comes in contact with, and freezes to, the leading surfaces of an aircraft. This can significantly alter the aircraft's aerodynamic properties by increasing the amount of drag on the aircraft, and reducing the lift. Since practical airborne remote detection hardware has not yet been developed, a ground-based detection system that can provide information to all aircraft entering and departing a terminal area (Fig. 1) is a key element in facilitating icing avoidance (Serke et al., 2010).

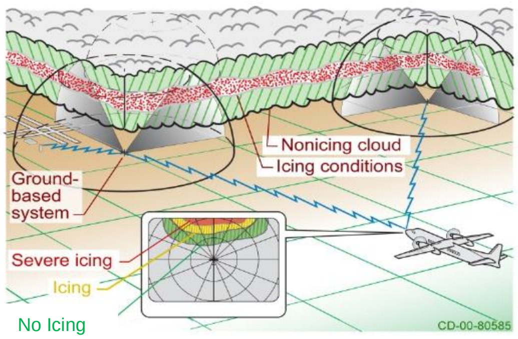

***Figure 1**. NIRSS in-flight icing detection concept.*

Currently there are two systems that are being developed for the detection of in-flight icing. The first detection system is the NASA Icing Remote Sensing System (NIRSS) a testbed that integrates three vertically pointing sensors; a Vaisala Laser Ceilometer, a Metek K_a-band radar, and a Radiometrics Corporation 23-channel radiometer (Fig. 2), (Reehorst et al., 2006).

**Corresponding author address: Christopher J. Johnston, National Center for Atmospheric Research, PO Box 3000, Boulder, CO 80307, chrisj@ucar.edu*

The multichannel microwave radiometer has the ability to derive integrated liquid water (ILW), atmospheric water vapor and temperature profiles (Solheim et al., 1998). A Vaisala laser ceilometer is used to define cloud base heights, and a Metek K_a-band radar is used to delineate cloud top and base heights. NIRSS combines the ILW, radar reflectivity, temperature profile, and cloud top and base heights to determine the presence of in-flight icing conditions in the atmosphere.

***Figure 2**. Image of the NIRSS hardware located at the NASA Glenn Research Center in Cleveland, Ohio.*

The second in-flight icing detection system is the Current Icing Product (CIP) which was developed at the National Center for Atmospheric Research. This system combines visible and infrared satellite imagery, radar reflectivity, lightning observations, Pilot Reports (PIREPs) and standard ground-based weather observations with numerical model output to produce a gridded, hourly, three dimensional representation of icing probability and severity (Bernstein et al., 2005).

Each horizontal grid point of CIP is based on a 20 km by 20 km Rapid Update Cycle (RUC) model grid point. First developed during the winter of 1997/98, CIP became an operational National Weather Service product in 2002.

Icing-related PIREPs are voluntary reports made by pilots to report on the presence or absence of in-flight icing conditions and other weather-related conditions. Both a subjective icing severity and icing type (rime, clear or mixed) are included. PIREP reports of no icing are useful as well, since the absence of icing is important information. The shortcomings of PIREPs are well documented and include non-uniformity in time or space and contamination by errors in location, altitude and time (Brown et al., 1997; Kelsch and Wharton, 1996). PIREPs can sometimes be inaccurate due to time lags before the pilot reports the observed icing condition, and whether he or she reports the correct altitude and location. The reported severity is also somewhat subjective as it can vary based on aircraft type, phase of flight, and pilot experience. Nevertheless, PIREPs are our only means of in-situ diagnoses of actual atmospheric conditions encountered by pilots and their aircraft in the absence of expensive icing research flights or specially instrumented fleet aircraft. The objective of this study is to examine how the testbed NIRSS icing severity product and the operational CIP severity product compare to PIREPs of icing severity, and how the NIRSS and CIP compare to each other.

2. METHODOLOGY

A three-year database of CIP, NIRSS and PIREP data was compiled focusing on winter periods from early November 2008 to late March 2010. During these three winter seasons, 917 icing PIREPS were collected within 40 km of the NIRSS system located in Cleveland, Ohio. CIP icing severity output from the nearest RUC gridpoint to the NIRSS location was archived and icing severity values were extracted at the time and height of each icing PIREP. A similar process was conducted for NIRSS severity output. If there was no icing data at the exact PIREP altitude, a vertical search was performed for the nearest altitude with icing severity data. Once this temporal and spatial matching was completed for all PIREPs, a statistical comparison was begun. Analysis occurred from the ground level to ~ 30,000 feet (or 9,144 meters). For this study, a PIREP reported over a range of heights is treated as multiple PIREPs spread over 1000 foot increments (Wolff et al., 2010).

3. ANALYSIS AND DISCUSSION

3.1 Case Study Comparison - December 15th, 2009

An example icing case study is presented here to illustrate how the comparison of NIRSS, CIP and PIREP severity looked for a single event. In the next section, statistics for three years of such cases are discussed.

On December 15th, 2009 at 0000 UTC, a surface low was dominant over the central Great Lakes region (Fig. 3). The warm front extended from the southeast portion of Lake Huron eastward into southeastern New York. The cold front was oriented from north central Ohio through southwest Ohio. The cloud top temperatures over the Cleveland area were between -5°C and -15°C (color scale). This temperature range has been shown in previous research to be conducive to supercooled liquid water (Rogers and Floyd, 1989).

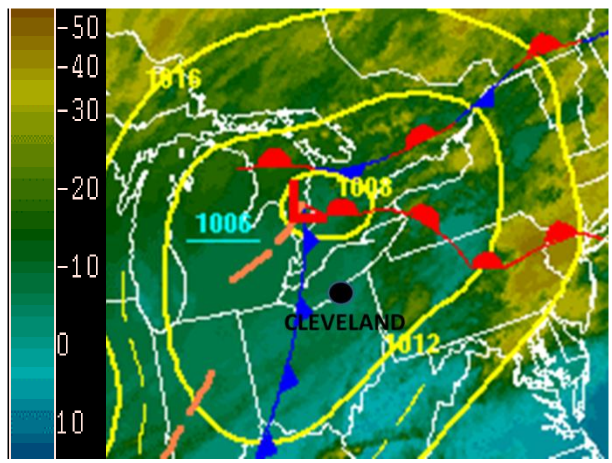

***Figure 3**. 0000 UTC surface pressure (yellow lines, [mb]), cloud top temperature (color bar, [°C]) and frontal analysis for December 15th, 2009.*

At approximately 0300 UTC, the cold front passed through Cleveland, Ohio. In the hours following the cold frontal passage, drizzle and rain fell over the metropolitan area, which changed to snow by 1500 UTC.

Figure 4 shows time versus height of icing severity from NIRSS (top), PIREPs (top numerals) and CIP (bottom) for December 15th, 2009. A zero to eight scale for icing severity was used for these plots, and the comparisons throughout the rest of this study where zero is no icing, one is trace amount of icing, two to three is light, four to five is moderate and six to eight is heavy.

Twenty-eight positive icing PIREPs were recorded between 1100 and 2200 UTC around Cleveland. In addition, three negative PIREPS were recorded during this time period. NIRSS diagnosed significant icing between 0900 and 2400 UTC from 1 - 6 kft AGL. CIP diagnosed icing from 0900 to 2400 UTC as well, from roughly 1 - 9 kft AGL. For this case, the temporal variability and the magnitude of the severities generally match between the two products despite the fact that CIP has a one-hour time resolution and NIRSS has one-minute resolution. CIP has a conservative cloud top and base estimate scheme (done purposely to insure thorough warnings).

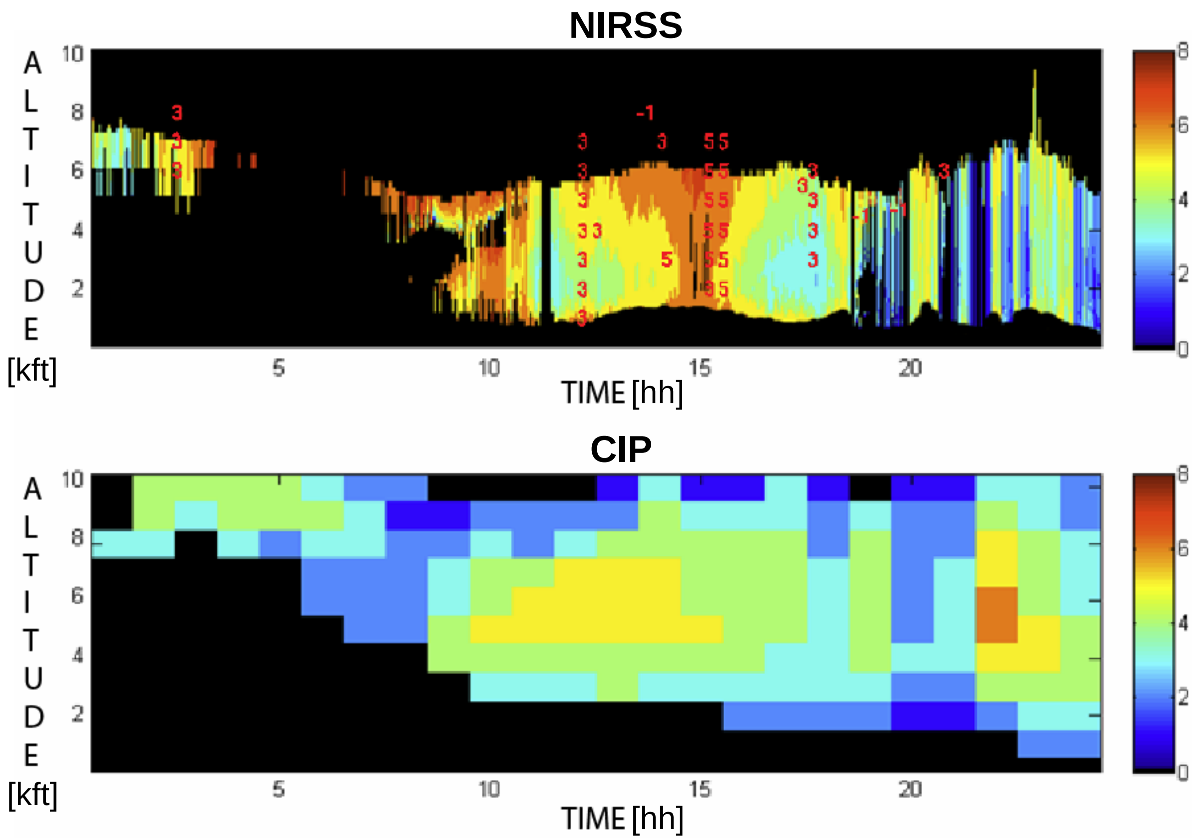

Figure 4*. Time [hh] versus height [kft] plots of NIRSS (top, color scale), PIREP (top, red numerals) and CIP (bottom, color scale) icing severity from December 15th, 2009 from 0000 UTC to 2400 UTC.*

3.2 Three-Year Archive Comparison

In the previous section we explored the comparison of icing products for a day-long icing case. This section will be a statistical intercomparison of all three icing detection methods for the full three-year study period. Similar to the case study presented above, the closest NIRSS and CIP icing severity measurements were found to each of the 917 PIREPs within 40 km of the NIRSS location. The matched severity categories were plotted in Figures 5 - 7.

NIRSS versus PIREP severity is shown in Figure 5, with one-to-one severity correlation bins highlighted in orange. The numbers in each bin represent the total number of icing severity matchups recorded for the three-year period. A linear best-fit line is overlain in blue. Taking the square root of the resulting R-squared value gives NIRSS a severity category correlation coefficient of 0.35 to the PIREP severity category. NIRSS appears to do well locating negative PIREPs, as well as finding moderate icing PIREPs. Very few severe or heavy PIREPs were reported during this time period. There are a significant number of positive PIREPs that NIRSS identifies as negative severity, possibly due to the high time resolution of NIRSS's ILW algorithm when viewing localized SLW cases.

CIP versus PIREP severity is shown in Figure 6. A linear best-fit line is again overlain in blue, with CIP having a severity category correlation of 0.21 to the PIREP severity category. CIP and NIRSS both do well at finding PIREPs from light to moderate values. CIP seems to correctly identify a much smaller fraction of negative PIREPs than NIRSS. This could be a result of the conservative cloud base and top diagnosis in CIP.

NIRSS									
8	0	0	0	0	0	0	0	0	0
7	0	0	0	13	0	0	0	0	0
6	1	1	0	18	4	29	0	0	0
5	1	7	0	28	8	23	0	0	0
4	1	5	0	46	7	15	0	0	0
3	6	18	0	95	18	22	0	0	0
2	9	20	0	71	3	28	0	0	1
1	28	10	1	71	12	14	0	0	2
0	115	19	0	93	16	34	4	0	0
PIREP	0	1	2	3	4	5	6	7	8

Figure 5. *Overall PIREP severity versus NIRSS severity.*

CIP									
8	0	0	0	0	0	0	0	0	0
7	1	1	0	5	0	4	0	0	0
6	4	0	0	1	1	1	0	0	0
5	5	10	0	21	7	13	0	0	3
4	12	13	0	78	20	43	0	0	0
3	22	14	0	94	20	35	0	0	0
2	32	12	1	97	13	40	0	0	0
1	38	23	0	92	6	7	0	0	0
0	47	4	0	47	0	22	4	0	0
PIREP	0	1	2	3	4	5	6	7	8

Figure 6. *Overall PIREP severity versus CIP severity.*

NIRSS versus CIP severity is shown in Figure 7. The operational CIP product is treated as 'truth' in this comparison, as NIRSS is still a testbed. A linear best-fit line is again shown in blue, with NIRSS having a severity category correlation of 0.18 to the CIP severity category. There seems to be a significant spread in the collocated severity values between the two products during the three-year study period. This spread is likely due to the difference in temporal resolution of the products, and the fact that the two products arrive at hazard estimates based on different input datasets.

NIRSS									
8	0	0	0	0	0	0	0	0	0
7	2	6	5	0	0	0	0	0	0
6	5	4	8	11	15	7	1	2	0
5	12	7	7	15	14	8	1	3	0
4	7	9	13	17	21	6	1	0	0
3	9	32	36	37	35	9	0	1	0
2	14	18	41	30	23	6	0	0	0
1	3	29	42	27	20	10	3	4	0
0	76	61	43	0	38	13	1	1	0
CIP	0	1	2	3	4	5	6	7	8

Figure 7. *Overall CIP versus NIRSS severity.*

Another useful statistic is the probability that each product will detect negative and positive icing PIREPs, or the *Probability of Detection* (POD). These statistics are termed POD_n and POD_y, respectively. To get POD_n, the fraction of negative icing PIREPs that the respective product identifies as negative icing is determined and then divided by the total cases where PIREP severity is equal to zero (Eqn. 1).

$$POD_n = \frac{\text{(Total Cases product icing severity = 0 when PIREP severity = 0)}}{\text{(Total cases PIREP icing severity = 0)}} \quad (1)$$

Similarly, POD_y is the fraction of positive icing PIREPs of any category (one through eight) that the respective product identifies as positive icing divided by the total cases where PIREP severity is greater than zero (Eqn. 2).

$$POD_y = \frac{\text{(Total cases product icing severity > 0 when PIREP severity >0)}}{\text{(Total cases PIREP icing severity >0)}} \quad (2)$$

For the three-year study period, POD_y and POD_n were calculated for the product comparisons shown in Figures 5 - 7. The results are shown in Table 1.

	N vs. P	C vs. P	N vs. C
POD_y	0.78	0.90	0.79
POD_n	0.71	0.29	0.59

Table 1. *Overall POD_n and POD_y statistics, N vs. P (NIRSS vs. PIREPS), C vs. P (CIP vs. PIREPS), N vs. C (NIRSS vs. CIP).*

NIRSS detected greater than 70% of both positive and negative PIREPs. CIP detected 90% of positive PIREPs but only 29% negative PIREPs. The percentage of the time averaged vertical profile that a product has identified a positive severity value is termed the warning volume. The warning volume for this study is calculated from the surface to the average height of the tropopause. A successful product must find an optimal balance between POD yes and no and warning volume because it would not be very useful for a product to have a POD_y of 1.0 (perfect icing detection) if the entire column is warned on at all times. Ideally, a product would have a maximized POD_y and POD_n with a minimized warning volume. For the three-year study period, NIRSS had a mean warning volume of 13%, and CIP had a mean warning volume of 34%. CIP detected 10% more positive PIREPs in over twice the warning volume. Furthermore, CIP's high warning volume causes it to classify regions as positive icing where they should be devoid of icing, based on negative PIREPs.

4. SUMMARY

In-flight icing detection is crucial to achieving a high safety standard for the national fleet of commercial and general aviation aircraft. In this study a comparison was done showing the quantitative severity categories of negative and positive icing PIREPs to the quantitative icing severity derived from the prototype NIRSS icing detection algorithm and operational CIP icing algorithm. An icing case study from December 15th, 2009 over the NIRSS location in Cleveland, Ohio was discussed to illustrate how the PIREP and icing product severities were compared. A statistical analysis over the full three-year study period found that NIRSS detected in-flight icing and negative icing at least as good as CIP when compared to all PIREPs within a 40 km radius. NIRSS detected almost 80% of positive PIREPs and over 70% of negative PIREPs in a relatively smaller warning volume. CIP detected slightly more positive PIREPs than NIRSS but did fairly poor in detecting negative PIREPs. This occurred in a warning volume over twice the percent of NIRSS's warning volume. CIP did very well at detecting positive PIREPs. NIRSS displayed respectable probabilities of icing detection with lower warning volumes than CIP. This is due to NIRSS having a higher time resolution and utilizing physically based vertical profiles of ILW, temperature and radar reflectivity. Therefore, the NIRSS testbed in-flight icing severity product seems to be at least as good as CIP. A shortfall of NIRSS is that it currently lacks volumetric scanning capability. This is being addressed by the addition of a one degree beamwidth multichannel scanning radiometer (Serke et al., 2010). Future work with NIRSS will include exploring Doppler fall velocities to detect possible freezing drizzle and freezing rain, and comparing NIRSS hazard detection to the polarimetric data from future upgraded NEXRAD.

5. ACKNOWLEDGEMENTS

This research is supported by the NASA Aviation Safety Program under the Atmospheric Environment Safety Technologies (AEST) project. The views expressed are those of the authors and do not necessarily represent the official policy or position of the NASA.

6. REFERENCES

Bernstein, B., McDonough, F., Politovich, M., Brown, B., Ratvasky, T., Miller, D., Wolff, C., and Cunning, G., "Current Icing Potential: Algorithm description and comparison to aircraft observations", Journal of Applied Meteorology, 44, pp. 969-986, 2005.

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. *Wea. Forecasting*, **12,** 890-914

Kelsch, M., and L. Wharton, 1996: Comparing PIREPs with NAWAU turbulence and icing forecasts: Issues and results. *Wea. Forecasting*, 11, 385-390

Reehorst, A., Politovich, M. K., Zednik, S., Isaac, G. A., & Cober, S., 2006: Progress in the Development of Practical Remote Detection of Icing Conditions (NASA/TM 2006-214242). NASA.

Rogers, R. R., M. K. Floyd: A short course in cloud physics. Pergamon Press, Oxford England.

Serke, D. J., Beaty, P., Reehorst, A. L., Kennedy, P., Solheim, F., Ware, R., Politovich, M. K., Brunkow, D. and Bowie, R. "A New Narrowbeam Multi-Frequency Scanning Radiometer and its Application to In-Flight Icing Detection," *AMS 14th Conference on Aviation, Range, and Aerospace Meteorology (ARAM)*, 2010.

Solheim, F., Godwin, J., Westwater, E., Han, Y., Keihm, S., Marsh, K., and Ware, R., “Radiometric profiling of temperature, water vapor and cloud liquid water using various inversion methods,” Radio Sci., 33, pp. 393-404, 1998.

Wolff, C. A., McDonough, F. “A Comparison of WRF-RR and RUC Forecasts of Aircraft Icing Conditions,” *AMS 14th Conference on Aviation, Range, and Aerospace Meteorology (ARAM),* 2010