PROGRESS IN THE DEVELOPMENT OF PRACTICAL REMOTE DETECTION OF ICING CONDITIONS

Andrew Reehorst
National Aeronautics and Space Administration
Glenn Research Center
Cleveland, Ohio

Marcia K. Politovich and Stephan Zednik
National Center for Atmospheric Research
Boulder, Colorado

George A. Isaac and Stewart Cober
Meteorological Service of Canada
Toronto, Ontario, Canada

1. INTRODUCTION

The NASA Icing Remote Sensing System (NIRSS) has been under definition and development at NASA Glenn Research Center since 1997. The goal of this development activity is to produce and demonstrate the required sensing and data processing technologies required to accurately remotely detect and measure icing conditions aloft. As part of that effort NASA has teamed with NCAR to develop software to fuse data from multiple instruments into a single detected icing condition product. The multiple instrument approach, which is the current emphasis of this activity, utilizes a X-band vertical staring radar, a multi-frequency microwave, and a lidar ceilometer. The radar data determine cloud boundaries, the radiometer determines the sub-freezing temperature heights and total liquid water content, and the ceilometer refines the lower cloud boundary. Data is post-processed with a LabVIEW program with a resultant supercooled liquid water profile and aircraft hazard depiction.

Additional ground-based, remotely-sensed measurements and in-situ measurements from research aircraft were gathered during the international 2003-2004 Alliance Icing Research Study (AIRS II). Comparisons between the remote sensing system's fused icing product and the aircraft measurements are reviewed here. While there are areas where improvement can be made, the cases examined suggest that the fused sensor remote sensing technique appears to be a valid approach.

2. DESCRIPTION OF SENSORS

The NIRSS is made up of three sensor components: a radar; a microwave radiometer; and a ceilometer (Fig. 1, a thorough description of the system is provided by Reehorst et al., 2001). The radar used for the NIRSS during winter 2003-2004 was a modified Honeywell WU-870 airborne X-band radar (Reehorst and Koenig, 2001). The radar provides reflectivity measurements that are used to define cloud boundaries. The microwave radiometer is a Radiometrics, Inc. TP/WVP 3000 Temperature and Water Vapor Profiler (Solheim et al., 1998). Among other parameters, this instrument provides a temperature profile and integrated

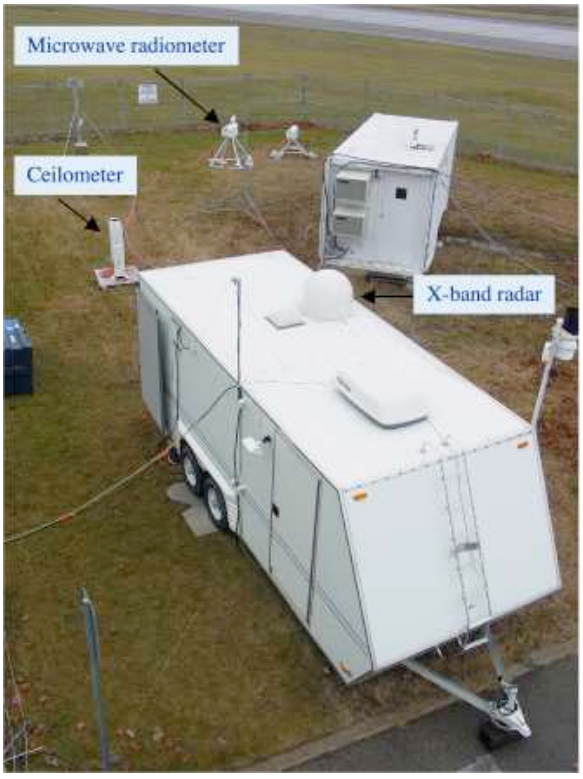

Figure 1: NIRSS components as configured during AIRS II.

liquid water. Finally, the ceilometer is a standard Vaisala CT25K Laser Ceilometer, which is used to refine the definition of the lower cloud boundary since it is less susceptible to precipitation than the radar.

3. DESCRIPTION OF SOFTWARE

The measurements from the three instruments are fused to produce a single indication of aircraft icing hazard. To date, two forms of fusion processing were utilized. The first generation (Gen1) fusion technique was quite simple. The radar reflectivity data was used to define the boundaries of cloud layers. The lower boundary of the lowest cloud layer is refined with the ceilometer data to correct for precipitation effects or for the close-range radar blind spot and side-lobe noise. Once the cloud boundaries were defined, the temperature profile was used to determine the portion of the clouds likely to be supercooled liquid. The test for supercooled liquid cloud was that the temperature must be below 0ºC and above -20ºC. A further test determined the range of all cloud that was liquid (both above and below 0ºC). The integrated liquid water measured by the radiometer was then evenly distributed over the liquid cloud region to determine the cloud liquid water content (LWC). The LWC cloud boundaries were then further limited by the range of supercooled cloud. If the resultant supercooled liquid cloud had an LWC greater than 0.1 gm^{-3}, then it was defined as being an aircraft icing hazard.

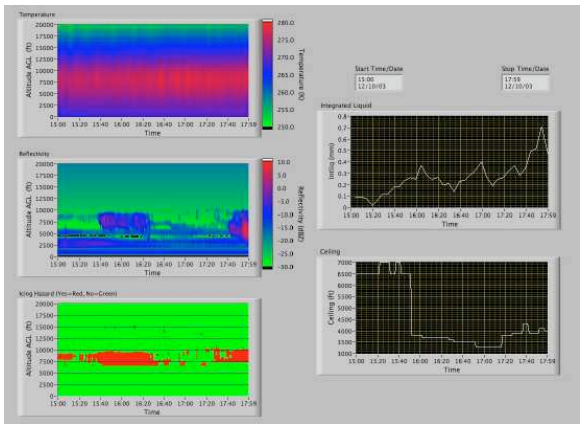

Figure 2: NIRSS Gen1 graphical interface.

The output graphic (Fig. 2) included a temperature profile history (top left), cloud reflectivity history (center left), integrated cloud liquid (liquid water path) history (top right), ceiling history (bottom right), and the resultant icing hazard profile history (bottom left).

An upgrade to a second generation fusion system (Gen2) is currently underway. LabView software was retained for data ingest and display, but improvements were made to the cloud identification, liquid distribution, and icing hazard components. More information was also included in the user display (Fig. 3). The data inputs were synchronized to run the fusion logic once per minute to avoid the noisy cloud boundaries observed in Gen1. Total integrated liquid water from the radiometer retrieval was distributed to cloud layers depending upon their depth and coldest temperature (thicker clouds have more liquid, colder clouds have less). Within each cloud layer, a fuzzy logic technique was used to distribute liquid. Four LWC profiles were calculated: uniform (as in Gen1); wedge (a linear increase with height to cloud top); reflectivity-dependent (proportional to radar reflectivity); and temperature-dependent (inversely proportional to temperature). An estimate was made, based on temperature and maximum radar reflectivity, on the composition of the cloud: either all-liquid or mixed-phase (at temperature <-20ºC, glaciated cloud was assumed). Different weights were applied to the profiles to reflect different compositions. For an all-liquid cloud, the wedge and reflectivity-weighted profiles were most heavily weighed, assuming that these tended to follow an adiabatic profile and the radar reflectivity was from cloud or perhaps drizzle-sized drops. For mixed-phase clouds, inverse temperature weighting dominated, assuming that colder temperatures implied more ice crystals and less liquid water. A third category made up of conditions that did not satisfy either our all-liquid or mixed phase criteria well, gave equal weighting to all four profiles.

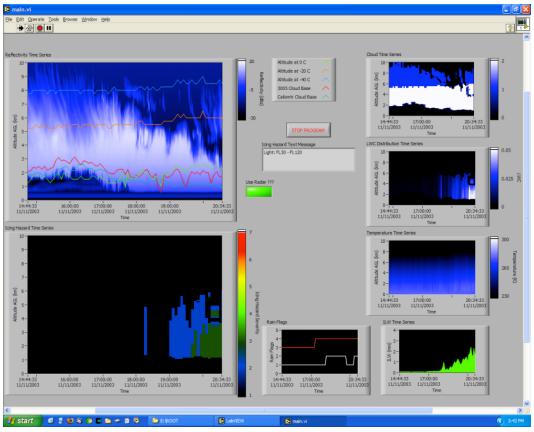

Figure 3: NIRSS Gen2 graphical interface.

4. FIELD TEST PROGRAM

The NIRSS was operated as part of the Second Alliance Icing Research Study (AIRS II), which was conducted November 2003 through February 2004. AIRS II was a collaborative scientific project involving numerous research organizations from Canada, the United States and Europe. The central research theme was aircraft icing, with operational objectives to test and evaluate remote sensing technologies, improving icing forecast technologies, further characterize the icing environment, and better characterize the aerodynamic effects of ice accretions.

Several research aircraft operated out of Ottawa, Ontario, Cleveland, Ohio, and Bangor, Maine. A large array of instrumentation, including NIRSS, was located at Mirabel Airport, Montreal, Quebec. The NASA Twin Otter Icing Research Aircraft and the National Research Council Canada (NRC) Convair 580 operated out of Ottawa during the test period. Besides other research activities, the Twin Otter and Convair performed spiral descents and missed approaches near the test site to obtain atmospheric soundings to compare to ground instrumentation (Strapp, 2003).

5. COMPARISONS WITH AIRCRAFT DATA

The output of the Gen1 software was compared to Convair and Twin Otter data. Gen2 comparisons will be done in more detail to examine the liquid water profiles and the compositions of the clouds. Case-by-case Twin Otter comparisons are discussed in detail in Reehorst et al. (2004). For both the Twin Otter and Convair datasets, comparison cases were selected based on availability of data: the aircraft was maneuvering (descending or climbing spirals) near the ground instrument site; the NASA X-band radar was operating (it was typically operated only when attended); and there was no indication of rain from the radiometer's rain sensor. Table 1 shows the seven Convair cases that satisfy these criteria and have been compared to NIRSS output.

5.1 Flight case comparisons

For each case in Table 1 a detailed comparison between Convair and Gen1 data was made. Figures 4 through 10 present the comparisons for vertical profiles of temperature, LWC, and occurrence-of-icing. Occurrence-of-icing was quantified with the aircraft Rosemount Ice Detector (RID) data and Gen1 Icing Hazard (binary yes=1 or no=0) data. Each of the three plots in the figures has aircraft data (blue lines),

Table 1. Convair cases used in the comparisons

Date (2003)	Time (UTC)	Maneuver type	Maneuver number
13 Nov	1441-1526	Descending Spiral	F03-1
13 Nov	1539-1551	Climbing Spiral	F03-2
13 Nov	1613-1626	Descending Spiral	F03-3
25 Nov	0044-0105	Descending Spiral	F07-1
25 Nov	0105-0113	Climbing Spiral	F07-2
25 Nov	0130-0140	Descending Spiral	F07-3
25 Nov	0140-0146	Climbing Spiral	F07-4

Gen1 results from remote sensor data acquired at the beginning of the aircraft maneuver (solid red line) and Gen1 results from sensor data acquired at the end of the aircraft maneuver (dashed red line).

Temperature and LWC data on these figures are self-explanatory. However, the occurrence-of-icing plots require some interpretation. The output from the aircraft RID is a voltage inversely proportional to the vibrating frequency of the ice detector's sensing element. As ice accretes on this surface, increasing its mass, the frequency decreases and the output voltage increases. When the frequency reaches a threshold value, the detector's vibrating element is heated, clearing the ice, returning the vibrating frequency and output voltage to the original values. Therefore cycling of the RID indicates icing conditions. For this effort a simple threshold of 1.7 V was used to determine occurrence-of-icing. The baseline voltage of the detector is 1.4 V, so this simple threshold is considered adequate for this initial comparison. However, this is probably too much of an oversimplification for an ideal comparison. In the future a technique needs to be developed which examines the time history of the RID signal for increasing voltage as a positive indication of icing conditions. The Gen1 results use a LWC threshold of 0.1 gm^{-3} of supercooled water to determine icing hazard.

For the first case, maneuver number F03-1, temperatures agreed reasonably well and the Gen1 results did a good job of bounding the LWC measured by the Convair (Fig. 4). The aircraft measured LWC spikes of 0.27 gm^{-3} around

1500 m, 0.1gm^{-3} at 2000 m, 0.14 gm^{-3} near 2600 m and <0.06 gm^{-3} elsewhere in the range 1400-4000 m. The remotely-measured integrated liquid water levels must have been changing through the time of the maneuver, because the distributed LWC calculated for the two times are dramatically different. Gen1 indicated no icing hazard for the beginning of the maneuver, but icing hazard existed at the end of the maneuver. The RID indicated icing conditions at the same altitudes with little or no accretion elsewhere (in regions of LWC <0.1 gm^{-3}). Compared to the aircraft data, the Gen1 algorithm underpredicted at the beginning of the maneuver by indicating no icing hazard and overpredicted the vertical extent of icing at the end of the maneuver. Judging the quality of the remote icing product based upon these results is difficult however since the conditions were obviously quickly changing during the maneuver. This issue of how to validate profiling remote sensing technologies to slower in-situ profiling techniques (aircraft and radiosondes) will need further attention in the future. It also hints at the larger issue of how best to relate measurements of a rapidly-changing environment to the user community.

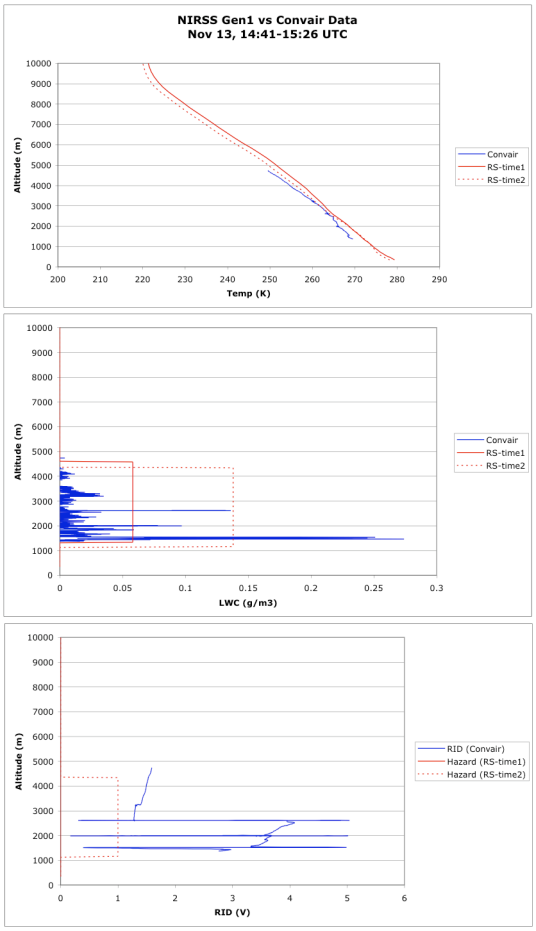

Figure 4: Comparison between Gen1 and Convair data for flight manuver F03-1.

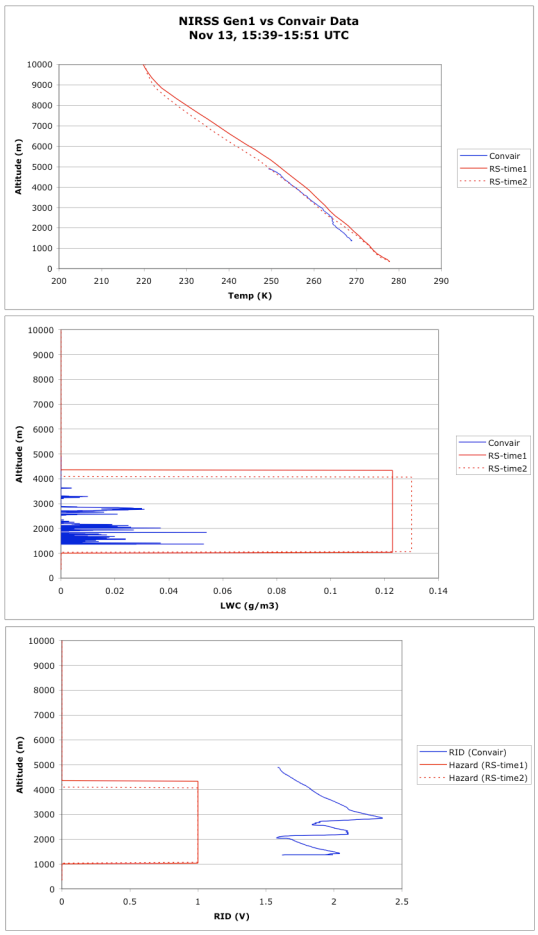

Figure 5: Comparison between Gen1 and Convair data for flight manuver F03-2.

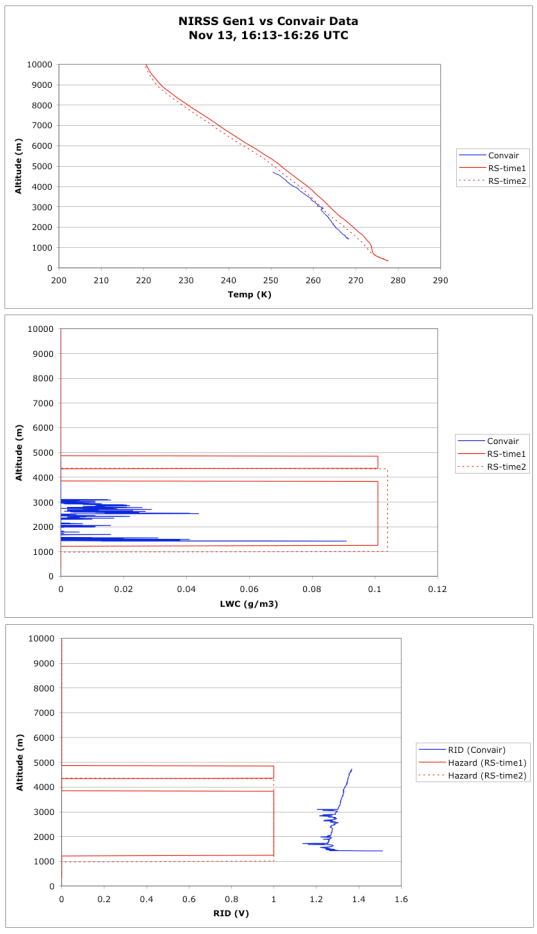

Figure 6: Comparison between Gen1 and Convair data for flight manuver F03-3.

Flight manuver F03-2 was performed 13 min after F03-1. Since it occurred a short time after the previous case, the results are quite similar. Again the temperature agreement was quite good (Fig. 5), the remote sensor-derived LWC bounded that measured by the aircraft, and remote sensor derived icing hazard bounded the area of RID activity on the aircraft. The RID seemed to be less active than the prior maneuver, and the aircraft-measured LWC was missing the higher values.

Maneuver number F03-3 came 22 min after F03-2, and again is fairly similar to the earlier cases (Fig. 6). The temperature agreement remained about the same and the aircraft recorded low LWC.

This time, the RID showed no activity except for a jump that corresponded to a LWC maximum of 0.09 gm^{-3}. The Gen1 results extended the icing hazard region up to ~5000 m, but the aircraft measured no LWC above ~3000 m. Further, based upon the RID activity, Gen1's indication of hazard for this case was probably incorrect.

Four flight maneuvers were conducted within ~1 h of one another on 25 November 2003. For these cases the remotely-measured temperature profiles were smoothed through an inversion at ~2000 m. This is a common response of profiling radiometers (Westwater, 1997). Unlike 13 November, these cases exhibited significant LWC.

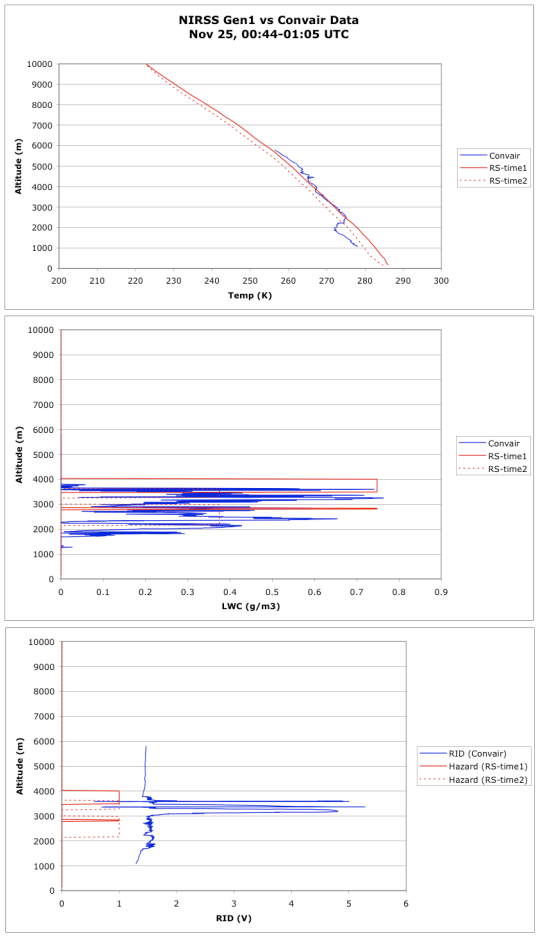

Figure 7: Comparison between Gen1 and Convair data for flight manuver F7-1.

For F07-1, Gen1 captured the upper bound of the LWC and icing hazard, but missed the lower boundary (Fig. 7). This was caused by Gen1 missing the temperature inversion, which in turn prevented it from indicating supercooled liquid at the lower altitudes. Oddly Gen1 placed a dry layer right in the region of highest RID activity. Upon closer review of the radar data, the investigators found that the reflectivity profiles changed rapidly throughout the manuver. As previously discussed, the dynamic nature of icing conditions makes the comparison of aircraft data to remotely-sensed data a challenge. It will also make appropriate reporting of conditions to flight crews a challenge. In this case, the aircraft data shows a single region of RID activity that appeared to be at the altitude where Gen1 indicated no icing hazard. However, if averaged over the whole maneuver time period, Gen1 indicate icing for the appropriate altitudes, which suggests that averaging times should be carefully considered for the system .

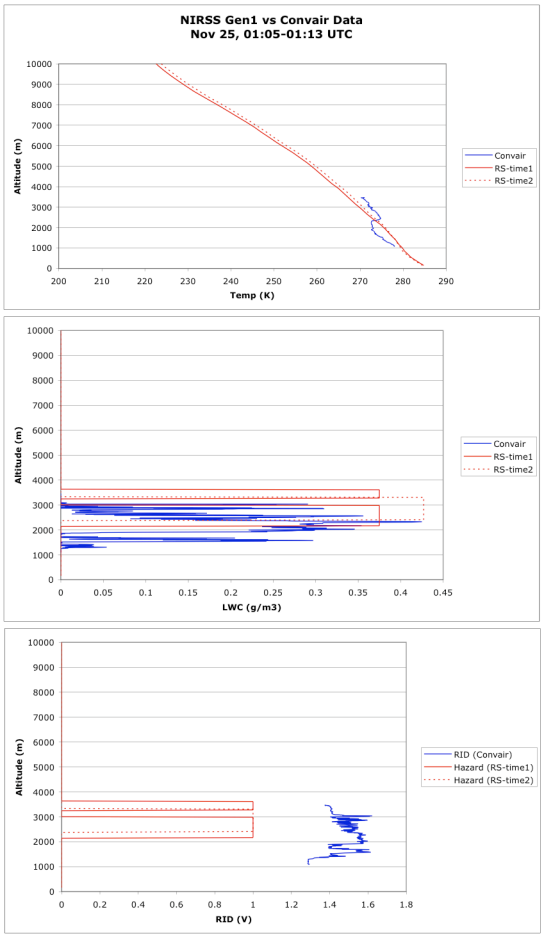

Figure 8: Comparison between Gen1 and Convair data for flight manuver F07-2.

Maneuver F07-2 took place immediately following the F07-1. Again the temperature profiles showed how Gen1 smoothed the inversion, which in turn prevented an indication of icing below ~2000 m (Fig. 8). Also, Gen1 indicated a supercooled liquid water layer above 3300 m, which the aircraft did not find. In spite of these problems, the LWC calculated by Gen1 generally agreed well with aircraft-measured values. But even with high LWC (up to >0.4 gm^{-3}), the aircraft-measured RID activity indicated little ice accumulation during this maneuver. The particular cloud conditions of this flight should be more closely examined to determine the cause of some of this seemingly contradictory behavior.

Comparisons for maneuvers F07-3 and F07-4 (Figs. 9 and 10). are similar to those for the two earlier flight maneuvers on 25 November. For F07-3, Gen1 indicated a supercooled liquid layer above ~5000 m. The aircraft did not sample this

portion of the atmosphere, so the accuracy of the Gen1 indication of this layer cannot be assessed.

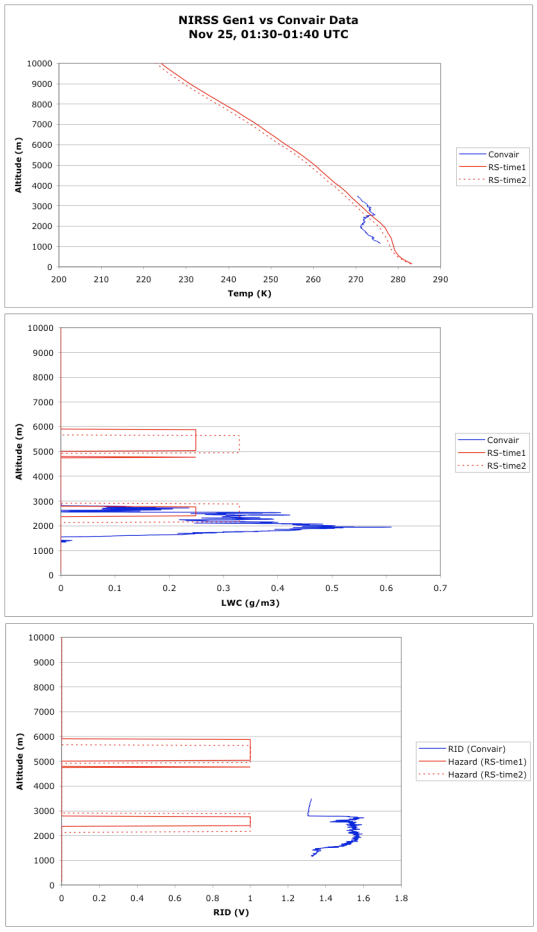

Figure 9: Comparison between Gen1 and Convair data for flight manuver F07-3.

The biggest difference between F07-3 and F07-4 is that the aircraft- measured LWC was much higher (with a peak value near 1 gm^{-3}) and there was greater RID activity, indicating significant aircraft ice accretion. This rapid change reinforces just how dynamic the icing environment can be. This cloud's LWC nearly doubled in that time and the RID indicated that the conditions changed from light icing to moderate or heavy icing. It is interesting to note that this higher LWC and icing hazard cloud has the "wedge" LWC structure of increasing LWC with altitude, while the earlier cloud had a maximum value mid-layer.

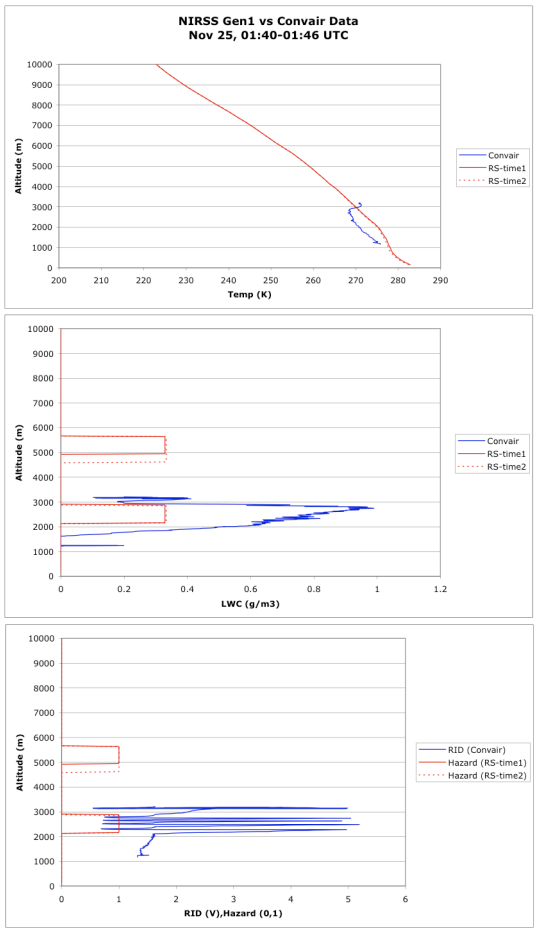

Figure 10: Comparison between Gen1 and Convair data for flight manuver F07-4.

5.2 Ensemble statistical comparisons

The remote sensor data were analyzed at both the beginning and ending time of each flight maneuver. For each comparison case, the remote sensor data and flight data were broken down into 100-m altitude increments. In each increment the data were analyzed for agreement (both positive and negative), for a remote sensor false alarm, or for a remote sensor missed detection. For all cases flight data were assumed to be truth. Due to the nature of flight data, it may be desirable in the future to refine the definition of a positive icing event.

Figure 11 shows the first set of quantitative comparisons between the NIRSS Gen1 algorithm and the corresponding Convair in-situ data. This chart shows the results when analyzed over the range of the remote sensors (100 elements, 0-10,000 m). The NIRSS Gen1 algorithm correctly categorized the icing hazard better than 80% of the time except for case F03-3. As seen earlier, the Gen1 LWC for F03-3 was only very slightly above the 0.1 g m^{-3} threshold for icing hazard, thus causing the large percentage of false alarm for this comparison case.

Figure 12 shows the results of the same comparison as Figure 11, except now using the smaller set of samples of the aircraft altitude range for the statistics. When analyzed over the smaller range of aircraft data (sample sizes listed on chart), the Gen1 algorithm accuracy drops to between 51% and 94% if case F03-3 is excluded. As discussed above, case F03-3 resulted in Gen1 LWC values just above the hazard threshold of 0.1 gm^{-3}. This comparison demonstrates how the analysis over the smaller altitude range results in more variation of the statistics and therefore points out a problem area – scaling – for a remote sensing system. However, since the sample sizes varied from case-to-case, the statistics should not be directly compared between cases. While the algorithm accuracy is much lower for this set of comparisons compared to the one using the full remote sensor altitude range, it should be noted that the missed detection values is still held relatively low. Except for case F03-1, missed detections are less than 14%. And when examined more closely, the F03-1 results are not particularly unfavorable. Between the beginning of the aircraft maneuver to the end of the maneuver, the agreement statistics changed dramatically, so it is very likely that the atmospheric conditions were rapidly changing during the time it took the aircraft to fly the maneuver. This points out an issue that will need to be addressed in the future as remote sensing data is disseminated to flight crews. Which result from a remote sensing system should be transmitted? Should the time-averaged condition be sent, or should the most severe condition be relayed? As conditions change, what should be the time period before a hazard level can be lowered? The most conservative and safest method of transmitting the most severe condition measured and then holding that value for some period of time after conditions have subsided may cause pilots to expect false alarms and learn to ignore warnings. However the other extreme of averaging conditions so that the comparison statistics are optimized will result in aircraft flying into conditions more severe than are being transmitted. More human factors work will need to be performed to answer these questions.

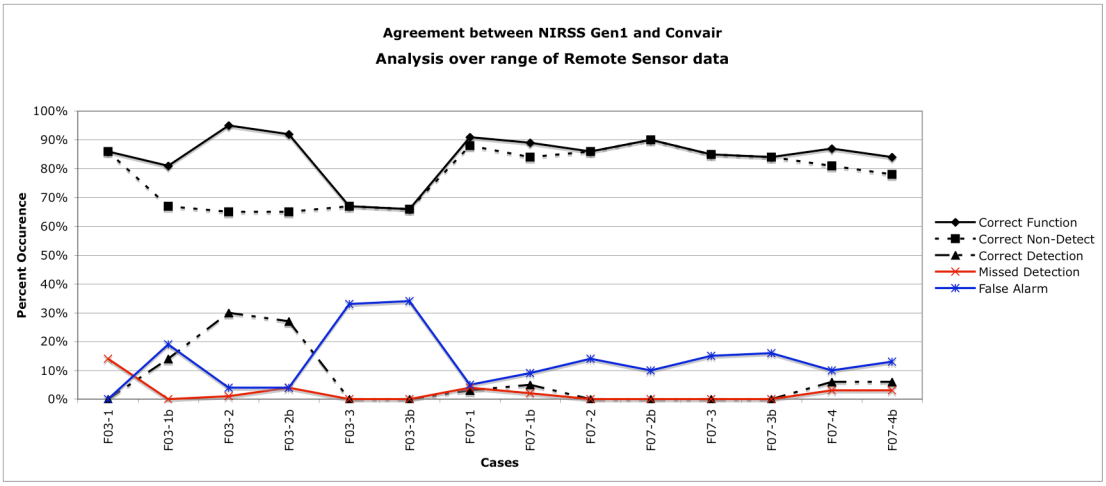

Figure 11: Quantitative agreement between Gen1 and Convair data analyzed over the altitude range of the remote sensor measurements.

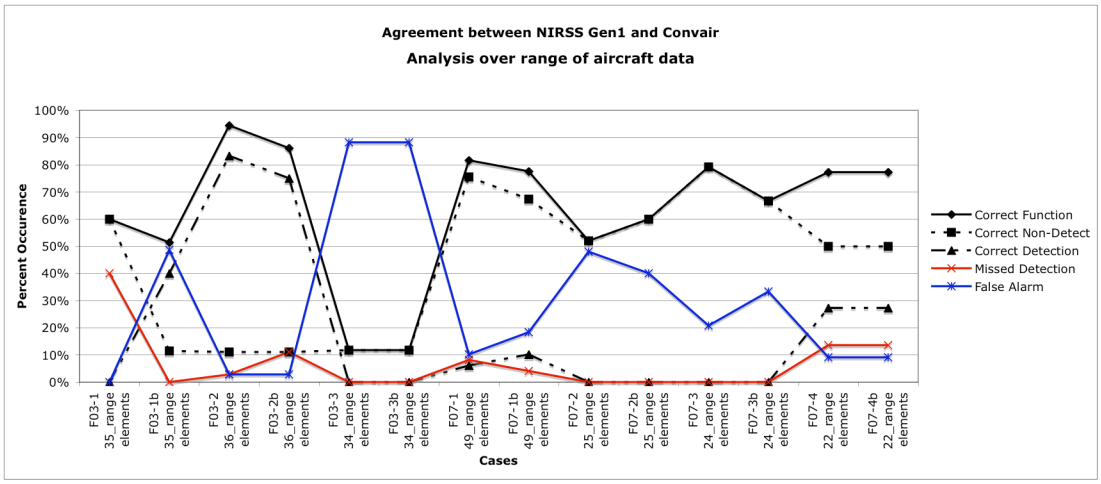

Figure 12: Quantitative agreement between Gen1 and Convair data analyzed over the altitude range of the aircraft data

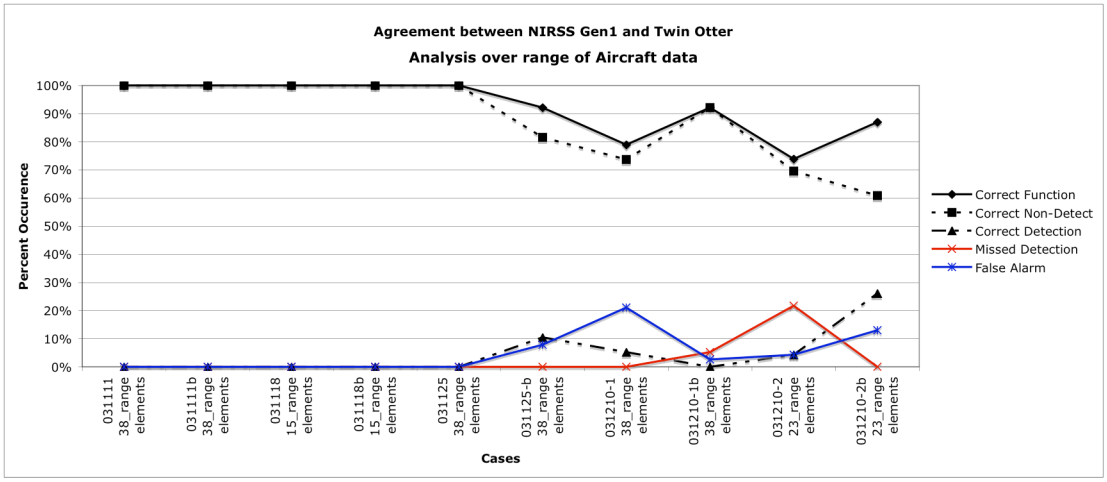

Figure 13: Quantitative agreement between Gen1 and Twin Otter data analyzed over the altitude range of the aircraft data

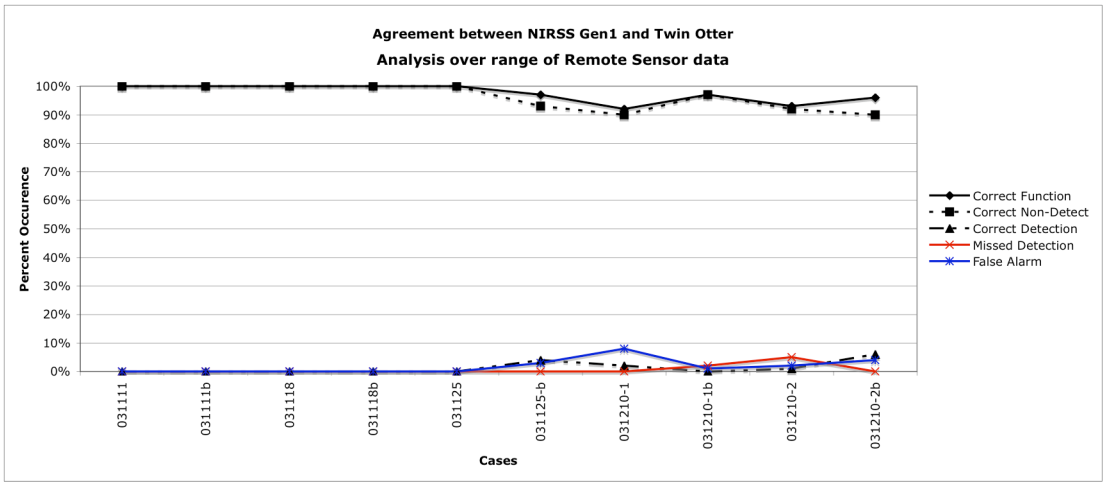

Figure 14: Quantitative agreement between Gen1 and Twin Otter data analyzed over the altitude range of the remote measurements.