2nd International Symposium on AI and Fluid Mechanics

AIFLUIDs 2026 • Chania, Crete • 31 August – 3 September 2026

Session 11

Industrial & Applied ML in Fluid Systems

S11-A2 Jialei Yan — Deep learning for the dynamic characteristics and stability of floating structures under complex marine environments

S11-A3 Xiao Xue — Uni-Flow: A Unified Autoregressive-Diffusion Framework for Stable Long-Horizon and High-Resolution Multiscale Flow Modeling

S11-A5 Xiao Xiao — Development of an ANN–ISAT Framework for Real-Fluid Modeling with Application to NH■–H■ injection system.

S11-A6 Louis Fliessbach — An automated CFD workflow for the generation of high-fidelity AI training data

S11-A7 Clemens Naumann — Exergy loss prediction in stratified thermal energy storage systems using ANNs

S11-A8 Mirza Popovac — , Numerical Strategies for Improving the Accuracy of ANN-Based Surrogate Model for Airflow Predictions within Rectangular Wall-Bounded Enclosures

S11-A9 Justo Matheus — Identifying Arrhythmic Conditions Through Shock Signatures in Rotary Steerable Drilling Systems

S11-A10 Haohan Chen — Multi-Agent Collaborative CAE Industrial Software Platform for Fluid Mechanics

S11-A11 Armand de Villeroché — Anchored-Branched Steady-state WInd Flow Transformer (AB-SWIFT): a metamodel for 3D atmospheric flow in urban environments

S11-P1 Ron Barron — Improving PINN Stability for Convection-Dominated Boundary Value Problems via Artificial Damping Regularization

S11-P4 Jing Li — Sparse Rational Learning for Aerodynamic Prediction

Deep learning for the dynamic characteristics and stability of floating structures under complex marine environments

[1, 2] Jialei Yan*; [1, 2] Ruicong Wu; [1, 2] Chenyu Lu;

[1]*Key Lab of Smart Prevention and Mitigation of Civil Engineering Disasters of the Ministry of Industry and Information Technology, School of Civil Engineering, Harbin Institute of Technology, Harbin, 150090, China*

[2]*Ministry-of-Education Key Laboratory of Structural Dynamic Behavior and Control, School of Civil Engineering, Harbin Institute of Technology, Harbin 150090, China*

Abstract

This study proposes a stability analysis method for floating structures under complex marine environments, focusing on pontoon submergence caused by water elevation changes owing to external excitations. The results indicate that the roll motion is more prone to reaching the critical submergence angle, with increasing wave heights considerably elevating exceedance risks at vulnerable locations. Numerous "flat zones" exhibiting minimal water elevation variations were identified around the pontoon, thereby expanding under multi-hazard structural radiation effects. Although earthquakes induce severe water vibrations, causing considerable translational responses, they are unlikely to submerge the pontoon within the surrounding flow field. Using the peak ground acceleration (PGA), wave height, and wave period under combined earthquake–wave actions as target features, a random forest machine learning approach is developed to predict the water elevation of the target structure, thereby establishing a rapid stability prediction method. The model achieves excellent accuracy, with the mean squared error (MSE) approaching zero, relative error (RE) below 5%, coefficient of determination (R^2) reaching 0.99, and a nearly vertical cumulative distribution function (CDF) slope at zero error, thereby confirming its reliable predictive capability.

Keywords:

Deep-water floating structures, spatial rotational motion, stability analysis, machine learning, random forest.

Uni-Flow: A Unified Autoregressive-Diffusion Framework for Stable Long-Horizon and High-Resolution Multiscale Flow Modeling

Xiao Xue*[1]; Tianyue Yang[1]; Mingyang Gao[1]; Maida Wang[1]; Kewei Zhu[1]; Peter V. Coveney[1,2]

1. University College London, London, UK

2.University of Amsterdam, Amsterdam, the Netherlands

**Corresponding author: x.xue@ucl.ac.uk*

Abstract

Accurately modeling multiscale spatiotemporal flows remains a central challenge in fluid mechanics, particularly when long-horizon temporal stability and fine-scale spatial fidelity must be achieved simultaneously. Existing operator-learning models often suffer from autoregressive instability, while diffusion-based approaches lack explicit mechanisms for causal time evolution.

We introduce **Uni-Flow**, a unified autoregressive-diffusion framework that explicitly decouples temporal evolution from spatial refinement. A physics-informed autoregressive module evolves low-resolution latent dynamics to preserve large-scale structure and ensure stable long-horizon rollouts. A conditional diffusion refinement module then reconstructs high-resolution physical fields, recovering fine-scale structures through a small number of deterministic denoising steps. This separation resolves the long-standing tension between temporal stability and spatial accuracy in data-driven flow modeling. We validate Uni-Flow across progressively complex systems: (i) two-dimensional Kolmogorov turbulence at Re = 1000, (ii) three-dimensional turbulent channel inflow generation at = 180, and (iii) patient-specific pulsatile aortic flow simulated using a high-fidelity lattice Boltzmann solver. Across all cases, Uni-Flow preserves invariant statistics, including energy spectra, probability density functions, and autocorrelation structures, while significantly outperforming purely autoregressive baselines in long-term stability. In the cardiovascular case, Uni-Flow reduces a multi-hour, multi-GPU hemodynamic simulation to second-scale inference on a single GPU, achieving task-level over five orders of magnitude wall-clock acceleration while maintaining physiologically consistent pressure dynamics. These results establish Uni-Flow as a general and architecture-agnostic framework for generative modeling of complex multiscale flows, enabling stable, high-fidelity, and computationally efficient surrogate simulations across turbulence and physiological fluid dynamics.

Keywords: Generative AI, Neural Operators, Koopman Operator, Turbulent flows, Hemodynamics, Diffusion models

Development of an ANN–ISAT Framework for Real-Fluid Modeling with Application to NH_3–H_2 injection system.

[1]Xiao Xiao *; [1]Bruno Delhom 2; [1]Chaouki Habchi 3; [1]Julien Bohbot 4

[1]IFPEN Rueil-Malmaison – France

Abstract

Two-phase flow configurations pose significant challenges for numerical modeling, necessitating the development of robust Real Fluid Models (RFM) capable of accurately describing flows under subcritical, transcritical, and supercritical thermodynamic conditions. In such regimes, strong property gradients, phase change phenomena, and non-ideal thermodynamic effects must be captured with high fidelity to ensure predictive simulations. To address these complexities, a dedicated RFM framework has recently been developed at IFPEN, providing a consistent formulation for real-fluid thermodynamics within compressible flow solvers. In the present work, a neural network (NN) strategy is introduced to overcome the inherent limitations associated with conventional thermodynamic property lookup tables, which often suffer from high memory requirements, interpolation errors, and limited scalability when extended to multi-component mixtures. The proposed data-driven approach enables continuous and differentiable predictions of thermodynamic properties, thereby improving numerical stability and computational efficiency. Furthermore, a systematic methodology is proposed for generating an optimized training dataset through the coupling of a vapor–liquid equilibrium (VLE) thermodynamic solver with the In-Situ Adaptive Tabulation (ISAT) technique. By selectively enriching the dataset in highly non-linear regions, the approach enhances the robustness of the NN training process while preventing spurious oscillations and noisy predictions. The proposed NN framework has been implemented in a three-dimensional CFD solver and applied to an NH_3–H_2 injection system. The results demonstrate that coupling the NN-based thermodynamic model with the RFM framework enables stable and accurate simulations of NH_3–H_2 spray dynamics. This validation constitutes a crucial step toward extending the methodology to more complex reactive and combustion configurations, which are currently under active investigation.

Keywords:

Subcritical, supercritical, transcritical flow, NH3-H2 injection, real fluid model, VLE, Neural network, ISAT

An automated CFD workflow for the generation of high-fidelity AI training data

[1]Louis Fliessbach*; [1]Marian Fuchs; [1]Hendrik Hetmann; [1]Charles Mockett; [1]Nilakantha Sahoo;

[1]*Upstream CFD GmbH, Berlin / Germany*

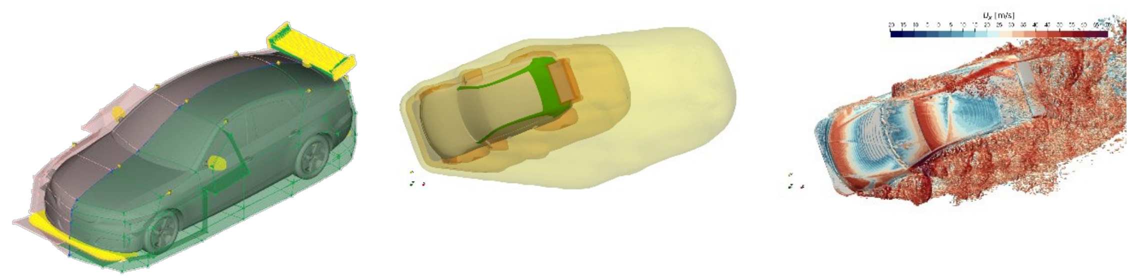
Abstract

Machine learning approaches have the potential to significantly accelerate aerodynamic analysis and design via rapid prediction models. Data-driven methods are, however, fundamentally constrained by the accuracy and consistency of the training data. High-quality CFD datasets, alongside validated and automated workflows for their generation, therefore remain a key prerequisite for reliable AI models.

This contribution presents an end-to-end automated workflow for the generation of high-fidelity CFD datasets targeting AI surrogate modelling. The workflow integrates geometry parametrisation and morphing, fully automated meshing, scale-resolving CFD simulation, and cost-optimised data handling on cloud-based infrastructure, allowing for systematic variation of geometry and flow conditions. A core component is the high-fidelity simulation methodology, which combines methods and best practices developed and validated over many years.

An early version of this workflow was developed for the creation of the open-source DrivAerML dataset [1]. More than 31 000 downloads in January 2026 alone underline the strong demand for such datasets. For DrivAerML, project partners provided geometries and meshes as external inputs to the customised OpenFOAM-based simulation framework. Since then, a fully automated snappyHexMesh-based meshing pipeline and an integrated ANSA-based geometry morphing process have been added, resulting in a single automated workflow covering all essential steps of dataset generation.

The contribution will discuss the individual workflow components and their interaction, as well as its application to a modified DrivAer-hpF geometry [2] for the generation of a proprietary dataset, demonstrating transferability to different base geometries. A brief review of the usage of the DrivAerML dataset since its release will also be given.



a) Parameterised morphing boxes

b) Volume refinement boxes for meshing

c) Instantaneous flow structures of scale-resolving simulation

Figure 1: Overview of different workflow components applied to DrivAer-hpF geometry.

Keywords: high-fidelity CFD, dataset generation, automation, AI training, training data

References

[1] N. Ashton, C. Mockett, M. Fuchs, L. Fliessbach, H. Hetmann, T. Knacke, N. Schonwald, V. Skaperdas, G. Fotiadis, A. Walle, B. Hupertz and D. Maddix, “DrivAerML: High-Fidelity Computational Fluid Dynamics Dataset for Road-Car External Aerodynamics,” *arXiv,* vol. 2408.11969, 2025.

[2] S. Rijns, T.-R. Teschner, K. Blackburn, A. R. Proenca and J. Brighton, “Experimental and numerical investigation of the aerodynamic characteristics of high-performance vehicle configurations under yaw conditions.,” *Physics of Fluids,* vol. 36 (4), no. 045112, 2024.

Acknowledgement

The authors would like to thank Vangelis Skaperda and Grigoris Fotiadis from ANSA and acknowledge their valuable support and input regarding the ANSA morphing workflow. In addition, we would like to thank Steven Rijns from Cranfield University for the fruitful exchange on his experimental and numerical studies of the DrivAer-hpF geometry.

The further development of the workflow was partially funded by an anonymous AI startup.

Other parts of this work were conducted within the EXASIM ("Exascale-fähige Softwarewerkzeuge zur Strömungssimulation im industriellen Design- und Optimierungsprozess") project (2022-2025 https://exasim-project.com) funded by the Federal Ministry of Education and Research under grant number 16ME0677 and the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Commission. Neither the European Union nor European commission can be held responsible for them.

Exergy loss prediction in stratified thermal energy storage systems using ANNs

[1]Clemens Naumann*; [1]Christian Cierpka

[1]*Institute of Thermodynamics and Fluid Mechanics, Technische Universität Ilmenau, Am Helmholtzring 1, 98693 Ilmenau, Germany*

Abstract

Energy storage facilities play a crucial role in storing surplus energy by volatile, renewable energy sources. A promising approach for that is the stratified thermal energy storage (TES), which exploits temperature-dependent density differences of the liquid storage medium to store a hot fluid layer over a cold fluid layer. This increases the total exergy of the system compared to a tank at uniform temperature. However, this is limited due to exergy losses by mixing of both fluid layers. This can be caused by heat transfer to the surrounding, or near-wall parasitic convection as a consequence of the high thermal conductivity of the storage tank wall, even when the system is neither charged nor discharged. To develop a precise performance model of such systems, the prediction of exergy destruction due to these effects is crucial, which is the overall aim of the presented work.

For that, a parametrized 2D axisymmetric model of a stratified TES for numerical CFD simulations is developed, where the total exergy content and the degree of stratification is calculated over time during the standby phase for different storage configurations. A configuration set of 200 samples with varying storage height, aspect ratio, heat loss to the surrounding, tank wall thickness and initial degree of thermal stratification is generated using Latin hypercube sampling. With these configurations, transient simulations with a total simulation duration of 24 hours are performed using the commercial software ANSYS fluent. The resulting time-dependent normalized exergy content and degree of stratification in dependence of the mentioned parameters act as training data for a multilayer perceptron artificial neural network. The final contribution will include a hyperparameter optimization of the neural network, including an analysis of the prediction accuracy, as well as a sensitivity analysis of the investigated system parameters in relation to the predicted exergy loss.

Keywords: Thermal energy storage, computational fluid dynamics, multilayer perceptron, exergy loss prediction

Numerical Strategies for Improving the Accuracy of ANN-Based Surrogate Model for Airflow Predictions within Rectangular Wall-Bounded Enclosures

M. Popovac*, T. Bäuml and D. Šimić

AIT-Austrian Institute of Technology, Center for Transport Technologies, Vienna, Austria

Abstract

Accurate and efficient prediction of airflow characteristics inside electric vehicle (EV) cabins is essential for thermal comfort assessment and HVAC design. While high-fidelity computational fluid dynamics (CFD) simulations provide detailed flow information, they remain computationally expensive for real-time or iterative design applications. The surrogate models based on artificial neural network (ANN) offer a promising alternative, provided that sufficient accuracy and physical consistency can be achieved. This paper presents numerical considerations aimed at improving the predictive accuracy of an ANN surrogate model for fluid flow quantities inside a simplified electric vehicle cabin.

The numerical investigation presented here builds upon previous work in which the ANN architecture and training framework were developed using CFD data generated with OpenFOAM for rectangular wall-bounded enclosures that resemble a simplified electric vehicle cabin. While the baseline model demonstrated encouraging performance, inaccuracies were observed in regions dominated by strong shear layers and flow separation.

To address the noted shortcomings, the present study introduces targeted modifications to the loss function used during training. First, a weighted loss treatment is applied to flow quantities in shear-layer regions, which are treated as statistical outliers due to their steep gradients and high sensitivity. This approach improves the network's ability to capture localized flow features without degrading global accuracy. Second, an additional physics-based loss component is incorporated by recasting the ANN-predicted velocity and pressure fields into residuals of the Navier–Stokes equations. These residuals are then compared to corresponding residuals computed from the CFD reference data, resulting in a supervised, physics-informed learning strategy. Unlike classical physics-informed neural networks, this approach retains full supervision while enforcing consistency with governing equations.

The paper details the numerical formulation of the modified loss functions and their Python-based implementation within the ANN training pipeline. Quantitative results for the simplified EV cabin flow demonstrate that the proposed numerical treatments significantly reduce prediction errors, particularly in shear-dominated regions, thereby enhancing the robustness and physical fidelity of the surrogate model

*Corresponding author: mirza.popovac@ait.ac.at

Identifying Arrhythmic Conditions Through Shock Signatures in Rotary Steerable Drilling Systems

[1, 2]Justo Matheus*;[2]Francisco-Javier Granados-Ortiz;

[1] *SLB, Oslo Innovation Factori, Norway*

[2]*Department of Engineering, University of Almeria, Spain*

Abstract

Rotary Steerable Systems (RSS) operate in increasingly aggressive drilling environments where high-energy mechanical shocks have become frequent rather than exceptional. While shocks are traditionally treated as undesirable disturbances to be mitigated or filtered out, recent advances in high-frequency downhole sensing make it possible to observe these events in unprecedented detail. This presentation introduces a novel diagnostic framework for Rotary Steerable Tools based on the concept of Drilling Electrocardiography (D-ECG), originally proposed by the author as an analogy between physiological health monitoring and the dynamic behavior of drilling systems.

In the D-ECG paradigm, high-frequency shock signatures acquired while drilling are treated as diagnostic waveforms that reflect the internal "health state" of the RSS. By continuously capturing and analyzing shock events at high sampling rates, the method distinguishes between benign, operationally acceptable shocks and anomalous shock patterns that are statistically and dynamically correlated with tool degradation or imminent failure. Rather than relying on threshold-based alarms or post-failure forensics, the proposed approach focuses on pattern recognition, temporal evolution, and repeatability of shock signatures.

The presentation demonstrates how shock morphology, energy distribution, and recurrence can be used to build a diagnostic fingerprint of normal RSS operation, against which deviations are evaluated in real time. The results suggest that shocks, when properly characterized, can serve as an early-warning signal for failure mechanisms such as mechanical looseness, bearing degradation, or structural fatigue. The D-ECG framework provides a systematic and physics-informed approach to RSS diagnostics, enabling condition-based monitoring and opening the door to predictive maintenance strategies while drilling.

Keywords: Rotary Steerable Systems, Oil Drilling, Shocks, Machine Learning, Pattern Recognition, Health Monitoring

Multi-Agent Collaborative CAE Industrial Software Platform for Fluid Mechanics

[1]Haohan Chen; [1, 2]Hao Ma; [3]Xiaojing Wu; [1]Weiwei Zhang*;

[1]*School of Aeronautics, Northwestern Polytechnical University, 710072 Xi'an, China*

[2]*School of Aerospace Engineering, Zhengzhou University of Aeronautics, 450046 Zhengzhou, China*

[3]*Institute of AI for Industries, Chinese Academy of Science, 211135 Nanjing, China*

Abstract

Computational Fluid Dynamics (CFD) plays a pivotal role in engineering simulation, yet its conventional methodologies remain constrained by prohibitive computational costs and limited efficiency. While the integration of artificial intelligence (AI) has substantially accelerated simulation workflows, current AI-augmented tools remain fragmented, lacking a unified platform capable of systematically inheriting domain knowledge and substantial engineering experience. AI-enhanced CFD currently confronts two persistent challenges: insufficient inheritance of modeling expertise and practical know-how, and the difficulties in multi-tool integration and low automation within engineering workflows. To address these issues, this paper proposes an intelligent simulation platform based on multi-agent collaboration. The platform encapsulates core functional modules—including mesh generation, physics solving, surrogate modeling, optimization, and post-processing—into autonomously interacting and cooperative agents. By embedding domain knowledge, enabling seamless tool integration, and facilitating dynamic coordination, the system supports intelligent optimization and adaptive execution of simulation workflows. The platform has been validated across multiple representative engineering scenarios, such as aerodynamic design of aircraft, rapid design optimization, and fluid–structure interaction analyses. This research establishes an efficient, adaptable, and scalable intelligent CFD system, offering a unified and advanced solution for addressing complex multiphysics engineering challenges.

Keywords: Multi-Agent Systems, Large Language Models, Computer-Aided Engineering, Fluid Mechanics, Simulation and Design, Industrial Software

Anchored-Branched Steady-state WInd Flow Transformer (AB-SWIFT): a metamodel for 3D atmospheric flow in urban environments

[1]Armand de Villeroché*; [1]Rem-Sophia Mouradi;

[2]Vincent Le Guen; [1]Sibo Cheng; [1]Marc Bocquet; [1]Alban Farchi;

[3]Patrick Armand; [1]Patrick Massin

[1]*CEREA, ENPC, EDF R&D, Institut Polytechnique de Paris, Île-de-France, France.*

[2]*EDF R&D, Île-de-France, France.*

[3]*CEA, DAM, DIF, F-91297 Arpajon, France.*

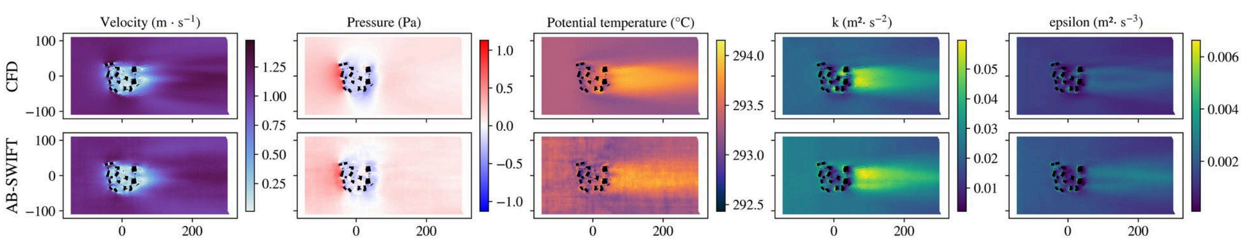
Abstract

In atmospheric environment studies, many applications rely on detailed 3D models of the atmospheric flow at a local scale, which can be simulated using Computational Fluid Dynamics (CFD). As CFD simulations can be computationally intensive, deep learning surrogates are appealing to speed up computation and reduce costs.

However, the air flow at a local scale depends on the local geometry, i.e. topography and buildings affecting the flow, as well as meteorological conditions such as atmospheric stability. Additionally, CFD simulations of this flow can use refined meshes of tens of millions of cells, which can be burdensome to scale up with surrogate models.

To tackle these challenges, we present Anchored Branched Steady-state WInd Flow Transformer (AB-SWIFT), a transformer-based surrogate model for three-dimensional air flow at a local scale. AB-SWIFT has an internal branch structure specifically designed for atmospheric flow modeling, with dedicated branches to encode geometry and meteorological conditions. AB-SWIFT also relies on anchor attention, enabling scaling to large meshes.

We challenge our model on a specially designed database of atmospheric simulations around randomized urban geometries and with a mixture of unstable, neutral, and stable atmospheric stratifications. Our model reaches the best accuracy on all predicted fields compared to state-of-the-art transformers and graph-based models.



Horizontal slice at *h=2* m above ground of an AB-SWIFT prediction on an unseen geometry under stable atmospheric conditions.

Keywords:

Computational fluid dynamics, deep learning, geometric deep learning, transformers, atmospheric dispersion, atmospheric stratification stability, atmospheric boundary layer

Improving PINN Stability for Convection-Dominated Boundary Value Problems via Artificial Damping Regularization

[1]Love Fadia; [1]Vatsal Shah [2,3]Ronald Barron*; [4]Mohammad Hassanzadeh

[1]*Department of Electrical and Computer Engineering, University of Windsor, Canada*

[2]*Department of Mechanical, Automotive and Materials Engineering, University of Windsor, Canada*

[4] *Department of Electrical and Computer Engineering / Mathematics and Computer Science, Lawrence Technological University, U.S.A.*

Keywords: artificial damping, PINN solvers, convection–diffusion equation

Abstract. Physics-Informed Neural Networks (PINNs) have emerged as a flexible framework for solving partial differential equations (PDEs) by embedding the equations into the training objective of neural networks. PINNs provide a mesh-free alternative to classical discretization methods and are especially attractive when generating labeled solutions by other means is expensive or solutions are required on multiple geometries. However, standard PINNs can be difficult to train for stiff PDEs, particularly in convection-dominated regimes where sharp gradients form and small numerical errors can amplify into oscillatory predictions. This research develops a stabilized PINN workflow for steady two-dimensional PDEs and investigates how artificial damping improves robustness and accuracy. Poisson's equation is used to validate the procedure under controlled conditions. The framework is then extended to a convection–diffusion setting where the solution exhibits boundary-layer behaviour for small diffusion. To address instability and spurious oscillations, the physics loss is augmented with artificial damping stabilization terms, analogous to classical 2nd and 4th-order artificial dissipation used in numerical methods. A systematic sweep over multiple damping coefficient combinations is performed to quantify the effect of stabilization on convergence and predictive accuracy, enabling comparison between convection-dominated and diffusion-dominated behaviour using the same network architecture and training procedure.

1. Introduction

Since the early work by Lagaris et al [1], physics-informed neural networks (PINNs) have become an active area of research because of their potential to compete with traditional mesh-based partial differential equation (PDE) solvers. PINNs incorporate the governing PDEs, boundary conditions, and available data directly into the training objective of a neural network [2-5]. Unlike conventional mesh-based solvers, PINNs are flexible across geometries, naturally support inverse as well as forward problems [6], and are especially attractive when labeled solution data are scarce or expensive to generate. These features have made PINNs an important component of the broader physics-informed machine learning landscape. In a recent review article, Ren et al. [7] demonstrated how a hybrid architecture embedding finite difference modules into neural networks mitigates non-physical oscillations, outperforming pure data-driven models in complex geometry simulations. Such hybrid strategies inherit important features of numerical methods while leveraging the mesh-free advantage of deep learning.

Despite these advantages, convection-dominated boundary value problems remain difficult for both classical numerical methods and neural PDE solvers. When convection overwhelms diffusion, the solution often develops thin boundary or interior layers with very large gradients. In such regimes, standard central difference discretization is well known to produce nonphysical oscillations unless additional stabilization is introduced. PINNs face a related difficulty: the governing equations become numerically stiff, the loss terms can become poorly balanced during optimization, and localized sharp features are often under-resolved by standard training procedures. Classical stabilization strategies for convection-dominated convection-diffusion equations include upwind differencing, residual-based streamline stabilization, and artificial dissipation. Wei et al [8] introduced linear inviscid damping and Zhang [9] has discussed stabilization and convergence of PINNs.

*Corresponding Author, R. Barron, az3@uwindsor.ca

Motivated by these challenges, this work investigates whether artificial damping can improve the robustness of PINN solvers for steady two-dimensional PDEs. The study first considers a manufactured Poisson problem to verify the implementation in a controlled setting and to test whether higher-order damping preserves solution accuracy when the target field is smooth. The framework is then extended to a convection-diffusion benchmark with known analytical solution and increasingly sharp gradient behaviour as the Reynolds number increases. In this setting, the physics loss is augmented with 2nd-order and 4th-order damping terms, analogous to classical artificial dissipation operators, and the damping coefficients are systematically varied to study their effect on stability, residual reduction, and predictive accuracy.

2. Artificial (Linear) Damping

Guermond et al. [10] introduced the concept of entropy viscosity to stabilize mesh-based numerical simulations of conservation laws. Wang et al. [11] applied the idea to large eddy simulations and Wang et al. [12] extended the entropy viscosity method to improve the performance of PINNs. Hoffman and Chiang [13] have discussed the modified PDE in finite difference formulations and related it to suppressing oscillations produced by finite difference solutions of the Navier-Stokes equations. For convection-diffusion equations, the concept of the modified PDE can be used to rationalize the addition of artificial implicit (2nd-order) and/or explicit (4th-order) linear damping, in the form

$$u\frac{\partial \emptyset}{\partial x} + v\frac{\partial \emptyset}{\partial y} + \varepsilon_i\left[(\Delta x)^2\frac{\partial^2\emptyset}{\partial x^2} + (\Delta y)^2\frac{\partial^2\emptyset}{\partial y^2}\right] = \Gamma\nabla^2\emptyset + S(x,y) - \varepsilon_e\left[(\Delta x)^4\frac{\partial^4\emptyset}{\partial x^4} + (\Delta y)^4\frac{\partial^4\emptyset}{\partial y^4}\right] \quad (1)$$

where the values taken for the damping coefficients ε_i and ε_e depend on the discretization method and are problem-dependent. In this research, we consider two examples using equation (1) to investigate the effects of linear damping on the accuracy and stability of PINN solutions.

3. Results and Discussion

3.1 Poisson equation: (a manufactured BVP)

The exact solution of the Poisson equation

$$\nabla^2\emptyset = -\{2 + \pi^2 x(1-x)\}cos(\pi y) \quad (2)$$

on the domain $-1 \le x \le 1$, $-1 \le y \le 1$, subject to the Dirichlet boundary conditions

$$\emptyset(-1,y) = -2\cos(\pi y),\ \emptyset(1,y) = 0,\ \emptyset(x,-1) = -x(1-x),\ \emptyset(x,1) = -x(1-x) \quad (3)$$

is given by

$$\emptyset = x(1-x)cos(\pi y). \quad (4)$$

Setting $u = 0, v = 0, \Gamma = 1$ and $\varepsilon_i = 0$ in equation (1) gives the modified Poisson equation with explicit 4th-order linear damping,

$$\nabla^2\emptyset = -S(x,y) + \varepsilon_e\left[(\Delta x)^4\frac{\partial^4\emptyset}{\partial x^4} + (\Delta y)^4\frac{\partial^4\emptyset}{\partial y^4}\right]. \quad (5)$$

Figure 1a shows the difference between the exact solution and the predicted solution from the PINN on a 129x129 grid, with $\varepsilon_e = 0.15$. The relatively larger errors near (-1, -1) are expected since the solution exhibits its steepest gradients in this region. However, the solution is symmetric about y = 0, and the PINN fails to capture this feature of the solution. A line plot of the solution along the line $x = 0.5$ is shown in figure 1b, indicating excellent agreement between the predicted and true solution along this line, except perhaps near the lower boundary (y = -1) where a soft boundary constraint has been implemented. These figures and other similar results demonstrate that 4th-order artificial damping does not degrade the overall accuracy of the solution prediction.

In mesh-based simulations, the damping coefficient is known to play a significant role in determining the stabilization and accuracy of the approximate solution [11]. A sensitivity analysis was performed over a range of ε_e values, several of which are shown on table 1. The solution values along the vertical line $x = 0.5$ are tabulated for different y-values and compared to a finite difference solution obtained on a 129x129 Cartesian grid. It appears that 4th-order damping has an insignificant effect on the solution accuracy, except perhaps near $y = 0.5$, which lies on the boundary line $x + y = 1$. This may be attributed

*Corresponding Author, R. Barron, az3@uwindsor.ca

to a decrease in accuracy for the evaluation of the 4th-derivatives in the damping term, but needs further investigation.

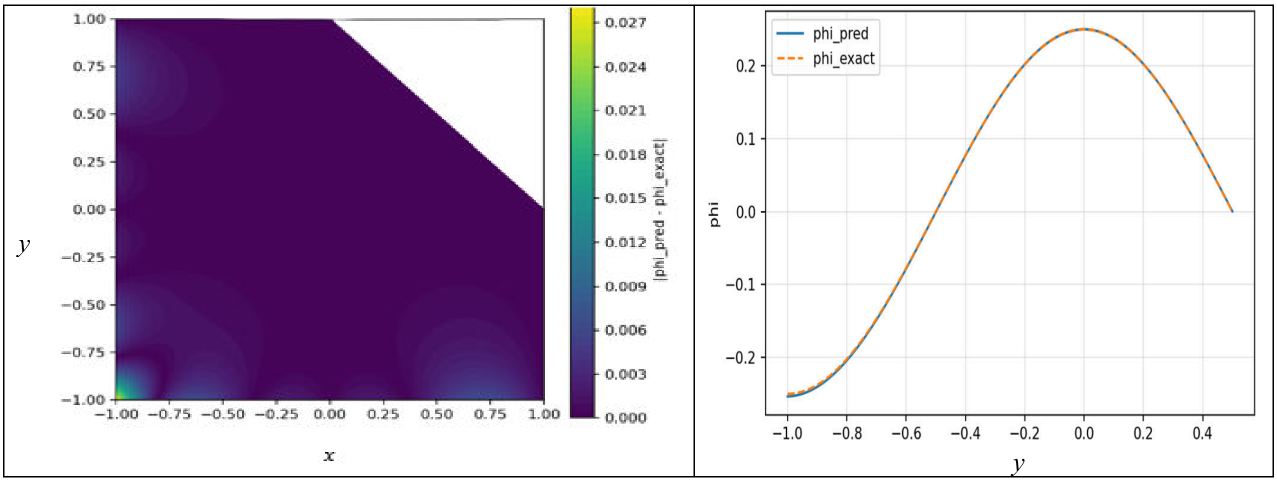


Figure 1 a) Error plot for PINN solution of Poisson equation on a 129x129 grid, b) comparison of predicted and exact solutions along $x = 0.5$, ($\varepsilon_i = 0, \varepsilon_e = 0.15$).

Table 1. Solution along $x = 0.5$, and metrics, illustrating the effect of the 4th-order linear damping coefficient on the PINN solution of Poisson equation.

y	FD, $\varepsilon_e = 0.0$	$\varepsilon_e = 0.0$	$\varepsilon_e = 0.15$	$\varepsilon_e = 0.5$	Exact soln.
-1.0	-0.25	-0.25002	-0.25347	-0.25248	-0.25
-0.2969	0.14893	0.14893	0.14862	0.14872	0.14893
0.0	0.25002	0.25001	0.25	0.25	0.25
0.5	-2.17e-6	-1.73e-6	4.87e-4	4.12e-4	0.0
Metrics					
Loss	n/a	-	7.41e-5	7.69e-4	n/a
MSE	n/a	-	1.76e-6	8.68e-7	n/a
MAE	2.23e-5	-	6.24e-4	4.29e-4	n/a

3.2 Ge & Zhang's convection-diffusion problem

To investigate the combined effect of 2nd-order implicit damping and 4th-order explicit damping, consider the manufactured convection-diffusion problem posed by Ge and Zhang [14], with the additional damping terms,

$$u\frac{\partial\emptyset}{\partial x} + v\frac{\partial\emptyset}{\partial y} + \varepsilon_i\left[(\Delta x)^2\frac{\partial^2\emptyset}{\partial x^2} + (\Delta y)^2\frac{\partial^2\emptyset}{\partial y^2}\right] - \frac{1}{Re}\nabla^2\emptyset = S(x,y) - \varepsilon_e\left[(\Delta x)^4\frac{\partial^4\emptyset}{\partial x^4} + (\Delta y)^4\frac{\partial^4\emptyset}{\partial y^4}\right] \quad (6)$$

with the convective velocity given by

$$u(x,y) = x(x-1)(1-2y), \qquad v(x,y) = -y(y-1)(1-2x).$$

The exact solution is taken as

$$\emptyset = [1 - b(x)][1 - b(y)] \quad (7)$$

where $b(\xi) = \left(e^{Re(1-\xi)} - 1\right)/\left(e^{Re} - 1\right)$ for $\xi = x$ or y, and *Re* is the Reynolds number. Substituting this solution into the non-damped PDE yields the corresponding source term,

*Corresponding Author, R. Barron, az3@uwindsor.ca

$$S(x, y) = -(u + 1)b'(x)[1 - b(y)] - (v + 1)b'(y)[1 - b(x)].$$

Figure 2 uses a bar graph of the mean absolute error (MAE) at *Re* = 10, to illustrate the effects due to different combinations of the damping coefficients and various differentiation methods for the damping terms. Values of MAE range from 9.04e-1 for $\varepsilon_i = 0.04, \varepsilon_e = 0.08$ and a combination of finite differencing and automatic differentiation, to 5.17e-5 for no artificial damping.

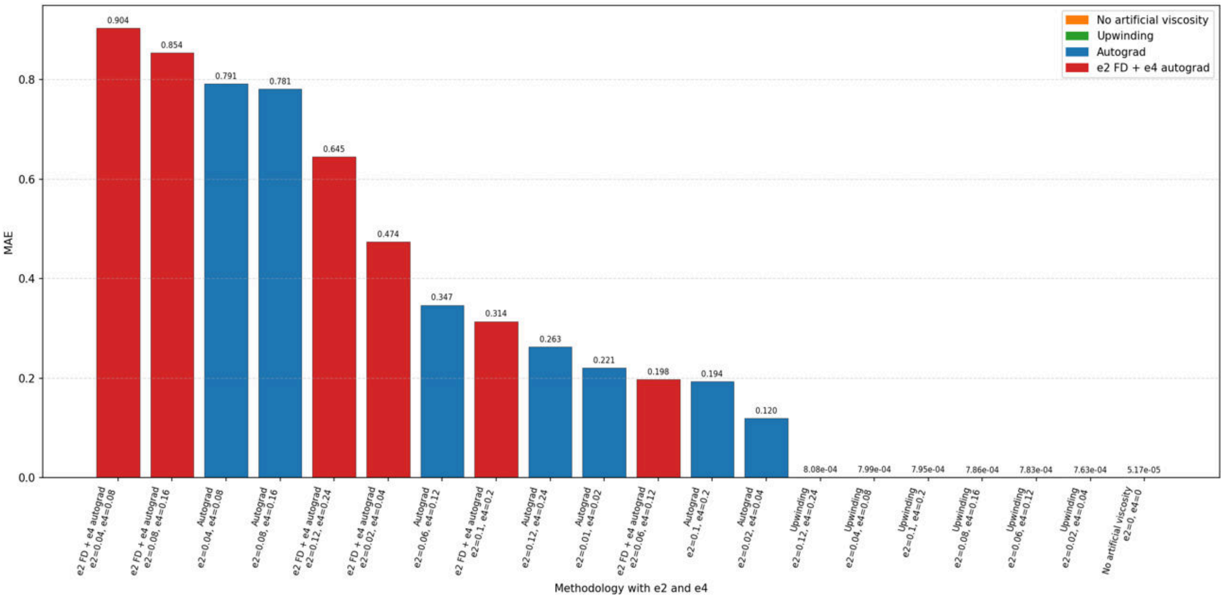


Figure 2 Bar chart showing the effect of artificial linear damping, and the method of evaluating the damping terms. (*Re* = 10)

The MAE is further tabulated in table 2, for different *Re* and combinations of ε_i and ε_e. At *Re* = 1, the PINN with explicit 4th-order damping only ($\varepsilon_i = 0.0, \varepsilon_e = 0.14$) and automatic differentiation of the 4th-derivatives yields the best result with an MAE of 1.05e-6. On the other hand, the addition of 2nd-order damping appears to increase the MAE by a factor of at least 24. For *Re* = 10, the lowest MAE corresponds to no artificial damping. The *Re* = 100 case is problematic for all combinations of damping coefficients, as well as the finite difference method. This is due to the extremely large solution gradients near the boundaries $x = 0$ and $y = 0$ for this case, where the solution rises from zero at the boundary to its maximum value of one within very thin boundary layers on these boundaries.

Table 2. MAE, illustrating the effect of the 2nd- and 4th-order linear damping coefficients on the PINN solution of the convection-diffusion equation of Ge and Zhang [14] at different Reynolds numbers. (FD = finite difference, AD = automatic differentiation)

Model / MAE	*Re* = 1	*Re* = 10	*Re* = 100
FD, $\varepsilon_i = 0.0, \varepsilon_e = 0.0$	1.18e-5	7.72e-4	2.53e-2
PINN, $\varepsilon_i = 0.0, \varepsilon_e = 0.0$	1.92e-5	5.34e-5	8.14e-3
PINN + AD			
$\varepsilon_i = 0.0, \varepsilon_e = 0.14$	1.05e-6	6.93e-4	2.19e-1
$\varepsilon_i = 0.0, \varepsilon_e = 0.8$	1.05e-6	7.10e-4	2.40e-1
$\varepsilon_i = 0.02, \varepsilon_e = 0.04$	2.50e-5	7.45e-4	4.32e-3
$\varepsilon_i = 0.04, \varepsilon_e = 0.08$	3.75e-3	--	--
$\varepsilon_i = 0.06, \varepsilon_e = 0.12$	--	6.66e-3	--
$\varepsilon_i = 0.1, \varepsilon_e = 0.2$	--	--	2.71e-1

*Corresponding Author, R. Barron, az3@uwindsor.ca

PINN + FD(ε_i and ε_e)			
$\varepsilon_i = 0.04, \varepsilon_e = 0.08$	1.75e-5	7.80e-4	8.47e-3
PINN + FD(ε_i) + AD(ε_e)			
$\varepsilon_i = 0.1, \varepsilon_e = 0.2$	6.32e-4	1.88e-1	9.56e-2

4. Conclusions

The central hypothesis of this paper is that controlled artificial dissipation can regularize PINN training in convection-dominated flow regimes without destroying the physically relevant structure of the solution. The implementation is reproducible and scalable, using deterministic seeding and a two-stage optimizer strategy. Model performance is evaluated using field-wise error metrics and qualitative visualization. By comparing unstabilized and stabilized models under the same network architecture and optimization strategy, this study aims to clarify when artificial damping acts as a beneficial regularizer and when it becomes overly diffusive. In doing so, this paper connects ideas from stabilized numerical PDE solvers with current efforts to improve the reliability of physics-informed learning for stiff transport problems. Preliminary tests suggest that adding controlled artificial dissipation to PINN residuals can improve training stability and solution quality for stiff PDEs, while maintaining generality across both Poisson-type and convection–diffusion formulations.

References

1. I. Lagaris, A. Likas, D. I. Fotiadis, ‘Artificial neural networks for solving ordinary and partial differential equations’, IEEE Transactions on Neural Networks. 9, 987-1000, 1998. http://dx.doi.org/10.1109/72.712178
2. J. Wu, J. Wang, H. Xiao, J. Ling, ‘Physics-informed machine learning for predictive turbulence modeling: A priori assessment of prediction confidence’, Physical Rev. Fluids. 2016. DOI: https://arxiv.org/abs/1607.04563
3. M. Raissi, P. Perdikaris, G. E. Karniadakis, ‘Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations’, J. Comp. Phys. 378, 686-707, 2019. DOI: 10.1016/j.jcp.2018.10.045.
4. G. E. Karniadakis, I. G. Kevrekidis, L. Lu, P. Perdikaris, W. Wang, L. Yang, ‘Physics-informed machine learning’, Nat. Rev. Phys. 3, 422–440, 2021. https://doi.org/10.1038/s42254-021-00314-5
5. S. Wang, H. Wang, P. Perdikaris, ‘Learning the solution operator of parametric partial differential equations with physics-informed DeepONets’, Sci. Adv. 7, eabi8605, 2021.
6. Q. He, A. M. Tartakovsky, ‘Physics-informed neural network method for forward and backward advection-dispersion equations’, Water Resources Research. 57, 2021. e2020WR029479. https://doi.org/10.1029/2020WR029479
7. Z. Ren, S. Zhou, D. Liu, Q. Liu, ‘Physics-informed neural networks: A review of methodological evolution, theoretical foundations, and interdisciplinary frontiers toward next-generation scientific computing’, Appl. Sci. 15, 8092-8109, 2025. https://doi.org/10.3390/app15148092
8. D. Wei, Z, Zhang, W. Zhao, ‘Linear inviscid damping for a class of monotone shear flow in Sobolev spaces’, Comm. Pure Appl. Math. 71, 617-687, 2018.
9. X. Zhang, ‘Enhancing stability and convergence of physics-informed neural networks’, Proc. of the Nat. Acad. of Sci. 119, 11, 2022.
10. J. L. Guermond, R. Pasquetti, B. Popov, ‘Entropy viscosity method for nonlinear conservation law’, J. Comp. Phys. 230 (11), 4248–4267, 2011.
11. Z. Wang, M. S. Triantafyllou, Y. Constantinides, G. E. Karniadakis, ‘An entropy-viscosity large eddy simulation study of turbulent flow in a flexible pipe’, J. Fluid Mech. 859, 691-730, 2019. doi:10.1017/jfm.2018.808
12. Z. Wang, X. Meng, X. Jiang, ‘Solution multiplicity and effects of data and eddy viscosity on Navier-Stokes solutions inferred by physics-informed neural networks’, arXiv2309.06010v1, 2023.
13. K. A. Hoffmann, S. T. Chiang, Computational Fluid Dynamics, Vol. 1, 4th Edition, Engineering Education Systems, Wichita, Kansas, U.S.A., 2000.
14. L. Ge, J. Zhang, ‘Accuracy, robustness and efficiency comparison in iterative computation of convection diffusion equation with boundary layers’, Num. Methods in PDEs. 16(4), 379-394, 2000. https://doi.org/10.1002/1098-2426(200007)16:4<379::AID-NUM3>3.0.CO;2-I.

*Corresponding Author, R. Barron, az3@uwindsor.ca

Proceedings of the 2nd international Symposium on AI and Fluid Mechanics
Paper No [leave blank]

Rational Function Symbolic Regression for Interpretable Aerodynamic Modeling: A Unified Algebraic Linearization Framework

[1]Jing Li; [1]Changtong Luo*;

[1]State Key Laboratory of High Temperature Gas Dynamics, Institute of Mechanics, Chinese Academy of Sciences, Beijing 100190, China; School of Engineering Science, University of Chinese Academy of Sciences, Beijing 100049, China; Corresponding author email: luo@imech.ac.cn

Abstract. Interpretable aerodynamic surrogate models are essential for design, control, and physical understanding of modern flight vehicles, yet existing approaches face a structural dilemma: polynomial regression cannot represent the rapid response variation that characterizes transonic flow; deep neural networks are accurate but opaque; genetic-programming-based symbolic regression introduces stochasticity and often returns deeply nested expressions. Rational functions offer a natural middle ground, but direct nonlinear fitting is ill-conditioned and combinatorial term selection is NP-hard. This work presents a unified algebraic linearization framework for rational function symbolic regression, and shows that two seemingly different rational learning approaches, enumerative Rational Learning and Sparse Rational Learning, are in fact two complementary algorithmic routes that share this linearization and differ only in how they select the sparse structure of the rational expression. The drag coefficient prediction of a NACA 0012 airfoil from the public NASA PALMO database is used as the validation case. Numerical results show that both routes independently select the squared sine of the angle of attack as the dominant numerator contribution and place Mach-dependent terms inside the denominator, which is consistent with the even drag symmetry of a symmetric airfoil and provides a rational modulation of the upper transonic drag rise. The refined enumerative Rational Learning can get the most accurate model while the Sparse Rational Learning usually results in more concise rational expressions.

Keywords: Symbolic regression; Rational function; Aerodynamic surrogate modeling; Algebraic linearization; Sparse optimization; Interpretability

1. Introduction

Aerodynamic surrogate models compress expensive flow simulations and wind tunnel measurements into compact mathematical representations that underpin trajectory design, control system synthesis, configuration optimization, and certification analyses for modern flight vehicles. For most of these uses, a surrogate must not only deliver accurate predictions but also remain interpretable. Interpretability allows engineers to verify physical consistency, transfer the model to neighboring configurations, and extract design rules from the observed aerodynamic response.

Statistical and machine learning based surrogates such as Kriging, radial basis functions, support vector machines, and deep neural networks deliver high accuracy on dense datasets but yield black box mappings that resist physical inspection. Polynomial regression is fully transparent yet fundamentally limited. A fixed degree polynomial cannot reproduce the rapid response variation that characterizes the transonic drag rise, and increasing the degree quickly produces spurious oscillations. Genetic-programming-based symbolic regression [1, 2] offers explicit expressions, but widely used implementations such as PySR [3] and Operon [4] remain stochastic and often return deeply nested expressions.

Rational functions occupy a natural middle ground. They have also been studied as practical multivariate approximation models for data-driven surrogate construction [5]. Many aerodynamic relations take rational form by construction, including the oblique shock and Prandtl Meyer relations in compressible flow and the Prandtl lifting line correction for finite aspect ratio wings [6]. Rational forms are also able to express rapid response modulation through a denominator that varies with the input, which makes them well suited for phenomena such as the transonic drag rise that occurs as the free stream Mach number approaches the drag divergence regime.

*Corresponding Author, Changtong Luo, luo@imech.ac.cn

Despite these structural advantages, fitting a rational model from data is difficult for two reasons. First, the denominator turns ordinary least squares into a nonconvex problem, and direct nonlinear optimization is poorly initialized and sensitive to local minima. Second, choosing which monomials enter the numerator and the denominator is a combinatorial selection problem whose cost scales exponentially with the candidate library size.

This short paper presents a unified algebraic linearization framework that transforms the nonlinear rational fitting into a linear least squares problem on an augmented feature space and leaves the discrete structure selection as a separate step. Two rational learning approaches recently developed by the authors fit naturally into this framework as complementary search strategies that share the same linearization but select the sparse structure at different stages. The enumerative route, hereafter referred to as Rational Learning [7], fixes a candidate structure in advance, applies the linearization to a small per structure linear system, and refines each candidate through nonlinear least squares. The sparse route, hereafter referred to as Sparse Rational Learning [8], applies the linearization once to the full candidate library and uses an L1 or an L0 penalty to select the support and the coefficients in a single optimization step. The framework is validated on a new aerodynamic case, the drag coefficient of a NACA 0012 airfoil over the transonic range from the public NASA PALMO database. This case is independent from the experimental evidence in the two companion manuscripts and directly examines whether the framework can identify aerodynamically meaningful structure in NACA 0012 transonic drag modeling. The leading discovered terms are consistent with the even drag symmetry of a symmetric airfoil and with the observed upper transonic drag rise, which together serve as the interpretability signature of the framework.

Section 2 develops the algebraic linearization that defines the framework. Section 3 describes the two algorithmic routes and contrasts their structural and computational properties. Section 4 reports the validation on the NACA 0012 drag coefficient. Section 5 discusses route selection and limitations.

2. Unified Framework: Algebraic Linearization

Consider a physical system with input and scalar output . The goal is to discover a multivariate rational function of the form

where and are nonconstant monomial basis functions drawn from a candidate library, is the constant term of the numerator, and is fixed to remove the scaling ambiguity.

2.1. Candidate Library

For a problem of dimension and a chosen maximum total degree , the candidate library

contains monomials, which serve as the common basis pool for both the numerator and the denominator in equation (1). Let denote the size of this library. The numerator and denominator subsets and in equation (1) are then drawn from the nonconstant part of with , and the total number of free coefficients is . In the maximum support case , every nonconstant monomial in the library appears in both the numerator and the denominator, and the coefficient count reaches , which grows combinatorially with both and . For and , already and , so a rational model in three physical inputs already admits a coefficient vector of comparable size to a typical experimental aerodynamic dataset. A useful rational surrogate must therefore activate only a sparse subset of , which is the structural selection problem that the rest of the framework addresses.

2.2. Algebraic Linearization

Multiplying both sides of equation (1) by and rearranging yields a linear identity in the unknown coefficients:

*Corresponding Author, Changtong Luo, luo@imech.ac.cn

Collecting the basis evaluations into an augmented feature vector

the rational fitting task becomes the linear system

The algebraic identity in equation (3) is the linearization trick shared by both algorithmic routes considered in this paper, and is closely related to implicit sparse identification strategies such as SINDy-PI [9]. It holds for any and any choice of monomial subsets and drawn from the nonconstant part of . The two routes in Section 3 differ only in how this freedom is exploited: the enumerative route fixes and the chosen subsets in advance and then invokes equation (5) on the resulting small system, whereas the sparse route always takes and lets a sparsity penalty drive most coefficients in to zero. The nonconvexity that the denominator introduces into the original rational regression is thereby converted into a discrete structure selection problem on top of a linear regression.

2.3. Two Search Strategies on the Same Backbone

Both algorithmic routes considered in this paper search for a sparse rational function of the form in equation (1), with the numerator and denominator monomials drawn from . They differ only in when the sparse support is chosen relative to the algebraic linearization. The first strategy fixes and the chosen monomial subsets in advance, by enumerating admissible combinations under explicit sparsity caps on the numerator and the denominator; the linearization is then applied to each resulting small per structure linear system, which is solved in closed form and refined on the original nonlinear objective. The second strategy takes the maximum support in equation (5), attaches an L1 or an L0 sparsity penalty to the resulting dimensional , and lets the active support and the nonzero coefficients emerge from a single optimization step. The remainder of the paper develops these two strategies in detail and shows that they yield complementary tradeoffs between accuracy, runtime, and complexity within the same rational function search space.

3. Two Algorithmic Routes

3.1. Enumerative Route: Rational Learning

The enumerative route, Rational Learning [7], controls the structure of the rational model through three integer hyperparameters. A maximum total degree , which plays the role of in equation (2), bounds the degree of every monomial drawn from the candidate library. Two route specific caps and then bound the number of monomials retained in the numerator and the denominator respectively. For each admissible pair with and , the route enumerates every choice of numerator monomials and denominator monomials from the library, and evaluates each choice in two stages. In the first stage, the algebraic linearization in equation (3) is applied to the chosen structure, producing a small linear system with design matrix in whose closed form least squares solution provides an initial estimate of the coefficients. This linearized solution is optimal in the augmented feature space, but, as noted in [7], neglects the nonlinear dependence of the denominator on and is therefore generally biased with respect to the original nonlinear loss. In the second stage, the linearized estimate is used as a warm start for a Levenberg Marquardt refinement that minimizes the residual of the original rational model. The route then returns a Pareto frontier indexed by , from which the user selects the operating point that best balances accuracy and complexity.

Algorithm 1. Enumerative Rational Learning with two-stage fitting.
Input: Dataset ; maximum total degree ; sparsity caps and .
Output: Pareto frontier .

Construct the candidate library and initialize .

*Corresponding Author, Changtong Luo, luo@imech.ac.cn

```
for  and  do
    .
    for each numerator support  and denominator support  do
        Assemble .
        Solve .
        Starting from , refine coefficients by Levenberg–Marquardt on the original nonlinear rational loss.
        if the refined loss is smaller than  then
            Update the best structure and coefficients for the current .
        end if
    end for
    Store the best model for the current  in .
end for
Construct the Pareto frontier  from  and return .
```

3.2. Sparse Route: Sparse Rational Learning

The sparse route, Sparse Rational Learning [8], takes the maximum support in equation (5), so that the augmented design matrix contains every nonconstant monomial in , and applies a single sparse optimization to the full that recovers both the support pattern and the coefficients without explicit enumeration. Two variants follow from the choice of sparsity penalty. The L1 variant follows the Lasso [10] and solves the convex problem

which is fast to solve but introduces a shrinkage bias that retains many small coefficients. The L0 variant adopts an data fit, introduces binary indicator variables for each coefficient, and ties each coefficient magnitude to its indicator through a Big M constraint:

The data fit keeps every term in equation (7) linear, so that the program becomes a mixed integer linear program in the spirit of modern best subset and L0 constrained regression [11, 12]. It returns an exactly sparse solution but scales unfavourably with the library size and is therefore restricted to low degree libraries in practice. A final thresholding step with tolerance is applied to both variants, but is most consequential for the L1 solution because the convex shrinkage tends to leave many coefficients with small but nonzero values.

3.3. Side-by-Side Comparison

Table 1. Comparison of the two algorithmic routes that share the algebraic linearization.

Aspect	**Enumerative RL**	**Sparse SRL (L1 / L0)**
Hyperparameters		(and for L0)
Structure selection	Explicit grid search over	Implicit via sparse penalty on
Coefficient solver	Linear LS init + LM refine	Convex QP (L1) or MILP (L0)
Determinism	Yes	Yes
Nonlinear refinement	Yes	No
Scaling with	Combinatorial in	Combinatorial in library size
Best suited for	Compact dictionaries, accuracy critical tasks	Large libraries, fast prototyping

The two routes are complementary rather than competing. Rational Learning provides a guaranteed nonlinear refined fit for each admissible structure, but pays the cost of enumerating every admissible combination, and benefits most when the candidate library is small enough that this enumeration remains affordable. Sparse Rational Learning replaces enumeration with a single sparse optimization on the full augmented system, and is most effective when the library is large and a fast first pass over the design space is preferred to a per structure nonlinear refinement. Section 4 confirms this

*Corresponding Author, Changtong Luo, luo@imech.ac.cn

complementarity on the NACA 0012 drag coefficient by comparing the best configuration of each route on a common dataset and metric.

4. Validation on the NACA 0012 Drag Coefficient

4.1. Dataset and Experimental Setup

The framework is validated on the NACA 0012 subset of the public NASA PALMO database [13]. The subset contains 3280 samples on a tensor grid in angle of attack (41 values), Mach number (10 values), and Reynolds number (8 values), and covers both the subsonic regime and the upper transonic regime where the drag rise occurs. The samples are split into 80 percent training (2624 points) and 20 percent test (656 points) using a fixed random seed.

The input variables are and the target is the drag coefficient . For the main comparison, all routes operate on min max normalized inputs , so that each variable enters on a comparable scale. We write , , and .

Hyperparameter selection is performed through deterministic sweeps on the training set. For Rational Learning, we use a degree three candidate library and the sparsity caps , which evaluate all nontrivial structures with and . Each admissible structure is initialized by the algebraic linearization and then refined by Levenberg Marquardt. For both Sparse Rational Learning variants, we follow a degree–regularization grid search with and . The reported sparse expressions use a coefficient truncation threshold of . The L0 mixed integer runs use and a 60 s time limit per configuration; nonconverged L0 configurations are counted in the sweep runtime but are not eligible as selected models.

4.2. Error Metrics

All errors are evaluated on the original drag coefficient, not on a normalized target. We report the root mean squared error

the coefficient of determination , and a dimensionless relative error

which is more informative than when the target spans nearly two orders of magnitude as in the present dataset. The mean absolute target value on the test set is , the maximum is 0.689, and the minimum is 0.0053.

4.3. Quantitative Comparison

For the quantitative comparison, we report the best configuration from each algorithmic route, selected on the training set, on the 656 sample held-out test set. The selected configurations are with for the refined Rational Learning model, with for the L1 sparse route, and with for the L0 sparse route. Higher degree L0 configurations did not converge within the 60 s mixed integer time limit and are therefore excluded from the main comparison.

table 2 summarizes the test set performance and the symbolic complexity of each model after denormalization. The complexity is the total node count of the SymPy expression tree, which captures both the number of monomial terms and the number of arithmetic operations required to evaluate the model.

Table 2. Quantitative comparison on the NACA 0012 test set (656 samples). Errors are evaluated on the original drag coefficient. Complexity is the SymPy expression tree node count of the denormalized model.

Method	Setting	Complexity		RMSE	rRMSE
RL, refined	,	36	0.9913	0.01343	0.0794

*Corresponding Author, Changtong Luo, luo@imech.ac.cn

Method	Setting	Complexity		RMSE	rRMSE
L1 SRL	,	67	0.9894	0.01486	0.0878
L0 SRL	,	28	0.9841	0.01818	0.1074

Table 3. Runtime and convergence statistics for the hyperparameter sweeps. Selected time is the runtime of the reported configuration: the selected (k,l) enumerative subproblem plus its nonlinear refinement for RL, and the single selected D, λ solver run for SRL. Sweep time is the total runtime over all configurations evaluated by the selection protocol.

Method	Sweep space	Converged / total	Rate	Selected time (s)	Sweep time (s)
RL, refined	15 structures	15 / 15	100%	115.14	165.07
L1 SRL	36 settings	36 / 36	100%	0.147	3.91
L0 SRL	36 settings	18 / 36	50%	9.316	1194.97

The three routes occupy complementary corners of the accuracy complexity time space. The refined Rational Learning model attains the smallest test RMSE at moderate complexity, but its selected enumerative subproblem is substantially slower than a single sparse solve. The L1 sparse route reaches comparable accuracy with the lowest sweep cost, at the price of a noticeably larger expression because the convex shrinkage retains many small coefficients in a degree four library. The L0 sparse route gives the most compact denormalized expression at degree two, but table 3 shows that its full sweep is dominated by the mixed integer scaling and by higher degree configurations that reach the 60 s time limit. Figure 1 places all 69 candidate models from the three routes on a single accuracy complexity plane using $\log10(1 - R^2)$ as the vertical axis.This scale expands the high accuracy region near $R^2 = 1$, so lower points correspond to more accurate models and leftward points correspond to simpler symbolic expressions.The connected markers trace the Pareto front of each route, making the accuracy complexity tradeoff visible beyond the three configurations reported in Table 2.

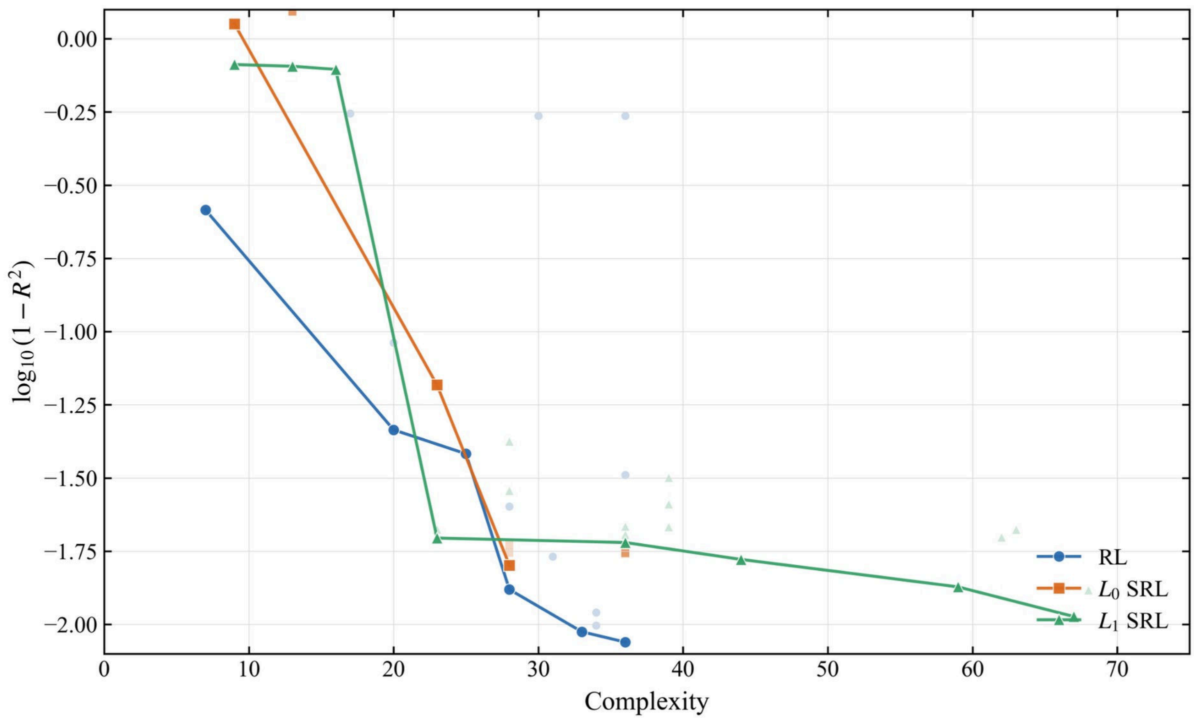


Figure 1. Accuracy complexity tradeoff for all candidate models from the three algorithmic routes on the NACA 0012 drag coefficient. The horizontal axis is the denormalized SymPy expression complexity and the vertical

*Corresponding Author, Changtong Luo, luo@imech.ac.cn

axis is log10(1 − R^2). Light markers show all 69 denormalized candidates (15 Rational Learning, 36 L1 sparse, 18 L0 sparse).

4.4. Interpretable Drag Model

For the interpretability discussion, we report representative expressions from all three routes rather than only the lowest error model in table 2. The representatives are chosen to expose both accuracy and sparsity: a refined Rational Learning model with (complexity 36,), an L1 Sparse Rational Learning model with and (complexity 44,), and an L0 Sparse Rational Learning model with and (complexity 28,). In physical variables, these three representative models are

The contrast between the two sparse representatives illustrates the expected sparsity tradeoff. The L1 model retains a moderately dense degree three expression, whereas the L0 model keeps a smaller support in the normalized coordinates, with two numerator terms and three denominator corrections, and reduces the expression complexity from 44 to 28.

Several properties of equations (10)–(12) agree with established aerodynamic intuition and with the trends visible in the PALMO data. First, all three numerators are dominated by , which is consistent with the even drag symmetry of the symmetric NACA 0012 airfoil. In normalized variables, the corresponding numerator coefficients are 0.388, 0.330, and 0.394 for the RL, L1, and L0 representatives, respectively. The data confirm this trend: the mean drag coefficient grows from 0.027 at to 0.397 at , and remains nearly even about . Second, all three models place Mach-dependent terms inside the denominator, which allows the rational form to amplify the response as approaches the upper end of the sampled transonic range. This denominator modulation reproduces the upper transonic drag rise observed in the data: the mean drag at climbs from 0.022 at to 0.057 at , and to 0.116 at . Third, the Reynolds number dependence is weak compared with the Mach number and angle of attack effects. The RL representative includes , but only through a small numerator correction , or in normalized variables. The two sparse representatives still reach high test accuracy without selecting any term, with for L1 SRL and for L0 SRL. This variable selection pattern indicates that Reynolds number has a secondary influence on in this subset, while and account for most of the observed variation. The structural agreement across three algorithmically different routes provides direct evidence that this decomposition is a property of the data rather than an artefact of any single search strategy.

5. Discussion and Outlook

The results suggest that the two routes should be viewed as complementary tools built on the same algebraic core. Rational Learning is useful when the user wants direct control over the numerator and denominator supports and can afford an explicit search over a modest library. L1 Sparse Rational Learning is attractive for fast screening because it solves the full library problem cheaply, although the resulting expressions may retain more small terms. L0 Sparse Rational Learning is more suitable when compactness is the primary objective, but its mixed integer formulation can become expensive as the library grows.

This unified view also clarifies the connection between rational symbolic regression and data-driven multivariate Pade approximation [5]. Because the denominator is learned together with the numerator, the resulting models can represent amplification, saturation, and asymptotic trends that are difficult to capture with polynomial expressions of comparable size. In the NACA 0012 example, this advantage

*Corresponding Author, Changtong Luo, luo@imech.ac.cn

appears through the denominator modulation of the transonic drag rise and through the consistent selection of the angle of attack contribution in the numerator.

Several limitations remain. The candidate monomial library still grows rapidly with input dimension and polynomial degree, and all methods depend on the quality of the prescribed basis. In addition, when the underlying response is not well represented by a rational form, the recovered expression should be interpreted as an accurate approximation rather than an exact governing structure. Future work will therefore focus on accelerating high dimensional libraries, for example through variable separation and block building strategies [14, 15], incorporating physics informed sparsity priors, and extending the framework to coupled multi output aerodynamic responses.

CRediT authorship contribution statement

Jing Li: Conceptualization, Methodology, Software, Validation, Writing - original draft. **Changtong Luo**: Supervision, Funding acquisition, Writing - review & editing.

Acknowledgements

This work has been supported by the National Natural Science Foundation of China (Grant Nos. 12072353, 12132017).

Companion manuscripts

The detailed derivations, full benchmark suites, and extended case studies underlying this short paper are reported in two companion manuscripts currently under review [7, 8].

References

1. Koza, J. R. (1994). Genetic Programming as a Means for Programming Computers by Natural Selection. Statistics and Computing, 4, 87–112.
2. La Cava, W., Orzechowski, P., Burlacu, B., de França, F. O., Virgolin, M., Jin, Y., Kommenda, M., and Moore, J. H. (2021). Contemporary Symbolic Regression Methods and their Relative Performance. In Proceedings of the NeurIPS Datasets and Benchmarks Track.
3. Cranmer, M. (2023). Interpretable machine learning for science with PySR and SymbolicRegression.jl. arXiv preprint arXiv:2305.01582.
4. Burlacu, B., Kronberger, G., and Kommenda, M. (2020). Operon C++: An Efficient Genetic Programming Framework for Symbolic Regression. Proceedings of the Genetic and Evolutionary Computation Conference Companion.
5. Austin, A. P., Krishnamoorthy, M., Leyffer, S., Mrenna, S., Müller, J., and Schulz, H. (2021). Practical Algorithms for Multivariate Rational Approximation. Computer Physics Communications, 261, 107663.
6. Anderson, J. D. (2016). Fundamentals of Aerodynamics (6th ed.). McGraw-Hill Education.
7. Li, J., and Luo, C. (2026). A Rational Learning Method for Aerodynamic Modeling. Manuscript under review.
8. Li, J., and Luo, C. (2026). Sparse Rational Learning for Interpretable Aerodynamic Surrogate Modeling. Manuscript under review.
9. Kaheman, K., Kutz, J. N., and Brunton, S. L. (2020). SINDy-PI: A Robust Algorithm for Parallel Implicit Sparse Identification of Nonlinear Dynamics. Proceedings of the Royal Society A, 476(2242).
10. Tibshirani, R. (1996). Regression Shrinkage and Selection via the Lasso. Journal of the Royal Statistical Society B, 58(1), 267–288.
11. Bertsimas, D., King, A., and Mazumder, R. (2016). Best Subset Selection via a Modern Optimization Lens. The Annals of Statistics, 44(2), 813–852.
12. Willis, M. J., and von Stosch, M. (2017). Simultaneous Parameter Identification and Discrimination of the Nonparametric Structure of Hybrid Semi-parametric Models. Computers and Chemical Engineering, 104, 366–376.
13. Cornelius, J. K., et al. (2025). NASA PALMO Aerodynamic Database. AIAA SciTech 2025 Forum, AIAA Paper 2025-0038; NASA Technical Memorandum NASA/TM-20240014546.
14. Chen, C., Luo, C., and Jiang, Z. (2018). A Multilevel Block Building Algorithm for Fast Modeling Generalized Separable Systems. Expert Systems with Applications, 109, 25–34.
15. Chen, C., Luo, C., and Jiang, Z. (2018). Block Building Programming for Symbolic Regression. Neurocomputing, 275, 1973–1980.

*Corresponding Author, Changtong Luo, luo@imech.ac.cn