9.3 CURRENT ICING POTENTIAL (CIP) SEVERITY INDEX

Tressa L. Fowler*[1], Michael Chapman[1], Barbara G. Brown[1], Jennifer Luppens Mahoney[2]

[1]National Center for Atmospheric Research, Boulder, Colorado
[2] Forecast Systems Laboratory, NOAA, Boulder, Colorado

1 INTRODUCTION

While the Current Icing Potential (CIP) has been designated as an operational product, the current operational version does not include a severity field, which is needed in order for the icing diagnoses to be comparable to the current operational advisories [i.e., the Airmens' Meteorological Advisories (AIRMETs)]. Both the icing severity and icing potential fields are needed to make flight decisions. New developments by the Federal Aviation Administration Aviation Weather Research Program's Inflight Icing Product Development Team (FAA/AWRP/IFIPDT) have led to creation of an icing severity field, and an early version of this field is considered in this paper (Bernstein et al, 2004).

Performance of the icing severity index provided by the CIP is evaluated for the winter of 2003, and compared to the performance of the AIRMETs and Significant Meteorological Advisories (SIGMETs).

Section 2 of this report describes the data used in these analyses. In section 3, the evaluation methodology is discussed. The results are presented in section 4. Finally, section 5 contains the conclusions.

2 DATA

For this study, CIP severity index values and PIREPs over the CONUS during the period January 1, 2003 through March 31, 2003 are examined. Additionally, the AIRMET and SIGMET statistics for all lead times from the same time period taken from the Real Time Verification System (RTVS; http://www-ad.fsl.noaa.gov/fvb/rtvs/index.html) are included for comparison (i.e., this paper is not intended as an evaluation of icing AIRMETs or SIGMETs).

2.1 CIP Severity Index

CIP is an icing diagnostic that combines observational data from satellite, NEXRAD radar mosaics, surface observations, a lightning detection network and recent icing PIREPs with RUC model output (McDonough and Bernstein, 1999; Benjamin et al, 1997). The current operational version of CIP produces only values representing icing potential. These values range from 0-1, with 0 indicating no icing potential and 1 indicating high icing potential, with increasing potential as the CIP values increase. The CIP icing potential field has been extensively evaluated. Recently, the CIP has been upgraded to produce an additional field, icing severity. The severity value should not be used alone, but instead in conjunction with the potential value to determine the nature of the icing conditions. The severity field is not independent of the potential field, as some of the same inputs are used to determine the value of each. However, additional information is used in determining a severity value, including PIREP severity values, while the RUC vertical velocity and super-cooled liquid water (SLW) fields are applied differently. More information on the CIP severity and its physical basis can be found in Politovich et al. (2004).

* *Corresponding author address*: Tressa L. Fowler, National Center for Atmospheric Research, P.O. Box 3000, Boulder, CO 80307 e-mail: tressa@ucar.edu

2.2 PIREPs

PIREPs, the most widely available observations of icing conditions, are used to verify all of the icing forecasts and diagnoses. During the 3 month study period, 15,254 PIREPs were received. The CIP severity values are compared to the PIREP icing intensity values, which are on a 0 to 8 scale. Table 1 shows the description of each icing intensity level, from no icing (0) to severe icing (8) along with the total number of PIREPs in each category for the 3 month study period. It is important to note that icing levels 2, 6, 7, and 8 have very low counts of PIREPs. Thus, interpretation of results for these categories must be tempered by an awareness that the sample size may be too small for the results to be meaningful.

Table 1: Icing Intensity Levels and count of PIREPs for each intensity.

Intensity of Icing	Level	Total PIREPs
No icing	0	6065
Trace	1	842
Trace to light	2	51
Light	3	4646
Light to moderate	4	1079
Moderate	5	2459
Moderate to severe	6	55
Heavy	7	2
Severe	8	55

PIREPs have known deficiencies as verification data (Kane et al., 1998). They are not systematic and they are biased in both time and space. More importantly for this analysis, the reported level of icing intensity is very subjective and may be somewhat arbitrary. Frequently, light icing is reported in the same location as moderate icing. While it might seem that the reported icing intensity would be related to the capabilities of the aircraft, previous studies have indicated that separating out larger aircraft from smaller has no effect on the level of reported icing intensity. These issues must be kept in mind when interpreting the comparisons between CIP severity values and PIREP severity values.

2.3 Research Aircraft Data

Observations of icing conditions from research aircraft are more limited both spatially and temporally than PIREPs. However, these observations are considered to be more consistent and reliable and of higher quality than the information provided by PIREPs.

For this analysis, liquid water content (LWC) observations (in grams/m^3) from the King probe on the NASA-Glenn research center's modified DeHavilland DHC-6 Twin Otter are compared to the CIP severity index (Bernstein et al., 1999; Miller et al. 1998). Eight Twin Otter flights from the winter 2003/2004 are matched to the CIP values valid within 1 hour of the flight measurements. Because each flight took multiple measurements, a total of 66 cases are available for analysis.

2.4 AIRMETs and SIGMETs

AIRMETs and SIGMETs are the operational forecasts if icing severity. AIRMETs forecast moderate or greater icing while SIGMETs forecast severe icing. These forecasts are produced by AWC forecasters every four or six hours, respectively, and may be valid until the next forecast is issued (NWS 1991). They may be amended as needed between the standard issue times. The forecasts are in a textual form that can be decoded into latitude and longitude vertices, with tops and bottoms of the icing regions defined in terms of altitude. Unfortunately, some other more descriptive elements of the AIRMETs and SIGMETs cannot be decoded and thus are not considered in verification analyses.

3 METHODS

After an initial step of mapping the CIP severity values to flight levels (i.e., every 1,000 ft) the PIREPs and research aircraft measurements are matched to the CIP severity values at the closest 12 surrounding grid points: the 4 surrounding grid points at the closest level, and the 4 grid points at the levels above and below that level. The maximum CIP severity value at any of those 12 grid points is calculated.

The research aircraft data are collected every second. For this study, data are conglomerated over an approximately 20 km flight distance (roughly 5 minutes). The location (latitude, longitude, and altitude) is averaged, and the maximum King liquid water content is kept for comparison to the CIP. When the average temperature is above freezing or the altitude changes by more than 1,000 feet over the 20 km, then the research aircraft measurements are excluded from analysis.

The pairs of PIREP intensities and corresponding CIP 12-point maximum severity values are used to compute some standard verification statistics. The CIP severity index and observations are treated here as dichotomous (i.e., Yes/No) values. AIRMETs and SIGMETs essentially are dichotomous (i.e., a location is either inside or outside the defined region). The severity field is converted to a variety of Yes/No forecasts by application of thresholds.

From a standard 2 x 2 contingency table (Table2), the verification statistics are calculated according to the formulas in Table 3. Included are the probability of detection (POD), probability of detection for non-events (PODNo) and the True Skill Statistic (TSS).

Table 2: Standard 2x2 contingency table.

	Observed	
Forecast	YES	NO
YES	YY	YN
NO	NY	NN

Table 3: Formulas for verification statistics.

Statistic	Formula
POD	YY/(YY+NY)
PODNo	NN(NN+YN)
TSS	POD + PODNo -1

In evaluating an algorithm or forecast, it is important to compare the quality of forecasts to the quality of one or more standards of reference. Thus, the quality of the CIP severity index is compared to the quality of the operational forecasts (i.e., AIRMETs and SIGMETs). However, it is important to emphasize that the index values and the AIRMETs are very different types of products, with different objectives. CIP severity index values generally are understood to be valid at a particular time. The AIRMETs and SIGMETs, on the other hand, are valid over longer periods and are designed to capture icing conditions as they move through the area over the period. Due to the differences between these products, it is difficult to clearly compare their performance. However, in order to understand the quality of the CIP severity index, it is necessary for the index values to be compared to the operational standard, especially since both types of information will be available to users. The comparisons are made in such a way as to be as fair as possible to both the operational products and CIP, while still obtaining the information needed. Nevertheless, users of these statistics should keep these assumptions in mind when evaluating the strengths and weaknesses of each type of product.

4 RESULTS

4.1 CIP Severity Statistics

Some verification results for the CIP severity index are presented in Table 2. Corresponding AIRMET statistics for the same time period are also presented for the purposes of comparison. The PODy value for moderate or greater (MOG) PIREPS for the CIP severity index with a threshold of 0.45 is 0.815. The PODn value is 0.784. The corresponding TSS value for the same set of CIP severity values is 0.599. For the same time period, the AIRMETs had a POD (MOG) of 0.685 and a PODNo of 0.554, with a TSS value of 0.239. With respect to each of these three measures, the CIP severity index outperformed the AIRMETs in correctly classifying MOG and no icing PIREPs.

The 0.45 threshold was selected because the CIP severity index was designed to produce continuous values similar to the discrete values of PIREP intensity. Thus, CIP severity values between 0.45 and 0.55 are intended to approximately correspond to a PIREP intensity of 5. However, these thresholds should not be considered absolute, as the index may not be precisely calibrated.

Table 4: CIP severity and AIRMET statistics.

	CIP severity (0.45)	AIRMETs
PODy (MOG)	0.84	0.685
PODn	0.784	0.554
TSS	0.599	0.239

For more severe icing, i.e. PIREP categories of 6, 7 or 8, comparing the CIP severity with a threshold of 0.65 to the SIGMETs makes more sense. Table 3 shows the resulting statistics. The CIP severity has a considerably higher POD (0.26 vs. 0.02) and a slightly lower PODNo (0.91 vs. 0.99). These results likely are affected by the use of PIREPs as verification data. In particular, pilots rarely fly into SIGMETs, and thus few PIREPs are available to report the conditions within a SIGMET area. This is not the case for areas with a high CIP severity, so PIREPs are more likely to be available in these areas.

Table 5: CIP severity and SIGMET statistics.

	CIP severity (0.65)	SIGMETs
PODy (SEV)	0.26	0.02
PODn	0.91	0.99

4.2 CIP Severity characteristics

Figure 1 shows box plots[1] of the maximum CIP severity value for the 12 grid points surrounding each PIREP at the nearest vertical level for each category of PIREP intensity. Ideally, the boxes would be small (i.e. show low variability) and progress along the diagonal from low to high as the PIREP intensity increases. For PIREP intensities of 0 and 1, the CIP severity would not be expected to agree well with the PIREP intensity because of the use of the maximum CIP intensity value. However, for PIREPs reporting no icing (category 0), more than half of the corresponding 12 point maximum CIP severity values were 0, and over 75% were less than 0.4. For PIREP severities of 1 through 5, the typical (median) CIP severity value is between 0.4 and 0.6. For PIREPs with an intensity of 6, the CIP severity value was typically about 0.6. The median CIP severity value increased only slightly from PIREP intensity 3 to PIREP intensity 5. The lack of a distinct increase may be due to the nature of the severity product. It's also possible that reported icing intensities of 3, 4, and 5 all represent similar icing conditions or that pockets of light and moderate icing conditions co-exist in a small area. Further, small sample sizes in categories 2, 6, 7, and 8 may lead to anomalous results.

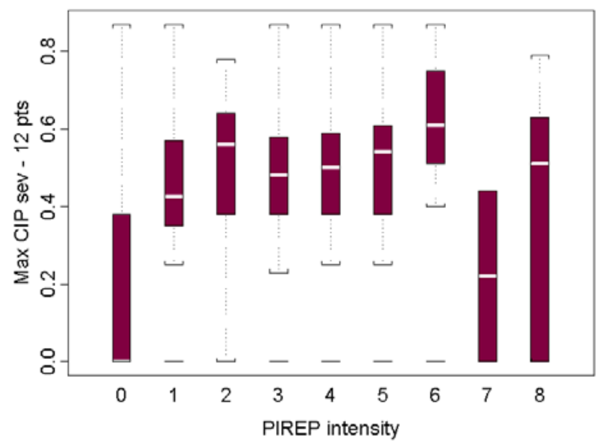

Figure 1: Boxplots showing the distribution of CIP severity values by PIREP icing intensity.

Figure 2 shows a scatter plot of the **non-zero** maximum CIP potential and severity values in the 12 grid points surrounding each PIREP. Figure 3 shows the same information in a box plot format, where the potential values have been divided into deciles (i.e. the box for 3 represents potential values greater than or equal to 0.3 and less than 0.4). The scatter plot shows the details of the edges of the data, but the center has so many points that it is difficult to decipher. The box plot better displays the bulk of the observations. Only the positive values are included in these figures because when potential is zero then severity must be zero and vice versa. The minimum severity value is always greater than 0.2 when it is positive (i.e. non-zero). When the potential is low, then the severity values are restricted to a smaller range than when the potential is large. This plot indicates that the values of icing severity are "sensible" given the values of icing potential. It is certainly possible to have a high icing potential for either low-severity icing or high severity icing. However, it is not really reasonable to expect a high icing severity value with a low potential.

[1] Box plots show the distribution of values. The line at the center of each box is the median while the top and bottom of the box represent the 75th and 25th percentiles, respectively. The whiskers extend to the maximum and minimum values that are not outliers. Any lines above or below the whiskers are outliers.

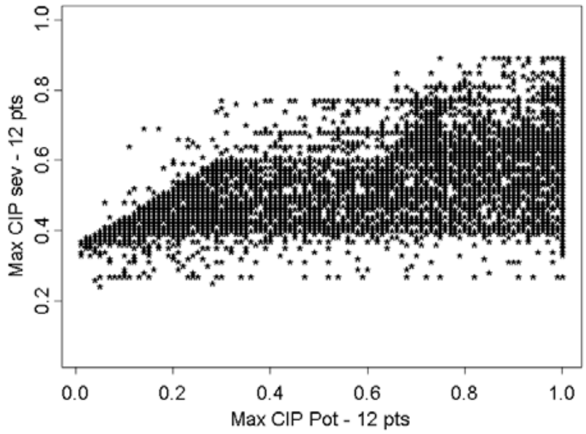

Figure 2: Scatter plot of non-zero CIP potential values versus CIP severity values.

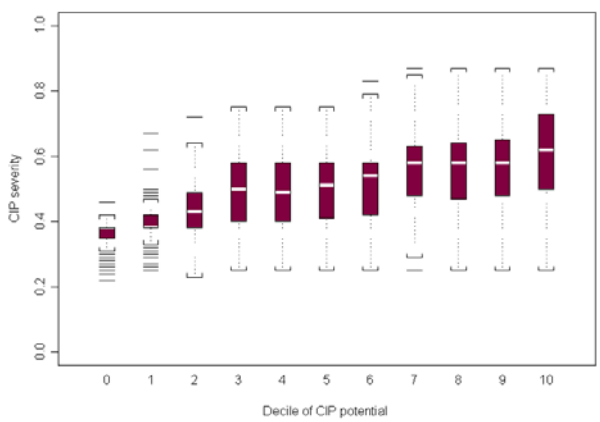

Figure 3: Box plot of non-zero CIP potential values versus CIP severity values.

Conditional histograms of the maximum CIP severity values in the 12 points surrounding a PIREP are displayed in Figure 4. The three histograms show the distribution of CIP severity values for PIREPs of no icing, PIREPs of less than moderate icing, and PIREPs of moderate or greater icing, respectively. While the first histogram clearly differs from the others, the remaining two look quite similar. If the distributions of CIP severity values do not differ for different PIREP types, then the CIP will not discriminate well between those PIREP types. Thus, Fig. 4 indicates that CIP can clearly discriminate between icing and no icing PIREPs, but it appears to have less ability to discriminate between LTM and MOG PIREPs.

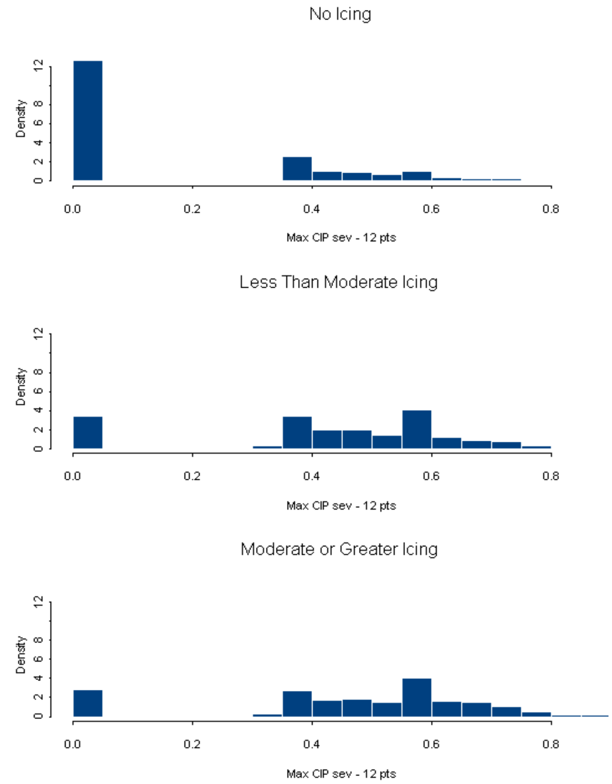

Figure 4: Conditional histograms of 12 point maximum CIP severity values for 3 types of icing PIREPs. These histograms are a probability density function in that the area of the bars (width times height) sums to 1.

The distribution of the CIP severity index looks very similar for PIREPs of MOG icing and LTM icing. However, it is possible that the apparent lack of discrimination ability is due at least in part to characteristics of the PIREPs rather than the severity index. Previous studies have shown that the icing severity reported in PIREPs is less meaningful than one would hope. For example, as shown in Figures 5 and 6, reports of trace, light, moderate, and severe conditions can be received in the same location. In Figure 5, near ORD, a light rime PIREP is shown neighboring two mixed-icing PIREPs, one moderate, the other light. Figure 6 has examples of co-located PIREPs of different severities in 4 locations, near Chicago, Kansas City, Green Bay and Columbus. The CIP severity index has a single value in these locations that may match one of the reported PIREP severities, but then cannot match the others. To determine if the CIP severity index can discriminate between different severities of icing, better observations must be obtained.

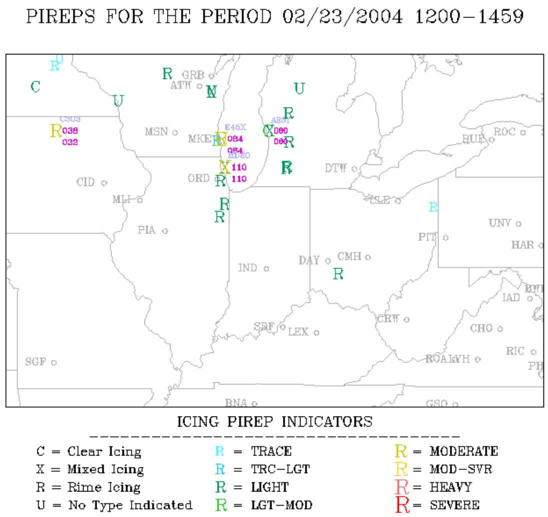

Figure 5: Map showing co-located PIREPs with different reported severities from 2/23/2004.

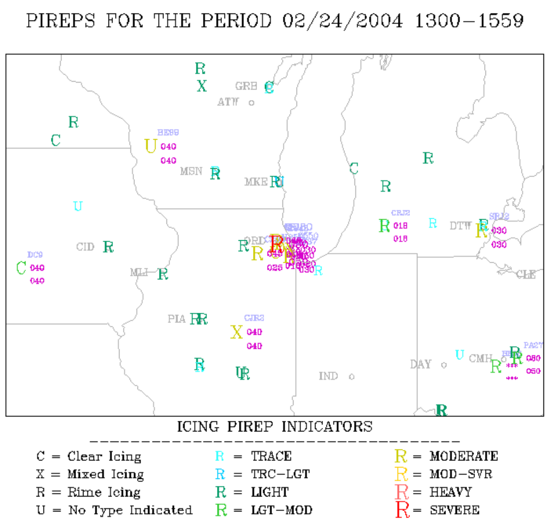

Figure 6: Map showing co-located PIREPs with different reported severities from 2/24/2004.

4.3 Research aircraft data compared to CIP

The maximum CIP severity value in the twelve grid point surrounding the 20-km average location of the NASA Glenn Twin Otter are compared to the maximum liquid water content measured by the King probe in Figure 7. Ideally this figure should show a linear relationship between CIP severity and liquid water at temperatures below freezing. The range of severity values produced by CIP is probably too narrow.

Figure 8 shows the same plot as Figure 7, but for the CIP potential rather than severity. The potential field has a wider range of values than the severity. When the research aircraft identifies areas of liquid water below freezing, the CIP severity and potential fields always have a corresponding positive value (i.e. there are no missed events). Some cases have no measured liquid water but positive CIP severity and potential values. These cases would appear to be CIP false alarms. However, here the resolution of the CIP is 1,000 feet in the vertical and 20 km in the horizontal while the resolution of the aircraft measurements is on the order of feet. Thus, if the aircraft flies just above or outside an icing area, the King probe will measure no liquid water but the CIP will indicate that icing conditions exist. Though both the King probe and the CIP are correct, the measurements are not directly comparable because of the differing scales.

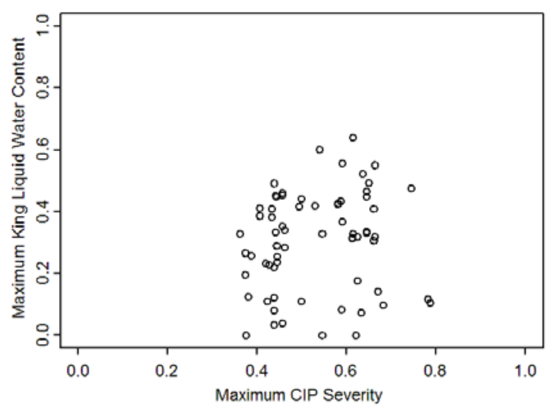

Figure 7: CIP severity versus maximum liquid water content measured by the King probe on the NASA Glenn Twin Otter for below freezing cases.

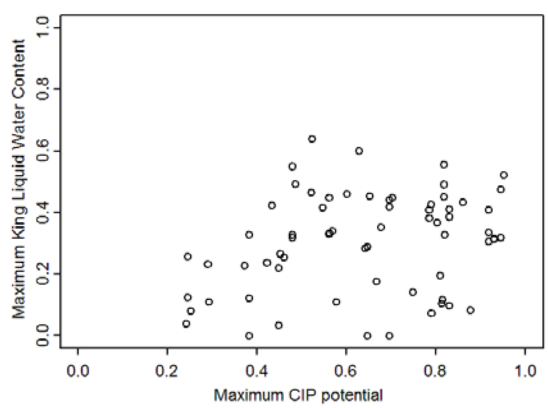

Figure 8: CIP potential versus maximum liquid water content measured by the King probe on the NASA Glenn Twin Otter for below freezing cases.

5 CONCLUSIONS AND FUTURE WORK

The CIP severity shows some ability to discriminate between no icing and MOG icing and between no icing and severe icing. However, it appears less able to discriminate between LTM and MOG icing. AIRMETs do not attempt to discriminate between LTM and MOG icing. Instead, they are issued only for MOG icing. All other areas are assumed to be either LTM or no icing. By converting the CIP severity index into categories and calculating standard verification statistics, it is clear that the CIP severity performs at least as well as the current operational icing severity indicator, the AIRMETs. Additionally, the CIP severity index has at least as much skill as the SIGMETs for detecting the more severe PIREPs (6 and above).

The relationship between values of CIP severity and CIP icing potential is sensible. In particular, the range of severity values increases with increases in the CIP icing potential values. Low severity CIP values occur for all values of CIP icing potential, while high severity values are restricted to intervals of higher icing potential.

While the CIP severity index does not seem to discriminate well between PIREPs of different severities, this may be due to the nature of PIREPs rather than a failing of the index.

When the research aircraft identifies areas of liquid water below freezing, the CIP severity and potential fields always have a corresponding positive value. In some cases, they also have a positive value when the King probe finds no water, but this is more likely an issue of scale than a true false alarm. The CIP potential field appears to correlate better to LWC measurements than the CIP severity field. Perhaps information from the potential field could be used to improve future versions of the CIP severity index.

ACKNOWLEDGEMENTS

NCAR is sponsored by the National Science Foundation. This research is in response to requirements and funding by the Federal Aviation Administration (FAA). The views expressed are those of the authors and do not necessarily represent the official policy of the FAA.

REFERENCES

Benjamin, 1997: Improvements in aviation forecasts from the 40-km RUC. Preprints: 7th Conference on Aviation, Range and Aerospace Meteorology, 2-7 February, Phoenix, p. 411-416.

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 *J. Appl. Meteor.*

Bernstein, B.C., T.P. Ratvasky, D.R. Miller and F. Mcdonough, 1999: Freezing rain as an in-flight icing hazard. Preprints, 8th Conference on Aviation, Range and Aerospace Meteorology, Dallas TX, 10-15 January. Amer. Meteor. Soc., Boston, 38-42. **Also NASA/TM-2000-210058.

Brown, B.G. and A.H. Murphy, 1996: Verification of aircraft icing forecasts: The use of standard measures and meteorological covariates. Preprints, 13th Conf. on Proability and Statistics in the Atmospheric Sciences, San Francisco, 21-23 February. Amer. Meteor. Soc., Boston, 251-252.

Kane, T. L., B.G. Brown, and R. Bruintjes, 1998: Characteristics of pilots reports of icing, Preprints: 14th Conference on Probability and Statistics, 11-16 January, Phoenix, p. 90-95.

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 Conference on Aviation, Range and Aerospace Meteorology, 10-15 January, Dallas, p. 467-471.

Miller, D., T. Ratvasky, B. Bernstein, F. McDonough, and J.W. Strapp, 1998: NASA/FAA/NCAR Supercooled Large Droplet Icing Flight Research: Summary of Winter 96-97 Flight Operations. 36th Aerospace Science Meeting and Exhibit, Reno, NV, 12-15 January. American Institute of Aeronautics and Astronautics, Washington, D.C.

NWS, 1991: National Weather Service Operations Manual, D-22. National Weather Service. (Available at Website http://www.nws.noaa.gov).

Politovich, M.K., B.C. Bernstein and F. McDonough, 2004: CIP Severity Algorithm Scientific and Technical Description. Report to Aviation Weather Technology Transfer Technical Review Board.

Politovich, M.K., B.C. Bernstein and F. McDonough, 2002: Issues in forecasting icing severity. Preprints, 10th Conference on Aviation, Range and Aerospace Meteorology, Portland OR, 13-17 May, Amer. Meteor. Soc., Boston, 85-88.