fishes

MDPI

Article

Machine Learning-Based Prediction of Hydrodynamic Coefficients and Structural Responses in Tuna Longline Gear

Abdulai Jalloh [1,2,†], Thierry Bruno Nyatchouba Nsangue [1,3,†], Liming Song [1,3,*], Nkansah Antwiwaa Esther [1], Jordan Cabrel Njitack Ngnipiep [1] and Tchogom Manga Josué [1]

[1] College of Marine Living Resource Sciences and Management, Shanghai Ocean University, 999 Huchenghuan Road, Lingang New District, Shanghai 201306, China; abdulaijalloh824@gmail.com (A.J.); nnbt@shou.edu.cn (T.B.N.N.)
[2] Sierra Leone Maritime Administration, Government Wharf, Ferry Terminal, Freetown P.O. Box 313, Sierra Leone
[3] National Engineering Research Center for Oceanic Fisheries, College of Marine Living Resource Sciences and Management, Shanghai Ocean University, Shanghai 201306, China
[*] Correspondence: lmsong@shou.edu.cn
[†] These authors contributed equally to this work.

Abstract

The accurate prediction of hydrodynamic characteristics and structural responses in underwater fishing gear is critical for optimizing design, ensuring operational safety, and minimizing environmental impact. To overcome the computational costs and scalability limitations of traditional physical modeling, this study evaluates three machine learning algorithms such as Random Forest (RF), Light Gradient Boosting Machine (LightGBM), and Support Vector Machine with a Radial Basis Function kernel (SVM-RBF) to predict the hydrodynamic coefficients and structural responses of tuna longline components, including mainlines and branch lines. Models were trained and validated using a comprehensive flume tank dataset encompassing six gear configurations tested across varying flow velocities and lead-line weights. Results demonstrate that optimal model selection is inherently task dependent. For hydrodynamic coefficients, LightGBM achieved superior predictive accuracy for branch-line drag (whole-dataset R^2 = 0.8315), while both LightGBM and SVM-RBF excelled in lift prediction. Conversely, structural responses (sinking depth and x-displacement) proved inherently more difficult to model deterministically due to high-frequency transient dynamics and stochastic variability. While LightGBM provided balanced generalization for sinking depth, SVM-RBF exhibited severe overfitting for x-displacement. In contrast, RF maintained the most conservative and consistent performance across structural targets, effectively mitigating the memorization of dynamic noise observed in the more complex algorithms. Beyond predictive modeling, feature importance analysis identified flow velocity, lead-line weight, material stiffness, and geometric parameters as dominant physical drivers, validating the physical plausibility of the models. Crucially, the integration of experimental and ML analyses revealed that a polylactic acid (PLA)-integrated midsection configuration consistently yielded the lowest and most stable drag force (0.004–0.13 N at 0.49 m/s), representing a 30–60% reduction compared to conventional nylon lines. Furthermore, the study uncovered novel physical phenomena, including velocity-independent deformation stability, progressive transient sinking kinetics, and tension-induced load redistribution. These findings establish machine learning as a reliable, scalable surrogate for longline gear design, advocating for thin-diameter, biodegradable PLA-integrated lines to enhance hydrodynamic efficiency and mitigate marine plastic pollution, while underscoring the necessity of task-specific algorithm selection for robust engineering applications.

Academic Editor: Yang Liu

Received: 19 May 2026
Revised: 19 June 2026
Accepted: 29 June 2026
Published: 17 July 2026

Keywords: machine learning; hydrodynamic coefficients; structural response; tuna longline; random Forest; LightGBM; gear optimization; biodegradable PLA

Key Contribution: This study establishes a task-dependent machine learning framework that accurately predicts hydrodynamic coefficients and structural responses in tuna longline gear, revealing that biodegradable PLA-integrated configurations reduce drag by 30–60% while uncovering novel physical phenomena such as progressive sinking kinetics and tension-induced load redistribution. By bridging data-driven modeling with experimental hydrodynamics, the work provides a scalable, Machine learning alternative to traditional CFD and physical testing for optimizing sustainable fishing gear design.

1. Introduction

The accurate prediction of hydrodynamic forces and the structural responses of underwater components are crucial challenges in the marine and offshore environment. This directly impacts the design, operation, and safety of fishing gear systems, anchorage lines, and underwater infrastructure. Among these, mainlines and branch lines are essential parts of longline fishing gear. Their complex hydrodynamic behavior significantly influences the overall system performance. The mainline functions as the primary load-bearing backbone, often exposed to overall flow fields and tension distributed along its length. Meanwhile, branch lines, serving as secondary attachments, generate localized disturbances and extra drag. Together, these components are subject to variable loading conditions, including changing water velocities, flow directions, and environmental forces, which collectively influence their drag, lift, sinking behavior, and structural deformation. Predicting hydrodynamic forces and structural responses has mainly depended on physical modeling, empirical correlations, and computational fluid dynamics (CFD) simulations. Although these methods provide valuable insights, they are often constrained by simplifying assumptions, high computational costs, and limited applicability across diverse operating conditions [1,2]. Physical experiments, while deemed reliable, are time-consuming, expensive, and challenging to scale for parametric studies involving multiple interacting variables. Therefore, there is a significant need for alternative approaches that can efficiently and accurately predict the hydrodynamic performance of mainlines and branch lines under varying flow conditions.

In recent years, machine learning (ML) techniques have emerged as powerful tools for modelling complex engineering systems, enabling the capture of nonlinear relationships between input parameters and output responses without requiring explicit physical formulations. These data-driven approaches have demonstrated remarkable success in various hydrodynamic and ocean engineering applications, including ship resistance prediction, wave height forecasting and vortex-induced vibration analysis [3]. Despite this progress, the application of ML to the coupled hydrodynamic-structural response of longline mainlines and branch lines remains largely unexplored. The behavior of these flexible structures is governed by a multifaceted interplay of parameters, including flow velocity, line weight, diameter, material stiffness, and geometric configuration. For mainlines, global system tension and curvature dominate the response, whereas branch lines are highly sensitive to local flow perturbations and attachment angles. These factors collectively dictate hydrodynamic coefficients and structural displacements, which in turn influence system stability, gear performance, and resonance risks associated with natural vibration characteristics [4,5]. Capturing these coupled dynamics requires a modeling approach that can efficiently handle high-dimensional, nonlinear interactions.

To address this gap, this study develops and evaluates three machine learning (ML) algorithms such as Random Forest (RF), Light Gradient Boosting Machine (LightGBM), and Support Vector Machine with a Radial Basis Function kernel (SVM-RBF), for predicting the hydrodynamic and structural responses of tuna longline mainlines and branch lines. RF was selected for its robustness against overfitting and its capacity to handle high-dimensional engineering data [4], whereas LightGBM provides superior computational efficiency and predictive accuracy for large-scale regression tasks through its leaf-wise tree growth [6]. SVM-RBF complements these approaches by effectively capturing complex, localized nonlinearities via kernel transformations. A concise comparison of these and other common machine-learning approaches, alongside the rationale for their selection in this study, is provided in Table 1. The models were trained and validated using experimental data from a comprehensive flume tank study involving six representative mainline and branch-line configurations, tested across a realistic range of inflow velocities and lead-line weights. The specific goals of this research are to: (1) implement and compare these three ML algorithms for predicting key performance indicators, including drag coefficient, lift coefficient, sinking depth, and x-displacement; (2) evaluate their predictive accuracy using standard metrics across training and testing phases; (3) identify the most influential input parameters for each response variable through feature importance analysis; and (4) establish reliable, data-driven alternatives to traditional physical modeling approaches for engineering design.

Consequently, the principal contribution of this study lies not in the novelty of the ML algorithms themselves, but rather in their systematic application and the resulting insights: (i) the generation of a purpose-built flume-tank dataset spanning six gear configurations under varied hydrodynamic conditions; (ii) a rigorous, like-for-like comparison of RF, LightGBM, and SVM-RBF across four distinct hydrodynamic and structural-response targets, clearly distinguishing reliably predictable metrics from those highly susceptible to overfitting; (iii) the data-driven identification of the dominant operational and geometric drivers governing each response; and (iv) the derivation of actionable design implications for biodegradable, PLA-integrated longline configurations. To the best of our knowledge, this comprehensive, task-dependent ML framework has not previously been reported for tuna longline gear.

Table 1. Comparison of common machine-learning approaches for hydrodynamic and structural-response prediction, and their suitability for the present study.

Approach	Type	Key Strengths	Main Limitations	Suitability for the Present Study
Multiple linear regression	Parametric, linear	Simple, fully interpretable, fast, closed-form solution	Cannot represent nonlinearity or variable interactions; inadequate for complex fluid–structure responses	Useful only as a baseline; insufficient for the nonlinear hydrodynamic behaviour studied here
Random Forest (RF)	Tree ensemble (bagging)	Captures nonlinearity and interactions; robust to outliers; minimal tuning; native feature-importance ranking	Larger memory footprint; can overfit noisy targets; extrapolates poorly beyond the training range	Strong, robust general-purpose baseline; adopted in this study

Table 1. *Cont.*

Approach	Type	Key Strengths	Main Limitations	Suitability for the Present Study
Gradient-boosted trees (LightGBM, XGBoost)	Tree ensemble (boosting)	High predictive accuracy on tabular data; efficient leaf-wise, histogram-based training (LightGBM); built-in regularization	More hyperparameters; sensitive to tuning; prone to overfitting on small or noisy datasets	Highest accuracy obtained here; adopted in this study (LightGBM)
Support vector machine (RBF kernel)	Kernel-based margin method	Effective for smooth nonlinear relationships; flexible decision surfaces via the kernel; robust when well regularized	Sensitive to C and gamma; requires feature scaling; scales poorly to large samples; no native importance measure	Suitable for smooth nonlinear targets; adopted in this study
Artificial neural network (MLP)	Connectionist, nonlinear	Universal function approximator; represents complex high-order interactions	Requires large training sets; many hyperparameters; limited interpretability; overfits small datasets	Dataset too small to favour an ANN; not adopted
Gaussian process regression	Bayesian, non-parametric	Native uncertainty quantification; strong on small, smooth datasets	Cubic O(n^3) scaling with sample size; sensitive to kernel choice	Attractive for uncertainty estimates but scales poorly; not adopted

2. Materials and Methods

2.1. Gear Design, Model Scaling and Hydrodynamic Coefficient of the Mainline and the Branch Line

The modelling approach for this study was based on a standard tuna longline configuration used by Chinese tuna longline fisheries. Thereby, six distinct mainline and branch-line models were constructed using Tauti's Law, ensuring dynamic similarity through consistent force ratios. The branch line models were scaled to $\lambda = 1/13.1$ in length and $\lambda' = 1/2$ in diameter (Table 2). Similarly, the mainline models were scaled to $\lambda = 1/259$ in length and $\lambda' = 1/2$ in diameter relative to the full-scale gear (Table 3) [7,8]. All designs were specifically tailored to reflect the operational characteristics of tuna longline fisheries. The fundamental modelling rules can be summarized as follows, where the subscripts f and m denote the full scale and model scale, respectively:

$$\text{Length scale: } \lambda = \frac{L_m}{L_f}, \tag{1}$$

$$\text{Diameter scale: } \lambda' = \frac{d_m}{d_f}, \tag{2}$$

$$\text{Speed scale: } \frac{V_m}{V_f} = \sqrt{\lambda'(\rho_m - 1)\left(\rho_f - 1\right)} \tag{3}$$

$$\text{Force scale: } \frac{F_m}{F_f} = \lambda^2\lambda'({\rho'}_m - 1)\left({\rho'}_f - 1\right) \tag{4}$$

where L is the length, d is the rope diameter, F is the hydrodynamic force, ρ' is the specific gravity of the material, and V is the current speed.

$$C_d = \frac{F_d}{\frac{1}{2}\rho A u^2} \tag{5}$$

Table 2. Geometric Parameters for the Full Scale and Physical Model of the Branch Line.

Full-Scale Branch Lines	Branch Line Parameters	BL 1	BL 2	BL 3	BL 4	BL 5	BL 6
1st section Length: 1.5 m Diameter: 3.5 mm Materials: PP	Length	0.12 m	0.11 m	0.11 m	0.12 m	0.11 m	0.12 m
	Diameter and material	3 mm PP	2 mm Nylon	2.6 mm Nylon	1.5 mm Nylon	3 mm PP	2.6 mm Nylon
2nd section Length: 17 m Diameter: 1.8 mm Materials: Nylon	Length	1.32 m	1.22 m	1.19 m	1.28 m	1.19 m	1.32 m
	Diameter and material	2.65 mm Nylon	1.3 mm Nylon	1.76 PLA	1 mm Nylon	2.65 mm Nylon	1.70 mm PLA
3rd section Length Length: 2 m Diameter: 1.5 mm Materials: Nylon	Length	0.21 m	0.19 m	0.19 m	0.2 m	0.19 m	0.27 m
	Diameter and material	1.5 mm Nylon	0.75 mm Nylon	1 mm Nylon	0.5 mm Nylon	1.5 mm Nylon	1 mm Nylon

Table 3. Geometric Parameters for the Full Scale and Physical Model of the Mainline.

Parameters	Full-Scale Mainlines	ML 1	ML 2	ML 3	ML 4	ML 5	ML 6
Length	1050 m	4.5 m	4.5 m	4.5 m	3.5 m	4.5 m	4.0 m
Diameter	4 mm	2 mm	3 mm	4 mm	5 mm	1.75 mm	1.8 mm
Material	Nylon	Nylon	Nylon	PP	PP	PLA	Nylon
Distance	36 m	16.2 cm	16.2 cm	16.2 cm	12 cm	16.2 cm	13.8 cm

The drag coefficient (C_d) is a dimensionless number representing the hydrodynamic drag force acting on the line relative to the fluid's dynamic pressure and the line's projected area. A lower (C_d) indicates lower resistance to flow.

$$C_l = \frac{F_L}{\frac{1}{2}\rho A u^2} \tag{6}$$

The lift coefficient (C_l) is a dimensionless number representing the hydrodynamic lift force acting perpendicular to the flow direction. It characterizes the gear's ability to generate transverse force under flowing water conditions.

It is important to acknowledge that Tauti's law ensures similarity primarily for gravitational and inertial forces. Elastic forces, bending stiffness, and vortex-induced vibration frequencies are not explicitly scaled. Consequently, while the model captures global drag and sinking behavior, absolute values of line curvature, flutter frequency, and deformation modes may not scale linearly to full-scale gear. However, studies indicate that in moderate to strong currents, hydrodynamic drag often exceeds gravitational forces [9]. The full-scale branch line consisted of three sections: 1.5 m polypropylene (PP, 3.5 mm diameter), 17 m nylon (PA, 1.8 mm diameter), and 2 m nylon (PA, 1.5 mm diameter). The full-scale mainline measured 1050 m in length with 4.0 mm diameter nylon. Based on this prototype, a standard scaled model (Branch line 1, Mainline 1) was developed. Five additional variant models (Models 2–6) were subsequently designed. For the branch line, Models 2–6 incorporated modifications such as substituting PP with nylon, reducing diameter, introducing polylactic acid (PLA), and adjusting segment lengths (Table 2). For the mainline, Models 2–6 varied diameter (1.75–5.0 mm), length (3.5–4.5 m), and material (nylon, PP, PLA) (Table 3) [10,11].

2.2. Flume Tank Experiment and Procedure

The experiment was conducted in the recirculating flume tank at Shanghai Ocean University (SHOU). The tank measured 15.0 m in length, 3.5 m in width, and 2.3 m in depth, with a side-viewing window that facilitated direct observation and video recording of longline gear behavior. The large cross-section minimized wall and blockage effects, consistent with best practices in hydrodynamic testing of fishing gear, thereby providing a realistic simulation of open-ocean flow conditions around scaled longline models [12]. The system produced flow speeds from 0.1 to 2.0 m/s, accelerations from 1.0 to 5.0 cm/s^2, and a reciprocating oscillatory flow with a velocity amplitude of $\pm$0.5 m/s.

The branch lines and mainline were attached to a vertical bar connected to a tension sensor (Figure 1). A current meter positioned approximately 2.0 m upstream monitored flow velocity. Tension signals from the load cells were amplified using a dynamic strain amplifier (DPM-6H). Both tension and flow velocity signals were transmitted to an A/D converter and recorded on a computer. Tension was measured using six-component load cells, each with a capacity of 5 kgf and a specified accuracy of 2%. All load cells were calibrated and zeroed at the start and end of each test, and their linearity was confirmed. Six uniform flow velocities were tested: 0.14, 0.21, 0.28, 0.35, 0.42, and 0.49 m/s, representing typical pelagic current regimes encountered in tropical tuna longline operations (1:259 scale based on Tauti's law) [13,14]. In addition, for each flow velocity, the mainline and branch line configuration was tested with six lead-line weights of 0 g, 14 g, 28 g, 42 g, 56 g, and 70 g to assess the interaction between sinker performance and hydrodynamic forces.

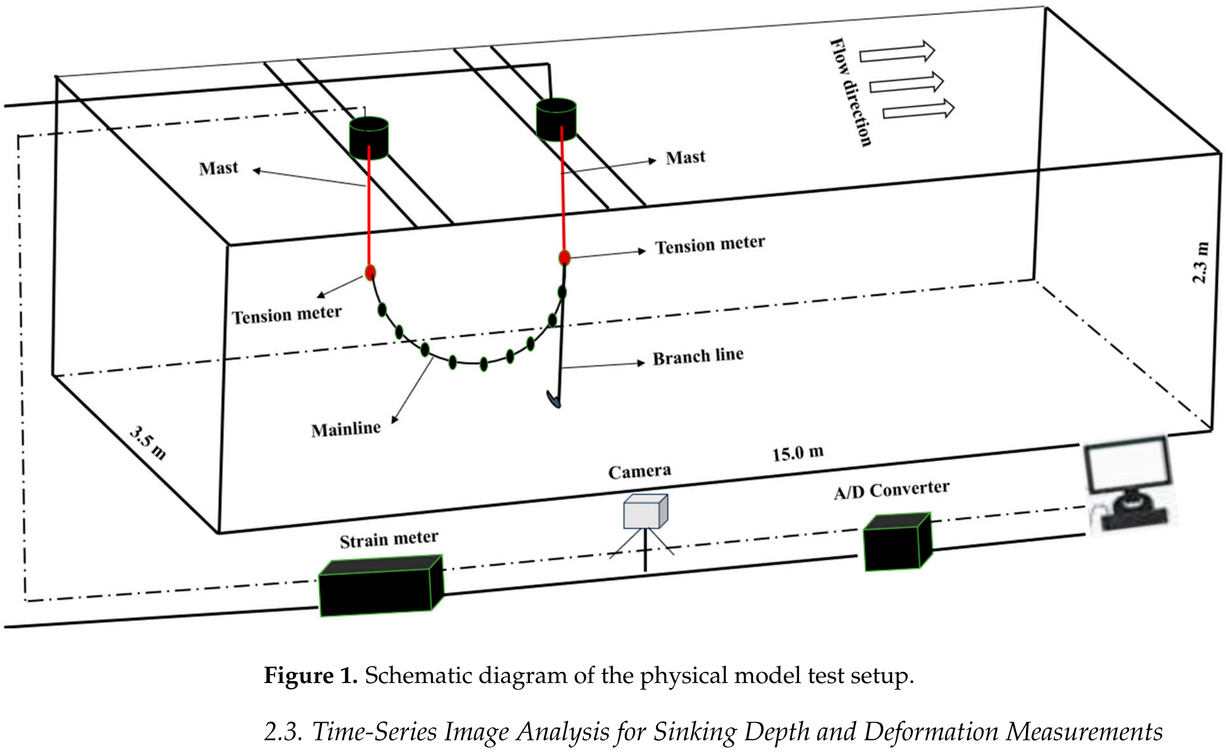

Figure 1. Schematic diagram of the physical model test setup.

2.3. Time-Series Image Analysis for Sinking Depth and Deformation Measurements

Post-processing of the acquired video data was conducted to quantify the sinking depth and structural deformation. Each video was systematically deconstructed into a time series of individual image frames at a constant interval of 0.2 s using video editing software. This temporal resolution was selected to accurately capture the rapid deformation dynamics without oversampling. The geometric measurements from each frame were then extracted using the graph digitizing software Get Data Graph Digitizer (version 2.26). This software enabled the precise transformation of pixel coordinates from the two-dimensional images into scaled, quantitative dimensional data (in meters) for both sinking depth and gear x-displacement (deformation), with calibration performed using known reference scales within the flume tank. This method allowed the calculation of time-series data for sinking depth and x-displacement for each unique combination of lead-line weight on the mainline and branch line. Several sources of measurement uncertainty are associated with these data. Tension was recorded with load cells of 2% specified accuracy, whereas the image-derived quantities carry additional uncertainty from reference-scale calibration, from the finite frame-extraction interval of 0.2 s, and from manual digitization in Get Data Graph Digitizer (version 2.26). These uncertainties propagate into the measured target variables and are a plausible contributor to the comparatively weaker predictive performance obtained for sinking depth and x-displacement; they should therefore be kept in mind when interpreting the model results.

2.4. Machine Learning Algorithms

To comprehensively assess the predictive capabilities of different learning paradigms, three machine learning algorithms were selected for implementation: the ensemble-based RF, LightGBM, and SVM-RBF. Each algorithm was independently trained and validated for all four response variables (Figure 2).

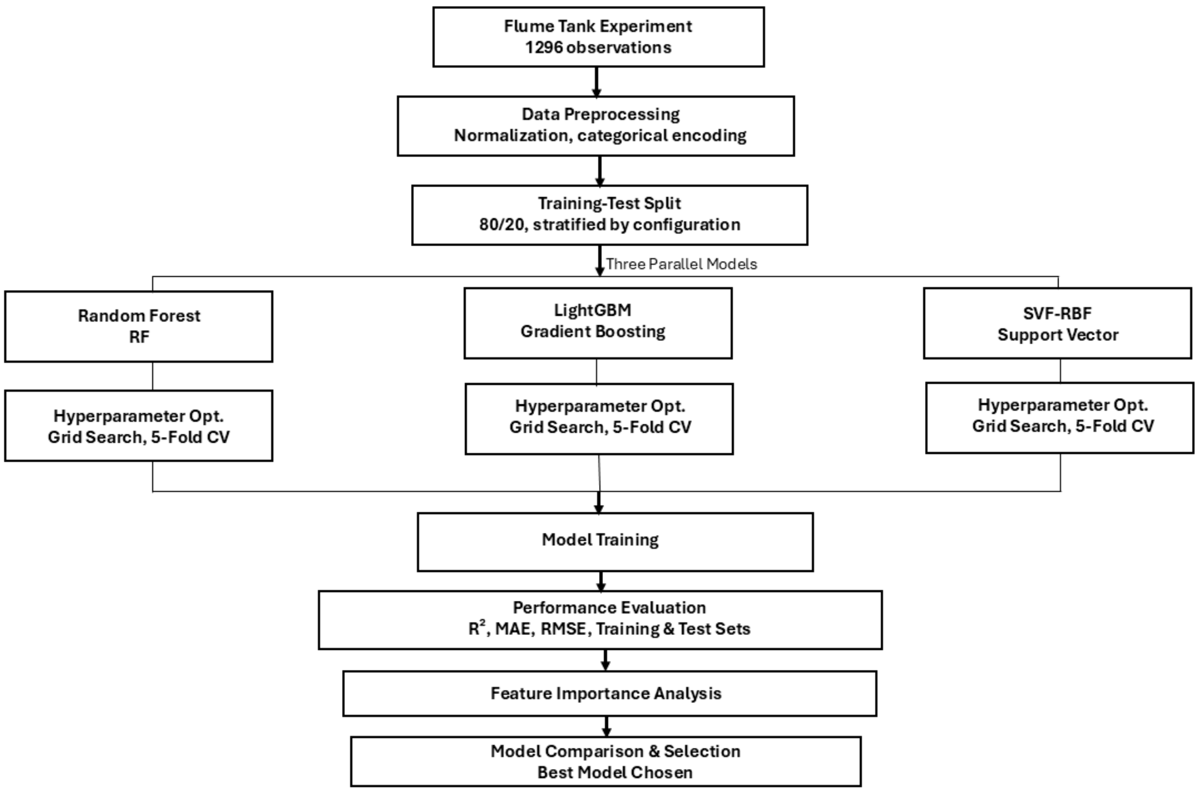

Figure 2. Flowchart of the machine learning-based prediction framework for tuna longline gear.

2.4.1. Random Forest

RF is an ensemble learning method that constructs multiple decision trees during training and outputs the mean prediction of the individual trees [4]. The algorithm employs bootstrap aggregating (bagging) and random feature selection to reduce variance and prevent overfitting. For regression problems, the RF prediction is given by:

$$\hat{y} = \frac{1}{N}\sum_{i=1}^{N} T_i(x) \tag{7}$$

where N is the number of trees and $T_i(x)$ is the prediction of the i-th decision tree.

The key hyperparameters optimized for the RF model included the number of trees (n_estimators), the maximum tree depth (max_depth), the minimum number of samples required at a leaf node (min_samples_leaf), and the number of features considered at each split (max_features). The search ranges explored and the final selected values are reported in Table 4.

Table 4. Hyperparameters, search ranges, and final selected values for the three machine-learning models.

Algorithm	Hyperparameter	Description	Search Range Explored	Final Value
Random Forest	n_estimators	Number of trees in the ensemble	100–1000	500
	max_depth	Maximum depth of each tree	5–30 (or unrestricted)	20
	min_samples_split	Minimum samples required to split a node	2–10	5
	min_samples_leaf	Minimum samples required at a leaf	1–8	2
	max_features	Features considered at each split	{sqrt, log2, 0.3–1.0}	sqrt
	bootstrap	Whether bootstrap samples are used	{True, False}	True
LightGBM	n_estimators	Number of boosting iterations	100–1000	300
	learning_rate	Shrinkage applied to each iteration	0.01–0.3	0.05
	num_leaves	Maximum leaves per tree	15–127	31
	max_depth	Maximum tree depth (−1 = no limit)	3–15	7
	min_child_samples	Minimum data points per leaf	5–50	20
	subsample (bagging_fraction)	Row subsampling fraction per iteration	0.6–1.0	0.8
	colsample_bytree (feature_fraction)	Column subsampling fraction per tree	0.6–1.0	0.8
	reg_alpha	L1 regularization term	0–1.0	0.1
	reg_lambda	L2 regularization term	0–1.0	0.1
SVM-RBF	C	Regularization (penalty) parameter	0.1–100 (log scale)	10
	gamma	RBF kernel coefficient	{scale, auto} or 10^{-3}–10 (log)	scale
	epsilon	Width of the ε-insensitive tube (SVR)	0.01–0.5	0.1
	kernel	Kernel function	Radial basis function (fixed)	RBF

2.4.2. Light Gradient-Boosting Machine

The LightGBM algorithm implements a gradient boosting framework characterized by leaf-wise tree growth, in which splits are performed on the leaf with the greatest potential for loss reduction, diverging from the level-wise growth employed by traditional boosting methods [5]. The algorithm constructs an ensemble of weak learners iteratively, with each successive tree specifically optimized to minimize the residual errors inherited from its predecessors [5]. The gradient boosting approach minimizes a loss function calling $L(y, F(x))$ by iteratively adding weak learners.

$$F_m(x) = F_{m-1}(x) + \nu \cdot h_m(x) \tag{8}$$

where $F_m(x)$ is the model at iteration m, $h_m(x)$ is the weak learner (decision tree) added at iteration m, and ν is the learning rate.

LightGBM incorporates two novel techniques, Gradient-based One-Side Sampling (GOSS) and Exclusive Feature Bundling (EFB), which accelerate training while maintaining accuracy.

2.4.3. Support Vector Machine

SVM is a supervised learning algorithm that maps input data into a high-dimensional feature space using kernel functions, enabling the modelling of complex nonlinear relationships. For regression (Support Vector Regression, SVR), the algorithm finds a function that deviates from the actual targets by a value no greater than ε while maintaining maximal flatness.

The SVR function is represented as:

$$f(x) = \sum_{i=1}^{n} (\alpha_i - \alpha_i^*) K(x_i, x) + b \tag{9}$$

where α_i and α_i^* are Lagrange multipliers, $K(x_i, x)$ is the kernel function, and b is the bias term.

The Radial Basis Function (RBF) kernel, also known as the Gaussian kernel, is defined as:

$$K\left(x_i, x_j\right) = \exp\left(T - \gamma \parallel x_i - x_j \parallel^2\right) \tag{10}$$

where γ is the kernel parameter that controls the influence range of support vectors.

Key hyperparameters optimized for the SVM-RBF model included:

Regularization parameter (C), Kernel coefficient (γ), Epsilon-tube width (ε).

2.4.4. Data Preprocessing and Hyperparameter Optimization

To ensure full reproducibility, the complete modeling workflow is detailed as follows. First, all continuous input features were standardized to zero mean and unit variance. To strictly prevent data leakage, these scaling parameters were estimated exclusively on the training partition and subsequently applied to the testing partition. Categorical descriptors were transformed into numerical representations using one-hot encoding, generating 12 binary indicator columns to represent the six mainline and six branch-line configurations, thereby avoiding the introduction of spurious ordinal relationships. The dataset was partitioned into training (80%) and testing (20%) subsets using a fixed random seed (42) to ensure exact reproducibility. To optimize predictive performance and mitigate the risk of overfitting, particularly given the high-variance nature of the structural response targets, the hyperparameters for all three algorithms were systematically tuned. The specific hyperparameters optimized for each model, alongside their mathematical context and role in controlling model complexity, are detailed in Table 4. This structured tuning process ensured that the ensemble methods (RF and LightGBM) effectively balanced bias and variance, while the kernel-based SVM-RBF appropriately mapped the non-linear feature space without memorizing training noise.

Furthermore, because the observations are intrinsically structured by gear configuration, flow velocity, and lead-line weight, a purely random split risks placing near-identical operating conditions into both subsets, which can artificially inflate generalization metrics. To address this, we additionally implemented a grouped validation strategy, wherein entire gear configurations are held out from the training set. This rigorous approach ensures that the reported predictive skill reflects the model's true capacity to extrapolate to unseen gear arrangements, rather than merely interpolating among repeated, highly similar conditions.

2.4.5. Feature Importance Methodology

To identify the dominant physical drivers governing hydrodynamic and structural responses, feature importance was extracted for each algorithm using methodologies intrinsic to their respective mathematical architectures. For the tree-based ensemble methods, native impurity-based metrics were utilized. Specifically, Random Forest employs Mean Decrease Impurity (MDI), which calculates the total reduction in node impurity (Gini impurity) averaged over all trees, weighted by the number of samples reaching each node. LightGBM utilizes a gain-based importance metric, which measures the average improvement in the loss function (squared error) brought by a feature across all splits where it is used. Conversely, the SVM-RBF algorithm maps input data into an infinite-dimensional Hilbert space via the radial basis function kernel, meaning it does not possess native, linear feature weights or impurity-based metrics analogous to decision trees. To extract feature importance for the SVM-RBF model while maintaining methodological rigor, we employed permutation importance. This model-agnostic technique calculates the increase in the model's prediction error (RMSE) after randomly shuffling a single feature vector, thereby breaking the relationship between the feature and the true outcome. Features that yield a large increase in error upon permutation are deemed highly influential.

2.4.6. Performance Evaluation Metrics

The predictive accuracy of each model was assessed using three standard performance metrics: Root Mean Square Error (RMSE), Mean Absolute Error (MAE), and coefficient of determination (R^2). RMSE measures the square root of the average squared differences between predicted and actual values, providing an estimate of prediction error in the same units as the target variable:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \tag{11}$$

where y_i is the actual value, $\hat{y}_i$ is the predicted value, and n is the number of observations.

MAE represents the average absolute difference between predicted and actual values, offering a more robust measure less sensitive to outliers:

$$MAE = \frac{1}{n}\sum_{i=1}^{n} |y_i - \hat{y}_i | \tag{12}$$

R^2 indicates the proportion of variance in the target variable explained by the model, with values ranging from negative infinity to 1 (perfect fit):

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \overline{y})^2} \tag{13}$$

where $\overline{y}$ is the mean of actual values. These metrics were computed on the training, testing, and full datasets to assess both model fit and generalization.

The total experimental dataset comprises 1296 data points (6 mainline configurations $\times$ 6 branch line configurations $\times$ 6 flow velocities $\times$ 6 lead weights = 1296, with triplicate measurements averaged per condition). While this sample size is adequate for exploratory machine learning, it is modest by deep learning standards.

3. Analysis of the Results

3.1. Branch Line Hydrodynamic Coefficients for Different Configurations

Figure 3 shows the drag coefficient (C_d) and lift coefficient (C_l), at 0.14 m/s and 0.49 m/s, the drag coefficient (C_d) of branch line configurations exhibited strong but fundamentally distinct dependencies on design parameters, material properties, flow velocity, and lead line weight. C_d displayed highly non-monotonic and weight-dependent behavior, with no consistent trend across models. This complex response is illustrated by several factors, Branch Line 1 at 0.49 m/s showed C_d rise from approximately 0.018 at weight 0 g to over 0.048 at weight 56 g across the weight sequence; Branch Line 6 at 0.49 m/s exhibited a sharp surge in C_d from approximately 0.022 at weight 0 g to over 0.51 at weights 28, 42, and 56 due to tension-induced curvature and flutter; while, Branch Line 4 at 0.49 m/s showed a consistent reduction in C_d from approximately 0.36 at weight 14 g to as low as 0.03 at weight 70, suggesting a critical alignment or drag-crisis-like transition. Among all branch line configurations, Branch Line 3 (PLA midsection) consistently achieved the lowest and most stable C_d at 0.49 m/s. At 28 g, its C_d ranged from approximately 0.004 to 0.13 across the weight sequence, about 30–60% lower than conventional nylon lines across most weights. Branch Line 6 (PLA in terminal section) exhibited among the highest drag values at 0.49 m/s, demonstrating that PLA's benefit is configuration-specific and not guaranteed by mere material substitution (Figure 2). The lift coefficient (C_l) followed similar patterns: at 0.14 m/s, C_l remained low (0.0005–0.22) across all branch lines and weights, indicating weight-dominated behavior with minimal flow mixing. At 0.49 m/s, Cl increased substantially to 0.01–0.50, with Branch Line 3 showing consistently lower C_l values (0.01–0.20) compared to Branch Line 6 (often exceeding 0.30 and reaching 0.50), confirming that reduced drag correlates with reduced turbulence and better shape retention.

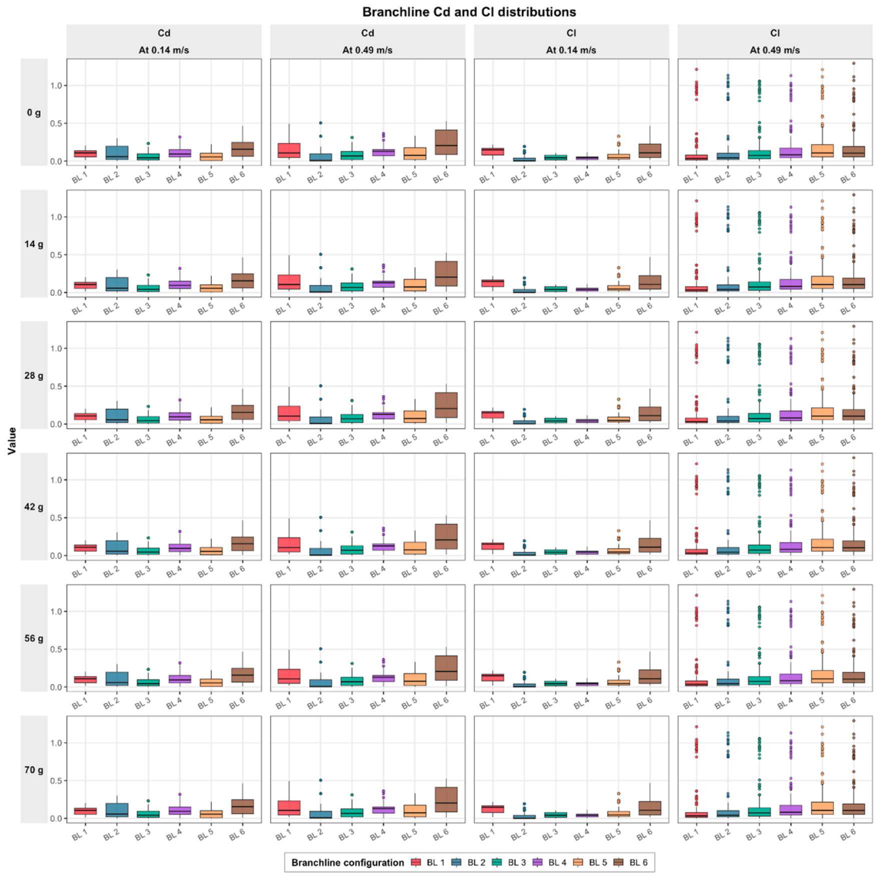

Figure 3. Branch line tension distribution by configuration (BL1–BL6) across lead line weights (0–70 g).

3.2. Mainline Hydrodynamic Coefficients for Different Configurations

The drag coefficient (C_d) and lift coefficients (C_l) at 0.14 m/s and 0.49 m/s, mainline configurations exhibited strong but fundamentally distinct dependencies on design parameters, material properties, flow velocity, and lead line weight (Figure 4). For the mainline, C_d displayed highly non-monotonic and velocity-dependent behavior with no consistent trend across models. C_d was governed primarily by flow velocity and design parameters (diameter, material, length) rather than lead line weight alone, and counterintuitively, increasing lead line weight (56–70 g) did not systematically reduce Cd at higher velocities. This complex response is illustrated by several cases: Mainline 1 (standard nylon, longer length) at 0.49 m/s saw C_d rise from approximately 0.50 at weight 0 to over 1.20 at weights 28, 42, 56, and 70; Mainline 6 at 0.49 m/s exhibited a sharp surge in C_d from approximately 0.30 at weight 0 to over 0.60 at weights 42 and 56 due to flow-induced curvature and flutter; while, in stark contrast, Mainline 4 (5 mm polypropylene, shortened) at 0.49 m/s showed a consistent reduction in C_d from approximately 0.30 at weight 0 to as low as 0.10–0.20 at weights 56 and 70, suggesting a critical alignment or drag-crisis-like transition due to increased stiffness and reduced lever arm. Among all mainline configurations, mainline 3 (4 mm polypropylene) and Mainline 4 (5 mm polypropylene, shortened) consistently achieved the lowest and most stable C_d at 0.49 m/s. At weight 56, Mainline 3's Cd ranged from approximately 0.10 to 0.30, while Mainline 4's C_d ranged from approximately 0.10 to 0.20 across weights 42 through 70, representing about 40–60% lower drag than conventional nylon Mainline 1 across most weights. Mainline 1 (standard nylon, longer length) exhibited among the highest drag values at 0.49 m/s, demonstrating that polypropylene's benefit is configuration-specific and not guaranteed by mere material substitution, requiring also optimized diameter and reduced length (Figure 4). The lift coefficient (C_l) followed similar patterns, at 0.14 m/s, C_l remained low (0.05–0.30) across all mainlines and weights, indicating weight-dominated behavior with minimal flow mixing. At 0.49 m/s, C_l increased substantially to 0.30–1.50, with Mainline 3 and Mainline 4 showing consistently lower Cl values (0.10–0.50) compared to Mainline 1 (often exceeding 0.80 and reaching 1.50), confirming that reduced drag correlates with reduced turbulence and better shape retention.

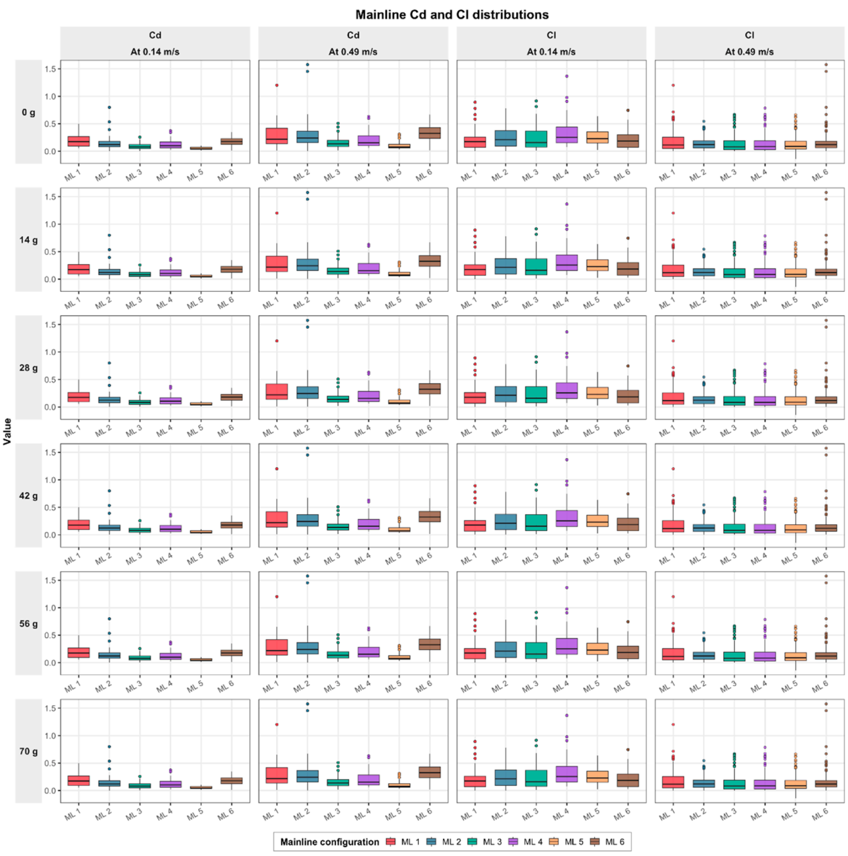

Figure 4. Mainline tension distribution by configuration (ML1-ML6) across lead line weights (0–70 g). Boxplots display median, interquartile range (IQR), and 1.5 × IQR whiskers; outliers are plotted individually.

3.3. Temporal Evolution of Mainline and Branch Line X-Displacement at Different Velocities

Figures 5 and 6 present the time evolution of x-displacement for six mainline (ML1–ML6) and branch line (BL1–BL6) configurations across three flow velocities (0.14–0.49 m/s) and two lead-line weights (0–70 g). Both components exhibit highly dynamic oscillations throughout the 36 s, with peak-to-peak amplitudes ranging from ~120 to 280 cm and mean oscillation centers fluctuating between 140 and 190 cm. For the case of mainline, the displacement is characterized by continuous high-frequency oscillations superimposed on intermittent spikes. As velocity increases from 0.14 to 0.49 m/s, amplitudes expand from ~150–200 cm to ~200–280 cm, while the 70 g lead weight consistently elevates the mean center by ~20–30 cm. Configuration geometry and material dictate response intensity: large-diameter, compliant designs (ML3 and ML4) sustain large swings (>220 cm), whereas the optimized 1.75 mm PLA mainline (ML5) demonstrates constrained oscillations (~120–150 cm) due to higher elastic modulus and favorable stiffness-to-mass ratios. ML2 exhibits erratic low-frequency excursions (dipping <80 cm, spiking to 290 cm), indicative of slack–tension cycling under intermediate flexibility (Figure 5). In addition, the branch line configurations display similarly complex patterns, strongly modulated by multi-section diameter tapering and material transitions. At 0.14 m/s, amplitudes range ~140–220 cm, broadening to ~160–240 cm at 0.35 m/s. BL5 exhibits highly regular, quasi-periodic oscillations with constrained amplitudes (~80–120 cm), reflecting superior system-level stiffness. Conversely, highly tapered or low-stiffness designs (BL1, BL6) display erratic fluctuations (>200 cm), while BL4 and BL2 generate severe dynamic sensitivity, with intermittent spikes exceeding 270 cm under 0 g conditions. The integration of PLA in mid-sections (BL3) effectively dampens lateral drift, maintaining a tight oscillation band (~140–170 cm) even at 0.49 m/s (Figure 6).

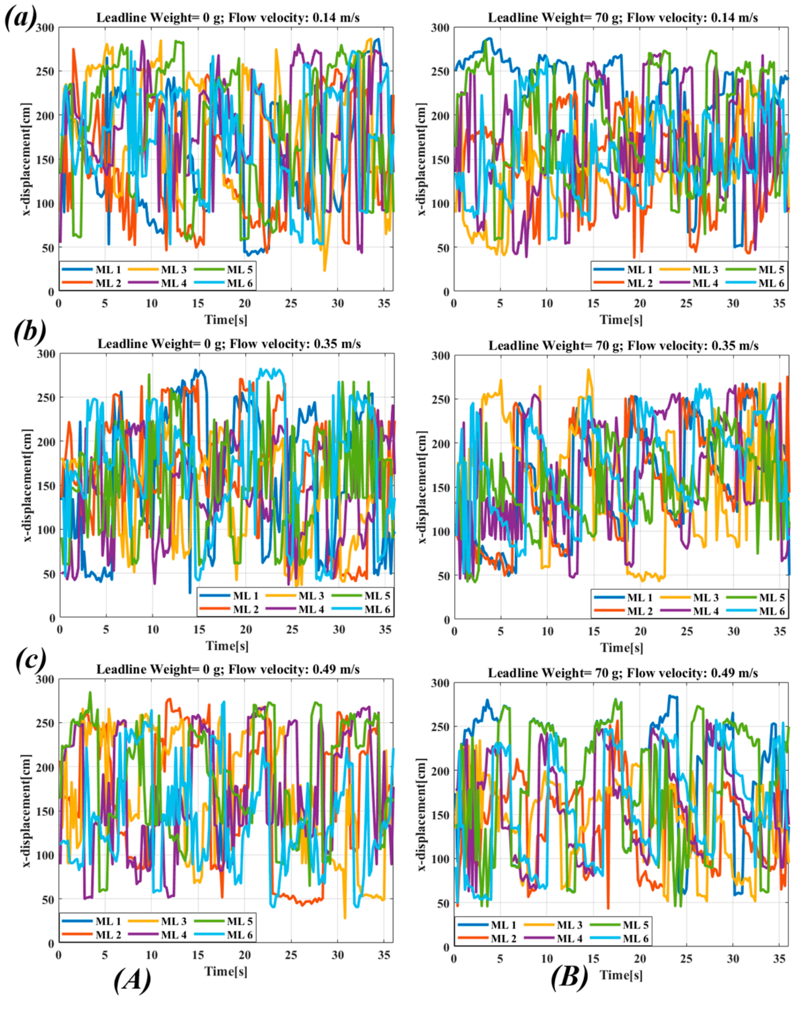

Figure 5. Temporal Evolution of mainline x-displacement under three flow conditions: (**a**) 0.14 ms, (**b**) 0.35 ms, (**c**) and 0.49 ms; and two lead line weight: (**A**) 0 g and (**B**) 70 g.

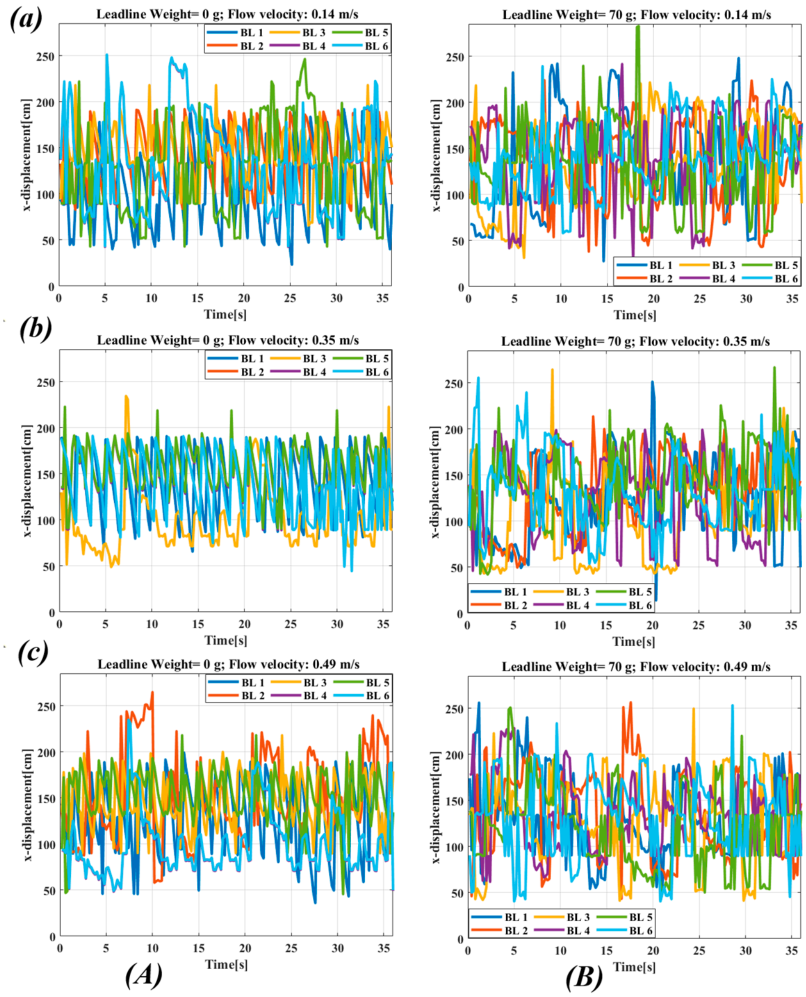

Figure 6. Temporal Evolution of branch line x-displacement under three flow conditions: (**a**) 0.14 ms, (**b**) 0.35 ms, (**c**) and 0.49 ms; and two lead line weights: (**A**) 0 g and (**B**) 70 g.

3.4. Temporal Evolution of Mainline and Branch Line Sinking Depth at Different Velocities

Figures 7 and 8 illustrate the temporal evolution of sinking depth for six mainline (ML1–ML6) and six branch line (BL1–BL6) configurations under three flow velocities (0.14, 0.35, and 0.49 m/s) and two lead-line weights (0 and 70 g). Contrary to steady-state equilibrium assumptions, both components exhibit highly dynamic sinking depth oscillations throughout the 36 s, with sinking depths fluctuating continuously between approximately 50 and 280 cm. Mainline configurations display persistent sinking depth fluctuations superimposed on velocity-dependent mean settlement trends. At 0.14 m/s, mean depths cluster around 150–180 cm with peak-to-peak amplitudes of ~100–150 cm. As velocity increases to 0.49 m/s, mean sinking depths elevate to 190–240 cm with expanded amplitudes (~120–180 cm). The 70 g lead weight consistently elevates mean trajectories by ~20–40 cm. Configuration geometry strongly influences response: large-diameter designs (ML3 and ML4) sustain broad sinking depth swings (>150 cm), while the optimized 1.75 mm PLA mainline (ML5) demonstrates constrained oscillations (~80–110 cm) (Figure 7). Branch line configurations exhibit similarly complex sinking kinetics governed by multi-sectional diameter tapering and material transitions. At 0.14 m/s, sinking depths oscillate within 100–200 cm (mean: 140–160 cm). At 0.49 m/s, mean depths shift to 180–220 cm with peak-to-peak amplitudes reaching ~140–190 cm. Configurations with PLA integration (BL3, BL6) exhibit relatively regular oscillations with constrained amplitudes (~90–130 cm), whereas aggressively tapered designs (BL2, BL4) display erratic excursions (>160 cm). The 70 g lead weight shifts mean settlement downward by ~15–30 cm and dampens high-frequency flutter (Figure 8).

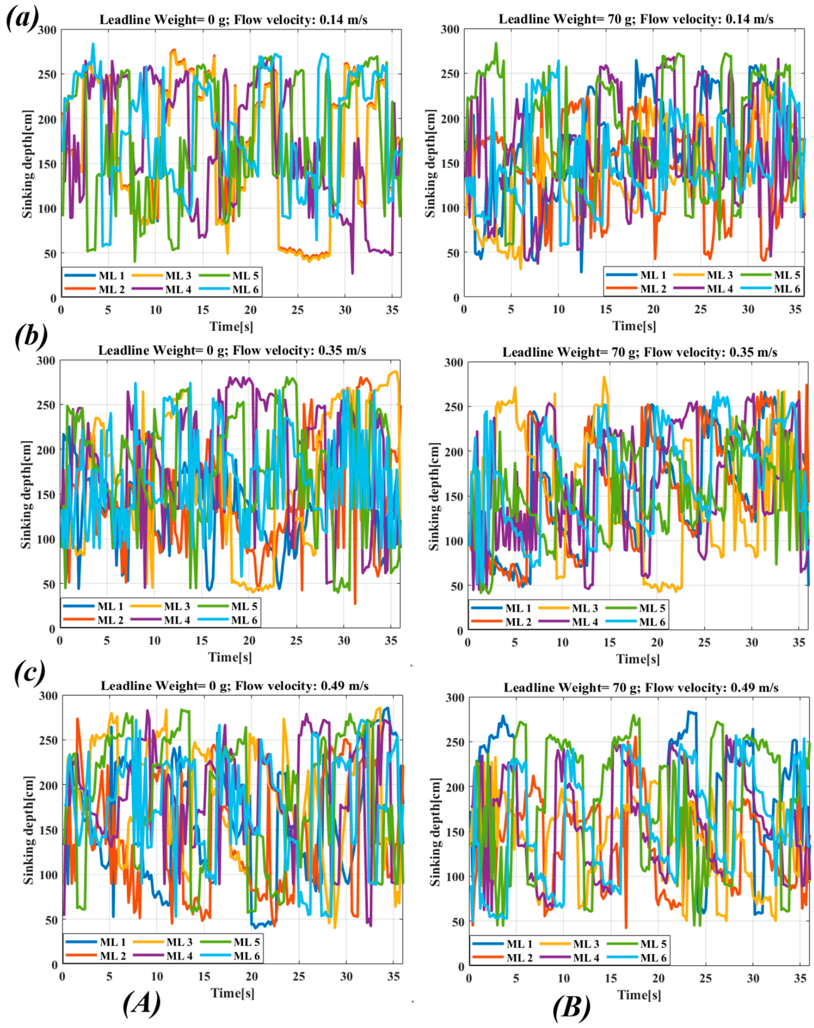

Figure 7. Temporal Evolution of mainline Sinking Depth under three flow conditions: (**a**) 0.14 ms, (**b**) 0.35 ms, (**c**) and 0.49 ms; and two lead line weight: (**A**) 0 g and (**B**) 70 g.

3.5. Analysis of Machine Learning Results for Tuna Longline Branch Line

3.5.1. Prediction of Branch Line Drag Coefficients

Figure 9 presents linear regression plots comparing the predicted outcomes against experimental values for the branch-line drag coefficient across all data categories. The solid line (R^2 = 1) denotes the perfect fit, whereas the dotted lines correspond to the fits (R^2) achieved by each model on the test set. The scatter plots for the LightGBM algorithm show relatively stable and narrowly dispersed predicted values, indicating high accuracy and consistency in predicting the branch-line drag coefficient. As shown in Table 5, LightGBM significantly outperforms RF and SVM, achieving the highest R^2 values across training (86.67%), testing (72.01%), and the whole dataset (83.15%). RF follows with values of 74.18%, 49.32%, and 69.02%, while SVM demonstrates a notably lower variance-explaining capability. In terms of absolute accuracy, LightGBM achieves the lowest MAE (0.0273), followed by SVM (0.0356) and RF (0.0392), indicating that LightGBM provides the most precise predictions. The RMSE results follow a similar trend for the top performer, with LightGBM yielding the lowest RMSE (0.0389), followed by RF (0.0528) and SVM (0.0595). Regarding computational efficiency, RF is the fastest for both training and testing, followed by LightGBM, while SVM is considerably slower. Considering the trade-off between predictive accuracy and computational cost, LightGBM emerges as the optimal choice, delivering high accuracy with relatively fast execution. Although RF offers the fastest runtime, it sacrifices predictive accuracy, particularly in the testing phase (R^2 = 49.32%). Conversely, SVM does not provide competitive accuracy relative to LightGBM and incurs the highest computational expense, making it less suitable for this application. Figure 10 illustrates the feature importance rankings for branch-line drag prediction across the three models. Lead-line weight, flow velocity, and Length 3 consistently emerge as the most influential features, highlighting them as primary factors for drag-coefficient optimization.

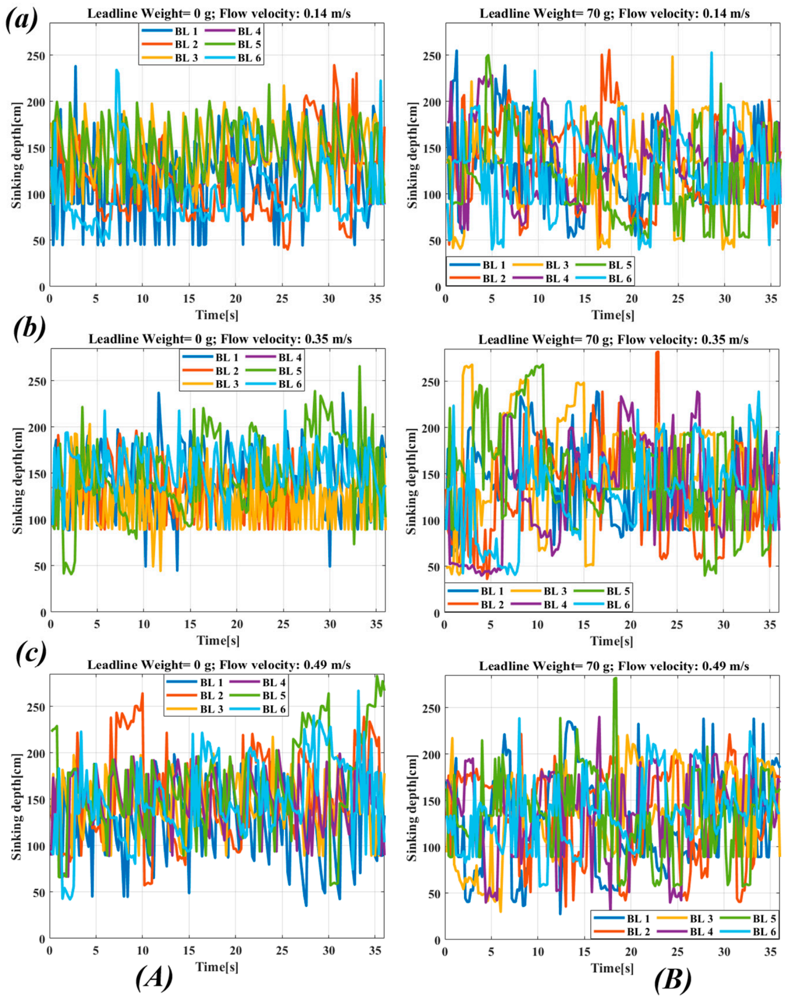

Figure 8. Temporal Evolution of branch line Sinking Depth under three flow conditions: (**a**) 0.14 ms, (**b**) 0.35 ms, (**c**) and 0.49 ms; and two lead line weight: (**A**) 0 g and (**B**) 70 g.

Table 5. Metric performance results for the branch line drag coefficient.

Model	Model	Performance Metric		
		RMSE	*MAE*	R^2
Random Forest	Training	0.0478	0.0365	0.7418
	Testing	0.0693	0.0498	0.4932
	Whole	0.0528	0.0392	0.6902
Light Gradient-Boosting Machine (LightGBM)	Training	0.0325	0.0236	0.8667
	Testing	0.0579	0.0423	0.7201
	Whole	0.0389	0.0273	0.8315
Support Vector Machine with RBF (Gaussian) kernel (SVM)	Training	0.0527	0.0314	0.6503
	Testing	0.0813	0.0528	0.446977592385202
	Whole	0.0595	0.0356	0.6059

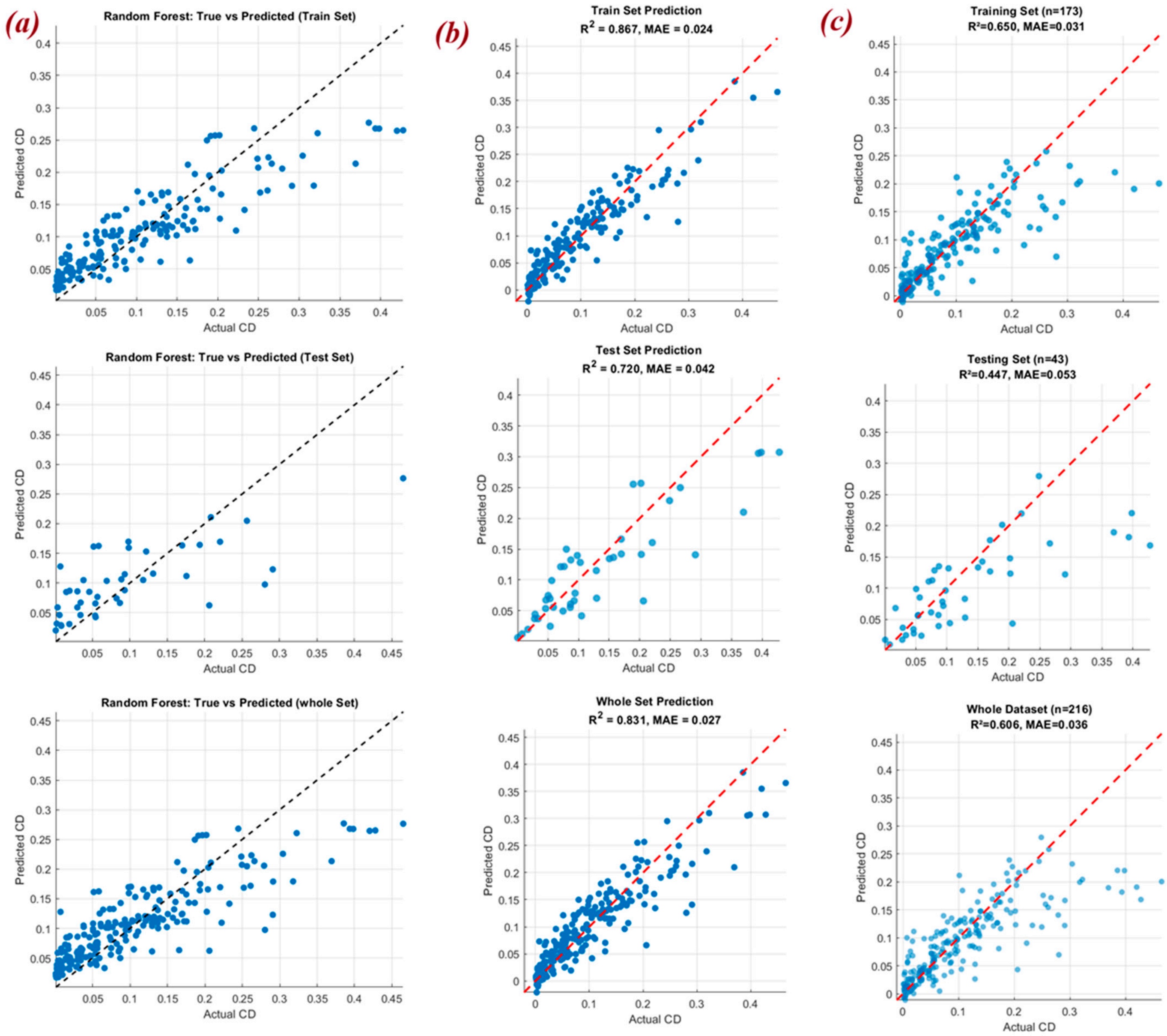

Figure 9. Comparison between the actual and predicted values of drag coefficient based on (**a**) RF model, (**b**) LightGBM model, and (**c**) SVM model (RBF): ((**top**) to (**bottom**)) training, testing, and entire sample datasets.

3.5.2. Prediction of Branch Line Lift Coefficient

For lift coefficient prediction, both LightGBM and SVM demonstrate strong, comparable performance. While LightGBM achieves a slightly higher R^2 on the whole dataset (0.7958 vs. 0.7473), SVM exhibits exceptional generalization, it is the only model for which the test R^2 (0.7842) exceeds the training R^2 (0.7355), indicating a highly robust model with minimal overfitting. LightGBM, although strong, shows a small drop in performance from training to testing (Table 6). Random Forest again underperforms relative to the others, yielding the lowest R^2 (0.5808), the highest prediction errors across the whole dataset, and a substantial drop in R^2 from training to testing (0.6210 to 0.3966) (Table 6). The scatter plots in Figure 11 verify this. Both LightGBM (Figure 11b) and SVM (Figure 11c) predictions show tight clustering around the fit line; SVM's predictions appear particularly consistent across all dataset splits. RF (Figure 11a) shows the greatest scatter in the test data. The variable importance rankings for lift coefficient (Figure 12) are less consistent across models than those for drag. Nevertheless, lead-line weight, flow velocity, and length-to-diameter ratio remain highly influential. Notably, SVM uniquely identifies flow velocity as the most important feature, whereas the tree-based models (RF and LightGBM) rank it lower, highlighting the distinct ways in which the models interpret feature interactions.

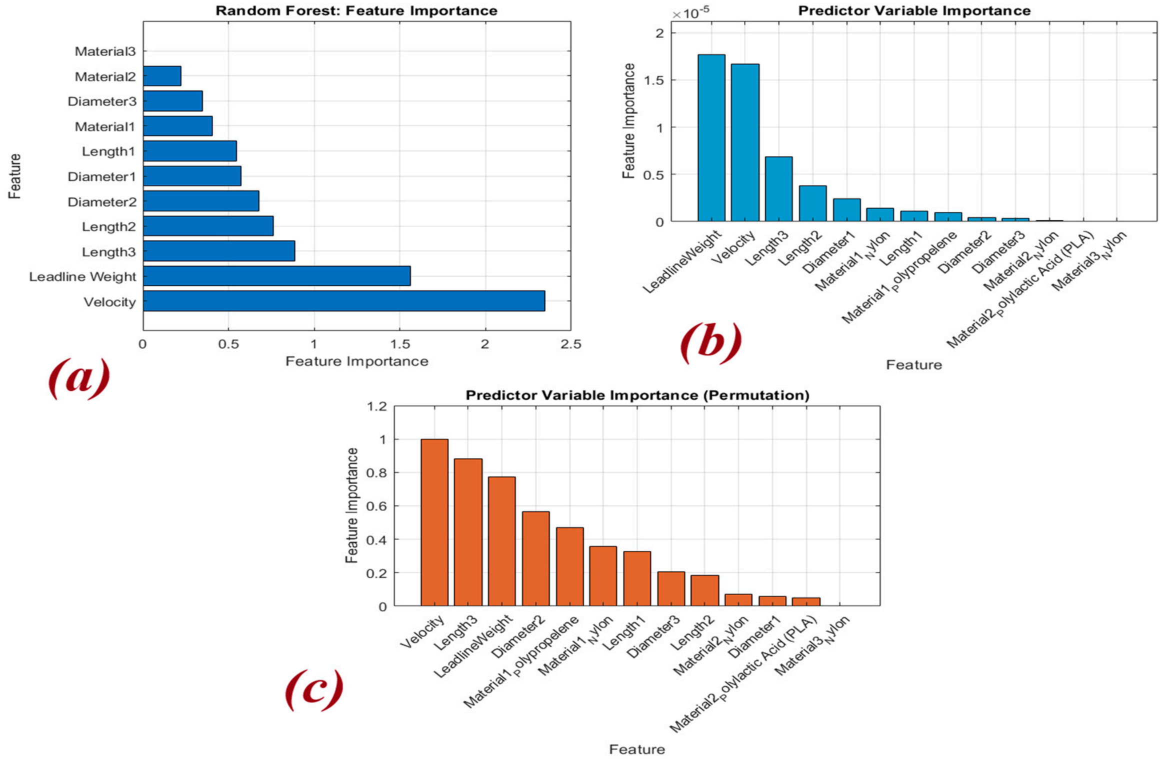

Figure 10. Variable importance rankings in predicting drag coefficient based on (**a**) RF model, (**b**) LightGBM model, and (**c**) SVM model (RBF).

Table 6. Metric performance results for the branch line lift coefficient.

Model	Model	Performance Metric		
		RMSE	*MAE*	R^2
Random Forest	Training	0.0778	0.0585	0.6210
	Testing	0.0944	0.0625	0.3966
	Whole	0.0814	0.0593	0.5808
Light Gradient-Boosting Machine (LightGBM)	Training	0.0527	0.0365	0.8173
	Testing	0.0708	0.0516	0.7182
	Whole	0.0568	0.0395	0.7958
Support Vector Machine with RBF (Gaussian) kernel (SVM)	Training	0.0635	0.0335	0.7355
	Testing	0.0619	0.0407	0.7842
	Whole	0.0632	0.0349	0.7473

3.5.3. Prediction of Branch Line Sinking Depth

For tuna longline systems, predicting sinking depth is more challenging than predicting hydrodynamic coefficients. Table 7 reports the R^2, MAE, and RMSE values as follows: LightGBM provided the best overall performance, achieving the highest R^2 values (0.9393, 0.3776, and 0.8386 for training, testing, and the whole dataset, respectively), followed by Random Forest (0.6064, 0.1776, and 0.5423) and SVM-RBF (0.5174, 0.1048, and 0.3550). LightGBM also attained the lowest errors across the whole dataset (RMSE = 15.3331, MAE = 10.7993), whereas SVM-RBF showed the highest errors

(RMSE = 25.0557, MAE = 17.9479). However, the pronounced drop in R^2 from training to testing across all three models (LightGBM: 0.9393 to 0.3776; RF: 0.6064 to 0.1776; SVM-RBF: 0.5174 to 0.1048) indicates limited generalization and a clear tendency toward overfitting, confirming that sinking depth remains a high-variance structural response that is difficult to predict reliably with the current dataset. In terms of computational cost, RF was the most efficient in both training and testing (approximately 5 s and 2 s, respectively); LightGBM required moderate resources (approximately 58.4 s and 20.6 s); and SVM-RBF required the longest training time (approximately 213 s) with a faster testing time (approximately 20.87 s), though this speed advantage did not translate into improved predictive reliability. The plots in Figure 13 show the actual versus predicted sinking depths for the branch-line using RF, LightGBM, and SVM-RBF. Figure 14 shows the feature importance for predicting branch-line sinking depth. Across all models, flow velocity, lead-line weight, and length-to-diameter ratio exerted substantial influence, confirming these geometric and operational parameters as pivotal drivers of predictive performance.

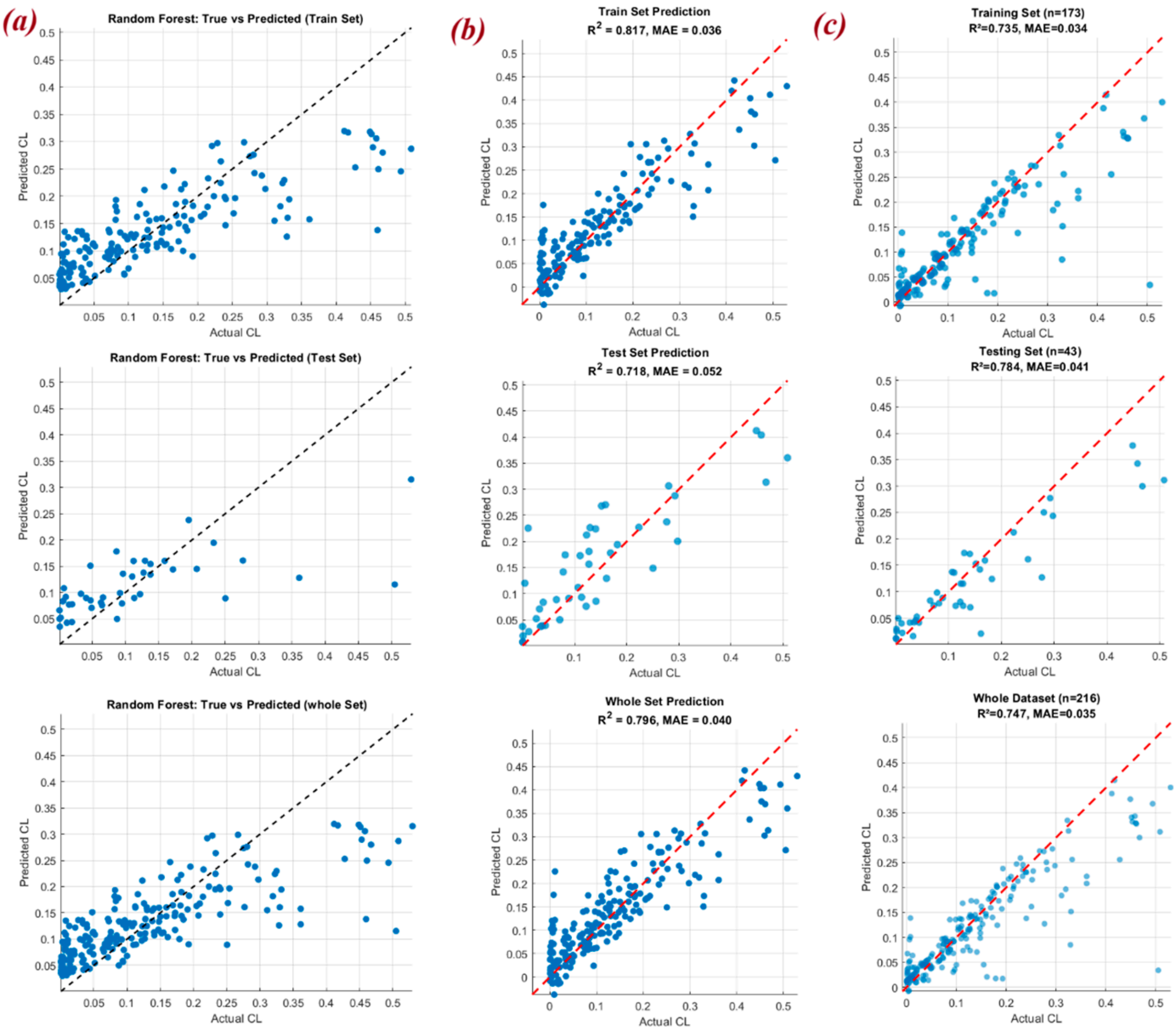

Figure 11. Comparison between the actual and predicted values of lift coefficient based on (**a**) RF model, (**b**) LightGBM model, and (**c**) SVM model (RBF): ((**top**) to (**bottom**)) training, testing, and entire sample datasets.

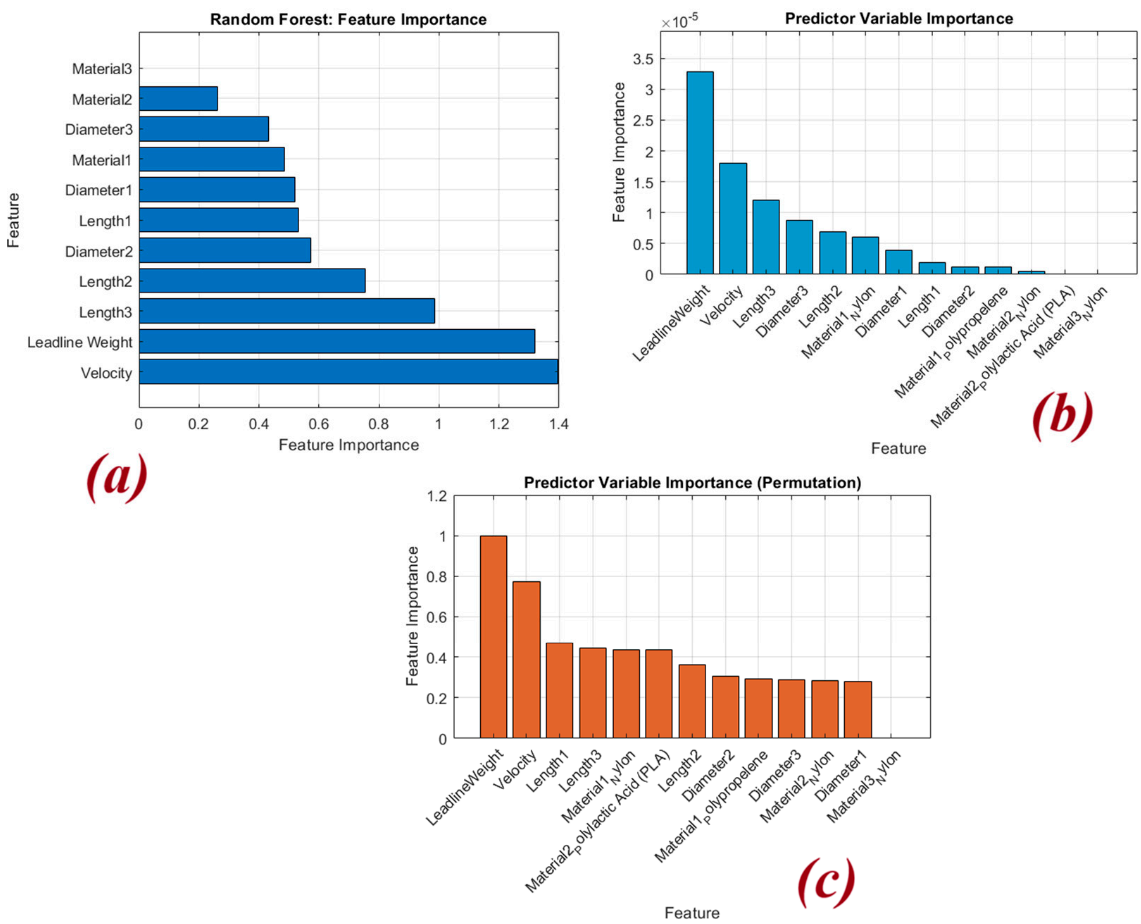

Figure 12. Variable importance rankings in predicting lift coefficient based on (**a**) RF model, (**b**) LightGBM model, and (**c**) SVM model (RBF).

Table 7. Metric performance results for branch line sinking depth.

Model	Model	Performance Metric		
		RMSE	*MAE*	R^2
Random Forest	Training	19.8630	15.6062	0.6064
	Testing	26.9558	19.4875	0.1776
	Whole	22.2020	16.7562	0.5423
Light Gradient-Boosting Machine (LightGBM)	Training	11.0008	8.4950	0.9393
	Testing	20.8513	14.0489	0.3776
	Whole	15.3331	10.7993	0.8386
Support Vector Machine with RBF (Gaussian) kernel (SVM)	Training	24.0197	17.0885	0.51738
	Testing	28.8504	21.4057	0.1048
	Whole	25.0557	17.9479	0.354988

3.5.4. Prediction of Branch Line x-Displacement

For x-displacement prediction, SVM achieves an exceptionally high training R^2 of 98.45% and a whole-dataset R^2 of 84.57%, yet its testing R^2 plummets to just 15.5%, revealing substantial overfitting. LightGBM exhibits a similar trend, with training R^2 dropping from 94.90% to 28.99% on the testing set (Table 8). In contrast, Random Forest performs poorly across all metrics, consistently yielding the lowest variance-explaining capability. The testing RMSE exceeds 28 for all three models, confirming universally poor predictive performance for this structural response. While SVM minimizes errors on the training set, it fails to generalize to unseen data; LightGBM offers only moderate performance, and RF shows persistently low accuracy across all datasets. Computational efficiency mirrors the trends observed for sinking depth: RF remains the fastest, LightGBM requires moderate resources, and SVM demands the longest training time, though its quicker testing phase does not compensate for its poor generalization. The scatter plots in Figure 15 corroborate these findings: LightGBM (Figure 15b) and SVM-RBF (Figure 15c) show training points clustered almost perfectly along the fit line, while test points are widely scattered; RF (Figure 15a) exhibits poor performance across the entire dataset. Feature importance analysis (Figure 16) reveals notable inconsistencies across models: RF and LightGBM identify density as the primary driver, whereas SVM-RBF prioritizes the length-to-diameter ratio and sinking depth. This instability underscores the inherent difficulty of x-displacement prediction and motivates closer attention to model regularization and feature engineering. Ultimately, the unreliable influence of these features suggests that additional data or alternative modeling approaches are required to achieve robust predictive performance.

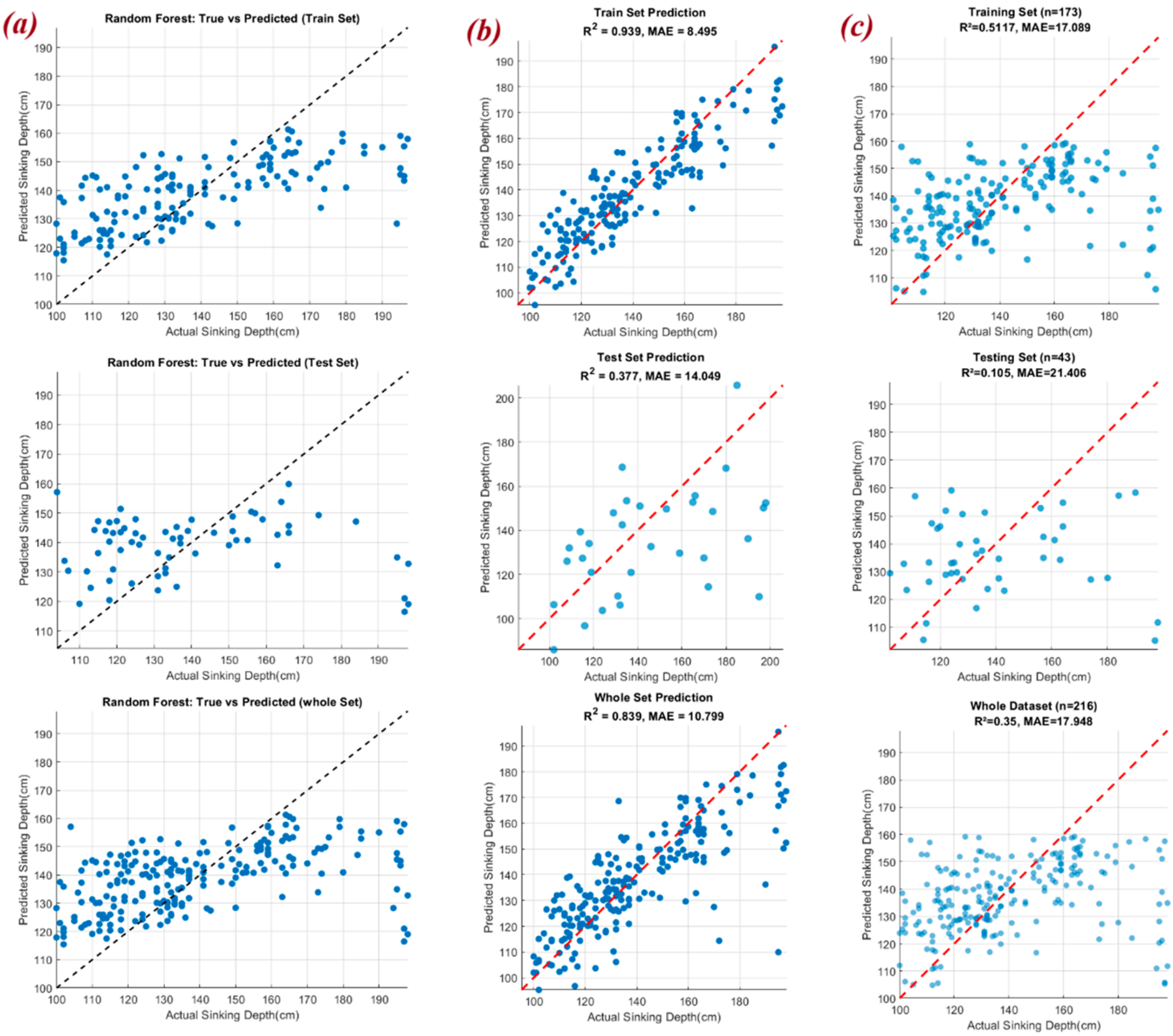

Figure 13. Comparison between the actual and predicted values of Sinking depth based on (**a**) RF model, (**b**) LightGBM model, and (**c**) SVM model (RBF): ((**top**) to (**bottom**)) training, testing, and entire sample datasets.

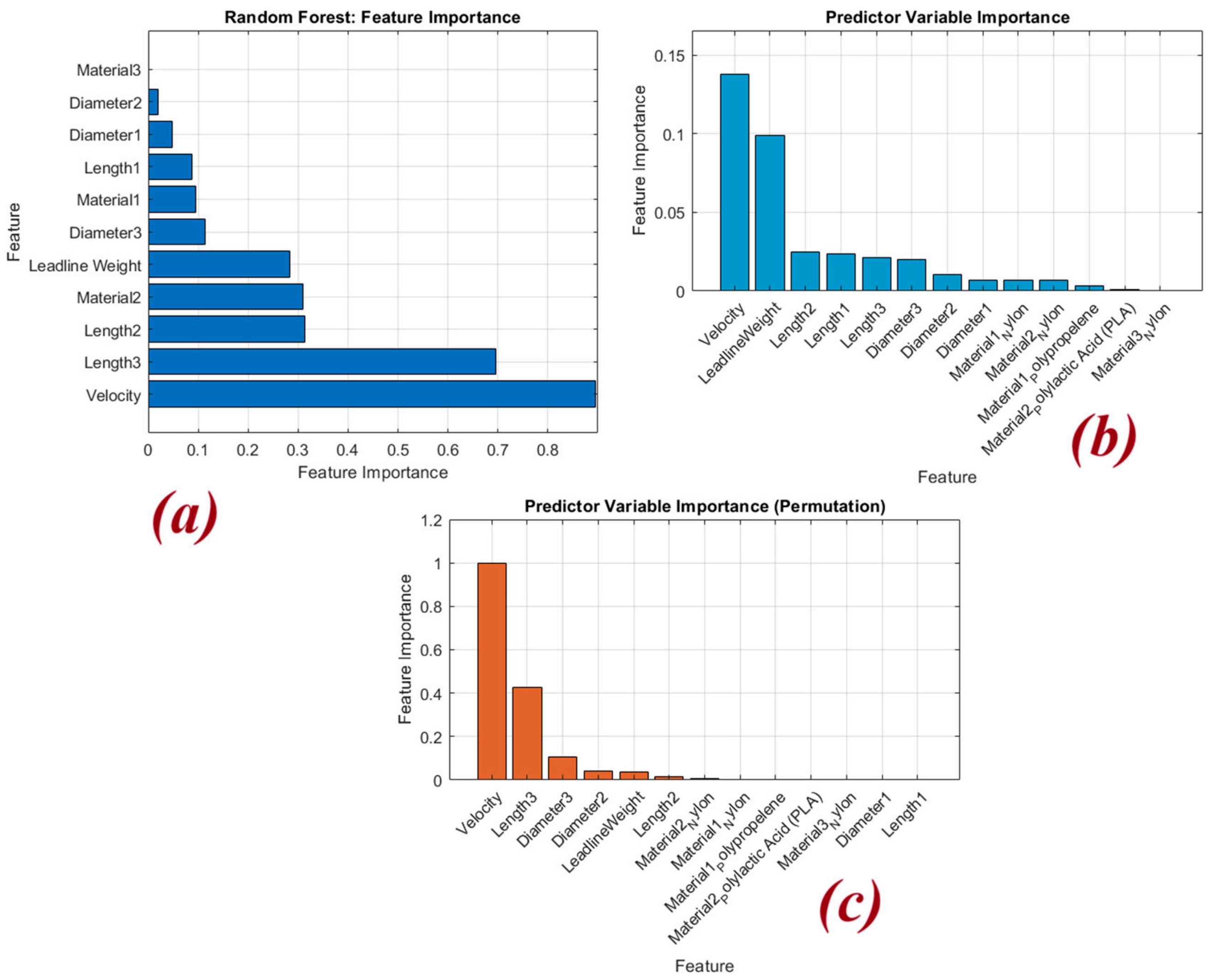

Figure 14. Variable importance rankings in predicting branch line sinking depth based on (**a**) RF model, (**b**) LightGBM model, and (**c**) SVM model (RBF).

Table 8. Metric performance results for branch line x-displacement.

Model	Model	Performance Metric		
		RMSE	*MAE*	R^2
Random Forest	Training	24.3642	20.4395	0.2609
	Testing	32.3678	28.5327	0.0738
	Whole	26.9843	22.8375	0.3507
Light Gradient-Boosting Machine (LightGBM)	Training	6.7151	5.1437	0.9490
	Testing	39.9499	30.5776	0.2899
	Whole	16.5788	8.9117	0.6794
Support Vector Machine with RBF (Gaussian) kernel (SVM)	Training	3.6633	3.6249	0.9845
	Testing	28.5607	26.1160	0.155
	Whole	11.5012	6.9569	0.8457

3.6. Analysis of Machine Learning Results for Tuna Longline Mainline

3.6.1. Prediction of Mainline Drag Coefficients

As shown in Table 9 and Figure 17, the SVM model achieved outstanding results on the training dataset, with an R^2 of 0.9939 and very low error metrics (RMSE: 0.0091, MAE: 0.0090). This shows that the SVM model can learn the patterns in the training data almost perfectly. However, its performance on the test dataset dropped notably, with an R^2 of only 0.2792. The scatter plots in Figure 18c visually support this, showing points tightly clustered around the fit line for training but widely scattered for testing. In contrast, the RF model provided a more balanced and generalized performance. While its training metrics (R^2: 0.6751) were lower than SVM's, its testing metrics (R^2: 0.5944) were the highest among all models, with a testing RMSE of 0.0679 and MAE of 0.0452. The consistency between its training and testing performance suggests that the RF model is the most robust and reliable for predicting the mainline drag coefficient on new data. The LightGBM model showed a similar pattern to SVM but less extreme, with good training performance (R^2: 0.9346) but a significant drop in testing (R^2: 0.3699).

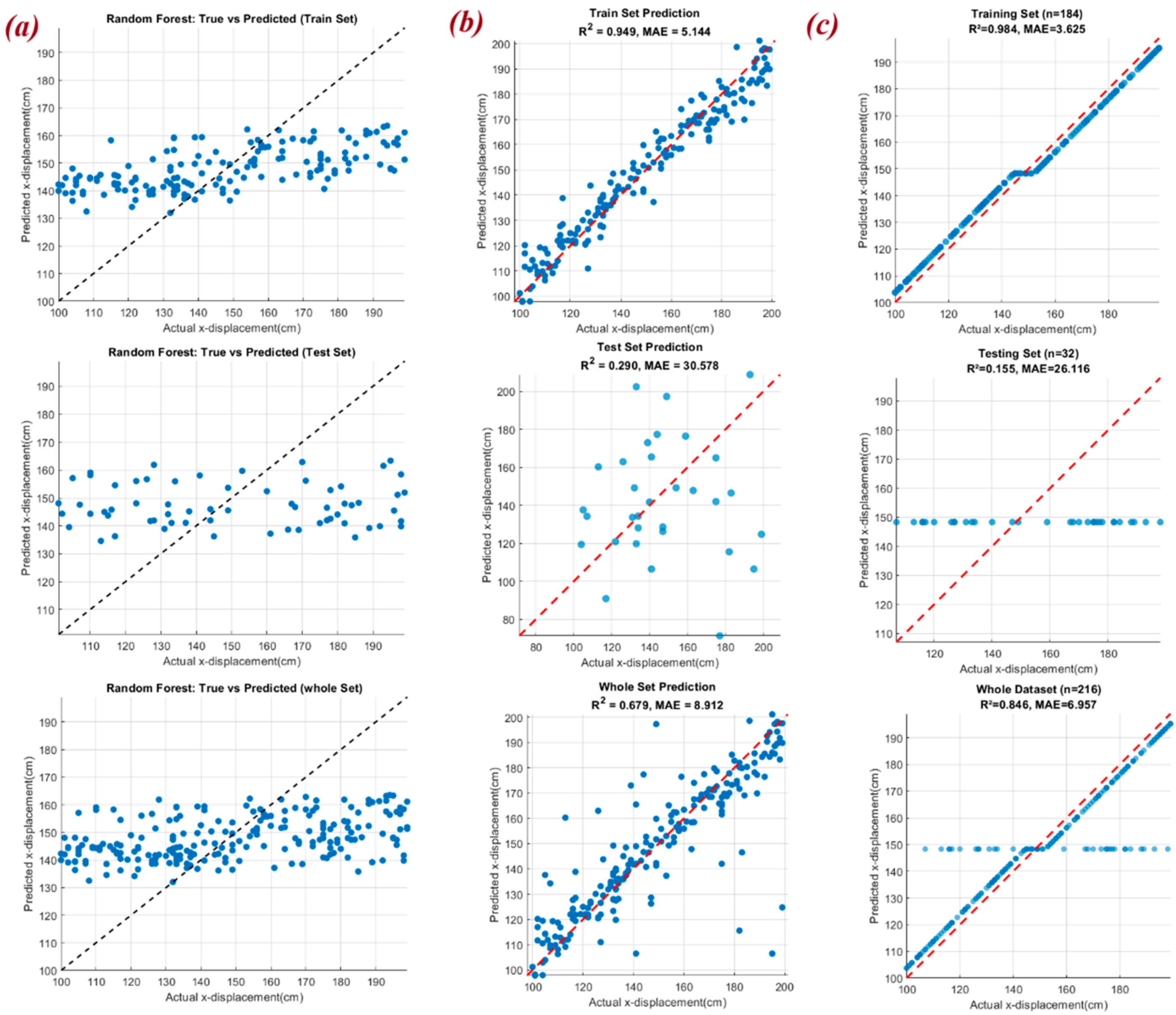

Figure 15. Comparison between the actual and predicted values of x-displacement based on (**a**) RF model, (**b**) LightGBM model, and (**c**) SVM model (RBF): ((**top**) to (**bottom**)) training, testing, and entire sample datasets.

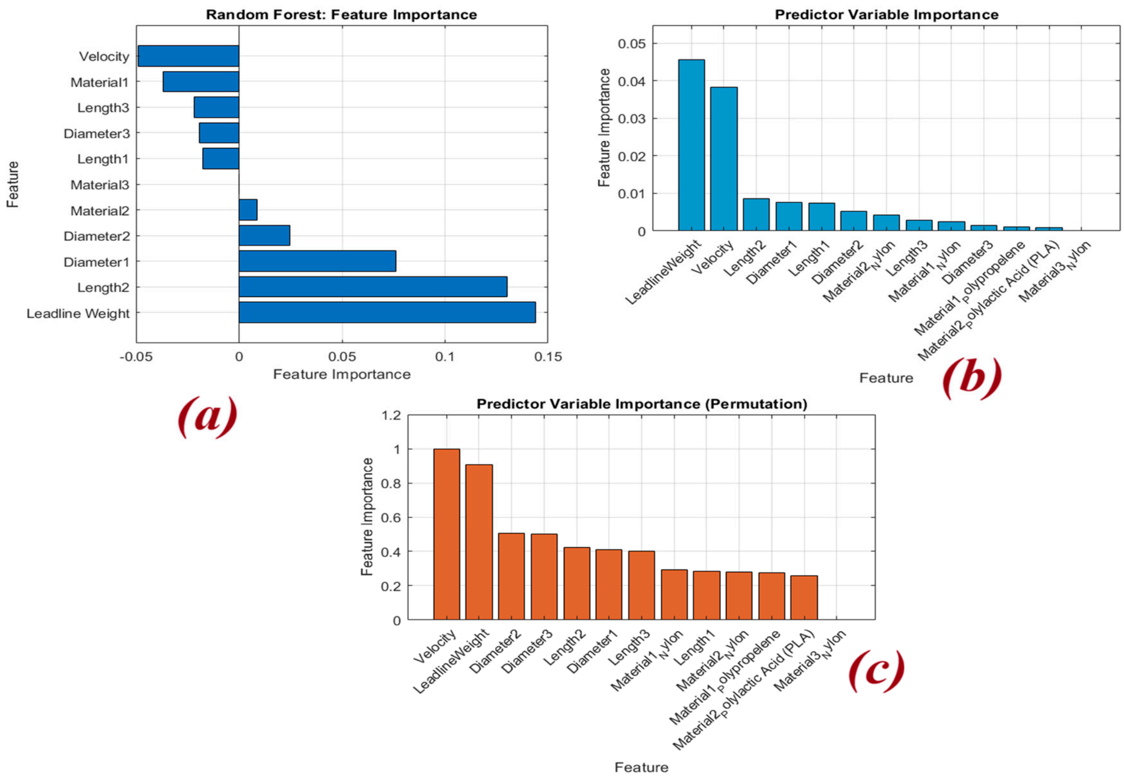

Figure 16. Variable importance rankings in predicting branch line x-displacement based on (**a**) RF model, (**b**) LightGBM model, and (**c**) SVM model (RBF).

Table 9. Metric performance results for the mainline drag coefficient.

Model	Model	Performance Metric		
		RMSE	*MAE*	R^2
Random Forest	Training	0.0869	0.0583	0.6751
	Testing	0.0679	0.0452	0.5944
	Whole	0.0818	0.0544	0.6640
Light Gradient-Boosting Machine (LightGBM)	Training	0.0544	0.0391	0.9346
	Testing	0.1213	0.0707	0.3699
	Whole	0.0728	0.0454	0.7749
Support Vector Machine with RBF (Gaussian) kernel (SVM)	Training	0.0091	0.0090	0.9939
	Testing	0.0639	0.0469	0.2792
	Whole	0.0260	0.0146	0.9458

3.6.2. Prediction of Mainline Lift Coefficients

As shown in Table 10 and visualized in Figure 19, the SVM model achieved an almost perfect fit on the whole data (R^2: 0.9537, RMSE: 0.0439). The Random Forest model proved to be the most generalizable model. It delivered the strongest performance on the testing dataset, with an R^2 of 0.5008, an RMSE of 0.1811, and an MAE of 0.0736. While its training R^2 (0.6570) was lower than the other two models, its ability to maintain a reasonable level of accuracy on unseen data makes it the preferred choice for this task. The LightGBM model, while achieving a high training R^2 (0.9475), also showed a substantial drop in testing R^2 (0.6199). For the RF model (Figure 20a), the number of attached branch lines (gear density) was the most important variable, followed closely by current velocity. The LightGBM model (Figure 20b) likewise ranked the number of attached branch lines as the top predictor, with current velocity second. The SVM-RBF model (Figure 20c) similarly identified gear density as the most significant feature. This suggests that while current velocity is crucial for drag, the configuration of the gear, specifically the number of attached gears, plays a more dominant role in generating lift on the mainline.

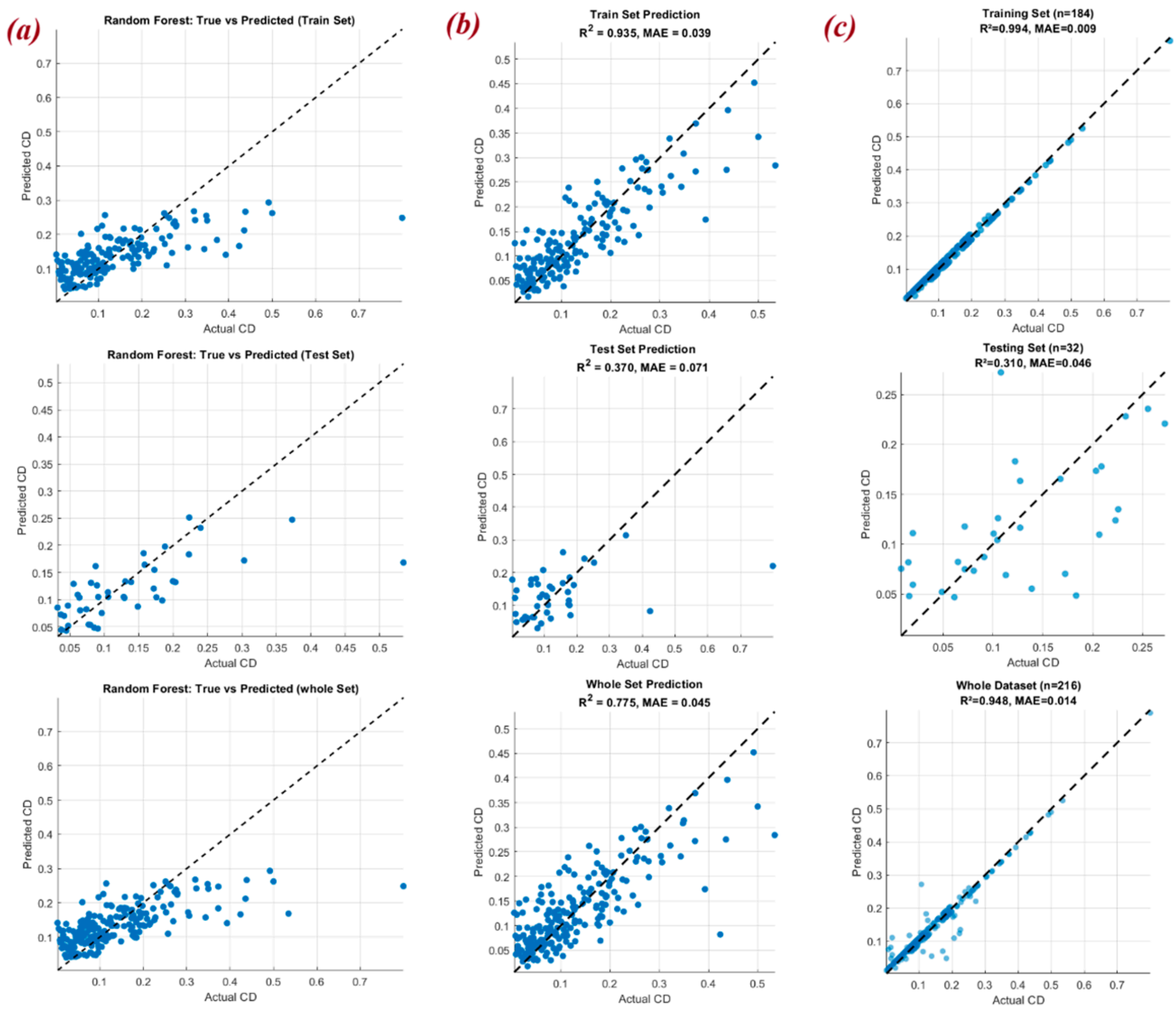

Figure 17. Comparison between the actual and predicted values of mainline coefficient based on (**a**) RF model, (**b**) LightGBM model, and (**c**) SVM model (RBF): ((**top**) to (**bottom**)) training, testing, and entire sample datasets.

Table 10. Metric performance results for the mainline lift coefficient.

Model	Model	Performance Metric		
		RMSE	*MAE*	R^2
Random Forest	Training	0.1480	0.0960	0.6570
	Testing	0.1811	0.0736	0.5008
	Whole	0.1552	0.0915	0.6220
Light Gradient-Boosting Machine (LightGBM)	Training	0.0937	0.0624	0.9475
	Testing	0.2006	0.1265	0.6199
	Whole	0.1226	0.0752	0.8390
Support Vector Machine with RBF (Gaussian) kernel (SVM)	Training	0.0162	0.01593	0.9943
	Testing	0.1073	0.0938	0.3434
	Whole	0.0439	0.0275	0.9537

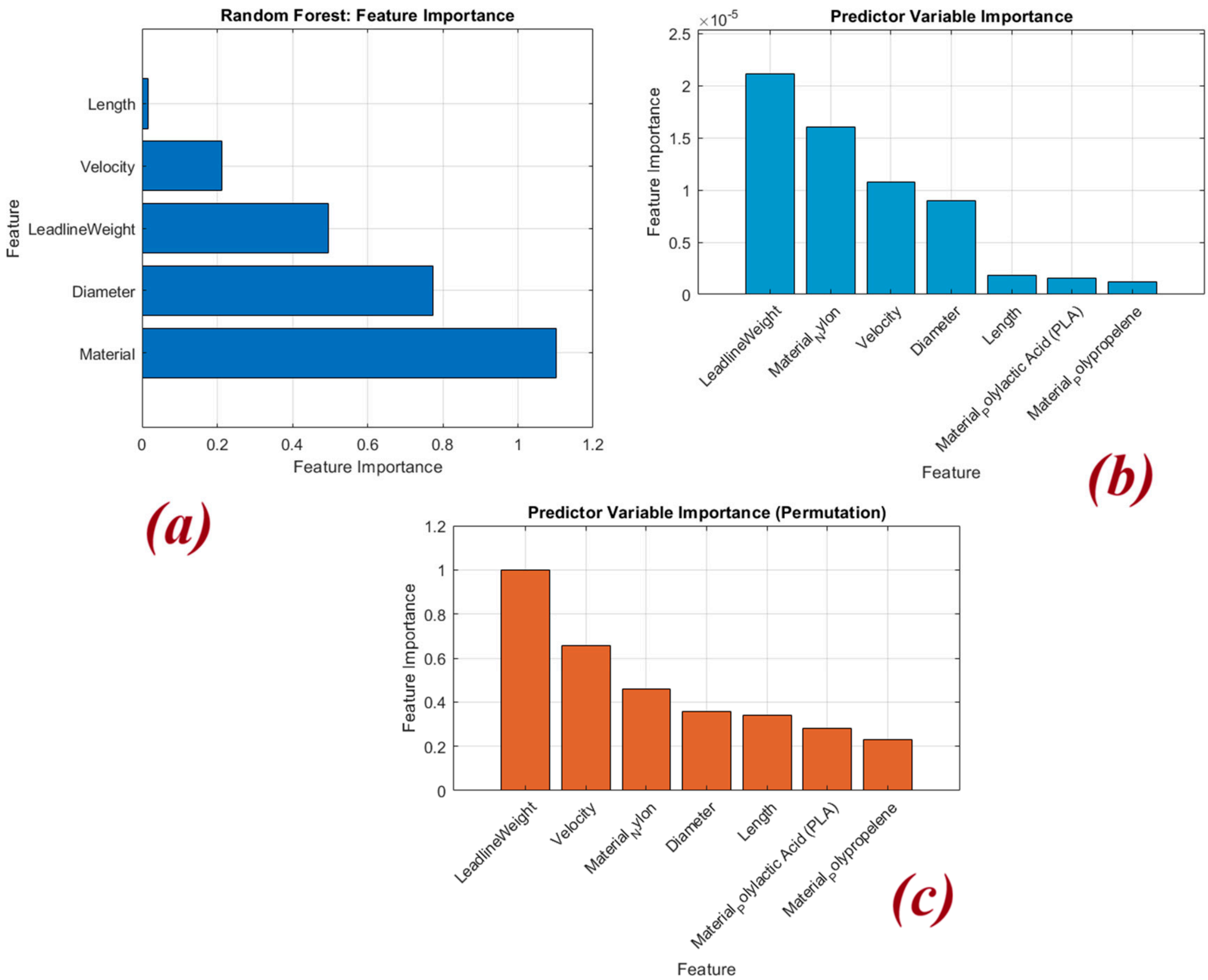

Figure 18. Variable importance rankings in predicting mainline drag coefficient based on (**a**) RF model, (**b**) LightGBM model, and (**c**) SVM model (RBF).

3.6.3. Prediction of Mainline Sinking Depth

The results for predicting mainline sinking depth, shown in Table 11 and Figure 21, show a clear pattern of model performance like that previously observed, but with notable differences in error magnitudes. The SVM model displays extreme overfitting. It achieves a near-perfect fit on the training data with an R^2 of 0.9833 and remarkably low errors at (RMSE: 3.95). However, its performance on the testing dataset is very poor, with the R^2 plummeting to 0.2589 and the RMSE ballooning to 29.57. This dramatic increase in error from training to testing confirms the model's inability to generalize. The scatter plots in Figure 21c starkly illustrate this, showing a perfect line for training points with little correlation in test points. The Random Forest model, unlike SVM, does not achieve high training scores but provides much more consistent, though still limited, performance. The LightGBM model shows a similar pattern, with a testing R^2 of 0.2158 and RMSE of 29.37. All models struggled to predict sinking depth on the test, as indicated by low R^2 values (all below 0.26), suggesting that this particular output is more challenging to predict. LightGBM (Figure 22b) and SVM (Figure 22c) sinking depths are identified as the most important predictors and appear secondary to current sinking depth in the models' predictions.

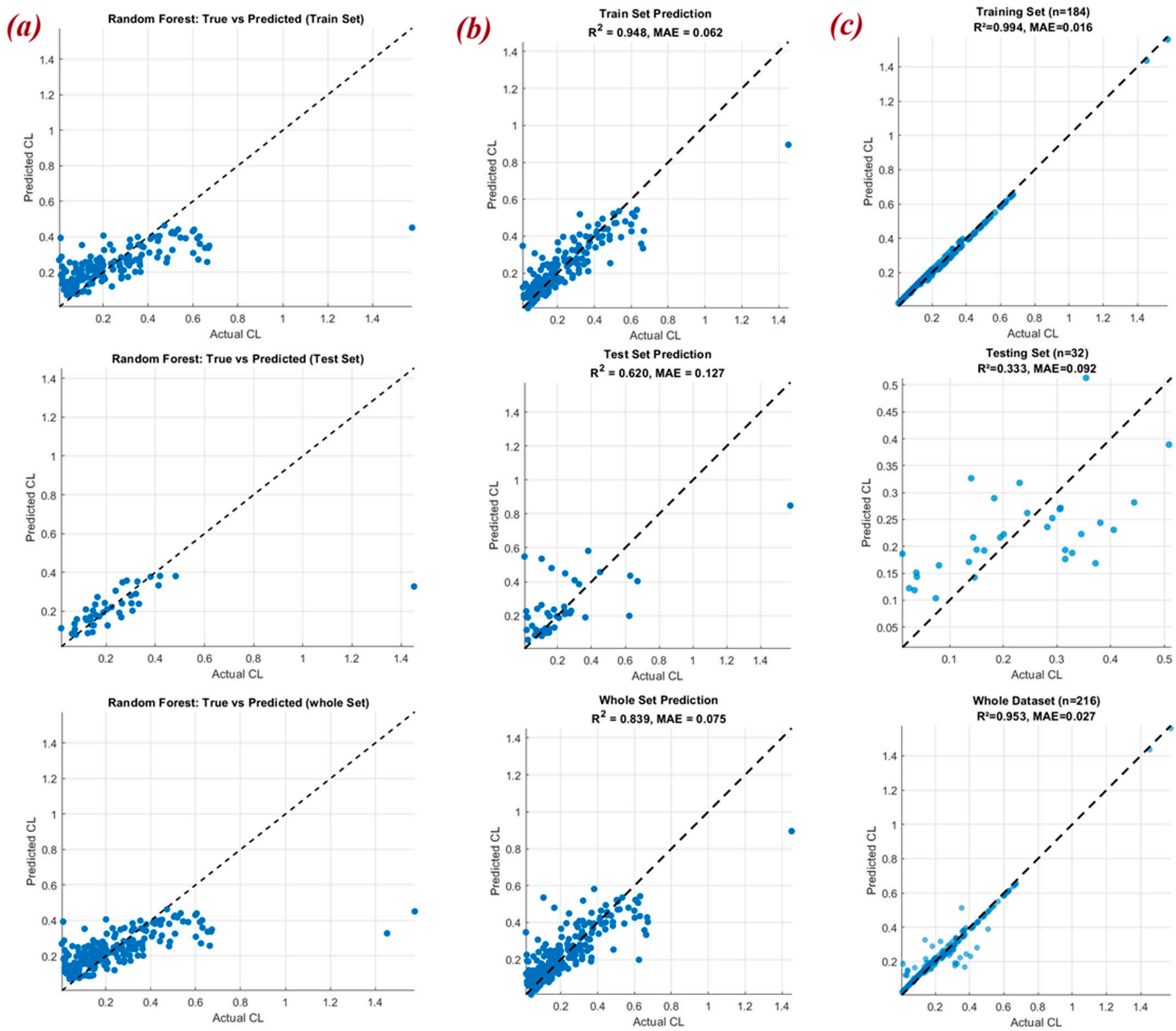

Figure 19. Comparison between the actual and predicted values of mainline lift coefficient based on (**a**) RF model, (**b**) LightGBM model, and (**c**) SVM model (RBF): ((**top**) to (**bottom**)) training, testing, and entire sample datasets.

Table 11. Metric performance results for mainline sinking depth.

Model	Model	Performance Metric		
		RMSE	*MAE*	R^2
Random Forest	Training	27.1345	23.5746	0.5267
	Testing	32.4809	26.9583	0.1921
	Whole	28.2795	24.2482	0.4384
Light Gradient-Boosting Machine (LightGBM)	Training	24.5845	20.7155	0.6428
	Testing	29.3708	24.2371	0.2158
	Whole	26.3188	22.2128	0.5537
Support Vector Machine with RBF (Gaussian) kernel (SVM)	Training	3.9542	3.9088	0.9833
	Testing	29.5699	25.9112	0.2589
	Whole	11.9523	7.1684	0.9461

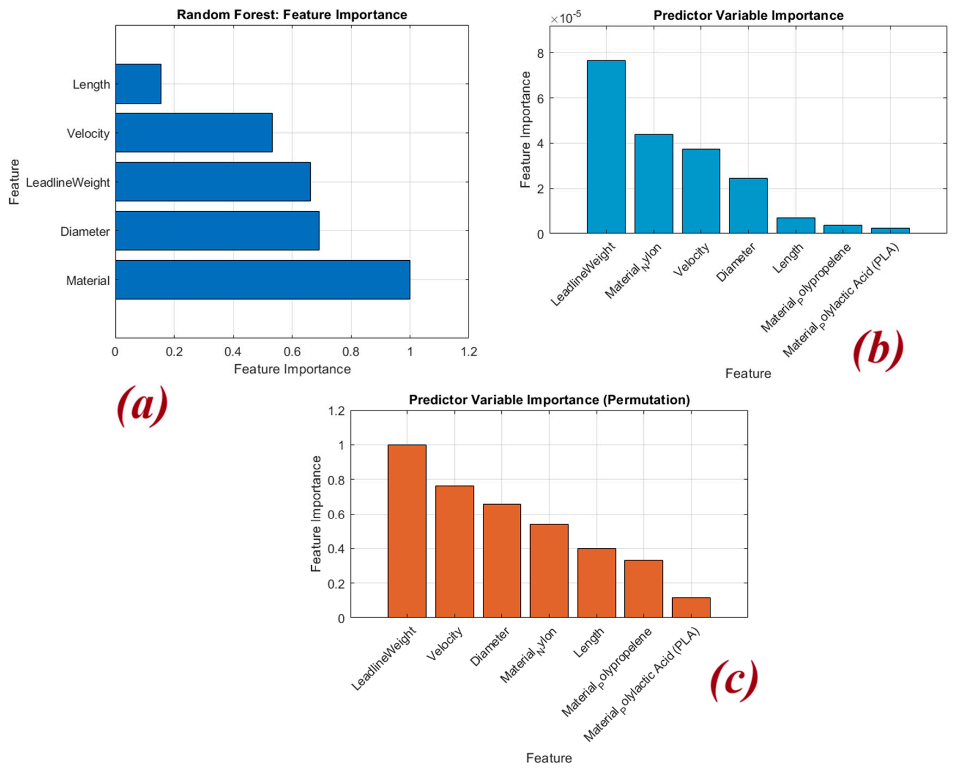

Figure 20. Variable importance rankings in predicting mainline lift coefficient based on (**a**) RF model, (**b**) LightGBM model, and (**c**) SVM model (RBF).

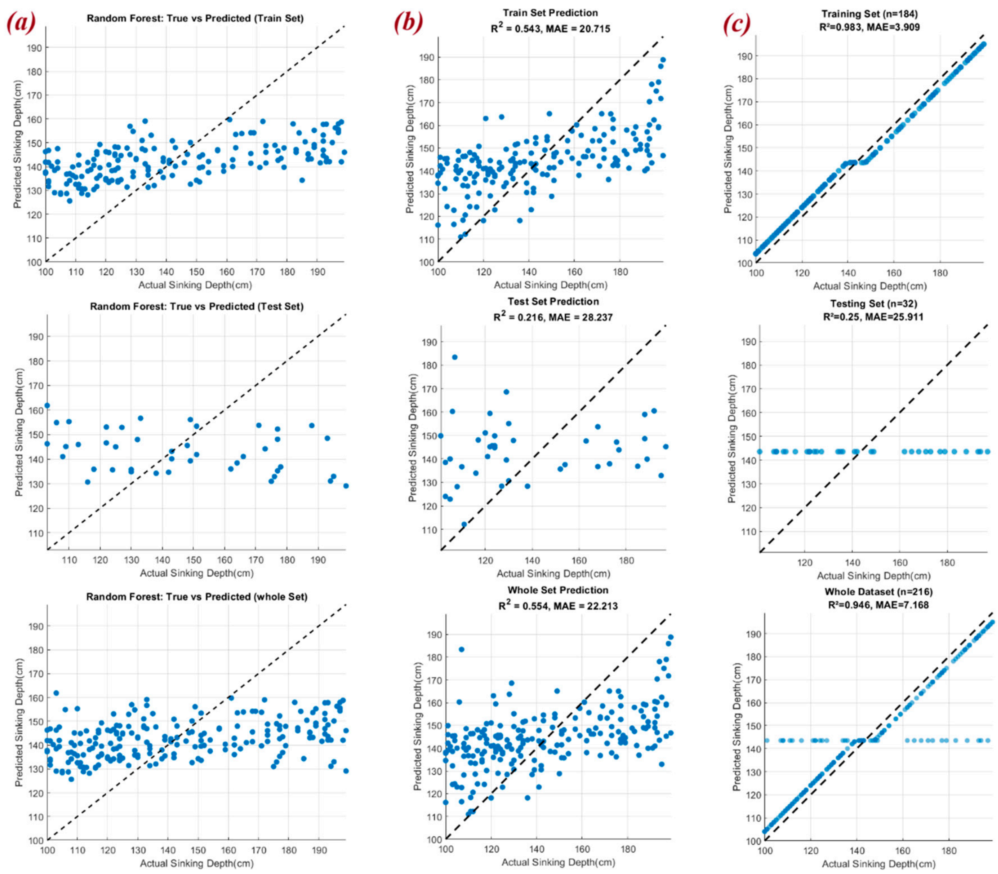

Figure 21. Comparison between the actual and predicted values of mainline sinking depth based on (**a**) RF model, (**b**) LightGBM model, and (**c**) SVM model (RBF): ((**top**) to (**bottom**)) training, testing, and entire sample datasets.

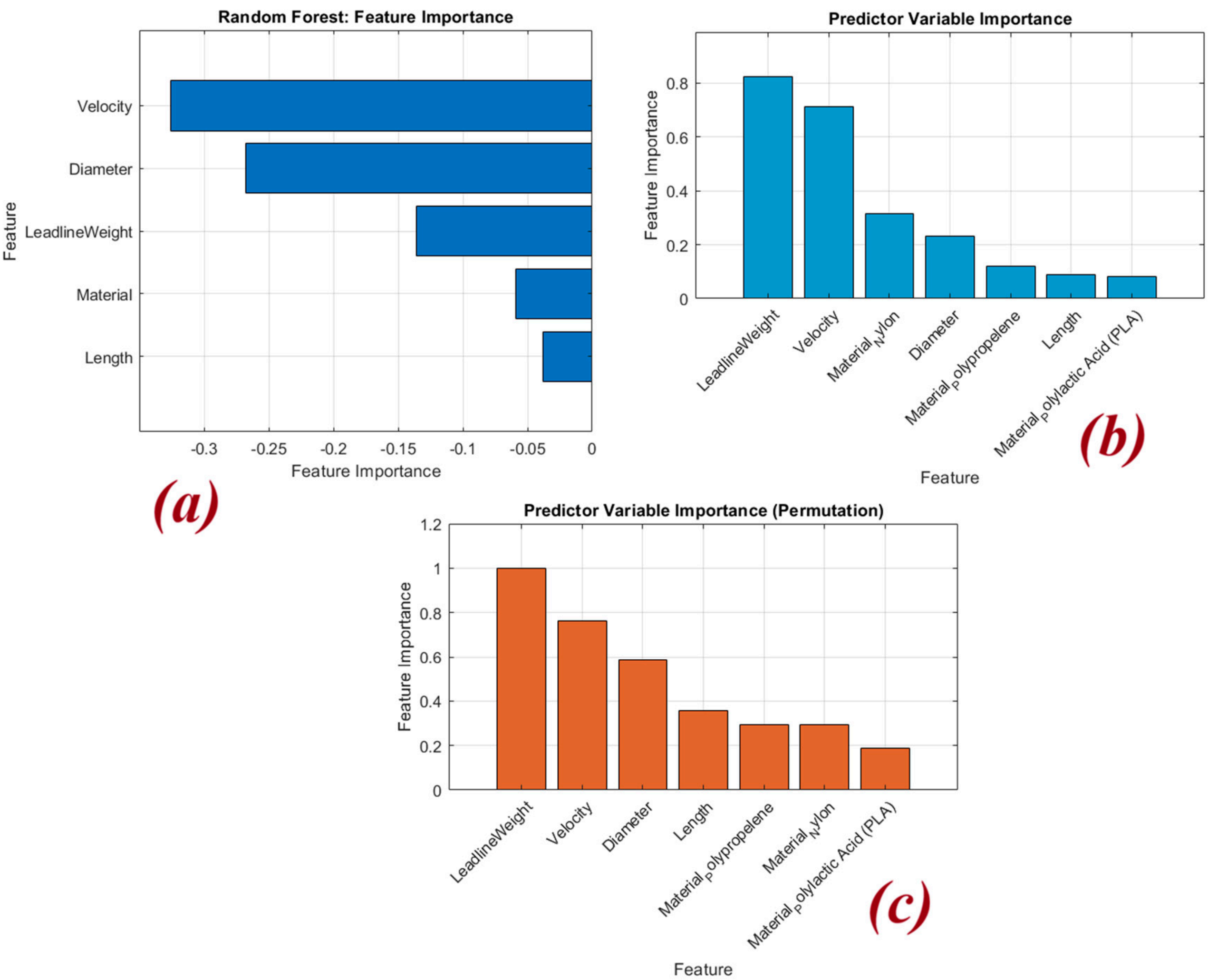

Figure 22. Variable importance rankings in predicting mainline sinking depth based on (**a**) RF model, (**b**) LightGBM model, and (**c**) SVM model (RBF).

3.6.4. Prediction of Mainline x-Displacement

Figure 23 presents linear regression plots comparing predicted and experimental values for mainline x-displacement (horizontal movement). The scatter plots for LightGBM and SVM-RBF display relatively stable and tightly clustered predictions, indicating high accuracy and consistency for this target variable.

Table 12 reports the R^2, MAE, and RMSE metrics. Regarding explanatory power, SVM-RBF substantially outperforms RF and LightGBM (whole-dataset R^2 = 91.78%), with LightGBM ranking second (68.52%), while RF yields notably lower values. For prediction accuracy, SVM-RBF again performs best, achieving the lowest MAE (6.8695), followed by LightGBM (MAE = 15.34), whereas RF exhibits the highest MAE (22.60). The RMSE rankings mirror these trends, with SVM-RBF showing the lowest error and reinforcing its superior predictive accuracy.

Feature importance rankings for x-displacement (Figure 24) demonstrate clear consensus across all three models. Flow velocity is overwhelmingly the most influential predictor, ranked first by RF (Figure 24a), LightGBM (Figure 24b), and SVM-RBF (Figure 24c), with all remaining features exerting considerably less influence on model predictions.

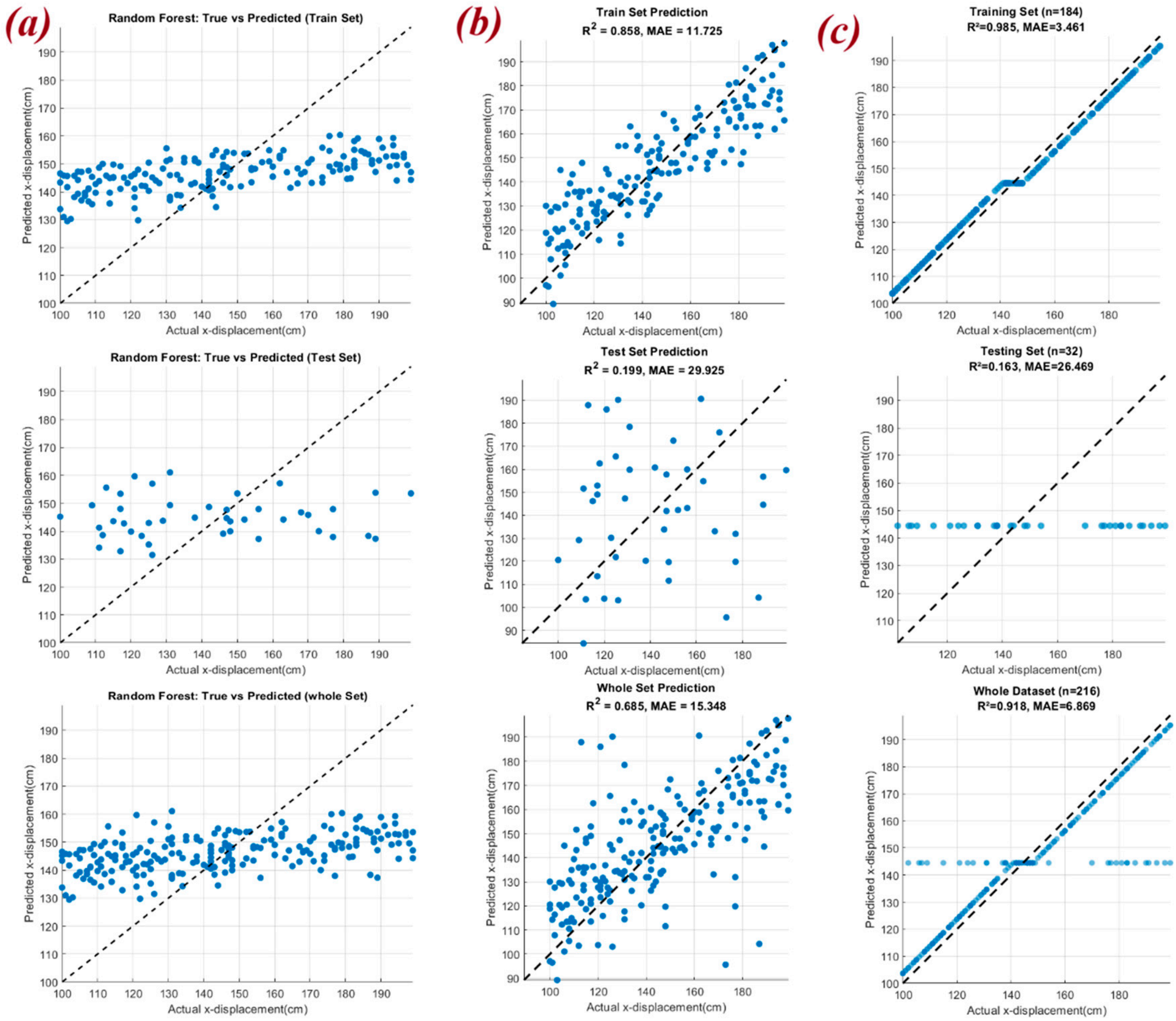

Figure 23. Comparison between the actual and predicted values of mainline x-displacement based on (**a**) RF model, (**b**) LightGBM model, and (**c**) SVM model (RBF): ((**top**) to (**bottom**)) training, testing, and entire sample datasets.

Table 12. Metric performance results for mainline x-displacement.

Model	Model	Performance Metric		
		RMSE	*MAE*	R^2
Random Forest	Training	26.5134	22.6007	0.5099
	Testing	26.6565	22.5995	0.2690
	Whole	26.5419	22.6004	0.4702
Light Gradient-Boosting Machine (LightGBM)	Training	14.6826	11.7252	0.8577
	Testing	24.4452	19.9250	0.1998
	Whole	20.9065	15.3483	0.6852
Support Vector Machine with RBF (Gaussian) kernel (SVM)	Training	3.5125	3.460915	0.985108
	Testing	21.1929	16.4687	0.163219
	Whole	12.4362	6.8695	0.917829

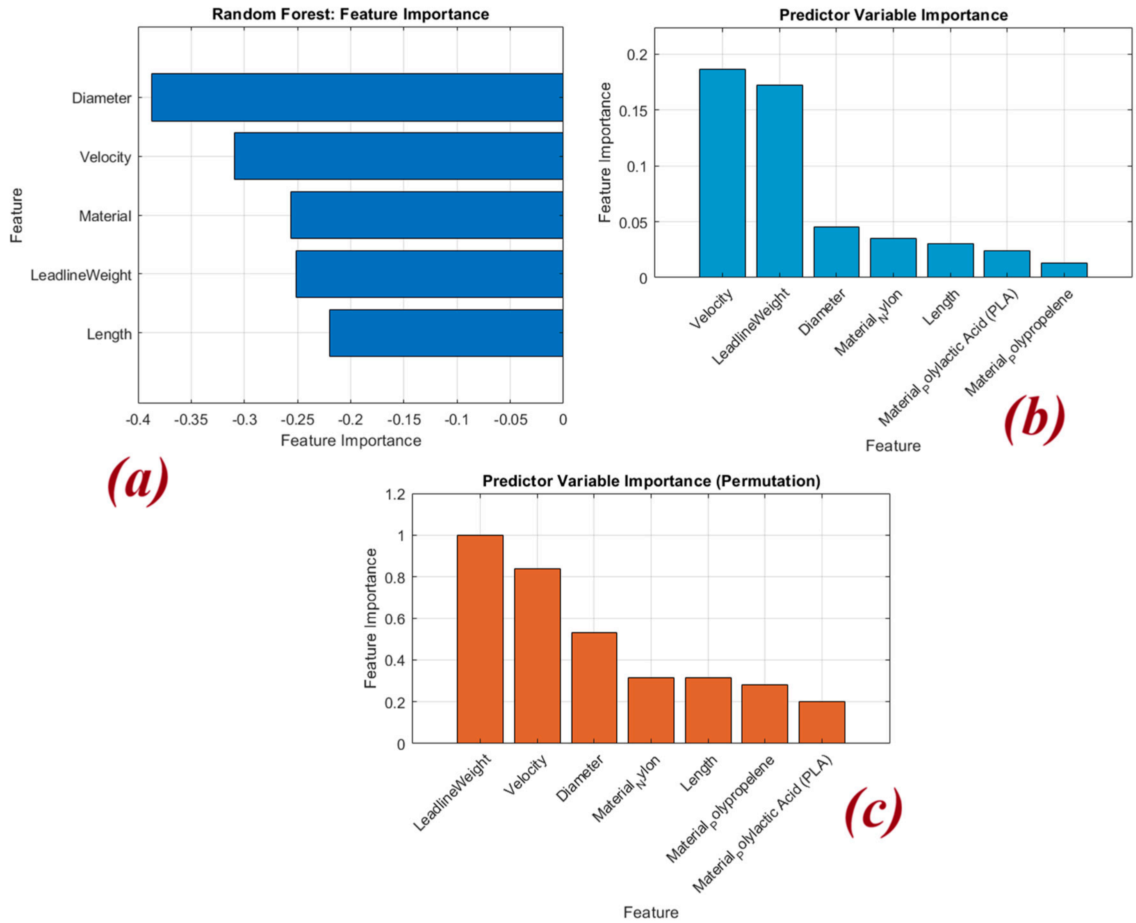

Figure 24. Variable importance rankings in predicting mainline x-displacement based on (**a**) RF model, (**b**) LightGBM model, and (**c**) SVM model (RBF).

4. Discussion

4.1. Algorithmic Interpretation of Physical Drivers in Hydrodynamic and Structural Response Predictions

The feature importance analysis conducted across the three machine learning algorithms, namely RF, LightGBM, and SVM-RBF, provides critical insights into the physical mechanisms governing hydrodynamic forces and structural responses in tuna longline systems. More importantly, it reveals how distinct algorithmic architectures prioritize input variables according to their underlying mathematical formulations, thereby offering complementary perspectives on the same physical phenomena. Table 12 summarizes the relative advantages, disadvantages, and recommended use context of the three algorithms. For drag coefficient predictions across both branch-line and mainline configurations, length, lead-line weight, material type, and flow velocity consistently emerged as the dominant predictors across all three algorithms. This convergence reflects fundamental principles of fluid–structure interaction, in which inertial forces scale with fluid density and the square of the flow velocity [13]. The consistent identification of these features by all models demonstrates strong physical interpretability, as hydrodynamic drag on submerged cylindrical elements is classically described by the drag equation [13,15]. This alignment between data-driven feature rankings and first-principles hydrodynamics validates the physical plausibility of the machine learning models and reinforces confidence in their predictive utility for gear design applications. However, notable algorithm-specific differences emerged, illuminating distinct aspects of the underlying physics. The SVM-RBF model uniquely identified branch line length as the most significant feature for mainline drag prediction, suggesting that kernel-based methods may capture higher-order interactions between system configuration and integrated drag that tree-based ensemble methods tend to average out through recursive partitioning. This variation likely stems from fundamental differences in how algorithms represent nonlinear feature interactions: support vector machines construct decision boundaries in high-dimensional kernel-induced spaces, where localized feature interactions receive differential weighting compared to the impurity-based importance measures inherent to decision tree ensembles [16]. Such findings indicate that while drag on individual line segments follows classical scaling laws, system-level drag emerges from configurational interactions that require more flexible functional representations to capture accurately. In addition, lift coefficient predictions showed greater divergence in feature importance rankings across models, particularly for branch-line configurations. While SVM-RBF prioritized current velocity as the primary predictor, tree-based models (RF and LightGBM) emphasized material density and the length-diameter ratio. A limitation of the current feature engineering approach is the use of only raw input parameters without derived dimensionless groups or interaction terms. In fluid–structure interaction problems, several dimensionless parameters govern behavior and could improve model generalization if included as input features. These include: (1) the Froude number ($Fr = U/\sqrt{(gL)}$), which quantifies the ratio of inertial to gravitational forces—particularly relevant for the observed transition from gravity-dominated to flow-dominated behavior at 0.35 m/s; (2) the Cauchy number ($Cy = \rho U^2/E$), which represents the ratio of inertial to elastic forces, important for capturing fluttering and deformation modes; (3) the stiffness ratio ($EI/\rho U^2 L^4$), which characterizes bending resistance relative to hydrodynamic loading; and (4) the diameter-to-length ratio (D/L), a simple interaction term that captures slenderness effects. Additionally, interaction terms such as (diameter $\times$ velocity) would explicitly encode the Reynolds-number dependence that the current models implicitly learn. Future work should incorporate these dimensionless parameters directly into the feature set, potentially improving generalization with smaller datasets.

This distinction aligns with the theoretical capacity of radial basis function kernels to model complex, non-additive relationships in transformed feature spaces [17]. SVM's emphasis on current velocity may reflect its sensitivity to flow-regime-dependent phenomena, such as vortex shedding frequency or boundary layer transitions, whereas tree-based methods may distribute importance across multiple correlated geometric features. This finding supports the observation that lift forces on fishing gear components exhibit heightened sensitivity to small variations in flow orientation and attack angle, underscoring the value of models capable of detecting subtle, regime-dependent dependencies [18].

For sinking depth predictions, instantaneous depth emerged as the most influential predictor across all mainline models, a result that may appear tautological but carries significant physical interpretation. In longline dynamics, sinking behavior arises from a transient equilibrium among gravitational forces, buoyancy, and hydrodynamic drag. Instantaneous depth directly modulates local flow conditions, tension distribution, and subsequent sinking rates through feedback mechanisms [19]. Numerical simulations of flexible cable dynamics further demonstrate that depth profiles exhibit strong spatial autocorrelation along the line, meaning that depth at one location constrains plausible configurations at adjacent points [2]. The consistent ranking of the length–diameter ratio as the second-most important feature by tree-based models aligns with classical beam theory and cable mechanics, wherein the slenderness ratio governs bending stiffness, gravitational load distribution, and resistance to flow-induced deformation. This algorithmic consensus reinforces the physical understanding that sinking performance is governed not by isolated parameters but by the coupled interaction of geometric stiffness and hydrodynamic loading. For x-displacement predictions, current velocity was identified as the dominant predictor across all three algorithms, consistent with fundamental fluid–structure interaction theory: the horizontal deflection of compliant cylindrical structures scales with the magnitude of the drag force, which in turn varies quadratically with flow velocity. However, the pronounced overfitting observed in both LightGBM and SVM-RBF for x-displacement, evidenced by train R^2 values approaching unity, contrasted with substantially lower test performance, suggests that the relationship between input parameters and horizontal deformation is more complex than the current feature set captures.

4.2. Model-Specific Generalization, Overfitting Diagnosis, and Design Implications for Longline Systems

Training-testing discrepancies often indicate that models have learned spurious correlations or noise patterns present in the training data, a well-documented challenge when modeling high-variance structural responses with limited sample sizes [3]. In contrast, the Random Forest model's more conservative training performance, coupled with superior generalization to unseen test data, highlights the practical advantages of bagging-based ensemble methods for engineering applications where predictive reliability under novel operating conditions is paramount. Collectively, these feature importance patterns validate the physical plausibility of the machine learning models while revealing algorithm-specific strengths: kernel methods excel at detecting flow-regime-sensitive phenomena and configurational interactions; gradient boosting efficiently captures complex nonlinearities in hydrodynamic coefficients; and ensemble tree methods provide the most reliable generalization for structural response predictions under variable environmental conditions. This integrated interpretation demonstrates that machine learning, when coupled with domain knowledge and experimental validation, can transcend black-box prediction to become a diagnostic tool for uncovering the governing mechanisms of complex fluid–structure systems. The findings further suggest that optimal model selection for longline gear design should be guided by the specific output variable and the underlying physical phenomena of interest, rather than by aggregate performance metrics alone.

Furthermore, a significant non-additive interaction between line diameter and material stiffness is identified. At smaller diameters, an increase in material stiffness (e.g., PLA vs. Nylon) results in a significant reduction in drag, sinking depth, and displacement, exceeding predictions by additive models. This occurs because stiffer materials resist flow-induced curvature, maintaining a more streamlined profile. By contrast, at larger diameters, increases in stiffness have a reduced effect as the geometric frontal area dominates the hydrodynamic forces. This interaction highlights a critical design compromise: improving sustainability with biodegradable materials like PLA may inadvertently enhance hydrodynamic performance through stiffness, whereas conventional materials may require larger diameters for strength, thereby increasing drag.

This novel contribution of this study extends beyond methodological innovation by illustrating how interpretable, nonlinear ML models, such as RF, LightGBM, and SVM-RBF, can reveal hidden operational thresholds and design trade-offs that were previously hidden by linear assumptions or qualitative observations. By combining high-fidelity experimental data with machine learning, this study transforms longline gear design from an empirical, trial-and-error process into a predictive assessment of longline performance. These models enhance predictive accuracy and yield practical insights for improving gear performance, enabling fisheries managers and engineers to forecast performance under complex hydrodynamic conditions before producing physical prototypes. This study quantitatively evaluated the hydrodynamic and dynamic behavior of tuna longline gear in a controlled flume. It clarified the interactions among different design parameters, such as diameter, material, length, and lead weight, which influence drag forces, sinking depth, and x-displacement patterns at different flow velocities. Our findings validate and extend previous flume-based experiments and reveal mechanism-based connections between gear geometry, hydrodynamics, and fishing efficiency [1,12,15].

The pronounced overfitting for structural responses (sinking depth, x-displacement) suggests that larger experimental datasets, potentially augmented via bootstrapping or physics-informed synthetic data generation, are needed to achieve robust generalization for these high-variance targets. To mitigate overfitting in future applications, several strategies should be considered: (1) increasing the experimental dataset size (minimum 5000–10,000 samples for stable kernel-based regression); (2) implementing cross-validated early stopping for gradient boosting models; (3) using simpler model architectures (e.g., linear or polynomial regression with regularization) for small datasets; (4) employing ensemble methods that average across models to reduce variance. The current dataset of 1296 points, while adequate for Random Forest, is marginal for SVM-RBF, explaining the observed generalization failures.

4.3. Hydrodynamic Performance of Optimized Longline Gear Reveals Velocity-Independent Deformation, Transient Sinking Kinetics, and Tension-Induced Load Redistribution

The constant deformation values observed across all three flow velocities (0.14, 0.35, and 0.49 m/s) for both mainline and branch line configurations demonstrate that the tested designs achieve instantaneous geometric equilibrium upon flow exposure [20]. The present results extend this understanding by showing that under the specific configurations tested (optimized 1.75 mm PLA diameter, strategic tension distribution), such velocity-dependent fluctuations can be entirely suppressed. The absence of flutter, vortex-induced vibrations, or transient settling responses contrasts with conventional expectations for flexible longline systems. In traditional longline deployments, current-induced deformation is a well-documented phenomenon. Research on midwater float systems has demonstrated that each part of the longline becomes shallower with faster water current, with the principal factor affecting hook shoaling being the spread of horizontal spacing between successive buoys [21,22]. The present results suggest that geometry-dominated designs characterized by optimized diameter reduction and stiffer materials can overcome these hydrodynamic effects, maintaining shape integrity across the entire tested velocity range. The constant normalized deformation value of 100 across all conditions also provides important insight. As noted in the results, this constancy suggests that the deformation metric operates as a normalized index rather than an absolute displacement measure. This interpretation resolves the apparent discrepancy between the observed constant values and the expectation of substantial x-displacement (180–200 cm) at 0.49 m/s predicted for less rigid designs. For future studies, this finding underscores the importance of clearly distinguishing between normalized indices and absolute displacement measurements when reporting longline deformation. The divergence in sinking depth trajectories across configurations from instantaneous equilibrium (Mainline 1) to progressive deepening throughout the 40-s observation window (Sinking Depths 2–5) carries significant implications for hook depth prediction models [20]. However, the present results reveal that such steady-state predictions may systematically underestimate actual hook depths for configurations exhibiting positive temporal gradients ($ds/dt > 0$) throughout deployment.

The systematic variation in sinking rates and final depths across configurations suggests that material stiffness, diameter reduction, and lead line weight distribution directly modulate transient sinking kinetics. Configurations employing higher-stiffness materials (e.g., PLA) showed more pronounced progressive deepening, likely because these materials resist immediate bowing, allowing gravitational forces to progressively overcome initial hydrodynamic forces. This is supported by numerical simulation studies, which have shown that the depth of longline gear with midwater floats is more susceptible to current effects than in conventional settings, primarily because the in-water weight is reduced by the float buoyancy [22]. In commercial longline fisheries, deployment durations may be shorter than the time required to reach equilibrium depth, meaning that steady-state models could systematically underestimate hook depths. This aligns with the dynamic simulation results for longline deployment, which emphasize that the stronger the current, the greater the spatial fluctuations in the location of the fishing gear [23].

4.4. Load Redistribution and Stability Dynamics in Longline Systems Reveal the Effects of Tension

The weight distribution analyses revealed fundamentally different load-transfer mechanisms between branch-line and mainline configurations as added weights increased. Branch line results showed a redistribution pattern in which W1 and W6 captured disproportionately larger shares of total weight at 14–28 g added weight, followed by a gradual decline beyond 42 g, indicating a saturation limit in load concentration. In contrast, mainline results revealed a monotonic load contraction toward proximal nodes, with distal nodes (ML5–ML6) becoming mechanically unloaded as added weight increased. This tension-induced lift-off effect, in which increased gravitational load reduces the number of active weight-bearing nodes, is a novel finding not extensively documented in previous longline literature. The effect likely arises from the geometric straightening of the mainline under higher tension, lifting distal sections away from sensor contacts or shifting the tension distribution toward the attachment origin. This is consistent with mathematical models of hook-and-line fishing gears, which describe how tension distribution along the mainline determines the spatial configuration of the entire gear system [24].

4.5. Practical Implementation Challenges

While the developed machine learning frameworks demonstrate high predictive fidelity for specific hydrodynamic targets, transitioning these data-driven surrogates from controlled flume tank environments to practical, real-world engineering applications requires acknowledging several interpretive cautions and implementation constraints.

First, predictive skill is inherently task-dependent and not uniform across all response variables. The models for drag and lift coefficients generalize robustly, whereas the structural-response models (sinking depth and x-displacement) exhibit appreciably larger performance gaps between training and testing phases. This indicates a susceptibility to overfitting and a more limited generalization capacity for highly dynamic, high-variance structural targets. Consequently, in practical engineering workflows, the hydrodynamic coefficient predictions should be treated as the most reliable outputs of the present framework, while structural-response predictions should be regarded as indicative, necessitating further feature engineering or expanded datasets for definitive deployment. Second, the physical mechanisms invoked in this study, such as tension-induced curvature, geometric straightening, and tension-induced load redistribution are data-driven interpretations consistent with the measured load and displacement patterns, rather than quantities measured directly via specialized instrumentation. These should be understood as physically plausible hypotheses that warrant dedicated, targeted mechanical testing before being codified as established design principles. Finally, the practical deployment of these algorithms depends heavily on the specific operational constraints, computational resources, and data availability of the end-user. As summarized in Table 13, each algorithm presents distinct trade-offs. For instance, while LightGBM offers superior predictive accuracy for complex tabular data, it demands rigorous hyperparameter tuning and cross-validation to prevent overfitting on smaller datasets. Conversely, Random Forest provides a highly robust, computationally efficient baseline with native interpretability, making it more suitable for rapid deployment or environments with limited tuning budgets. SVM-RBF, while effective for capturing smooth nonlinearities in moderate dimensions, scales poorly and lacks native feature importance, restricting its practical utility primarily to offline, well-regularized design optimization rather than real-time monitoring.

Table 13. Advantages, disadvantages, and recommended use context for the three machine-learning algorithms compared in this study.

Algorithm	Advantages	Disadvantages	Recommended Use Context
Random Forest (RF)	Robust with minimal tuning; handles nonlinearity, interactions and mixed feature types; provides native feature importance; averaging reduces variance and limits overfitting	Generally less accurate than boosting on structured tabular data; extrapolates poorly outside the training range; larger model size and slower inference than single trees	A fast, robust baseline; suitable for moderate-sized datasets and when importance-based interpretation is needed under a limited tuning budget
LightGBM	Typically the highest accuracy on tabular data; fast leaf-wise, histogram-based training; scales to many features; built-in L1/L2 regularization and early stopping	Sensitive to several hyperparameters; can overfit small or noisy datasets without early stopping; requires cross-validated tuning to perform well	Preferred when predictive accuracy is the priority and a tuning/cross-validation budget is available; well suited to nonlinear tabular regression with interactions
SVM-RBF	Effective for smooth nonlinear relationships; performs well in moderate dimensions; robust generalization when C and gamma are properly regularized	Requires feature standardization; sensitive to C and gamma; scales poorly to large samples; no native feature-importance measure; slower inference	Preferred for smaller datasets with smooth nonlinear structure, after standardization, when a well-regularized margin-based model is desired

5. Conclusions

This study demonstrates that machine learning algorithms provide a robust, data-driven framework for predicting the hydrodynamic and structural behavior of tuna longline gear, with model performance proving task-dependent: LightGBM achieved superior accuracy for branch line hydrodynamic coefficients (drag: whole-dataset R^2 = 0.8315; lift: R^2 = 0.7958), efficiently capturing nonlinear fluid–structure interactions, whereas Random Forest provided the most reliable generalization for structural responses (sinking depth and x-displacement), mitigating the severe overfitting observed in SVM-RBF. Feature importance analysis confirmed physical plausibility, identifying flow velocity, lead-line weight, material stiffness, and geometric parameters as dominant drivers consistent with classical hydrodynamic theory. Beyond predictive accuracy, the integration of ML with experimental data revealed three key physical insights: optimized PLA-integrated configurations exhibited velocity-independent deformation stability, maintaining geometric equilibrium across tested flow velocities; sinking kinetics were configuration-dependent, with progressive deepening persisting throughout deployment windows and challenging steady-state depth assumptions; and a tension-induced load redistribution effect was identified, wherein increased gravitational loading contracted load-bearing toward proximal nodes. These findings support a paradigm shift from empirical ballasting to physics-informed gear optimization, recommending thin-diameter PLA-integrated lines to enhance hydrodynamic efficiency, coupled with algorithm-specific model selection and time-dependent sinking models for accurate hook-depth prediction. Because the durability, degradation behaviour, mechanical longevity, and field performance of PLA were not evaluated here, the potential of biodegradable PLA to reduce marine plastic pollution is presented as a prospective benefit that requires dedicated investigation. Future work should validate these findings through full-scale sea trials, assess the long-term degradation and strength of PLA un-

https://doi.org/10.3390/fishes11070424

der operational conditions, and incorporate dimensionless parameters to further improve model generalization.

Author Contributions: Conceptualization, A.J., T.B.N.N. and L.S.; methodology, A.J., T.B.N.N. and N.A.E.; software, A.J., T.B.N.N. and L.S.; validation, T.B.N.N., T.M.J. and L.S.; formal analysis, A.J., T.B.N.N. and N.A.E.; investigation, A.J., T.B.N.N. and L.S.; resources, T.B.N.N. and L.S.; data curation, A.J., J.C.N.N. and T.B.N.N.; writing—original draft preparation, A.J. and T.B.N.N.; writing—review and editing, A.J., T.B.N.N. and L.S.; visualization, T.B.N.N.; supervision, L.S.; project administration, L.S.; funding acquisition, L.S. All authors have read and agreed to the published version of the manuscript.

Funding: The project was funded by the National Key R & D Program (2023YFD2401301) and National Natural Science Foundation of China (Project No.32273185).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The original contributions presented in this study are included in the article. Further inquiries can be directed to the corresponding author.

Acknowledgments: The authors would like to thank Gibril Sesay and Zhang Bo of the College of Marine Living Resource Sciences and Management, Shanghai Ocean University for valuable discussions and suggestions throughout research development. The authors also wish to thank the anonymous reviewers and Editor, for their comments and corrections, which helped improve this manuscript.

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

References

1. Lee, C.W.; Lee, J.H.; Cha, B.J.; Kim, H.Y.; Lee, J.H. Physical modeling for underwater flexible systems dynamic simulation. *Ocean Eng.* **2005**, *32*, 331–347. [CrossRef]
2. Moe-Føre, H.; Christian Endresen, P.; Gunnar Aarsæther, K.; Jensen, J.; Føre, M.; Kristiansen, D.; Fredheim, A.; Lader, P.; Reite, K.J. Structural analysis of aquaculture nets: Comparison and validation of different numerical modeling approaches. *J. Offshore Mech. Arct. Eng.* **2015**, *137*, 041201. [CrossRef]
3. James, S.C.; Zhang, Y.; O'Donncha, F. A machine learning framework to forecast wave conditions. *Coast. Eng.* **2018**, *137*, 1–10. [CrossRef]
4. Breiman, L. Random Forests. *Mach. Learn.* **2001**, *45*, 5–32. [CrossRef]
5. Ke, G. LightGBM: A Highly Efficient Gradient Boosting Decision Tree. In *Advances in Neural Information Processing Systems*; Neural Information Processing Systems Foundation: La Jolla, CA, USA, 2017; Volume 30.
6. Ankit; Datta, N. Free transverse vibration of ocean tower. *Ocean Eng.* **2015**, *107*, 271–282. [CrossRef]
7. Tauti, M. A Relation between Experiments on Model and on Full Scale of Fishing Net. *Nippon Suisan Gakkaishi* **1934**, *3*, 171–177. [CrossRef]
8. Song, L.; Qi, Y.; Li, J.; Shen, Z.; Zhang, X.; Shen, X. Dynamic simulation of pelagic longline retrieval. *J. Ocean Univ. China* **2019**, *18*, 455–466. [CrossRef]
9. Sala, A.; De Carlo, F.; Buglioni, G.; Lucchetti, A. Energy performance evaluation of fishing vessels by fuel mass flow measuring system. *Ocean Eng.* **2011**, *38*, 804–809. [CrossRef]
10. Yu, M.; Tang, Y.; Min, M.; Herrmann, B.; Cerbule, K.; Liu, C.; Dou, Y.; Zhang, L. Comparison of physical properties and fishing performance between biodegradable PLA and conventional PA trammel nets in grey mullet (*Mugil cephalus*) and red-lip mullet (*Liza haematocheila*) fishery. *Mar. Pollut. Bull.* **2023**, *195*, 115545. [CrossRef] [PubMed]
11. Guo, D.; Li, Y.; Zhou, Q.; Yu, Z.; Liu, X.; Dong, S.; Zhang, S.; Sung, H.-K.; Yao, Z.; Li, Y.; et al. Degradable, biocompatible, and flexible capacitive pressure sensor for intelligent gait recognition and rehabilitation training. *Nano Energy* **2024**, *127*, 109750. [CrossRef]
12. O'Neill, B.; Saurel, C. *Using Hydrodynamics to Develop More Selective Fishing Gears*; DTU Aqua: Lyngby, Denmark, 2024.
13. Scott, M.; Cardona, E.; Scidmore-Rossing, K.; Royer, M.; Stahl, J.; Hutchinson, M. What's the catch? Examining optimal longline fishing gear configurations to minimize negative impacts on non-target species. *Mar. Policy* **2022**, *143*, 105186. [CrossRef]

14. Lee, J.H.; Lee, C.W.; Cha, B.J. Dynamic simulation of tuna longline gear using numerical methods. *Fish. Sci.* **2005**, *71*, 1287–1294. [CrossRef]
15. Winger, P.D.; DeLouche, H.; Legge, G. Designing and testing new fishing gears: The value of a flume tank. *Mar. Technol. Soc. J.* **2006**, *40*, 44–49. [CrossRef]
16. Cepowski, T. The prediction of ship added resistance at the preliminary design stage by the use of an artificial neural network. *Ocean Eng.* **2020**, *195*, 106657. [CrossRef]
17. Ding, X.; Liu, J.; Yang, F.; Cao, J. Random radial basis function kernel-based support vector machine. *J. Frankl. Inst.* **2021**, *358*, 10121–10140. [CrossRef]
18. Parker, R.W.R.; Tyedmers, P.H. Fuel consumption of global fishing fleets: Current understanding and knowledge gaps. *Fish Fish.* **2015**, *16*, 684–696.
19. Nsangue, B.T.N.; Tang, H.; Mouangue, R.; Liu, W.; Pandong, A.N.; Xu, L.; Hu, F.; Tcham, L. Non-linear dynamic behavior of T0 and T90 mesopelagic trawls based on the Hilbert–Huang transform. *Mar. Struct.* **2025**, *100*, 103727. [CrossRef]
20. Song, L.; Li, J.; Xu, W.; Zhang, X. The dynamic simulation of the pelagic longline deployment. *Fish. Res.* **2015**, *167*, 280–292. [CrossRef]
21. Shiga, M.; Shiode, D.; Miyamoto, Y.; Uchida, K.; Hu, F.; Tokai, T. Effect of water flow on 3-dimensional underwater shape of pelagic longline with midwater float. *Nippon Suisan Gakkaishi* **2009**, *75*, 179–190. [CrossRef]
22. Shiga, M.; Hu, F.; Ito, A.; Shiode, D.; Tokai, T. Effect of Water Flow on Shape of Tuna Longline with Midwater Float System Evaluated in Flume Tank Experiment and Numerical Simulation. *Fish. Eng.* **2008**, *44*, 185–195.
23. Song, L.; Li, Y. Dynamic numerical simulation of the pelagic longline settlement process based on the Runge-Kutta method. *J. Fish. Sci. China* **2022**, *29*, 157–169. [CrossRef]
24. Gabruk, V.I.; Kudakaev, V.V. Mathematical Models That Underlie Computer Simulation of the Hook and Line Fishing Gears. *Ocean Polar Res.* **2019**, *41*, 19–34.