NOAA Technical Memorandum OAR GSD-52

https://doi.org/10.7289/V5/TM-OAR-GSD-52

Assessment of the Current Icing Product (CIP) and Forecast Icing Product (FIP) Version 1.1

August 2014

Brian J. Etherton
Matthew S. Wandishin
Joan E. Hart
Geary J. Layne
Michael H. Leon
Melissa A. Petty

Earth System Research Laboratory
Global System Division
Boulder, Colorado
August 2014

NATIONAL OCEANIC AND ATMOSPHERIC ADMINISTRATION / Office of Oceanic and Atmospheric Research

https://doi.org/10.7289/V5/TM-OAR-GSD-52

NOAA Technical Memorandum OAR GSD-52

Assessment of the Current Icing Product (CIP) and Forecast Icing Product (FIP) Version 1.1

Brian J. Etherton[1]
Matthew S. Wandishin[2]
Joan E. Hart[2]
Geary J. Layne[2]
Michael H. Leon[3]
Melissa A. Petty[3]

[1] National Oceanic and Atmospheric Administration, Earth System Research Laboratory, Global Systems Division (NOAA/ESRL/GSD)

[2] Cooperative Institute for Research in Environmental Sciences (CIRES) and NOAA/ESRL/GSD

[3] Cooperative Institute for Research in the Atmosphere (CIRA) and NOAA/ESRL/GSD

UNITED STATES DEPARTMENT OF COMMERCE

Wilbur Ross
Secretary

NATIONAL OCEANIC AND ATMOSPHERIC ADMINISTRATION

RDML Tim Gallaudet, Ph.D., USN Ret.
Under Secretary for Oceans
And Atmosphere/NOAA Administrator

Office of Oceanic and Atmospheric Research

Craig N. McLean
Assistant Administrator

TABLE OF CONTENTS

TABLE OF TABLES

TABLE OF FIGURES

1 INTRODUCTION

This assessment compares the High-Resolution Current Icing Product (CIP) and Forecast Icing Product (FIP) algorithms version 1.1 developed by the National Center for Atmospheric Research. These 1.1 products are minor upgrades to the current CIP/FIP version 1.0, and will be combined with 1.2 upgrades released in fall 2014, for deployment to replace the current operational version. There are only two differences in 1.1 versus 1.0, and both are in the CIP: a correction to an aliasing with surface data and the use of METAR reports of cloud data.

The assessment incorporates output from the operational CIP/FIP 1.0 algorithms, the experimental CIP/FIP 1.1 provided by NCAR, as well as observations (METARs and PIREPs), to establish a performance baseline. The assessment has three main areas of investigation:

1. Characteristics of the product fields
2. Overall performance and meteorological accuracy of the CIP/FIP 1.1 as compared to the CIP/FIP 1.0
3. A geographical comparison of the Hires CIP/FIP 1.1 and Hires CIP/FIP 1.0 fields.

2 DATA

The time period for this study is approximately one month, January 6th through February 11th 2014 (with a brief outage in that date range, the result being 30 days of data). CIP/FIP 1.1 data were provided by NCAR for this assessment, whereas CIP/FIP 1.0 was ingested via a non-operational feed from the Aviation Weather Center (AWC).

2.1 ANALYSES/FORECASTS

The output from the grid-based CIP/FIP algorithms includes: calibrated icing probability, icing severity, and potential for SLD (including freezing drizzle and freezing rain). The methodology used for producing CIP can be found in Bernstein et al., 2005. References for FIP methodologies can be found in McDonough et al. (2003), Brown and Bernstein (2006), and Wolff et al. (2008). The spatial and temporal attributes of the CIP/FIP are outlined below (Table 2.1) and apply to both versions 1.0 and 1.1.

Table 2.1: Attributes of the CIP/FIP.

Issues	Every hour
Leads	CIP: 0 FIP: 1-3, 6, 9, 12
Horizontal Resolution	13km
Altitudes	500–30,000 ft; 500 ft increments

2.2 OBSERVATIONS

2.2.1 Voice Pilot Reports (PIREPs)

PIREPs are reported irregularly at the pilot's discretion and include a subjective assessment of many meteorological variables including the existence/absence of icing and a subjective measure of the icing intensity. Included in the icing reports are the location, altitude or range of altitudes, type of aircraft, air temperature, intensity, and type of icing (NWS 2007). The full range of intensity values are listed below. The 'clear' type is used to indicate the possibility of SLD.

Icing intensity

1. Trace: Ice becomes perceptible. The rate of accumulation is slightly greater than sublimation. Deicing/anti-icing equipment is not utilized unless encountered for an extended period of time (over one hour).
2. Light: The rate of accumulation may create a problem if flight is prolonged in this environment (over one hour). Occasional use of deicing/anti-icing equipment removes/prevents accumulation. It does not present a problem if deicing/anti-icing is used.
3. Moderate: The rate of accumulation is such that even short encounters become potentially hazardous, and use of deicing/anti-icing equipment or diversion is necessary.
4. Severe/Heavy: The rate of accumulation is such that deicing/anti-icing equipment fails to reduce or control the hazard. Immediate diversion is necessary.

Icing types

1. Rime: Rough, milky, opaque ice formed by the instantaneous freezing of small super-cooled water droplets.
2. Clear: A glossy, clear or translucent ice formed by the relatively slow freezing of large super-cooled water droplets.
3. Mixed: This is a combination of rime and clear.

2.2.2 METAR Observations

Routine surface report (METAR) data are used to provide observations of icing conditions at the surface and to infer SLD events between the surface and the cloud ceiling. For instance, when freezing rain or freezing drizzle is recorded in the METAR, an SLD event is then inferred to exist between the surface and the cloud base (lowest cloud layer of at least "broken" coverage) (Madine 2008). This information is used to assess the quality of the CIP/FIP SLD parameter.

2.3 Stratifications

Performance results are stratified spatially, temporally, and according to certain icing intensity thresholds, specified below.

Altitude Bins

Results are aggregated into the altitude ranges shown in Table 2.2.

Table 2.2: Altitude bin stratifications.

Stratification	Altitudes
Low	500 – 10,000 ft
Middle	10,500 – 20,000 ft
High	20,500 – 30,000 ft

ICING PROBABILITY STRATIFICATIONS

Consistent with information provided by the Aviation Digital Data Service (ADDS), CIP and FIP icing severity are masked using three probability thresholds: > 5%, ≥ 25%, and ≥50%.

TEMPORAL STRATIFICATION

Forecast performance is stratified by forecast issue and lead times.

INTENSITY STRATIFICATION

The majority of the focus of the evaluation of icing intensity is on the Moderate-or-Greater level, but all CIP/FIP categories are considered. Table 2.3 shows how severity categories are related between PIREPS and the CIP/FIP products, as they are different between the two datasets.

Table 2.3: Mappings of icing severity categories.

(ADDS) PIREP	CIP/FIP category
Neg Neg-Clr	None
Trace	Trace
Trace-Light Light	Light
Light-Mod Mod	Moderate
Mod-Severe Heavy Severe	Heavy

SLD STRATIFICATIONS

Consistent with information provided by the Aviation Digital Data Service (ADDS), values of SLD potential are masked using three thresholds: <0.0 (unknown), between 0.0 and 0.05 (no SLD), and ≥0.05 (SLD present).

3 METHODS

The mechanics of the assessment include: 1) evaluating distributions of field values in CIP and FIP, 2) a neighborhood-based approach for verifying CIP/FIP severity and SLD using PIREP observations, and 3) a verification of CIP/FIP SLD using METAR observations, and 4) evaluating spatial distributions of CIP/FIP 1.0 and 1.1. The techniques described below are applied to both 1.0 and 1.1 versions of CIP/FIP.

3.1 FORECAST AND OBSERVATION PAIRING TECHNIQUES

To enable forecast comparisons and evaluation of quality, forecasts and observations are matched spatially and temporally using the mechanics described in the following subsections. In all techniques, the CIP/FIP severity field is masked using probability values of 0.05, 0.25, and 0.50.

3.1.1 CIP/FIP SEVERITY TO PIREPS

It is known that PIREPs have location error. Pearson and Sharman (2013) reported that the median horizontal error of turbulence PIREPs was 35km, and the interquartile range of the vertical error is -176 to 79ft or -25 to 28ft depending on the airline. To account for location uncertainty of PIREPs, a neighborhood approach using forecast grid points around the location of the PIREP will be used to match the forecast to the PIREP.

The neighborhood is defined by a 22-point horizontal neighborhood, essentially a circle of radius approximately 35km centered at the grid point closest to the PIREP location (figure 3.1), which is included at each flight level. The severity value within the neighborhood that best matches the PIREP intensity is taken as the associated forecast value.

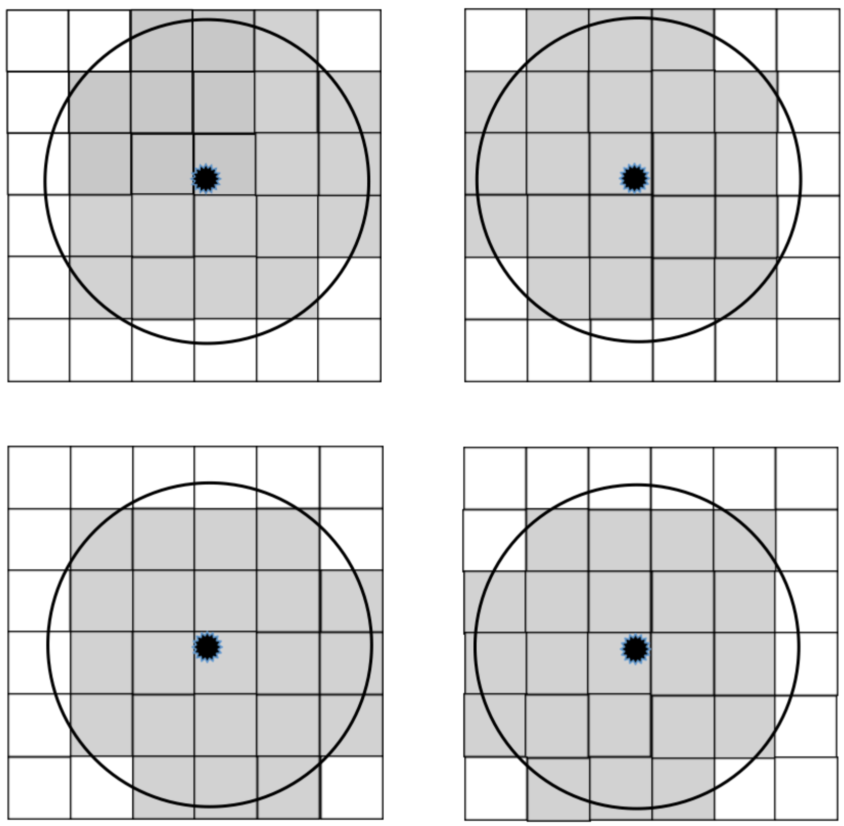

Figure 3.1: A schematic of the four possible 22-point neighborhoods surrounding a PIREP.

The selection of neighborhood size is a balance between accounting for PIREP location uncertainty and weakly representing the resolution of information for a product. The choice of 35km is based on the median error, and as such it is expected that only half of the PIREPs will actually be in their respective neighborhood.

The grid points chosen for the neighborhood are determined by the quadrant of the grid box in which the PIREP lies. On the boundaries of the CIP/FIP grid, the subset of points available in the neighborhood is used for a best match. Grid points located below the model surface elevation are also excluded. In the 'best-match' approach, if there is not a perfect match (a CIP/FIP intensity directly matching the PIREP intensity) the closest match is determined by first searching all higher intensities for the closest higher, then searching all lower intensities.

The vertical extent of the neighborhood for a report of icing at a single altitude is defined by including the grid points within 35km of the PIREP at the CIP/FIP level closest to the PIREP flight level, along with the level above and below (+/- 500 feet), resulting in three vertical levels around the PIREP. Though this is greater vertical extent than the value from Pearson and Sharman (2013), that work investigated turbulence PIREPS, not icing. Given that icing is more likely on ascent/descent than at cruise altitude, we expect greater vertical uncertainty in icing PIREPs than in turbulence PIREPs, hence, the broader range. The left-most image of figure 3.2 shows this neighborhood.

For PIREPs in which a top and base of icing is reported, the neighborhood is defined by including the grid points within 35km of the PIREP from the CIP/FIP level above the level closest to the PIREP's reported 'icing top', to the level below the level closest to the PIREP's reported 'icing base'. For PIREPs that report 'no' icing over a layer where the top level is 'unlimited' (also known as 'clear and above'), the neighborhood extends from one level above the icing base to the top level of CIP/FIP. In this case, rather than the best-match approach, a correct CIP/FIP forecast is one in which there is a column of no-icing forecasts that extends through the depth of the neighborhood column. The center- and right-most images of figure 3.2 show the two neighborhoods used when an icing PIREP reports a top and bottom of the layer of icing (or of no icing).

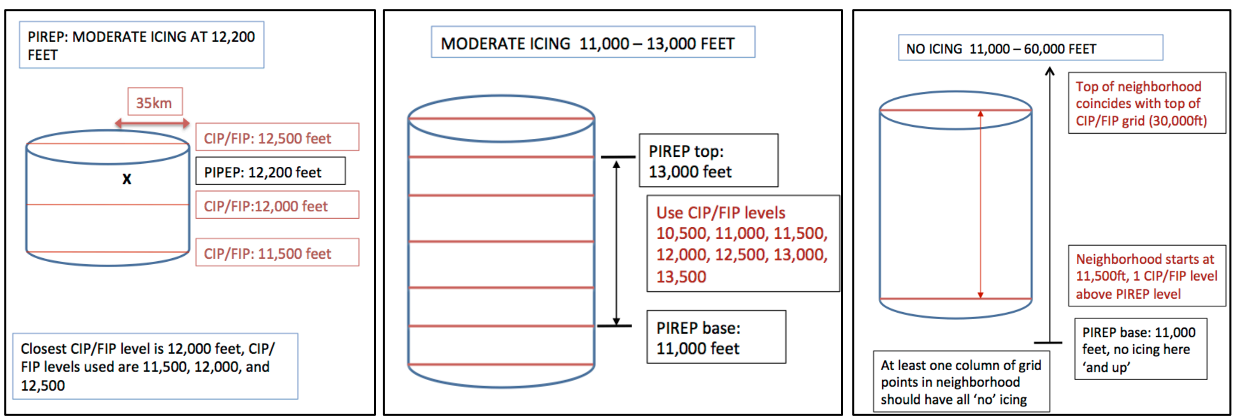

Figure 3.2: A schematic of three possible neighborhoods surrounding a PIREP. On the left is a neighborhood surrounding a report for a single vertical level. The other two images are for a report that has a top and base for the icing conditions (center) or no icing (right) at and above a given altitude.

For temporal matching, all PIREPs within a time window of [-30, 30) minutes around the forecast valid time are used to verify FIP. Because PIREPs prior to the analysis time are incorporated in CIP, only a time window of [0, 30) minutes around the analysis time are used to verify CIP.

3.1.2 CIP/FIP SLD

3.1.2.1 SLD PIREP Based

A PIREP is considered an observation of SLD conditions if it indicates severe icing intensity and clear icing type. PIREPs of all other types and intensities are considered negative reports of SLD, with one exception: PIREPs include a 30 character "Weather String" that provides information on the weather observed. In this string, it is possible to have reports of Freezing Rain or Freezing Drizzle. If such a report exists within the Weather String, that PIREP is considered to be a positive report of SLD, regardless of what icing type and severity was reported.

PIREPs are matched to the CIP/FIP forecast grid using the same spatial and temporal neighborhoods as for severity. The SLD forecast field is categorized into 'yes', 'no', and 'unknown' using the thresholds indicated in the Stratifications section.

3.1.2.2 METAR Based

METARs are included as an observation set for verification of SLD. Positive SLD observations are established using reports of freezing rain (FZRA) or freezing drizzle (FZDZ) that also reported a cloud layer of at least "broken". The ceiling value from the METAR is used to estimate the depth of the observed SLD layer, with the ceiling value being the top of the SLD layer and the ground being the bottom. A METAR is considered a report of 'no' SLD if it reports clear skies or snowfall.

For METARs that indicate SLD, SLD is assumed present from the ground to cloud base. For METARs that indicate no-SLD, the observation is assumed valid from the ground to either cloud base (if the METAR indicates snow) or to 30,000 feet (if the METAR indicates clear skies).

The grid box that contains the METAR location, from the lowest vertical level up to the chosen top level, is compared to the METAR report. For METARs that indicate SLD, at least one of the vertical levels in the column of forecast grid boxes above the METAR site is expected to contain SLD. For the METARs that indicate no SLD, it is expected that all grid boxes above the site, up to the chosen top, will not contain SLD.

3.2 Evaluations

Terminology and score definitions are first provided for reference in the subsequent sections:

MOG:	Moderate-or-Greater Icing
POD:	(Probability of Detection) Proportion of all observed events that are correctly forecast to occur, in this case, of detecting icing at a specific threshold
POFD:	(Probability of False Detection) Proportion of all observed non-events that are mistakenly forecast to be events, in this case, detecting icing less than the specified threshold
PSS:	Peirce Skill Score, aka True Skill Score (POD – POFD); Forecast skill relative to an unbiased random forecast; Provides a measure of the product's ability to separate 'yes' events from the 'no' events

4 Results

4.1 Distributions

The makeup of the CIP/FIP fields is first evaluated using value-based distributions, which are generated for each field: bins for CIP/FIP severity are generated per severity category, and the probability and SLD values are binned from 0 to 1.0 using a bin size of 0.01. Note that for all fields, only values greater than zero are shown. Given the relatively large number of zero values in the SLD and probability fields, their exclusion from the plots makes the plots easier to interpret.

Comparing the distributions of icing severity from CIP and FIP, 1.1 and 1.0 (figures 4.1 and 4.2), the two products are similar. There is approximately 10% more severity field being produced in the CIP 1.1 than CIP 1.0. The increase is not evenly distributed— there is a greater ratio of trace and light icing amounts than severe icing. There is more severe icing in CIP 1.0 than in CIP 1.1. For FIP, the amount (number of grid boxes) of icing of trace or greater is the same, but there is a slight shift towards higher severity values. The differences between the distributions of icing severity is more pronounced when the probability of icing is greater than 50% (figure 4.2) than when the probability of icing is 5% or greater (figure 4.1). Given that there were no explicit changes to the FIP algorithm, it is possible these differences could be due to round-off error between the two systems processing the data.

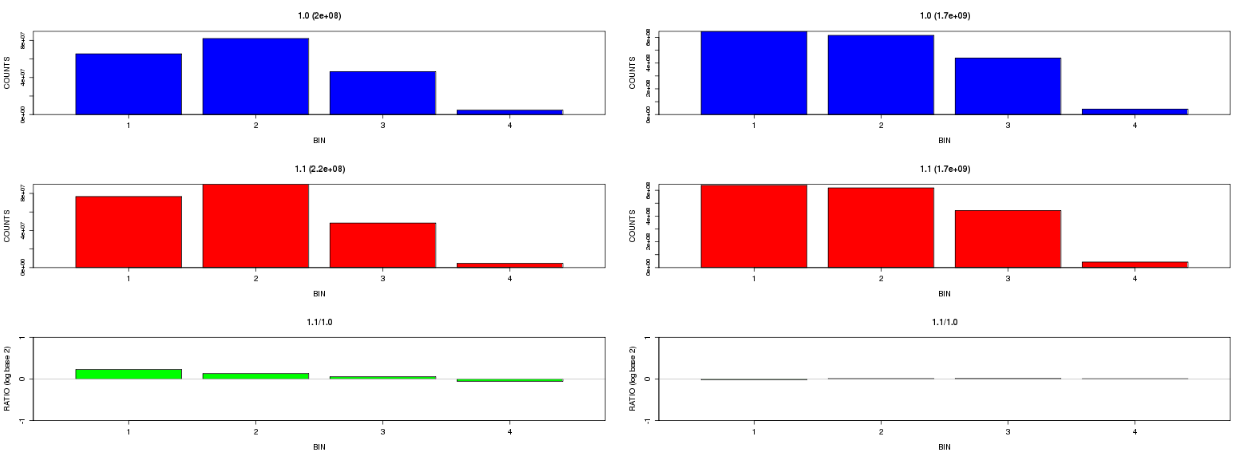

Figure 4.1: The number of gridboxes when the value of severity was trace (1), light (2), moderate (3), or severe (4) from the two products, when the probability of icing was greater than 5%. CIP is on the left, FIP is on the right. CIP and FIP 1.0 are in blue, CIP and FIP 1.1 are in red, and the ratio of the two products (log base 2) is in green.

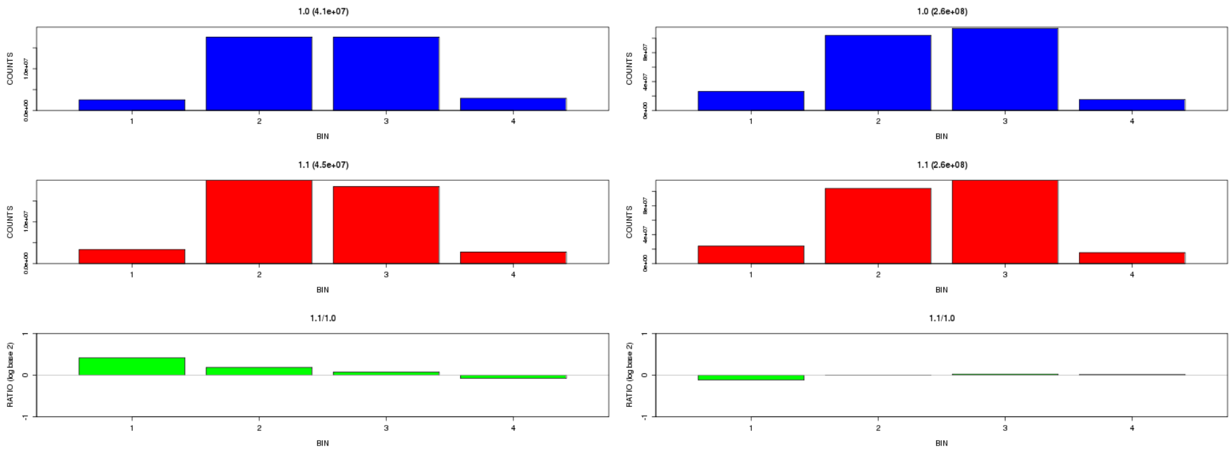

Figure 4.2: The number of gridboxes when the value of severity was trace (1), light (2), moderate (3), or severe (4) from the two products, when the probability of icing was greater than 50%. CIP is on the left, FIP is on the right. CIP and FIP 1.0 are in blue, CIP and FIP 1.1 are in red, and the ratio of the two products (log base 2) is in green.

When comparing the distributions of the probability of icing (figure 4.3), and of SLD (figure 4.4), FIP 1.1 and FIP 1.0 are identical. The CIP are different, with CIP 1.1 having a greater amount of probability of icing and SLD as compared to CIP 1.0. These differences are evenly distributed through the range of values. On average, there is about 10% more probability of icing in CIP 1.1 than in CIP 1.0 (figure 4.3). For SLD (figure 4.4), there is roughly double the amount in CIP 1.1 than CIP 1.0. The ratio of the distributions of SLD between CIP 1.1 and CIP 1.0 is the largest of any field, which is not surprising, as the SLD field is a subset of the total icing and the smallest of all fields investigated. It is noted that the SLD distributions have spikes at 1 in the CIP and 0.67 in the FIP. These results are similar to those noted in past assessments, both the large number of values at certain thresholds, and the differences between products produced at NCAR and AWC, which is the case with 1.1 and 1.0 in the present study.

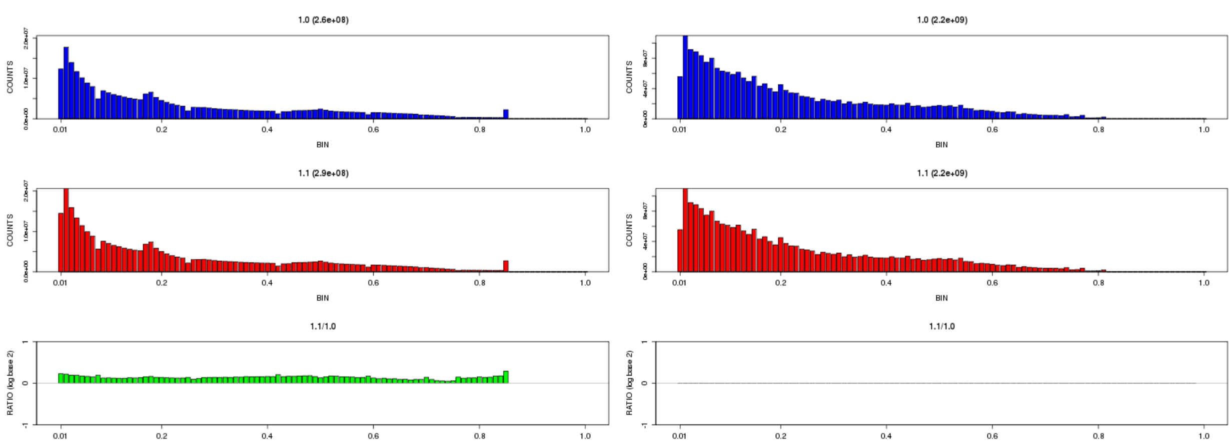

Figure 4.3: The distribution of values of probability of icing, from 0.01 to 1. CIP is on the left, FIP is on the right. CIP and FIP 1.0 are in blue, CIP and FIP 1.1 are in red, and the ratio of the two products (log base 2) is in green.

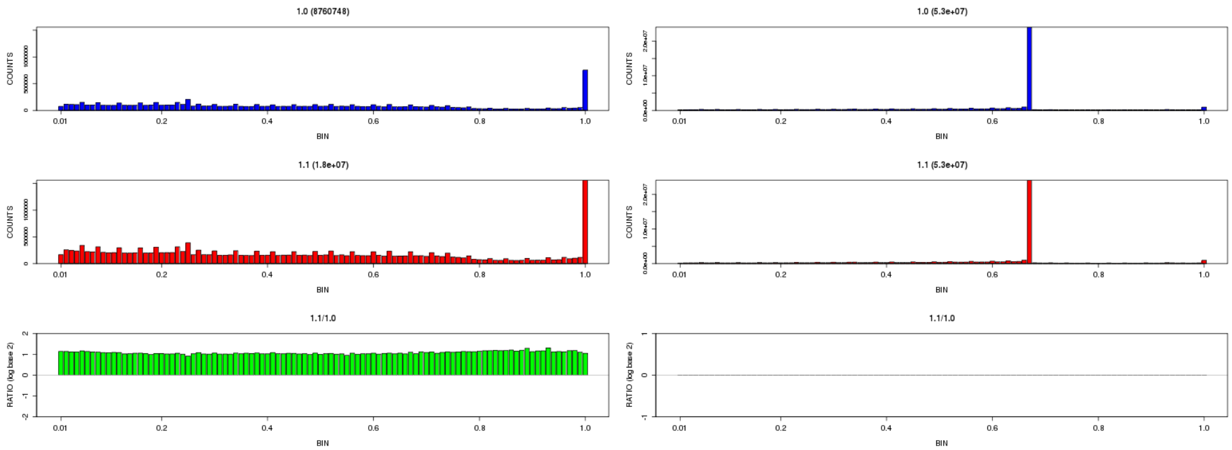

Figure 4.4: The distribution of values of SLD, from 0.01 to 1. CIP is on the left, FIP is on the right. CIP and FIP 1.0 are in blue, CIP and FIP 1.1 are in red, and the ratio of the two products (log base 2) is in green.

4.2 SKILL – SEVERITY

Due to the non-systematic nature of PIREPs, the "yes" observations and "no" observations must be treated separately (Carriere et al. 1997). As a result, it becomes inappropriate to compute several common statistics that would otherwise be computed and analyzed (e.g. Critical Success Index, Bias, and False Alarm Ratio). The rationale for this is well documented by Brown and Young (2000) and Carriere et al. (1997).

The association of the CIP/FIP product to PIREPs as described in section 3.1.1 yields the following contingency table:

Hit:	forecast = yes; obs = yes
False alarm:	forecast = yes; obs = no
Miss:	forecast = no; obs = yes
Correct no:	forecast = no; obs = no

'Yes' signifies that the forecast or observation equals or exceeds a given threshold, and 'no' signifies that the forecast or observed value is less than the threshold. POD, POFD, and PSS are computed from the contingency table.

The analyses and forecasts of MOG icing conditions, as shown in figure 4.5, show that the POD and POFD of CIP 1.1 are better than those of CIP 1.0, though the differences are likely not statistically significant. The differences in POD and POFD of FIP 1.1 and 1.0 are negligible, with slight improvement in FIP 1.1.

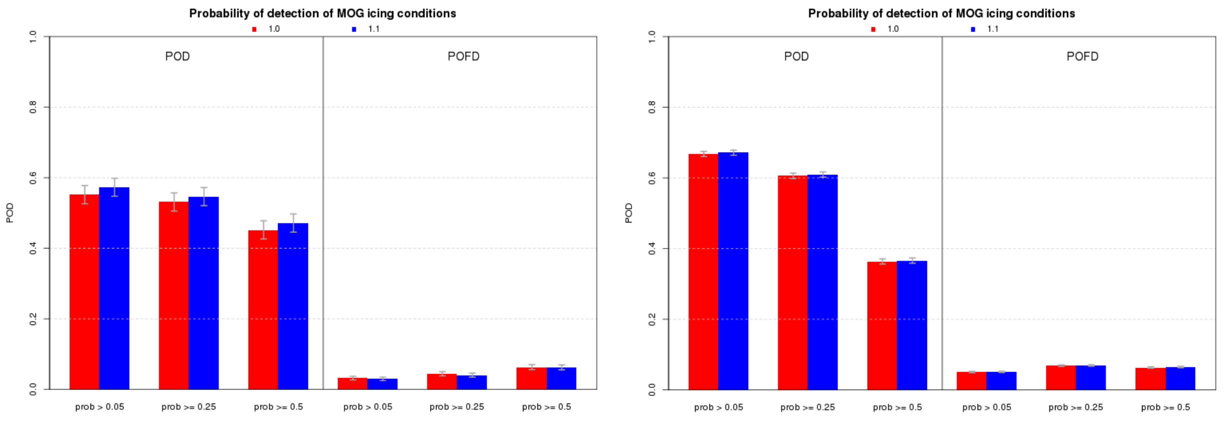

Figure 4.5: The probability of detection (POD) and the probability of false detection (POFD) of moderate or greater icing from CIP 1.1 and 1.0 (left panel) and FIP 1.1 and 1.0 (right panel). CIP and FIP 1.1 are in blue, CIP and FIP 1.0 are in red.

When comparing the analyses and forecasts of any icing conditions (figure 4.6), there is no difference in the FIP 1.1 and FIP 1.0, as expected. CIP 1.1 performs slightly better than CIP 1.0, but performs worse than the one-hour forecast from FIP. The characteristic of higher skill for the one-hour forecast has been noted in prior assessments.

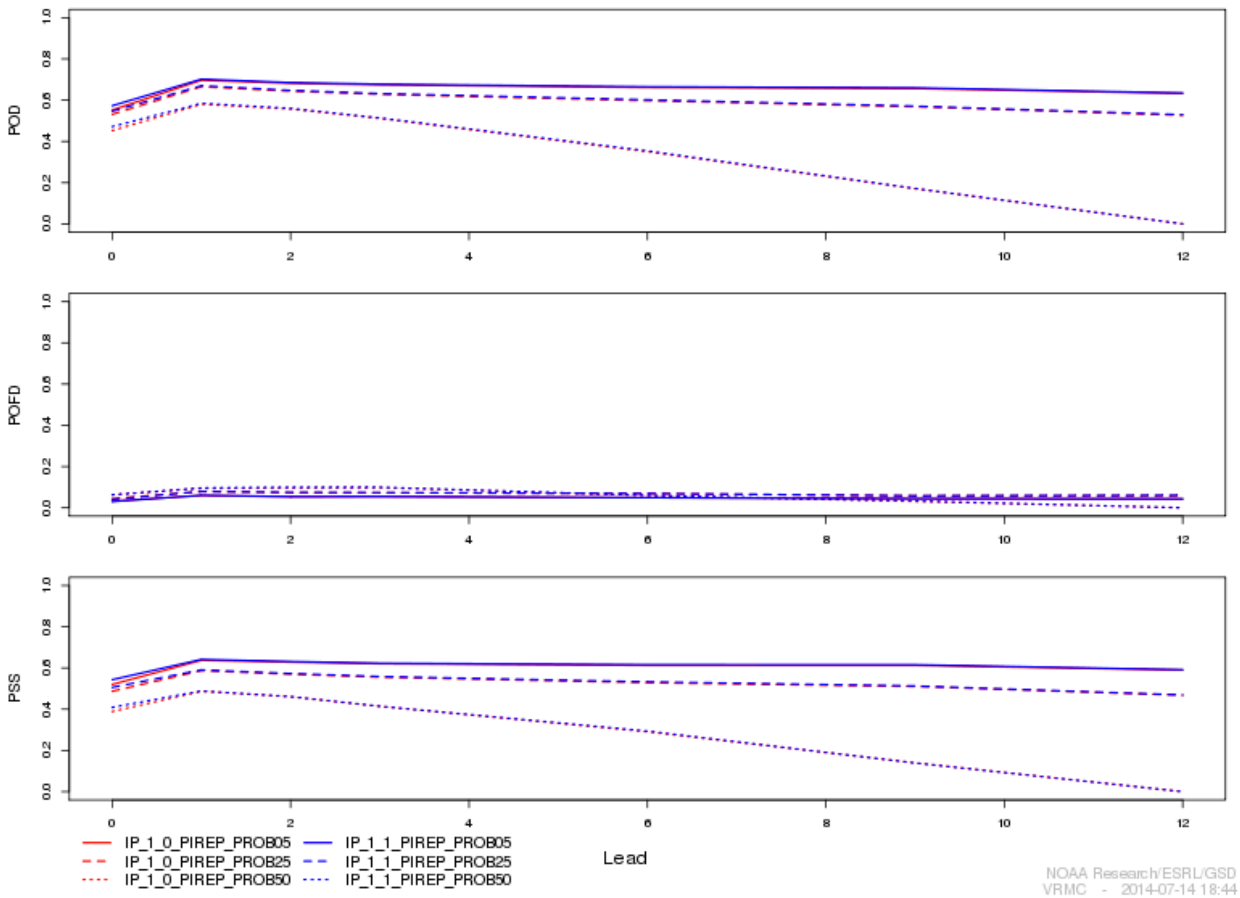

Figure 4.6: The probability of detection (POD, top), probability of false detection (POFD, middle), and Pierce skill score (PSS, bottom) of icing conditions from CIP 1.1 and 1.0 and FIP 1.1 and 1.0, by lead time. CIP and FIP 1.1 are in blue, CIP and FIP 1.0 are in red. Scores are presented for three different probability forecasts: 5% (solid line), 25% (dashed line), and 50% (dotted line).

4.3 Skill – SLD

Both PIREPs and METARs are used to verify SLD. It is acknowledged there are likely to be few PIREPs with SLD (as this is an area aircraft avoid). Skill scores are provided separately for each observation set.

For the FIP ADDS display, SLD potential is converted to a yes/no, where all SLD potential ≥ 0.05 is defined as a 'yes' forecast of SLD. The SLD field has an additional category, considered 'unknown'. For the ADDS display these grid points are treated as 'no' forecasts; however, the performance of the CIP/FIP SLD forecasts will be assessed considering each of the three possible treatments of the 'unknown' points. POD, POFD, and PSS scores will be calculated for CIP/FIP three ways: 1) considering the 'unknown' points as 'yes' forecasts 2) considering them as 'no' forecasts, and 3) leaving them as 'unknown'—essentially removing those points from the verification.

As expected from the results of the comparison of the distributions of the 1.1 and 1.0, the FIP forecasts, leads 1–12, show identical skill, as shown in figure 4.7 and 4.8. The only difference between the products is at a lead time of zero, the CIP. When compared to PIREPs (figure 4.7) CIP 1.1 is slightly better than CIP 1.0, and that difference is most pronounced when an SLD forecast of 'unknown' is treated as 'no'. Treating a CIP or FIP SLD forecast of 'unknown' as 'yes' increases both the POD and POFD but does not substantially change skill score. Treating SLD forecast of 'unknown' as unknown (not using it in generating the results) yields the best skill score; however, in forcing a definitive treatment of the SLD 'unknown'—as users of CIP/FIP do not have the option of treating a forecast of 'unknown' as unknown—considering the SLD 'unknown' areas as 'yes' forecasts yields greater skill than considering them as 'no' forecasts.

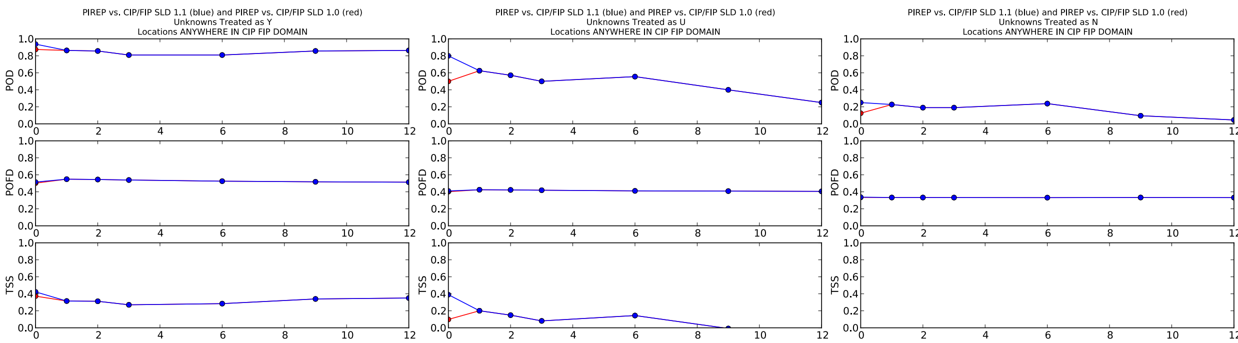

Figure 4.7: The probability of detection (POD), probability of false detection (POFD), and skill score (TSS) of CIP and FIP 1.1 (blue) and 1.0 (red) forecasts of SLD as compared to PIREP observations of SLD. Forecasts of SLD as 'unknown' are treated as 'yes' forecasts in the left panel, are ignored in the center panel, and treated as 'no' forecasts in the right panel

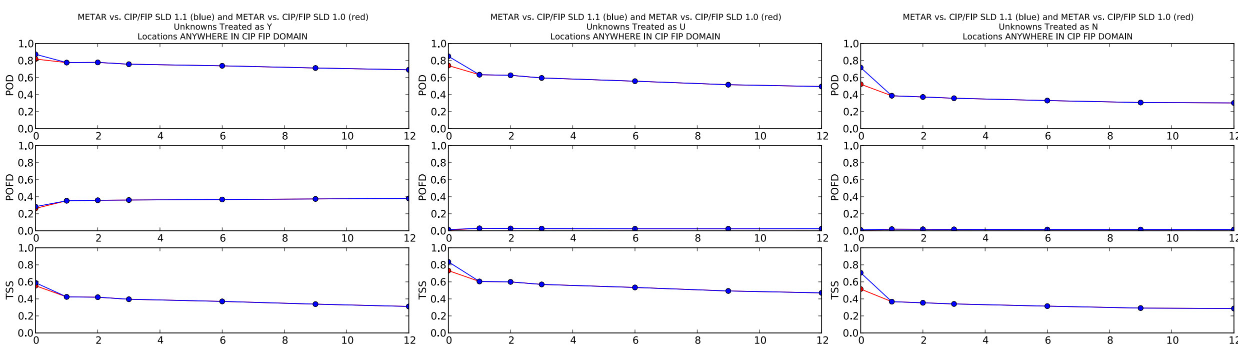

Figure 4.8: The probability of detection (POD), probability of false detection (POFD), and skill score (TSS) of CIP and FIP 1.1 (blue) and 1.0 (red) forecasts of SLD as compared to METAR observations of SLD. Forecasts of SLD as 'unknown' are treated as 'yes' forecasts in the left panel, are ignored in the center panel, and treated as 'no' forecasts in the right panel.

4.4 DIFFERENCE MAPS

A domain-wide comparison of CIP/FIP versions 1.0 and 1.1 is accomplished using difference maps. Spatial distributions are derived by aggregating counts of CIP and FIP field values exceeding a threshold (e.g., 5% for SLD, 25% for probability, MOG for severity) over a date range, issue and lead times, and vertical layers as defined in section 2.3, for each grid point. In addition, difference maps of the counts are generated where positive values indicate a higher HiRes CIP/FIP 1.1 count and negative values indicate a higher HiRes CIP/FIP 1.0 count.

There is no difference between the FIP 1.1 and FIP 1.0 probability of icing forecasts. For CIP 1.1 and CIP 1.0, the differences in the CIP fields of probability of icing show characteristics of differing processes for their respective construction. Figure 4.9 shows that for the lowest altitudes, the impact of Canadian surface data leads to differences between 1.1 and 1.0. At higher altitudes, satellite data over the ocean seems to give the biggest differences. At middle altitudes, it is a combination of these two effects.

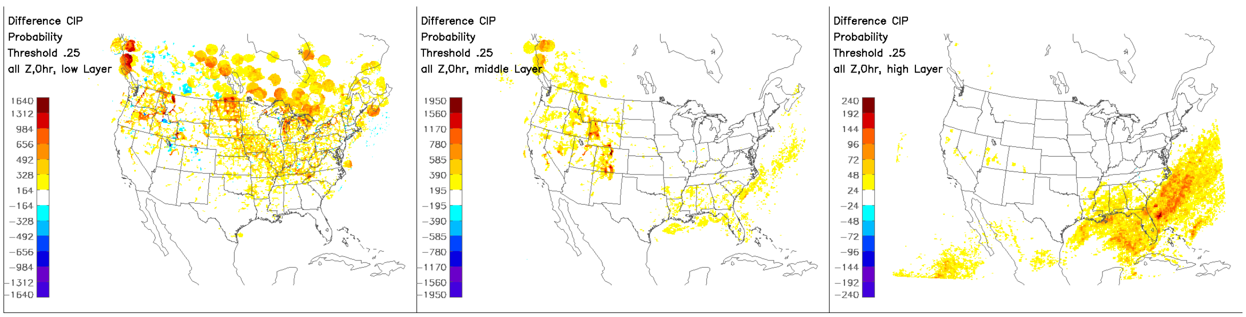

Figure 4.9: Differences between counts of grid points with probability of icing greater than 25% of CIP 1.1 run at NCAR and CIP 1.0 run at AWC as a function of altitude. Altitudes are 500–10,000 feet (left), 10,500–20,000 feet (center) and 20,500–30,000 feet (right)

When comparing the severity of icing in 1.1 to 1.0, there are differences in both the FIP and CIP. The FIP differences are a modest increase in severity level, typically in areas of higher terrain. There is no immediately obvious reason for this difference in FIP severity, as the probability of icing and SLD fields are identical, and as mentioned earlier could be attributed to round-off error in computing the severity categories. As for the CIP severity fields, the differences between 1.1 and 1.0 are, as with probability of icing, likely attributable to differences in the use of Canadian surface observations and of satellite data at NCAR and at AWC.

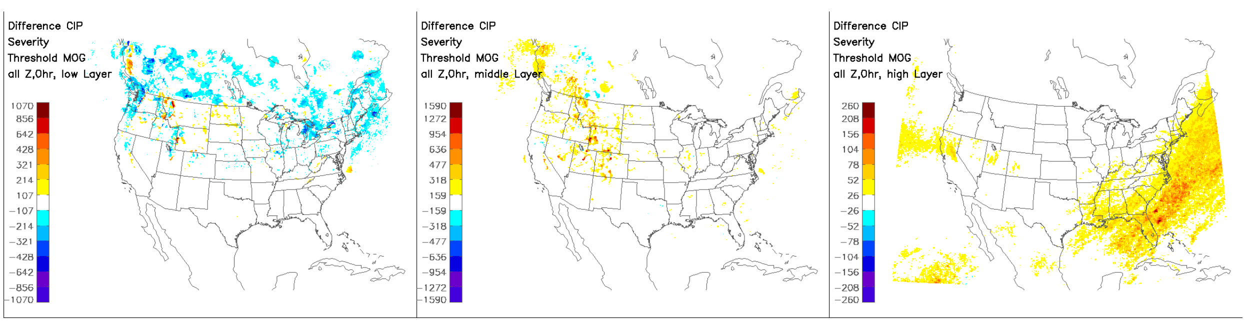

Figure 4.10: Differences between counts of grid points for MOG severity between CIP 1.1 and CIP 1.0 as a function of altitude. Altitudes are 500–10,000 feet (left), 10,500–20,000 feet (center) and 20,500–30,000 feet (right).

The differences between the CIP 1.1 provided by NCAR, and CIP 1.0 provided by AWC, are similar to those seen between CIP 1.0 provided by NCAR and CIP 1.0 provided by AWC. Figure 4.11 shows the differences between 1.1 and 1.0, and the differences between 1.0 (from a prior assessment) run at the two different locations. Though the images are not identical, there are similarities in the locations and characteristics of the differences.

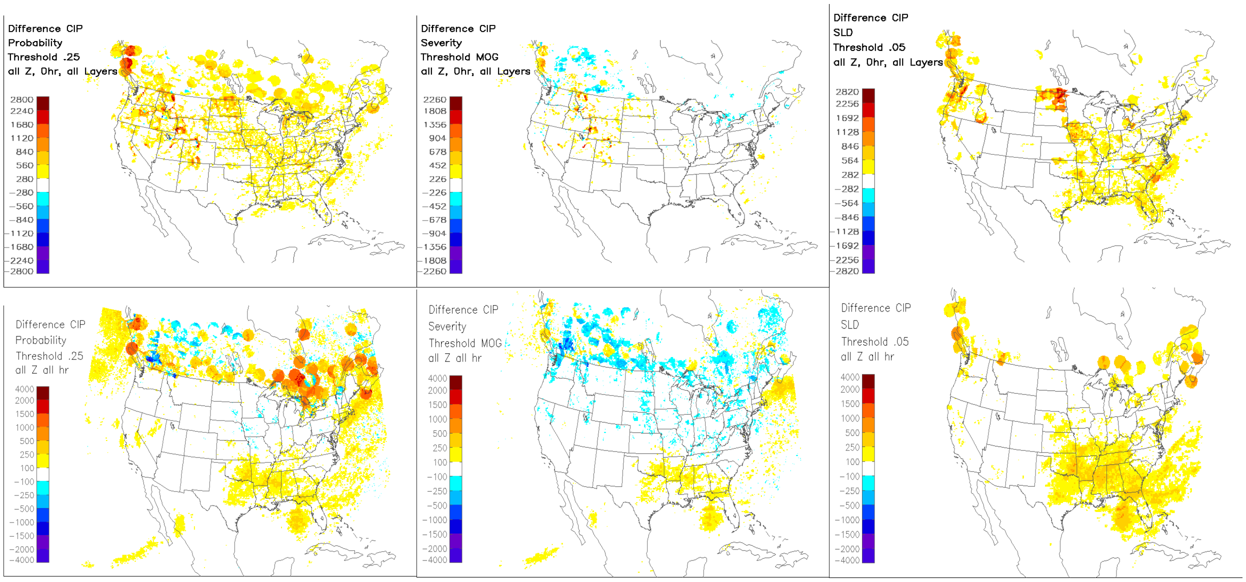

Figure 4.11: Differences computed for CIP 1.1 (NCAR) vs CIP 1.0 (AWC) for current assessment (top), as compared to differences computed for CIP 1.0 (NCAR) vs CIP 1.0 (AWC) from a previous assessment (bottom).

The differences between 1.1 and 1.0 are likely not simply differences in the CIP algorithm – they also arise from running CIP at two different locations.

5 Conclusions

Skill evaluation of severity indicates slight improvement (though not likely statistically significant) in CIP 1.1 as compared to 1.0, with FIP skill essentially identical. For SLD, the skill was identical for FIP, with improvement for the CIP in 1.1, most notably for cases when CIP SLD 'unknown' is treated as unknown (not using it in generating performance results) or 'no'. In general, treating SLD forecasts of 'unknown' as unknown yields the best skill score; however, in forcing a definitive treatment of the SLD 'unknown'—as users of CIP/FIP do not have the option of treating a forecast of 'unknown' as unknown—considering the SLD 'unknown' areas as 'yes' forecasts yields greater skill than considering them as 'no' forecasts.

The most notable difference in field distributions was with CIP SLD, which nearly doubled in amount. This was expected, due to the known changes in the algorithm, as well as the fact that the SLD field is smallest of all fields investigated. Difference maps indicated slight differences between fields in specific regions, which were also consistent with differences identified in a previous assessment of CIP/FIP data produced at AWC as compared to NCAR, and could possibly be due to the differences in observation set availability (METARs, satellite) for the two organizations. While the slight improvements in skill therefore cannot be entirely attributed to algorithm changes, findings strongly indicate that the algorithm changes have caused no detriment to performance and have resulted in slight improvements for CIP, with FIP skill remaining the same.

6 REFERENCES

Benjamin, S.G., T.G. Smirnova, K.J. Brundage, S.S. Weygandt, T.L. Smith, B. Schwartz, D. Devenyi, J.M. Brown, and G.A. Grell, 2004: A 13-km RUC and beyond Recent developments and future plans. 11th Conference on Aviation, Range, and Aerospace Meteorology, Hyannis, MA, October 2004, American Meteorological Society (Boston).

Brown, and G.A. Grell, 2004: A 13-km RUC and beyond Recent developments and future plans. 11th Conference on Aviation, Range, and Aerospace Meteorology, Hyannis, MA, October 2004, American Meteorological Society (Boston).

Bernstein, B.C., F. McDonough, M.K. Politovich, B.G. Brown, T.P. Ratvasky, D.R. Miller, C.A. Wolff and G. Cunning, 2005: Current Icing Potential (CIP): Algorithm description and comparison with aircraft observations. J. Appl. Meteor., 44, 969-986.

Brown, B.G., and G.S. Young, 2000: Verification of icing and icing forecasts: Why some verification statistics can't be computed using PIREPs. Preprints, 9th conference on Aviation, Range, and Aerospace Meteorology, Orlando, FL, Sep. 11-15, American Meteorological Society (Boston), 393-398.

Carriere, J.M., S. Alquier, C. LeBot, and E. Moulin, 1997: Statistical Verification of Forecast Icing Risk Indicies, Meteorological Applications, Vol 4, Issue 2, p.115-130.

Chapman, M., M. Pocernich, A. Holmes, P. Boylan, P. Kucera, B.G. Brown, J.L. Mahoney, and J.T. Braid, 2007: Quality Assessment Report: Forecast Icing Product (FIP) – Severity. Report to the FAA Aviation Weather Technology Transfer Board. Available from the Quality Assessment Research Team (Jennifer.Mahoney@noaa.gov).

Kay, M.P., C. Lu, S. Madine, J. L. Mahoney, and P. Li, 2009: Detecting cloud icing conditions using CloudSat datasets. Preprints, 23rd Conference on Weather Analysis and Forecasting, 1-5 June, Omaha, NE, Amer. Met. Soc.

Madine, S., S. A. Lack, S. A. Early, M. Chapman, J. K. Henderson, J. E. Hart, and J. L. Mahoney, 2008: Quality Assessment Report: Forecast Icing Product (FIP).

McDonough, F., B.C. Bernstein, and M.K. Politovich, 2003: The Forecast Icing Potential (FIP) Technical Description. Report to the FAA Aviation Weather Technology Transfer Board. Available from M.K. Politovich (NCAR, P.O. Box 3000, Boulder, CO, 80307), 30 pp.

Murphy, M. P., 2010: Product Description Document, Graphical Airman's Meteorological Advisory (G-AIRMET). Available at: http://aviationweather.gov/static/docs/gairmet/G-AIRMETPDD_2010.pdf

NWS, 2007: Aviation Weather Services, Advisory Circular AC 00-45F. U.S. Department of Commerce, National Oceanic and Atmospheric Administration, National Weather Service, and U.S. Department of Transportation, 393 pp.