JOURNAL OF AEROSPACE INFORMATION SYSTEMS

Check for updates

Technical Notes

Data-Driven Machine Learning Model for Aircraft Icing Severity Evaluation

Sibo Li*
University of Illinois at Chicago, Chicago, Illinois 60607
Jingkun Qin†
China Literature Limited, 100105 Beijing, People's Republic of China
and
Roberto Paoli‡
University of Illinois at Chicago, Chicago, Illinois 60607

https://doi.org/10.2514/1.I010978

Downloaded by Roberto Paoli on July 14, 2021 | http://arc.aiaa.org | DOI: 10.2514/1.I010978

I. Introduction

AIRCRAFT icing represents a serious hazard to aviation that caused many tragic fatalities over the past decades [1]. The physical formation process of ice is affected by a variety of aerodynamic and environmental factors, such as flight speed, angle of attack, exposure time, liquid water content (LWC), droplet median volumetric diameter (MVD), and freestream temperature; hence, ice accretion remains an issue far from being completely resolved [2,3]. Instead of explicitly modeling the icing formation process, recent studies proposed the use of regression analysis and machine learning models to predict the aircraft icing severity based on icing data collected in experimental campaigns and/or numerical simulations. This approach is motivated by the fact that significant computing resources are required for numerical simulations to calculate the ice accretion along the wing. Data-driven methods can provide quick evaluations of critical ice features such as the ice global coverage or maximum ice thickness [4]. For example, Li et al. [4] introduced a purely data-driven approach to forecast the aircraft icing severity level. Ogretim et al. [5] developed the methodology for ice accretion prediction by incorporating the Fourier series expansion and using neural network. McCann [6] built models based on neural networks to make icing forecasts of different icing intensities. Zhan et al. [7] proposed a framework of local reduced-order modeling using machine learning algorithms to explore the in-flight icing certification envelopes.

For data-driven aircraft icing forecasting, the mapping relationship between the input flight conditions and the output aircraft icing severity features is likely to be strongly nonlinear. Machine learning models are capable of addressing strong nonlinearity with the aid of constructing a black-box input–output mapping [8]. The black-box mapping can be represented as a forest of trees that could correspond to the machine learning algorithm gradient boosting (GBoost) [9]. The extreme gradient boosting model (XGBoost) [10] is applied in the current work. XGBoost is a state-of-the-art GBoost model, designed to be more computational efficient and flexible. The XGBoost model is able to account for the complex interactions and correlations among the features and have been applied in many data-driven applications. In the previous study [4], we employed XGBoost model to explore the complex pattern between LWC, MVD, and exposure time. It was a first step toward the development of a data-driven approach to predict the aircraft icing severity. However, it ignored the aerodynamic factors such as flight speed, angle of attack, and static temperature, which have direct impact on the aircraft icing results. In this paper, we studied adapting XGBoost for aircraft icing severity evaluation based on six flight conditions (flight speed, angle of attack, exposure time, LWC, MVD, and freestream temperature) to represent real flight situation. We aim to address the following issues: predicting the icing severity level [3], area size covered by ice and maximum ice thickness, assessing the importance of the flight conditions toward the icing severity level, and understanding the effect of each flight condition. To the best of our knowledge, this paper represents the first study of applying XGBoost in aircraft icing severity evaluation in real flight condition. The multiple linear regression (MLR) [11] and ordinal logistic regression (OLR) [12] serve as the benchmark models. The models are trained and evaluated on a database of available flow data obtained from previous simulations [13,14]. To evaluate the accuracy of the predictions quantitatively, performance error analysis method containing various components is established. Applications to the two most important icing features further demonstrate that the proposed approach can provide a suitable alternative to numerical simulation methods with reasonable accuracy while saving computational time. Furthermore, by coupling with computational fluid dynamics (CFD) codes, the proposed approach can be used to estimate the degradation of the aircraft aerodynamic performance. In the hybrid system, CFD can provide a detailed simulation based on machine learning predictions.

II. Data-Driven Methods

A. Data Collection

The NACA0012 airfoil is studied in this paper. For each case study, the numerical simulations were run by applying the ice accretion modeling solver developed by the authors in the previous work [13,14] to generate the training observations and test data. Based on the theoretical modeling of aircraft icing, we consider six flight conditions: flight speed, angle of attack, exposure time, LWC, MVD, and freestream temperature. Each of these flight conditions has several levels. An example of the flight conditions with their corresponding maximum and minimum values and step size are given in Table 1. Combinations of different values for the six parameters represent different sets of icing experiments, and 1890 samples are selected to form the dataset. From the set of full generated samples, testing and training samples are partitioned randomly following the same design principle to have the same population distribution. Testing and training datasets have 567 and 1106 samples, respectively.

The icing severity level (Table 2) based on ice thickness described by Cao et al. [3] is firstly predicted. Four levels are introduced to describe the icing severity. The pilots could use the standard as a reference to assess the severity of the flight condition [6]. It is reasonable to establish the standard based on the maximum ice thickness rather than the rate of accretion because in reality the airplane flight performance will only be little affected if the time spent in severe icing state is limited. Besides the icing severity level (classification problem), the prediction models are also trained to predict the size of the area on the airfoil covered by ice and maximum ice thickness (regression problems). The larger they are, the more damage will be caused to the aerodynamic performance [3].

Received 12 February 2021; revision received 8 May 2021; accepted for publication 3 June 2021; published online 14 July 2021. Copyright © 2021 by the American Institute of Aeronautics and Astronautics, Inc. All rights reserved. All requests for copying and permission to reprint should be submitted to CCC at www.copyright.com; employ the eISSN 2327-3097 to initiate your request. See also AIAA Rights and Permissions www.aiaa.org/randp.

*Ph.D. Candidate, Department of Mechanical and Industrial Engineering.

†Data Scientist.

‡Research Assistant Professor, Department of Mechanical and Industrial Engineering; also Leadership Computing Facility, Argonne National Laboratory, Lemont, Illinois 60439. Senior Member AIAA.

Downloaded by Roberto Paoli on July 14, 2021 | http://arc.aiaa.org | DOI: 10.2514/1.I010978

Table 1 Statistics of flight conditions

Feature	V_∞, knots	T_∞, K	AOA, °	LWC, g/m^3	MVD, μm	t_{ice}, min
Maximum	250.0	265	9	1.50	50	30.0
Minimum	100.0	253	0	0.50	5	1.0
Step size	37.5	3	3	0.25	5	9.5

Table 2 Icing severity level based on icing thickness

Icing severity level	Maximum thickness, mm
Light	0.1–5.0
Moderate	5.1–15
Heavy	15.1–30
Severe	>30

B. Extreme Gradient Boosting Model

XGBoost uses an ensemble of classification and regression trees (CARTs) as the mapping to fit the training data samples. CARTs are predictive models, which explain how an outcome variable's values can be predicted based on other values [15]. It is a supervised machine learning algorithm for structured or tabular datasets on classification and regression predictive modeling problems. For the sake of brevity, the reader is referred to [10] for more penetrating insights into the XGBoost.

The XGBoost algorithm involves fitting a large number of decision trees to a training data set. Each tree is constructed based on the information from all previously built trees, enabling the model to learn gradually. Shallow trees capture few details of the problem and often yield poor performance, whereas deeper trees might lead to overfitting. Thus, it is important to tune the number of trees and interaction depth with XGBoost. Additionally, the hyperparameter shrinkage factor also needs to be tuned to prevent the model from quickly fitting and then overfitting the training dataset. Other tuning parameters considered in this work include subsample ratio (which means that XGBoost randomly samples a certain ratio of the training data before growing trees) and minimum child weight (which is the minimum number of instances needed to be in each node). In the current work, we firstly created a grid that contains all the possible combinations of tuning parameters. Multiple values of the tuning parameters are chosen within reasonable ranges. To identify the optimal hyperparameter set to improve the accuracy of the resulting models, we use grid search during 10-fold cross-validation (CV) to evaluate the independent sets of hyperparameters from the prespecified grid. The best set of hyperparameters is obtained using a scikit-learn class called "GridSearchCV." Subsequently, the best set of hyperparameters is selected and used in final models to make predictions on the testing dataset.

C. Performance Evaluation Measures

To evaluate the performance of the developed models to predict the icing area and maximum ice thickness, multiple error analysis measures were employed, including root mean squared error, R^2, and error distribution. For predicting the icing severity levels, the model is quantitatively evaluated by using several model evaluation indicators, such as precision, recall rate, F1 score, and confusion matrix.

III. Results and Discussion

This section has two parts. Performance evaluations are given in Sec. III.A. Performance comparisons between MLR and XGBoost for evaluating icing area and maximum ice thickness, and between OLR and XGBoost for evaluating icing severity level are conducted. Error analyses are performed to demonstrate the effectiveness of the proposed approach. Aircraft icing severity prediction results analysis is given in Sec. III.B.

A. Performance Evaluation

Icing area and maximum ice thickness on the test dataset are predicted, and the performance comparison between MLR and XGBoost is summarized in Table 3. It has been shown that the XGBoost has superior performance to MLR in both cases. The MLR yields much higher root mean square error (RMSE) and lower R^2 relative to XGBoost. It indicates that the linear input–output mapping given by MLR is not suitable in the current work due to the strong nonlinearity between the flight conditions and icing severity features. The MLR model simply serves as a benchmark model and should not be used for prediction. Indeed, if a commonly used MLR model gives satisfying results, there would be no use in predicting icing severity results using a sophisticated model such as XGBoost.

To further demonstrate the effectiveness of the XGBoost model, we present the comparison between the observed results and predicted results in the test dataset in Fig. 1. The predicted results are the ones predicted by the model and the observed results are the ones prepared in the dataset. The scatter plot is the predicted maximum ice thickness versus the observed maximum ice thickness, and predicted icing area versus the observed icing area in the test dataset. The red line with unit slope represents a perfect prediction. It indicates that the sufficient agreement is achieved between the predicted and observed results.

Figure 2 presents the histograms of the test errors to examine the nature of the error distributions. To gauge prediction performance, we use the median error instead of mean due to the highly skewed nature of the error distributions. For predicting maximum ice thickness, the median error is 0.0266. For predicting the icing area, the median error is 0.0061. Both of them are lower than 0.0500, which is deemed satisfactory [16].

Table 4 shows the performance comparison between OLR and XGBoost model in predicting the icing severity level. We can see that the accuracy of XGBoost model is significantly higher than OLR. The precision, recall rate, and F1 score generated by XGBoost for the four categories are all above 90%.

To further compare the performance between the OLR and XGBoost, we summarized the confusion matrix generated by the two models in Table 5. Each row of the matrix represents the predicted category, and each column represents the actual category. It can be seen that in the matrix, the diagonal values are much higher than the nondiagonal value for the XGBoost model. However, we observe a large number of extreme error cases for OLR. It is concluded that the OLR model simply serves as a benchmark model and should not be used for prediction. XGBoost model vastly improves the prediction accuracy.

B. Aircraft Icing Severity Evaluation Based on XGBoost

In this section, the prediction results from the XGBoost model are presented. Six flight conditions, that is, flight speed, angle of attack (AOA), exposure time, LWC, MVD, and environmental temperature, are given to the model. The effects of MVD and exposure time on the aircraft icing severity are studied by holding other flight conditions constant. The feature importance rankings [10] from XGBoost show that the MVD and exposure time have the highest importance scores among the six flight conditions. It indicates that MVD and exposure time have the most significant effect on the icing severity features (icing area, maximum ice thickness, and icing severity level). It should be noted that the built XGBoost model was able to make predictions in seconds on a 3.5 GHz Intel Core i7 processor. However, numerical simulation approach usually takes a few hours of CPU time.

Table 3 Performance comparison between MLR and XGBoost

	Maximum ice thickness		Icing area	
Model	RMSE	R^2	RMSE	R^2
MLR	0.1031	0.654	0.5112	0.641
XGBoost	0.0011	0.995	0.0600	0.995

Downloaded by Roberto Paoli on July 14, 2021 | http://arc.aiaa.org | DOI: 10.2514/1.I010978

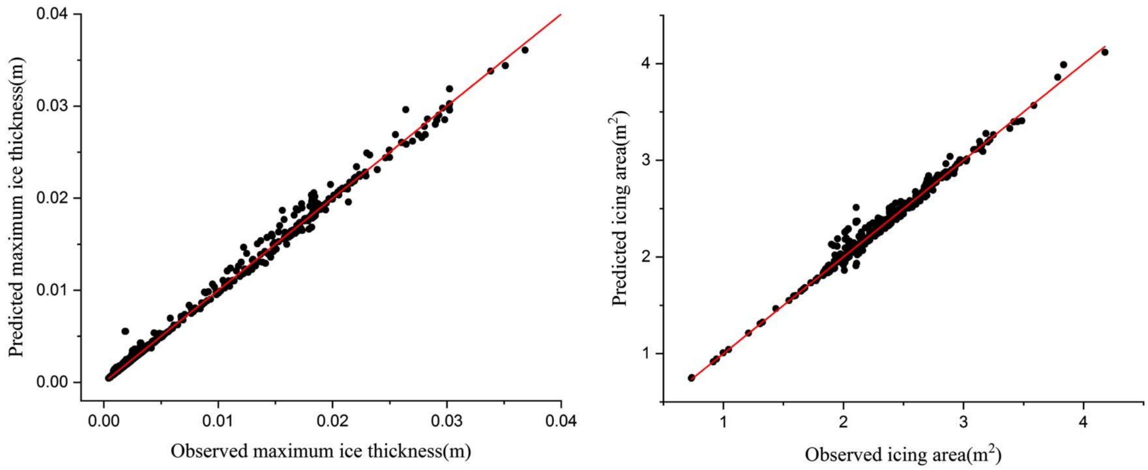

Fig. 1 Scatter plot of observed results vs predicted results. Left panel: maximum ice thickness; right panel: icing area.

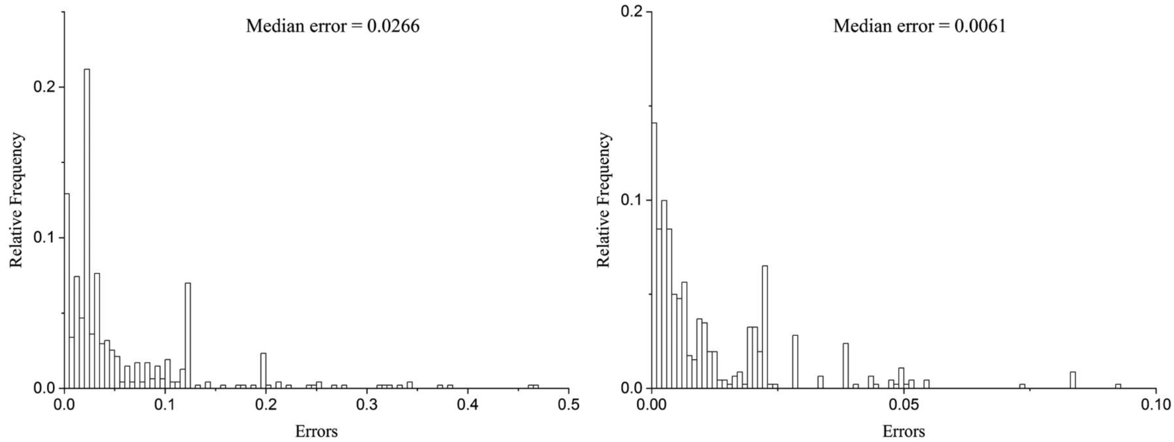

Fig. 2 Histogram showing distribution of test errors. Left panel: maximum ice thickness; right panel: icing area.

The effect of MVD on the icing severity results is presented in Fig. 3. The corresponding environmental temperature, exposure time, LWC, angle of attack, and flight speed are set to be 253 K, 1 min, 1.0 g/m^3, 0°, and 130 knots, respectively. It can be concluded that fairly good agreement is achieved for the model. The icing area and maximum ice thickness increase with the MVD. It can be accounted for by the inertia increment. It has been shown that the droplet trajectory depends strongly on its inertia K, whose expression is given as [17]

$$K = \frac{1}{18}\frac{D^2 V_\infty \rho_w}{c\mu_A} \tag{1}$$

where D is the droplet diameter, V_∞ is the free steam velocity, ρ_w is the water droplet density, c is the characteristic chord length, and μ_A is the air viscosity.

A droplet's size, and therefore mass and inertia, directly affect how and where the ice forms on an aircraft surface. The larger the diameter of water droplets, the greater the inertia [see Eq. (1)]; it is more likely that water droplets penetrate the surface streamlines and impact on the aircraft surface and hence increase the icing area. As can be seen from Fig. 3, the larger the water droplets, the thicker the ice layer, and the larger the icing area, the greater impact on the aircraft safety. However, it should be noticed that the droplets with diameter less than 10 μm do not contribute much to the icing severity. Indeed, droplets smaller than 15 μm are not included in the FAR [18] because the droplets are so small that they are convected around aircraft surface.

Figure 4 shows that the built model can successfully capture the icing area and maximum ice thickness with exposure time with the corresponding environmental temperature, angle of attack, LWC, MVD, and flight speed are set to be 253 K, 0°, 1.0 g/m^3, 20 μm,

Table 4 Performance results of OLR and XGBoost in predicting icing severity level

	OLR			XGBoost		
Category	Precision	Recall rate	F1 score	Precision	Recall rate	F1 score
Light	0.80	0.84	0.82	1.00	0.99	0.99
Moderate	0.68	0.64	0.66	0.96	0.99	0.97
Heavy	0.70	0.66	0.68	0.97	0.94	0.96
Severe	0.69	0.79	0.73	0.93	0.98	0.95

Table 5 Confusion matrix results of OLR and XGBoost in predicting icing severity level

	OLR				XGBoost			
Category	Light	Moderate	Heavy	Severe	Light	Moderate	Heavy	Severe
Light	187	36	0	0	244	3	0	0
Moderate	43	119	23	0	0	147	2	0
Heavy	4	17	74	17	0	3	113	4
Severe	0	2	8	37	0	0	1	50

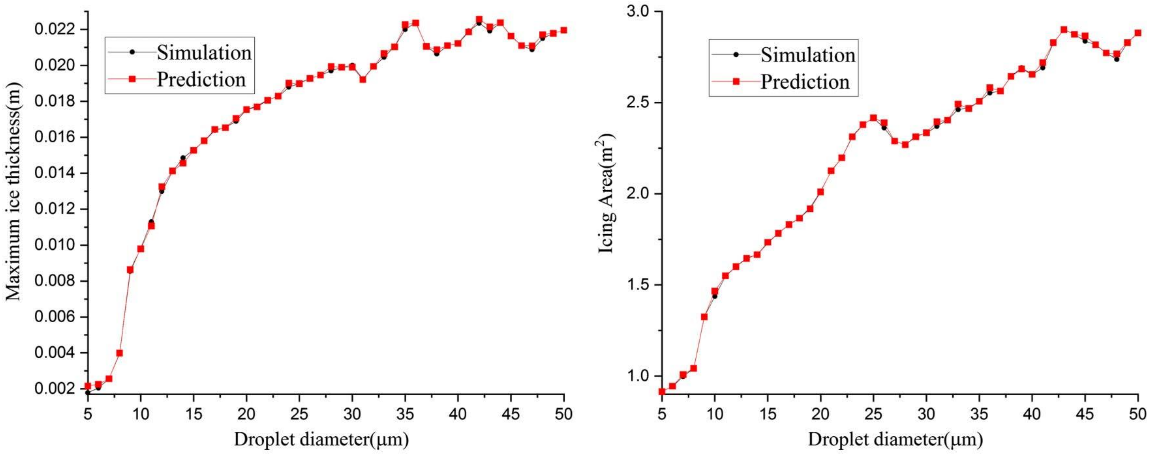

Fig. 3 Effect of medium droplet diameter on icing severity results. Left panel: maximum ice thickness; right panel: icing area.

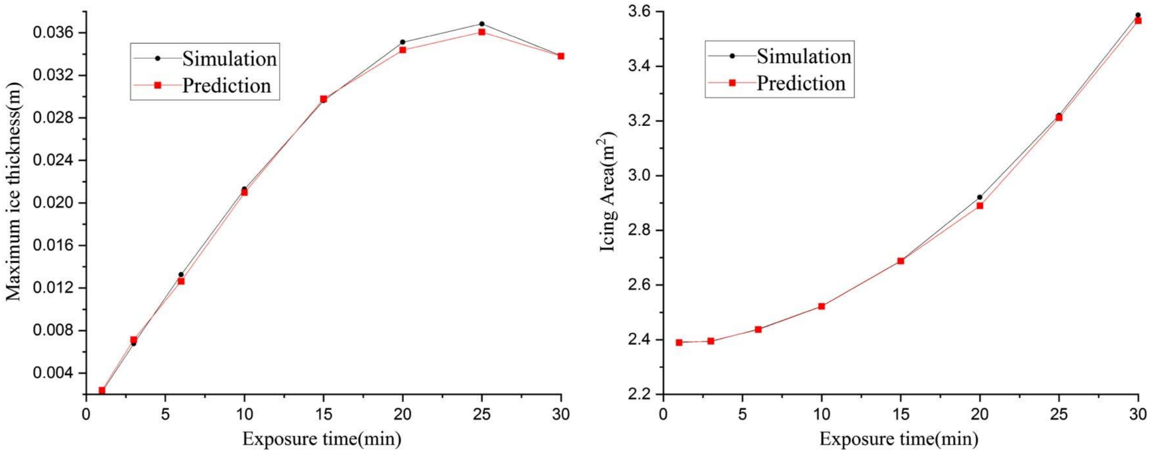

Fig. 4 Effect of exposure time on icing severity results. Left panel: maximum ice thickness; right panel: icing area.

Downloaded by Roberto Paoli on July 14, 2021 | http://arc.aiaa.org | DOI: 10.2514/1.I010978

and 130 knots, respectively. Exposure time is the time it takes to travel through the icing conditions [3]. It should be noticed that the icing area growth ratio increases with time, which means that the longer a flight stays in the icing cloud, the faster the icing area increases. This observation might be attributed to the runback water effect [19]. Under certain icing conditions, in the early icing stage where very thin ice layer created, all the water droplets freeze immediately upon the impact on the aircraft surface and there exists no overflow water. As the ice layer thickness increases, the conductive heat loss gets weaker. The overflow water first appears as the ice layer has grown to a certain extent, which is referred to as critical ice thickness [19]. Because the overflow water is mainly driven by the air–water interface friction, it moves closely following the wall air streamlines [20]. Hence, it is appropriate to consider that the overflow water might contribute to the icing area increment. Intuitively, the longer the aircraft stays in icing condition, the thicker the ice layer, which is also confirmed by Fig. 4. Additionally, it is worth mentioning that thickness growth ratio decreases with time. Again, it can be accounted for by the runback water effect. As the ice layer grows to the critical ice thickness, only part of the water droplet gets frozen and the control volume consists both ice and water layers. This phenomenon slows down the ice thickness growing speed.

IV. Conclusions

This paper proposed a method for aircraft icing severity prediction at different flight conditions based on machine learning model XGBoost. MLR and OLR serve as the benchmark models; the performance measures show that they are not suitable for the icing severity evaluation. In the application to predict icing area, maximum ice thickness, and icing severity level, well-defined performance measures are carried out through the training and testing process. The applications demonstrate that the proposed approach can provide an attractive alternative to traditional numerical simulation approach due to its limited computational resources requirement, fast performance, and reasonable accuracy. The effects of different flight conditions on the aircraft icing severity results are studied. The feature importance rankings from the trained model indicate that the flight conditions droplet diameter and exposure time have the most significant effect on the icing severity. The limitation found on the use of the current method is that the range of the predictions is limited to the range of the dataset. The built model will not be able to make accurate predictions if the given flight conditions are out of the current dataset range. The hybrid machine learning and CFD methods can be applied to the estimation of the degradation of the aircraft performance. The coupling between the proposed methodology and other CFD codes is currently being explored. The comprehensive investigation about different machine learning models (conventional methods and ensemble methods) on the icing thickness and area prediction will be presented in the future work, including the comparison in building process, effectiveness, and computational cost.

Acknowledgments

This work was supported by Argonne National Laboratory through grant number ANL 4J-30361-0030A, titled "Multiscale Modeling of Complex Flows," and by National Science Foundation through grant number 1854815, titled "High-Performance Computing and Data-Driven Modeling of Aircraft Contrails," awarded to R. Paoli.

References

[1] Mclean, J., "Determining the Effects of Weather in Aircraft Accident Investigations," *24th Aerospace Sciences Meeting*, AIAA Paper

1986-0323, 1986.
https://doi.org/10.2514/6.1986-323

[2] Bourgault, Y., Boutanios, Z., and Habashi, W. G., "Three-Dimensional Eulerian Approach to Droplet Impingement Simulation Using FENSAP-ICE, Part 1: Model, Algorithm, and Validation," *Journal of Aircraft*, Vol. 37, No. 1, 2000, pp. 95–103.
https://doi.org/10.2514/2.2566

[3] Cao, Y., Tan, W., and Wu, Z., "Aircraft Icing: An Ongoing Threat to Aviation Safety," *Aerospace Science and Technology*, Vol. 75, April 2018, pp. 353–385.
https://doi.org/10.1016/j.ast.2017.12.028

[4] Li, S., Qin, J., He, M., and Paoli, R., "Fast Evaluation of Aircraft Icing Severity Using Machine Learning Based on XGBoost," *Aerospace*, Vol. 7, No. 4, 2020, pp. 36–54.
https://doi.org/10.3390/aerospace7040036

[5] Ogretim, E., Huebsch, W., and Shinn, A., "Aircraft Ice Accretion Prediction Based on Neural Networks," *Journal of Aircraft*, Vol. 43, No. 1, 2006, pp. 233–240.
https://doi.org/10.2514/1.16241

[6] McCann, D. W., "NNICE—A Neural Network Aircraft Icing Algorithm," *Environmental Modelling and Software*, Vol. 20, No. 10, 2005, pp. 1335–1342.
https://doi.org/10.1016/j.envsoft.2004.09.027

[7] Zhan, Z., Habashi, W. G., and Fossati, M., "Local Reduced-Order Modeling and Iterative Sampling for Parametric Analyses of Aero-Icing Problems," *AIAA Journal*, Vol. 53, No. 8, 2015, pp. 2174–2185.
https://doi.org/10.2514/1.J053654

[8] de Mello, R. F., and Ponti, M. A., "Statistical Learning Theory," *Machine Learning*, Springer, New York, 2018.
https://doi.org/10.1007/978-3-319-94989-5

[9] Friedman, J. H., "Greedy Function Approximation: A Gradient Boosting Machine," *Annals of Statistics*, Vol. 29, No. 5, 2001, pp. 1189–1232.
https://doi.org/10.1214/aos/1013203451

[10] Chen, T., and Guestrin, C., "XGBoost: A Scalable Tree Boosting System," *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, Assoc. for Computing Machinery, New York, 2016, pp. 785–794.
https://doi.org/10.1145/2939672.2939785

[11] Darlington, R. B., and Hayes, A. F., "Multicategorical Regressors," *Regression Analysis and Linear Models: Concepts, Applications, and Implementation*, Guilford Publ., New York, 2016, https://www.guilford.com/books/Regression-Analysis-and-Linear-Models/Darlington-Hayes/9781462521135.

[12] Harrell, F., "Ordinal Logistic Regression," *Regression Modeling Strategies*, Springer Series in Statistics, Springer, New York, 2015, pp. 311–325.
https://doi.org/10.1007/978-3-319-19425-7

[13] Li, S., and Paoli, R., "Modeling of Ice Accretion over Aircraft Wings Using a Compressible OpenFOAM Solver," *International Journal of Aerospace Engineering*, Vol. 2019, June 2019, p. 11.
https://doi.org/10.1155/2019/4864927

[14] Li, S., and Paoli, R., "Numerical Study of Ice Accretion over Aircraft Wings Using Delayed Detached Eddy Simulation," Bulletin of the American Physical Soc., No. Q23, 2019, https://meetings.aps.org/Meeting/DFD19/Session/Q23.9.

[15] Grajski, K. A., Breiman, L., Prisco, G. V. D., and Freeman, W. J., "Classification of EEG Spatial Patterns with a Tree-Structured Methodology: CART," *IEEE Transactions on Biomedical Engineering*, Vol. BME-33, No. 12, 1986, pp. 1076–1086.
https://doi.org/10.1109/TBME.1986.325684

[16] Nwachukwu, A., Jeong, H., Pyrcz, M., and Lake, L. W., "Fast Evaluation of Well Placements in Heterogeneous Reservoir Models Using Machine Learning," *Journal of Petroleum Science and Engineering*, Vol. 163, April 2018, pp. 463–475.
https://doi.org/10.1016/j.petrol.2018.01.019

[17] Heinrich, A., Ross, R., Zumwalt, G., Provorse, J., and Padmanabhan, V., "Aircraft Icing Handbook," Vols. 1–3, Federal Aviation Administration TR DOT/FAA/CT–88/8–1, 1993, http://www.tc.faa.gov/its/worldpac/techrpt/ct888-1.pdf.

[18] "Part 25—Airworthiness Standards: Transport Category Airplanes," Federal Aviation Administration, 2013, https://www.ecfr.gov/cgi-bin/text-idx?node=14:1.0.1.3.11.

[19] Cao, Y., Huang, J., and Yin, J., "Numerical Simulation of Three-Dimensional Ice Accretion on an Aircraft Wing," *International Journal of Heat and Mass Transfer*, Vol. 92, Jan. 2016, pp. 34–54.
https://doi.org/10.1016/j.ijheatmasstransfer.2015.08.027

[20] Hedde, T., and Guffond, D., "ONERA Three-Dimensional Icing Model," *AIAA Journal*, Vol. 33, No. 6, 1995, pp. 1038–1045.
https://doi.org/10.2514/3.12795

D. Allaire
Associate Editor

Downloaded by Roberto Paoli on July 14, 2021 | http://arc.aiaa.org | DOI: 10.2514/1.I010978