4.4 FORECASTING AIRCRAFT ICING IN COMPLEX TERRAIN

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

Ingrid Gotchel and Gordon R. Brooks
Air Force Weather Agency
AFWA/DNXT
Offutt AFB, NE 68113

Clinton E. Wallace and David Knapp
Aviation Weather Center
NCEP/NWS/NOAA
Kansas City, MO 64153

1. INTRODUCTION

Aircraft icing has been the focus of many studies due to its importance with regard 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 study aircraft icing forecast techniques in a region of complex terrain to determine the accuracy, and possibly develop more effective methods of forecasting icing.

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 was gathered continuously on the summit and hourly summit observations were archived. Icing nowcasts and forecasts were compiled for the study period from three sources, the first used was an algorithm from the March 1997 FYI Icing publication (Air Weather Service, 1997). This algorithm will be referred to throughout this paper as the "AWSP algorithm". Input data for the AWSP algorithm was gathered from ETA grid interpolations for KMWN and nearby radiosonde data. The Aviation Weather Center (AWC) and the Air Force Weather Agency (AFWA) archived icing nowcasts and forecasts for comparison. 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 because the terrain induces additional adiabatic cooling and increased humidity. These factors 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. 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 using three methods. Hourly weather observations on the summit were archived, and two quantity-sensitive icing detectors, a Rosemount sensor and a multi-cylinder, were used to compile in situ icing data. Pilot Reports (PIREPs) that indicate positive or negative icing below 10,000ft in the vicinity of KMWN were also gathered.

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. It was positioned to face the northwest, the mean climatological wind direction for the winter months. The Rosemount ice detector is a small device with a cylindrical probe protruding out its top. The probe oscillates at a fixed frequency, A. Ice collects on the probe during icing conditions and the added weight dampens the oscillation of the probe. When the oscillation decreases to a predetermined frequency B, then a heater is activated to melt the ice from the probe.

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

When the probe is ice free, the heater shuts off and the probe begins oscillating again at frequency A. One cycle is measured as the time it takes for the frequency to decrease from A to B. During every cycle the same weight of ice accretes on the probe, and an icing intensity rate can be calculated from the cycle interval (Claffey et al., 1995). These data were archived and sent to CRREL for analysis.

The multi-cylinder is an electrically powered rotating device placed on the summit tower during icing conditions. The cylinder is oriented to point into the wind and remains outside for an average time of 20 minutes. The cylinder shaft is composed of six different sized cylinders that decrease in size as you move away from the base. Ice accretes on the cylinders, and the rotation of the device ensures equal ice distribution over the circumference of each cylinder. Once adequate ice coverage has formed the multi-cylinder is taken inside and each individual cylinder is weighed and measured to determine the mass and thickness of ice that has built up. Using these measurements liquid water content and droplet size of the atmosphere can be determined so that an icing severity level can be estimated. CRREL also analyzed these data.

The hourly Mount Washington observations are manually taken and are in standard National Weather Service METAR format and provide data including temperature, dewpoint, relative humidity, wind speed and direction, and present weather conditions. Observations containing reports of freezing fog were manually extracted, and assumed to be synonymous with icing conditions on the summit. This provided data for a “yes/no” verification of the icing forecasts based on summit surface observations, and are described later in this paper.

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. This generated only four applicable PIREPs for the entire period of study.

2.1 *AWSP Algorithm*

The AWSP icing algorithm is an empirical forecasting device developed by the Air Force and shown in figure 1. The algorithm requires initial input of surface frontal position to begin the decision tree process. Input of some or all of the following environmental variables are also needed: precipitation, temperature, dewpoint depression (labeled as ‘spread’ in Figure 1), temperature advection, and cloud type. Following the branch system will yield a forecast of icing type and intensity. Two separate sets of nowcasts were completed with this algorithm, one using only data from the Gray Maine (GYX) radiosonde and the other from the ETA model interpolated sounding for KMWN. In addition, four sets of forecasts were completed using ETA interpolated soundings from 12, 24, 36, and 48 hrs.

Gray Maine is the closest radiosonde station to Mt. Washington, located approximately 100km to the southeast of KMWN. GYX has negligible topographic influence, with a station height of 115.2m above mean sea level, yet is susceptible to marine influence being less than 20km from the Atlantic Ocean. Data was gathered from the 850hPa level, which is the closest significant level to the summit of KMWN.

The ETA model data was taken from the sounding for KMWN, interpolated from the local gridpoints. This interpolated sounding is tainted by the inherent errors of the ETA model, the most important being the inability to capture realistic terrain features. The ETA topography for the area is so smooth for the KMWN region that the summit appears at less than half of its actual height.

The surface frontal situation was manually determined primarily using the National Weather Service (NWS) surface analysis charts. NWS depicted fronts were further analyzed to ensure they represented a substantial temperature gradient and/or wind shift boundary. Additional surface plots of the region depicting temperature and wind were referred to when analyzing more ambiguous situations such as stationary or occluded fronts. An effort was made to keep the process as objective as possible and the same procedures were followed for each analysis.

Some of the variables needed to make a forecast using the AWSP algorithm are more straightforward than others. An effort was made to determine each as objectively as possible. Below is a description of how each was determined. All values collected from the GYX sounding or the ETA interpolated soundings were done so for the 850hPa level.

The presence of precipitation was determined by relative humidity values greater than or equal to 95%. This was an input used only in non-frontal situations. No effort was made to determine freezing drizzle or freezing rain, and therefore that portion of the algorithm was ignored.

The temperature and dewpoint depression were taken from the sounding or interpolated sounding at each time period. They were taken to the nearest tenth of a degree. Nowcasts for the month of February, and all forecast periods, had temperature and dewpoint depression values rounded to the nearest whole degree as a result of archiving complications.

The temperature advection is a quantitative value from the ETA model expressed in °c/10,000sec. It was fit to the qualitative categories in the algorithm of warm, neutral, weak cold and strong cold as follows: An advection value of -1 (units of °c/10,000sec) to 1 inclusive was considered neutral advection, -1 to -2 represents weak cold, less than -2 is strong cold, and anything greater than 1 is classified as warm advection. The ETA nowcast advection values were used as the advection value for the GYX data set, since no comparable objective value was available from the GYX sounding.

Cloud type was based on the vertical velocity values gathered from the ETA model output. A positive vertical velocity value represents an upward vertical motion and was correlated to a cumuliform cloud. A

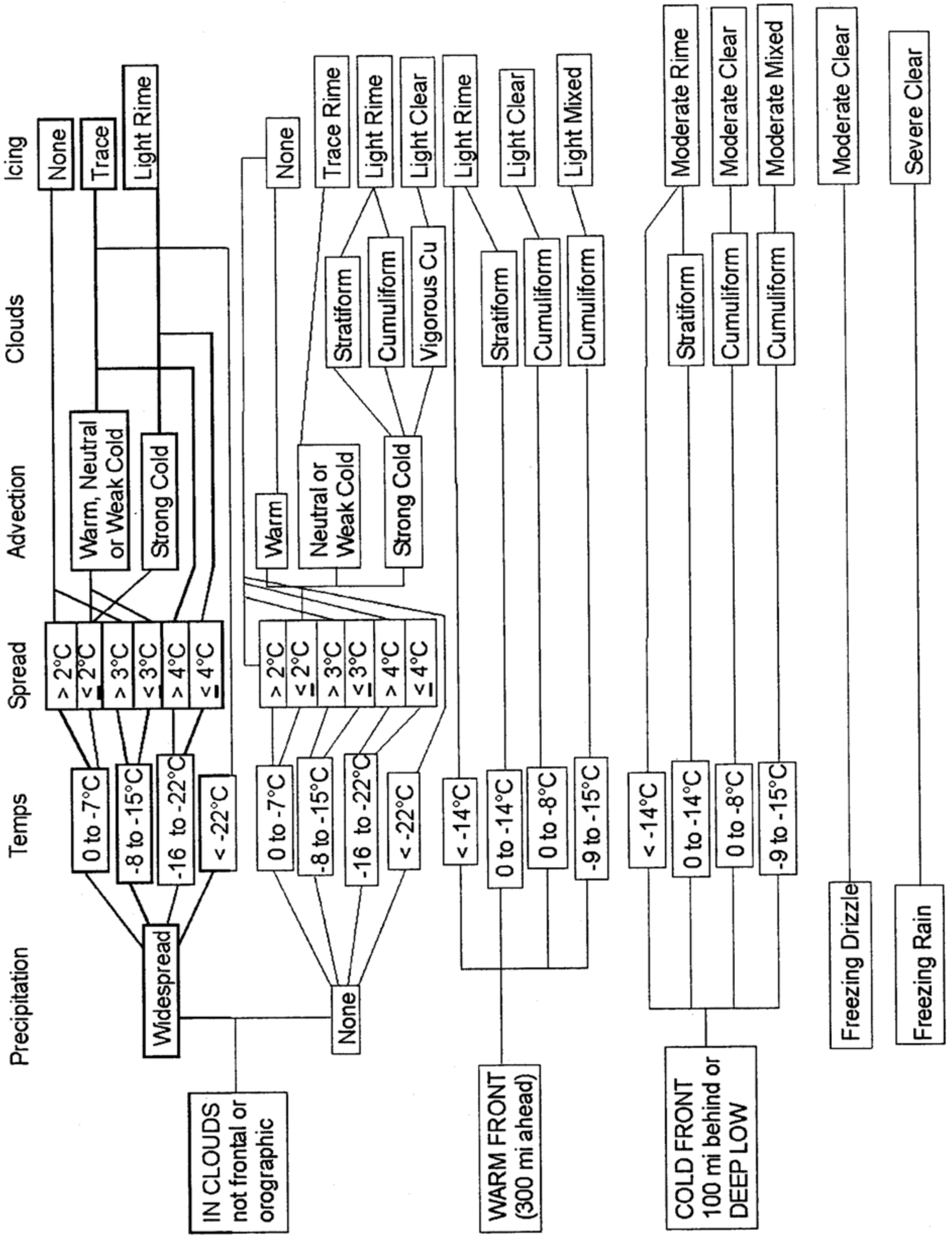

Figure 1. AWSP Algorithm (Air Weather Service, 1997)

negative value indicated a stratiform cloud. This ETA cloud-typing scheme was used for both the GYX and ETA nowcasts.

2.2 *AFWA and AWC Forecasts*

The AFWA and AWC icing predictions were archived by their respective organizations for comparison purposes. 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 inserted into an algorithm developed by NCAR/RAP, and further adjusted by AFWA. The MM5 was run at 15km and 45km resolution, yielding two sets of nowcasts.

AFWA forecasts were available at 06hr, 18hr, 36hr, 42hr, and 54hr using 45km resolution, and 06hr, 18hr, 36hr, and 42hr using the 15km resolution. The 06hr was used comparatively as the nowcast, since no 00hr prediction was available. The data set began on 16 December 2000, and had very few missing data throughout the remaining period.

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 to determine liquid water content and droplet size. This is then related to the aerodynamic effect on an airplane and from that a numerical output is given that corresponds to an icing intensity scale (none, trace, light, moderate, severe). 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. This could help explain some of the performance issues to be discussed later.

The AWC forecasts were available at 00, 06, and 12hrs for both NNICE and VVICE. The AWC data set had a significant amount of missing data. Predictions were available for 15 December 2000 through 26 January 2001, and then 22 March 2001 through the end of the period, with very few missing data among these dates. NNICE and VVICE had 145 and 162 available periods, respectively, out of a possible 302.

3. VERIFICATION

Upon completion of nowcasts and forecasts, a preliminary icing verification was done using a “yes/no” format. This comparison views the icing forecasts versus icing occurrences on the summit; intensity was not considered for these verifications. The nowcasts and forecasts were compared against summit surface observations as recorded by Mt. Washington Observatory staff (summit observation verification), and also to summit icing instrument data from CRREL (instrument data verification). Summit observations of freezing fog within three hours before or after the 00UTC or 12UTC time periods were considered positive icing occurrences. The summit icing instruments detected 19% fewer icing periods than the summit surface observations (when comparing time periods with no missing data). Causes of this discrepancy can be attributed to instances where fog was present yet the temperature was below the threshold (-22°C in the AWSP algorithm) for presence of super-cooled liquid water, or light winds prevented impaction and ice accretion on instruments.

Figure 2 shows the accuracy, and distribution of correct icing forecasts, correct non-icing forecasts, false positive and false negative results for all data sets for which we had data. The pie charts in Figure 3 show the graphical distribution of the same categories just for the individual nowcasts of the various methods with a legend in the upper right. These charts help to visualize the correct nowcasts and incorrect nowcasts, which are further divided to show how many were correct icing (plain) and correct non-icing (dots) predictions and how many resulted in false positive (vertical lines) and false negative (horizontal lines) events.

Some interesting conclusions are observed from this data and from Figure 2. Looking at each forecasting method separately we see some interesting strengths and weaknesses of each process.

The AWSP algorithm shows increased accuracy when using ETA model data versus GYX data. In fact, the 12hr ETA forecasts are more accurate than GYX nowcasts. An interesting and important observation is the very few number of false positives, which is beneficial in two ways. These results show agreement with the original hypothesis of not capturing the extent of the icing, and the lack of false positives make this process a good candidate for algorithm manipulation to capture terrain-induced effects. Second, this algorithm has good practical use, since icing forecasts almost always are indicative of icing and overall accuracy percentages are fairly high.

The AFWA forecasts also show low numbers of false positives, and accuracy percentages are good, but not great. Surprisingly, the 15km MM5 input data did not appear to yield overall improved icing forecasts compared to the 45km data.

The AWC NNICE and VVICE forecasts were of two extremes, as one tended to overpredict icing events, while the other generally underpredicted them. The NNICE nowcast predicted icing for all but 5 periods for the total period of study. It is essentially as accurate as a constant positive icing prediction all the time, in which the accuracy percentage is directly related to how much icing occurs. Although the accuracy percentages for NNICE are fairly high, the usefulness in day-to-day forecasting operations is not as high, as the NNICE accuracy percentage is essentially a climatological percentage of icing occurrence.

The VVICE showed the other extreme with very few overall icing predictions. Although the accuracy percentage is only around 50%, the usefulness of VVICE may exceed that of NNICE. AWC was missing data for about half of the time periods in

	AWSP						AFWA									AWC					
Entire Study Period Totals	ETA	GYX	ETA 12	ETA 24	ETA 36	ETA 48	45KM 06 HR	45KM 18 HR	45KM 30 HR	45KM 42 HR	45KM 54 HR	15KM 06 HR	15KM 18 HR	15KM 30 HR	15KM 42 HR	NNICE 00	NNICE 06	NNICE 12	WICE 00	WICE 06	WICE 12
Verification using KMWN Surface Obs																					
Correct Icing Forecasts	107	79	83	79	71	72	72	80	83	77	69	76	88	85	84	103	104	104	28	19	16
Correct Non-Icing Forecasts	87	83	86	84	71	71	66	68	62	58	59	65	67	57	54	4	3	5	54	55	48
Total Correct Forecasts	194	162	169	163	142	143	138	148	145	135	128	141	155	142	138	107	107	109	82	74	64
False Positives	3	5	2	4	6	6	12	8	13	19	17	14	10	17	20	55	56	54	5	4	11
False Negatives	98	119	118	118	106	93	104	97	94	97	105	96	85	89	89	1	0	0	75	84	87
Total Incorrect Forecasts	101	124	120	122	112	99	116	105	107	116	122	110	95	106	109	56	56	54	80	88	98
% Correct Forecasts	66%	57%	58%	57%	56%	59%	54%	58%	58%	54%	51%	56%	62%	57%	56%	66%	66%	67%	51%	46%	40%
Verification using Instrument Obs																					
Correct Icing Forecasts	94	67	71	68	58	61	61	66	67	63	56	61	69	68	68	74	75	75	19	16	11
Correct Non-Icing Forecasts	90	88	88	86	72	69	83	83	80	78	82	79	77	75	73	4	3	5	66	73	65
Total Correct Forecasts	184	155	159	154	130	130	144	149	147	141	138	140	146	143	141	78	78	80	85	89	76
False Positives	9	10	10	8	12	10	22	21	26	30	26	26	27	30	32	75	76	74	13	6	14
False Negatives	61	80	79	81	74	65	76	71	67	68	74	73	65	63	62	1	0	0	55	58	63
Total Incorrect Forecasts	70	90	89	89	86	75	98	92	93	98	100	99	92	93	94	76	76	74	68	64	77
% Correct Forecasts	72%	63%	64%	63%	60%	63%	60%	62%	61%	59%	58%	59%	61%	61%	60%	51%	51%	52%	56%	58%	50%

Figure 2. Number of occurrences of correct and incorrect forecasts is listed in the table. The percentages are adjusted for missing data.

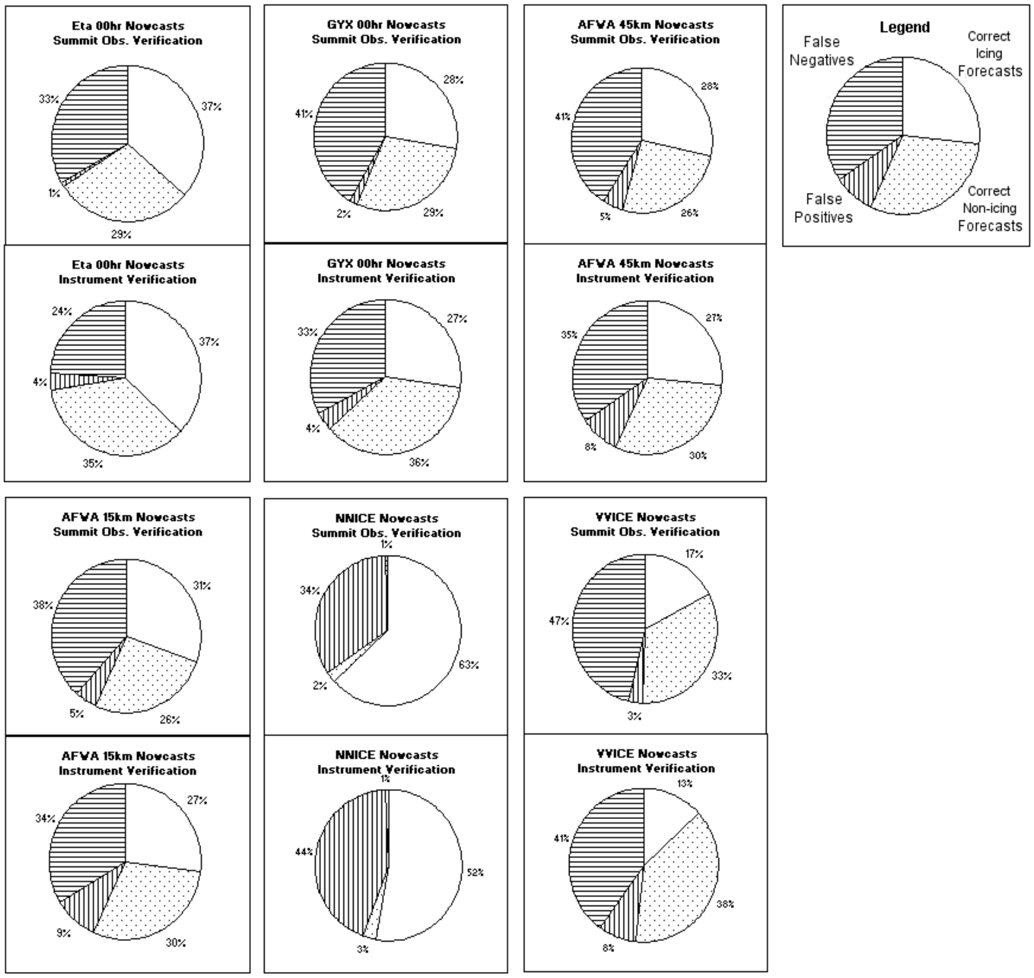

Figure 3. Legend is located in the upper right hand corner. Numbers corresponding to pie slices represent total percent of predictions in the specified category during the study.

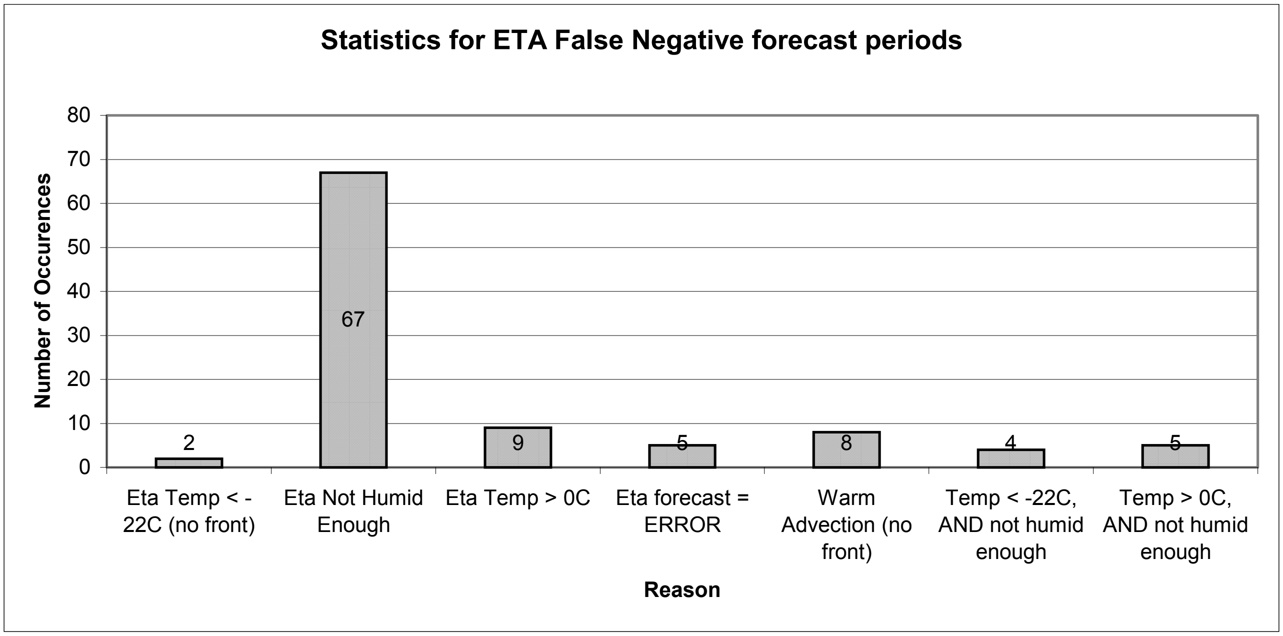

Figure 4. Reasons for ETA false negative forecasts during study period, based on the AWSP algorithm.

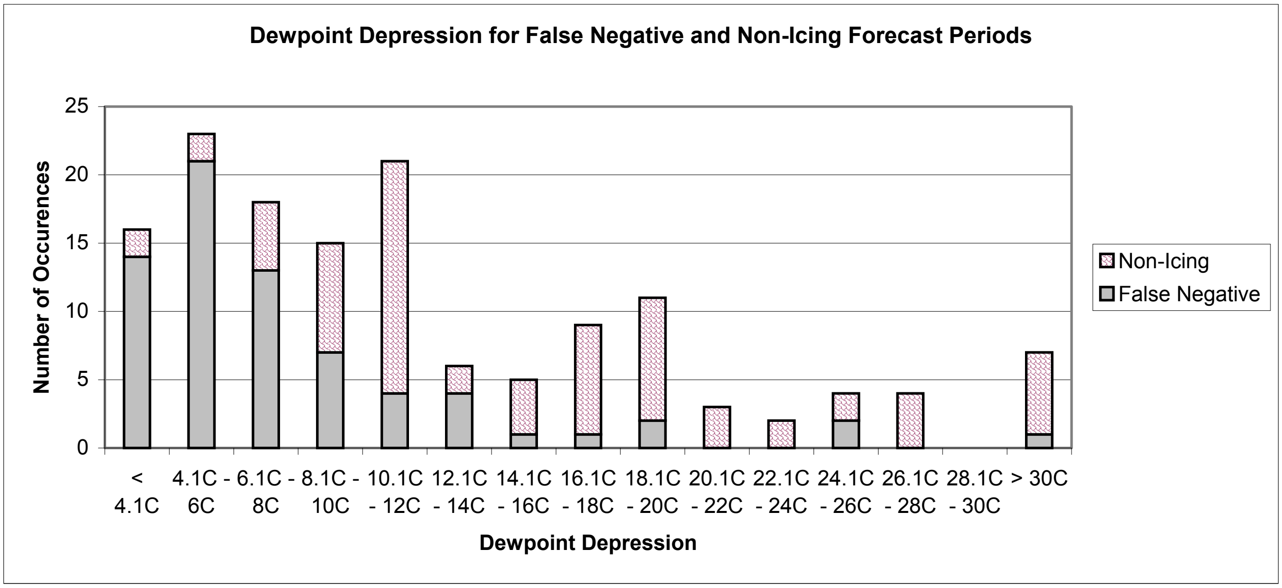

Figure 5. Dewpoint depression statistics for AWSP algorithm false negative and non-icing periods during the study period.

this study, and for equal comparison to other forecasting methods it would have been advantageous to have a complete set of forecasts. However, most of the missing data came during the months of February and March, which fall between the other data we received, and therefore the results would not likely be very different.

At the final writing of this paper, we received the NCAR IIDA model diagnostics and these results will be discussed at the conference.

4. ANALYSIS

As mentioned before, the ETA AWSP predictions had almost no occurrence of false positives. Although the ETA falls short of predicting all icing events, overprediction is not a problem. For this reason, the ETA nowcasts were further analyzed for the purpose of algorithm improvement for regions of complex terrain.

The ETA AWSP false negative nowcasts were examined and separated based on the most likely reason the algorithm led to a negative icing forecast (using summit observation verification). Figure 4 shows the results of this analysis. It becomes quite clear from looking at the graph that lack of humidity in the ETA prediction is the primary cause for false negative icing forecasts. 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 probably common in models over regions with complex terrain.

The dewpoint depressions for those 67 times were graphed in a frequency distribution (Figure 5—lower part of bar). This shows a distribution skewed towards lower dewpoint depression values. Shown on the same graph are the distributions of dewpoint depressions (upper part of bar) for non-icing periods.

One method of AWSP algorithm adjustment that could improve forecasts is to incorporate larger dewpoint depressions to be indicative of icing, such as, raising the icing threshold criteria to 8C or lower. Two changes would result from this adaptation. First, a number of false negative forecasts would become correct icing forecasts, and secondly, a smaller number of correct non-icing forecasts would become false positives. Again from Figure 5, we can see the number of forecasts that would be affected by such an adjustment. The lower portion of the three bars to the left shows how many false negatives would become correct forecasts, and the hatched (upper portion of the bar) shows how many correct non-icing forecasts would become false positives. This means that a large number of false negative forecasts could be avoided and relatively few false positives would result. Further research and additional data are needed to determine whether this updated criteria for complex terrain environments should be established.

Another 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 adiabatically to the height of the actual summit. Our software permits us to obtain a value called the parcel LCL from the ETA model data, which is the pressure level of the LCL location when using averaged temperature and dewpoint over the first 100mb above the surface. This parcel LCL, which we only have for 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.

5. CONCLUSION

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 72% for the study period as a whole. However, there were many differences noted in each of the forecasts examined.

The NNICE forecast produced high accuracy percentages, although it overpredicts icing occurrences. On the other hand, VVICE tended to underpredict icing events, with accuracies on average around the 50% mark. The AFWA MM5-based forecasts and AWSP algorithm both produced forecasts that had a very high ratio of false negatives to false positives. The AFWA algorithm did not produce a significantly improved forecast when grid spacing was reduced from 45km to 15km. AFWA accuracy percentages were in general higher than AWC.

The AWSP algorithm was generally more accurate than either the AWC or AFWA methods. Using ETA soundings interpolated for KMWN as input data produced forecasts of higher accuracy than using nearby GYX radiosonde data. There is potential for manipulation of this forecasting algorithm for improvement in regions of complex terrain. Initial research shows that adaptation of the algorithm by adjusting the input data is a way to improve forecasts.

Additional work is underway to evaluate the icing severity predictions of the various forecast methods. These results will be reported in a future paper and/or report.

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 room for improvement.

6. 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.

7. 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.

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 metrievals. *Wea. Forecasting*, **11**, 66-77.

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