9.7 EVALUATION OF LOW-LEVEL AIRCRAFT ICING FORECASTS FROM MT. WASHINGTON SUMMIT OBSERVATIONS

Rebecca J. Stanley and James P. Koermer
Plymouth State College
Plymouth, NH 03264

Charles C. Ryerson
US Army Cold Regions Research and Engineering Laboratory
Hanover, NH 03755

1. INTRODUCTION

Aircraft icing has been the focus of many studies due to its importance to flight safety. Past research has been conducted to explain the mechanism of aircraft icing (Politovich, 1989), the synoptic conditions or climatology associated with aircraft icing (Bernstein et al., 1997; Curry & Liu, 1992), as well as the forecasting of aircraft icing (Kelsch & Wharton, 1996; Shultz & Politovich, 1992; Thompson et al., 1997; Tremblay et al, 1995; Tremblay et al., 1996). In the majority of these previous studies, complex topographic regions were neglected in order to draw more general conclusions. The goal of this project was to examine the accuracy of aircraft icing forecasts in a region of complex terrain, and possibly develop more effective methods of forecasting icing in these regions.

This study focuses on Mt. Washington (KMWN) in northern New Hampshire, from December 2000 through April 2001 at the 00UTC and 12UTC time periods. In situ icing intensity data were gathered continuously on the summit for use in evaluating the accuracy of four different icing forecasting methods. Icing nowcasts and forecasts were compiled from the following sources: an algorithm from the March 1997 FYI Icing publication (throughout this paper referred to as the "AWSP" algorithm) (Air Weather Service, 1997), Aviation Weather Center (AWC), Air Force Weather Agency (AFWA), and the National Center for Atmospheric Research (NCAR). Icing forecasts examined were limited to altitudes below 10,000ft for the KMWN area.

The authors hypothesized that most forecasts would not predict icing to the extent that it occurs on the summit. Terrain induced lifting creates additional adiabatic cooling and increased humidity, which would help to produce more frequent and heavier icing events in the lower layers of the atmosphere. Comparisons of the forecast accuracies were completed, as well as a statistical analysis of the icing intensity correlations. Adaptations to the AWSP algorithm to account for the temperature and humidity variations over complex terrain are also suggested.

2. OBSERVATIONAL DATA

Mt. Washington stands as the highest peak in New England at 1,905m (6,288ft), and is staffed year round as an observation and research facility. Icing occurrences on or near the summit were thoroughly documented during the study period from December 2000 through April 2001 by archiving hourly weather observations, Pilot Reports (PIREPs) in the vicinity, and using two icing detectors that measure in situ icing data.

Atop the tower of Mt. Washington's summit building, a Rosemount ice detector was mounted and maintained by the Army's Cold Regions Research and Engineering Laboratory (CRREL) to gather continuous ice accretion rates. A rotating multi-cylinder device was also placed on the summit tower only during icing conditions (Stanley, et al.). CRREL compiled and analyzed these data from both instruments.

The hourly KMWN summit observations are manually taken in standard National Weather Service METAR format. These observations provided data for a quality check the icing instrument data.

PIREPs were archived for the entire US for the period of study by AWC. PIREPs that were recorded in the vicinity of KMWN (approximately within a 50km radius) were extracted, and then manually sorted to find only reports below 10,000ft that reported positive or negative icing. However, this generated only four applicable PIREPs for the entire period of study.

3. AWSP, AFWA, AWC & NCAR FORECASTS

The AWSP icing algorithm is an empirical forecasting decision tree developed by the Air Force and shown in figure 1. The algorithm uses input of several environmental variables and yields a forecast of icing type and intensity. The algorithm was evaluated using ETA model interpolated sounding for KMWN at 12-hour intervals from 00hrs to 48 hrs out (Stanley et al., 2002). This interpolated sounding is tainted by the inherent errors of the ETA model, the most important being its inability to capture realistic terrain features. The ETA topography for the KMWN area has been smoothed so that the summit appears at less than half of its actual height.

The AFWA, AWC and NCAR icing predictions were archived by their respective organizations for comparison purposes.

Corresponding author address: James P. Koermer, MSC#48, Natural Science Dept., Plymouth State College, Plymouth, NH 03264; e-mail: koermer@mail.plymouth.edu.

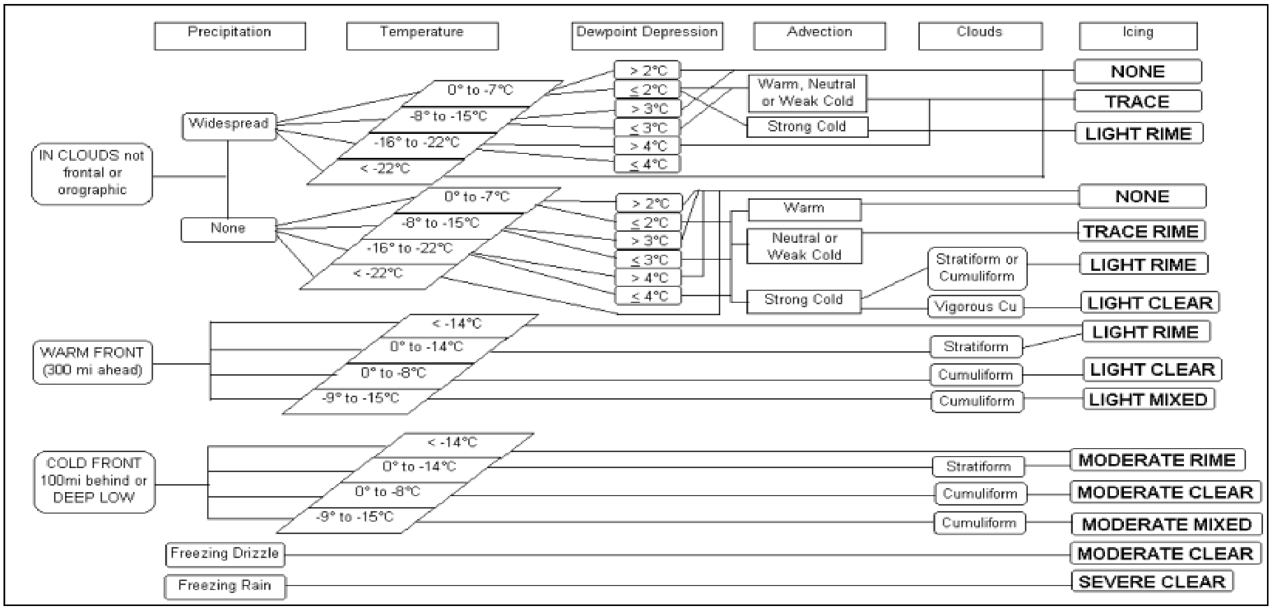

Figure 1. AWSP Algorithm adapted from Air Weather Service (1997).

AFWA produced extracted model forecast results specifically for KMWN that gave icing intensity (none, light, moderate or severe), but was not type specific. AFWA uses MM5 model data, run at both 15km and 45km resolution for this study, yielding two sets of predictions (hereafter the 15km data set is used for comparison). Forecast data was available from 06hrs to 54hrs out at 12-hour intervals.

The AWC archived two sets of icing predictions as well. The first is NNICE, a neural network artificial intelligence program that recognizes patterns of upper air data (specifically relative humidity, temperature, and stability) that correlate to an icing intensity scale. The second is an experimental VVICE that uses a cloud physics model combined with an aerodynamic effect analysis to produce icing intensity. AWC forecasts were extracted from the NNICE and VVICE grids using the nearest grid point to KMWN and taking the highest icing value below 700hPa. The thresholds provided by AWC for specifying a particular forecast from VVICE were quite subjective and may need some adjustment. Forecast data out to 12hrs were available.

NCAR's Integrated Icing Diagnostic Algorithm (IIDA) differs slightly from the other algorithms in this study in that it evaluates icing potential rather than predicting icing intensity. The algorithm uses RUC model data integrated with satellite, surface and radar observations. Since it uses real-time input data, it is available only as a nowcast. In this study it is compared only in the 'yes/no' verification because it does not evaluate icing intensity; icing potential only indicates the chance of icing being present regardless of intensity.

4. VERIFICATION

Upon completion of nowcasts and forecasts, a preliminary icing verification was done using a "yes/no" format. This evaluation compares the icing forecasts and icing occurrences on the summit, neglecting icing intensity.

Figure 2 summarizes the results of the nowcast initial verification. Icing predictions fall into one of three categories: correct icing forecast (includes correct icing and non-icing forecasts), false positive forecasts (where icing was predicted, yet was not observed), and false negative forecasts (where no icing was predicted, yet icing was observed). It is clearly shown that the AWSP and IIDA predictions were the most accurate through the study period. It is also interesting to note the distribution of false positive and false negative forecasts, which will be discussed later in this paper.

This initial verification described above was very useful in gathering a general impression of the accuracy of the icing algorithms compared in this study. However, most of the icing forecasts given were in terms of icing intensity, and the evaluation of this became our focus for a secondary verification. A statistical analysis of the icing intensity accuracies was performed by CRREL, and the results are summarized below.

Forecasts proved to be almost as accurate as nowcasts for most prediction sets. The ETA AWSP algorithm had a slow decline in accuracy as forecast time increased; between the 00hr nowcast and 48 hour forecast there was about a 10% decrease in accuracy. The AFWA, NNICE and VVICE had quite consistent accuracy percentages even as the forecast times increased, i.e. the longer term forecasts were nearly as accurate as the nowcasts.

In order to compare the in situ data to the icing severity predictions, liquid water content (LWC) calculated from the ice detector and rotating multicylinders was used for analyses. We computed an average LWC for each icing intensity of each forecast method and used the statistical analyses to assess if significantly different LWCs were associated with each

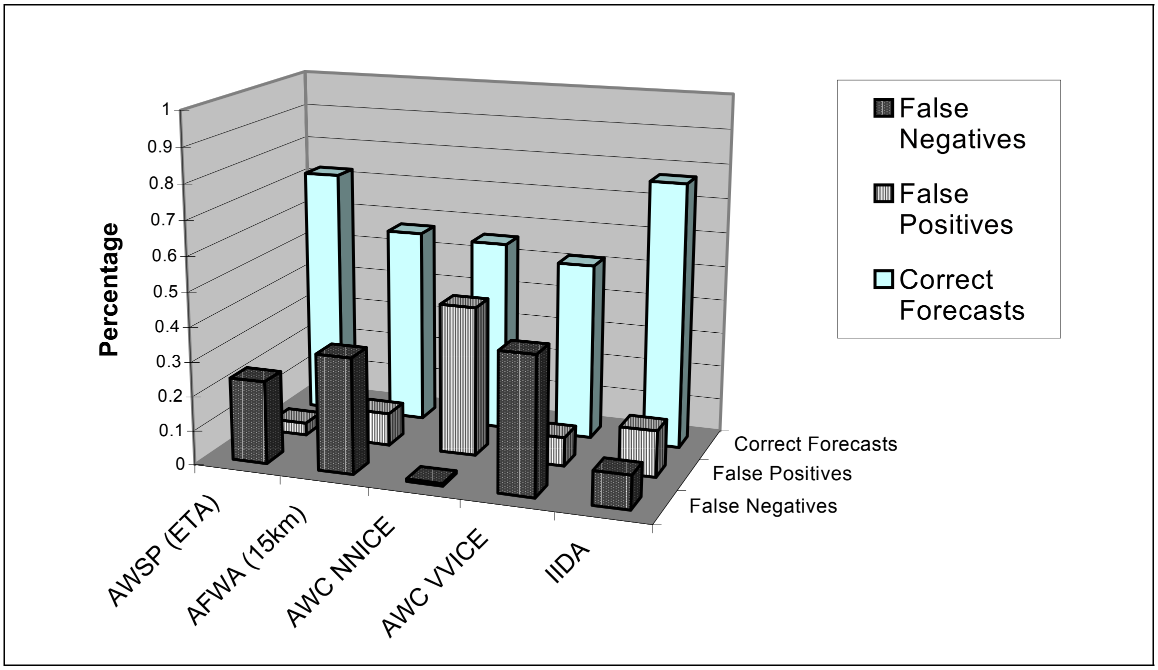

Figure 2. Nowcast Results from Preliminary Verification.

icing intensity. That is, if two icing intensities have the same, or nearly the same, LWC for a given forecast method, then that method does not discriminate well between those two intensities.

The analysis was performed using a Mann-Whitney U test (similar to a parametric T test) due to the small sample size and non-gaussian distribution of the data. Values of the Mann-Whitney U test that approach zero signify that there is a significant difference in the LWC between the two categories evaluated. Figure 3 shows the results of the Mann-Whitney U test for each forecast source, between each level of icing intensity. [Note: the NCAR IIDA data were not included since it predicts probability of icing rather than icing intensity and therefore is not comparable.]

It was shown that for the ETA AWSP, AFWA, and AWC VVICE, there was a significant difference in the LWC of the atmosphere between the severity levels of none compared with all trace, light and moderate. No forecast method predicted severe icing during the observation period. Therefore, there is high degree of accuracy when determining icing vs. no icing conditions from these icing prediction sources. The AWC NNICE was unable to distinguish LWC levels described above. However, NNICE was the only algorithm to establish a clear difference in the atmospheric LWC values between the severity levels of moderate versus none, trace and light.

5. ANALYSIS

Through further analysis, we found that adaptation of the AWSP algorithm could lead to improved forecasts in regions of complex terrain. The ETA AWSP predictions had almost no occurrence of false positives, indicating overprediction is not a problem.

The ETA AWSP false negative nowcasts were examined and separated based on the most likely reason the algorithm led to an incorrect negative icing forecast. The results of this analysis clearly show that lack of humidity in the ETA prediction is the primary cause for false negative icing forecasts (Stanley et. al., 2002), since 67 out of 100 false negative predictions were due to dewpoint depressions that were too large. This confirms that the false negative forecasts are primarily due to low humidity values, an underestimation most likely common in models over regions with complex terrain.

Two methods of modifying the AWSP algorithm were pursued: the first was to adjust the humidity threshold to allow great dewpoint depressions to indicate icing, and the second is to adjust the ETA input data before using the algorithm.

Frequency distributions were constructed to show the result of expanding the dewpoint depression threshold on the AWSP algorithm. It was shown that raising the threshold up to a certain point would be beneficial in improving the overall accuracy. The false negative forecasts were lowered significantly, without causing a large increase in the number of false positives.

The other method to improve the icing forecasts would be to alter the input data by adjusting it for additional adiabatic lifting. One way to handle this is by using the ETA model conditions for the model summit, and then adjusting them for dry adiabatic ascent to the height of the actual summit. Our software permits us to obtain the parcel LCL from the ETA model data, which is the pressure level of the LCL location

ETA	none	trace	light	moderate
none				
trace	*<0.0001*			
light	*<0.0001*	0.0396		
moderate	*<0.0001*	0.3526	0.2109	

AFWA	none	light	moderate
none			
light	*0.0002*		
moderate	0.0027	0.01942	

NNICE	none	trace	light	moderate
none				
trace	0.4015			
light	0.735	0.098		
moderate	0.0075	*<0.0001*	*<0.0001*	

VVICE	none	trace	light	moderate
none				
trace	*0.0004*			
light	0.0256	0.1928		
moderate	0.051	>.9999	0.4636	

Figure 3. Statistical Results from Icing Intensity Verification.

when using averaged temperature and dewpoint over the first 100mb above the surface. This parcel LCL, which was available for only two months due to archiving difficulties, was examined and compared to icing occurrences on the summit An initial analysis showed that it was highly accurate in placing the height of the LCL at or below that of the actual summit when icing was occurring. Further study may reveal this to be a good method of adjusting the model dewpoint input for an improved AWSP algorithm for regions with complex terrain.

6. CONCLUSIONS

Through verification of different aircraft icing forecasts for the period from December 2000-April 2001, it appears that forecasts of aircraft icing over the complex terrain could be improved. No forecasting method produced accuracies higher than 77% for the study period as a whole, and the average accuracy was around 60%. Forecast data showed slow deterioration of accuracy (ETA AWSP), or no deterioration (AFWA, NNICE, VVICE) as forecast time increased.

There is potential for manipulation of the ETA AWSP algorithm for improvement in regions of complex terrain. Initial research shows that adaptation of the algorithm by adjusting the input data shows success in improving accuracy percentage as well as improving the accuracy of icing intensity forecasts.

Further research over more than a single season is needed to study the adaptation of icing forecasting methods over complex terrain. It is shown that there is potential for an algorithm to be developed that can incorporate the atmospheric dynamics over terrain regions more successfully than current methods. With input data coming from the ETA model, forecasts can be made at 6-hour intervals out to 60 hours. Advanced warning and accurate icing forecasts over topographic regions has widespread importance, and there appears to be promise for improvement.

7. ACKNOWLEDGMENT

This research was performed 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 or position of the FAA.

8. REFERENCES

Air Weather Service, 1997: Icing. *FYI*, **38**, A-1.

Bernstein, B., et al., 1997: The relationship between aircraft icing and synoptic scale weather conditions. *Wea. Forecasting*, **12**, 742-762.

Claffey, K., K. Jones, and C. Ryerson, 1995: Use and calibration of Rosemount ice detectors for meteorological research. *Atmos. Research,* **36**, 321-337.

Curry & Liu, 1992: Assessment of aircraft icing potential using satellite data. *J. Appl. Meteor.*, **31**, 605-621.

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

Politovich, M.K., 1989: Aircraft icing caused by large supercooled droplets. *J. Appl. Meteor.*, **28**, 856-868.

Shultz, P., and M. K. Politovich, 1992: Toward the improvement of aircraft icing forecasts for the continental United States. *Wea. Forecasting*, **7**, 491-500.

Stanley, R., et al., 2002: Forecasting Aircraft Icing in Complex Terrain. *Third Symp. Of Environ. Applications*, Orlando, FL, 42-49.

Thompson, G., et al., 1997: Intercomparison of in-flight algorithms Part I: WISP94 real-time icing prediction and evaluation program. *Wea. Forecasting*, **12**, 878-889.

Tremblay, A. et al., 1996: An intercomparison of mesoscale forecasts of aircraft Icing using SSM/I retrievals. *Wea. Forecasting*, **11**, 66-77.

_____, et al., 1995: Forecasting of supercooled clouds. *Mon. Wea. Rev.*, **123**, 2098-2113.