High Resolution Numerical Weather Model Forecasts of Icing at the Ground and in the Air

Gregory Thompson

National Center for Atmospheric Research, Boulder, Colorado, USA

gthompsn@ucar.edu

Abstract— **Recent advances in high-performance computing have enabled higher-resolution numerical weather models with increasingly complex data assimilation and more accurate physical parameterizations. With respect to aircraft and ground icing applications, a weather model's cloud physics scheme is responsible for the direct forecasts of the water phase and amount and is a critical ingredient to forecasting future icing conditions. Numerical model results using the Weather Research and Forecasting (WRF) model are compared with aircraft observations taken during icing research flights and operational icing pilot reports (PIREPs), and the general characteristics of liquid water content, median volume diameter, droplet concentration, and temperature within aircraft icing environments were evaluated. The comparison reveals very promising skill by the model in predicting these characteristics consistent with observations. In addition to the explicit icing application, a ten-year analysis of surface weather conditions reveals good skill for many types of weather situations but also reveals areas that can still be improved.**

Keywords— ***Microphysics, numerical weather prediction, cloud physics.***

I. Introduction

As compared to 25 or more years ago, numerical weather prediction models have dramatically improved their skill. Global weather model skill is typically measured by the 500 hPa geopotential height anomaly, which has rather consistently shown about one forecast day of improvement per decade of model development/improvement. In other words, the 5-day forecast of today has about the same skill that a 4-day forecast had a decade ago. Perhaps less attention has been placed on surface sensible weather or explicit cloud-phase predictions, especially for something as specialized as atmospheric icing.

Global weather models have relatively low resolution and unsophisticated cloud physics schemes that do not produce very detailed forecasts of amount of supercooled liquid water content (LWC), droplet concentration, or mean size of water droplets. In contrast, a regional mesoscale numerical weather prediction (NWP) model with high resolution can accurately resolve small terrain features and use sophisticated cloud physics that can be used to derive moderately to highly accurate icing forecasts. The Weather Research and Forecasting (WRF) model [1] is a good example of a NWP model that has proven skill in predicting supercooled liquid water (SLW) that is responsible for aircraft and ground icing conditions [2].

This paper will show an analysis of the WRF-model explicit prediction of SLW against pilot reports (PIREPs) of icing as well as surface observations of specific weather conditions, including ground icing. Where appropriate, specific mention of prior work will highlight the weather model improvements gained over the last decade or two. The next section will summarize the observational data used in the analysis whereas Section 3 contains a brief description of the numerical weather model experiments. Section 4 contains the statistical verification, while Section 5 contains preliminary results of ultra-high resolution prototype forecasts during a recent field campaign as well as future research ideas.

II. Observational Data

A. FAA Icing Database

The Federal Aviation Administration (FAA) has supported aircraft icing research studies for roughly 50 years. As part of this effort, aircraft measurement campaigns spanning multiple decades have collected temperature, liquid water content, droplet concentration, and median volume diameter (MVD) data when research aircraft encounter icing conditions. Icing scientist Richard Jeck at the FAA William J. Hughes Technical Center compiled most of these data into a single icing database [3] of approximately 9000 observations that were used in the analysis herein.

B. Pilot Reports (PIREPs)

The use of research aircraft to study icing conditions aloft occurs rarely, however, routine private and commercial aircraft often make "voice pilot reports" (PIREPs) over the United States. While these data do not provide the highly valuable measurements of LWC, MVD, temperature, or other data, they are a moderately good indication of the presence or absence of SLW.

PIREPs of icing consist of somewhat subjective intensity categories of "trace," "light," "moderate," or "severe/heavy" icing along with frequent mention of "rime," "clear," or "mixed" icing type. Often times, a PIREP will also explicitly state that "no icing" was found, which are nearly as important as the affirmative reports. The altitude(s) of the icing encounter (or lack thereof) and occasional temperature or other conditions are sometimes reported. Therefore, a typical statistical contingency table can be constructed from the binary, yes/no, PIREPs of SLW to compare against a numerical model prediction of any SLW. Attempting to correlate the icing intensity or icing type with LWC or other metrics would be ill-advised.

For the analysis presented below, there were approximately 280.000 icing PIREPs reported in total for the 6-month periods of 15 Oct to 15 Apr in the years from 2001 to 2011. Warm-season PIREPs of icing were ignored since icing in convection is common but predicting the exact location of thunderstorms in a numerical model is unrealistic. Nearly half of these PIREPs indicated the lack of icing conditions. In performing the analysis, the distinction between rime, clear, and mixed icing was ignored and recorded PIREPs that specifically mentioned a range of icing intensity such as light-to-moderate were simplified to the more severe category. Table 1 below shows the number of each category of icing PIREPs. The category called "Neg-1" are those reports that pilots declared "no icing," whereas the "Neg-2" category represents the assumption of no icing due to other recorded data such as unlimited visibility or a report of "clear above" the flight altitude. Note that severe icing is rarely reported.

TABLE I. NUMBER OF PIREPS USED IN THIS STUDY

Category	Number of observations	Percent of all affirmative icing
Neg-1	43.241	-
Neg-2	72.060	-
Trace	14.456	8.8%
Light	96.091	58.6%
Moderate	51.416	31.4%
Severe	1.910	1.2%

C. Surface Weather Observations (METARs)

Many airports report surface weather observations of temperature, pressure, humidity, wind, visibility, clouds, and any form of precipitation. In the United States, most airport weather stations are automated, some of which have weather observers to augment the automated data. The fully automated sites have great difficulty detecting freezing drizzle, sometimes being mis-reported as light snow, but freezing rain reports have much greater reliability [4].

Surface observations of freezing drizzle or freezing rain are obvious indicators of a ground icing event. Freezing fog could be an indicator of presence of SLW, although in the U. S., the visibility alone is used to report either fog or freezing fog (depending on above/below 0 ºC) when visibility is less than 5/8 statute miles (approx. 1 km). In other words, a direct observation of presence of small cloud droplets is not assured when automated stations report fog or freezing fog.

Similar to the PIREPs data, the surface weather reports (METARs) between Oct. 15 and Apr. 15 for the same 10 years were used in this analysis. Individual or various forms of precipitation type were grouped together for verification purposes. These are: rain or drizzle (RA/DZ), freezing rain or freezing drizzle (FZRA/FZDZ), snow grains, graupel, or hail (SG/GR/GS), ice pellets (PL), snow (SN), fog (FG), and freezing fog (FZFG). For reasons mentioned above with the reliability of observed FG and FZFG as direct indicators of liquid droplets present, caution should be used when interpreting the results.

In total, roughly 3,4 million METAR observations were used in the model verification for the 60 cool-season months. The raw number of each type of weather condition is shown in Table 2.

III. METHODOLOGY

The categories of RA/DZ or FZRA/FZDZ were directly validated by presence of any model-predicted rain (amount > 0) at the lowest model level with temperature (T) either above or below 0 ºC. More specifically though, an observation of FZRA with a model prediction of rain with T > 0 ºC is considered an unsuccessful forecast. The direct validation of snow occurs regardless of temperature, so melting snow both in model and observations are verified. Ice pellets, snow grains, graupel, and hail were included in this study as they are an indicator of presence of SLW aloft. In validating this latter group of weather types, if the WRF model contained any graupel reaching the surface, then it was considered a successful forecast.

TABLE III. NUMBER OF METARS USED IN THIS STUDY

Category	Number of observations	Percent (not including FG/FZFG)
RA/DZ	1.561.299	46.0%
SN	1.747.399	51.5%
FZRA/FZDZ	62.493	1.8%
PL	12.186	0.4%
SG/GR/GS	7.279	0.2%
FG	314.505	-
FZFG	148.222	-

For each and every METAR or PIREP from 30 minutes prior to 30 minutes after the top-of-the-hour, the model data were interrogated for the nearest 6 x 6 WRF-model grid points (approximately 24 x 24 km2 area) to evaluate as either a "hit" or a "miss." If any of the WRF 36 points had the corresponding condition, it was scored a hit; otherwise a miss. To validate fog or freezing fog, the model cloud liquid water needed to be non-zero, whereas the model rain, snow, or graupel were used to validate the other categories. Validating the WRF model versus METARs required only the model lowest level data while validating the PIREPs was done using any model vertical level found within 1000 feet (305 m) of the reported altitude. Most often this resulted in 2 or 3 model vertical levels being inspected; although it varied with height of the PIREP due to the vertical spacing of model levels increasing with height. In all, the time window and spatial constraints of this verification exercise were considerably shorter/smaller than prior similar work [5]. As in Ref. [5], the probability of detection (POD), which is the ratio of hits to the number of observed events was calculated. Similarly, the false alarm ratio (FAR) was computed as one minus the number of correctly forecast negative event divided by the number of negative observations.

Full details of the WRF model set-up, physical parameterizations, and forcing data are found in [6]. The WRF simulation covered the entire contiguous U.S. with 4-km grid spacing and was started on 1 Oct 2000 and simulated 13 years. Thus, the simulation could be considered a hindcast or re-analysis as it did not perform a large series of short-term forecasts that re-started at a regular interval. Most pertinent to the point of icing/SLW forecasts is the use of the "aerosol-aware" microphysics scheme described in [7] since few of the available WRF microphysics choices predict aerosols and their impacts on cloud physics directly.

IV. RESULTS

A. WRF Versus FAA Icing Database

A comprehensive validation of the WRF model results versus the FAA Icing Database for a multitude of parameters

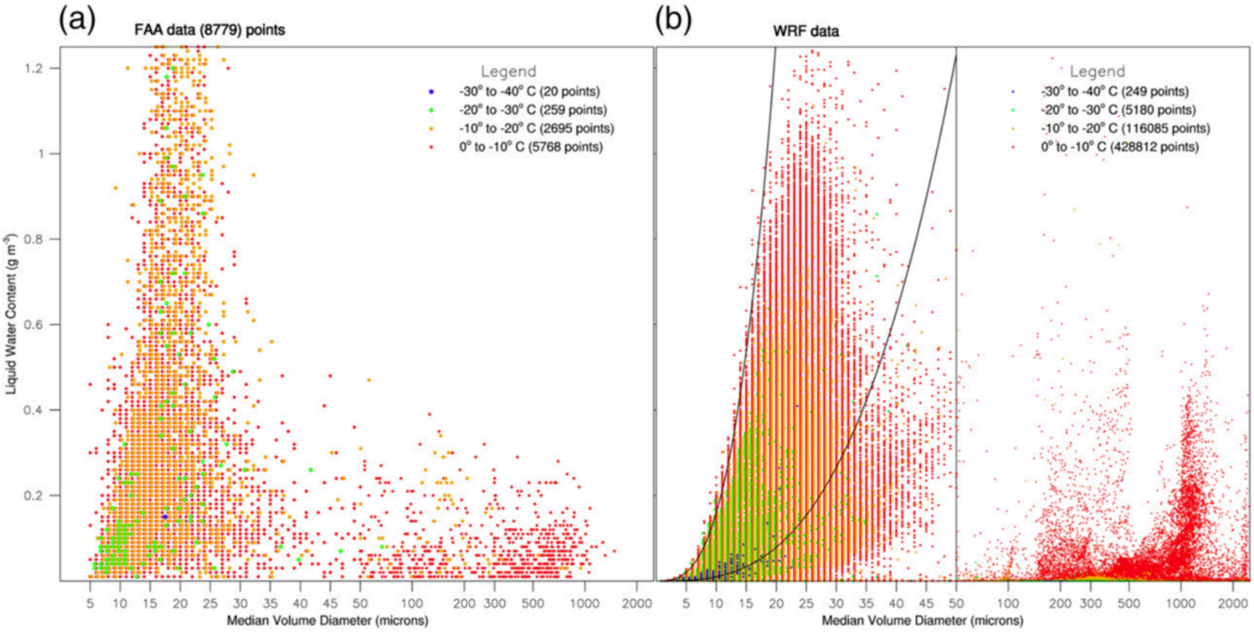

Fig. 1 Scatterplots of (a) FAA observations and (b) WRF Model results of MVD vs LWC color-coded by temperature. The left portion of each panel shows the typical cloud droplet size and is displayed with a linear scale of MVD, whereas the right portion of each panel shows the SLD and uses a logarithmic scale to capture the full range of values in the dataset.

was given in [2]. A brief overview of the most important parameters of LWC versus MVD as a function of temperature is summarized in their Fig. 1 repeated below. The same collection of PIREPs shown in Table 1 was used to retrieve the WRF model values of LWC and other data shown in the figure.

The FAA icing database was collected over decades of infrequent icing research campaigns going as far back as the 1970s. Therefore, the ~9000 points of FAA data shown in Fig. 1a is effectively a climatology of the variables shown during icing events in general. Since it would be rather difficult to run the WRF model for a plethora of dates and times to match exactly the icing flights captured in the FAA icing database, the WRF model simulation in this analysis was a continuous run starting 1 Oct. 2000 for 13 years using 4-km grid spacing. The analysis here extracted 60 cool-season months from this simulation. To gather icing environment statistics from WRF, the time and location of icing PIREPs were used to retrieve the same parameters from the 36 (6 x 6) points of WRF at matching PIREP locations. The environmental characteristics at points in which the model contained any SLW were used to create a nearly random sampling of model data from the millions of total WRF points containing SLW. Therefore, the evaluation should be considered as an observed climatology of LWC, MVD, temperature and other data along with the same from the WRF model. The one-to-one correspondence with PIREPs and METARs shown in the next two subsections, on the other hand, are a direct time and space comparison from the same WRF simulation.

Note the general trend of lower temperatures being correlated with lower LWC and smaller MVD – both are consistent in the observations and the model. This trend agrees with laboratory observations of the freezing of water drops [8] based on their size and the temperature with the smallest drops being less likely to freeze until very low temperatures. Also note the near lack of any large drops in the freezing rain size at T < -10 ºC. The cluster of points in the WRF data (Fig. 1b) centered near MVD = 1 mm are not easily seen in the observations with high LWC, but this is also a region of potentially very severe icing, so a lack of observations could be attributable to the danger of collecting aircraft icing in these conditions and not a true lack of occurrence.

Two additional areas of Fig. 1 are noteworthy. The general gap of observed data points in the MVD range of approximately 35 to 100 μm, which is most likely due to collision-coalescence as relatively large cloud drops attain appreciable fallspeeds and begin collecting the much smaller cloud droplets in the "warm-rain process." In other words, drops in this size rather are likely to grow rapidly into larger rain drops and transit this section of size range very rapidly. The WRF model results appear quite similar. The second noteworthy area is related as there appears to be too many points in the WRF results with MVD approximately 30 – 50 μm and LWC > 0.2 g m^{-3}. It is suspected this is due to the cloud droplet number concentration possibly having fewer droplets than observations at times or else their MVD would be somewhat smaller (for the same LWC value).

B. *WRF Versus Icing PIREPs*

A comparison of icing PIREPs to WRF model results is found in Fig. 2, broken down for each month. Recall that the start and end months are only ½ months in contrast to the others. The box diagram for the Severe category is so large because the sample size is very small (see Table 1). Overall, it is obvious from the figure that more intense reported icing is more likely to be captured by the WRF forecast. This is especially good since brief or trace/light icing encounters are unlikely to produce a serious concern to pilots – otherwise the

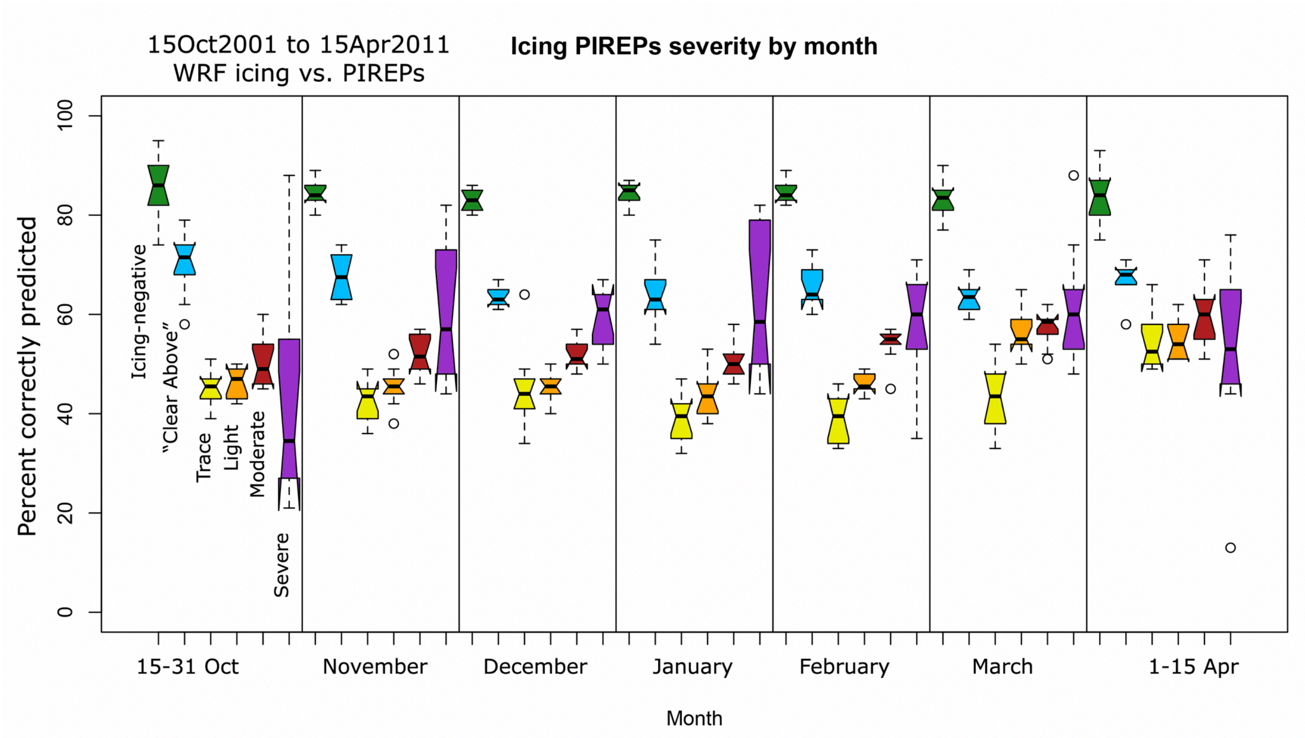

Fig. 2 Monthly box/whisker plots of correctly-predicted icing pilot report (PIREP) by. category from WRF model results. Analysis contains a total of approximately 280.000 individual icing pilot reports (PIREPs).

encounter would be defined as moderate or severe. Also, with the POD of negative icing (green boxes in Fig. 2) near 85%, the false alarm ratio (FAR) is near 15%. The cyan boxes labelled "Clear Above" is when a PIREP included ancillary information such as a cloud deck with tops at a specified altitude and clear skies above. For the statistic shown here, the WRF model must also show zero cloud hydrometeors above the same altitude. So clearly the WRF model is not as frequently entirely showing the same zero cloud condition at higher altitudes as it is showing the lack of SLW/icing. Regardless, the skill is rather impressive given the spatial/temporal constraints.

C. WRF Versus Surface Weather Observation

Similar to the PIREP comparison, Fig. 3 shows the same statistics for the various categories of surface weather condition as reported in METARs. As in Fig. 2, note that the Spring/Fall months reveal lower skill than the other months, most likely due to the rapidly changing weather conditions in these months making them a greater challenge to get correct.

Considering the largest majority of all reports are either rain or snow, and a good fraction of the time it is likely obvious due to temperature, the skill of WRF to predict the correct condition is very high (nearly 90% correct). What is less obvious here is how often the model predicts rain or snow when nothing occurs in reality. This analysis for a much shorter but representative time period is found in [9] that showed a similar WRF-model analysis indicated an over-forecast of rain; although most other categories had much lower FAR.

FZRA or FZDZ are predicted right by the WRF model roughly 65% of the time through the cold season, with lower skill in the "shoulder" months of Oct/Apr. Ice pellets are also predicted with similar but lower skill to FZRA/FZDZ even though it is rarely reported. Fog, either from actual cloud droplets or simply very low visibility, is predicted by the model with a fair amount of success (~60% of the time). Freezing fog, on the other hand, appears to have far lower skill although getting a relatively low visibility in cool temperatures of winter is likely to be more frequent than the fog conditions. If U.S. reporting of fog actually required the presence of droplets rather than the visibility, it would be interesting to see how well the model can predict this condition. Regardless, the exactly correct prediction of ground fog is extremely difficult due to myriad feedbacks in model physical parameterizations including the highly important land-surface model, radiation, microphysics, and turbulence.

V. Conclusions

The WRF model is highly capable of explicitly predicting icing conditions aloft and at the ground. The more intense the reported icing, the more likely the model captured the presence of SLW. The model results indicate less success with Trace and Light icing encountered by aircraft, but perhaps these are less concerning than more serious intensities. Not surprisingly, the Spring and Fall season are more highly variable than the cold months of Nov – Mar in which case the model performs better overall.

Surface precipitation falling as either rain or snow are predicted with extremely good skill, although areas of light rain are often seen in the model more frequently than observed. Ground icing events of freezing drizzle and freezing rain are predicted correctly roughly 65% of the time, however it is

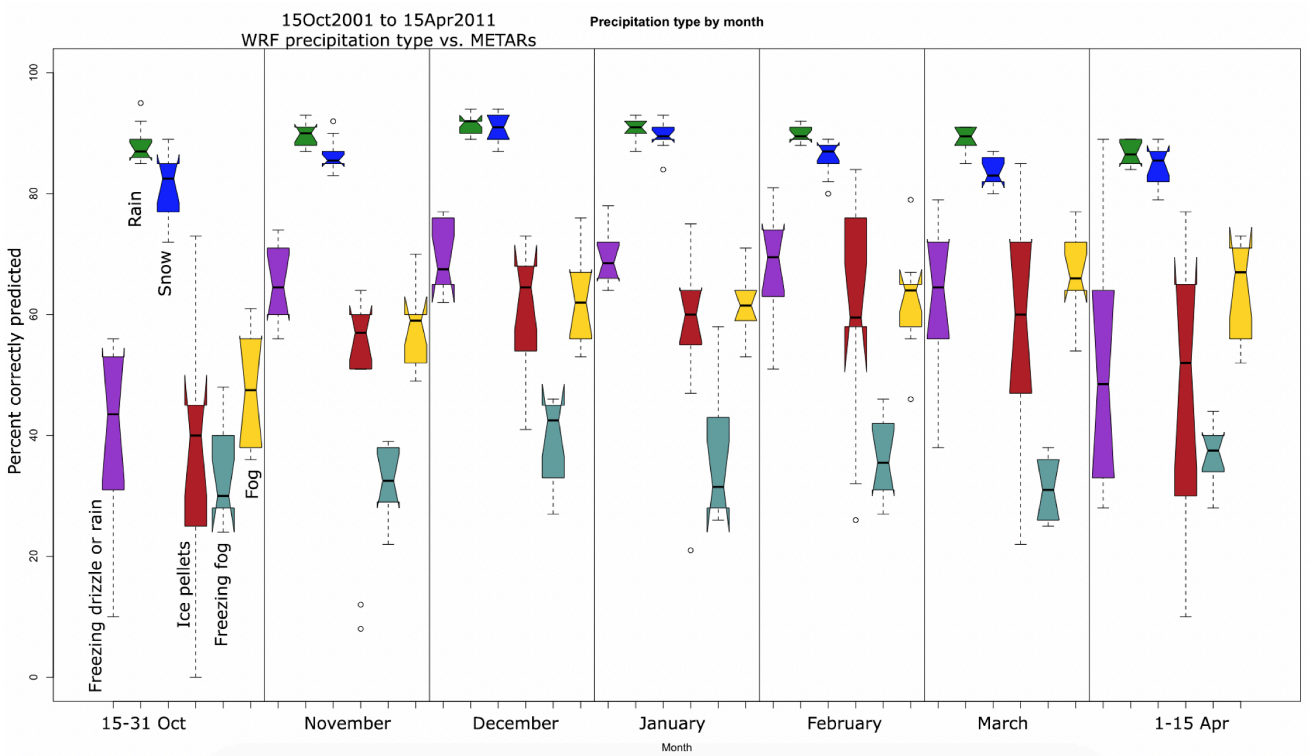

Fig. 3 Monthly box/whisker plots of correctly-predicted surface weather category from WRF model results. Analysis contains a total of approximately 4.3M individual reports (METARs).

noteworthy to consider the temporal and spatial constraints of this study as well as the fact that freezing drizzle can often be mis-reported in U.S. METARs as light snow. Even the infrequent occurrence of ice pellets is correctly predicted about 60% of the time and yet only observer-based augmented sites can report PL since this is not part of the fully automated reports. With higher quality-controlled human-based observations of precipitation type, it would be very interesting to see how these statistics may change.

VI. Future Work

This past winter, the Federal Aviation Administration (FAA) conducted a field program called ICICLE (In-Cloud Icing and Large-Drop Experiment) based in Rockford, Illinois to study events of freezing drizzle, freezing rain, and other icing conditions at the surface and aloft. A unique data set was collected including enhanced surface measurements as well as particle size distributions of water and ice in the clouds and precipitation above the ground by a research aircraft. Preliminary results are currently being compiled as well as comparisons to a unique WRF model simulation using 600-meter grid spacing that was run in support of the project.

Acknowledgment

The Federal Aviation Administration (FAA) is gratefully acknowledged for their long-term support of icing research. James Riley of the FAA William J. Hughes Technical Center provided the icing database of Dr. Jeck. 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 position of the FAA.

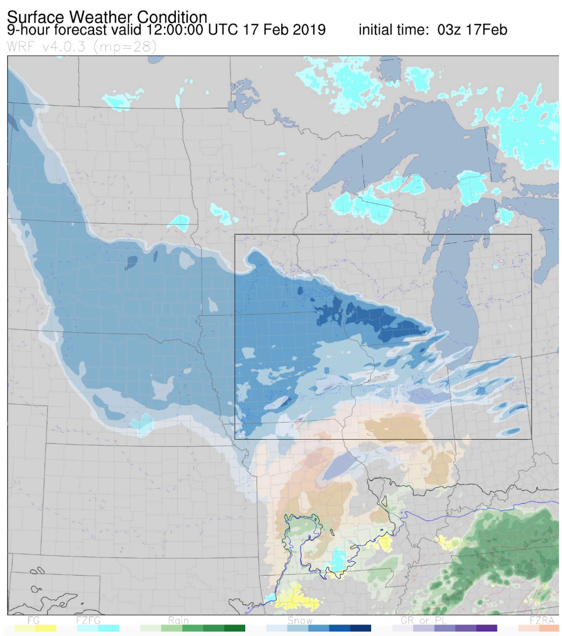

Fig. 4: Example WRF 9-h forecast of surface weather conditions during ICICLE experiment valid at 1200 UTC 17 Feb 2019.

References

[1] Skamarock. W. C., and J. B. Klemp, 2008: A time-split non-hydrostatic atmospheric model for weather research and forecasting. *J. Comput. Phys*., **227,** 3465–3485.

[2] Thompson, G., M. K. Politovich, and R. M. Rasmussen, 2017: A Numerical Weather Model's Ability to Predict Characteristics of Aircraft Icing Environments. *Wea. Forecasting,* **32,** 207–221.

[3] Jeck, R. K., 2008: Advances in the characterization of supercooled clouds for aircraft icing applications. FAA Tech. Rep. DOT/FAA/AR-07/4, 54 pp. 1 appendixes. [Available online http://www.tc.faa.gov/its/worldpac/techrpt/ar074.pdf].

[4] Landolt, S., A. Gaydos, D. Porter, S. DiVito, D. Jacobson, A. J. Schwartz, G. Thompson, J. Lave, 2019: A New Method for Diagnosing Freezing Drizzle, Ice Pellets, and Frost using the Automated Surface Observing System (ASOS). *J. Appl. Meteo. Clim.,* submitted.

[5] Brown, B., G. Thompson, R. Bruintjes, R. Bullock, and T. Kane, 1997: Intercomparison of in-flight icing algorithms. Part II: Statistical verification results. *Wea. Forecasting*, **12,** 890–914.

[6] Rasmussen R., C. Liu, K. Ikeda, D. Gochis, D. Yates, F. Chen, M. Tewari, M. Barlage, J. Dudhia, W. Yu, K. Miller, K. Arsenault, V. Grubisic, G. Thompson, E. Gutmann, 2011: High-resolution coupled climate runoff simulations of seasonal snowfall over Colorado: A process study of current and warmer climate. *J. Clim.,* **24,** 3015–3048

[7] Thompson, G., and T. Eidhammer, 2014: A study of aerosol impacts on clouds and precipitation development in a large winter cyclone. *J. Atmos. Sci.,* **71,** 3636–3658.

[8] Bigg, E. K., 1953: The supercooling of water. *Proc. Phys. Soc.* London, 66B, 688–694.

[9] Xu, M., G. Thompson. D. R. Adriaansen, and S. Landolt, 2019: On the Value of Time-Lag-Ensemble Averaging to Improve Numerical Model Predictions of Aircraft Icing Conditions. Wea. Forecasting, 34, 507–519.