

Article

Study on Icing Environment Judgment Based on Radar Data

Jinhu Wang ^{1,2,3,*} , Binze Xie ^{1,*}, Jiahua Cai ¹, Yuhao Wang ¹ and Jiang Chen ¹

¹ Collaborative Innovation Center on Forecast and Evaluation of Meteorological Disasters, Key Laboratory for Aerosol-Cloud-Precipitation of China Meteorological Administration, Nanjing University of Information Science and Technology, Nanjing 210044, China; 20181204001@nuist.edu.cn (J.C.); 20201248028@nuist.edu.cn (Y.W.); 20201248002@nuist.edu.cn (J.C.)

² Key Laboratory of Middle Atmosphere and Global Environment Observation, Institute of Atmospheric Physics, Chinese Academy of Sciences, Beijing 100029, China

³ Nanjing Xinda Institute of Safety and Emergency Management, Nanjing 210044, China

* Correspondence: goldtigerwang@nuist.edu.cn (J.W.); 20181204022@nuist.edu.cn (B.X.); Tel.: +86-138-1451-2847 (J.W.)

Abstract: As a major threat to aviation flight safety, it is particularly important to make accurate judgments and forecasts of the ice accumulation environment. Radar is widely used in civil aviation and meteorology, and has the advantages of high timeliness and resolution. In this paper, a variety of machine learning methods are used to establish the relationship between radar data and icing index (Ic) to determine the ice accumulation environment. The research shows the following. (1) A linear model was established, based on the scattering rate factor (Zh), radial velocity (v), spectral width (w), velocity standard deviation (σ) detected by 94 GHz millimeter wave radar, and backward attenuation coefficient (β) detected by 905 nm lidar, so linear regression was carried out. After principal component analysis (PCA), the correction determination coefficient of the linear equation was increased from 0.7127 to 0.7240. (2) Ice accumulation was unlikely for samples that were significantly off-center. By clustering the data into three or four categories, the proportion of icing lattice points could be increased from 18.81% to 33.03%. If the clustering number was further increased, the ice accumulation ratio will not be further increased, and the increased classification is reflected in the classification of pairs of noises and the possibility of omission is also increased. (3) Considering the classification and nonlinear factors of ice accumulation risk, the neural network method was used to judge the ice accumulation environment. Two kinds of neural network structures were established for quantitative calculation: Structure 1 first distinguished whether there was ice accumulation, and further calculated the icing index for the points where there was ice accumulation; Structure 2 directly calculated the temperature and relative humidity, and calculated the icing index according to definition. The accuracy of the above two structures could reach nearly 60%, but the quantitative judgment of the ice accumulation index was not ideal. The reasons for this dissatisfaction may be the small number of variables and samples, the interval between time and space, the difference in instrument detection principle, and the representativeness of the ice accumulation index. Further research can be improved from the above four points. This study can provide a theoretical basis for the diagnosis and analysis of the aircraft ice accumulation environment.

Keywords: aviation safety; icing index; radar data; linear regression; clustering analysis; neural network



Citation: Wang, J.; Xie, B.; Cai, J.; Wang, Y.; Chen, J. Study on Icing Environment Judgment Based on Radar Data. *Atmosphere* **2021**, *12*, 1534. <https://doi.org/10.3390/atmos12111534>

Academic Editor: Jiping Liu

Received: 15 September 2021

Accepted: 16 November 2021

Published: 20 November 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Common causes of aircraft icing include: (1) encountering clouds with super-cooled water droplets during flight; (2) being contaminated prior to takeoff; and (3) encountering high concentrations of ice crystals during flight [1]. Ice accumulation in different parts of the aircraft can easily damage the aerodynamic characteristics or power system of the aircraft, or cause the pilot to misjudge the flight status of the aircraft, thus causing accidents. In 1994, American Eagle Flight 4184 crashed due to ice pack, but the investigation report noted that

there was no significant weather report of ice pack at the time of the accident [2]. In 2009, when Air France Flight 447 was flying over the Atlantic Ocean, the autopilot disconnected due to ice jam in the Pitot tube, and then the pilot misjudged the flight status and made an improper operation, resulting in the plane crash [3]. An investigation of 803 aviation accidents in the United States between 1975 and 1988 found that these accidents were partly or wholly related to ice accumulation on aircraft [4]. Although there are many anti-icing and deicing devices installed on modern aircraft [5–7], air accidents caused by ice pack still occur from time to time. Therefore, it is of great significance to establish a time-efficient and high-resolution algorithm for judging an ice pack environment.

There are many classical icing diagnosis and prediction algorithms abroad [8–11]. The commonly used instruments for remote sensing data inversion of aircraft ice pack include dual-wavelength radar, weather radar, meteorological satellite, microwave radiometer, etc. [12–14]. GOES (Geostationary Operational Environmental Satellite) is used to identify cloud features and super-cooled water to classify ice accumulation [15,16]. Javier et al. [17] demonstrated that the WRF (weather research and forecasting) model is capable of correctly detecting icing associated with mountain waves. Merino et al. [18] used in-cloud micro-physics measurements taken during 10 flights of a C-212 research aircraft under winter conditions. Bolgiani et al. [19] studied the risk of ice accumulation near Adolfo Suárez Madrid-Barajas International Airport based on a WRF model and satellite data. Pereira [20], based on ECMWFs (European Center for Medium-Range Weather Forecasts), presented a verification of different icing algorithms for two periods. Part 25: Airworthiness Standards for Transport Aircraft of the FAA (Federal Aviation Administration) provides conditions for atmospheric icing in Appendix C and conditions for freezing of super-cooled large droplets in Appendix O [21].

Some scholars have studied the shape and severity of body ice accumulation using a neural network algorithm [22–24]. Dong [25] proposed an ice accumulation fault detection system based on deep learning. Li et al. [26] established the relationship between flight conditions and aircraft ice accumulation. Kolbakir et al. [27] found that plasma brakes could prevent the wing from icing. Zhang et al. [28] proposed a probabilistic method under considering subjective and objective parameters uncertainties based on the Monte Carlo method for estimating the airplane icing severity. Zhou et al. [29] presented the reachability analysis based on the level set method. Cao et al. [30] established an engineering prediction method of aerodynamic derivatives based on an individual component CFD (computational fluid dynamics) calculation and narrow strip theory. Liu et al. [31] presented a comprehensive aircraft icing simulation tool implemented in a Navier–Stokes parallel multi-block solver. Wang et al. [32] established a safety analysis model to quantify the safety degree of aircraft icing under the fuzzy inputs and fuzzy state.

In conclusion, the data from various scholars' studies on the aircraft ice accumulation environment mainly focused on the conventional observation data, and there are few studies based on radar data. In the past research, the authors have conducted a preliminary qualitative study on the aircraft ice accumulation environment in China [33]. Therefore, in this paper, based on the detection data of millimeter wave radar and Lidar, linear regression, principal component analysis, cluster analysis, and neural network methods are used to study the direct inversion of ice index to realize the quantitative judgment of the ice environment by radar data. This method can improve the resolution and timeliness of the environmental judgment of ice accumulation.

2. Materials and Methods

2.1. Data

The Science and Technology Facilities Council (STFC) at Chilbolton Observatory in the UK is home to many meteorological and atmospheric science research observation systems. The Chilbolton Facility for Atmospheric and Radio Research (CFARR) has multiple radar and Lidar as well as observation equipment such as radiometers at different bands, weather sensors, rain gauges, and cloud cameras. The sources of radar data used in this

paper include cloud radar and Lidar, among which the 94 GHz cloud radar (Galileo radar) is located at 51.145° N, 1.437° W, and the Lidar (905 nm Vaisala CT75K Lidar Ceilometer) is located at 51.145° N, 1.427° W [34]. The millimeter-wave radar data used in this paper includes radar scattering rate factor (Zh), radial velocity (v), spectral width (w), and velocity standard deviation (σ), and the Lidar data used was backward attenuation coefficient (β). The website is <http://data.ceda.ac.uk/badc/chilbolton/data> (accessed on 1 September 2020).

The sounding data are from the Department of Atmospheric Science, University of Wyoming (UW, <http://www.weather.uwyo.edu/upperair/sounding.html>, accessed on 1 September 2020). The radiosonde data of Herstmonceux Station (position: 50.90° N, 0.32° E, station number: 03882) from several stations near Chilbolton were selected as the radiosonde data corresponding to the radar data in this paper, and the geographical location is shown in Figure 1. According to the altitude, temperature, and relative humidity information of the station, the ice accumulation index was calculated as the reference of the radar data. Since the grid points of the radiosonde data and radar data do not match perfectly, the linear interpolation method was used to interpolate the radiosonde data to the corresponding height according to the height of each grid point in the radar data, and then the icing index was calculated to generate training set and test set to realize the training and test of the model. Obviously, there will be errors in linear interpolation, and the interpolation result may not be equal to the real temperature and humidity at the corresponding height. This error may be greater when interpolation is interpolated outward.

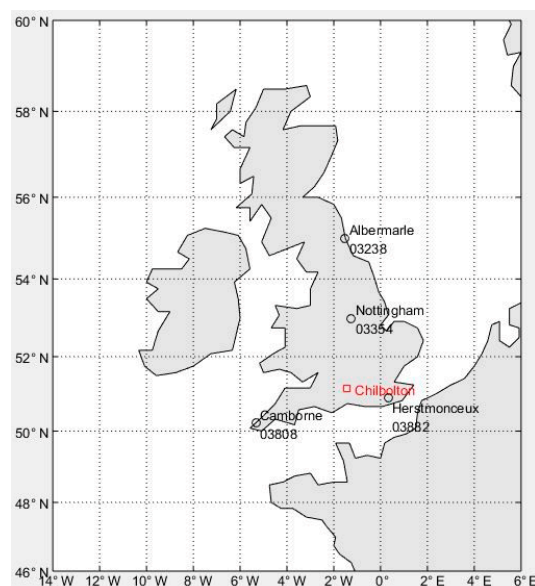


Figure 1. Location of Chilbolton Station and sounding stations. The figure shows the location and number of multiple sites around Chilbolton, with Herstmonceux being the closest.

2.2. Icing Index

The ICAO (International Civil Aviation Organization) recommends using the icing index ' I_C ':

$$I_C = 2(RH - 50) \times \frac{T(T + 14)}{-49} \quad (1)$$

where RH is the relative humidity (%), T is the temperature (°C), RH is less than 50%, or T is outside the range of −14 °C to 0 °C, so no ice accumulation occurs. The output range of I_C is 0–100. The larger the value, the stronger the ice accretion [35].

2.3. Methods

Equation (2) is used for standardized calculation:

$$X_i^* = \frac{X_i - \bar{X}}{S} \quad (2)$$

\bar{X} is the sample mean, and S is the sample standard deviation. After processing, the mean value of each variable is 0, and the standard deviation is 1. This processing can eliminate the influence of dimension and order of magnitude on the data analysis and algorithm model. Variables with * in this paper refer to those after standardized processing.

PCA is a commonly used dimensionality reduction algorithm for high-dimensional data. Its idea is to construct a new hyperplane, that is, a linear combination of existing coordinates to express all samples. The eigenvalues and eigenvectors of the covariance matrix were solved, the eigenvalues were sorted from largest to smallest, and the eigenvectors corresponding to the first q eigenvalues were taken as the solutions of the PCA [36].

The clustering algorithm used in this paper was the K-means algorithm. In this algorithm, K clustering centers are selected to cluster all samples into K classes, and the sum of the distances between each sample and its clustering center is calculated to describe the closeness of the samples around the mean vector in the cluster. When sum is the minimum, this clustering method is considered to be reasonable. Since it is difficult to find the optimal solution directly, the iterative method is adopted. Given K centers of the primary election, the iterative method is used to gradually generate the optimal solution [37].

Neural networks can be used for regression and classification to build possible relationships between inputs and outputs. The most fundamental component of a neural network is the neuron model. The neuron receives input signals from N neurons in the previous layer, and the signals are transmitted through weighted connections. The total input value received by the neuron is compared with the threshold value and processed through the “activation function” set by the network to obtain the output signal. Many such neurons are connected according to a certain hierarchical structure to obtain the neural network [38].

2.4. Feasibility Analysis

Since the discussion in this paper was mainly based on cloud radar data, the relationship between cloud radar data and radiosonde data was first analyzed to illustrate the feasibility of using cloud radar data to judge ice accumulation. Statistics were conducted on whether the radar measured data and whether the grid point was interpolated to obtain ice accumulation. The results are shown in Table 1.

Table 1. Correspondence between cloud radar data and icing.

Data Type	Sample Size	Sample Proportion
Class 0 (Lack of Radar Data, no Risk of Icing)	5379	73.48%
Class 1 (Lack of Radar Data, Risk of Icing)	431	5.888%
Class 2 (with Radar Data, Risk of Icing)	284	3.880%
Class 3 (with Radar Data, no Risk of Icing)	1226	16.75%

According to the statistics in Table 2, among the total 7320 samples, only 431 samples had an ice accumulation risk, which is not reflected in the radar data. Meanwhile, considering the geographical deviation of Chilbolton and Herstmonceux and the interference of interpolation on the sounding data, the data volume of such samples may be even smaller in the real situation. This indicates that it is meaningful to establish the relationship between radar data and icing index to judge ice accumulation.

Table 2. Correlation coefficient matrix.

	I_C	Zh^*	v^*	w^*	σ^*	β^*
I_C	1	0.4429	0.1959	0.5262	0.0834	−0.1005
Zh^*	0.4429	1	−0.3858	−0.0933	−0.5473	−0.2310
v^*	0.1959	−0.3868	1	0.0163	0.5585	0.1874
w^*	0.5262	−0.0933	0.0163	1	0.4268	0.3043
σ^*	0.0834	−0.5473	0.5585	0.4268	1	0.6059
β^*	−0.1005	−0.2310	0.1874	0.3043	0.6059	1

3. Contributions of Radar Data to Icing Index

The 0 h data collected by Galileo’s cloud radar and Lidar in November 2003 consisted of 192 altitudes per day for 26 days. Among them, cloud radar and Lidar obtained data, and a total of 48 data points with icing were obtained according to interpolation. These 48 samples are the training set discussed in this section.

First, variables were standardized according to Equation (2). After standardization, the correlation coefficients between dependent variables and independent variables and between independent variables are given in Table 2.

The linear equation between the icing index and five standardized variables was established:

$$I_C = 30.7516 + 15.6657Zh^* + 10.1979v^* + 14.5997w^* + 2.6934\sigma^* - 6.7355\beta^* \quad (3)$$

According to the data in Table 2, the icing index was more closely correlated with the reflectivity factor and spectral width, but it could also be seen that these data had strong autocorrelation and a large amount of information overlap. The principal component analysis method was used to process the data below, and the obtained variance contribution of variables and the correlation coefficient between the converted variables and ice accumulation index are given in Table 3, while the maximum absolute value of the autocorrelation coefficient between the new independent variables was also $<10^{-15}$, suggesting that each variable does not contain overlapping information.

Table 3. Variable contribution and transformation matrix.

Variable	X_1	X_2	X_3	X_4	X_5
Contribution	48.7108%	22.1684%	13.1940%	12.1635%	3.7633%
Coefficient	0.0358	0.3731	−0.3005	0.7156	0.0203

It can be found in Table 3 that X_1 , which had the largest variance contribution, had very little correlation with the ice accumulation index, which may indicate that X_1 mainly represents a large amount of noise contained in radar data. In general, all new variables with a cumulative contribution rate of 85% are selected after principal component analysis. Therefore, the expressions of X_1 , X_2 , X_3 and X_4 are given below:

$$X_1 = -0.4300Zh^* + 0.4103v^* + 0.3044w^* + 0.6000\sigma^* + 0.4405\beta^* \quad (4)$$

$$X_2 = 0.3825Zh^* - 0.5184v^* + 0.6799w^* + 0.0273\sigma^* + 0.3492\beta^* \quad (5)$$

$$X_3 = 0.3296Zh^* - 0.0355v^* - 0.5909w^* + 0.0204\sigma^* + 0.7352\beta^* \quad (6)$$

$$X_4 = 0.7078Zh^* + 0.6648v^* + 0.1794w^* + 0.0666\sigma^* - 0.1429\beta^* \quad (7)$$

Considering the variance contribution and correlation coefficient, the following three linear regression equations were established:

$$I_C = 30.7516 + 0.5410X_1 + 8.3538X_2 - 8.7225X_3 + 21.6289X_4 \quad (8)$$

$$I_C = 30.7516 + 8.3538X_2 - 8.7225X_3 + 21.6289X_4 \quad (9)$$

$$I_C = 30.7516 + 21.6289X_4 \quad (10)$$

Table 4 shows the test statistics of the four regression equations:

Table 4. Test statistics.

	R^2	$\overline{R^2}$	F	p	s^2
Equation (3)	0.7433	0.7127	24.3171	2.06×10^{-11}	159.6470
Equation (8)	0.7428	0.7189	31.0531	3.54×10^{-12}	156.1843
Equation (9)	0.7416	0.7240	42.0839	5.51×10^{-13}	153.3960
Equation (10)	0.5120	0.5014	48.2668	1.09×10^{-8}	277.0431

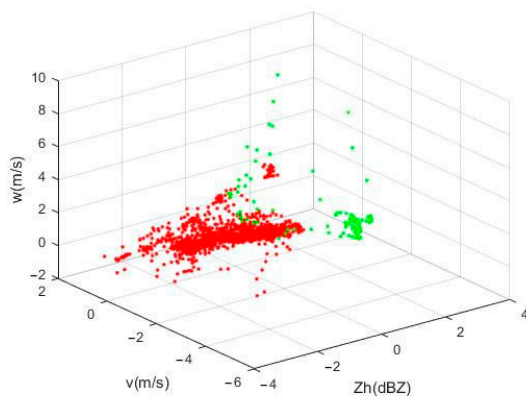
Through the calculation of F statistics, the hypothesis test of $\alpha = 0.001$ under the corresponding degree of freedom was passed, that is, there was a 99.9% confidence that the linear relationship of the above four formulas was significant. In terms of statistics, Equation (3) had the largest coefficient of determination, Equation (9) had the largest corrected coefficient of determination and the smallest p value and error variance, and Equation (10) had the largest F value. This statistic can illustrate the following points:

- (1) It is significant to use principal component analysis to process data in this study;
- (2) Although the variable X_1 contains a lot of information, introducing X_1 into the linear regression will reduce the correction determination coefficient and linear significance of the equation, and increase the error variance, so it is reasonable not to introduce X_1 , which also confirms that X_1 mainly represents the noise in the radar data; and
- (3) Although the correlation coefficient of X_2 and X_3 is small, and the opposite value of their correlation is similar, it will cause the loss of a lot of the main information in the sample if they are removed.

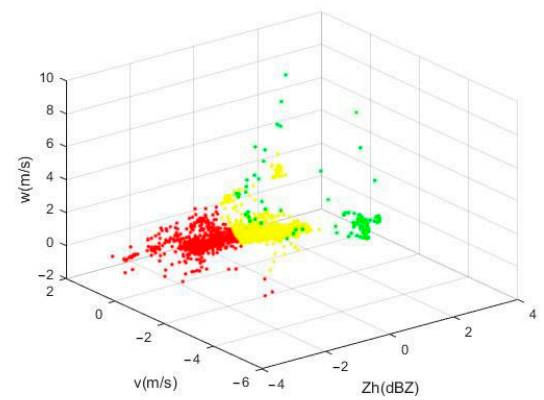
4. Qualitative Classifications of Radar Data

As can be seen from Table 1, for the samples with radar data, only 18.81% of them had icing risk, while 81.19% had no icing risk, indicating that those with and without ice accumulation should be separated as far as possible to reduce the false alarm rate. Millimeter wave has a better ability of penetrating clouds, but Lidar detection is easily affected by clouds and aerosols and may cause errors [39–41]. In the discussion in Section 3, it was found that the velocity standard deviation had little correlation and contribution to the icing index. Meanwhile, in order to increase the number of training samples, millimeter-wave radar data were used in the discussion of the following paper. In this part, the method of cluster analysis adopted was based on the three variables of Z_h , v , and w . A total of 1510 daily 0 h data were collected by the Galileo cloud radar in November 2003. First, these 1510 groups of variables were standardized. As shown in Figure 2, cluster analysis was conducted on the standardized variables, which were grouped into two to seven categories, respectively.

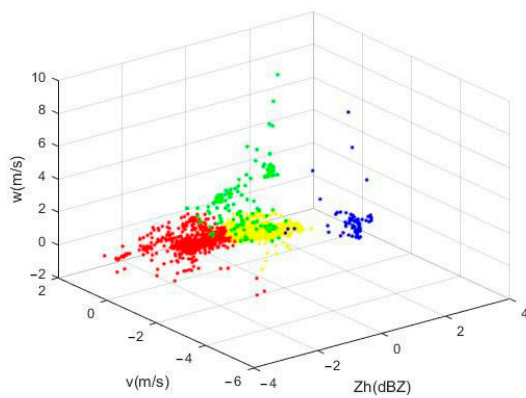
After many experiments, it was proven that the algorithm was convergent, regardless of how the initial clustering center was selected. The final clustering center was the same. For the convenience of illustration, we arranged the clustering centers in the order as shown in Table 5, and named them from category A to category G.



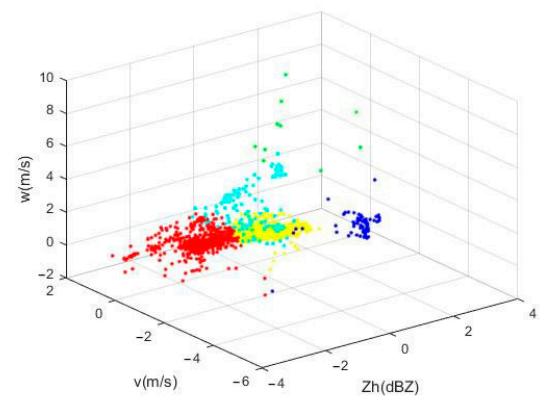
(a) 2 categories



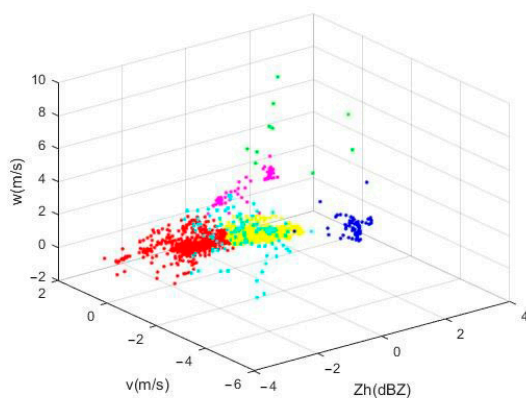
(b) 3 categories



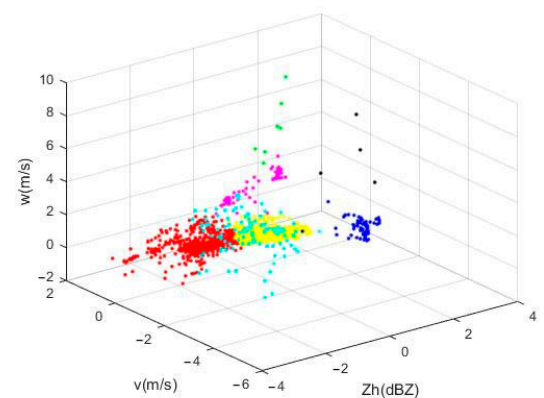
(c) 4 categories



(d) 5 categories



(e) 6 categories



(f) 7 categories

Figure 2. Radar data clustering analysis. From (a–f), the number of clusters increases from two to seven, respectively. In (a), red is category A and green is category B. After the number of categories is increased, the color and category number remain unchanged, and the newly added color is a new category (category C to category G).

Table 5. Cluster analysis results.

Class Number	Clustering Centroid Coordinates		
2	A: −0.0641, 0.1995, −0.2051	B: 0.8320, −2.5901, 2.6630	
3	A: −0.9357, 0.5370, −0.4123	B: 0.8336, −2.6728, 2.6869	C: 0.6812, −0.0907, −0.0173
4	A: −0.9288, 0.5410, −0.4338	B: 0.3101, −0.0073, 1.8097	C: 0.7175, −0.1236, −0.2475
	D: 1.0626, −3.3536, 2.4532		
5	A: −0.9272, 0.5579, −0.4562	B: 1.0696, −0.7167, 6.1073	C: 0.7279, −0.1161, −0.2576
	D: 1.0190, −3.3857, 2.2587	E: 0.1607, −0.0542, 1.4754	
6	A: −0.9137, 0.5984, −0.4645	B: 1.0696, −0.7167, 6.1073	C: 0.7347, −0.0896, −0.2614
	D: 1.0795, −3.4246, 2.3176	E: −0.5444, −0.9664, 1.0476	F: 0.9914, 0.8563, 1.7943
7	A: −0.9137, 0.5984, −0.4645	B: 0.9658, 0.0421, 5.9582	C: 0.7347, −0.0896, −0.2614
	D: 1.0688, −3.4310, 2.2907	E: −0.5444, −0.9664, 1.0476	F: 0.9914, 0.8563, 1.7943
	G: 1.4486, −2.6048, 5.9101		

The classification results were compared with the icing index corresponding to the sample to discuss the proportion of the samples with icing risk in each category. The statistical results are shown in Table 6.

Table 6. Risk proportion of each type of sample icing.

Class Number	Ratio of Icing Risk (Number of Samples with Icing Risk/Total Number of Samples in This Category)						
	Class A	Class B	Class C	Class D	Class E	Class F	Class G
2	284/1402	0/108					
3	65/646	0/104	219/760				
4	67/641	1/136	216/654	0/79			
5	67/629	0/10	215/646	0/77	2/148		
6	61/617	0/10	215/641	0/74	8/107	0/61	
7	61/617	0/7	215/641	0/73	8/107	0/61	0/4

It can be seen from Table 6 that ice accumulation information was mainly contained in Class A and Class C, but it was difficult to extract obvious features, and impossible to show ice accumulation for samples that deviated significantly from the center value. According to Table 7, when the number of classifications increased from two to three, the proportion of ice accumulation risk of newly generated Class C reached 28.82%; when the number of classifications increased to four, the proportion of ice accumulation risk of Class C was further increased to 33.03%, but there was already one missing sample in Class B. With the further increase in the number of classifications, overfitting began to appear. Further classifications only improved the classification of noise and increased the possibility of omission. To sum up, it is more appropriate to take the classification number of clustering samples as three or four categories.

Table 7. Statistics of judgment results of icing accretion.

	Neural Network Structure 1	Neural Network Structure 2
COR	49.80%	76.52%
WRO	50.20%	23.48%
FOH	37.06%	56.88%
FAR	62.94%	43.12%
DFR	0%	7.97%

5. Quantitative Judgment of Icing Index

According to the previous discussion, linear regression has good significance, but cannot judge whether there is ice accumulation. The clustering algorithm can improve the ice accumulation proportion, but the effect is not obvious. The difficulty of this problem lies in the following two points: first, it is necessary to distinguish whether there is an ice accumulation risk at this point, and it is necessary to further calculate the icing index only for points with ice accumulation risk; second, ice accumulation is a complex nonlinear problem, and the linear equation given by linear regression is difficult to fully reflect the nonlinear relationship between the physical quantities. A neural network approach has the advantage of solving these two difficulties. Therefore, two kinds of neural network structures were designed, as shown in Figure 3.

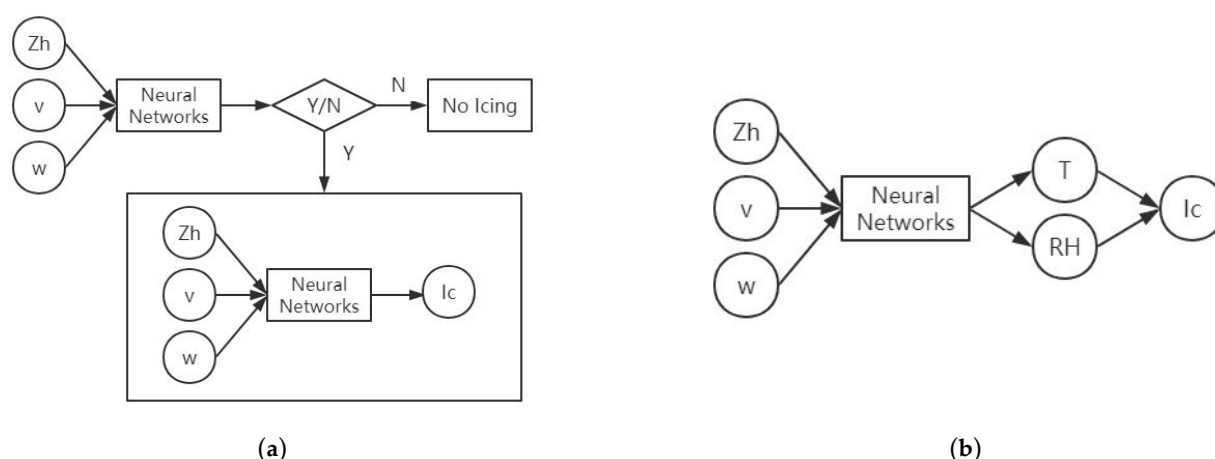


Figure 3. Calculation of icing index based on neural network. Structure 1 takes the Z_h , v , and w as input to judge whether there is icing. For points with icing, another network is used to calculate the icing index, and for points without icing, it is directly output no-icing. Structure 2 takes the Z_h , v , and w as input, takes temperature and relative humidity as output, and calculates the icing index according to Equation (1). (a) Neural Network Structure 1; (b) Neural Network Structure 2.

Here, the training set was 0 h data in November 2003, and the test set was 0 h data in November 2008. The results were calculated respectively according to the linear model and two kinds of neural network structures. All neural networks adopted the BP (back propagation) neural network, which is divided into three layers. Except for the input layer and output layer shown in Figure 3, the hidden layer had 30 neurons. If the network calculation times reaches 10,000 times, no iteration will be performed, and other parameters are default values. The statistics and analysis of the results include the following two parts. First, the statistical algorithm can judge whether there is ice accumulation or not, aiming at the grid point of data detected by millimeter wave radar. We defined the following four kinds of results as: A is the “observation that may be ice, may actually think ice”; B is the “observation that can’t be ice, may actually think ice”; C is the “observation that may be icing, actually thought impossible to ice”; and D is the “observation that impossible to ice, actually thought impossible to ice”. The second is the quantitative error of the Icing Index.

The following statistics are defined [42]:

Correct rate

$$\text{COR} = \frac{A + D}{A + B + C + D} \times 100\% \quad (11)$$

Error rate

$$\text{WRO} = \frac{B + C}{A + B + C + D} \times 100\% \quad (12)$$

Hit rate

$$\text{FOH} = \frac{A}{A + C} \times 100\% \quad (13)$$

Empty rate

$$FAR = \frac{C}{A + C} \times 100\% \quad (14)$$

Non-response rate

$$DFR = \frac{B}{B + D} \times 100\% \quad (15)$$

The results of the test are presented in Table 7 and Figure 4.

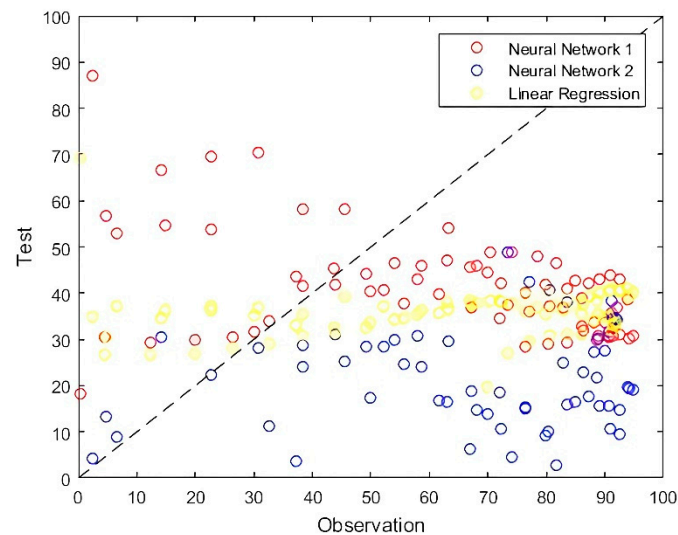


Figure 4. Test errors.

As can be seen from Table 7 and Figure 4, the two networks had similar results. From the judgment of ice accumulation, both could achieve almost or more than 50% of the judgment accuracy, the error was mainly an empty report, and the possibility of missing reports was controlled to a very low level. Linear regression and two kinds of neural networks had poor effects on the quantitative determination of index, and it is difficult to provide an accurate icing index. We provide a reasonable explanation of this problem in the next section.

6. Discussions

6.1. The Correspondence between Radar Data and Sounding Data

According to Table 1, 5.888% of the data lacked correspondence, and this phenomenon was more obvious in Figure 5.

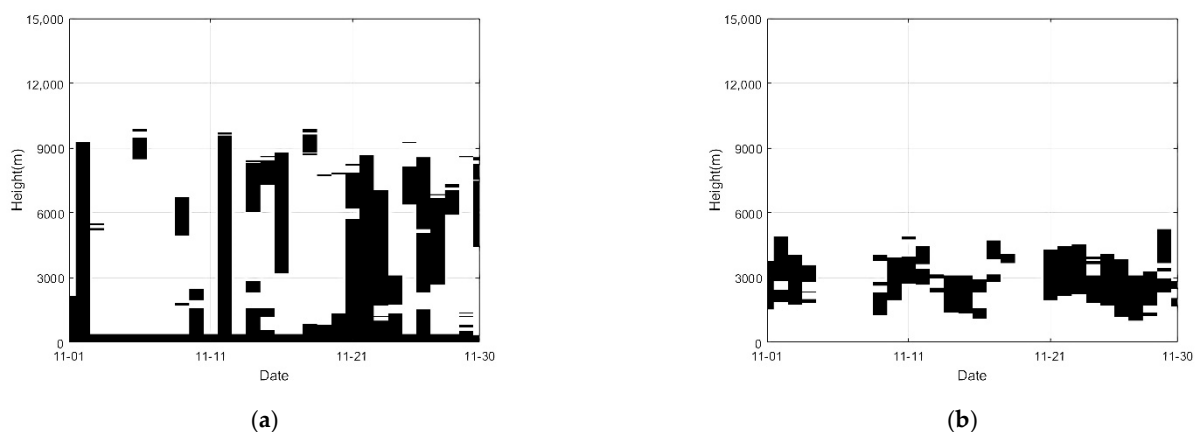


Figure 5. Comparison of millimeter-wave radar data and radiosonde Data in November 2003. (a) Galileo Radar; (b) Sounding.

As shown in Figure 5, due to the mismatch between radar data and sounding data, some data lacked correspondence. There are two reasons for this mismatch. First, there is a certain distance between Chilbolton and Herstmonceux, so there is a certain difference in the atmospheric environment between the two places, resulting in the mismatch. Second, the millimeter-wave radar, Lidar, and radiosonde used in this paper had some differences in the detection principles. The radar was set up on the ground or on the top of the building to carry out measurement and inversion by means of remote sensing. Restricted by the band they use, the objects of detection are different. Millimeter-wave radars mostly detect cloud particles [39], while Lidars are susceptible to aerosols near the ground [40,41]. Most radiosondes are set up on balloons or rockets for direct detection [43,44].

6.2. Cause Analysis of Test Results

The changes in various weather processes and the overall climate environment may bring significant influences in a short time, which is reflected in the rapid changes in the microphysical characteristics of liquid water content, particle size, rainfall rate, and so on. In this study, due to the limitation of the data volume of CFARR, data in the same month of two years were selected for training and testing, but the temperature and humidity environment of these two months were different to some extent.

As can be seen from Figure 6, compared with the average temperature and average relative humidity in November 2003 and November 2008, the temperature and relative humidity in 2008 were significantly lower than those in 2003 at a height below 5 km. The temperature was about 2–3 °C lower, and the relative humidity was about 5–10% lower. Combined with Figure 5, the icing index also changed significantly within a month. The icing threat in November 2003 was relatively more stable, and was strongest in late November 2003. In the early part of November 2008, when the air humidity was higher, the icing index was higher, so the corresponding liquid water content may increase, while the humidity was significantly lower in the late part of November 2008, and the icing index was smaller. The millimeter-wave radar also did not detect the corresponding clouds or liquid water. In addition, studies have shown that, in some cases, relative humidity cannot completely represent the threat possibility of icing, and it is currently difficult to accurately conduct numerical weather prediction on the ice [45,46], and the relative humidity is defined as vapor pressure and the ratio of saturated vapor pressure, so cannot be used to directly determine the air liquid water content, particle size, and other factors. Therefore, the inaccurate judgment may be caused by this factor.

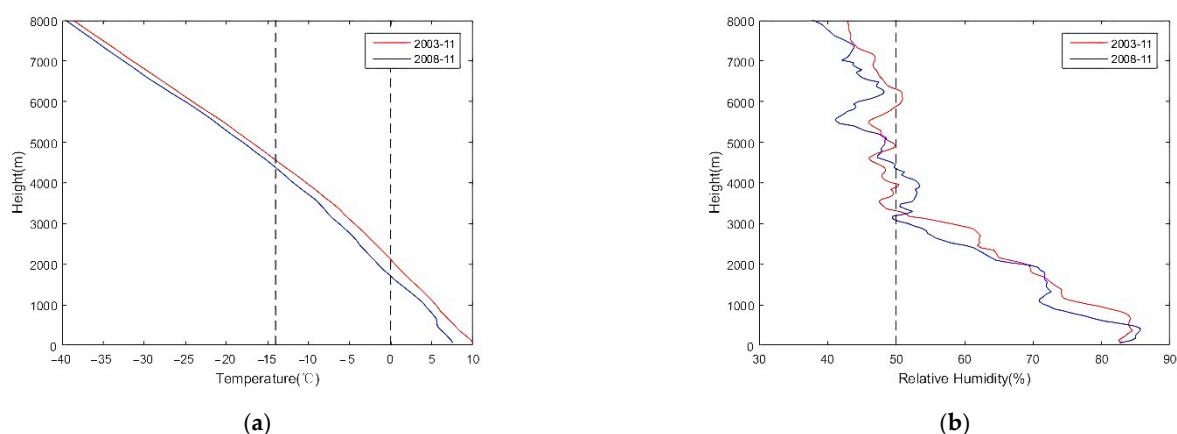


Figure 6. Profiles of average temperature and relative humidity in November 2003 and November 2008. (a) Monthly average temperature profile; (b) Monthly average RH profile. Considering the definition of Icing Index, vertical dash lines are added to illustrate the range of possible icing.

7. Conclusions

Based on CFARR millimeter-wave radar and Lidar data and the basis of the ICAO icing index using a variety of machine learning and statistical methods, from the perspective of both qualitative and quantitative, gradually thorough methods of using radar data are given to judge expression of aircraft icing conditions and methods. The training set and testing set confirmed that PCA, cluster analysis, and neural network methods had certain significance, respectively, and some of the problems and possible improvements for discussion and analysis. This current study is a first step for the development of a diagnosis tool of an aircraft icing environment.

- (1) Combined with the data of Lidar and millimeter-wave radar, the principal component analysis method was used to improve the correction determination coefficient to 0.7240, and the noise in the data was effectively eliminated.
- (2) Clustering analysis can increase the proportion of ice accumulation samples from 18.81% to 33.03%. If the classification number continues to increase, there will be overfitting, so it is difficult to further improve this proportion. However, the samples that significantly deviate from the central value can be considered as impossible for ice accumulation and excluded.
- (3) Two kinds of neural networks are constructed, which have similar performance on the judgment results of the test set, and can reach more than a 50% accuracy rate. The error is mainly shown as a false report, and the omission rate is very low, but it is difficult to calculate the ice accumulation index quantitatively.
- (4) Possible reasons for inaccurate quantitative judgment include inconsistency between the location of the radar station and the sounding station, a great difference between the samples of the training set and the test set, and the ice accumulation index cannot fully represent the ice accumulation environment, etc.

This study can be further improved from the following points: (1) select more matching data from the same site for training; (2) introduce more new variables that may improve the significance and accuracy of the judgment; (3) select indexes that are more objective and accurate for the expression of icing threat to carry out the above experiment; and (4) modify the algorithm to consider the physical properties of the model.

Author Contributions: B.X. conceived and designed the experiments, performed the experiments, analyzed the data, and wrote the paper; J.W., J.C. (Jiahua Cai), Y.W. and J.C. (Jiang Chen) helped in the discussion and revision; J.W. reviewed the manuscript. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Natural Science Foundation of China (41905026); the 63rd Batch of China Postdoctoral Science Foundation in General (2018M631554); the Open Fund by Key Laboratory of Aerosol-Cloud-Precipitation of CMA-NUIST (KDW1703); and the Key Laboratory of Middle Atmosphere and Global Environment Observation (LAGEO-2019-05).

Data Availability Statement: Not applicable.

Acknowledgments: The authors would like to acknowledge the CFARR and the UW for providing the observation data and precise products. Thanks to Jiang Qinglin and Xu Junhui for their help in the revision process.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Cao, Y.; Tan, W.; Wu, Z. Aircraft icing: An ongoing threat to aviation safety. *Aerosp. Sci. Technol.* **2018**, *75*, 353–385. [[CrossRef](#)]
2. NTSB. *In-Flight Icing Encounter and Loss of Control Simmons Airlines, D.B.A. American Eagle Flight 4184 Avions De Transport Regional (ATR) Model 72–212, N401AM, Roselawn, Indiana 31 October 1994*; National Transportation Safety Board: Washington, DC, USA, 1996; pp. 44–61.
3. BEA. *Final Report On the Accident on 1 June 2009 to the Airbus A330–203 Registered F-GZCP Operated by Air France Flight AF 447 Rio de Janeiro—Paris*; French Civil Aviation Safety Investigation Authority: Paris, France, 2012; pp. 76–77.

4. Cole, J.; Sand, W. Statistical study of aircraft icing accidents. In Proceedings of the 29th Aerospace Science Meeting, Reno, NV, USA, 7–10 January 1991; American Institute of Aeronautics: Washington, DC, USA, 1991. [\[CrossRef\]](#)
5. Zilio, C.; Patricelli, L. Aircraft anti-ice system: Evaluation of system performance with a new time dependent mathematical model. *Appl. Therm. Eng.* **2014**, *63*, 40–51. [\[CrossRef\]](#)
6. Grzesik, N.; Kowalik, K. *Fuzzy Controller for Aircraft Anti-Icing System—Initial Design and Analysis*. Solid State Phenom; Trans Tech Publications Ltd.: Stafa-Zurich, Switzerland, 2016; Volume 251, pp. 218–223. [\[CrossRef\]](#)
7. Su, Q.; Chang, S.; Zhao, Y.; Zheng, H. A Review of Loop Heat Pipes for Aircraft Anti-Icing Applications. *Appl. Therm. Eng.* **2017**, *130*, 528–540. [\[CrossRef\]](#)
8. Thompson, G.; Brintjes, R.T.; Brown, B.G. Intercomparison of in-flight icing algorithms: Part I: WISP94 realtime icing prediction and evaluation program. *Weather Forecast* **1997**, *12*, 878–889. [\[CrossRef\]](#)
9. Kelsch, M.; Wharton, L. Comparing PIREPs with NAWAU Turbulence and Icing Forecast: Issues and Results. *Weather Forecast* **1996**, *11*, 385–390. [\[CrossRef\]](#)
10. Rotondo, D.; Cristofaro, A.; Johansen, T.A. Diagnosis of Icing and Actuator Faults in UAVs Using LPV Unknown Input Observers. *J. Intell. Robot. Syst.* **2017**, *91*, 651–665. [\[CrossRef\]](#)
11. Carriere, M.; Alquier, S.; Bot, L. Statistical Verification of Forecast Icing Risk Indices. *Meteor. Appl.* **1997**, *4*, 115–130. [\[CrossRef\]](#)
12. Gosset, M.; Sauvageot, H. A Dual-Wavelength Radar Method for Ice-Water Characterization in Mixed-Phase Clouds. *J. Atmos. Ocean. Technol.* **1992**, *9*, 538–547. [\[CrossRef\]](#)
13. Vivekanandan, J.; Zhang, G.; Politovich, M.K. An Assessment of Droplet Size and Liquid Water Content Derived from Dual-Wavelength Radar Measurements to the Application of Aircraft Icing Detection. *J. Atmos. Ocean. Technol.* **2001**, *18*, 1787–1798. [\[CrossRef\]](#)
14. Gaussiat, N.; Sauvageot, H.; Illingworth, A.J. Cloud Liquid Water and Ice Content Retrieval by Multiwavelength Radar. *J. Atmos. Ocean. Technol.* **2003**, *20*, 1264–1275. [\[CrossRef\]](#)
15. Rauber, R.; Tokay, A. An Explanation for the Existence of Supercooled Water at the Top of Cold Clouds. *J. Atmos. Sci.* **1991**, *48*, 1005–1023. [\[CrossRef\]](#)
16. Ellrod, G.P.; Bailey, A. Assessment of Aircraft Icing Potential and Maximum Icing Altitude from Geostationary Meteorological Satellite Data. *Weather Forecast* **2007**, *22*, 160–174. [\[CrossRef\]](#)
17. Javier, D.F.; Lara, Q.H.; Pedro, B. Mountain Waves Analysis in the Vicinity of the Madrid-Barajas Airport Using the WRF Model. *Adv. Meteorol.* **2020**. [\[CrossRef\]](#)
18. Merino, A.; Ortega, E.G.; González, S.F. Aircraft Icing: In-Cloud Measurements and Sensitivity to Physical Parameterizations. *Geophys. Res. Lett.* **2019**, *46*, 11559–11567. [\[CrossRef\]](#)
19. Bolgiani, P.; González, S.F.; Martín, M.L. Analysis and numerical simulation of an aircraft icing episode near Adolfo Suárez Madrid-Barajas International Airport. *Atmos. Res.* **2017**, *200*, 60–69. [\[CrossRef\]](#)
20. Pereira, M.B. Comparison of in-flight aircraft icing algorithms based on ECMWF forecasts. *Meteorol. Appl.* **2016**, *22*, 705–715. [\[CrossRef\]](#)
21. Federal Aviation Administration. Part 25—Airworthiness Standards: Transport Category Airplanes. Available online: <https://www.ecfr.gov/cgi-bin/ECFR?page=browse> (accessed on 1 September 2020).
22. Matthew, D.J.; Kamran, R. Using artificial neural networks and self-organizing maps for detection of airframe icing. *J. Aircr.* **2001**, *38*, 224–230. [\[CrossRef\]](#)
23. Ogretim, E.; Huebsch, W.; Shinn, A. Aircraft ice accretion prediction based on neural networks. *J. Aircr.* **2006**, *43*, 233–240. [\[CrossRef\]](#)
24. Dong, Y. An application of Deep Neural Networks to the in-flight parameter identification for detection and characterization of aircraft icing. *Aerosp. Sci. Technol.* **2018**, *77*, 34–49. [\[CrossRef\]](#)
25. Dong, Y. Implementing Deep Learning for Comprehensive Aircraft Icing and Actuator/Sensor Fault Detection/Identification. *Eng. Appl. Artif. Intell.* **2019**, *83*, 28–44. [\[CrossRef\]](#)
26. Li, S.; Qin, J.; He, M. Fast Evaluation of Aircraft Icing Severity Using Machine Learning Based on XGBoost. *Aerospace* **2020**, *7*, 36. [\[CrossRef\]](#)
27. Kolbakir, C.; Hu, H.; Liu, Y. An experimental study on different plasma actuator layouts for aircraft icing mitigation. *Aerosp. Sci. Technol.* **2020**, *107*, 106325. [\[CrossRef\]](#)
28. Zhang, F.; Huang, Z.; Yao, H. Icing severity forecast algorithm under both subjective and objective parameters uncertainties. *Atmos. Environ.* **2016**, *128*, 263–267. [\[CrossRef\]](#)
29. Zhou, C.; Li, Y.; Zheng, W. Safety Analysis for Icing Aircraft during Landing Phase Based on Reachability Analysis. *Math. Probl. Eng.* **2018**, *14*. [\[CrossRef\]](#)
30. Cao, Y.; Tan, W.; Su, Y. The Effects of Icing on Aircraft Longitudinal Aerodynamic Characteristics. *Mathematics* **2020**, *8*, 1171. [\[CrossRef\]](#)
31. Liu, T.; Cai, J.; Qu, K. In-flight icing simulation for two-dimensional configurations. *Int. J. Mod. Phys. B* **2020**, *34*, 2040068. [\[CrossRef\]](#)
32. Wang, J.; Lu, Z.; Shi, Y. Aircraft icing safety analysis method in presence of fuzzy inputs and fuzzy state. *Aerosp. Sci. Technol.* **2018**, *82*, 172–184. [\[CrossRef\]](#)

-
33. Wang, J.; Xie, B.; Cai, J. The Distribution of Aircraft Icing Accretion in China—Preliminary Study. *Atmosphere* **2020**, *11*, 876. [[CrossRef](#)]
 34. CEDA Achieve. Dataset Collection: Chilbolton Facility for Atmospheric and Radio Research (CFARR): Surface, Radar and Lidar Measurements (1998-Present). Available online: <https://catalogue.ceda.ac.uk/uuid/7cbc3fc19bfa037a48ba4cba4b93544d> (accessed on 1 September 2020).
 35. Liu, F.L.; Sun, L.T.; Li, S.J. Study on Methods of Aircraft Icing Diagnosis and Forecast. *Meteorol. Environ. Sci.* **2011**, *34*, 26–30. [[CrossRef](#)]
 36. Scholkopf, B.; Smola, A.; Muller, K. Nonlinear Component Analysis as a Kernel Eigenvalue Problem. *Neural. Comput.* **1998**, *10*, 1299–1319. [[CrossRef](#)]
 37. Aloise, D.; Deshpande, A.; Hansen, P. NP-hardness of Euclidean Sum-of-squares Clustering. *Mach. Learn.* **2009**, *75*, 245–248. [[CrossRef](#)]
 38. Kohonen, T. An Introduction to Neural Computing. *Neural Networks*. **1988**, *1*, 3–16. [[CrossRef](#)]
 39. White, A.B.; Fairall, C.W.; Frisch, A.S. Recent Radar Measurements of Turbulence and Microphysical Parameters in Marine Boundary Layer Clouds. *Atmos. Res.* **1996**, *40*, 177–221. [[CrossRef](#)]
 40. Ghazal, F.; Robert, S.; Joseph, D.M. Classification of lidar measurements using supervised and unsupervised machine learning methods. *Atmos. Meas. Tech.* **2021**, *14*, 391–402. [[CrossRef](#)]
 41. Romain, C.; Matthew, B. Aerosol light extinction and backscattering: A review with a lidar perspective. *J. Quant. Spectrosc. Radiat. Transf.* **2020**, *262*, 107492. [[CrossRef](#)]
 42. Brown, B.G.; Thompson, G.; Bruintjes, R.T. Intercomparison of in-flight icing algorithms: Part II: Statistical verification results. *Weather Forecast* **1997**, *12*, 890–914. [[CrossRef](#)]
 43. Thorne, P.W.; Lanzante, J.; Peterson, T. Tropospheric temperature Trends: History of an Ongoing Controversy. *Wires. Clim. Chang.* **2011**, *2*, 66–88. [[CrossRef](#)]
 44. Chernykh, I.V.; Aldukhov, O.A. Temperature and Humidity Trends in the Lower Atmospheric 2-km Layer over the Russian Arctic According to Radiosonde Data. *Russ. Meteorol. Hydrol.* **2020**, *45*, 615–622. [[CrossRef](#)]
 45. Gultepe, I.; Agelin-Chaab, M.; Komar, J. A Meteorological Supersite for Aviation and Cold Weather Applications. *Pure Appl. Geophys.* **2018**, *176*, 1977–2015. [[CrossRef](#)]
 46. Gultepe, I.; Sharman, R.; Williams, P.D. A Review of High Impact Weather for Aviation Meteorology. *Pure Appl. Geophys.* **2019**, *176*, 1869–1921. [[CrossRef](#)]