Aerospace Science and Technology 176 (2026) 112103

Contents lists available at ScienceDirect

Aerospace Science and Technology

journal homepage: www.elsevier.com/locate/aescte

Machine learning prediction of aircraft anti-icing fluid endurance: A comparative study using natural and artificial snow data

Andres Cardona [a,b,*], Haïfa Nakouri [c,d], Jean-Denis Brassard [a,b], Gelareh Momen [a,e]

[a] University of Quebec at Chicoutimi, Department of Applied Sciences, Chicoutimi, Québec, Canada
[b] Anticing Materials International Laboratory (AMIL), UQAC, Chicoutimi, Québec, Canada
[c] University of Quebec at Chicoutimi, Mathematics and Computer Science, Chicoutimi, Québec, Canada
[d] University of Tunis, Institut Supérieur de Gestion (ISG), LARODEC, Tunis, Tunisia
[e] École de Technologie Supérieure (ETS), Department of Aerospace Engineering, Chicoutimi, Montréal, Canada

A R T I C L E I N F O

Communicated by Dr Pinqi Xia

Keywords:
Machine learning
Aircraft anti-icing fluid endurance
Natural snow
Artificial snow

A B S T R A C T

Ground ice poses a critical safety risk to winter aviation operations, primarily because the protective endurance of aircraft anti-icing fluids fluctuates unpredictably with volatile weather conditions. To mitigate these risks, this study develops a robust machine learning framework designed to forecast fluid endurance times using high-fidelity data from both artificial and natural snow environments. Research was conducted via endurance tests following the SAE ARP5485B protocol to train and evaluate multiple decision tree-based architectures, specifically the Decision Tree, Random Forest, and Extreme Gradient Boosting (XGBoost) models. Model reliability was ensured through rigorous k-fold cross-validation and systematic hyperparameter optimization. Results demonstrate that the optimized XGBoost model achieved superior predictive accuracy, yielding an RMSPE of 15.38% and a MAPE of 8.88%. Furthermore, SHAP analysis identified snow deposition rate and ambient temperature as the most influential determinants of fluid failure. These findings suggest that integrating machine learning into endurance testing can revolutionize current safety protocols and facilitate precise, data-driven decision-making for winter ground operations.

1. Introduction

Ground icing constitutes a considerable and frequently overlooked risk to aviation safety, especially during pre-takeoff procedures. A thin layer of snow, frost, or freezing rain on aerodynamic surfaces, such as wings, tail, or fuselage, can significantly disrupt airflow, resulting in increased drag and decreased lift [1]. Research indicates that even minimal contamination can significantly compromise aerodynamic performance, increasing the likelihood of accidents during critical phases of flight [2]. The crash of Air Ontario Flight 1363 in 1989 exemplified significant hazards: insufficient de-icing resulted in ice buildup on the wings, impairing lift and leading to the aircraft's crash, which caused 24 fatalities [3,4]. This event highlighted the critical significance of comprehensive ground de-icing and anti-icing procedures.

To mitigate such risks, Transport Canada Civil Aviation (TCCA) enforces compliance with SAE ARP5485 procedures, which determine the endurance time of anti-icing fluids under controlled and real-world conditions [5–7]. These fluids–classified as Type I (AMS1424) and Type II–IV (AMS1428)–exhibit non-Newtonian, pseudoplastic behaviour, with viscosity dependent on shear rate [6]. Endurance testing is adding fluid to a sloping aluminum plate that simulates an aircraft surface, followed by exposure to controlled snow conditions until failure, characterized by the commencement of surface contamination [8]. Field testing conducted under natural precipitation subsequently corroborate laboratory findings and develop Holdover Time (HOT) guidelines, assisting operators in determining safe time limits prior to departure [9].

Nevertheless these standardized methods, significant disparities between laboratory and natural test outcomes endure, underscoring the divergence between controlled and real-world settings. A recent study suggests that snow particles interacting with surfaces under various conditions undergo thermal and hydrodynamic interactions [10–12]. Villeneuve et al [13] assert that ethylene-based fluids have unique cooling and rebalancing phases correlated with the type of snow and saturation level. This underscores the significance of considering snow unpredictability when evaluating endurance length. Although widely utilized, standardized laboratory and outdoor endurance tests are limited by practical, budgetary, and environmental constraints. Natural testing is inherently opportunistic, expensive, and geographically limited, leading to sparse datasets that inadequately represent the complete spectrum of operating winter conditions. In contrast, laboratory-based ar-

* Corresponding author.
E-mail address: acpaniagua@etu.uqac.ca (A. Cardona).

https://doi.org/10.1016/j.ast.2026.112103
Received 3 November 2025; Received in revised form 12 February 2026; Accepted 8 March 2026
Available online 10 March 2026

tificial snow experiments provide controlled repeatability but fail to completely replicate the stochastic variability of natural precipitation, encompassing variations in snow density, particle shape, and thermal interactions at the fluid–surface interface [14,15]. Consequently, endurance time evaluations based exclusively on a single testing environment sometimes lack generalizability in real-world applications.

Beyond conventional endurance testing, recent research has investigated a broad range of anti-icing and surface protection strategies designed to limit ice and snow accumulation under challenging winter conditions. These efforts include the development of active and passive anti-icing systems, advances in surface engineering, and hybrid thermal–material solutions, reflecting the inherently multidisciplinary nature of aircraft icing mitigation [10,16,17]. Collectively, these studies indicate that anti-icing effectiveness is governed not only by fluid formulation, but also by the coupled influence of environmental forcing, surface characteristics, and operational conditions [12,18]. As a result, there is an increasing need for predictive frameworks capable of integrating heterogeneous variables to support the assessment and optimization of anti-icing strategies under realistic and variable winter scenarios.

Machine learning provides an additional framework that may integrate various data from both controlled laboratory and natural testing to overcome these limitations. Data-driven models can interpolate in untested scenarios and improve forecast reliability by identifying non-linear relationships among environmental variables, fluid properties, and endurance performance, therefore exceeding the constraints of experimental validation alone. In this regard, machine learning attempts to improve current SAE and Transport Canada testing methodologies by enabling scalable, adaptive, and interpretable forecasts of anti-icing fluid endurance length under diverse winter conditions.

A variety of machine learning algorithms have been tested to simulate ice formation and predict anti-icing effectiveness. Deep learning frameworks, such as Deep Belief Networks (DBN) and Stacked Auto-Encoders (SAE), have achieved notable accuracy in simulating icing phenomena; however, they often require considerable training and computational resources [16,19–21]. Tree-based models, especially XGBoost, combine interpretability and computational efficiency, providing strong classification and regression results with moderate data demands [22]. This study introduces a cohesive and interpretable machine learning framework for aircraft icing, trained and validated on both artificial and natural snow datasets, in contrast to previous research that utilized only laboratory or field data. It explicitly assesses generalization, optimization strategies, and feature-level physical interpretability via SHAP analysis. The methodology seeks to uphold a Mean Absolute Percentage Error (MAPE) under 10–15% and a Root Mean Square Error (RMSE) within 5% of the Holdover Time (HOT), in accordance with engineering accuracy standards and validation criteria established in Transport Canada and SAE documents [23–27].

2. Methodology

The study is structured into two phases. Endurance time data is initially collected through standardized research in both natural and artificial snow settings. A machine learning model is subsequently developed to predict endurance time using environmental and operational data.

Three algorithms based on decision trees were assessed: Decision Tree, Random Forest, and XGBoost, utilizing the same input qualities. The model's reliability was evaluated using 10-fold cross-validation. Hyperparameter optimization was conducted using Bayesian Optimization and Genetic Algorithms, as depicted in Fig. 1. To verify applicability in operational contexts, model outputs were tested against FAA and Transport Canada Holdover Time (HOT) rules.

SHAP analysis indicates that endurance is influenced by a limited set of interpretable, physics-based variables, particularly when accounting for non-linear interactions between temperature and snow.

2.1. Data collection

Data preparation is the first stage in developing the prediction model. To recollect the data, two separate tests must be performed, both with the same parameters, for the Society of Automotive Engineering (SAE) [28]. To collect the snow in both tests, an aluminium support is installed, as depicted in Fig. 2, which features a right-angle support structure to replicate the inclination of aircraft wings at 10°.

Natural snow endurance tests were conducted at two sites to maximize variability: APS Aviation Inc. (under contract with Transport Canada's Transportation Development Centre) and the Anti-Icing Materials International Laboratory (AMIL). Both facilities followed the SAE ARP5485B protocol of the Aircraft Ground Anti-Icing Fluid Holdover and Endurance Time Testing Program. The setup presented in Fig. 3 employed an external aluminum structure to simulate an aircraft surface. Test fluid was released during snowfall, with environmental factors documented at one-minute intervals. Standardized protocols enabled uniformity across locations, facilitating the accurate assessment of fluid endurance duration under natural snow circumstances.

Artificial snow testing was conducted at the AMIL Laboratory under controlled conditions, in accordance with the SAE ARP5485B standard. The experimental setup consists of a cold room capable of maintaining ambient conditions with a thermal stability of ±0.5°C. Test plates were fitted with Pt100 temperature detectors made of platinum, located at mid-span and linked to an automated data collection system for ongoing monitoring and traceability. All temperature sensors and weighing instruments were calibrated semiannually in accordance with the ISO 10,012 requirements.

For reference, the combined standard uncertainty for derived quantities (e.g., snow load) can be expressed in accordance with Gaussian error propagation as:

$$u_y = \sqrt{\sum_{i=1}^{n} \left(\frac{\partial y}{\partial x_i} u_{x_i} \right)^2}$$

where u_{x_i} represents the standard uncertainty of each measured variable. However, because the objective of this study is predictive modeling using experimentally observed endurance times rather than deterministic analytical calculation, uncertainty propagation was not explicitly computed for each endurance prediction. Instead, measurement variability is inherently embedded within the experimental dataset through repeated trials and varying environmental conditions. The use of k-fold cross-validation further ensures that the reported performance metrics (RMSE, MAE, and MAPE) reflect averaged predictive deviations incorporating both modeling and experimental uncertainties across the full dataset.

The AMIL snow machine features an automated feeding system operating on a 10° sloped test plate (30cm × 50cm). Snow is produced by a top-mounted supply box and released onto a rotating cylinder equipped with small cavities that control the deposition rate and uniformity of artificial snowfall [29]. To ensure the quality and conformity of the artificial snow with the characteristics of natural snowfall, its density is measured twice daily. Volumetric sampling is performed using a calibrated snow dropper, and the deposited mass is recorded using a VWR-203P2 digital precision balance with 0.001g readability and ±0.004g accuracy. For each volume setting, six successive measurements are averaged to ensure repeatability. The SLF SNOWPRO-4040 sensor (FPGA GmbH, Switzerland) is utilized to determine the bulk density of snow. This measurement method adheres to the protocols established in recent research on indoor snow testing [30].

As illustrated in Fig. 4, the snow machine was initiated until approximately 30% of the test surface was covered by unmelted white snow, and the cold chamber was preconditioned several days in advance to reach the target ambient temperature. During testing, the snow machine performs a back-and-forth motion to uniformly cover the test

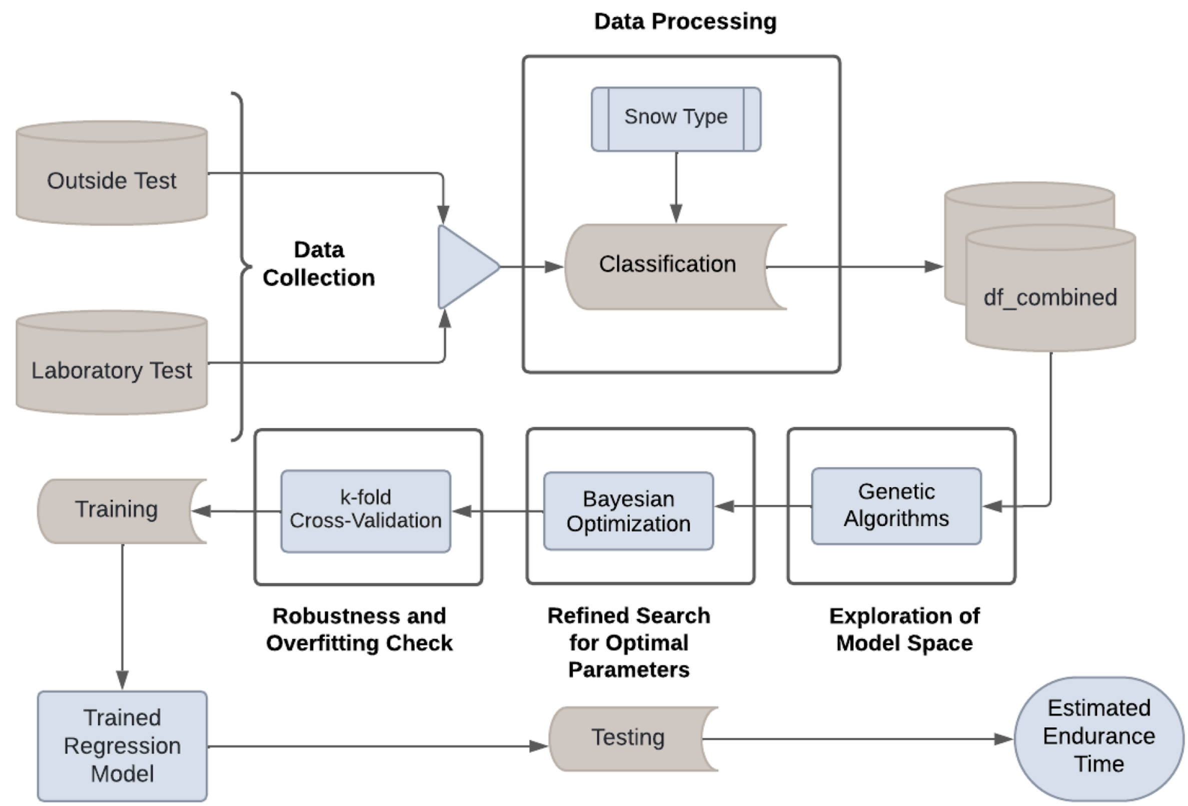


Fig. 1. Data processing and optimization pipeline for snow classification and fluid endurance prediction.

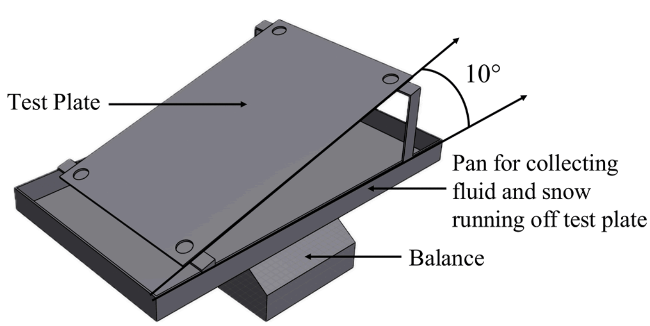


Fig. 2. Support structure for test plate and snow deposit.

plate. Snowfall is directed directly beneath the machine, where a scale positioned under the test plate measures the accumulated mass. The data is relayed to a computer that processes the measurements and provides them to the snowfall distribution control system. The testing occurs in a temperature-controlled cold room, with data gathering and control computers positioned externally to prevent heat interference.

Fluid failure, according to the SAE ARP5485B criteria, is considered to occur when approximately 30% of the test surface is visually obscured by white snow, indicating that the fluid has reached melt saturation and has lost its anti-icing effectiveness [30]. To reduce subjectivity, the evaluation is carried out by two trained observers following the defined protocol. The visual assessment is reinforced by infrared thermography, which confirms the absence of further temperature decrease upon snow deposition, as well as by the visible accumulation of unmelted snow on the plate surface. In practice, the grid lines drawn on the plate surface facilitate estimation of the 30% coverage threshold, as shown in the images. The implementation of a defined SAE ARP5485B visual technique, alongside independent assessment by two qualified observers and validation by infrared thermography, guaranteed consistency and repeatability in failure detection across all testing.

Fig. 5 illustrates a representative test used to build the ML training dataset. The left image shows the start of the test, with the surface completely covered by the fluid and snow melting upon contact. After 138 minutes, as depicted in the right image, around 30% of white snow, particularly in the central and lower areas, although some fluid persists at the top. This evident alteration in surface coverage illustrates the systematic identification and documentation of failures for dataset construction.

The dynamics of snowflake deposition and their interaction with the test surface are known to affect the fluid's stability and failure threshold. Such interactions, including restitution and flake deformation on impact, have been recently modeled using ML and fluid-mechanical frameworks [11,31].

2.2. Computational methods

To provide a clearer overview of the experimental setup and dataset structure, Table 1 summarizes the key characteristics of the samples collected under both artificial and natural snowfall conditions, as well as the modeling strategies applied in this study.

The combined dataset demonstrates an evident mismatch between artificial snow samples (478 cases) and real snow samples (295 cases), highlighting the practical limitations of field testing under genuine snowfall settings. This disparity reflects practical realities, since controlled laboratory testing facilitates greater data throughput compared to opportunistic outdoor experiments. To alleviate potential bias resulting from disparate sample sizes, various solutions were implemented. Initially, all models conducted evaluation by k-fold cross-validation to guarantee consistent performance across various data segments. Secondly, model generalization was methodically evaluated by training on artificial, natural, and hybrid datasets, while ensuring assessment on natural snow data to represent authentic operational settings. Ultimately, performance indicators were analyzed across datasets to ensure that predicted accuracy did not unduly favor the predominant artifi-

Fig. 3. Setup for natural snow test.

Table 1
Dataset Characteristics and Testing Conditions.

Characteristic	Details
Total Samples	773
Artificial Snow Samples	478 (collected in controlled laboratory environment)
Natural Snow Samples	295 (collected under real snowfall conditions)
Samples per Test (Lab)	Approximately 3 samples per fluid
Samples per Test (Natural)	Approximately 9 samples per fluid
Anti-Icing Fluids Tested	4 fluids (labeled A, B, C, D for confidentiality)
Snow Intensity Range	1.24 to 35.8 $g/dm^2/h$
Temperature Range	−26.2°C to −1.97°C
Modeling Techniques	Decision Tree, Random Forest, XGBoost
Prediction Target	Endurance time (y) of anti-icing fluids
Research Objective	Predict anti-icing fluid performance under diverse environmental conditions

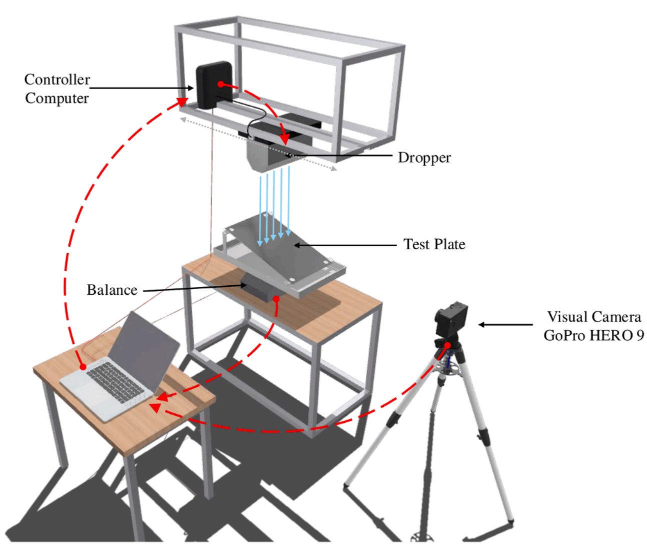


Fig. 4. Setup for artificial snow test.

cial snow category. These techniques jointly mitigate sampling-induced bias and enhance the reliability of the proposed machine learning system for predicting endurance duration under varying snowfall circumstances.

In this study, three supervised ML algorithms were implemented to model the endurance time of anti-icing fluids: Decision Tree, Random Forest, and Extreme Gradient Boosting (XGBoost). These algorithms were selected due to their capacity to handle nonlinear interactions between environmental variables such as temperature, snow concentration, and fluid type. Ensemble models like Random Forest and XGBoost further enhance predictive performance by aggregating multiple weak learners and reducing overfitting [32,33]. Among them, XGBoost demonstrated the most stable and accurate results during cross-validation and was therefore chosen as the final model. Detailed theoretical descriptions of these models are available in prior literature [34–36].

2.2.1. Validation strategies for enhanced model training

K-fold cross-validation is one of the most popular and effective validation strategies employed. It is a technique leveraged to assess a model's generalization capability by partitioning the dataset into K subsets, commonly referred to as folds. During training, the model undergoes training for K iterations, utilizing $K-1$ folds for the training process while reserving the remaining fold for validation purposes [37]. The final performance is determined by calculating the average across all folds, providing a more dependable assessment of the model's accuracy. The formal representation of the cross-validation error calculation is as follows:

$$CV_{error} = \frac{1}{K}\sum_{i=1}^{K} E_i \quad (1)$$

In this context, E_i represents the error associated with the ith validation fold. This methodology facilitates the identification of overfitting

(a) Beginning of the test (b) After 138 minutes

Fig. 5. Visual fluid failure, according to SAE ARP5485B criteria.

and guarantees the robustness of the model, particularly in the context of limited datasets.

2.2.2. Hyperparameter tuning

Hyperparameter tuning is the process of selecting the optimal values for a ML model's hyperparameters. Bayesian Optimization and Genetic Algorithms are two of the most powerful techniques for hyperparameter tuning.

Bayesian Optimization represents a probabilistic approach to hyperparameter tuning, wherein a surrogate function, commonly modeled as a Gaussian Process (GP), is constructed to facilitate an efficient search process. The process achieves a balance between exploration and exploitation by determining the subsequent hyperparameter values through the application of an acquisition function, exemplified by Expected Improvement (EI):

$$EI(x) = (\mu(x) - f(x^{+}))\Phi(Z) + \sigma(x)\phi(Z) \tag{2}$$

where $\mu(x)$ and $\sigma(x)$ are the mean and standard deviation of the GP, $f(x^{+})$ is the best observed value, and Z is the standard normal variable. Bayesian Optimization presents significant advantages for the tuning of models such as XGBoost, Support Vector Machines, and deep learning networks, primarily by minimizing the number of required function evaluations [38].

Genetic Algorithms (GA) are derived from the principles of biological evolution and are employed to address complex optimization challenges. A population of candidate solutions is developed through successive generations utilizing selection, crossover, and mutation processes. The fitness of each individual is assessed through the application of a function denoted as $F(x)$, with subsequent generations being established in accordance with selection probabilities. The process of evolution can be represented as follows:

$$x_{new} = \alpha x_{parent1} + (1-\alpha)x_{parent2} + \text{mutation} \tag{3}$$

where α serves as a parameter that regulates the crossover rate. Genetic algorithms demonstrate efficacy in contexts characterized by high-dimensional or non-continuous search spaces, including applications such as deep learning architecture search and feature selection. Although it incurs significant computational costs, this approach offers global optimization in contexts where gradient-based methods encounter difficulties [39,40].

The amalgamation of these methodologies has the potential to improve the efficacy of predictive models, as they fulfill various roles, encompassing global search, hyperparameter optimization, and model assessment. The subsequent approach is characterized by a logical framework:

- Evaluate the robustness of the model and address overfitting by employing **K-Fold Cross-Validation**.
- Conduct an investigation into a diverse array of hyperparameter configurations and potential model architectures through the application of the **Genetic Algorithm**.
- Employ **Bayesian Optimization** to enhance the search process, concentrating on significant hyperparameter regions previously identified by the Genetic Algorithm.

In order to prevent overfitting while employing effective optimization methods, model building followed a methodical and sequential approach. K-fold cross-validation (k = 10) was employed as the external validation mechanism to assess model generalization across different data partitions. This validation system utilized Genetic Algorithms to thoroughly explore the hyperparameter space, identifying significant regions without dependency on gradient-based assumptions. Bayesian Optimization was subsequently employed to refine hyperparameters inside these domains, focusing on local exploitation and convergence efficiency. Optimization was exclusively conducted on training folds, while validation folds remained unexposed during the tuning process. The final model's efficacy was evaluated using natural snow data to confirm its reliability in real-world operational contexts. This stratified method integrates exploration, exploitation, and validation, thus reducing the danger of overfitting while enhancing prediction stability.

The optimization of hyperparameters, coupled with the establishment of robust validation protocols, is complemented by the implementation of advanced feature engineering techniques, which are crucial for enhancing the predictive capability of the model. The integration of polynomial and nonlinear features, feature interactions, and feature transformations enables the model to effectively capture the complex relationships that are inherent in the physical behaviour of snow-fluid interactions. For example, polynomial expansions, including the squaring of temperature values, facilitate the model's ability to incorporate nonlinear effects, wherein variations in temperature may not exert a uniform influence on endurance time [41]. In a similar vein, interaction terms (for instance, the product of temperature and snow rate) are essential for accurately depicting synergistic effects that arise solely when variables function simultaneously at extreme values [42]. Furthermore, log transformations significantly reduce skewness in distributions, hence improving numerical stability and promoting better model convergence during the optimization process [43].

The incorporation of Rule-based Features and Domain-Informed Features serves to augment the model's ability to accurately represent real-world phenomena. Binary indicators, which are informed by domain knowledge, such as a threshold-based cold condition flag, contribute to the identification of critical regimes in which fluid performance is adversely affected [44]. Furthermore, the calculated features, including the snow load obtained from density and mass, incorporate physics-based reasoning directly into the ML pipeline [37]. The integration of

Table 2
Descriptive statistics (minimum, maximum, mean, and standard deviation) for the input and engineered features used in model development.

Feature	Min	Max	Mean	Std Dev
Temperature	−26.08	−1.98	−14.20	7.13
Rate	1.40	35.78	18.56	10.12
Temp_Squared	3.91	680.03	252.33	206.20
Rate_Squared	1.96	1280.19	446.83	383.62
Temp_Rate_Interaction	−930.35	−5.03	−266.63	209.79
Abs_Temperature	1.98	26.08	14.20	7.13
Is_Very_Cold	0.00	1.00	0.86	0.34

domain-specific factors improves the model's interpretability and realism, essential for evaluating endurance time predictions in safety-critical contexts, such as the efficacy of aircraft anti-icing fluids. The fusion of these feature engineering techniques with the optimization framework yields a model distinguished by accuracy and grounded in both physical and statistical validity.

Table 2 presents the descriptive statistics (minimum, maximum, mean, and standard deviation) of the main input and engineered features used in model development.

These statistical ranges confirm the dataset's heterogeneity across environmental and operational conditions, supporting the model's ability to generalize under varied snow scenarios.

3. Results

3.1. Data analysis

Comprehending the distribution of critical variables, including temperature, snow application rate, and endurance time, is imperative prior to the interpretation of model outcomes. The configuration of these distributions elucidates fundamental trends within the dataset; for example, it indicates whether specific conditions were disproportionately represented or if outlier values might affect the model's performance. A skewed distribution in endurance time, for instance, suggests that a limited number of test scenarios result in considerably longer durations, which may signify optimal fluid performance under particular environmental conditions. Identifying these patterns serves to affirm the reliability of the data and establishes the framework for understanding the distribution of critical variables, including temperature, snow application rate, and endurance time, is crucial before interpreting the subsequent graphical results.

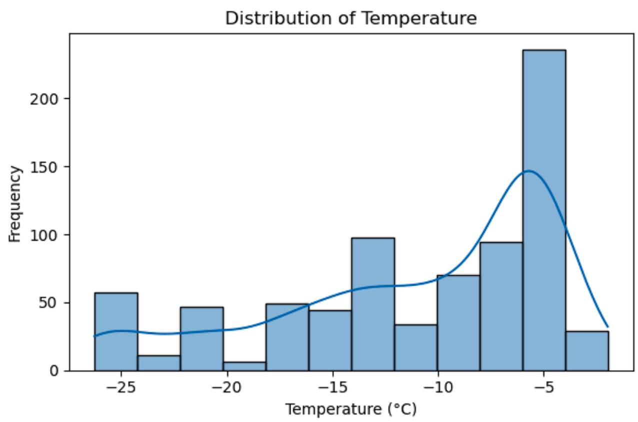
The distribution of key experimental variables, shown in Fig. 6, offers valuable insight into the operating conditions under which the performance of anti-icing fluid was assessed. These distributions were derived from the combined natural and artificial datasets, thereby capturing a comprehensive range of testing scenarios. The temperature variable exhibits a right-skewed distribution, characterized by a significant concentration of observations near −5 °C. This suggests that a significant percentage of the tests were conducted in proximity to the upper threshold of standard freezing conditions, wherein the efficacy of the fluid is anticipated to be of greater importance. The rate of snow deposition demonstrates a bimodal and positively skewed distribution, characterized by primary and secondary peaks located approximately at 5 to 10 units and 25 to 30 units, respectively. This indicates that the dataset encompasses distinct clusters that correspond to scenarios of low and high snowfall intensity. The distribution of fluid endurance time exhibits a right-skewed pattern, with the majority of values concentrated within the range of 30 to 70 minutes, accompanied by a gradual decrease in frequency as endurance time extends. This pattern underscores that, although the majority of fluids failed within a discernible time frame, a specific subset demonstrated prolonged endurance under particular environmental conditions. The distributions presented collectively illustrate the variability and heterogeneity inherent in real-world anti-icing conditions, thereby providing a foundational basis for the training of the predictive models.

Fig. 7 illustrates the differences in endurance time distributions among various fluid types and snow conditions, specifically emphasizing produced snow. The median endurance times exhibit significant stability across various snow types; however, controlled environments reveal more pronounced outliers and wider interquartile ranges, particularly for Fluids B and D. Fluid B exhibits endurance durations exceeding 200 minutes on artificial snow, suggesting notable variability rather than a uniform trend of enhancement. The visual evidence suggests that dispersion increases under controlled snow conditions, likely due to laboratory-specific factors such as flake uniformity or deposition rate. The observed variation may indicate that the laboratory setting successfully replicated a broader spectrum of natural variables. This observation does not imply that artificial snow inherently improves performance. Rather, the variability may reflect a wider range of environmental scenarios simulated in the laboratory setting. It is important to clarify that the identification of subsets showing prolonged endurance cannot be made solely from Fig. 7; further statistical or model-based analysis would be required. Therefore, interpretations based solely on interquartile range should be made cautiously.

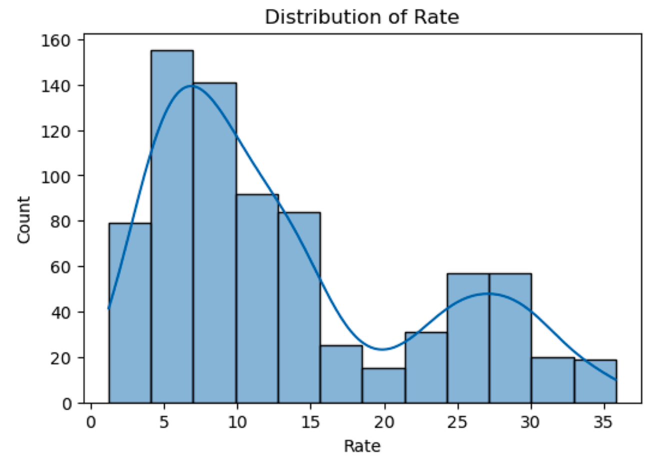
The correlation between endurance time and the rate of snow deposition demonstrates a distinct inverse relationship across all fluid categories. The scatter plot, Fig. 8, together with logarithmic regression fits, demonstrates a steady decrease in endurance as snow rate increases throughout the temperature range of 0 to -20°C. This observation suggests that increased snowfall contributes to an expedited occurrence of fluid failure. Although all fluids adhere to this overarching trend, the rate of decline exhibits variability, with Type A and Type D fluids demonstrating a more pronounced decrease in endurance as the rates increase. Conversely, Type B demonstrates a tendency to sustain comparatively elevated endurance values in low to moderate snowfall conditions, which is consistent with prior observations regarding its wider performance spectrum. The existence of high-endurance outliers at low snow rates, particularly for Type B and Type D, underscores the potential for prolonged protection during light snowfall conditions. The results of this study substantiate the significance of the snow rate as an essential predictive variable for the modelling of fluid endurance, This helps differentiate of fluid formulations in relation to environmental load conditions. This variability aligns with findings from recent snow impact studies, which highlight the non-linear influence of snow particle trajectory and surface restitution on retention and fluid endurance [10,31].

The correlation matrix (Fig. 9) depicts the relationships among key environmental and response variables. A moderate negative correlation ($r = -0.55$) was observed between snow deposition rate and endurance duration, indicating that higher snowfall rates reduce fluid efficacy. The temperature and snow rate exhibit a moderate positive correlation ($r = 0.49$), while the temperature and endurance time show a weak association ($r = 0.05$). The robustness of this link will be further evaluated using the SHAP analysis in Section 3.3.

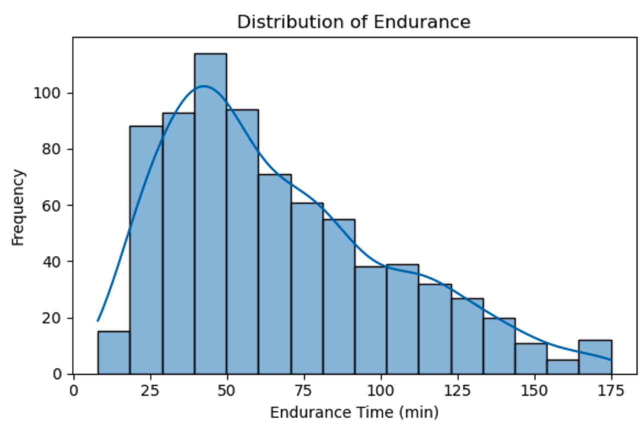
The exploratory analysis of variable distributions and their interrelationships reveals significant patterns that inform the modeling strategy. The considerable influence of snow deposition rate on endurance time, coupled with the observed variability among various fluid and snow types, highlights the necessity for models capable of accurately depicting the complex, non-linear interactions inherent in the data. The findings articulated in this study not only corroborate the experimental dataset but also offer direction for the selection and assessment of appropriate ML methodologies. Building upon the analytical framework previously established, the subsequent section assesses the predictive efficacy of three supervised learning models, Decision Tree (DT), Random Forest (RF), and Extreme Gradient Boosting (XGBoost), across different test conditions, with the aim of identifying the most effective approach for estimating fluid endurance time in relation to various environmental loads.



(a) Distribution of Temperature



(b) Distribution of Rate



(c) Distribution of Endurance

Fig. 6. Distribution of the variables.

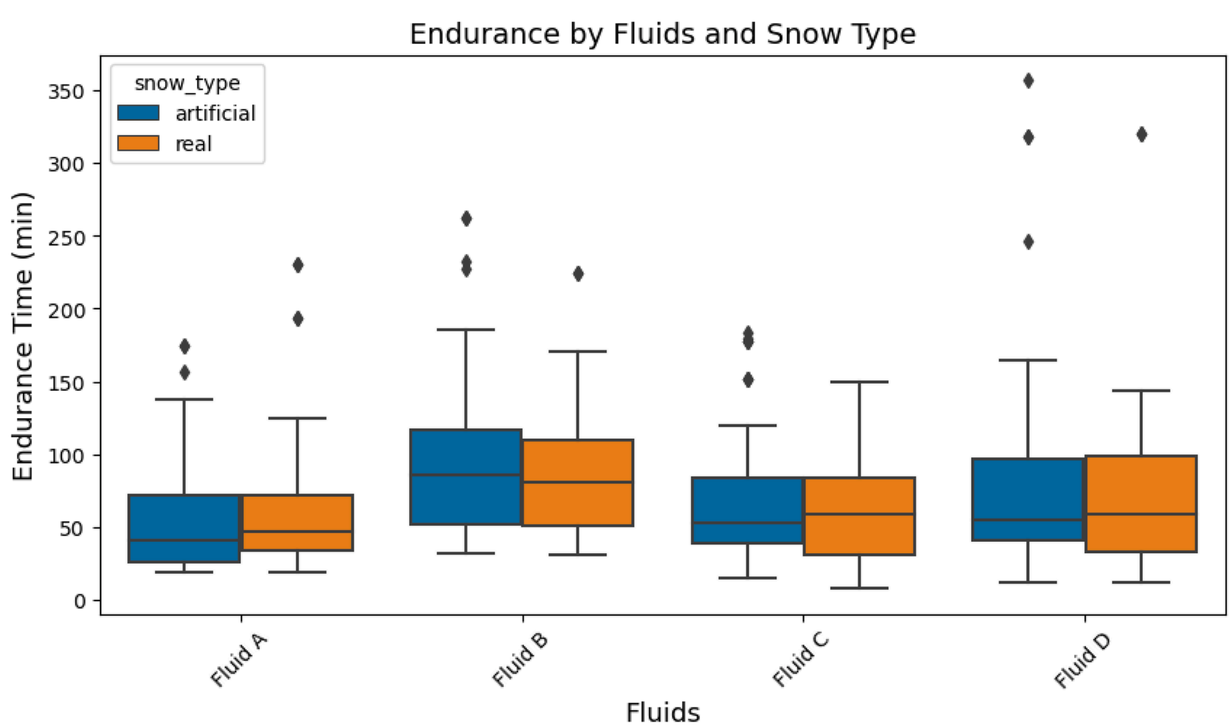


Fig. 7. Boxplot endurance time by fluids and snow type.

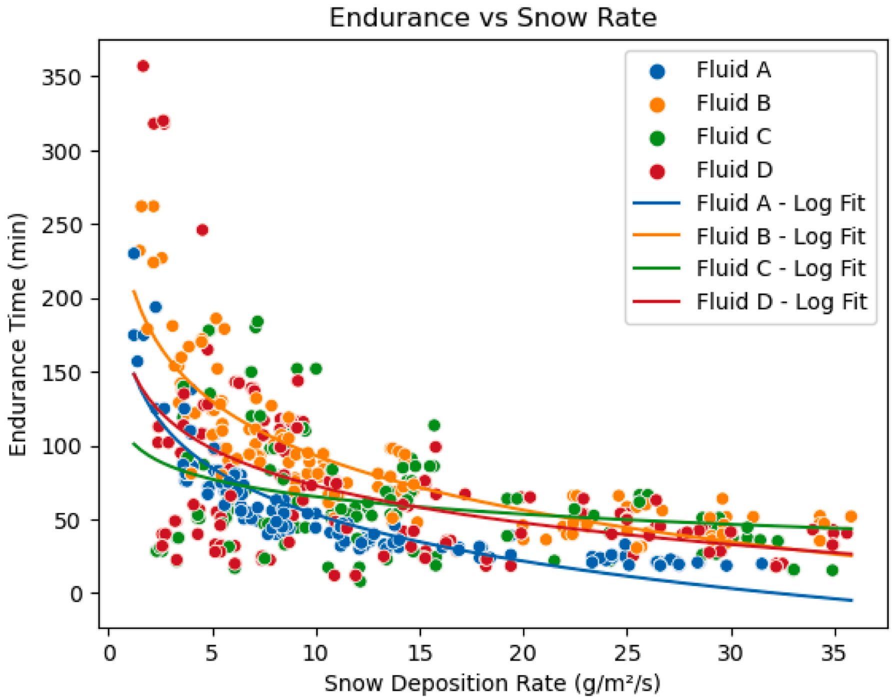


Fig. 8. Endurance time vs. snow deposition rate by fluids from temperatures 0 to -20°C.

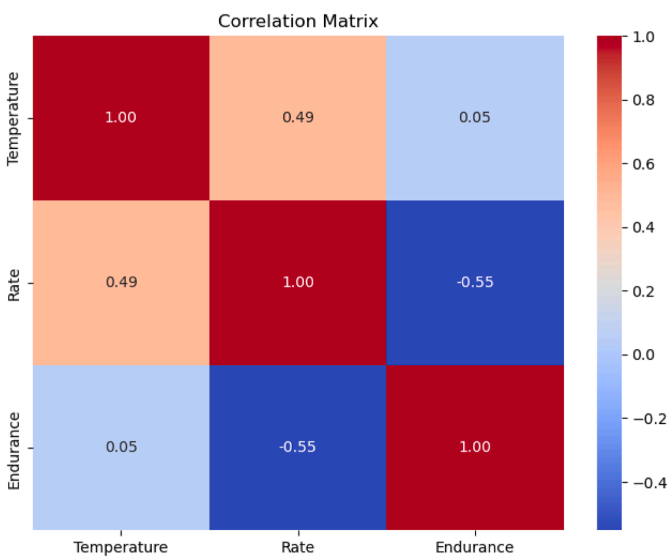


Fig. 9. Matrix of correlations: Temperature, snow rate, endurance.

3.2. Prediction performance

In order to evaluate the predictive capability of different algorithms for estimating endurance time, an assessment was conducted utilizing three supervised learning models: Decision Tree (DT), Random Forest (RF), and Extreme Gradient Boosting (XGBoost). The unrefined models, developed without feature engineering or post-processing, were evaluated across natural, artificial, and hybrid datasets. Table 4 delineates the performance metrics utilized: the coefficient of determination (R^2), Root Mean Square Error (RMSE), and Mean Absolute Error (MAE), offering an exhaustive assessment of the accuracy and reliability of each model. It is important to note that, for each model, separate training was conducted on the natural, artificial, and combined datasets. The performance metrics presented in Tables 4–6 correspond exclusively to predictions made on real (natural) snow data, allowing for a consistent evaluation of model generalization under authentic environmental conditions.

The configuration of hyperparameters for each model, along with training time and cross-validation details, is summarized in Table 3. These values were obtained either from default settings (for DT and RF) or through optimization procedures such as Bayesian and genetic algorithms in the case of XGBoost.

3.2.1. Evaluation metrics

The coefficient of determination (R^2) is used to measure how well a prediction model fits the actual data. A value closer to 1 indicates that the model captures most of the variability in the target variable, while values closer to 0 suggest poor predictive power. RMSE, which gives higher weight to larger errors, indicates how concentrated the prediction errors are around the true values; a lower RMSE reflects higher accuracy. Likewise, MAE measures the average absolute deviation between expected and actual values, enhancing its interpretability in real units. Lower MAE values are advantageous, signifying that the model exhibits reduced average prediction errors.

3.2.2. Model selection

Gradient Boosting, namely XGBoost, was selected for its ability to enhance fluid behavior identification [45] and its expertise in modeling complex, non-linear interactions between environmental variables and endurance time. Cross-validation and comparative assessments of Decision Tree and Random Forest models confirmed their robustness.

The consolidated dataset analysis revealed that XGBoost attained a $R^2 = 0.896$, an RMSE $= 12.105$, and a MAPE $= 5.135$. While both Decision Tree and Random Forest occasionally attained slightly improved R^2 or reduced RMSE values (see Table 4), XGBoost consistently produced lower MAE, demonstrating greater prediction stability. The model exhibited strong performance after implementing advanced feature engineering and optimization, confirming its selection for further refinement and interpretability assessment, as detailed in Section 3.3.

Table 3
Hyperparameter configurations and training details for each ML model. XGBoost parameters were optimized using Bayesian and genetic algorithms.

Hyperparameter	Decision Tree [35]	Random Forest [32]	XGBoost [33]
Max Depth	5	10	5
n Estimators	–	100	100
Learning Rate	–	–	0.1
Subsample	–	–	1.0
Colsample Bytree	–	–	1.0
Alpha (α)	–	–	0.0
Lambda (λ)	–	–	1.0
Random State	42	42	42
Cross-Validation Folds (k)	10	10	10
Training Time (s)	0.41	2.23	3.98

Table 4
Model Comparison – Coefficient of Determination (R^2), Root Mean Squared Error (RMSE), Mean Absolute Error (MAE) for Decision Tree (DT), Random Forest (RF), and XGBoost (XGBoost).

	DT			RF			XGBoost		
Dataset	R^2	RMSE	MAE	R^2	RMSE	MAE	R^2	RMSE	MAE
Natural	0.899	11.193	4.109	0.887	14.302	6.745	0.899	11.195	4.033
Artificial	0.897	11.795	4.409	0.888	13.826	6.340	0.896	11.953	4.688
Combined	0.898	11.591	4.290	0.886	13.874	6.440	0.896	12.105	5.135

Table 5
Model Comparison with Feature Engineering – for Decision Tree (DT), Random Forest (RF), and XGBoost (XGBoost).

	DT			RF			XGBoost		
Dataset	R^2	RMSE	MAE	R^2	RMSE	MAE	R^2	RMSE	MAE
Natural	0.899	11.193	4.099	0.887	14.229	6.702	0.899	11.194	4.123
Artificial	0.897	11.795	4.409	0.883	14.548	6.716	0.897	11.892	4.648
Combined	0.898	11.590	4.290	0.888	13.288	6.787	0.897	11.908	4.972

3.2.3. Feature engineering and model refinement

In order to improve the model's representation of physical interactions, several features were added: `Temp_Squared` (to describe the nonlinear temperature–endurance relationship), `Temp` × `SnowRate` (to account for the combined effects of temperature and snow deposition rate), `Log_SnowAge` (to rectify distributional skewness), `Is_Cold` (a binary indicator for essential temperature thresholds), and `Density_Mass` (a physics-informed feature that indicates snow load).

The inclusion of these engineered variables enhanced the input space of the XGBoost model. As shown in Table 5, the updated model achieved an R^2 = **0.897**, an RMSE = **11.908**, and MAPE = **4.972** on the combined dataset, compared with the previous R^2 = **0.896**, RMSE = **12.105**, and MAPE = **5.135**. Considering the slight numerical improvements, these findings validate that feature engineering helped improve model robustness and interpretability. The diminished prediction error and more consistent metrics suggest that improved physical representations, specifically snow load and temperature-snow interactions, augmented the model's ability to accurately reflect essential endurance dynamics. Subsequent enhancements may augment precision and applicability.

The current model exhibits considerable predictive power; however, it is crucial to further reduce residual errors to enhance resilience and generalizability, particularly in safety-critical applications. This project will employ sophisticated optimization methods, such as Bayesian optimization and evolutionary algorithms, along with an efficient k-fold cross-validation strategy to improve hyperparameter tuning, avoid overfitting, and minimize prediction error.

The R^2 and RMSE values displayed in Tables 4 and 5 provide complementary insights into model performance and reliability. High R^2 values (about 0.89–0.90 across datasets) indicate that the models accurately capture the primary variability in endurance time affected by environmental and operational factors. The minor variations in R^2 among natural, artificial, and mixed datasets suggest that the essential physical correlations learned by the models remain consistent across many testing environments.

RMSE values, expressed in minutes, denote the absolute deviation between predicted and actual endurance durations. In aviation ground operations, endurance times generally extend over several tens of minutes; hence, RMSE values of roughly 10 to 12 minutes align with forecast errors that remain within the permissible operating parameters established in SAE ARP5485B. The decrease in MAE noted for XGBoost compared to Decision Tree and Random Forest models further signifies enhanced stability and less sensitivity to outliers. Collectively, these measures indicate that the chosen modeling methods attain a balance between goodness-of-fit and operationally significant accuracy.

3.2.4. Optimized models prediction performance

The model went through additional optimization aimed at improving generalization and forecast accuracy, utilizing the feature collection that was developed earlier. The optimization phase focused on adjusting hyperparameters rather than introducing new variables, maintaining the same physical and statistical characteristics described in the preceding section. This included nonlinear and interaction terms: `Temp_Squared` (to capture the nonlinear temperature–endurance relationship), `Temp` × `SnowRate` (for the joint influence of temperature and snow deposition rate), `Log_SnowAge` (to correct distributional skewness), `Is_Cold` (a binary indicator for critical temperature thresholds), and `Density_Mass` (a physics-informed feature representing snow load).

The enhanced feature representation provided a solid foundation for the optimization phase, enabling the XGBoost algorithm to discern more complex relationships between environmental variables and endurance time. Four optimization methodologies were utilized to evaluate model performance across artificial, natural, and hybrid datasets. The initial

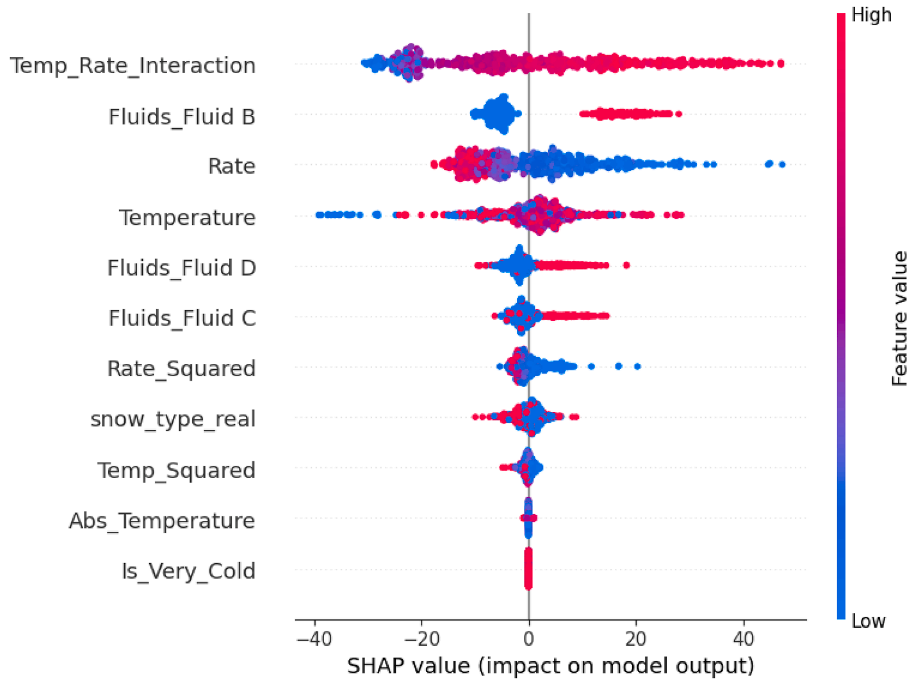


Fig. 10. SHAP summary plot for the tuned XGBoost model trained on combined natural and artificial snow data.

Table 6
Performance Metrics for Optimization Techniques: R^2, RMSE, RMSE (%), MSE, MAE, MAE (%).

	10-Fold CV			Bayesian			10-Fold CV + Bayesian			Model Refinement Method		
Training Dataset	Artificial	Natural	Combined	Artificial	Natural	Combined	Artificial	Natural	Combined	Artificial	Natural	Combined
R^2	0.90	0.90	0.90	0.92	0.81	0.79	0.92	0.91	0.92	0.92	0.91	0.92
RMSE	11.36	12.07	11.70	10.10	15.71	8.75	10.12	11.60	10.29	10.10	11.29	10.65
RMSPE(%)	16.99	18.21	17.48	17.51	18.81	18.02	15.13	17.51	15.38	15.11	17.04	15.91
MSE	129.16	145.69	136.90	629.49	661.48	351.75	102.45	134.59	105.91	102.11	127.46	113.33
MAE	7.16	6.84	7.86	14.87	8.04	13.56	5.63	6.41	5.94	5.78	6.52	6.45
MAPE(%)	10.71	10.33	8.88	23.07	8.16	22.50	8.64	9.68	8.88	8.64	9.83	9.64

method employed 10-fold cross-validation (10-Fold CV), selecting $k = 10$ for its equilibrium between bias, variance, and computational efficiency [46]. The second employed Bayesian optimization, the third merged 10-Fold cross-validation with Bayesian optimization, and the fourth amalgamated genetic algorithms with both methods–designated as the Model Refinement Method.

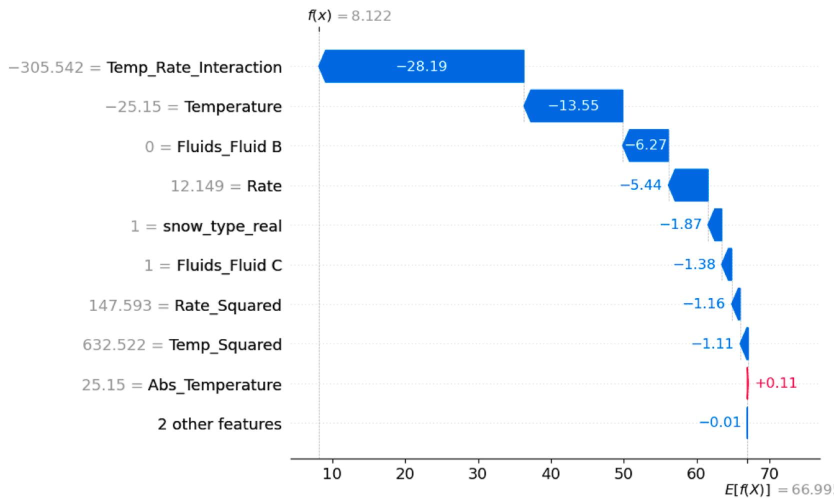
The evaluation of model generalization in artificial and natural snow environments was conducted by training and validating models on distinct datasets, as well as on a unified dataset. The results in Table 6 indicate that models trained on the combined dataset exhibit consistently robust performance when assessed on natural snow data, with minimal deterioration compared to artificial snow forecasts. The optimized XGBoost model has similar RMSPE and MAPE values across testing scenarios, suggesting that the established relationships are not confined to controlled laboratory settings. Artificial snow data offer more consistent sampling and reduced variability, but real snow data demonstrate more stochasticity resulting from environmental changes. The model's capacity to uphold expected accuracy under these circumstances validates its robust generalization and endorses the utilization of synthetic snow data as a dependable adjunct to natural snowfall testing within a cohesive machine learning framework.

The study targets a prediction accuracy within 10% of actual endurance time, consistent with industry standards in SAE Aerospace Recommended Practice ARP5485B [8]. The integrated optimization framework, incorporating Bayesian optimization, and k-fold cross-validation, yields superior overall performance. The consolidated dataset analysis revealed that XGBoost achieved a coefficient of determination (R^2) of 0.92, an RMSPE of 15.38%, and a MAPE of 8.88% after optimization, surpassing alternative methodologies. This hybrid methodology effectively succeeded by Bayesian optimization to concentrate on high-performance areas. The slight disparity in RMSE between artificial (10.12) and natural datasets (11.60) affirms robust generalization facilitated by the k-fold framework.

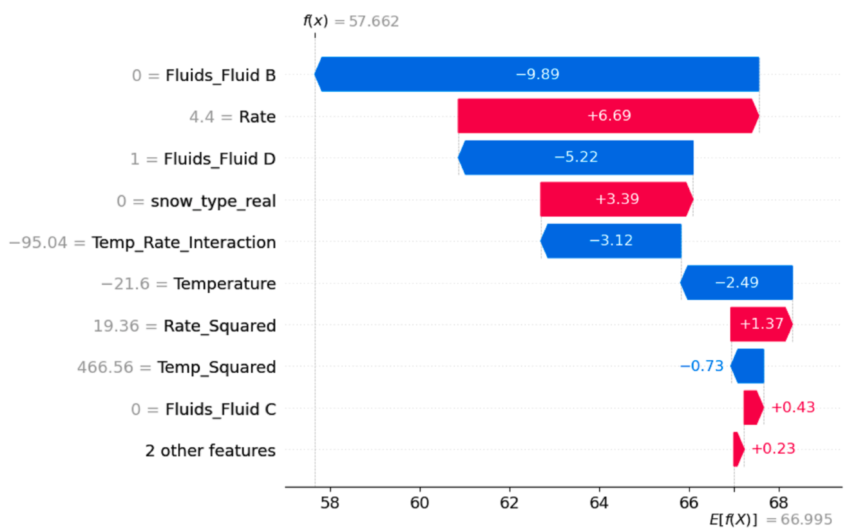
Bayesian Optimization alone performs well on artificial data (R^2 = 0.92) but exhibits limited generalization, as reflected by its high MAPE (23.07%). The Genetic Algorithm shows larger variability (MAE = 8.52 for artificial and 6.52 for natural data), whereas the integrated approach balances errors across all datasets (MAE = 5.78, 6.52, and 6.45, respectively). The 10-Fold Cross-Validation baseline produces reliable results (RMSE = 11.70, MAPE = 8.88%) but does not fully exploit the parameter space. Overall, the comprehensive optimization framework reduces RMSE by roughly 9% and increases R^2 from 0.90 to 0.92, demonstrating effective overfitting control and improved predictive stability under diverse snow conditions.

3.3. Model explainability using SHAP analysis

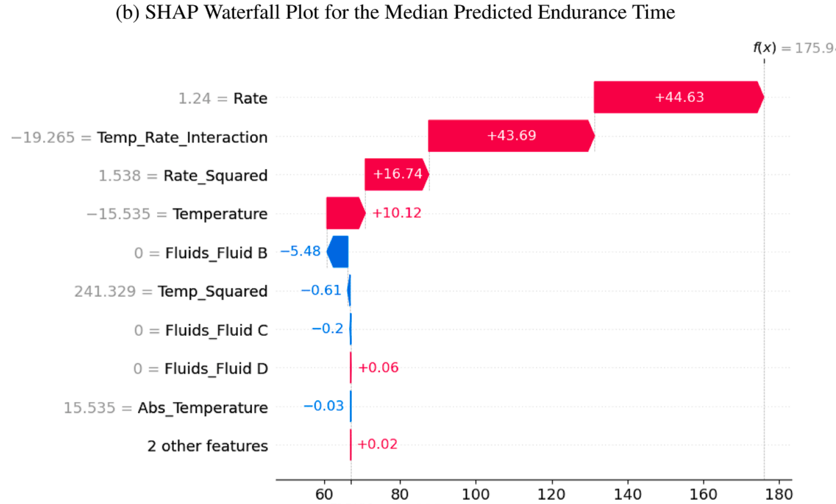
Comprehending the influence of each factor on model predictions is crucial for understanding outcomes and directing further experiments. Upon determining XGBoost as the optimal model, feature significance was assessed utilizing SHapley Additive exPlanations (SHAP) values [47].



(a) SHAP Waterfall Plot for the Lowest Predicted Endurance Time



(b) SHAP Waterfall Plot for the Median Predicted Endurance Time



(c) SHAP Waterfall Plot for the Highest Predicted Endurance Time

Fig. 11. SHAP waterfall plots illustrating feature contributions for representative low, median, and high endurance time predictions.

Before conducting SHAP analysis, continuous input variables (e.g., temperature, snow deposition rate, and designed interaction terms) were utilized in their original physical units, as tree-based models like XGBoost are unaffected by monotonic feature scaling. During model training, categorical variables denoting anti-icing fluid type were encoded via binary indicator (one-hot) encoding. SHAP values were derived directly from the trained XGBoost model via the TreeExplainer framework, hence ensuring alignment between the prediction model and the explainability study. This method enables the interpretation of both continuous and categorical characteristics in a physically significant way, without the bias introduced by scaling.

Fig. 10 illustrates the SHAP summary plot, emphasizing that the interaction between temperature and snow deposition rate (`Temp Rate Interaction`) is the predominant factor influencing endurance duration. Elevated snowfall rates coupled with temperature fluctuations substantially influence fluid dynamics. Additional significant contributors comprise `Rate`, `Fluid B`, and engineered variables such as `Rate Squared` and `Temp Squared`, highlighting the significance of non-linear and interaction effects.

SHAP waterfall plots were created for three representative cases of endurance time: low, median, and high, to effectively demonstrate individual forecasts (Fig. 11). These charts illustrate the positive or negative contributions of each attribute to certain predictions. In low-endurance settings, significant adverse effects from low temperatures and elevated deposition rates prevail. Conversely, high endurance forecasts are influenced by moderate snowfall rates, conducive temperatures, and beneficial interactions among environmental variables and fluid characteristics

SHAP analysis indicates that endurance is influenced by a limited set of interpretable, physics-based variables, particularly when accounting for non-linear interactions between temperature and snow.

The SHAP analysis indicates that endurance time is primarily affected by a combination of environmental, operational, and fluid-related factors. In both natural and artificial snow datasets, the primary influencing factors are the snow deposition rate and external temperature, with heightened snowfall intensities and diminished temperatures consistently impacting fluid failure. Operational characteristics are represented by interaction terms, particularly the temperature-snow rate interaction, which highlights the non-linear correlation between thermal conditions and snow loading. Fluid-specific properties also affect endurance performance, revealing differences in formulation reactions to the same environmental stimuli. The relative importance of these factors remains consistent in both natural and artificial snow environments; however, research on manufactured snow reveals a slightly less variability in SHAP contributions due to controlled deposition rates and uniform snow properties. In contrast, natural snow data demonstrate heightened variability, indicating a broader range of environmental interactions. The findings indicate that machine learning models accurately depict physically significant systems while maintaining robustness across many testing scenarios.

4. Discussion

This study illustrates that machine learning can yield precise and comprehensible forecasts of anti-icing fluid longevity under various climatic conditions. By merging natural and artificial snow datasets for training, the model bridges the disparity between controlled laboratory experiments and real-world settings.

The integration of XGBoost, Bayesian optimization, and k-fold cross-validation improved predictive accuracy and mitigated overfitting, particularly under fluctuating snow and temperature conditions. These findings align with recent work emphasizing the significance of snow-surface interactions and energy dissipation processes in endurance behavior [12,48].

Moreover, feature engineering significantly influenced the model's performance. The inclusion of polynomial terms, interaction effects, and domain-specific indicators (such as `Temp Rate Interaction` and `Density Mass`) improved the model's physical realism and reduced residual errors. These properties allowed the model to generalize across various contexts while maintaining interpretability.

The SHAP study further shows that endurance time is primarily influenced by the interaction of snow deposition rate and temperature, together with fluid-specific characteristics. The waterfall diagrams illustrated how these interactions might either accelerate or delay fluid collapse. Most operational scenarios often occur within a moderate endurance range, where environmental factors exert a measurable but manageable influence.

Despite its strong performance, the model is limited by the relatively small size of the real snow dataset, which affects its ability to encounter rare environmental edge cases. Future studies should expand natural testing and examine additional physical properties, including surface roughness and phase transitions of precipitation. The application of generative models like VAEs and GANs may assist in data augmentation.

The prevailing influence of snow deposition rate and ambient temperature on endurance duration aligns with recognized physical principles that dictate anti-icing fluid failure. Heightened snowfall intensity expedites fluid dilution and saturation, whereas diminished temperatures impair melting efficiency and augment snow accumulation, resulting in an earlier decline in fluid efficacy. Comparable dependencies have been documented in experimental and numerical investigations of snow-surface interactions and fluid degradation in winter conditions, wherein the non-linear interplay between thermal and hydrodynamic effects is pivotal to endurance behavior [10,12,13].

From a modeling standpoint, the robust efficacy of tree-based machine learning methods demonstrates their capacity to include non-linear relationships and interaction effects that are challenging to depict through empirical correlations alone. Prior research utilizing machine learning for aircraft icing and anti-icing systems has shown that ensemble techniques like XGBoost provide an advantageous equilibrium between predictive precision and interpretability under conditions of limited data availability [22]. The uniformity of model performance across artificial and natural snow datasets corroborates existing research that indicates controlled laboratory data, when suitably integrated, can augment prediction robustness without compromising physical relevance.

5. Conclusion

This study presented a machine learning methodology to forecast the endurance duration of airplane anti-icing fluids, integrating artificial and natural snow data to reconcile differences between controlled experiments and actual environmental conditions. Utilizing XGBoost in conjunction with sophisticated optimization methods, including Bayesian optimization, and cross-validation, the model attained robust predictive efficacy (R^2 = 0.92, RMSPE = 15.38%, MAPE = 8.88%). Engineered characteristics, such as temperature-snow rate relationships and domain-informed variables, enhanced interpretability and congruence with physical behavior. SHAP studies confirmed that fluid degradation is affected by temperature and the interplay between snowfall and deposition. These results highlight AI's potential to enhance operational safety and refine fluid decision-making in winter aviation settings.

CRediT authorship contribution statement

Andres Cardona: Writing – original draft, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization; **Haïfa Nakouri:** Writing – review & editing, Writing – original draft, Validation, Supervision, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization; **Jean-Denis Brassard:** Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Software, Investigation, Funding acquisition, Data curation, Conceptualization; **Gelareh Momen:** Writing – review & editing, Supervision, Project administration, Investigation, Funding acquisition, Conceptualization.

Data availability

The data that has been used is confidential.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Gelareh Momen reports financial support was provided by Transport Canada. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work was supported by Transport Canada. Also like to express their sincere gratitude to the Anti-icing Materials International Laboratory (AMIL) for providing the facilities, technical support, and valuable insights that made this research possible.

References

[1] D. Gagnon, J.-D. Brassard, H. Ezzaidi, C. Volat, Computer-assisted aircraft anti-icing fluids endurance time determination, Aerospace 7 (4) (2020) 39. https://doi.org/10.3390/aerospace7040039
[2] F.A. Administration, Aircraft Ground Deicing and Anti-icing Program (Advisory Circular 120-60), Technical Report, Federal Aviation Administration, 2020. https://www.faa.gov/documentLibrary/media/Advisory_Circular/AC_120-60.pdf.
[3] V.P. Moshansky, Commission of Inquiry into the Air Ontario Crash at Dryden, Ontario, Technical Report, Government of Canada, 1992. http://epe.lac-bac.gc.ca/100/200/301/pco-bcp/commissions-ef/moshansky1992-eng/moshansky1992-eng.htm.
[4] T.S. B.o. Canada, Aviation Investigation Report A17C0146, Technical Report, Transportation Safety Board of Canada, 2017. https://www.tsb.gc.ca/eng/rapports-reports/aviation/2017/a17c0146/a17c0146.html.
[5] SAE International, AMS1424: deicing/anti-icing fluid, aircraft, SAE Type I, 2018, (https://www.sae.org/standards/content/ams1424/). Standard issued by SAE International.
[6] SAE International, AMS1428K: fluid aircraft deicing/anti-icing, non-Newtonian (Pseudoplastic), SAE Types II, III and IV, 2018, (https://www.sae.org/standards/content/ams1428k/). SAE International, p. 26.
[7] Transport Canada, Advisory circular: degree-specific holdover times; civil aviation, standards, 2021, (https://tc.canada.ca/en/aviation/reference-centre/advisory-circulars).
[8] SAE International, ARP5485B: endurance time test procedures for SAE type II/III/IV aircraft deicing/anti-icing fluids, 2017, (https://www.sae.org/standards/content/arp5485b/).
[9] H.H. Oda, P. Adrian, M. Arriaga, L. Davies, J. Hille, T. Sheehan, E.T. Suter, Safe winter operations, Aero Magazine (2010) 5–13.
[10] E. Khoshbakhtnejad, F.B. Golezani, H. Sojoudi, Dynamics of snowflakes impacting superhydrophobic surfaces, Langmuir 40 (37) (2024). https://doi.org/10.1021/acs.langmuir.4c01903
[11] E. Khoshbakhtnejad, F.B. Golezani, B. Mohammadian, H. Sojoudi, A.H. Abou Yassine, Trajectory and impact dynamics of snowflakes: fundamentals and applications, Powder Technol. (2024).
[12] B. Mohammadian, E. Khoshbakhtnejad, H. Sojoudi, Interfacial phenomena in snow from its formation to accumulation and shedding, Colloids Surf., A 623 (2021) 126758. https://doi.org/10.1016/j.colsurfa.2021.126758
[13] E. Villeneuve, C. Charpentier, J.D. Brassard, G. Momen, A. Lacroix, Aircraft anti-icing fluids endurance under natural and artificial snow: a comparative study, Int. Rev. Aerospace Eng. (IREASE) 15 (1) (2022) 1–11.
[14] L. Fay, M. Akin, S. Wang, X. Shi, D. Williams, Correlating Lab Testing and Field Performance for Deicing and Anti–icing Chemicals (Phase 1), Technical Report, Montana State University, 2010.
[15] M. Chaput, Endurance Time Testing in Snow: Reconciliation of Indoor and Outdoor Data: Winter 2000-01, Technical Report, Publications du gouvernement du Canada, 2001.
[16] H. Mao, X. Lin, Z. Li, X. Shen, W. Zhao, Anti-icing system performance prediction using POD and PSO-BP neural networks, Aerospace 11 (6) (2024) 430.
[17] H.K. Pazarlioglu, et al., Thermohydraulic performance assessment of new alternative methods for anti-icing application against current application in an aircraft, Proc. Inst. Mech. Eng. Part G: J. Aerospace Eng. (2025). https://doi.org/10.1177/09544089231190182
[18] H.K. Pazarlioglu, A.U. Tepe, K. Arslan, Optimization of parameters affecting anti-Icing performance on wing leading edge of aircraft, Avrupa Bilim ve Teknoloji Dergisi 34 (2022) 19–27. https://doi.org/10.31590/ejosat.1062495
[19] X. Yi, Q. Wang, C. Chai, Prediction model of aircraft icing based on deep neural network, Trans. Nanjing Univ. Astronaut. 38 (2021) 535–544.
[20] E.S. Abdelghany, M.B. Farghaly, M.M. Almalki, H.H. Sarhan, M.E.S.M. Essa, Machine learning and IoT trends for intelligent prediction of aircraft wing anti-icing system temperature, Aerospace 10 (8) (2023) 676.
[21] Z. Chu, J. Geng, Q. Yang, X. Yi, W. Dong, Prediction of temperature distribution on an aircraft hot-air anti-icing surface by ROM and neural networks, Aerospace 11 (11) (2024) 930.
[22] S. Li, J. Qin, M. He, R. Paoli, Fast evaluation of aircraft icing severity using machine learning based on XGBoost, Aerospace 7 (4) (2020) 36.
[23] J. Yang, et al., Snow depth estimation and historical data reconstruction from satellite and in situ data, Cryosphere 14 (6) (2020) 1763–1783. https://doi.org/10.5194/tc-14-1763-2020
[24] S. Thapa, et al., Comparative analysis of snowmelt-driven streamflow forecasting using machine learning techniques, Water (Basel) 16 (15) (2024) 2095. https://doi.org/10.3390/w16152095
[25] SAE International, Aerospace recommended practice ARP5485C: endurance time tests for aircraft deicing/anti-icing fluids, Technical Report, SAE International, 2013. https://www.sae.org/standards/content/arp5485c/.
[26] SAE International, Aerospace recommended practice ARP5718: minimum qualification standards for aircraft deicing/anti-icing fluids, Technical Report, SAE International, 2016. https://www.sae.org/standards/content/arp5718/.
[27] Transport Canada, Guidelines for aircraft ground-icing operations (TP 14052E), 2021, https://tc.canada.ca/en/aviation/publications/guidelines-aircraft-ground-icing-operations-tp-14052e.
[28] SAE International, AS5900: standard test method for aerodynamic acceptance of SAE AMS1424 and SAE AMS1428 aircraft deicing/anti-icing fluids, 2020, (https://www.sae.org/standards/content/as5900/). Society of Automotive Engineers, p. 33.
[29] A. Beisswenger, J. Perron, Indoor Snow Testing of Aircraft Ground Anti-icing Fluids, Technical Report, AMIL/UQAC, AMIL/UQAC: Saguenay, Qc, 2006.
[30] S. Benaissa, D. Harvey, H. Khawaja, G. Momen, Thermographic analysis of ethylene glycol–based aircraft anti-icing fluid: investigation of fluid failure mechanisms during simulated snow endurance tests, Cold Reg. Sci. Technol. 234 (2025) 104472. https://doi.org/10.1016/j.coldregions.2025.104472
[31] E. Khoshbakhtnejad, H. Sojoudi, Predicting coefficient of restitution of ice particles: a machine learning–based approach, Powder Technol. (2025). https://api.semanticscholar.org/CorpusID:280509902.
[32] L. Breiman, Random forests, Mach. Learn. 45 (1) (2001) 5–32.
[33] T. Chen, C. Guestrin, XGBoost: a scalable tree boosting system, In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining 22 (2016) 785–794.
[34] Y. Wang, H. Wu, D. Nettleton, Stability of random forests and coverage of random–forest prediction intervals, in: Advances in Neural Information Processing Systems, 2023, https://openreview.net/forum?id=xPqINp0Eu1. https://doi.org/10.48550/arXiv.2310.18814
[35] B. Lantz, Machine Learning with R: Expert techniques for predictive modeling, Packt Publishing, Birmingham, UK, 4 edition, Birmingham, UK, 2021.
[36] A. Géron, Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow, O'Reilly Media, Sebastopol, CA, 3 edition, Sebastopol, CA, 2022.
[37] B. Bischl, M. Binder, M. Lang, T. Pielok, J. Richter, S. Coors, J. Thomas, T. Ullmann, M. Becker, A. Boulesteix, D. Deng, M. Lindauer, Hyperparameter optimization: foundations, algorithms, best practices and open challenges, Wiley Interdiscipl. Rev. Data Mining Knowl. Discov. 13 (2) (2023) e1484. https://doi.org/10.1002/widm.1484
[38] X. Wang, Y. Jin, S. Schmitt, M. Olhofer, Recent advances in bayesian optimization, ACM Comput. Surv. 55 (13s) (2023) 1–36. https://doi.org/10.1145/3582078
[39] Y. Xie, Q. Liu, B. Xue, Genetic algorithm for hyperparameter optimization of neural networks, Appl. Soft Comput. 101 (2021) 107050. https://doi.org/10.1016/j.asoc.2020.107050
[40] Z. Liu, Q. Zhang, Y. Wang, A comprehensive survey on genetic algorithms for hyperparameter optimization, IEEE Access 10 (2022) 27590–27610. https://doi.org/10.1109/ACCESS.2022.3156604
[41] G. James, D. Witten, T. Hastie, R. Tibshirani, An Introduction to Statistical Learning: with Applications in R, Springer, 2013.
[42] M.H. Kutner, C.J. Nachtsheim, J. Neter, W. Li, Applied Linear Statistical Models, McGraw-Hill Education, 2005.
[43] G.E.P. Box, D.R. Cox, An analysis of transformations, J. R. Stat. Soc. Series B (Methodol.) 26 (2) (1964) 211–243.
[44] C. Molnar, Interpretable Machine Learning, Christoph Molnar, 2022. https://christophm.github.io/interpretable-ml-book/.
[45] Y. Sun, J. Zhang, Y. Zhang, Fluid classification through well logging is conducted using the extreme gradient boosting model based on the adaptive piecewise flatness-based fast transform feature extraction algorithm, Phys. Fluids 36 (1) (2024).
[46] H. Ghasemzadeh, R.E. Hillman, D.D. Mehta, Toward generalizable machine learning models in speech, language, and hearing sciences: estimating sample size and reducing overfitting, (2023). https://doi.org/10.48550/arXiv.2308.11197
[47] S.M. Lundberg, G. Erion, H. Chen, A. DeGrave, J.M. Prutkin, B. Nair, R. Katz, J. Himmelfarb, N. Bansal, S. Lee, From local explanations to global understanding with explainable AI for trees, Nature Machi. Intell. 2 (1) (2020) 56–67. https://doi.org/10.1038/s42256-019-0138-9
[48] E. Khoshbakhtnejad, F.B. Golezani, H. Sojoudi, Dynamics of snowflakes impacting superhydrophobic surfaces, Langmuir 40 (37) (2024) 11265–11275. https://doi.org/10.1021/acs.langmuir.4c01122