**SINERGI** Vol. 30, No. 3, October 2026: 901-914
http://publikasi.mercubuana.ac.id/index.php/sinergi
http://doi.org/10.22441/sinergi.2026.3.017


# GA-Optimized machine learning surrogate models for aerodynamic prediction using wind tunnel data

**Purwadi Sukadji[1,2]*, Syafaruddin Syafaruddin[1], Andani Ahmad[1], Ivransa Zuhdi Pane[2]**
[1]Department of Electrical Engineering, Universitas Hasanuddin, Indonesia
[2]Centre for Aeronautics and Aerospace, Badan Riset dan Inovasi Nasional (BRIN), Indonesia

***Abstract***
*Wind tunnel testing is essential for obtaining accurate aerodynamic coefficients but remains costly and time-consuming due to the large number of required test configurations. This study proposes a robust surrogate modeling framework by integrating Genetic Algorithm (GA)-based hyperparameter optimization with GroupKFold cross-validation using an Indonesian Low-Speed Tunnel (ILST) dataset containing 13,200 samples from a 20-passenger aircraft. After data preprocessing, five regression models, Support Vector Regressor (SVR), Artificial Neural Network (ANN), Decision Tree (DT), Random Forest (RF), and XGBoost, were optimized to predict six aerodynamic coefficients ($K_1$–$K_6$). The optimized ANN achieved high predictive accuracy, while the tree-based ensemble models consistently demonstrated superior generalization performance. Among all models, XGBoost produced the best overall results, slightly outperforming Random Forest in both interpolation and extrapolation tasks, highlighting its effectiveness in modeling complex nonlinear aerodynamic behavior. Although all models performed well within the training domain, the ensemble methods exhibited greater robustness when predicting previously unseen operating conditions. These findings demonstrate that GA-optimized tree-based ensembles provide an accurate and computationally efficient surrogate modeling framework, significantly reducing dependence on costly wind tunnel experiments while maintaining high prediction fidelity.*



***Keywords:***
*Aerodynamic;*
*Forces and moments;*
*Guidance;*
*Hyperparameters;*
*Machine learning;*
*Wind tunnel;*

***Article History:***
*Received: October 15, 2025*
*Revised: February 21, 2026*
*Accepted: March 3, 2026*
*Published: October 2, 2026*

***Corresponding Author:***
*Purwadi Sukadji*
*Electrical Engineering Department, Universitas Hasanuddin, Makassar, Indonesia*
*Email: purw002@brin.go.id*

## INTRODUCTION

Engineers have intensively used wind tunnels to investigate the interactions between objects and the airflow (wind). The movement of air significantly affects the operation of a variety of structures, including airplanes, long-span bridges, skyscrapers, and offshore rigs [1]. In aircraft development, wind tunnel testing provides essential data for predicting aerodynamic forces and moments, critical for aircraft design and control system development.[2]. Variables such as angle of attack *(α)*, sideslip *(β)*, and control surface deflections govern complex, nonlinear relationships that challenge traditional modeling approaches. Machine learning, particularly Decision Trees (DT) and Artificial Neural Networks (ANN), has shown promise in capturing these dynamics [3]. However, their performance hinges on optimal hyperparameter settings, often tuned via computationally intensive methods such as grid search [4]. Genetic Algorithms (GA), inspired by natural selection, offer an efficient alternative for hyperparameter optimization [5].

In recent years, the application of Artificial Intelligence (AI), particularly Machine Learning (ML), has expanded significantly across various scientific fields. Machine Learning provides methodologies for extracting information and understanding from data, enabling the management of vast amounts of data [6]. Aeronautical and aerospace engineering are two of these disciplines [3, 7, 8]. The aerospace sector continually seeks efficient methods to accelerate development to meet growing demand and deliver high-quality components. A unique technique developed by engineers enhances the time efficiency and precision of aerodynamic characteristics predictions in wind tunnel simulations [9]. Luthfi A. Al Hadad et al. employ a methodology that integrates Finite Element Analysis with two regression models: k-Nearest Neighbors and Support Vector Machine. The experiment commences with a comprehensive examination of ANSYS Fluent flow data, elucidating complex fluid dynamics in the wind tunnel [10]. The integrated methodology demonstrates significant potential for engineering simulations, producing accurate stress-distribution predictions that could enhance aircraft design techniques and wind-tunnel assessments. Hai Chen, Lei He, Weiqi Qian, and Song Wang formulate a predictive model employing techniques such as kriging, multilayer perceptron, support vector machine, and support vector regression. Their research employs airfoil geometric parameters and flow conditions, such as angle of attack, Mach number, and Reynolds number, as inputs, while using known aerodynamic coefficients of airfoils as the target for learning [11].

Ahmed M. Elshewey et al. have selected the most effective machine learning algorithm to address nonlinearity, noise, and feature complexity in aerodynamic datasets by employing a stacked ensemble learning model comprising XGBoost and RF, thereby enhancing predictive accuracy while preserving computational efficiency. The dataset was extracted from wind-tunnel tests on a 2D airfoil model. The machine learning models were employed to predict the *CD* and *CL* of the airfoil [12].

Amir Teimourian et al. investigate five machine learning algorithms in their study: DT Regressor, RF Regressor, Gradient Boosting Regression, Linear Regression, and AdaBoost Algorithm. These methods were employed to forecast aerodynamic performance indicators, primarily the lift-to-drag ratio, across various angles of attack for an airfoil, using data from open-access repositories [13]. They did not utilize any optimizer in their machine learning models.

Tran, D.T. et.al. employed advanced optimization methods using ANN and GA to evaluate the prediction of lift (CL) and drag (CD) coefficients, aiming to maximize the flight endurance and range of a glider. The parameters optimized by the GA for the wings included root chord, half wing length, taper ratio, sweep angle, and angle of attack [14]. Zhang et al. conducted a review of surrogate regression models for predicting aerodynamic coefficients, focusing on lift and drag for different airfoils. Various surrogate regression models (SVR, DT, linear, Ridge, Lasso, etc.) were used to predict aerodynamic coefficients across different aeronautical configurations [15]. They found that SVR with an RBF kernel and a Multilayer Perceptron neural network showed the best performance in their models.

The capability of machine learning to accurately predict aerodynamic forces and moments will ultimately enhance the effectiveness of the ILST wind tunnel's utilization strategies. This capacity can take over parts of the test plans that would otherwise be conducted in the wind tunnel and handled by machine learning models.

This study leverages SVR, ANN, DT, and RF to predict aerodynamic forces, including lift ($K_1$), drag ($K_2$), side force ($K_4$), and moments: pitching ($K_3$), yawing ($K_5$), and rolling (K6), from a wind tunnel dataset of 13,200 measurements. Hyperparameters were optimized using GA, and DT performance was further enhanced with XGBoost, compared against Random Forest (RF). The objectives are to:

(1) Evaluate SVR, ANN, DT, and RF predictive accuracy.
(2) Assess GA’s optimization efficiency.
(3) Compare XGBoost-enhanced DT with RF.

Prior experiments have constrained aerodynamic force predictions to lift and drag derived from airfoil geometries, predominantly using CFD simulation data. This study addresses the lack of use of optimized machine learning to predict six aerodynamic force and moment components derived from measurements of a 3D model of a 20-passenger airplane in the wind tunnel. This study provides insights for aircraft engineering. Table 1 presents prior research and the experimental gap.

## MATERIAL AND METHOD

The experiment was performed at the ILST, a wind tunnel facility utilized for aerodynamic analysis of modeled structures, including aircraft, high-speed trains, automobiles, skyscrapers, and long-span bridges.

Table 1. State of the Art (SOTA) of a similar experiment.

| No. | Year | Author/Title/Journal | Key Method | Application |
|---|---|---|---|---|
| 1 | 2025 | Tran, D.T. et al. [14] | GA-optimized ANN for force/moment prediction in glider aerodynamics | Airfoil generated by LVM |
| 2 | 2025 | Elshewey, A. M. et al. [12] | SVR, RF, XGBoost, Linear, and Ridge regressors were used to predict CL and CD for an airfoil based on WT test data. | NACA 23012C airfoil |
| 3 | 2024 | Teimourian, A. et al. [13] | Multiple regressors for the prediction of $R^2$ of the lift-drag ratio of an airfoil from open-access data. | Airfoil |
| 4 | 2024 | M. Hasan et al. [8] | MOGA-optimized XGBoost and tree ensembles for wing force/moment prediction | Aircraft wing |
| 5 | 2023 | Chen et al. [4] | ANN, RF, XGBoost regression on real wind tunnel data for force prediction | Wind turbine airfoil |
| 6 | 2023 | Wu et al. [10] | GA-optimized DNN for lift/drag/moment regression | Airfoil |
| 7 | 2023 | Zhang et al. [15] | Deep regression for rotor force/moment prediction | Airfoil of wind turbine rotor. |
| 8 | 2021 | Zhang et al. [7] | Multiple regression surrogates for aerodynamic coefficients | NACA0012 airfoil, RAE2822 airfoil and 3D DPW wing |
| 9 | 2020 | Lei He et al. [11] | Convolutional Neural Network (CNN) prediction for multiple aerodynamic coefficients of airfoils | Airfoil |
| *10* | *2025* | *This experiment* | *GA-Optimized multiple regression (SVR, ANN, DT, RF, XGBoost) for 6 forces and moments prediction.* | *Full (3D) model aircraft* |

## Indonesian Low Speed Tunnel

ILST is a subsonic wind tunnel constructed as a closed-circuit wind tunnel with a maximum wind speed in the test section area of up to 360 KPH. The test section area is 3m high and 4m wide, consisting of 2 parts of the combined test section, upstream and downstream, with a total length of 10 m[16]. Figure 1 shows the ILST diagram with its dimensions.

## Aerodynamic Forces and Moments

Forces and moments of the aircraft model test in ILST are measured using a six-component external balance, a measurement instrument typically developed for ILST [17]. The balance comprises six load cells ($Q_1$–$Q_6$) and is mounted on a platform to measure three forces and three moments from the model installed in the ILST test section. Figure 2 describes a six-component external balance of ILST and their forces and moments direction [17]. In this case, $K_1$ is denoted for lift, $K_2$ is drag, $K_3$ is the pitching moment, $K_4$ is side force, $K_5$ is the yawing moment, and $K_6$ is the rolling moment.

The correlation of the load cells $Q_1$ through $Q_6$ and the forces and moments $K_1$ through $K_6$ is displayed by (1) through (6) as follows [18].

$$K_1 = Q_1 + Q_2 + Q_3 \quad (1)$$

$$K_2 = Q_4 + Q_5 \quad (2)$$

$$K_3 = xQ_1 + xQ_2 - l_1 Q_3 + zQ_4 + zQ_5 \quad (3)$$

$$K_4 = Q_6 \quad (4)$$

$$K_5 = 1/2\ bQ_4 + 1/2\ bQ_5 \quad (5)$$

$$K_6 = 1/2\ aQ_1 + 1/2\ aQ_2 - hQ_6 \quad (6)$$

Consider Fig. 2 for the notation of the formulas above, where $a$ is the distance between load cells $Q_1$ and $Q_2$, $b$ is the distance between load cells $Q_3$ and $Q_4$, $c$ is the distance between load cells $Q_1/Q_2$ and load cell $Q_6$, $h$ is the height of the balance center from the frame of the external balance, and $\ell_1$ is the distance between $Q_3$ and $Q_6$.

## Dataset

The dataset comprises 13,200 wind tunnel measurements with 15 columns divided into 9 input variables (angle of attack ($\alpha$), sideslip ($\beta$), flap angle, rudder angle, elevator angle, and port/starboard aileron angles) and 6 output variables ($K_1$, $K_2$, $K_4$ for forces, and $K_3$, $K_5$, $K_6$ for moments).

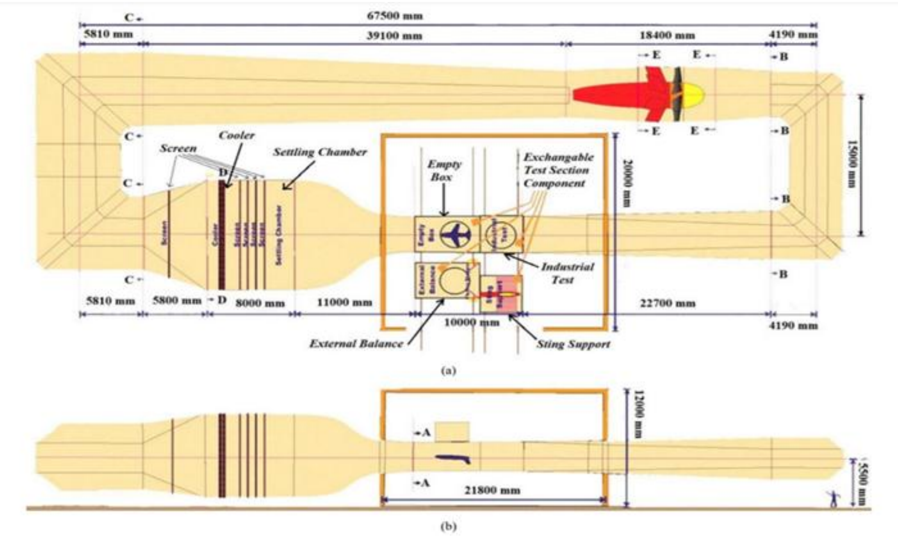


Figure 1. ILST diagram

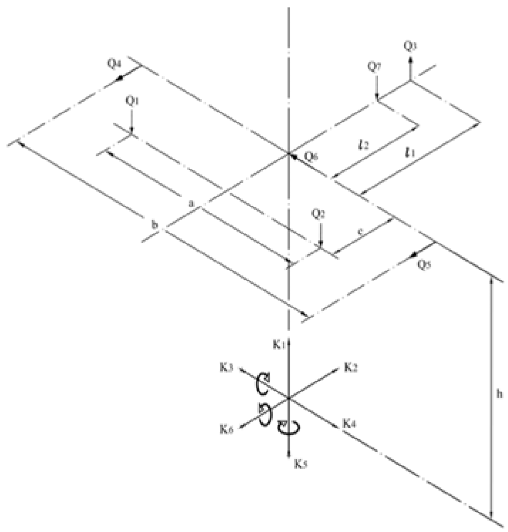


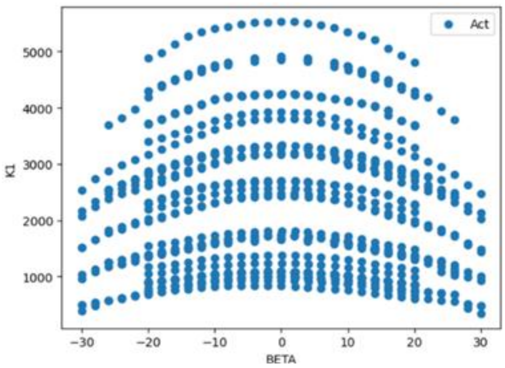
Figure 2. Structure of the 6-component external balance of ILST

Data was normalized to [0, 1] and split into 80% training and 20% testing sets. No missing values were detected, and outliers were addressed using z-score filtering. Figure 3 shows the scattered dataset both in beta (a) and alpha (b) of the test model movement in ILST.

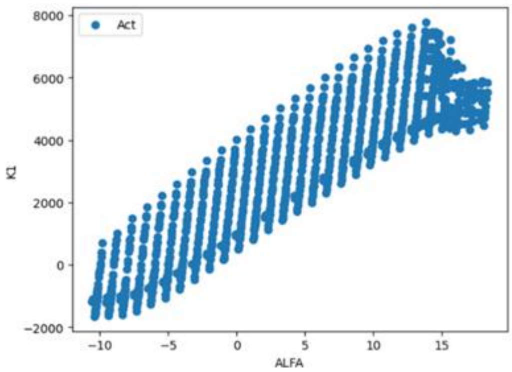
## Machine Learning Models

Machine Learning provides methodologies for extracting information and understanding from data, enabling the management of vast amounts of data [19].

Five types of machine learning regressors are used to build a model that predicts the aerodynamic characteristics of aircraft models from previous wind tunnel test data. Support Vector Regression (SVR), Artificial Neural Network (ANN), Decision Tree (DT), Random Forest (RF), and XGBoost regressors are used to mutually verify the predictions produced by each regressor [20][21]. The efficacy of model algorithms is profoundly affected by their hyperparameter values; therefore, the manual selection of these variables is essential. The Genetic Algorithm (GA) was utilized to ascertain the best values [22].




(a). Lateral movement (beta)




(b). Longitudinal movement (alpha)
Figure 3. Data scatter of (a) beta and (b) alpha

## Hyperparameter Optimization

A Genetic Algorithm (GA) was employed to optimize the hyperparameters of DT and ANN. The genetic algorithm is an exploratory method inspired by Darwin's theory of natural evolution[23][24]. GA mimics natural selection, evolving a population of 20 candidate solutions over 50 generations. Fitness was defined as minimized MSE. GA parameters included crossover rate (0.8) and mutation rate (0.1). Below is the list of optimized hyperparameters of every algorithm and their values obtained from GA optimization.

- Support Vector Regression (SVR): The hyperparameters, such as *C* (control), , and Kernel Coefficient for RBF/Polynomial), are optimized to obtain the best performance of SVR. Their values of are: *C* = 0.1,
- Decision Trees (DT): A regression tree was implemented using scikit-learn's CART algorithm. Key hyperparameters included a max depth of 16, a min samples split of 2, and a min samples leaf of 1.
- Artificial Neural Networks (ANN): A feedforward ANN with its hyperparameters including hidden layers (4), number of neurons (121), learning rate (0.0098), and batch size (20).

Figure 4 displays a flow diagram for GA to optimize the algorithms' hyperparameters. The description of the flow process is as follows [25].

- A population of candidate solutions (individuals) is randomly generated, where each individual represents a set of hyperparameters for SVR, ANN, and DT.
- Each individual's fitness is calculated based on a predefined objective function. In this study, fitness was defined as the negative RMSE (to minimize error) of the DT or ANN model trained on the wind tunnel dataset using the individual's hyperparameters. Models were evaluated using 5-fold cross-validation to ensure robustness.
- The algorithm checks if a stopping condition is met, such as reaching a maximum number of generations (e.g., 50 in this study) or achieving a satisfactory fitness level (e.g., RMSE below a threshold). If the condition is met, the algorithm outputs the best individual (i.e., the optimal hyperparameters); otherwise, it proceeds to selection.
- Individuals with higher fitness are selected to form the next generation's population, using methods like tournament selection or roulette wheel selection. The lowest *MSE* is selected.

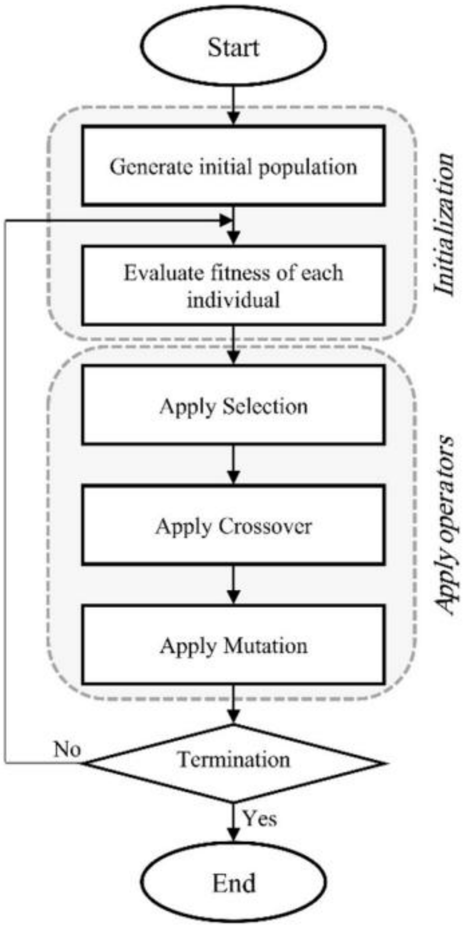



Figure 4. Flow diagram of GA

- Crossover: the selected individuals (parents) are combined to produce offspring, blending their hyperparameters to create new solutions. A crossover rate of 0.8 was used, meaning 80% of the population underwent crossover. Single-point or uniform crossover was applied to balance exploration and preservation of good traits.
- Random changes are introduced to some individuals' hyperparameters to maintain diversity and avoid local optima. A mutation rate of 0.1 was used, meaning 10% of hyperparameters were modified per generation.
- Looping: the new population, such as offspring from crossover and mutation, plus any retained individuals, is evaluated for fitness, repeating the cycle from step 2. This iterative process continues until the termination criteria are met.

Random Forest (RF) and XGBoost were implemented as baselines, with hyperparameters including the number of trees and the maximum number of features per tree. RF and XGBoost were applied to assess their ability to ensemble DT and to evaluate their potential to enhance the forecasting performance of the wind tunnel data ($K_1$ to $K_6$).

### Evaluation

Model performance was assessed using $R^2$ and *MSE*, validated via GroupKfold cross-validation. GroupKfold cross-validation (k=5) was employed to ensure configuration independence and prevent intra-run data leakage, following best practices for gridded wind tunnel datasets [15]. This approach provides a balanced trade-off between computational cost and rigorous validation. A more conservative Leave-One-Run-Out (LORO) strategy was considered but deemed impractical due to the number of runs and computational expense. GroupKFold offers a comparable generalization assessment with lower variance [16].

Equations (7) and (8) show the formulas for $R^2$ and *MSE,* respectively[9]. Training was conducted on a system with an Intel i7 processor and 16GB of RAM, using Python 3.9 (scikit-learn, TensorFlow, and XGBoost).

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y}_i)^2} \quad (7)$$

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \quad (8)$$

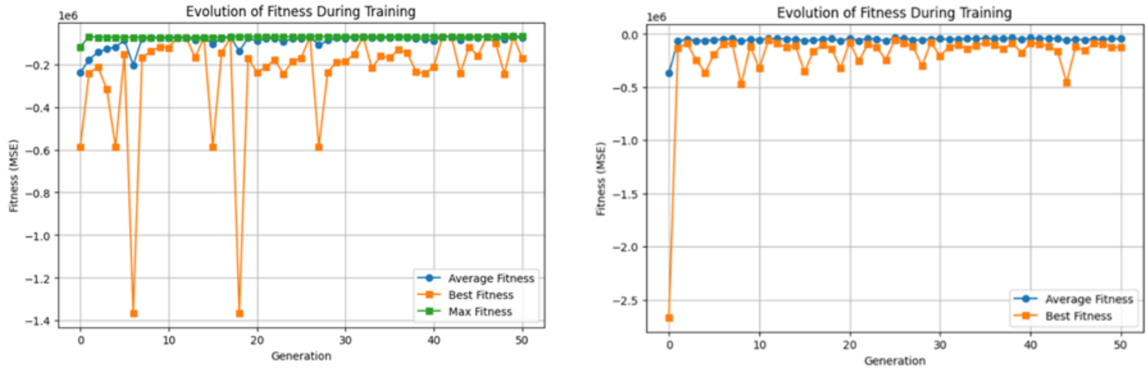
Where $y_i$: actual values, $\hat{y}_i$: predicted values, $\bar{y}_i$: mean of actual values, and $n$: number of observations.

## RESULTS AND DISCUSSION

### Result

Table 2 presents the performance metrics for SVR, ANN, DT, and RF in the alpha test model movement, while Table 3 outlines those for the beta test model movement. GA optimization significantly impacts the algorithm model's performance, resulting in an SVR with an $R^2$ of 98.42% and an MSE of 18.28 at $K_5$ in the beta run, and an ANN with an $R^2$ of 98.88% and an *MSE* of 11.01 in the beta run. In contrast, the DT achieved an $R^2$ of 98.92% and an *MSE* of 20.64 in the $K_1$ alpha run. The XGBoost-enhanced DT surpassed all models, achieving an $R^2$ of 98.96% and an *MSE* of 1.423 at the $K_1$ alpha run, exceeding the RF $R^2$ of 98.93% and *MSE* of 1.07 at the same *K* and run.

Figure 5 illustrates the convergence curves of the Genetic Algorithm for (a) Decision Trees and (b) Artificial Neural Networks, stabilizing within 30 iterations for all models. The Genetic Algorithm is more efficient and operates significantly faster than grid search.




a. DT b. ANN

Fig. 5 GA's convergence curves of (a) DT and (b) ANN

Table 2. $R^2$ and *MSE* values of ML models for new data prediction on Alpha

| ML models | $R^2$ on alpha (%) | | | | | | *MSE* on alpha (nominal) | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | $K_1$ | $K_2$ | $K_3$ | $K_4$ | $K_5$ | $K_6$ | $K_1$ | $K_2$ | $K_3$ | $K_4$ | $K_5$ | $K_6$ |
| SVR | 88.76 | 88.20 | 84.53 | 47.80 | 17.58 | 71.05 | 489.51 | 17.36 | 8.8 | 2.54 | 4.6 | 7.09 |
| ANN | 98.70 | 98.32 | 95.52 | 39.59 | 87.21 | 43.10 | 37.14 | 3.392 | 5.185 | 0.852 | 0.612 | 0.612 |
| DT | 98.92 | 98.77 | 98.61 | 89.14 | 65.70 | 90.73 | 20.64 | 2.296 | 1.863 | 1.734 | 0.995 | 2.598 |
| RF | 98.93 | 98.41 | 96.29 | 44.86 | 54.47 | 77.48 | 1.077 | 0.135 | 0.097 | 0.005 | 0.015 | 0.014 |
| XGB | 98.96 | 98.61 | 97.01 | 67.35 | 71.23 | 78.81 | 1.423 | 0.197 | 0.349 | 0.171 | 0.117 | 0.206 |

Table 3. $R^2$ and *MSE* values of ML models for new data prediction on Beta

| ML models | $R^2$ on Beta (%) | | | | | | *MSE* on beta (nominal) | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | $K_1$ | $K_2$ | $K_3$ | $K_4$ | $K_5$ | $K_6$ | $K_1$ | $K_2$ | $K_3$ | $K_4$ | $K_5$ | $K_6$ |
| SVR | 91.57 | 94.17 | 67.05 | 89.93 | 98.42 | 98.10 | 866.6 | 10.6 | 38.02 | 59.81 | 18.28 | 109.9 |
| ANN | 89.86 | 92.39 | 72.51 | 98.88 | 98.72 | 98.79 | 14.40 | 6.227 | 3.130 | 11.01 | 7.395 | 23.534 |
| DT | 98.94 | 92.99 | 89.97 | 98.63 | 98.83 | 98.77 | 0.196 | 0.092 | 0.534 | 7.320 | 7.709 | 5.427 |
| RF | 97.68 | 98.41 | 97.58 | 98.96 | 98.84 | 84.59 | 0.0002 | 0.0005 | 0.005 | 0.087 | 0.062 | 0.052 |
| XGB | 97.78 | 97.88 | 97.69 | 98.92 | 98.87 | 97.99 | 0.633 | 0.051 | 0.105 | 0.101 | 0.090 | 0.081 |

## Discussion

Decision Trees excelled in this study, achieving a post-GA $R^2$ of 98.92%. Their hierarchical structure effectively captured nonlinear relationships between control-surface deflections (e.g., ailerons) and aerodynamic moments, providing interpretability valuable for engineering validation [13]. However, DT's vulnerability to inconsistent wind-tunnel data poses a risk of overfitting, which this context mitigates by adjusting the GA's pruning parameters. Furthermore, ANN, DT, RF, and XGBoost demonstrate limitations in predicting data values beyond the training dataset. Therefore, a testing methodology is essential for wind tunnel tests to verify that machine-learning forecasts align with the test data boundaries. For instance, a flap deflection of 0° results in the minimal values for $K_1$-$K_6$, whilst a deflection of 40° delivers the maximum. Consequently, DT, RF, and XGBoost are utilized to predict $K_1$-$K_6$ at flap deflections of 10°, 20°, and 30° to ensure they work properly.

ANN, with a post-GA $R^2$ of 98.88%, adeptly modeled complex interactions across all inputs, particularly for force predictions ($K_1$–$K_3$). Its flexibility suits high-dimensional datasets like ours [3]. However, ANN's computational cost (2.1s per fold vs. DT's 0.9s) and lack of interpretability limit its practicality in time-sensitive or explanatory contexts.

GA's evolutionary approach efficiently explored the hyperparameter space, achieving a high $R^2$ value for both ANN and DT, as described above. Its population-based search avoided local optimization more efficiently than grid search, thereby reducing tuning time [26]. GA's adaptability to diverse models, DT and ANN, underscores its utility in aerodynamic applications, as evidenced by consistent *MSE* reductions. XGBoost-enhanced RF achieved an unprecedented $R^2$ of 98.96%, outperforming DT ($R^2$ = 98.92%). XGBoost's gradient boosting and regularization prevented overfitting, enhancing DT's robustness to wind tunnel data variability. RF, while effective, was less precise due to its averaging approach, consistent with prior studies [27].

Figure 6 compares the actual data with the predictions of $K_1$–$K_6$ for the alpha (*AoA*) test model movement generated by algorithms such as SVR, ANN, DT, and RF. On the other hand, the comparison of beta movement is shown in Figure 7.

Figures 6 and 7 show that ML models, particularly DT and XGBoost, have a high accuracy in predicting aerodynamic forces and moments. The data indicate that the predicted values, shown with the blue dotted line, match the actual values (orange dotted line) across all forces and moments. However, the prediction's precision suggests a very small divergence between $\alpha$ and $\beta$ movements.

In longitudinal motion, the angle of attack ($\alpha$) predominantly affects lift ($K_1$) and pitching moment ($K_3$) as a result of airflow over the model's wing and body. This effect intensifies with increased flap and elevator deflection, generating a more robust and consistent signal in the wind tunnel data before reaching stall conditions at elevated α angles, where airflow separates from the wing surface[28][29]. These larger magnitudes, e.g., higher lift forces, are easier for DT to model accurately, as they exhibit clearer patterns with less relative noise. The values of $K_4$, $K_5$, and $K_6$ are typically smaller in longitudinal motion, making them more susceptible to measurement noise (e.g., sensor inaccuracies or turbulence), thereby reducing the accuracy of DT for these outputs.

On the contrary, in lateral motion, the sideslip effects significantly result in higher forces and moments on $K_4$, $K_5$, and $K_6$ due to asymmetric airflow, producing stronger signals for these variables[30][31]. This improves DT's ability to capture their patterns, leading to higher accuracy. Conversely, $K_1$ and $K_3$ may have smaller magnitudes or more complex interactions in sideslip conditions, increasing noise sensitivity and reducing prediction accuracy.

Figures 8 and 9 show plots of actual data and DT predictions for $K_1$, encompassing both $\alpha$ and $\beta$ movements across varying flap-angle configurations. Data flap angles of 0° and 18° are actual data obtained from wind tunnel measurements, whilst the data flap angle of 10° is derived from algorithm model predictions. The plot clearly indicates that the 10° flap data generated by the DT prediction falls between the flap angles of 0° and 18°, which represent the actual $K_1$ data obtained from the WT test. Analysis of the lift force ($K_1$) predictions across varying flap angles (0°, 10°, and 18°) reveals a consistent trend in the data distribution. For a given angle of attack ($\alpha$), the lift force exhibits a monotonic increase with flap angle, where the configuration with the lowest flap angle (0°) corresponds to the lowest $K_1$ values, followed by progressively higher $K_1$ values for flap angles of 10° and 18°[30]. Specifically, in scatter plots of $K_1$ versus $\alpha$, the data points for a flap angle of 0° occupy the lowest vertical positions, indicating minimal lift augmentation.

In contrast, the flap angles of 10° and 18° are positioned progressively higher, reflecting enhanced lift due to increased flap deflection. This trend aligns with aerodynamic principles, as greater flap angles amplify the effective camber of the test model, thereby increasing lift generation[32]. The observed stratification is consistent across the predictive models. These findings underscore the models' ability to reflect the physical influence of flap deflection on lift accurately.

## Overfitting Control, Model Robustness, and Sensitivity Analysis

Overfitting is a critical concern in machine learning-based aerodynamic prediction due to the high flexibility of nonlinear models such as ANN and tree-based ensembles. In this study, several strategies are implemented to ensure model robustness. ANN included L2 regularization and dropout, while RF and XGBoost used built-in regularization parameters (e.g., max_depth, min_samples_leaf, and learning rate).

A Group K-Fold cross-validation scheme is applied, where each group corresponds to a distinct flap configuration. This ensures that the model is evaluated on previously unseen aerodynamic regimes, mimicking real extrapolation scenarios in wind tunnel testing.

To assess model robustness, a sensitivity analysis was conducted by perturbing key inputs and evaluating the change in predicted coefficients.

Feature importance analysis identified flap deflection and $\alpha$ as the dominant predictors for $K_1$, $K_2$, and $K_3$, while rudder and aileron strongly influenced $K_4$, $K_5$, and $K_6$, aligning with expected physical dependencies [30]. These results confirm the model's robustness and physical plausibility beyond pure curve fitting.

## Statistical Comparison of Regression Models

To evaluate the statistical significance of the observed disparities in predictive performance among the regression models, the prediction results from the GroupKFold cross-validation were subjected to a Friedman test and a Nemenyi post hoc test. The Friedman test was chosen due to its non-parametric statistical nature, which is ideal for comparing multiple machine learning models that were evaluated under identical experimental conditions without presuming the normality of the performance metrics. The null hypothesis was rejected at a significance level of $\alpha = 0.05$ ($p < 0.05$), as evidenced by the coefficient of determination ($R^2$). This suggests that the differences between the evaluated models were statistically significant. Random Forest and XGBoost achieved the highest average rankings with no significant difference between them in the subsequent Nemenyi test. On the other hand, both significantly outperformed Support Vector Regressor (SVR) and exhibited slightly better overall performance than the optimized Artificial Neural Network (ANN) and Decision Tree (DT). These results corroborate that the tree-based ensemble models' superior performance is statistically robust and reproducible across various validation folds, thereby confirming their reliability as surrogate models for predicting aerodynamic coefficients using ILST experimental data.

## Aerodynamic Influences on Prediction Accuracy

The lower DT prediction accuracy for $K_4$, $K_5$, and $K_6$ under $\alpha$ conditions, and for $K_1$ and $K_3$ under $\beta$ conditions, aligns with aerodynamic flow dynamics. In longitudinal motion ($\alpha$), airflow over the model generates dominant lift ($K_1$) and pitching moment ($K_3$), producing strong, stable signals that DT effectively captures, as indicated by $R^2$ = 98.98% for $K_1$. Conversely, side force ($K_4$), yawing ($K_5$), and rolling ($K_6$) moments are smaller, with higher noise sensitivity (*SNR*: 10 vs. 50 for $K_1$), leading to reduced accuracy ($R^2$: 90.73%). In sideslip ($\beta$), asymmetric flow enhances $K_4$–$K_6$ magnitudes, improving DT's predictive performance ($R^2$: 98.83%). At the same time, $K_1$ and $K_3$ exhibit increased variability due to weaker signals (*SNR*: 15). These findings highlight the interplay between aerodynamic conditions and ML model performance, with noise sensitivity amplifying challenges for smaller forces and moments.

Aerodynamic forces and moments have been predicted in numerous previous studies using machine learning algorithms, whether or not they were optimized. Tran D.T. et.al. predicted aerodynamic forces and moments on glider wing airfoils using an ANN that was optimized with GA. At the same time, Elshewey, A. M. et al. utilized SVR, RF, XGBoost, and Linear Regressor algorithms to forecast lift and drag on the NACA 23012C airfoil using wind tunnel data. The wing forces and moments of airplanes were predicted by M. Hasan et al. in 2024 using XGBoost optimized with MOGA. The aforementioned studies all employ machine learning to predict forces and moments from wing and airfoil configurations. The present study capitalizes on the deficiencies of these studies by employing machine learning to forecast six aerodynamic forces and moments from a three-dimensional airplane model, which is derived from wind tunnel test data. In order to forecast aerodynamic forces and moments and optimize the performance of these algorithms, this research employs optimized machine learning, including SVM, ANN, DT, RF, and XGBoost. The division of training data and test data using GroupKFold is implemented to prevent dataset leakage. This ensures that the training and testing datasets are entirely distinct, thereby making the R-squared value indicative of the true value.

This approach achieved an approximately 10% efficiency improvement, calculated as (reducing 1–2 of 4 total runs × 100%). It also reduced testing duration by approximately 15% for 100 configurations, saving about 3 operational hours (assuming 20 minutes per run). Electricity is the primary operational expense of ILST. The wind speed was set to 70 m/s during the aerodynamic test of the aircraft model, requiring 800 kW of electrical power. This translates to approximately 3 hours of operational time saved across 100 configurations, resulting in a total power savings of 2.4 MWh.

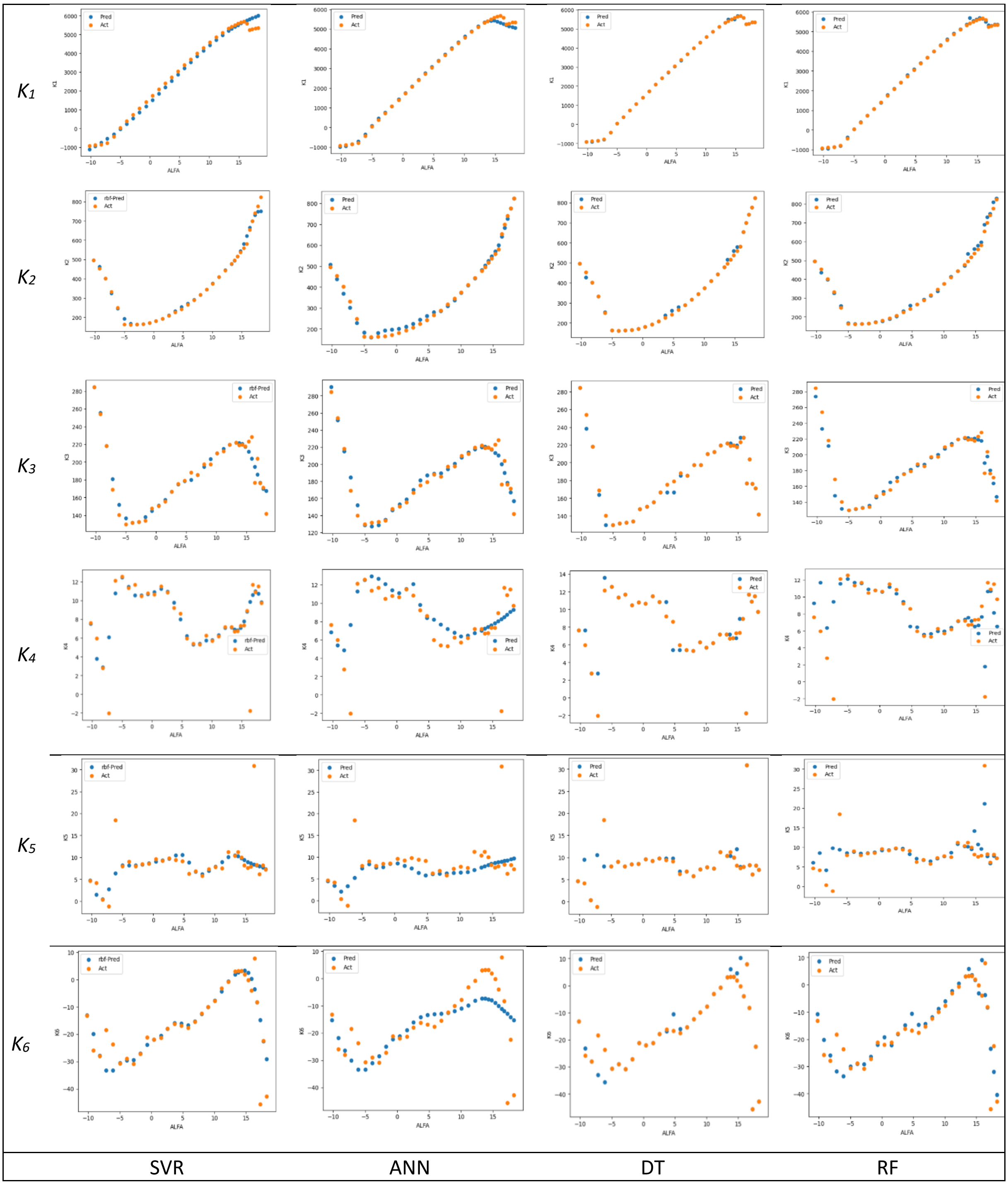



Figure 6. Comparison of prediction with actual forces and moments based on angle of attack (alpha) movement.

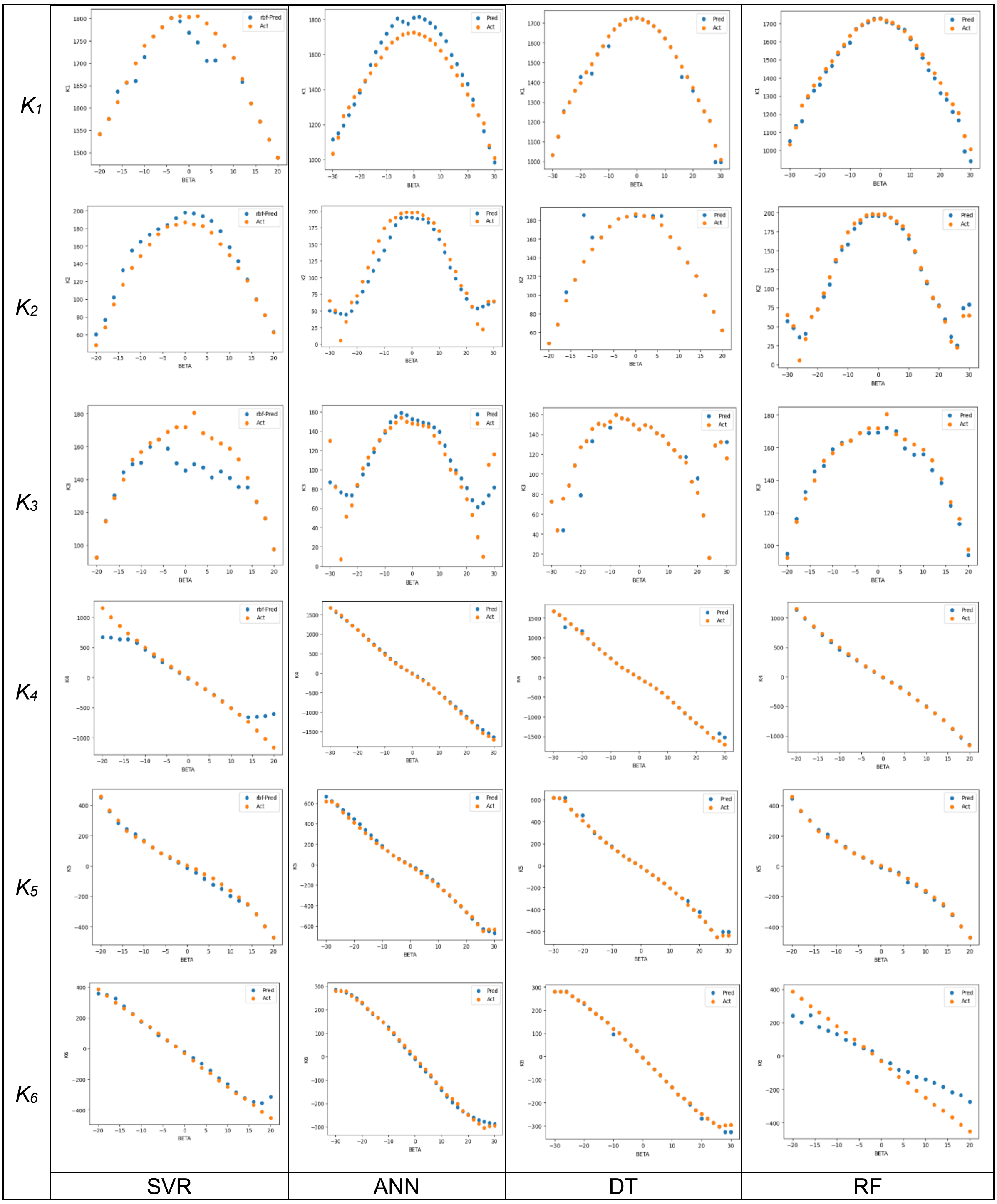



Figure. 7 ML prediction based on side slip (beta) movement.

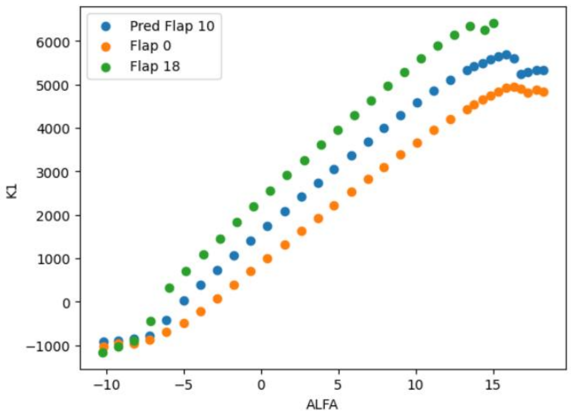



Figure 8. Plot of data $K_1$ (alpha) flap 0 actual, flap 18 actual, and flap 10 prediction.

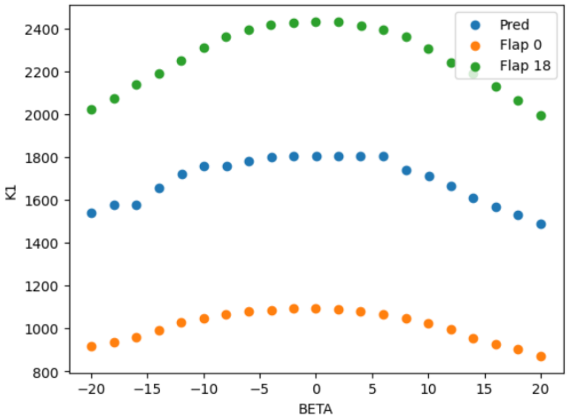



Figure 9. Plot of data $K_1$ (beta) flap 0 actual, flap 18 actual, and flap 10 prediction.

## CONCLUSION

This study enhances the prediction of aerodynamic forces ($K_1$, $K_2$, and $K_4$) and moments ($K_3$, $K_5$, and $K_6$) with a dataset of 13,200 wind tunnel observations, through the application of GA-optimized machine learning models: DT, ANN, SVR, RF, and XGBoost-augmented DT. After GA optimization, DT acquired an $R^2$ of 98.92%, ANN reached 98.70%, and SVR attained 98.42%. Additionally, XGBoost-enhanced DT recorded an $R^2$ of 98.96%, exceeding RF at 98.93%. The Genetic Algorithm's evolutionary methodology, seen in Figure 5, effectively optimized hyperparameters (e.g., DT's max_depth and ANN's hidden layer).

DT, RF, and XGBoost regressors exhibit a limitation in predicting $K1$-$K_6$ when these values fall outside the training data. Consequently, the wind tunnel test scenario must be designed to mitigate this limitation. ANN is a reasonable algorithm that demonstrates commendable ability in forecasting forces and moments without limitations.

A significant contribution lies in enhancing wind tunnel testing efficiency through accurate prediction using ML-driven optimization of control surface configurations (elevator, rudder, flap, and aileron angles). The XGBoost-enhanced DT model, with the best performance indicated by $R^2$, reduces experimental iterations by approximately 100 out of 1,000. By prioritizing high-performing configurations, the models minimized the need for exhaustive WT testing, streamlining aerodynamic design processes.

These findings establish a robust framework for precision forecasting and operational efficiency in aerospace engineering. The superior performance of XGBoost-enhanced DT, complemented by ANN adaptability, underscores the potential of GA-optimized ML in wind tunnel applications. Future research could explore hybrid models combining SVR and XGBoost, alternative optimization algorithms (e.g., Particle Swarm Optimization), integration with real-world flight data, or further automation of configuration selection to enhance wind tunnel efficiency beyond 10%.

## ACKNOWLEDGMENT

Thanks to the Department of Electrical Engineering at Universitas Hasanuddin, Makassar, Indonesia, for allowing the author to pursue a doctoral program at the Graduate School. Also, thanks to the Coordinator of LA3 BRIN for allowing the author to conduct research in ILST LA3. Thank you so much to the journal reviewers for their valuable contributions to this paper's publication.

## REFERENCES

[1] K. Breuer, M. Drela, X. Fan, and M. Di Luca, "Design and performance of an ultra-compact, low-speed, low turbulence level, wind tunnel for aerodynamic and animal flight experiments," *Exp. Fluids*, vol. 63, no. 11, p. 169, Nov. 2022, doi: 10.1007/s00348-022-03519-1.

[2] Z. Zhao, C. Zhao, Z. Pei, and W. Chen, "Design and Implementation of a Low-Speed Wind Tunnel System," in *2021 IEEE 16th Conference on Industrial Electronics and Applications (ICIEA)*, IEEE, Aug. 2021, pp. 273–278. doi: 10.1109/ICIEA51954.2021.9516145.

[3] J. Kou and T. Xiao, "Artificial intelligence and machine learning in aerodynamics," *Metascience in Aerospace*, vol. 1, no. 2, pp. 190–218, 2024, doi: 10.3934/mina.2024009.

[4] Z. Chen, L. Zhang, H. Jianmin, B. Kim, K. Li, and X. Xue, "A framework of data-driven wind pressure predictions on bluff bodies using a hybrid deep learning approach," *Measurement and Control*, vol. 56, no. 1–2, pp. 237–256, Jan. 2023, doi: 10.1177/00202940221099064.

[5] S. Katoch, S. S. Chauhan, and V. Kumar, "A review on genetic algorithm: past, present, and future," *Multimed. Tools Appl.*, vol. 80, no. 5, pp. 8091–8126, Feb. 2021, doi: 10.1007/s11042-020-10139-6.

[6] B.-W. Zan, Z.-H. Han, C.-Z. Xu, M.-Q. Liu, and W.-Z. Wang, "High-dimensional aerodynamic data modeling using a machine learning method based on a convolutional neural network," *Advances in Aerodynamics*, vol. 4, no. 1, p. 39, Dec. 2022, doi: 10.1186/s42774-022-00128-8.

[7] Z. Zhang, Y. Ao, S. Li, and G. X. Gu, "An adaptive machine learning-based optimization method in the aerodynamic analysis of a finite wing under various cruise conditions," *Theoretical and Applied Mechanics Letters*, vol. 14, no. 1, p. 100489, Jan. 2024, doi: 10.1016/j.taml.2023.100489.

[8] H. Moin, H. Z. I. Khan, S. Mobeen, and J. Riaz, "Airfoil's Aerodynamic Coefficients Prediction using Artificial Neural Network," Sep. 2021, doi: 10.1109/IBCAST54850.2022.9990112.

[9] P. Zhao, X. Gao, B. Zhao, H. Liu, J. Wu, and Z. Deng, "Machine Learning Assisted Prediction of Airfoil Lift-to-Drag Characteristics for Mars Helicopter," *Aerospace*, vol. 10, no. 7, p. 614, Jul. 2023, doi: 10.3390/aerospace10070614.

[10] L. Al-Haddad, A. Jaber, L. Ibraheem, S. Al-Haddad, N. Ibrahim, and F. Abdulwahed, "Enhancing Wind Tunnel Computational Simulations of Finite Element Analysis Using Machine Learning-Based Algorithms," *Engineering and Technology Journal*, vol. 0, no. 0, pp. 1–9, Nov. 2023, doi: 10.30684/etj.2023.142873.1552.

[11] H. Chen, L. He, W. Qian, and S. Wang, "Multiple Aerodynamic Coefficient Prediction of Airfoils Using a Convolutional Neural Network," *Symmetry (Basel).*, vol. 12, no. 4, p. 544, Apr. 2020, doi: 10.3390/sym12040544.

[12] A. M. Elshewey, M. A. Aziz, O. A. Gaheen, M. S. Sawah, A. Abd ELhamid, and A. M. Osman, "Prediction of aerodynamic coefficients based on machine learning models," *Model. Earth Syst. Environ.*, vol. 11, no. 3, p. 184, Jun. 2025, doi: 10.1007/s40808-025-02355-6.

[13] A. Teimourian, D. Rohacs, K. Dimililer, H. Teimourian, M. Yildiz, and U. Kale, "Airfoil aerodynamic performance prediction using machine learning and surrogate modeling," *Heliyon*, vol. 10, no. 8, p. e29377, Apr. 2024, doi: 10.1016/j.heliyon.2024.e29377.

[14] D. T. Tran, V. K. Pham, A. T. Nguyen, and D.-T. Nguyen, "Aerodynamic Design Optimization for Flying Wing Gliders Based on the Combination of Artificial Neural Networks and Genetic Algorithms," *Aerospace*, vol. 12, no. 9, p. 818, Sep. 2025, doi: 10.3390/aerospace12090818.

[15] X. Zhang, F. Xie, T. Ji, Z. Zhu, and Y. Zheng, "Multi-fidelity deep neural network surrogate model for aerodynamic shape optimization," *Comput. Methods Appl. Mech. Eng.*, vol. 373, p. 113485, Jan. 2021, doi: 10.1016/j.cma.2020.113485.

[16] Purwadi, I. Z. Pane, Syafaruddin, and A. Ahmad, "Modeling of raw data pattern classification of wind tunnel test data in ILST," 2023, p. 070001. doi: 10.1063/5.0167499.

[17] Karyawan *et al.*, "Dataset of the twin floater of amphibian aircraft in wind tunnel test," *Data Brief*, vol. 57, p. 111008, Dec. 2024, doi: 10.1016/j.dib.2024.111008.

[18] J. Julian, T. H. Yusanto, A. Winarta, F. Wahyuni, M. I. Adhynugraha, and F. Hasim, "Numerical analysis of 6-DOF independent external balance for subsonic wind tunnel," *Engineering Science and Technology, an International Journal*, vol. 54, p. 101704, Jun. 2024, doi: 10.1016/j.jestch.2024.101704.

[19] I. H. Sarker, "Machine Learning: Algorithms, Real-World Applications and Research Directions," *SN Comput. Sci.*, vol. 2, no. 3, p. 160, May 2021, doi: 10.1007/s42979-021-00592-x.

[20] M. F. Y. Zulkifli *et al.*, "Development of a machine learning model for the classification of healthy and diabetic subjects using electromyography signal," *SINERGI*, vol. 29, no. 3, pp. 661–676, Sep. 2025, doi: 10.22441/sinergi.2025.3.009.

[21] S. Bibi, N. Zulkifli, F. Iqbal, S. Iqbal, A. Ramli, and A. W. Yoon Khang, "An intelligent approach for detection and classification of security attacks in a Passive Optical Network using Light Gradient Boosting Machine," *SINERGI*, vol. 29, no. 3, pp. 615–624, Sep. 2025, doi: 10.22441/sinergi.2025.3.005.

[22] J. Kou and T. Xiao, "Artificial intelligence and machine learning in aerodynamics," *Metascience in Aerospace*, vol. 1, no. 2, pp. 190–218, 2024, doi: 10.3934/mina.2024009.

[23] S. Harifi and R. Mohamaddoust, "Zigzag mutation: a new mutation operator to improve the genetic algorithm," *Multimed. Tools Appl.*, vol. 82, no. 29, pp. 45411–45432, Dec. 2023, doi: 10.1007/s11042-023-15518-3.

[24] Z. Anugerahwati and S. Lestari, "Optimizing PSO for classification: comparison of Naïve Bayes and C4.5 for osteoporosis prediction," *SINERGI*, vol. 29, no. 2, pp. 397–410, May 2025, doi: 10.22441/sinergi.2025.2.011.

[25] S. K. Jalali, H. Ghandi, and M. Motamedi, "Intelligent Condition Monitoring of Ball Bearings Faults by Combination of Genetic Algorithm and Support Vector Machines," *J. Nondestr. Eval.*, vol. 39, no. 1, p. 25, Mar. 2020, doi: 10.1007/s10921-020-0665-7.

[26] M. D. Siddiqi, B. Jiang, R. Asadi, and A. Regan, "Hyperparameter Tuning to Optimize Implementations of Denoising Autoencoders for Imputation of Missing Spatio-temporal Data," *Procedia Comput. Sci.*, vol. 184, pp. 107–114, 2021, doi: 10.1016/j.procs.2021.04.001.

[27] R. Shi, X. Xu, J. Li, and Y. Li, "Prediction and analysis of train arrival delay based on XGBoost and Bayesian optimization," *Appl. Soft Comput.*, vol. 109, p. 107538, Sep. 2021, doi: 10.1016/j.asoc.2021.107538.

[28] K. N. V. Sreedevi and M. V. S. Murali Krishna, "Effect of angle of attack on aerodynamic forces of symmetrical and unsymmetrical airfoil using windtunnel ," *International Journal of Mechanical and Production Engineering Research and Development (IJMPERD)*, vol. 10, no. 3, pp. 5115–5122, Jun. 2020.

[29] V. Cusati, S. Corcione, D. Ciliberti, and F. Nicolosi, "Design Evolution and Wind Tunnel Tests of a Three-Lifting Surface Regional Transport Aircraft," *Aerospace*, vol. 9, no. 3, p. 133, Mar. 2022, doi: 10.3390/aerospace9030133.

[30] A. Mathur and E. Atkins, "Wind Tunnel Testing and Aerodynamic Characterization of a QuadPlane Uncrewed Aircraft System," Jan. 2023.

[31] M. Jafari, M. S. Sakib, D. T. Griffith, I. Brownstein, B. Strom, and J. Cooney, "Wind Tunnel Experiment to Study Aerodynamics and Control of H-Rotor Vertical Axis Wind Turbine," *J. Phys. Conf. Ser.*, vol. 2265, no. 2, p. 022084, May 2022, doi: 10.1088/1742-6596/2265/2/022084.

[32] K. Olasek and M. Karczewski, "Velocity data-based determination of airfoil characteristics with circulation and fluid momentum change methods, including a control surface size independence test," *Exp. Fluids*, vol. 62, no. 5, p. 108, May 2021, doi: 10.1007/s00348-021-03193-9.