

11.5 ICING RISK LEVEL DERIVED FROM GLOBAL/REGIONAL ICING POTENTIAL AND OCCURRENCE PROBABILITY

Yin-lam NG^{1,*}; Kin-chung AU-YEUNG¹, CHONG Pui Lam, Ivan², Monique Zoe WAH³, Wang CHOW¹

¹ Hong Kong Observatory, Hong Kong, China

² Chinese University of Hong Kong

³ Hong Kong University of Science and Technology

1. INTRODUCTION

Aircraft icing is the accretion of supercooled liquid onto an airplane during flight. Meteorologists, aerospace engineers, and pilots need and want information about airframe icing because it can adversely affect the flight performance of an aircraft. Icing can increase drag, decrease lift, and cause control problems. The added weight of the accreted ice is generally a factor only for light aircraft.

Certified aircraft are commonly equipped with devices that either prevent ice from adhering to the airframe or remove it once it has adhered. Such anti-icing or de-icing equipment may be deployed manually or through an automatic system triggered by an icing detection probe.

Icing is usually classified into four severity categories: trace, light, moderate, and severe. Severity reported depends on the state of the icing environment, the aircraft's response, and the pilot's assessment of the response.

The Forecast Icing Potential (FIP) algorithm was developed at the National Center for Atmospheric Research under the Federal Aviation Administration's Aviation Weather Research Program. It uses 20-km resolution Rapid Update Cycle (RUC) model output to determine the potential for in-flight aircraft icing conditions and supercooled large drops (SLD; defined as droplets with diameters greater than $50\mu m$). FIP combines the model input data using fuzzy logic membership functions and a decision tree to estimate the potential for icing. The membership functions are based on cloud physics principles, forecasting and research experience, and comparison of fields to icing pilot reports (PIREPs). They map data onto a 0-1 scale, which represents the expected likelihood of icing, given the value from that field (McDonough et al. 2004).

2. DATA

2.1. Observations

Icing reports by pilots are treated as the ground truth of areas with icing encounters. Locations with no icing encountered would also be needed for verification or constructing models to predict icing occurrence

probability, so that they could learn how to distinguish potential icing areas. However, areas with no icing reports do not guarantee the locations are not icing favourable, it could simply because no flights happened to have passed through the areas. In this study, locations with Civil Aviation Administration of China pilot reports (CAAC PIREPs) indicating light turbulence and null icing are considered as areas with no icing threats. The amount of bias introduced due to samples being only taken along common flight paths is considered to be negligible for developing meteorological products, which are aimed to aid aviation operations.

1) CIVIL AVIATION ADMINISTRATION OF CHINA PILOT REPORTS (CAAC PIREPs)

CAAC PIREPs during the period January 2019 to October 2023 have been collected for the study. They are pilot reports provided by the Civil Aviation Administration of China with turbulence or icing events location, height, time and intensity. Intensity is reported on a scale of 1–5 and 0 indicated no encounter. All icing PIREPs (i.e. intensity being 1–5) are used for this study, while turbulence report (with no icing encountered) with intensity 1–2 (i.e. light turbulence) are randomly selected as null icing reports. Reports with light turbulence and icing would be considered as icing reports. Turbulence reports are randomly selected such that the number of turbulence with no icing encountered and icing events is roughly 1:1 for this study.

2) SPECIAL AIR REPORTS (ARS)

To enlarge the observation data set, ARS on icing events during January 2021 to June 2023 are also collected from the World Meteorological Organization (WMO) Information System (WIS) at the Tokyo Global Information System Centre (GISC) for this study. Only "light/ moderate/ severe" would be reported by icing ARS. Based on the proportion of intensity reported, icing ARS are considered to be equivalent to CAAC PIREPs intensities: light = 1–2, moderate = 3, severe = 4–5.

2.2. Numerical Weather Prediction (NWP) model

1) GLOBAL NWP MODEL

The high resolution model by the European Centre for Medium-Range Weather Forecasts (ECMWF) has been

*Corresponding author address: Danice YL Ng, Hong Kong Observatory, 134A Nathan Road, Kowloon, Hong Kong, China.; E-mail: ylng@hko.gov.hk.

used to construct forecast icing potential. Parameters including temperature (T), relative humidity (RH), and vertical velocity (w - negative indicates updraft) for calculating FIP are available at pressure levels [950 hPa, 925 hPa, 900 hPa, 850 hPa, 800 hPa, 700 hPa, 600 hPa, 500 hPa, 400 hPa, 300 hPa, 250 hPa, 200 hPa].

2) REGIONAL NWP MODEL

The Hong Kong Observatory (HKO) has been running the HKO-WRF (Hon 2020), a large-area mesoscale NWP model covering the Asian Pacific region and supporting regional aviation applications, which has been demonstrated to possess positive skill in clear air turbulence (Hon et al. 2020) and ice crystal icing (Ng et al. 2017).

3. FORECAST ICING POTENTIAL (FIP)

With reference to McDonough et al. (2004) while adjusting slightly to fit the meteorological conditions for icing in this region, clouds are defined as $RH \geq 80\%$. If a layer in an air column has thickness ≥ 75 hPa and $RH < 50\%$, it is considered as a dry slot between cloud layers in that column of atmosphere. Cloud top temperature (CTT in K) is defined as the temperature of the nearest cloud top layer. An illustration is shown in Figure 1.

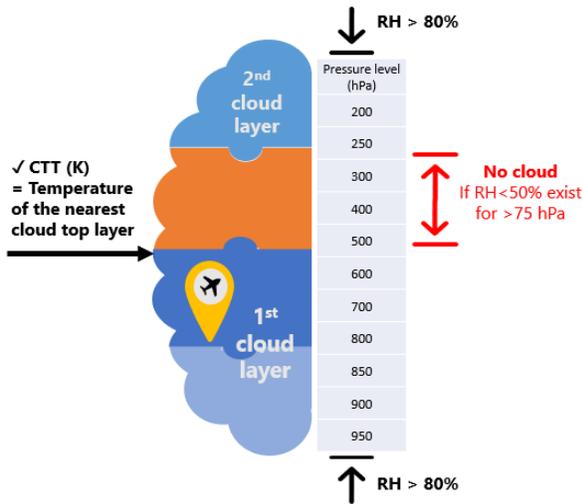


FIG. 1. Defining cloud layer(s) for each point of interest (a column of atmosphere).

Using membership functions to map the T, CTT and RH to T_{map} , CTT_{map} and RH_{map} respectively, FIP (FIP_{final}) is calculated as follow:

$$FIP_{initial} = T_{map} \times CTT_{map} \times RH_{map}$$

$$\text{If } w \geq 0, \quad FIP_{final} = FIP_{initial} + w_{map} (1 - FIP_{initial})$$

$$\text{If } w < 0, \quad FIP_{final} = FIP_{initial} - w_{map} (FIP_{initial})$$

3.1. FIP derived from global NWP model

Based on the locations of the PIREP, nearest grid points in ECMWF model data were identified as points of interest. The distributions of each ECMWF model elements (T, RH, CTT, w) were plotted in Figure 2. Note that there were only 4.5% points of interest were cloud free with the above-described definition of cloud layers. Based on McDonough et al. (2004), NCAR membership functions were given in red. It is observed that CTT and w histograms were not as well represented by NCAR membership functions in this region. All four membership functions for calculating FIP from ECMWF parameters were adjusted based on CAAC PIREP and shown in dandelion.

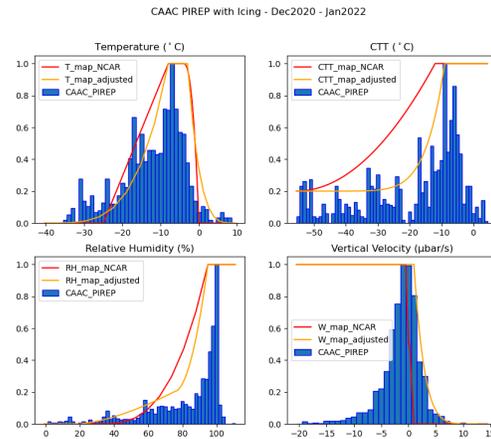


FIG. 2. Distributions of model elements extracted from CAAC PIREP points of interest with lines showing NCAR membership functions (red) and adjusted membership functions for ECMWF (dandelion).

3.2. FIP derived from regional NWP model

Similarly, nearest grid points associated with PIREP locations in HKO-WRF model data were identified as points of interest. The distributions of each HKO-WRF model elements (T, RH, CTT, w) were plotted in Figure 3. All four membership functions for calculating FIP from HKO-WRF parameters were adjusted based on CAAC PIREP and shown in pink.

4. ICING OCCURRENCE PROBABILITY

Ng (2023) compared three types of machines learning models to forecast icing occurrence or icing severity. Icing occurrence probability predicted by eXtreme Gradient Boosting (XGBoost) was found to be the most accurate prediction. XGBoost Python package was used to construct XGBoost model for icing occurrence (no icing vs icing) probability. XGBoost is a scalable

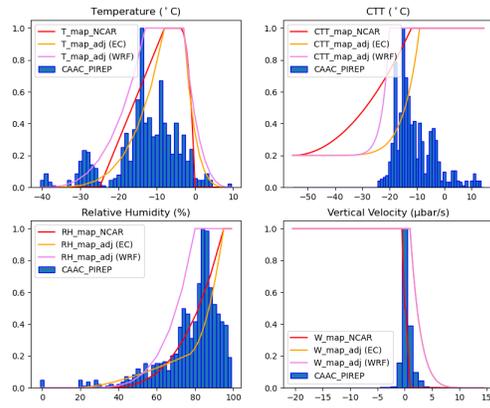


FIG. 3. Distributions of model elements extracted from CAAC PIREP points of interest with lines showing NCAR membership functions (red), adjusted membership functions for ECMWF (dandelion) and HKO-WRF (pink).

end-to-end tree boosting system developed by Chen and Guestrin (2016) which used different regularisation penalties to avoid overfitting. Chen and Guestrin (2016) also showed their system runs more than ten times faster than existing popular solutions on a single machine.

5. ICING RISK LEVEL

Figure 4 and Figure 5 showed boxplot for ECMWF FIP and icing occurrence probabilities against PIREP/ARS reported icing severity. The linear trends across icing severity were not statistically significant.

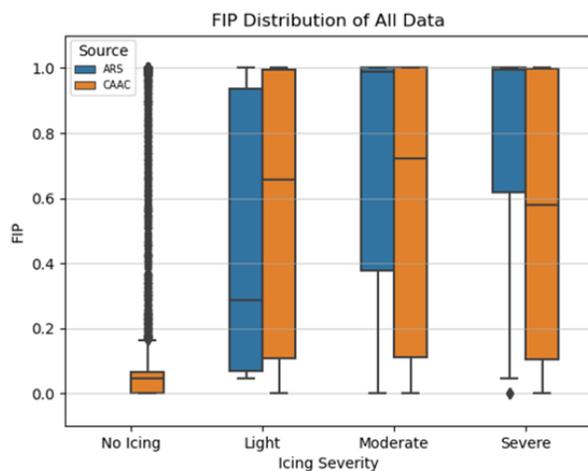


FIG. 4. ECMWF FIP plotted against PIREP/ARS reported icing severity.

FIP only considered the parameters T, CTT, RH and

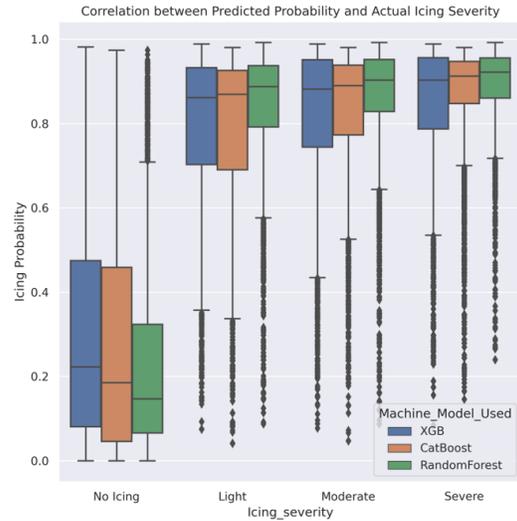


FIG. 5. Icing occurrence probabilities by three machine learning models (Ng 2023) plotted against PIREP/ARS reported icing severity.

w ; while icing occurrence probability took into account for T, RH, w , divergence, geopotential height, potential vorticity, specific humidity, U -component of wind, V -component of wind, and vorticity (relative). With meteorological conditions given by FIP and probability quantified by icing occurrence probability available, it is possible to multiply them together as icing risk level and construct an icing risk matrix (Table 1).

		Probability (Occurrence)		
		Unlikely	Probable	Likely
Potential (FIP)	High			
	Medium			
	Low			

TABLE 1. An illustration of icing risk matrix

As described in Ng (2023), two XGBoost models were fitted for ECMWF parameters based on CAAC PIREP or CAAC PIREP together with ARS. Similarly, two XGBoost models were fitted for HKO-WRF parameters. A total of four icing occurrence probabilities would be looked at in this study for constructing icing risk levels.

6. RESULTS

Several statistical analyses were done, including receiver operating characteristic curve (ROC curve), area under the ROC curve (AUC), hypothesis testings on linear trend across icing severity, and Analysis of Variance (ANOVA).

Thresholds for yes/no icing, moderate or above icing and severe icing were investigated. However, the amount of severe icing reports was too scarce for drawing meaningful conclusion. Therefore, results for yes/no icing and moderate or above icing would be presented. Based on the best thresholds, probability of detection (POD), false alarm ratio (FAR), and critical success index (CSI) were calculated for comparison.

6.1. FIP

1) GLOBAL NWP MODEL

Figures 6 and 7 showed the ROC curves for classifying yes/no icing and moderate or above icing respectively using FIP calculated from global NWP model parameters. The corresponding thresholds (and associated AUC given in brackets) are 0.29 (0.87) and 0.32 (0.77). The linear relationship between FIP derived from global NWP model and reported icing severity was not statistically significant.

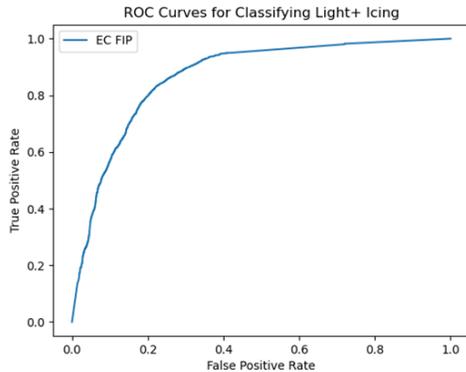


FIG. 6. ROC curve for classifying icing cases (yes/no icing) using FIP derived from global NWP model (ECMWF).

2) REGIONAL NWP MODEL

Figures 8 and 9 showed the ROC curves for classifying yes/no icing and moderate or above icing respectively using FIP calculated from regional NWP model parameters. The corresponding thresholds (AUC) are 0.30 (0.82) and 0.35 (0.72). The linear relationship between FIP derived by regional NWP model and reported icing severity was not statistically significant.

The performance of FIPs derived from global and regional NWP models is very similar as shown in Table 2 in terms of CSI (around 0.6). Regional FIP has slightly higher POD compared to global FIP.

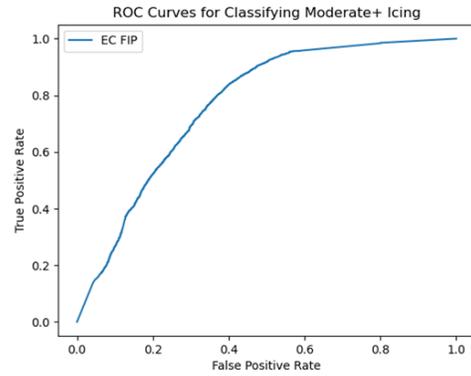


FIG. 7. ROC curve for classifying moderate or above icing cases using FIP derived from global NWP model (ECMWF).

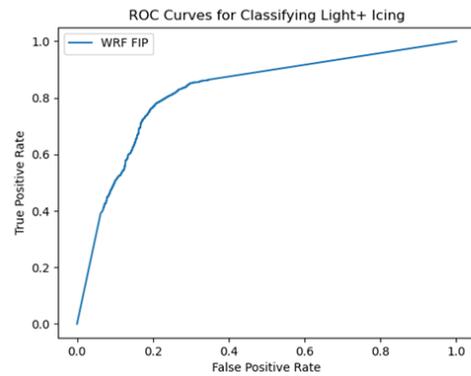


FIG. 8. ROC curve for classifying icing cases (yes/no icing) using FIP derived from regional NWP model (HKO-WRF).

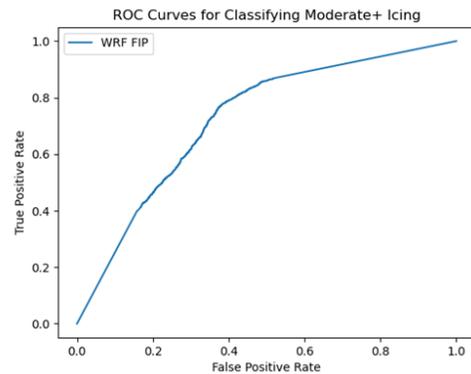


FIG. 9. ROC curve for classifying moderate or above icing cases using FIP derived from regional NWP model (HKO-WRF).

$FIP \geq 0.3$	POD	FAR	CSI
derived from global NWP model (ECMWF)	0.679	0.171	0.595
derived from regional NWP model (HKO-WRF)	0.709	0.189	0.608

TABLE 2. Skill scores for FIPs derived from global and regional NWP models.

6.2. Icing occurrence probability

1) GLOBAL NWP MODEL

Ng (2023) presented that ROC curves for FIP and icing occurrence probability derived from global NWP model were very similar. Also, the linear relationship between icing occurrence probability derived from global NWP model and reported icing severity was not statistically significant.

2) REGIONAL NWP MODEL

Four icing occurrence probability models were trained based on training dataset used:

- **caac prob**: CAAC PIREPs
- **combined prob**: CAAC PIREPs and ARS
- **caac (MOD+) prob**: CAAC PIREPs that reported moderate or above severity
- **combined (MOD+) prob**: CAAC PIREPs and ARS that reported moderate or above severity

Figures 10 and 11 showed the ROC curves for classifying yes/no icing and moderate or above icing respectively. The corresponding thresholds (AUC) for the four models to classify yes/no icing are 0.61 (0.88), 0.57 (0.92), 0.42 (0.90) and 0.63 (0.92). And that for classifying moderate or above icing cases are 0.46 (0.74), 0.61 (0.85), 0.40 (0.82) and 0.58 (0.90).

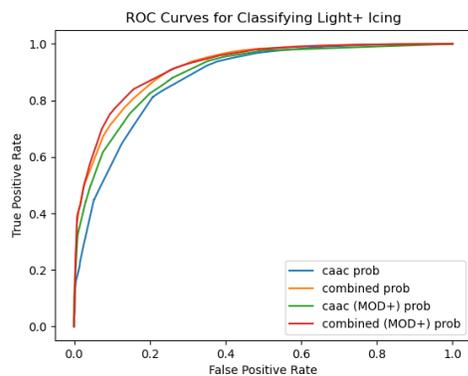


FIG. 10. ROC curve for classifying icing cases (yes/no icing) using icing occurrence probability derived from regional NWP model (HKO-WRF).

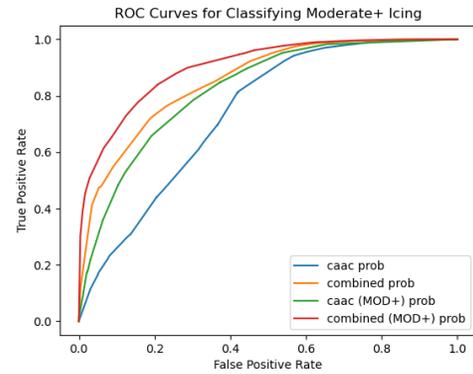


FIG. 11. ROC curve for classifying moderate or above icing cases using icing occurrence probability derived from regional NWP model (HKO-WRF).

Although Figures 10 and 11 showed AUCs for **combined (MOD+) prob** are greatest, Figure 12 showed that the median for severe cases was lower than that for moderate cases. Considering both AUC and linear trend in probability against icing severity, **caac (MOD+) prob** managed to provide the most reasonable classification thresholds for icing severity (light vs MOD+ categories).

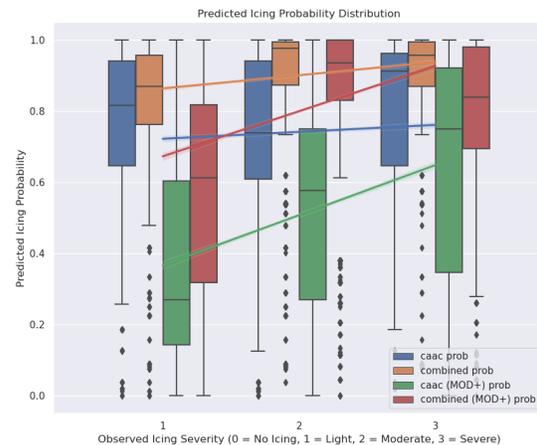


FIG. 12. Boxplots of icing occurrence probability derived from regional NWP model against observed severity.

6.3. Icing risk level

With FIP denoting meteorological condition favourable for icing and icing occurrence probability provided by XGBoost, this section explored icing risk level constructed by multiplying FIP with icing occurrence probability.

1) GLOBAL NWP MODEL

Since FIP and icing occurrence probability derived from the global NWP model did not show linear trend against icing severity, the icing risk did not show linear relationship with reported icing severity.

2) REGIONAL NWP MODEL

The corresponding thresholds (AUC) for the icing risk derived from **caac (MOD+) prob** model to classify yes/no icing is **0.27 (0.86)** and that for classifying moderate or above icing cases is **0.41 (0.78)**. The thresholds obtained are more distinguishable compared to corresponding thresholds derived from probability by **caac (MOD+) prob** model. In addition, the linear regression for icing risk against icing severity has significant positive slope being 0.283 with standard error 0.017 (Figure 13). Therefore, icing risk derived from **caac (MOD+) prob** model could potentially inform users of the impact of icing condition and icing severity (light vs MOD+ categories).

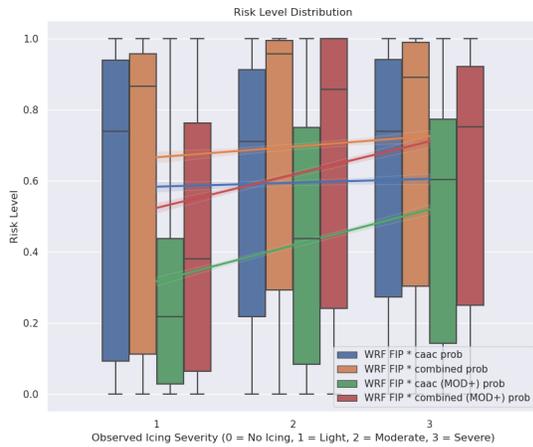


FIG. 13. Boxplots of icing risk derived from regional NWP model against observed severity.

7. DISCUSSIONS

This study showed that interest functions for constructing FIP would require adjustment based on regional observations for both global and regional NWP models. The icing occurrence probability by either global or regional NWP models had good AUC diagnostics for icing occurrence threshold, while the icing occurrence probability by regional NWP model (HKO-WRF) was positively correlated with icing severity.

Constructing icing risk via FIP and icing occurrence probability was also explored in this study. The idea is that FIP could provide the extend to which the meteorological condition favours airframe icing and icing occurrence probability was the likelihood of airframe

icing learnt by XGBoost models. Statistically, icing risk derived from regional NWP model (HKO-WRF) was more positively correlated with icing severity compared to that from global NWP model.

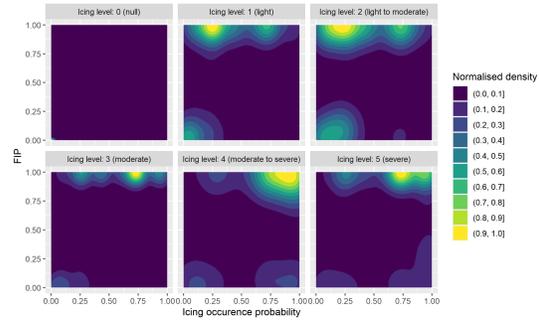


FIG. 14. For each observed icing severity level, normalised density is plotted for FIP against icing occurrence probability derived from regional model.

As shown in Table 1 and Figure 14, risk level is not simply linearly related to icing severity as there might be low to moderate risk for high FIP with unlikely occurrence. Therefore, icing risk might provide an useful overview of icing impact or summary of FIP and icing occurrence probability for specific locations, e.g. common way points for major flight routes, or small areas, such as holding areas. On the other hand, referencing both FIP and icing occurrence probability might provide a better spatial consideration for forecasting airframe icing hazard as combining FIP and icing occurrence probability into icing risk would collapse the two-dimension risk matrix information into a one-dimension information (the three colours: red, orange, and green).

Acknowledgments. Special acknowledgement goes to the Civil Aviation Administration of China (CAAC) for their provision and permission to use the pilot reports (PIREPs) for the study. Thank you all pilots who have recorded, compiled and disseminated special air reports on icing. Special thanks to the staff of the International Aviation Meteorological Collaboration Division in HKO for their valuable comments and technical support. We would also like to extent our thanks to the internship students, Mr. Lau Cheuk Yin, Ms. Ying Yui Wang, Mr. Wong Kwong Lung, and Mr. Dai Man Hei, who undertook side projects related to icing forecast with HKO. Last but not least, the authors would like to thank Mr. Chan Sai Tick, Mr. Kok Mang Hin, Marco and Ms. Leung Yan Yu, Christy for their valuable comments on the research project and the abstract.

REFERENCES

- Chen, T., and C. Guestrin, 2016: Xgboost: A scalable tree boosting system. *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining*, 785–794.
- Hon, K.-K., 2020: Tropical cyclone track prediction using a large-area wrf model at the hong kong observatory. *Tropical Cyclone Research and Review*, **9** (1), 67–74.
- Hon, K. K., C. W. Ng, and P. W. Chan, 2020: Machine learning based multi-index prediction of aviation turbulence over the asia-pacific. *Machine Learning with Applications*, **2**, 100 008.
- McDonough, F., B. C. Bernstein, M. K. Politovich, and C. A. Wolff, 2004: 3.2 the forecast icing potential (fip) algorithm.
- Ng, Y., H. Law, J. Lee, K. Hon, L. Li, and P. Li, 2017: Satellite detection and nowcasting high altitude ice crystals. *Proc. WMO Aeronautical Meteorology Scientific Conference 2017*.
- Ng, Y.-I., 2023: Development of artificial intelligence models for airframe icing forecast based on pilot reports and numerical weather prediction model parameters. *103rd AMS Annual Meeting*, AMS.