

## An Intercomparison of Mesoscale Forecasts of Aircraft Icing Using SSM/I Retrievals

ANDRÉ TREMBLAY, STEWART G. COBER, ANNA GLAZER, AND GEORGE ISAAC

*Cloud Physics Research Division, Atmospheric Environment Service, Dorval, Quebec, Canada*

JOCELYN MAILHOT

*Recherche en Prévision Numérique, Atmospheric Environment Service, Dorval, Quebec, Canada*

(Manuscript received 13 April 1995, in final form 14 November 1995)

### ABSTRACT

A technique for the detection of supercooled liquid water (SLW) from Special Sensor Microwave/Imager (SSM/I) data is discussed. For this study, these SLW retrievals depict areas of icing that are used to compare against different aircraft icing forecast algorithms. It is shown that currently used automated algorithms give incorrect distributions of SLW events with temperature and include systematically glaciated clouds in icing forecasts. This problem is eliminated when an explicit SLW scheme is used. With this scheme the Canadian forecast system can detect roughly 75% of the SLW signal within an accuracy of 100 km when compared to SSM/I retrievals. The scheme also detects 99% of nonicing events within the same accuracy.

### 1. Introduction

Aircraft icing remains a significant hazard to aviation, particularly for small planes and helicopters where anti-icing and deicing equipment may not be installed. Consequently, accurate forecasting of regions of potential icing is an important research objective and currently poses a challenging problem to the meteorologist. Correct prognosis of the occurrence, location, altitude, amount, and type of clouds and precipitation must be successfully completed to issue useful aircraft icing forecasts. Such forecasts must clearly discriminate whether the clouds are composed of innocuous ice crystals or the more dangerous supercooled liquid water droplets.

The use of numerical weather model output to generate forecasts of significant weather for aviation has grown through the years. Methods to predict aircraft icing are typically diagnostics that use criteria for the dewpoint depression, temperature, and vertical motion to delineate icing regions as derived from pilots reports (PIREPs) of icing. These procedures are based mostly on the criteria suggested by Appleman (1954), or by Schultz and Politovich (1992). Tremblay et al. (1995) proposed an alternative technique to forecast the occurrence of supercooled clouds based on a prognostic cloud scheme that can provide the cloud liquid water content. Recent studies have emphasized the impor-

tance of this parameter in the development of reliable aircraft icing forecast techniques that embody the crucial cloud physics processes (Curry and Liu 1992; Lee et al. 1994; Modica et al. 1994).

In this paper, the advantages, disadvantages, and accuracy of the Appleman (1954), Schultz and Politovich (1992), and Tremblay et al. (1995) aircraft icing (or equivalently SLW) forecast schemes are discussed. The Canadian regional finite element (RFE) model with an explicit cloud scheme (Pudykiewicz et al. 1992) was used to provide a comparison framework for these various forecast schemes. Comparisons of the predicted distributions of SLW were made with coincident Special Sensor Microwave/Imager (SSM/I) retrievals.

### 2. Aircraft icing forecasting algorithms

#### *a. Simple schemes*

Prior to the development of large-scale numerical models, aircraft icing predictions were based on statistical or empirical techniques. For example, the Air Weather Service (AWS) manual (1980) describes a number of forecast aids and parameters computed from synoptic or aircraft measurements, which forecasters could use in predicting areas of potential icing. Similarly, radar and radiosonde data coupled with curves of icing frequency have been used to generate icing threat forecasts. More recently, direct output of meteorological models has been used to forecast aircraft icing. [See Schultz and Politovich (1992) and references therein.] Typically, a diagnostic forecast of icing is pro-

---

*Corresponding author address:* Dr. André Tremblay, Cloud Physics Research Division, Atmospheric Environment Service, 2121 Trans Canada Highway, Dorval, PQ, Canada H9P 1J3.

duced using criteria for the dewpoint depression, temperature, and upward vertical motion. These parameters are tested at each grid point to verify if certain critical conditions are met.

The U.S. Air Force Global Weather Central, in their automated icing forecast program (Mansur 1984), uses criteria based on results from an aircraft icing survey as reported in AWS (1980). The criteria for the presence of icing at any altitude are temperatures between 0° and -16°C, upward vertical motion greater than -0.2 Pa s<sup>-1</sup>, and dewpoint depression less than a certain threshold.

Such an indicator for aircraft icing was first suggested by Appleman (1954). From frost-point theory and aircraft icing reports, Appleman concluded that when temperature was plotted against dewpoint as in Fig. 1, observations of fog or cloud lying between the lines  $T = T_d$  (saturation line) and  $T = 0.8T_d$  (Appleman line) represented cases of SLW, and icing was highly probable. Appleman verified this conclusion with only a relatively small sample (49 cases) of weather reconnaissance observations, and his conclusion was validated with a considerably larger sample of data (5483 observations in subfreezing stratiform clouds) by AWS (1980). Since these comparisons did not consider non-icing events, the results can suggest only that the Appleman criterion (hereafter referred as APP) may be a *necessary* condition to observe icing events. Unfortunately, it was not possible to demonstrate from these observations that this criterion is a *sufficient* condition for the occurrence of icing.

Using a database of icing reports over the continental United States and numerical model output, Schultz and Politovich (1992) identified temperature and relative humidity criteria for two icing threat classes. Class I

used temperatures extended from 0° to -20°C and relative humidity greater than 50%, which encompassed 87% of icing reports in the database. Class II (hereafter referred as the Schultz and Politovich standard algorithm or SPS) was more restrictive, with a temperature range of -2° to -15°C and relative humidity greater than 65%, which included only 71% of the icing reports.

Figure 1 illustrates the difference between the APP and SPS schemes for the specification of icing threat as a function of temperature and dewpoint. For clarity, the two schemes are shown over the same temperature interval. The area between the saturation line and the SPS line indicates icing. The SPS scheme distributes icing occurrence evenly with respect to temperature, because it is based on a uniform relative humidity threshold. The APP criterion is based on frost-point theory and tends to increase icing probability with decreasing temperature as indicated by the increasing distance of the APP line from the saturation line at lower temperatures.

Since any icing event is associated with the presence of SLW, the APP and SPS schemes can be considered indicators of SLW. Thus, when applied directly to numerical weather prediction model output, the SPS scheme should yield to a flat distribution of SLW events with respect to temperature and the APP procedure should tend to generate an increasing number of SLW forecasts with decreasing temperature. This should happen since relative humidity and temperature are independent meteorological variables. However, the occurrence of SLW, and consequently icing events, tends to decrease with decreasing temperature (Pruppacher and Klett 1978). This is partly due to the increase of the concentration of active ice nuclei capable

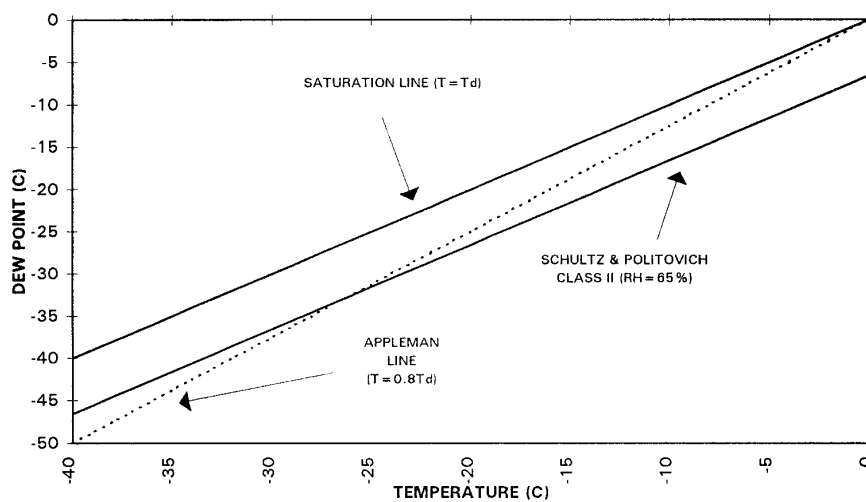


FIG. 1. Dewpoint vs temperature diagram showing icing occurrence according to Schultz and Politovich (1992) class II and Appleman (1954) schemes. Icing is predicted for regions between the saturation line and the Appleman line or for regions between the saturation line and the RH = 65% line for the Schultz and Politovich scheme.

of initiating the ice phase with decreasing temperature (Fletcher 1962). The depletion of SLW is also due to the spontaneous freezing of small water droplets at low temperature and to the efficiency of the Bergeron–Findeisen process (Cotton and Anthes 1989). Aircraft measurements (Moss and Johnson 1994; Cober et al. 1995) have explicitly confirmed the decrease of proportion of water in clouds with decreasing temperature. Thus, it is conceivable that the SPS and APP schemes may fail to reproduce correctly the observed distribution of SLW with temperature.

### b. A cloud microphysics scheme

To delineate the conditions for the existence of SLW, Tremblay et al. (1995) performed several simulations with the cloud microphysics model of Zawadzki et al. (1993). This model includes conservation equations for temperature, water vapor, cloud water, cloud ice, rain, snow, and graupel. It also has the capability of simulating explicitly the distribution of SLW.

Cloud-scale simulations were accomplished with this model for a number of cases during the second Canadian Atlantic Storms Program (CASP II) held from 15 January to 15 March 1992 (Stewart 1991), during which aircraft measurements of cloud microphysical properties were available. Detailed analyses of these microphysical simulations and comparisons with observations have shown that snow and SLW account for most of the hydrometeor mass at temperatures below 0°C. This result suggests that a simple model involving only these two types of hydrometeors accurately characterizes the subfreezing moisture distribution of winter storms.

By evaluating the relative importance of the various microphysical processes involved in the production of SLW, Tremblay et al. (1995) showed that the following condition was sufficient for the existence of SLW:

$$wG - \text{SDEP} > 0, \quad (1)$$

where  $w$  is the vertical velocity,  $G$  is the vertical gradient of the saturation mixing ratio, and SDEP is the vapor deposition on snow. Physically,  $wG$  represents the generation of water vapor excess over saturation from wet adiabatic cooling. Equation (1) indicates that if the amount of vapor condensed by wet adiabatic cooling exceeds the amount of vapor deposition occurring on snow, then SLW will be produced. The variable  $G$  is a function of temperature and pressure and is given by

$$G(T, p) = \rho r_s \left[ \frac{L_v}{R_v T^2} \Gamma_w(T, p) - \frac{g}{R_d T} \right], \quad (2)$$

where  $r_s$  is the saturation mixing ratio,  $\Gamma_w$  the wet adiabatic lapse rate,  $L_v$  the latent heat of vaporization,  $T$  the ambient temperature,  $g$  the gravitational acceleration,  $R_d$  and  $R_v$  the gas constants for dry air and water

vapor, respectively, and  $\rho$  the air density. Vertical velocity, temperature, and pressure are available on the RFE mesoscale grid and can be used to calculate  $wG$  at each point.

The microphysical process SDEP is a function of temperature, pressure, and snow content ( $q_s$ ) for supercooled clouds saturated with respect to water. A simple formula is obtained for SDEP by integrating the expression for crystal growth rate (Rogers 1979) over an assumed snow size distribution:

$$\text{SDEP}(q_s, T, p) \approx \frac{2\pi[S_i(T) - 1]}{\frac{L_s^2}{KR_v T^2} + \frac{R_v T}{e_{si} D}} f \times M_s^{(1)}, \quad (3)$$

where  $S_i = e_s/e_{si}$  is the ratio of the saturation vapor pressure over water and ice,  $K$  is the coefficient of thermal conductivity of air,  $D$  is the coefficient of diffusion of water vapor in air,  $L_s$  is the latent heat of sublimation, and  $f$  is a dimensionless adjustment factor, to be discussed below. The first moment of the snow particle size distribution  $M_s^{(1)}$  is related empirically to  $q_s$  (Zawadzki et al. 1993) by

$$M_s^{(1)} = 55.6 \left( \frac{6q_s}{\pi\rho_s} \right)^{0.58}, \quad (4)$$

where  $\rho_s$  is the density of snow. Therefore, the qualitative criterion (1) can be expressed in terms of vertical velocity ( $w$ ), air temperature ( $T$ ), pressure ( $p$ ), and snow content ( $q_s$ ).

With temperature as an independent variable at constant pressure, Fig. 2 shows an ensemble of functions  $wG$  indexed by vertical velocity and an ensemble of functions SDEP indexed by snow content. By fixing temperature (−15°C), pressure (700 mb), and vertical velocity (0.10 m s<sup>−1</sup>), giving a specific value of  $wG$ , the point X in this figure is obtained. The proposed criterion states that for all SDEP points on isotherm

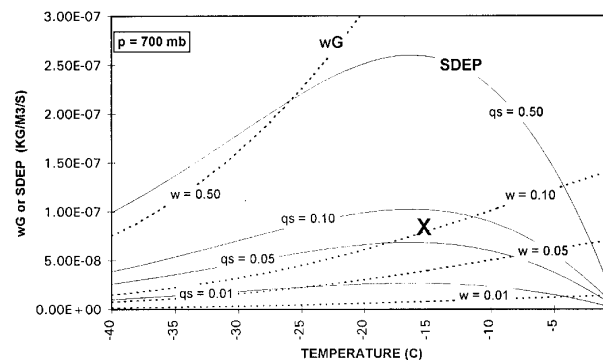


FIG. 2. A diagram illustrating the determination of supercooled liquid water at a given point as a function of temperature (°C), vertical velocity  $w$  (m s<sup>−1</sup>), and snow content  $q_s$  (g m<sup>−3</sup>) for the 700-mb pressure level. The  $wG$  and SDEP curves are shown for a few selected values of  $w$  and  $q_s$ ; however, an ensemble of curves is implied.

$-15^{\circ}\text{C}$  located below X there is a possibility of existence of SLW. These points correspond to a snow content lower than  $0.06\text{ g m}^{-3}$ . For all the points above X, corresponding to a snow content greater than  $0.06\text{ g m}^{-3}$ , an ice cloud exists. Alternatively, the point X may also be interpreted as characterizing fixed conditions of temperature ( $-15^{\circ}\text{C}$ ) and snow content ( $0.06\text{ g m}^{-3}$ ), giving a specific value of SDEP. In this manner, the criterion predicts a possibility of existence of SLW for vertical velocities exceeding  $0.10\text{ m s}^{-1}$  (above point X) and ice cloud for vertical velocities less than  $0.10\text{ m s}^{-1}$  (below point X).

The above technique can easily be combined with mesoscale prognostic cloud schemes such as the one proposed by Sundqvist et al. (1989) to generate aircraft icing forecasts. In this scheme, all hydrometeors are combined within a single variable  $m$ , which is calculated explicitly from a continuity equation. Because  $m$  combines all frozen hydrometeors while SDEP depends solely on snow content ( $q_s$ ), SDEP will likely be overestimated. Consequently, the regions with potential SLW are minimized and this criterion [i.e., (1)] will flag only the zones with the greatest potential probability for the occurrence of supercooled clouds. Thus, there is a possibility for missing weakly forced icing events, and there is a need for an optimization of the technique from observations. Moreover, since (1) is used as the criterion for SLW, the factor  $f$  must be empirically adjusted to account for large uncertainties in the determination of snow size distribution parameters and to account for the mesoscale nature of the vertical velocity field. The factor  $f$  should be dependent on model resolution (since the scheme uses the resolved-scale vertical velocity) and plays a role similar to the relative humidity, temperature, and vertical motion thresholds in the APP and SPS schemes. In the present work, an optimal adjustment for  $f$  is obtained from comparisons of SLW forecasts and satellite retrievals, and acceptable results are obtained with  $f = 0.1$  in (3). This choice of  $f$  was based on visual inspection of a typical forecast and the corresponding satellite retrieval of SLW prior the calculation of verification statistics discussed below. No attempts were made to minimize the error to obtain an optimal value for  $f$ . Although there are well-known mathematical tools for solving such problems, it is felt that the present dataset is too limited to explore this alternative. Larger values of  $f$  tend to produce localized regions of SLW that do not agree very well with the horizontal scales obtained from the SLW retrievals discussed below.

### 3. Mesoscale forecasts and SSM/I retrievals of SLW

#### a. Mesoscale icing forecasts

To provide a framework for validation, comparison, and optimization of the various icing prediction schemes, several numerical mesoscale simulations

were performed with the RFE model (Benoit et al. 1989; Mailhot et al. 1995) for selected North Atlantic winter storms observed during CASP II. An important feature of the model is the use of an explicit treatment of condensation and clouds, based on Sundqvist et al. (1989). This formulation explicitly treats the cloud content as a prognostic variable, which allowed the application of the cloud microphysics icing forecast scheme of Tremblay et al. (1995, hereafter referred as the T95 scheme).

For each simulation, the hemispheric grid of the RFE model was adjusted to match the footprint of the U.S. Defense Meteorological Satellite Program (DMSP) polar-orbiting satellites over the region of interest. All the integrations discussed here were done with a horizontal grid spacing of 50 km (the current operational setup of the RFE). To verify the convergence of the numerical solution, all integrations were also performed with a horizontal grid spacing of 25 km. The increase in model resolution did not change the general characteristics of the forecasts and does not alter the conclusions of this paper. The characteristics of the mesoscale forecasts are given in Table 1. The validation periods of these forecasts ranged from 6 to 29 h and were mainly dictated by available meteorological data for model initialization.

#### b. SSM/I retrievals

To evaluate the accuracy of any icing scheme, one should ideally know the distribution of SLW in the atmosphere. Such information is rarely available, and direct validation is a difficult task. For example, Tremblay et al. (1995) compared model predictions and aircraft measurements by interpolating relevant model variables along space-time trajectories where aircraft measurements were available. Time series of temperature and cloud water (model output) as well as ice and SLW (generated from the T95 scheme) were qualitatively compared with aircraft measurements of these quantities. Such comparisons have shown that the T95 scheme has some merits, but due to limited spatial coverage and large scaling differences between aircraft data and model output, a definitive mesoscale evaluation was not possible.

Curry and Liu (1992) analyzed *Nimbus-7* microwave radiances, U.S. Air Force three-dimensional nephelyses, the European Centre for Medium-Range Weather Forecasts analyses, and high-resolution infrared satellite radiances and deduced the distribution of SLW over the North Atlantic Ocean. More recently, Lee et al. (1994) employed a retrieval of the liquid water path (LWP) derived from the SSM/I in conjunction with infrared images and numerical model output, to generate short-term forecasts of marine aircraft icing potential.

In this study, SSM/I retrievals of the vertically integrated liquid water, satellite infrared measurements,

TABLE 1. Characteristics of mesoscale forecasts for validation of aircraft icing forecast schemes.

Case	Forecast time (h)	Validation time & date UTC (1992)	DMSP overpass	Comments
A	24	0000 22 Jan	F-10	Ahead of a warm front
B	14	2000 05 Feb	F-11	Comma cloud structure
C	6	1200 08 Feb	F-10	Ahead of a warm front
D	14	2000 08 Feb	F-11	Ahead of a warm front
E	23	2300 16 Feb	F-10	Low pressure center: cold and warm fronts
F	20	2000 16 Feb	F-11	Low pressure center: cold and warm fronts
G	17	2300 29 Feb	F-10	Warm sector clouds
H	14	2000 29 Feb	F-11	Warm sector clouds
I	28	2200 14 Mar	F-10	Low pressure center: cold and warm fronts
J	25	1900 14 Mar	F-11	Low pressure center: cold and warm fronts
K	6	1200 16 Jan	F-10	Low pressure center: cold and warm fronts
L	16	2200 16 Jan	F-10	Low pressure center: cold and warm fronts
M	17	2300 16 Jan	F-10	Low pressure center: cold and warm fronts
N	29	1100 17 Jan	F-10	Low pressure center: cold and warm fronts
O	13	1900 16 Jan	F-11	Low pressure center: cold and warm fronts
P	12	1200 8 Mar	F-10	Low pressure center: cold and warm fronts
Q	23	2300 8 Mar	F-10	Low pressure center: cold and warm fronts
R	9	0900 8 Mar	F-11	Low pressure center: cold and warm fronts

and analyses from the Canadian Meteorological Centre were used to obtain a diagnostic of the SLW distribution in the horizontal. At a given horizontal point, SLW was inferred if the cloud-top temperature was less than  $0^{\circ}\text{C}$ , SSM/I retrievals of LWP [computed from algorithms given in Petty (1990) and Petty and Katsaros (1990)] exceeded  $0.3 \text{ kg m}^{-2}$ , and the SSM/I signals were assessed as not being significantly influenced by ice crystal scattering. A minimum LWP of  $0.3 \text{ kg m}^{-2}$  was chosen because it was high enough to screen out liquid water regions associated with thin clouds, while being small enough to ensure that significant regions of liquid water within synoptic cloud were included. Since numerical forecast models are generally more accurate in reproducing high liquid water content clouds rather than low liquid water content clouds, the icing forecasts output from such models are expected to be highly correlated with major synoptic events. Thus, a minimum value of LWP eliminates low liquid water content or thin clouds from the satellite measurements and allows for a more reasonable and focused comparison between the model and satellite measurements.

More specifically, when the freezing level was at the surface with no layer exceeding  $0^{\circ}\text{C}$  aloft, each horizontal point with  $\text{LWP} \geq 0.3 \text{ Kg m}^{-2}$  was flagged as SLW. For situations where the freezing level was aloft with cloud-top temperature  $< 0^{\circ}\text{C}$ , SLW was also determined to be present, provided no ice particles were detected. When both solid and liquid phases coexist within a vertical column (i.e., ice particles are detected), there is a possibility to have either ice particles or mixed phase above the freezing level. In such a situation, there was an ambiguity, and the retrieval algorithms were unable to distinguish a warm liquid water from an SLW region. The presence of ice particles was inferred from the polarization-corrected brightness

temperature (PCT) depression at 85 GHz (Spencer et al. 1989). The PCT responds to volume scattering by ice particles and infers the concentration and size of frozen particles. Those SSM/I pixels where the PCT depression is above 10 K can be used to identify the regions of solid particles (Grody 1993; Petty 1990). By selecting meteorological situations where the fractional coverage of  $\text{PCT} > 10 \text{ K}$  was less than 10%, it was expected that the number of ambiguous points was not large enough to affect the SLW retrievals. Moreover, since these specific points were generally embedded within large-scale regions of SLW, they were expected to be positive contributors to the SLW signal.

Any uncertainties associated with the retrieval of PCT, LWP, or any other aspect of the algorithm package could affect the determination of SLW regions. This highlighted the difficulty in using current SSM/I algorithms to either predict icing regions as proposed by Lee et al. (1994) or to validate icing forecasts as in the present work. However, if the dataset is biased because some SLW points are incorrectly diagnosed, then this bias should affect each icing forecast procedure equally. Consequently, this technique should not be expected to validate absolutely the different icing forecast algorithms but should provide a reasonable framework for comparing the advantages and limitations of the different icing forecast procedures.

SLW retrievals are presented on a regular longitude–latitude grid with a grid spacing roughly equivalent to 25 km, which is near the nominal resolution of the SSM/I. As an example, Fig. 3a depicts an SLW retrieval (gray shading) associated with a warm frontal system based on the 8 February 1992 DMSP F10 overpass at 1200 UTC. The figure shows a widespread area of SLW southeast of Newfoundland over the ocean. The extent of solid phase remains confined to small

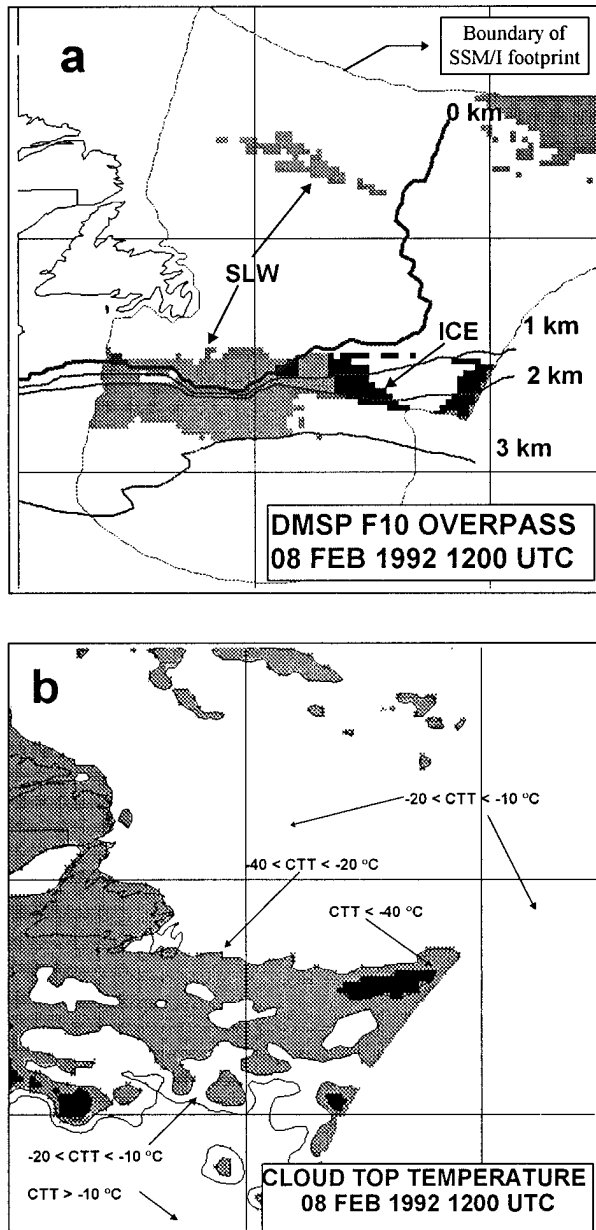


FIG. 3. (a) A supercooled liquid water retrieval (Cloud top temperature < 0°C; gray: LWP > 0.3 kg m<sup>-2</sup> and PCT < 10 K, black: LWP > 0.3 kg m<sup>-3</sup> and PCT > 10 K) based on the *DMSP F10* overpass of 1200 UTC 8 February 1992, superimposed with freezing level contours (km). (b) Cloud-top temperature (CTT) from *NOAA-12 IR* (gray: -20°C < CTT < -40°C; black: CTT < -40°C). Temperatures greater than -20°C are not shaded. The irregular contour delineates the SSM/I footprint.

regions where the freezing level was lower than 2 km (black shading). As discussed above, in such regions the SLW retrieval was ambiguous. Generally cloud-top temperatures were below -20°C with embedded convective cells less than -40°C (Fig. 3b).

**4. Verification procedure**

Since the retrieval procedure outlined in the previous section is based on vertically integrated values, it gives the horizontal distribution of SLW events. Thus, this technique can be used only within a two-dimensional framework to evaluate the accuracy of the forecast system for positioning SLW events in the horizontal. Strictly, a three-dimensional forecast verification can be done from a direct comparison between model output and research aircraft measurements, such as in T95; however, this approach is limited due to the relatively small size of the dataset. In principle, using a PIREPs dataset for forecast verification allows a three-dimensional validation but the nonsystematic and biased nature of these observations causes problems (Brown et al. 1995). No direct three-dimensional forecast verifications were attempted in this work but an indirect evaluation will be discussed in the next section.

The schemes outlined above (APP, SPS, and T95) produce yes/no categorical forecasts on the three-dimensional grid of the weather prediction model. To ensure a consistent comparison with SSM/I retrievals, these three-dimensional forecasts were collapsed to two dimensions by identifying horizontal points where an SLW forecast existed somewhere in the vertical. The signal detection theory (SDT) (Mason 1982) is a verification procedure applicable to two-state categorical weather elements, such as in the present case. With the SDT, the forecast skill can be evaluated from the probability of detection (POD) and the probability of false detection (POFD) [from this point, the nomenclature of Doswell et al. (1990) will be adopted]:

$$POD = \frac{A}{A + B} \tag{5}$$

$$POFD = \frac{C}{C + D}, \tag{6}$$

and *A*, *B*, *C* and *D* are defined in the 2 × 2 contingency table (Table 2). The POD can be described as the number of correct SLW forecasts divided by the total number of SLW observations. Conversely, the POFD is the number of erroneous SLW forecasts divided by the total number of non-SLW observations. A perfect forecast is described by POD = 1 and POFD = 0.

A straightforward application of these specific statistical parameters [or other related parameters, such as

TABLE 2. A contingency table to evaluate the icing forecasts.

Observed SLW	Forecast SLW	
	Yes	No
Yes	<i>A</i>	<i>B</i>
No	<i>C</i>	<i>D</i>

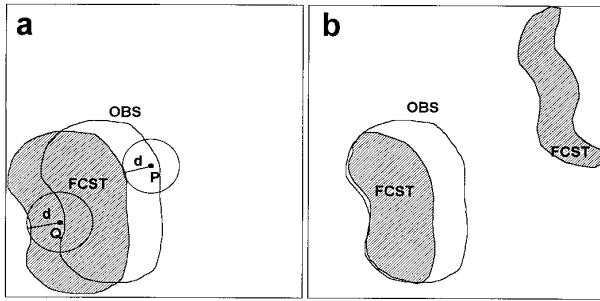


FIG. 4. Idealized situation showing two forecasts with the same number of hits and false alarms but with very different spatial characteristics.

those discussed in Doswell et al. (1990), Schaefer (1990 and references therein)] to icing forecast schemes may yield to incorrect evaluations. As an example, Fig. 4 depicts two forecasts with exactly the same POD and POFD. In an aircraft icing forecast perspective, Fig. 4a represents a much better prediction than Fig. 4b. The first forecast matches perfectly the shape and extent of observations but has a small geographical shift. The second prediction has issued false alarms over a region far from the threat zone, which represents a more serious error. Thus, there is a need to take into account the spatial structure when performing a forecast evaluation.

To consider this effect, a different approach is proposed. For each point with an SLW observation (as retrieved from SSM/I), the distance  $d$  to the nearest forecast is calculated (for example point P in Fig. 4a). Let  $N_1(d)$  represent the distribution of these distances (the number of SLW observations where  $d$  is within the interval  $d$  and  $d + \delta d$ ). Next, for each point where SLW is not detected, the distance  $d$  to the nearest non-SLW forecast is also calculated (point Q in Fig. 4a). Alternatively, let  $N_0(d)$  represent the distribution of these distances. Clearly,  $N_1(0) + N_0(0)$  delineates the number of perfect hits. For example in Fig. 4,  $N_1(0)$  is the number of points within the intersection of the forecast and observations.

The skill of forecasts can be discussed in terms of the cumulative frequency functions of these distributions:

$$f_0(d) = \frac{\sum_{x \leq d} N_0(x)}{\sum_x N_0(x)} \quad (7)$$

$$f_1(d) = \frac{\sum_{x \leq d} N_1(x)}{\sum_x N_1(x)}, \quad (8)$$

where  $f_1(d)$  can be interpreted as the POD within a circle of radius  $d$  of the observations, and  $1 - f_1(d)$  as the probability of misses (POM) (fraction of missed

events lying outside the validation circle). Equivalently,  $f_0(d)$  represents the fraction of correct forecasts of non-SLW events within the validation circle. In other words, it can be called the probability of correct null forecasts (POCN). Within this framework,  $1 - f_0(d)$  delineates the POFD and the SDT is a particular case obtained as  $d \rightarrow 0$ .

It should be noted that the present statistical environment provides an evaluation of forecasts relative to observations. However, for any joint distribution of forecast and observations there are two possible factorizations (Murphy and Winkler 1987). Thus, it is also possible to discuss the accuracy of forecasts in a frame of reference relative to forecasts, that is, by calculating the distances from a forecast to the nearest observation. It is felt that the former description is the user's point of view, while the later is related to the forecaster's perspective. To illustrate the difference between these two concepts, consider the following example. Let  $d$  the maximum distance required to validate a forecast and consider two stations P and Q, separated by a distance exceeding  $2d$ . Suppose now that two yes forecasts are produced within a distance  $d$  from P. If both stations observe an event, from the user's point of view, only station P has had a correct forecast, so the score is 50% of success. However, from the forecaster's side both forecasts are validated, since P is in their neighborhood, and a score of 100% must be inferred. Mathematically speaking, these two approaches are

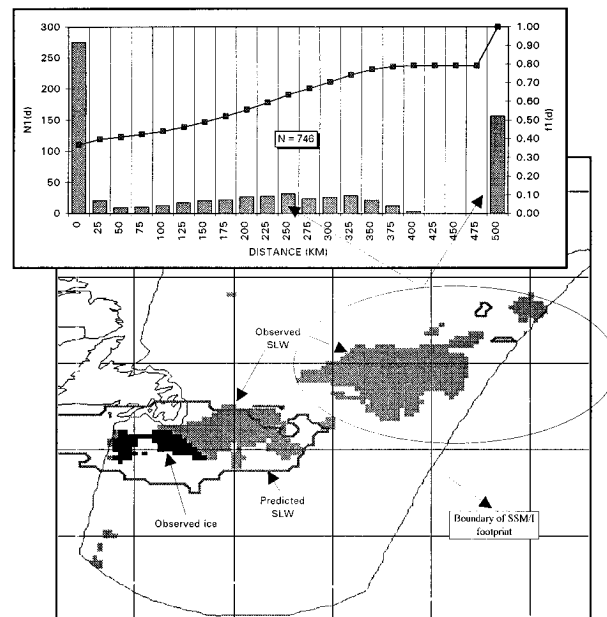


FIG. 5. A map showing the integrated SLW distribution retrieved from the SSM/I (shaded as in Fig. 3) and the outline (contoured) of forecast (12-h forecast valid 1200 UTC 8 March 1992) region of SLW. The embedded diagram shows the distribution of hits  $N_1(d)$  and the corresponding frequency function  $f_1(d)$ .

workable. However, it is felt that the user's interpretation is intuitively easier to understand and is more related to aviation purposes, so in the next section the forecast system will be evaluated from the user's side.

Figure 5 shows an example of the distribution  $N_1(d)$  calculated for a 12-h SLW forecast valid 1200 UTC 8 March 1992. The map in the figure shows that this prediction missed a large region of SLW over the ocean that contributed to the large values of  $N_1(d)$  at intermediate ( $100 \text{ km} \leq d \leq 375 \text{ km}$ ) and far ( $d \geq 500 \text{ km}$ ) distances, as shown in the embedded diagram in Fig. 5. The cumulative frequency function plotted on the same graph shows that only 45% of SLW events were detected within an accuracy of 100 km ( $2\Delta x$ ), emphasizing the poor quality of this specific forecast.

### 5. Forecast verifications

Table 1 describes the cases used for the comparison of the three forecast schemes. These cases encompassed fundamental aspects of winter storms, including warm frontal regions, low pressure areas, cold fronts, and warm sector icings. These are regions where the most severe icing was expected to be found and should provide a strong background for comparison of the three icing forecast schemes. Meteorological analyses, available every 6 h do not provide cloud water, the essential variable for the absolute verification of the T95 scheme. To verify the proposed scheme, we have to rely on the SSM/I analysis of SLW.

To match SSM/I retrievals, each forecast listed in Table 1 was mapped on a 25-km (latitude–longitude) grid and was evaluated with the verification technique outlined above. Results are presented in Fig. 6. The two diagrams in the figure show the frequency functions  $f_1(d)$  and  $f_0(d)$  for each forecast in the dataset. To help in the visualization of the results, the frequency functions are represented with a stacked bar format. The distance  $d$  was classified with 25-km bins that were close to the nominal resolution of the SSM/I at 37 GHz. In the figure, only distances  $\leq 100 \text{ km}$  ( $2\Delta x$  for model forecasts) are represented; forecasts with  $d > 100 \text{ km}$  were considered to have missed their target. Those forecasts enclosed within the first bin [ $f_1(0)$  and  $f_0(0)$ ] were hits and those within  $f_1(100)$  and  $f_0(100)$  were accurate at a  $2\Delta x$  level. Illustrating the extremes found in these situations, the model had  $f_1(0) = 0.9$  for the forecast L but  $f_1(0) = 0.05$  for the forecast G. Alternatively, the model had a fair degree of success in detecting non-SLW events. For example, the forecast showed  $f_0(0) = 0.95$  and  $f_0(100) = 0.99$  for case A, but for case I, which was a bad forecast for non-SLW events,  $f_0(0) = 0.70$  and  $f_0(100) = 0.94$ . In these diagrams, the category labeled TOT represents the same calculation for all data points. Overall, about 50% of SLW events were detected at their right location, 65% within a  $1\Delta x$  distance, and 75% within a  $2\Delta x$  distance. Roughly 25% of SLW events were absolutely missed

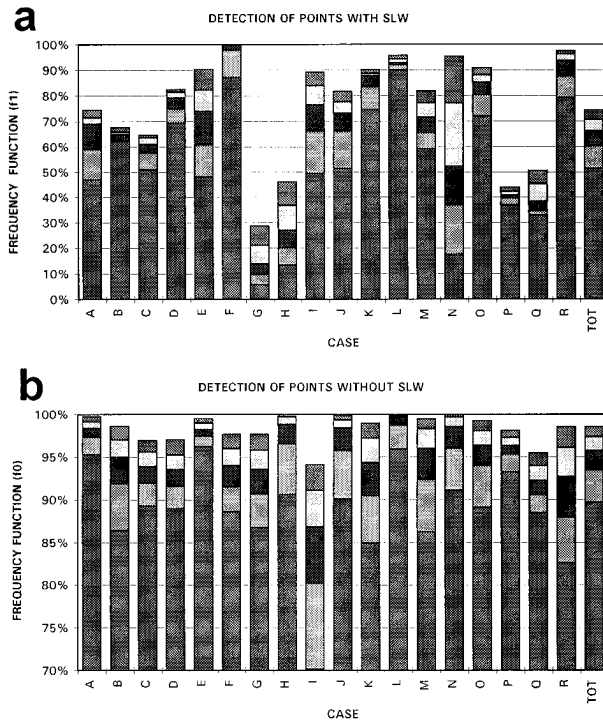


FIG. 6. (a) The frequency functions  $f_1(d)$  for SLW events and (b)  $f_0(d)$  for non-SLW events for each forecasts in the dataset, displayed with 25-km bins stacked bar format (shading corresponds to  $d = 0, 25, 50, 75$  and  $100 \text{ km}$ ). The category labeled TOT represents the entire dataset.

( $d > 100 \text{ km}$ ) by the forecasting system. About 90% of the non-SLW events were forecasted at their right location, and about 1% of these events were not forecasted within the  $2\Delta x$  accuracy, as shown in Fig. 6b.

From Fig. 6, the forecast J appears to be representative of the general behavior of the total dataset. The characteristics of this forecast, shown in Fig. 7, illustrate a typical product of the aircraft icing forecast system. The SLW was mainly associated with a developing low pressure center over the North Atlantic. In particular, the forecast captured the general structure of the SLW, although the forecast SLW field lagged behind the observed.

To examine, in detail, the behavior of the different SLW schemes, Fig. 8 depicts the frequency functions (as in Fig. 6) calculated for the entire dataset using four different SLW forecast algorithms. To ensure a consistent comparison, the APP and SPS criteria were applied within the same temperature interval  $-20^\circ\text{C} \leq T < 0^\circ\text{C}$ . The SPW scheme was identical to the SPS scheme but used the APP condition for vertical motion (i.e.,  $\omega \leq -0.2 \text{ Pa s}^{-1}$ ). Figures 8a and 8b show that the forecast performances are seriously impaired when  $d$  decreases, for each SLW algorithm. For example, for  $d = 0$ ,  $f_1$  falls below 0.60, which is close to the number of expected correct forecasts due purely to chance. This

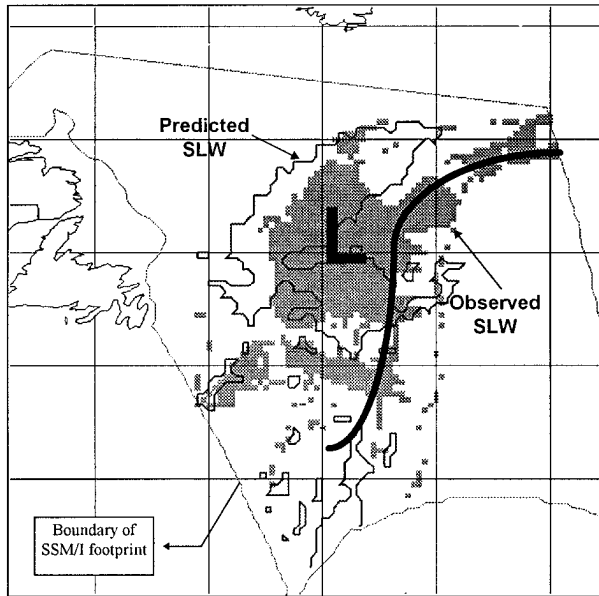


FIG. 7. Integrated SLW distribution for the 25-h forecast (contoured) valid 1900 UTC 14 March 1992 and retrieved (shaded), illustrating a typical product of the forecasting system. The observed frontal analysis and low pressure position are also displayed.

suggests that the system does not produce reliable point forecasts ( $d = 0$ ). However, as indicated by the values of  $f_0$  and  $f_1$  displayed in Fig. 8, it generates useful predictions within 100-km accuracy. Figure 8b demonstrates that the SPS scheme systematically flagged most of the mesoscale domain for aircraft icing, as indicated by the low value of  $f_0(100) = 0.30$ , implying its uselessness in the present context. However, if the SPS scheme is modified for vertical velocity (SPW), dramatic improvement occurs and  $f_0(100)$  increases to about 0.95.

Figure 8a shows that T95 ( $f_1(100) = 0.745$ ), APP ( $f_1(100) = 0.755$ ), and SPW ( $f_1(100) = 0.790$ ) schemes have about the same accuracy for detecting the vertically integrated SLW signal. This relatively weak dependence of  $f_1$  with respect to the forecasting technique used suggests that the quality of the SLW forecast was strongly related to model accuracy in predicting vertical velocity, temperature, moisture, and cloud content. Since there were 8822 SLW events in the dataset, these values of  $f_1$  imply that T95 has missed 2250 of these events, APP 2161 and SPW 1853.

Figure 8b indicates that  $f_0(100) = 0.987$  for T95,  $f_0(100) = 0.952$  for SPW, and  $f_0(100) = 0.963$  for APP. This implies a POFD of 0.013 for T95, 0.048 for SPW, and 0.037 for APP. Since there was 74 784 non-SLW observations in the dataset, the T95 scheme had erroneously flagged as SLW 972 data points. Equivalently, the SPW scheme erroneously flagged 3589 data points. Thus SPW had issued a number of false detections greater by a factor of 3.7 with respect to T95. The

performance of APP was superior to SPW with a total number of false detections of 2804. As a comparison, T95 missed 397 more events than SPW but removed false alarms at 2617 points.

From the above discussion, it can be established that T95 had issued 7544 yes forecasts, and SPW 10 558 and APP 9465. This implies that T95 tends to slightly *underforecast* the SLW extent (8822 observed events), while SPW and APP tend to slightly *overforecast* the SLW signal. Another useful statistical evaluator for forecasts comparisons is the frequency of hits (FOH). This parameter is defined as a proportion of yes forecasts matched by a yes observation. Its complement is the false alarm ratio (FAR) defined as the proportion of yes forecasts contraindicated by a no observation. It can easily be shown that FOH = 0.871 for T95, 0.660 for SPW, and 0.704 for APP. This corresponds to FAR = 0.129 for T95, 0.340 for SPW, and 0.296 for APP. Thus, according to this statistical evaluator, 87% of SLW forecasts issued by T95 were correct. This number falls to 70% for APP and to 66% for SPW.

The high number of false detections of the APP, SPS, and SPW schemes were likely related to the oversimplified cloud microphysics embodied in the construction of these algorithms. As discussed above, the physical principles underlying these icing forecast procedures may result in an erroneous distribution of SLW with temperature, implying that glaciated and super-cooled clouds were not discriminated correctly. Gla-

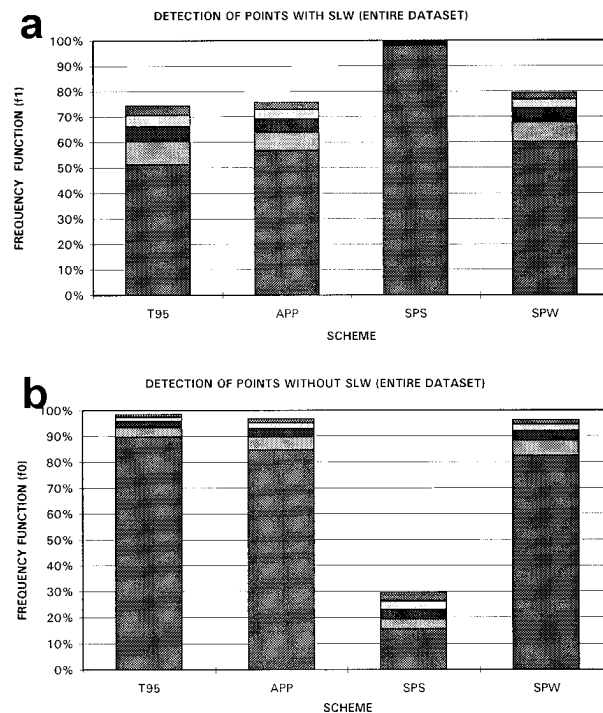


FIG. 8. (a) The frequency functions  $f_1(d)$  and (b)  $f_0$  as in Fig. 7 for the entire dataset and for the four forecast schemes.

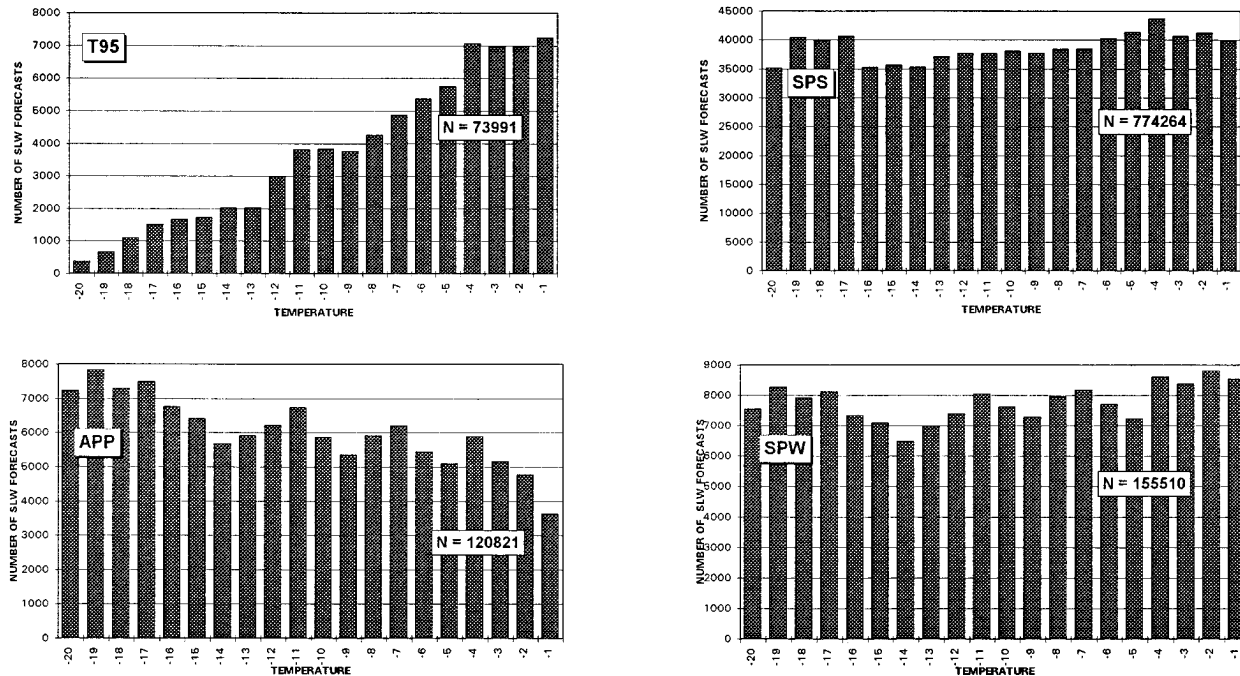


FIG. 9. Distribution of the number of forecasted icing events on the three-dimensional model grid with respect to temperature, for the four schemes. The total number of forecasts are also reported in each diagram.

ciated clouds may also have significantly contributed to the SLW horizontal hits of these procedures (Fig. 8a), emphasizing the need to discuss the vertical structure of the forecasts. Since the atmosphere is thermally stratified in the vertical, the distribution of SLW forecasts with temperature gives indirect information about their vertical structure and reveals their physical consistence. This is illustrated in Fig. 9, which displays the distribution of all predicted SLW events on the three-dimensional model grid, with respect to temperature (for the 18 cases). The T95 scheme has generated a sharp distribution of SLW events decreasing with

colder temperature. On the other hand, both SPS and SPW schemes generated relatively flat distributions with respect to temperature. This result suggests that the probability of an ice encounter is independent of temperature and contradicts observations (see Fig. 10). The APP scheme had a clear tendency to generate a decreasing frequency of SLW events with increasing temperature. In fact, such a tendency is just the reverse of observations, and this nonphysical behavior classifies the APP scheme as an inadequate forecasting tool for aircraft icing. As indicated in the diagrams of Fig. 9, the total number of forecasted SLW points is strongly dependent on the algorithm used. For example, the SPS scheme generated 10 times the number of SLW events compared with the T95 scheme, showing the high rate of false alarms characterizing this scheme. The SPW scheme generated approximately 2 times the number obtained from the T95 scheme. Figure 9 shows that colder temperatures contribute to this difference, suggesting that the SPW scheme issued icing forecasts within cold clouds classified as being glaciated by the T95 scheme.

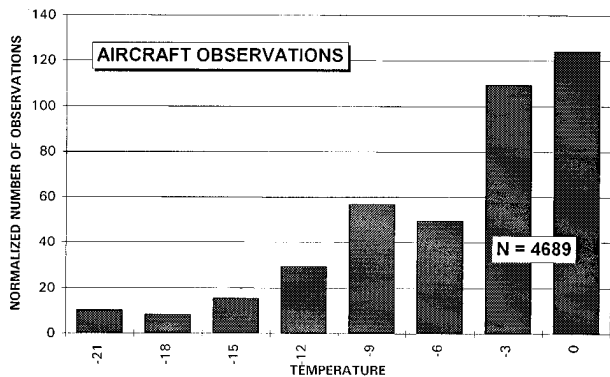


FIG. 10. Number of aircraft measurements with liquid water content greater than  $0.025 \text{ g m}^{-3}$ , normalized by the percentage of total number of measurements in a given temperature bin.

To provide a quantitative evaluation of the icing forecast scheme, Fig. 10 depicts the distribution of SLW events with temperature observed during the CASP II field experiment. This dataset was generated from a compilation of observations taken during 33 research flights (Cober et al. 1995) and averaged into  $3^\circ\text{C}$  bins. Figure 10 shows that the 4689 SLW observations were distributed with temperature in a fashion

closely related with the T95 scheme but unrelated to the other schemes. This reemphasizes the need for a physically based scheme to generate forecasts consistent with observations.

## 6. Summary

In this paper, the accuracy of three aircraft icing (or equivalently SLW) forecast schemes were evaluated. From integrations of the Canadian RFE model with an explicit cloud scheme, a dataset of mesoscale icing forecasts, representative of typical east coast Canadian winter storms, was constructed. The forecasts were evaluated by comparison with SLW distributions derived from satellite data.

It was argued, and subsequently demonstrated from numerical integrations, that the forecast procedures of Appleman (1954) and Schultz and Politovich (1992) resulted in erroneous distributions of SLW events as a function of temperature. The Schultz and Politovich (1992) forecast scheme yielded a flat distribution of SLW frequency with respect to temperature. For the Appleman (1954) procedure, the frequency of icing forecasts decreased as temperature increased. These contradict recent measurements (Cober et al. 1995) and imply that glaciated clouds were systematically included in SLW forecasts resulting in a large number of false detections. Therefore, the operational use of these schemes is questionable.

The procedure of Tremblay et al. (1995) was found to give a realistic distribution, with the number of SLW events sharply decreasing as temperature decreases. This distribution is consistent with atmospheric conditions, since the likelihood of supercooled cloud water tends to decrease with temperature (Pruppacher and Klett 1978) and suggests that the basic cloud physics embodied in the derivation of the Tremblay et al. (1995) scheme is fairly well representative of microphysical processes found in natural clouds. This resulted in the total number of SLW forecasts (on the three-dimensional model grid) decreasing by a factor of 2 when compared to the Appleman scheme and by a factor of 10 with respect to the Schultz and Politovich (1992) scheme.

To provide a verification of SLW forecasts, a retrieval technique for the SLW, based on infrared satellite imagery, microwave data from the SSM/I DMSP satellites, and weather analyses was suggested. An SLW retrieval, matching each numerical icing forecast, was then generated and used with a validation procedure based on the signal detection theory to evaluate the accuracy of forecasts.

The Tremblay et al. (1995) icing forecast system detected, with a perfect precision (i.e., at the right location), 50% of the model grid points where an SLW signal is retrieved from satellite analyses. However, 65% of the signal can be detected within one grid length  $\Delta x$  (50 km, in the present study) and 75%

within a  $2\Delta x$  range. The 25% of the forecasts outside the validation circle of radius  $2\Delta x$  were considered as missed events. The false alarm rate was rather low, with about 99% of non-SLW events detected by the system within a  $2\Delta x$  range (100 km). This number of false alarms was roughly equivalent to 13% of the observations of SLW. It was also established that T95 forecasts had a false alarm ratio of 13%, while APP and SPW forecasts had FARs of 30% and 34%, respectively. This outlines the limit of operational applicability of the forecasting system and should be used as a guideline to properly educate potential users. The Tremblay et al. (1995) scheme represents a significant improvement in accuracy over existing schemes and should be considered a good first guess for aviation forecasts.

The icing forecast scheme of Tremblay et al. (1995) shows a tendency to reduce the extent of the regions of potential icing, as compared to the schemes of Appleman and Schultz and Politovich. However, the forecast regions are only potential icing regions and may still not have any supercooled liquid water. The model predicts SLW regions on grid scales of 50 km, which are significantly larger than the average horizontal extent of supercooled liquid water in east coast winter storms. From aircraft measurements through 31 such storms, Cober et al. (1995) showed that the average horizontal extent of supercooled liquid water was 4 km. Consequently, until more substantial microphysical processes are included on a smaller grid scale, model forecasts should be considered only as a guidance or a best guess of regions of potential icing.

*Acknowledgments.* This work was funded by the Canadian National Search and Rescue Secretariat. Funding for the field project was also provided by the Institute for Aerospace Research (IAR) of the National Research Council of Canada, Boeing Commercial Airplane Group, and Airbus Industrie. The authors would like to acknowledge Prof. G. W. K. Moore for his help in processing the satellite infrared images. Special thanks to Drs. Godeleive Deblonde, Stephane Laroche, and Ashu Dastoor for constructive comments on the manuscript.

## REFERENCES

- Air Weather Service, 1980: Forecasting guide on aircraft icing. AWS Tech. Rep. AWS/TR-80/001, U.S. Air Force, Scott AFB, IL, 55 pp. [NTIS AD-A085 490/1.]
- Appleman, H., 1954: Design of a cloud-phase chart. *Bull. Amer. Meteor. Soc.*, **35**, 223–225.
- Benoit, R., J. Cote, and J. Mailhot, 1989: Inclusion of a TKE boundary layer parameterization in the Canadian Regional finite-element model. *Mon. Wea. Rev.*, **117**, 1726–1750.
- Brown, B., R. Bullock, G. Thompson, and R. Bruitjes, 1995: WISP94 real-time icing prediction and evaluation program (WI-REP): Statistical issues and forecast verification results. Preprints, *Sixth Int. Conf. on Aviation Weather Systems*, Dallas, TX, Amer. Meteor. Soc., 207–212.

- Cober, S. G., G. A. Isaac, and J. W. Strapp, 1995: Aircraft icing measurements in east coast winter storms. *J. Appl. Meteor.*, **34**, 88–100.
- Cotton, W. R., and R. A. Anthes, 1989: Storm and cloud dynamics. *International Geophysics Series*, No. 44, Academic Press, 883 pp.
- Curry, J. A., and G. Liu, 1992: Assessment of aircraft icing potential using satellite data. *J. Appl. Meteor.*, **31**, 605–621.
- Doswell, C. A., R. Davies-Jones, and D. L. Keller, 1990: On summary measures of skill in rare event forecasting based on contingency tables. *Wea. Forecasting*, **5**, 576–585.
- Fletcher, N. H., 1966: *The Physics of Rainclouds*. Cambridge University Press, 386 pp.
- Grody, N. C., 1993: Remote sensing of the atmosphere from satellites using microwave radiometry. John Wiley and Sons, 596 pp.
- Lee, T. F., J. R. Clark, and S. D. Swadley, 1994: Potential applications of the SSM/I cloud liquid water parameter to the estimation of marine aircraft icing. *Wea. Forecasting*, **9**, 173–182.
- Mailhot, J., R. Sarrazin, B. Bilodeau, N. Brunet, A. Methot, G. Pellerin, C. Chouinard, L. Garand, C. Girard, and R. Hogue, 1995: Changes to the Canadian regional forecast system. *Atmos.–Ocean*, **33**, 55–80.
- Mansur, M. V., 1984: Automated aircraft icing forecast technique. Air Force Global Central Tech. Rep. AFGWC-PR-84-001, U.S. Air Force, 108 pp.
- Mason, I., 1982: A model for assessment of weather forecasts. *Aust. Meteor. Mag.*, **30**, 291–303.
- Modica, G. D., D. W. Johnson, and R. M. Rasmussen, 1994: Application of an explicit microphysics mesoscale model to a regional icing event. *J. Appl. Meteor.*, **33**, 53–64.
- Moss, S. J., and D. W. Johnson, 1994: Aircraft measurements to validate and improve numerical model parameterization of ice to water ratios in clouds. *Atmos. Res.*, **34**, 1–25.
- Murphy, A. H., and R. L. Winkler, 1987: A general framework for forecast verification. *Mon. Wea. Rev.*, **115**, 1330–1338.
- Petty, G. W., 1990: On the response of the Special Sensor Microwave/Imager for atmospheric parameters retrievals. Ph.D. thesis, University of Washington, 291 pp.
- , and K. B. Katsaros, 1990: New geophysical algorithms for the Special Sensor Microwave Imager. *Fifth Int. Conf. on Satellite Meteorology and Oceanography*, London, United Kingdom, Amer. Meteor. Soc., 247–251.
- Pruppacher, H. R., and J. D. Klett, 1980: *Microphysics of Clouds and Precipitation*. D. Reidel, 714 pp.
- Pudykiewicz, J., R. Benoit, and J. Mailhot, 1992: Inclusion and verification of a predictive cloud-water scheme in a regional weather prediction model. *Mon. Wea. Rev.*, **120**, 612–626.
- Roger, R. R., 1979: *A Short Course in Cloud Physics*. International Series in Natural Philosophy, No. 96, 235 pp.
- Sand, W. R., W. A. Cooper, M. K. Politovich, and D. L. Veal, 1984: Icing conditions encountered by a research aircraft. *J. Climate Appl. Meteor.*, **23**, 1427–1440.
- Schaefer, J. T., 1990: The critical success index as an indicator of warning skill. *Wea. Forecasting*, **5**, 570–575.
- Schultz, P., and M. Politovich, 1992: Toward the improvement of aircraft icing forecasts for the continental United States. *Wea. Forecasting*, **7**, 491–500.
- Spencer, R. W., H. M. Goodman, and R. E. Hood, 1989: Precipitation retrieval over land and over ocean with the SSM/I: Identification and characteristics of the scattering signal. *J. Atmos. Oceanic Technol.*, **6**, 254–273.
- Stewart, R. E., 1991: Canadian Atlantic Storms Program: Progress and plans of the meteorological component. *Bull. Amer. Meteor. Soc.*, **72**, 364–371.
- Sundqvist, H., E. Berge, and J. E. Krisjansson, 1989: Condensation and cloud parameterization studies with a mesoscale numerical weather prediction model. *Mon. Wea. Rev.*, **117**, 1641–1657.
- Tremblay, A., A. Glazer, W. Szyrmer, G. Isaac, and I. Zawadzki, 1995: Forecasting of supercooled clouds. *Mon. Wea. Rev.*, **123**, 2098–2113.
- Zawadzki, I., L. Ostiguy, and R. Laprise, 1993: Retrieval of the microphysical properties in a CASP storm by integration of numerical kinematics model. *Atmos.–Ocean*, **31**, 201–233.