Session 5

Dimensionality Reduction and ROM

Modeling and sensing transonic airfoil buffet flows at high Reynolds numbers on a low-order submanifold

Kai Fukami*, Yuta Iwatani, Soju Maejima, Hiroyuki Asada, and Soshi Kawai

Department of Aerospace Engineering, Graduate School of Engineering, Tohoku University, Sendai, 980-8579, Japan

Abstract

We discuss the low-dimensionality of transonic airfoil buffet flows at high Reynolds numbers for their modeling and sensing through nonlinear machine learning. This study first examines flows over the OAT15A supercritical airfoil at a chord-based Reynolds number of $Re = 3\times 10^6$ with two Mach numbers, $M_\infty = 0.715$ and 0.730. We find that the dynamical behavior of seemingly complex transonic airfoil buffet flows is described in a three-dimensional subspace identified through lift-augmented manifold learning in a physically interpretable manner. While the current model can be leveraged for sensor-based reconstruction, the present submanifold extracts the similarity of transonic airfoil buffet across the Reynolds number, enabling the application of the model trained at $Re = 3\times 10^6$ to the level of a real aircraft operation of $Re = 3\times 10^7$. The current study provides data-driven insights into high-Reynolds number transonic airfoil buffet flows. Details on the current study are referred to Fukami et al. (J. Fluid Mech., 2025).

Keywords: Transonic airfoil buffet, Manifold learning, Low-order modeling, Sparse sensor reconstruction

Estimating phase-amplitude sensitivity fields of laminar airfoil wakes from sparse sensors with machine learning

[1]Daiki Beppu*; [2]Vedasri Godavarthi; [3]Koichiro Yawata; [3]Hiroya Nakao; [1]Kai Fukami;

[1]*Department of Aerospace Engineering, Tohoku University*

[2]*Department of Mechanical Engineering, Johns Hopkins University*

[3]*Department of Systems and Control Engineering, Institute of Science, Tokyo*

Abstract

Phase-amplitude reduction has recently been employed for reduced-order modeling of aerodynamic flows. The sensitivities for the reduced phase and amplitude variables can also be used to derive the optimal control strategies for frequency modification. However, accurate identification of the sensitivity functions is challenging due to the intrinsic high dimensionality of fluid flows. To address this issue, this study proposes a data-driven technique that estimates phase and amplitude sensitivities from sparse sensors placed on an airfoil. Equipped with sparse dynamical modeling and a nonlinear decoder, the proposed approach enables the identification of phase and amplitude sensitivity fields for cases with different airfoil parameters.

Keywords: Post-stall laminar wakes, Data-driven phase-reduction analysis

Data-driven Modal Decomposition of the Effect of Fins in Unsteady Channel Flow Dynamics

[1]F.-J. Granados-Ortiz*, [2]J. Garicano-Mena, [1]A. López-Martínez, [3]J. Ortega-Casanova

[1]*Department of Engineering, University of Almería, 04120 Almeria, Spain*

[2]*ETSI Aeronáutica y del Espacio, Universidad Politécnica de Madrid, 28040 Madrid, Spain*

[3]*Department of Mechanical, Thermal and Fluid Engineering, University of Málaga, 29071 Málaga, Spain*

Abstract

Channel flows are present in many applications, such as fluid mixing, microfluidics or medical applications, among others. When an object is present in the channel (something frequent in micromixing or harvesting devices), unsteady structures may appear, which propagate downstream at certain frequencies and amplitudes. In this context, the presence of fins at different locations, angles and lengths can be of interest to perturb their dynamics. To gain insights on the effect of fins on these vortical structures, data-driven modal decomposition methods are applied to CFD simulation data to extract frequencies, growth rates and/or the most energetic modes, useful to plan control strategies. It has been observed that depending on the set-up of the fins, at a fixed Reynolds number, the effect on the flow can be very notorius, often completely changing the flow dynamics downstream and upstream (see Fig. 1). This analysis is of strong interest in the development of fins-based micromixing devices or energy harvesting control.

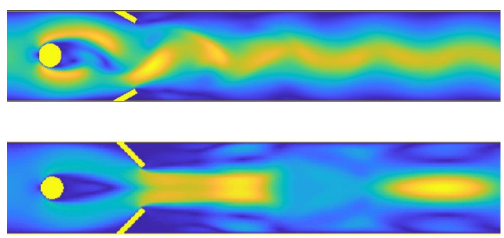


Figure 1. Different patterns in the unsteady flow dynamics after the aggregation of fins.

Keywords: Data-driven, fluid mechanics, Reduced Order Modelling, CFD, channel flows

NEURAL FIELD BASED SURROGATE MODEL FOR DITCHING LOAD PREDICTION

[1]Henning Schwarz*; [1]Pyei Phyo Lin; [2]Jens-Peter M. Zemke; [1]Thomas Rung

[1]Institute for Fluid Dynamics and Ship Theory

Hamburg University of Technology, Am Schwarzenberg-Campus 4

21073 Hamburg, Germany

[2]Institute of Mathematics,

Hamburg University of Technology, Am Schwarzenberg-Campus 3

21073 Hamburg, Germany

Abstract

Data-driven surrogate models for computational fluid dynamics are often based on dimension reduction using convolutional autoencoders (CAE). While such models have shown great potential in various tasks, they usually rely on the same regular grid that is used for all inputs. On the contrary, coordinate-based models like neural fields only process single coordinates at a time and are therefore independent of the discretization. This is a beneficial feature when data examples with different discretizations should be considered, as this does not require to map all examples to a meta mesh that is typically used for a CAE. Moreover, this enables super-resolution reconstructions as the output can be computed at arbitrary points.

This work deals with the development of data-driven surrogate models for the prediction of aircraft ditching loads and deformations. In previous works, aircraft ditching loads were predicted using CAE-LSTM combinations [1] and the interpretability of the CAE-based model was improved by disentangling the latent space [2]. In this work, a conditional neural field approach, where the latent variables are obtained with an auto-decoding strategy, is investigated. The ultimate goal of this approach is to consider different aircraft geometries with different discretizations within a single model.

Keywords: conditional neural field, aircraft ditching

References

[1] H. Schwarz, M. Überrück, J.-P. M. Zemke, and T. Rung. *Machine learning based prediction of ditching loads*, AIAA Journal, 63(5):1835–1854, 2025, https://doi.org/10.2514/1.J064086

[2] H. Schwarz, P. P. Lin, J.-P. M. Zemke and T. Rung. *Disentangled latent spaces for reduced order models using deterministic autoencoders*, Computers & Fluids, 302:106837, 2025, https://doi.org/10.1016/j.compfluid.2025.106837

A New Approach for Fast Multiscale POD Computation

[1,2,3]Marek Belda*; [3]Lorenzo Schena; [3]Romain Poletti; [3]Miguel Alfonso Mendez

[1]*Department of Fluid Dynamics and Thermodynamics, Faculty of Mechanical Engineering, Czech Technical University in Prague, Technická 1902/4, Prague 6 – Dejvice, 160 00, Czech Republic*

[2]*Institute of Thermomechanics, Czech Academy of Sciences, Dolejškova 1402/5, Prague 8 – Libeň, 182 00, Czech Republic*

[3]*The von Karman Institute for Fluid Dynamics, Waterloosesteenweg 72, Sint-Genesius-Rode, B-1640, Belgium*

Abstract

Data-driven decompositions have seen a sharp increase in popularity in recent years. This is especially true for hybrid decomposition methods, such as Multiscale Proper Orthogonal Decomposition (mPOD) discussed in this contribution. However, the classical way of computing the mPOD has unfavorably high computational complexity, with every resolved scale requiring an equivalent of full POD of the data. In this contribution, we introduce a new approach to mPOD computation, called fast mPOD. This approach provides all the benefits of the classical mPOD at a fraction of its computational cost. Both the computational gains and accuracy were assessed on multiple test cases, which confirmed the benefits of the proposed approach.

Keywords: mPOD, hybrid data decomposition methods, spectral domain, computational complexity

The AIAA Data-Driven Modeling of Fluid Flows Community Challenge

[1]Oliver Schmidt* and [2]Aaron Towne;
[3]Adrian Lozano-Duran; [4]Scott T. M. Dawson; [2]Ricardo Vinuesa;

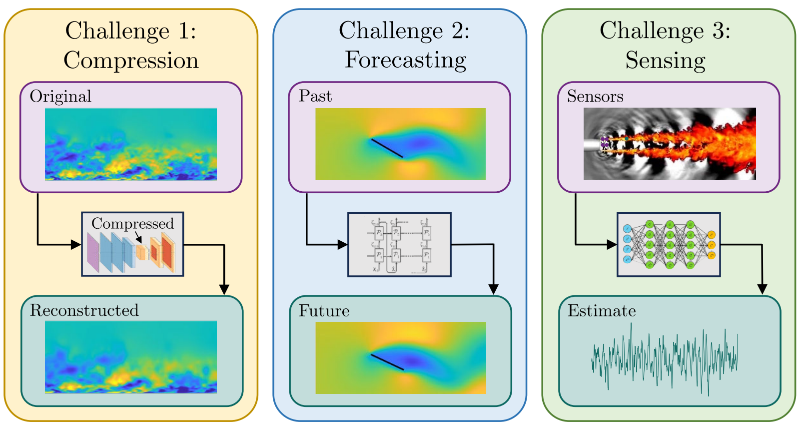
[1]*University of California San Diego, La Jolla, CA*

[2]*University of Michigan, Ann Arbor, MI*

[3]*California Institute of Technology, Pasadena, CA*

[4]*Illinois Institute of Technology, Chicago, IL*

Abstract



We introduce a community challenge designed to facilitate direct comparisons between data-driven methods for compression, forecasting, and sensing of complex aerospace flows. The challenge is organized into three tracks that target these complementary capabilities: compression (compact representations for large datasets), forecasting (predicting future flow states from a finite history), and sensing (inferring unmeasured flow states from limited measurements). Across these tracks, multiple sub-challenges span a diverse set of canonical and applied flow configurations, ranging from turbulent boundary layers and jets to cavity flows and pitching airfoils, thereby emphasizing different modeling requirements (e.g., dimensionality, nonlinearity, parametric variability, and measurement sparsity).The compression track focuses on turbulent boundary layers and includes (1.1) compression of 2D velocity fields from high-Reynolds-number DNS and (1.2) compression of fully 3D velocity fields from the same high-Re DNS database, enabling direct assessment of methods as the data dimensionality and scale increase. Baseline methods include proper orthogonal decomposition (POD) for the 2D case and a convolutional autoencoder (CNN-AE) for the 3D case. The forecasting track targets unsteady turbulent dynamics in two settings: (2.1) a time-resolved PIV dataset of a turbulent open-cavity flow, with dynamic mode decomposition (DMD) as a classical baseline, and (2.2) DNS of pitching airfoils with parametric variations across pitching angles and frequencies, with a long short-term memory (LSTM) network as a baseline sequence model. The sensing track addresses reconstruction from limited measurements in two regimes: (3.1) sparse-measurement sensing in an LES of an axisymmetric turbulent jet, with Wiener filtering and a multi-layer perceptron (MLP) as baselines, and (3.2) reconstruction of a turbulent boundary layer from wall-stress signals using low-Re DNS, with linear stochastic estimation (LSE) and a convolutional neural network (CNN) baseline. The challenge is open to anyone, and we invite broad participation to build a comprehensive and balanced picture of what works and where current methods fall short. To support fair comparisons, we provide standardized success metrics, evaluation tools, and baseline implementations, with one classical and one machine-learning baseline per sub-challenge. Final assessments use blind tests on withheld data, and we explicitly encourage negative results and careful analyses of limitations. Outcomes will be disseminated through an AIAA Journal Virtual Collection and invited presentations at AIAA conferences.

Keywords: community challenge, benchmarking, compression, forecasting, sensing, complex flows, blind test

Parametric Prediction of Turbulent Flows Using a Novel Condition-guided variational Auto-encoder

[1]Haoyan Li*; [2]Alberto Procacci; [3]Alessandro Parente; [4]Eliott Lumet; [5]Melanie C.Rochoux;

[1]Universite Libre de Bruxelles, Ecole Polytechnique de Bruxelles, Aero-Thermo-Mechanics Laboratory, Ixelles, 1050, Belgium

[2]Brussels Institute for Thermal-Fluid Systems and Clean Energy (BRITE), Universite Libre de Bruxelles and Vrije Universiteit Brussele, Ixelles, 1050, Belgium

[3] Universite de Toulouse, Meteo-France/CNRS, CNRM, Toulouse, France

[4]Universite de Toulouse, CNRS/Cerfacs/IRD, CECI, Toulouse, France

Abstract

Predicting turbulent flows under changing boundary conditions is important to atmospheric and urban fluid mechanics, but relying on high-fidelity approaches for this task is computationally expensive. Reduced-order models (ROMs) provide an efficient alternative by compressing the high-dimensional flow fields into a small set of dominant modes. However, existing ROMs generally treat dimensionality reduction and regression as two independent steps. The latent variables are first extracted without any knowledge of the physical parameters, and only afterward is a separate regression model trained to map boundary conditions onto this latent space. This decoupled procedure prevents the latent variables from aligning naturally with physically meaningful variations in the flow and propagates the regression error into the decoder during prediction.

To address these issues, we introduce a condition-guided beta-variational autoencoder (Beta-VAE) that integrates a regression module directly into the encoder-decoder architecture. This end-to-end design allows to couple dimensionality reduction, mapping between physical boundary conditions, and flow-field reconstruction. The model is trained using velocity fields from Large Eddy Simulations (LES) of the Mock Urban Setting Test (MUST) field campaign, covering a broad range of inflow wind directions and friction velocities representative of atmospheric boundary-layer conditions. Through parametric study of regularization and latent dimension, the proposed framework learns a compact and disentangled latent representation in which each latent variable varies smoothly with the underlying physical parameters. This structure allows the latent variables to correspond directly to interpretable flow features. Compared to ROMs developed using Proper Orthogonal Decomposition (POD) and a conventional Beta-VAE, the condition-guided formulation reduces the mean prediction from 0.32, 0.27 to 0.14 and increases the mean coefficient of determination from 0.79 and 0.88 to 0.96, respectively.

The resulting condition-guided ROM provides accurate, stable, and physically interpretable predictions of turbulent flow fields across different boundary conditions. Its combination of predictive performance and mechanistic insight makes it a promising tool for rapid parametric evaluation, uncertainty quantification, and physically grounded analysis of urban and atmospheric turbulent flows.

Keywords: Reduced-Order Model, Parametric Prediction, Turbulent Flows, Variational Autoencoder

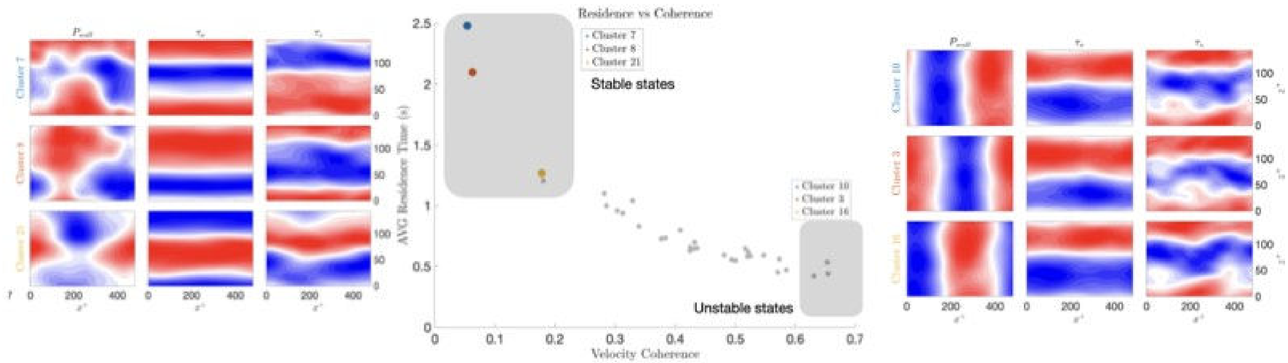
Latent-Space SINDy-Discovered ODEs for Reconstructed Turbulent Dynamics with Causal Analysis of Attractor Transitions

[1]Miguel Pérez Cuadrado*; [1]Giorgio Maria Cavallazzi; [1]Alfredo Pinelli;

[1]*Department of Engineering, School of Science and Technology, City St Georges, University of London, UK*

Near-wall turbulence sustains itself through a regeneration cycle whose dynamics are accessible, in principle, from wall measurements alone. We exploit this by building a reduced-order model driven exclusively by wall pressure and shear stress fluctuations from a minimal turbulent channel at $\Re_\tau=180$. Snapshot POD encodes the fluctuation field into 83 latent modes retaining 95% of the turbulent kinetic energy. A sparse identification framework then learns the latent dynamics from a Mahalanobis RBF library whose basis functions adapt to the local geometry of the phase-space trajectory, correcting both the coordinate bias introduced by POD anisotropy and the temporal sampling bias that arises from non-uniform residence times across the attractor. G-means clustering of the corrected trajectory recovers, without supervision, two dynamically distinct populations: high-residence, low-coherence states corresponding to stable streak attractors, and low-residence, high-coherence states associated with burst-initiating instabilities — a partition that maps cleanly onto the autonomous near-wall cycle of Jiménez & Pinelli (1999). The learned model reproduces DNS statistics faithfully over long integrations, matching wall-shear PDFs, energy spectra, and inter-state transition probabilities, while tracking individual trajectories accurately up to the Lyapunov predictability horizon of approximately $5t^+$. Applying the Liang–Kleeman information-flow framework to the POD temporal coefficients then reveals a sparse, state-dependent causal graph in which a small set of energetically dominant modes drives the regeneration cycle; transition-level analysis identifies the obligatory causal pathways connecting streak instability to turbulent bursting, and the minimal subset of wall-observable modes that sustains them.

Keywords: wall turbulence, reduced-order modelling, SINDy, RBF, attractor, causality, POD



Typical stable and unstable states corresponding to long and short residence time within the invariant regions (classified via clusters). On the left typical "streaky structure" footprint on. the wall (pressure and two wall shear stresses), in the centres the residence time for each sampled cluster. On the right the unstable states, from where the wall cycle escapes quickly.

Experimentally-based least-order model of the transient growth of a laminar separation bubble

[1,2]João V. N. Neiva*; [1]Luc Pastur; [2]Igor B. de Paula; [2]Omar E. H. Pinedo;

[1]*IENSTA Paris, Institute Polytechnique de Paris*

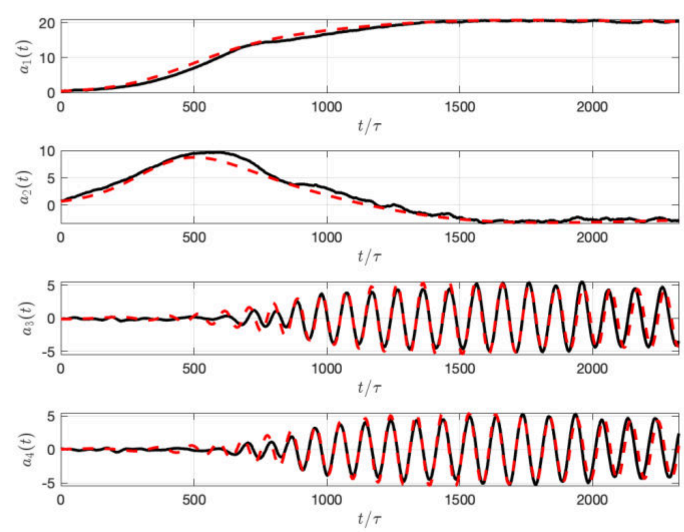
[2]*Pontifícia Universidade Católica do Rio de Janeiro*

Abstract

Laminar separation bubbles (LSBs) have a significant role in the aerodynamic performance of airfoils operating at low Reynolds numbers. In many applications, they are subject to a time-varying level of environmental disturbances, such as in turbine cascades and unmanned aerial vehicles. This time–varying regime is not yet fully described. In this scenario, reduced order models emerge as an useful tool to help describing and modeling the phenomenon. In this work, a four-dimensional and a three-dimensional Galerkin model using Proper Orthogonal Decompostion (POD) and Sparse Identification of Dynamical System (SINDy) are tested for modeling the formation of the LSB. The model is built using experimental data from time-resolved Particle Image Velocimetry measurements of a flow over a flat plate in a water tunnel with a convergent divergent geometry. To this end, the vorticity transport formulation of the incompressible Navier-Stokes equation is used to identify the steady solutions. Two flow states are identified and used as base flow for the modal decomposition, and therefore, for the models. The models rely on POD modes: one or two static modes u1 and u2, depending on whether the model is three or four-dimensional, and two oscillatory modes u3 and u4. The time dynamics of the third and fourth modes indicate an oscillating permanent regime reached after a transient. The resulting four-dimensional POD-Galerkin model correctly predicts the full transient dynamics, from the beginning of the boundary layer separation until the permanent regime with vortex shedding. The three-dimensional model retrieves the famous Landau equation and conjecture a Hopf bifurcation as the route for the dynamical system. Although the model only captures the transient part just before boundary layer transition, it correctly predicts the dynamics.

Keywords:

Fluid Mechanics; Laminar Separation Bubble; Reduced Order Model; POD; SINDy;



Transferable Scaling Function Learning Method for Cross Shape Aerodynamic Database Construction

[1,2,3]Haitao Lin , [1,4]Xu Wang , [1,2,3]Weiwei Zhang*

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

[2] *International Joint Institute of Intelligent Fluid Mechanics , Northwestern Polytechnical University , Xi'an 710072 , China*

[3] *National Key Laboratory of Aircraft Configuration Design , Xi'an 710072 , China*

[4] *Department of Mechanical Engineering , Hong Kong Polytechnic University , Hong Kong 999077*

Abstract

Accurate and comprehensive aerodynamic data is crucial for aircraft performance evaluation, however, the construction cost of high-dimensional and nonlinear aerodynamic databases is high. In order to achieve the construction of a full state nonlinear aerodynamic database across different shapes, this paper proposes a cross shape scaling function learning method. This method first targets aerodynamic datasets of multiple source domain aircraft, and combines symbolic regression method and shape parameter optimization strategy to mine the composite function expression of aerodynamic coefficients. The inner layer function is a scaling function with shape generalization ability, and the outer layer function is an explicit expression of aerodynamic performance with respect to the scaling function. In the modeling of the new shape aircraft database in the target domain, sparse reconstruction of the database is achieved by fine-tuning the parameters of the scaling function through aerodynamic samples. For the high-speed aircraft example, the scaling functions of axial force coefficient, normal force coefficient, and pitch moment coefficient of HB-2 and cone were extracted through pre training, and the aerodynamic model was obtained by fine-tuning sparse aerodynamic samples of different configurations (HBS, double ellipsoid, HyTRV waverider). By using only 16 state samples for each shape, a full state aerodynamic database can be constructed (with a relative error of less than 4%), and cross speed domain extrapolation can be achieved. Verified the appearance generalization ability and state extrapolation ability of the proposed method.

Keywords: Aerodynamic modeling; Scaling function; Nonlinear dimensionality reduction; Symbolic regression; Pre-trained model

Active Flow Control via Model-Based Reinforcement Learning

[1]Guillermo Suarez*; [1]Emre Özkaya; [1]Nico R. Gauger;

[1]*RPTU University Kaiserslautern-Landau*

Abstract

In this work we investigate reinforcement learning (RL) for active flow control (AFC) on the Schäfer and Turek benchmark: a 2D cylinder in a channel flow equipped with two synthetic jets operating under zero-net-mass-flux actuation. This configuration provides a well established test case for drag and vortex-shedding mitigation while retaining moderate computational cost for computational fluid dynamics (CFD).

Reinforcement learning is particularly attractive for AFC because it formulates the problem as a global optimization task and does not require gradient information from the Navier-Stokes solver. However, RL relies on extensive interaction with the environment. When the environment is a CFD solver, this results in an excessive computational time, making direct training impractical for realistic flow configurations.

To address this limitation, we explore replacing the CFD environment with surrogate models that approximate the flow dynamics. Several data-driven approaches are investigated, including Dynamic Mode Decomposition, and Long Short-Term Memory networks. While these methods can capture certain flow features, they are not robust enough to represent the nonlinear, transient dynamics, or they require a considerable amount of data for training.

We therefore investigate the model-based meta-policy optimization framework. Instead of relying on a single surrogate, it trains an ensemble of surrogate models to represent uncertainties in the learned dynamics. A meta-policy is then optimized across this ensemble, improving robustness and generalization before being deployed in the real CFD environment. The proposed approach aims to significantly reduce training cost while preserving a good performance in the real CFD environment.

Keywords: CFD; Reinforcement Learning; Model-Based

Fuzzy quantized local reduced models for chaotic flows.

[1]Antonio Colanera* and [2,3]Luca Magri;

[1]*SISSA, International School for Advanced Studies, Trieste, IT*

[2]*Imperial College London, London, UK*

[3]*Politecnico di Torino, Torino, IT*

Abstract

Chaotic and turbulent flows often evolve on low-dimensional attractors with intricate, curved geometry. To deal with this structure we recently introduced quantized local reduced-order models (ql-ROMs), which partition the solution manifold into K clusters and build local POD–Galerkin ROMs around each centroid. In their sharp form, ql-ROMs rely on discontinuous switching between clusters; these sharp transitions can induce spurious oscillations and degrade accuracy near cluster boundaries. Here we propose fuzzy ql-ROMs, a smooth and still interpretable alternative based on fuzzy c-means. Fuzzy clustering yields a membership vector that measures the affinity of a state to each local region. In the online stage, all local reduced states are advanced in parallel, and their full-state reconstructions are fused through the previous-step memberships, producing a convex mixture that evolves continuously across regimes. This parallel evolution plus the membership-based fusion removes boundary discontinuities, improves robustness, and preserves locality. We also provide a practical criterion to select the number of clusters, based on a pseudo-BIC score tailored to fuzzy c-means. We assess fuzzy ql-ROMs on three benchmarks of increasing complexity: the 1D Kuramoto–Sivashinsky (KS) equation, the 2D KS equation in an intermittent regime characterized by bursts of kinetic energy, and the fluidic pinball flow featuring multi-frequency shedding and nonlinear modal interactions. Across all cases, fuzzy ql-ROMs outperform global POD–Galerkin ROMs and sharp ql-ROMs in short-term prediction and in reproducing long-term statistics. The results indicate that soft clustering is a simple yet effective mechanism to stabilize local ROMs and to capture intermittent regime transitions in spatiotemporally chaotic flows.

Keywords:

Reduced Order Models, Clustering, Chaotic flows

Data-driven Manifold Charting using Persistent Homology for Chaotic Dynamics

[1]Alec J. Linot*; [1]Eli Ruminer

[1]*University of Massachusetts, Amherst*

Abstract

A major difficulty in experiments and simulations of turbulent flows is the high-dimensional nature of the state. Despite this apparent high-dimensionality, turbulent flows are dissipative, which causes long-term dynamics to attract to a finite invariant manifold. Many data-driven reduced-order modeling approaches have succeeded in parameterizing these manifolds using undercomplete autoencoders. However, finding a minimal dimensional parameterization for any given manifold, in general, requires cutting the manifold into coordinate domains and training an autoencoder (i.e., coordinate map) for each of these domains. This pair is called a chart.

In this work, we show that finding these coordinate domains with standard methods (e.g., K-means) becomes increasingly difficult as the dimension of the manifold increases. To solve this, we instead guide the coordinate domain selection using persistent homology. The central idea is that when coordinate domains have holes they cannot be represented in a minimal number of dimensions. For example, the circle is a 1D manifold with a hole. Globally the circle requires two-dimensions to represent, and locally the circle requires one-dimension to represent. Thus, our method optimizes the location of each coordinate domain such that the persistence diagrams of these domains do not show holes – including N-dimensional holes. Then we train autoencoders for each of these coordinate domains. We show that this approach allows us to find minimal dimensional representations for dynamical systems that lie on the N-torus. As we increase the value of N, our approach becomes more important. Currently, we are applying this approach to the chaotic dynamics of the Kuramoto-Sivashinsky equation as a step towards turbulent flows.

Keywords: Autoencoders, topological data analysis, persistent homology, chaotic dynamics, dimension estimate

Data-Driven Reduced Kinetic Modeling for Rarefied Kolmogorov Flow

[1]Zhibo Wang*; [1]Federico Toschi

[1]Department of Applied Physics and Science Education, Eindhoven University of Technology, 5600 MB Eindhoven, Netherlands

Abstract

Rarefied flow remains challenging for low-order lattice and kinetic models because non-equilibrium effects become increasingly important as the Knudsen number grows. In this work, we investigate a data-driven reduced-order modeling framework for rarefied Kolmogorov flow, with the aim of improving the accuracy and efficiency of low-order representations in non-equilibrium regimes.

Our study is based on high-fidelity kinetic simulations and focuses on how reduced models can be enhanced using data-driven information while remaining compatible with established numerical solvers. Preliminary results indicate that local state information alone may not always be sufficient to capture the effective dynamics of rarefied flow on reduced representations. This suggests that enriched modeling strategies, potentially incorporating additional spatial or kinetic context, may be beneficial.

The present work outlines the motivation, formulation, and initial progress of this framework, and aims to contribute to the development of reliable Machine-learning-enhanced computational tools for rarefied fluid dynamics.

Keywords: Rarefied flow, Reduced-order modeling, Kinetic modeling, Lattice Boltzmann method, Machine learning

Manifold learning of pitching airfoil dynamics from pressure and PIV measurements

[1]Lucas Schneeberger; [1]Guy Y. Cornejo Maceda; [1]Zhiyuan Wang; [1]Iacopo Tirelli; [1]Stefano Discetti; [1]Andrea Ianiro

[1]*Department of Aerospace Engineering, Universidad Carlos III de Madrid, Leganés, Spain*

Abstract

Despite occurring in various engineering disciplines (e.g., helicopters, drones, and wind turbines), dynamic stall still lacks a unified model that encompasses the numerous problem variables. In this work, we propose to characterize the aerodynamics of a NACA 0018 spanning a wide range of flow conditions from snapshot data. We take advantage of isometric feature mapping (ISOMAP, [1]), a non-linear data-driven manifold learning method that unveils the intrinsic structure of the dynamics. We experimentally investigate the pressure over the airfoil using surface pressure data synchronized with time-resolved particle image velocimetry (PIV). The airfoil undergoes a range of pitching motions with varying reduced frequency, mean angle of attack, and pitching amplitude, while being exposed to different inflow conditions (Reynolds number, turbulence intensity) generated by a fan array. The wing is instrumented with 59 surface pressure ports distributed along the chord, providing time-resolved pressure measurements, from which the unsteady lift coefficient and pitching moment can be retrieved. Simultaneously, PIV identifies coherent flow structures interacting with the airfoil. To reveal the intrinsic low-dimensional organization of the pitching airfoil dynamics, we apply dimensionality reduction techniques to both pressure and velocity datasets. The low-dimensional embedding obtained via ISOMAP is analyzed in terms of trajectory evolution in the latent space, and the latent coordinates are interpreted in relation to the underlying flow dynamics. Such a low-dimensional representation of the dynamics allows for a better estimation of the flow state, and paves the way for more accurate prediction of the pressure and therefore aerodynamic loads.

This work is a companion study to "Actuation manifold for dynamic flow conditions—A flapping wing story" by Cornejo Maceda *et al.*, presented at the same conference.

Keywords: Pitching airfoil; Manifold learning; Particle image velocimetry; Dynamic stall; Data-driven modeling.

Acknowledgements:

This activity is part of the project EXCALIBUR (Grant No PID2022-138314NB-I00), funded by MCIU/AEI/ 10.13039/501100011033 and by "ERDF A way of making Europe".

References:

[1] Farzamnik, E., Ianiro, A., Discetti, S., Deng, N., Oberleithner, K., Noack, B. R., & Guerrero, V. (2023). From snapshots to manifolds – a tale of shear flows. J. Fluid Mech., *955*, A34.

Physics-guided Temporal Classification of Unsteady Cavitation Stages Using Supervised Learning

[1]Letian Jiang*; [1] Steffen J. Schmidt; [1,2] Nikolaus A. Adams;

[1]*Chair of Aerodynamics and Fluid Mechanics, TUM School of Engineering and Design*

Technical University of Munich

[2] *Munich Institute of Integrated Materials, Energy and Process Engineering*

Technical University of Munich

Abstract

Accurate identification of transient cavitation stages is essential for understanding and modeling unsteady cavitation dynamics. However, stage classification remains challenging due to strong temporal coupling, nonlinearity, and ambiguous phase boundaries [1]. In this work, we propose a physics-guided machine learning framework for automatic stage identification in quasi-periodic unsteady cavitation over a NACA0015 hydrofoil. We first extract physically interpretable features from simulation data by computer vision method, including sheet cavity length, vapor volume fraction, temporal derivatives, and other indicators. A supervised Random Forest (RF) classifier [2] is trained using manually labeled stage data. To enforce physically consistent temporal evolution, the classifier is integrated with a cyclic Hidden Markov Model (HMM) which ensure sequential transitions between stages. The proposed approach achieves a macro F1-score of approximately 0.9 and demonstrates robust generalization to previously unseen cavitation cycles. Errors are primarily confined to adjacent stage boundaries, indicating physically meaningful predictions rather than random misclassification. To evaluate the necessity of supervision, we compare our framework against unsupervised baselines, including POD-based k-means clustering. While low-dimensional representations capture dominant energetic structures, they fail to reliably separate physically defined cavitation stages and maintain temporal consistency. Our results suggest that energy-optimal representations alone are insufficient for event-level phase discrimination. The study highlights the importance of combining physics-informed feature engineering with temporal consistency constraints for reliable stage identification in cavitation flows, providing a foundation for future data-driven cavitation analysis and control strategies.

Keywords: Unsteady cavitation; Physics-informed machine learning; Cavitation stage identification; Random Forest

Reference:

[1] Budich, Bernd, S. J. Schmidt, and Nikolaus A. Adams. "Numerical simulation and analysis of condensation shocks in cavitating flow." *Journal of Fluid Mechanics* 838 (2018): 759-813.

[2] Murphy, Kevin P. *Probabilistic machine learning: an introduction*. MIT press, 2022.

Learning from the past to predict the future: a general, data-driven and physically interpretable algorithm, applied to turbulent flow forecasting

[1]G. Papadakis* and [1]S. Lu

[1]*Dept of Aeronautics, Imperial College London, SW7 2AZ*

**Corresponding author*

Abstract

A data-driven algorithm is proposed that employs sparse data from velocity (or scalar) sensors to reconstruct the current state and also forecast the future evolution of turbulent flows. The algorithm combines time-delayed embedding together with Koopman theory and linear optimal estimation theory. It consists of 3 steps; dimensionality reduction (currently using POD but other approaches can be also employed), construction of a linear dynamical system for the latent space variables (currently POD coefficients), and system closure using sparse sensor measurements. The algorithm essentially establishes a mapping from current sparse data to the future state of the dominant structures of the flow over a specified time window. The method is scalable, physically interpretable, and provides sequential reconstruction as well as forecasting on a sliding time window of pre-specified length. It is applied to the 3D turbulent recirculating flow over a surface-mounted cube (with more than 10^8 degrees of freedom) and is able to forecast accurately the future evolution of the most dominant structures over a time window at least two orders of magnitude larger that the (estimated) Lyapunov time scale of the flow. Most importantly, increasing the forecasting window only slightly reduces the accuracy of the estimated future states.

Keywords: Data-driven methods, Low order models, Flow estimation and forecasting, Turbulent flows

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

ShipNet: A Geometric Deep Learning Surrogate for Real-Time Ship Hydrodynamic Prediction

[1]K. Odendaal*, and [2]G. Drakoulas*;

[1]Maritime Research Institute Netherlands (MARIN), Wageningen, Netherlands;
[2]Damen Research, Gorinchem, Netherlands

Abstract. Accurate prediction of hydrodynamic performance is central to ship design, yet high-fidelity computational fluid dynamics remains prohibitively expensive for large-scale parametric exploration. This motivates the development of data-driven surrogate models that provide rapid approximations to hydrodynamic predictions at substantially reduced cost. We present ShipNet, a geometric deep-learning surrogate that predicts both hull-surface pressure distributions and far-field free-surface wave patterns directly from hull geometry and speed. The network employs a regression-based dynamic graph convolutional backbone on hull point clouds, with a multi-head decoder for simultaneous near-body pressure and free-surface elevation outputs. Training data consist of 420 inviscid free-surface simulations generated using a potential-flow panel method for two parent yacht hulls, each parameterised into 70 variants and evaluated at three speeds. ShipNet predicts per-point pressure coefficient and two-dimensional wave elevation map using a composite loss that combines point-wise regression and image-structure terms. On a geometry-held-out test set, ShipNet achieves for hull pressure and for wave fields. Inference requires 0.15s per case, yielding approximately 1500 faster prediction than the RAPID solver. Limitations include the restricted geometry and speed ranges and the inviscid training data, while future work will extend the model to high-fidelity viscous simulations with physics-informed regularisation.

Keywords: geometric deep learning, ship hydrodynamics, surrogate modelling, graph neural networks

1. Introduction

The maritime industry is undergoing a shift toward more sustainable and efficient operation, driven by tightening environmental regulation and economic pressure. A central lever is hull-form design, since the hull strongly influences hydrodynamic performance and therefore fuel consumption [1]. In modern practice, computational fluid dynamics (CFD), particularly Reynolds-Averaged Navier–Stokes (RANS) simulations, is widely used as a high-fidelity reference method to evaluate candidate designs by resolving hull pressure and the associated free-surface wave pattern [2–4]. These quantities are directly linked to resistance components and are therefore valuable targets for early-stage screening. Despite its accuracy, CFD is difficult to use as an interactive design tool. High-fidelity simulations can require substantial computational resources and long turnaround times, especially when free-surface effects, trim, and sinkage are included [3,4]. In addition, reliable CFD results require expertise in geometry preparation, meshing, solver settings, and model selection. These practical constraints reduce the number of designs that can be evaluated and slow down the design spiral.

Data-driven surrogate models aim to improve this speed–accuracy trade-off by learning direct mappings from design inputs to simulation outputs [5]. Classical reduced-order modelling has been successful for parametric variation in boundary conditions and material properties, but typically assumes fixed spatial discretisation which can struggle to accommodate geometric changes [6,7]. Recent scientific machine learning methods have expanded surrogate modelling to field prediction, including encoder–decoder approaches, physics-informed training, and operator learning [8–11,17]. However, many approaches still require structured grids or implicit geometry encodings that are not ideal when the geometry itself is the dominant source of variability.

Geometric deep learning provides an attractive alternative because it operates directly on unstructured representations such as point clouds and meshes, allowing the model to learn features that align with local geometric structure [12]. In external aerodynamics, regression-based Dynamic Graph Convolutional Neural Networks (regDGCNN), as demonstrated in DrivAerNet and DrivAerNet++, have shown that dynamic graph construction and EdgeConv features can learn meaningful geometry-conditioned representations from point clouds [13–15]. This motivates adapting similar architectures to ship hydrodynamics, where geometry also governs pressure trends and wave generation.

This work adapts regression-based Dynamic Graph Convolutional Neural Networks from single-phase automotive aerodynamics to free-surface ship hydrodynamics, where the coupled air–water interface together with trim and sinkage makes the learning problem more challenging. Our hypothesis is that a geometry-aware backbone can learn a shared latent representation from hull point clouds and operating conditions that is rich enough to support both near-body pressure prediction and far-field wave prediction. Accordingly, ShipNet combines a graph-based backbone with a multi-head decoder that predicts hull-surface pressure and free-surface elevation jointly. The contributions are: (i) adaptation of regDGCNN to ship hydrodynamic field prediction from hull geometry and operating condition, (ii) a multi-head formulation for joint prediction of hull-surface pressure and free-surface

*Corresponding Author, K. Odendaal (k.odendaal@marin.nl); G. Drakoulas (george.drakoulas@damen.com)

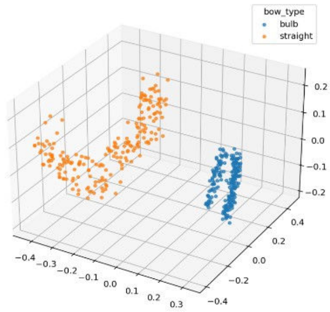
elevation, and (iii) geometry-held-out validation showing for hull pressure coefficient, for free-surface elevation, and approximately 1500 faster inference than potential-flow solver RAPID.

2. Methodology

2.1. Dataset and preprocessing

Two parent generic yacht hulls, one with a straight bow and one with a bulbous bow, were parametrically varied using a Latin hypercube sampling method, yielding 70 unique shape variants per family and 140 variants in total. All geometries were simulated at 12, 15, and 18 knots, corresponding to a Froude-number range of 0.197–0.322, using the RAPID potential-flow solver [16]. This produced 420 simulation cases in total. The parent geometries and the corresponding variants within a reduced dimensional-cluster space can be view in Figure 1. For each case, the hull-surface pressure coefficient () and a free-surface elevation () map on a 256 × 256 grid were exported. Data were split into train, validation, and test sets with a 70/15/15 ratio using geometry-grouped splits to avoid leakage across speeds.

(a)



(b)

Figure 1. (a) Renderings of the two parent yachts hulls: (top) Straight and (bottom) bulbous bow. (b) A three-dimensional Principle Component Analysis plot of the final 140 unique under geometries per operational speed, coloured by parent class.

RAPID was selected because it offers a practical trade-off between computational speed and hydrodynamic fidelity for early-stage design exploration. Although inviscid, it is based on a nonlinear ship-wave resistance formulation and is therefore suitable for capturing pressure and wave-making trends relevant to ranking candidate designs [16]. In contrast to the RANS-CFD methods discussed in the introduction, RAPID provides lower-cost potential-flow training data for this proof-of-concept surrogate study. A consistent RAPID setup was used for all simulations. The hull discretisation used approximately 6,000 panels for bulbous-bow variants and 2,000 panels for straight-bow variants, with a free-surface grid defined by 34 half-width strips and 79 half-width panels. The computational domain extended from 0.50 upstream to 0.807 downstream, with a half-width of 0.748, and dynamic trim treatment was enabled.

All hulls were placed in a consistent reference frame with the origin at midship, centerline, and the calm-water plane. No additional geometric scaling by perpendicular length () was applied because all variants belong to the same yacht class and share fixed principal dimensions. Each hull surface was represented by quasi-uniformly sampled points over the wetted surface using stratified random sampling, which reduces bias from non-uniform tessellation while preserving coverage of localised features such as bow curvature and shoulder transitions. The operating condition was appended to every point, such that each input point is represented by (x, y, z, Froude number). The free-surface elevation field was interpolated onto a fixed Eulerian grid, and both and were standardised using training-set statistics:

(1)

The corresponding values are .

2.2. Multi-head regDGCNN model

ShipNet uses a shared regression-based Dynamic Graph Convolutional Neural Network backbone to process hull point clouds directly [13,14]. The input is the sampled point set (x, y, z, Froude), where the Froude number is appended to every point so that a single model can represent multiple operating conditions. Three successive EdgeConv layers with dynamic k-nearest-neighbour graphs produce point-wise geometric features, which are symmetrically pooled into a 1024-dimensional latent embedding of the global hull form. The neighbourhood size was fixed at , following prior RegDGCNN practice [15] and the sensitivity study described in Section 2.4. A schematic visualisation of the modelling architecture can be seen in Figure 2.

*Corresponding Author, K. Odendaal (k.odendaal@marin.nl); G. Drakoulas (george.drakoulas@damen.com)

Two task-specific heads decode this shared representation. The pressure head combines the global embedding with final-layer point-wise features and applies one-dimensional convolutions to predict the pressure coefficient, Cp, at the sampled hull points. As a result, each pressure estimate depends on both local geometry and global hull context. The wave head reshapes the latent embedding into a low-resolution feature map and progressively upsamples it with transposed-convolution blocks to predict the free-surface elevation field in the ship-fixed frame, following a U-Net-like decoder structure [17]. This shared-backbone, multi-head formulation allows the model to couple geometry and speed information across both outputs while preserving task-specific decoding.

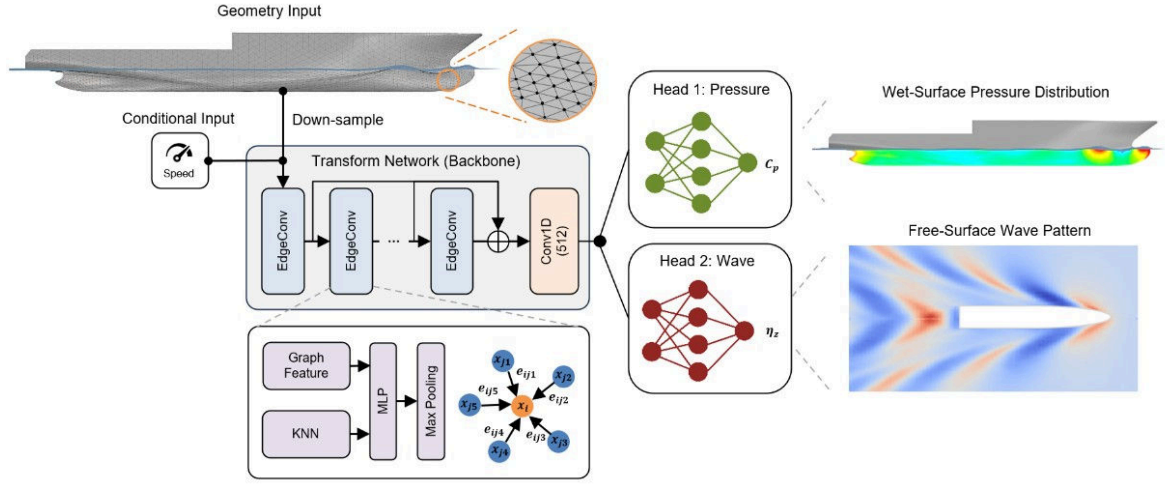


Figure 2. The proposed multi-head architecture adapted from DrivAerNet++ [15]. An input mesh, represented as a point cloud, is processed by a shared regression-based Dynamic Graph Convolutional Neural Network backbone to produce a latent feature representation. This representation is then decoded by two specialised heads to predict the hull-surface pressure and the two-dimensional free-surface wave map.

2.3. Training and evaluation

The model is trained end-to-end with a composite objective,

(2)

where is the mean squared error on hull-point pressure coefficients. For the wave field, we use,

(3)

which combines pixel-wise fidelity (L1), structural similarity (SSIM), and gradient consistency; Sobel filters are used for the gradient term. The full model was trained for 100 epochs using stochastic gradient descent (SGD) on an NVIDIA T4 GPU. The primary training configuration used a batch size of 4, learning rate , dropout of 0.4, and a ReduceLROnPlateau scheduler, following the DrivAerNet++ baselines [15]. Performance on the held-out test set is reported using per-sample MAE, MSE, and coefficient of determination (), summarised by mean and standard deviation across cases. MAE and MSE are computed on standardised outputs.

2.4. Hyperparameter selection

Model hyperparameters were selected through a targeted sensitivity study covering point sampling density (), training-data fraction, wave-grid resolution, and task-loss weighting. The detailed optimisation results, including mean test performance and confidence intervals across random seeds, are provided in appendix A. In brief, the study supported the use of sampled hull points, showed that performance saturated at roughly 50% of the available training data, indicated limited sensitivity to wave-grid resolution within the tested range, and confirmed that balanced task weighting is important for stable joint learning.

3. Results

3.1. Quantitative accuracy

After hyperparameter selection, the final model was evaluated on 63 held-out test cases. On standardised outputs, ShipNet achieved strong predictive accuracy for hull pressure, with an Mean Absolute Error (MAE) of and . For the free-surface wave field, the model achieved an MAE of and .

*Corresponding Author, K. Odendaal (k.odendaal@marin.nl); G. Drakoulas (george.drakoulas@damen.com)

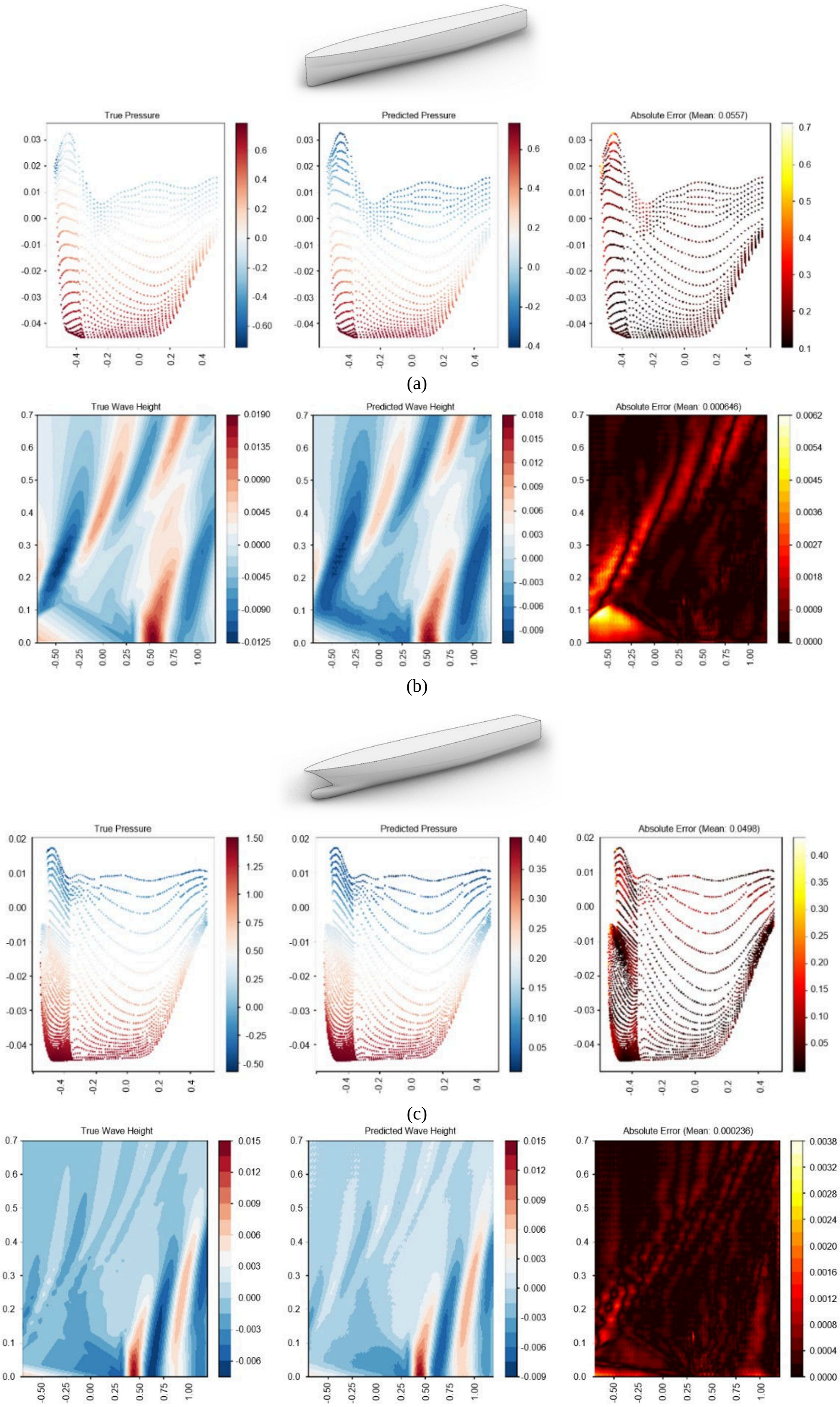


Figure 3. Qualitative comparison of ShipNet predictions for representative no-bulb and bulbous-bow test cases. The small schematic above each case illustrates the visualization procedure, where three-dimensional hull-surface pressure points are projected from the wetted hull surface onto a normalized – plane. For each hull, the first row shows reference, predicted, and absolute-error fields for hull-surface pressure coefficient , while the second row shows the corresponding free-surface elevation field on the two-dimensional ship-fixed wave grid. Panels (a, b) correspond to the no-bulb case, and panels (c, d) correspond to the bulbous-bow case.

*Corresponding Author, K. Odendaal (k.odendaal@marin.nl); G. Drakoulas (george.drakoulas@damen.com)

Figure 3 provides a visual comparison of the absolute error maps for two representative examples. For hull pressure, the model captures the bow stagnation region, suction at the shoulder, and downstream pressure recovery for both straight and bulbous bow variants. The largest errors are concentrated near the bulb leading edge and other regions of strong curvature and pressure gradient, while most of the wetted surface is predicted accurately. For the wave field, the model reproduces the main Kelvin-wake structure and the changes induced by bow-form variation; the largest discrepancies occur in the first trough aft of the bow, a region that is especially sensitive to stagnation pressure and steep local gradients. Although visualised in two-dimensional form for compactness, the predicted hull pressure remains a scalar field on the three-dimensional wetted hull surface, while the predicted free-surface elevation can be reconstructed as a ship-fixed height field for standard post-processing.

3.2. Ablations and loss design

The wave-loss formulation was the most important design choice for stable prediction. Replacing the hybrid loss with a standard MSE loss caused training failure, yielding a negative mean on the wave field Table 3. This reflects the sparsity of the wave-elevation target, where a purely quadratic loss favors over-smoothed solutions. In contrast, the hybrid L1 + SSIM + Sobel loss preserves both global structure and sharp wave features.

Table 3. Loss ablation for free-surface wave prediction

Wave Loss	MAE
Single: MSE	
Hybrid: L1 + SSIM + Sobel	

This behaviour is physically and numerically intuitive. The wave map is dominated by near-zero regions punctuated by relatively sharp crests and troughs, so a quadratic loss heavily penalises small spatial misalignments of high-amplitude features and drives the optimiser toward a blurred average solution. The added SSIM and gradient terms counteract this failure mode by rewarding structural agreement and crest localisation rather than only minimising pixel-wise amplitude error. In practice, the hybrid objective was necessary to recover realistic Kelvin-wake patterns rather than a diffuse mean-field prediction.

3.3. Computational cost

The engineering value of ShipNet lies in its efficiency. A RAPID simulation requires approximately 228s on a single CPU core, whereas ShipNet performs inference in 0.15s per case on an NVIDIA T4 GPU, corresponding to an approximate speedup. This makes near-interactive early-stage design screening feasible and enables dense parametric sweeps that would be impractical with solver-only workflows (see Table 4). It should be noted, this comparison reflects the practical solver-versus-surrogate workflow used in this study rather than a hardware-normalised benchmark, because RAPID and ShipNet were evaluated on different processor types.

Table 4. Computational performance summary

Comparison	Speed-up ratio	
GPU vs CPU (train iter)		0.15s vs. 0.40s per iter
RAPID vs CPU (inference)		228s vs. 0.40s
RAPID vs GPU (inference)		228s vs. 0.15s

4. Conclusion and Limitations

We presented ShipNet, a multi-head geometric deep learning surrogate that predicts hull-surface pressure and free-surface elevation directly from hull geometry and operating condition. On held-out geometries, the model achieved for pressure and for waves while reducing inference time to 0.15s per case, corresponding to an approximate speedup over the potential-flow solver RAPID. The present study is limited to inviscid potential-flow data and a narrow family of slender yacht hulls. In addition, the current dataset covers bare hulls only and the model does not explicitly enforce physical constraints such as conservation laws. Extending the approach to high-fidelity RANS data, broader vessel classes, appendages, and physics-informed regularisation is therefore the next step toward deployment in a wider ship-design setting.

5. Acknowledgements

This research was made possible through funding from the Cooperative Research Ships (CRS) fund. The authors wish to express their gratitude for this essential support.

*Corresponding Author, K. Odendaal (k.odendaal@marin.nl); G. Drakoulas (george.drakoulas@damen.com)

6. References

[1] M. Tadros, M. Ventura, and C. Guedes Soares. Review of the decision support methods used in optimizing ship hulls towards improving energy efficiency. Journal of Marine Science and Engineering, 11(4):835, 2023. doi: 10.3390/jmse11040835.

[2] L. Larsson and H. C. Raven. Ship Resistance and Flow. J. R. Paulling, editor. Principles of Naval Architecture Series. Society of Naval Architects and Marine Engineers, 2010. ISBN 9780939773763.

[3] N. J. van der Kolk, I. Akkerman, J. A. Keuning, and R. H. M. Huijsmans. Part 2: Simulation methodology and numerical uncertainty for RANS-CFD for the hydrodynamics of wind-assisted ships operating at leeway angles. Ocean Engineering, 201:107024, 2020. doi: 10.1016/j.oceaneng.2020.107024.

[4] L. Huang, B. Pena, and G. Thomas. Towards a full-scale CFD guideline for simulating a ship advancing in open water. Ship Technology Research, 70(3):222–238, 2023. doi: 10.1080/09377255.2023.2167537.

[5] G. Drakoulas, T. V. Gortsas, E. Polyzos, S. V. Tsinopoulos, L. Pyl, and D. Polyzos. An explainable machine learning-based probabilistic framework for the design of scaffolds in bone tissue engineering. Biomechanics and Modeling in Mechanobiology, 23(3):987–1012, 2024. doi: 10.1007/s10237-024-01817-7.

[6] P. Benner, S. Gugercin, and K. Willcox. A survey of projection-based model reduction methods for parametric dynamical systems. SIAM Review, 57(4):483–531, 2015. doi: 10.1137/130932715.

[7] G. I. Drakoulas, T. V. Gortsas, G. C. Bourantas, V. N. Burganos, and D. Polyzos. FastSVD-ML–ROM: A reduced-order modeling framework based on machine learning for real-time applications. Computer Methods in Applied Mechanics and Engineering, 414:116155, 2023. doi: 10.1016/j.cma.2023.116155.

[8] G. I. Drakoulas, T. V. Gortsas, and D. Polyzos. Physics-based reduced order modeling for uncertainty quantification of guided wave propagation using Bayesian optimization. Engineering Applications of Artificial Intelligence, 133:108531, 2024. doi: 10.1016/j.engappai.2024.108531.

[9] M. Raissi, P. Perdikaris, and G. E. Karniadakis. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. Journal of Computational Physics, 378:686–707, 2019. doi: 10.1016/j.jcp.2018.10.045.

[10] Z. Li, N. B. Kovachki, K. Azizzadenesheli, B. Liu, K. Bhattacharya, A. Stuart, and A. Anandkumar. Fourier neural operator for parametric partial differential equations. In International Conference on Learning Representations, 2021.

[11] L. Lu, P. Jin, G. Pang, Z. Zhang, and G. E. Karniadakis. Learning nonlinear operators via DeepONet based on the universal approximation theorem of operators. Nature Machine Intelligence, 3(3):218–229, 2021. doi: 10.1038/s42256-021-00302-5.

[12] M. M. Bronstein, J. Bruna, T. Cohen, and P. Veličković. Geometric deep learning: Grids, groups, graphs, geodesics, and gauges. arXiv:2104.13478, 2021. doi: 10.48550/arXiv.2104.13478.

[13] Y. Wang, Y. Sun, Z. Liu, S. E. Sarma, M. M. Bronstein, and J. M. Solomon. Dynamic graph CNN for learning on point clouds. ACM Transactions on Graphics, 38(5):146, 2019. doi: 10.1145/3326362.

[14] M. Elrefaie, A. Dai, and F. Ahmed. DrivAerNet: A parametric car dataset for data-driven aerodynamic design and prediction. Journal of Mechanical Design, 147(4):041712, 2025. doi: 10.1115/1.4068104.

[15] M. Elrefaie, F. Morar, A. Dai, and F. Ahmed. DrivAerNet++: A large-scale multimodal car dataset with computational fluid dynamics simulations and deep learning benchmarks. *Advances in Neural Information Processing Systems*, 37:499–536, 2024. doi: 10.48550/arXiv.2406.09624.

[16] H. C. Raven. A solution method for the nonlinear ship wave resistance problem. PhD thesis, Delft University of Technology, Delft, The Netherlands, 1996.

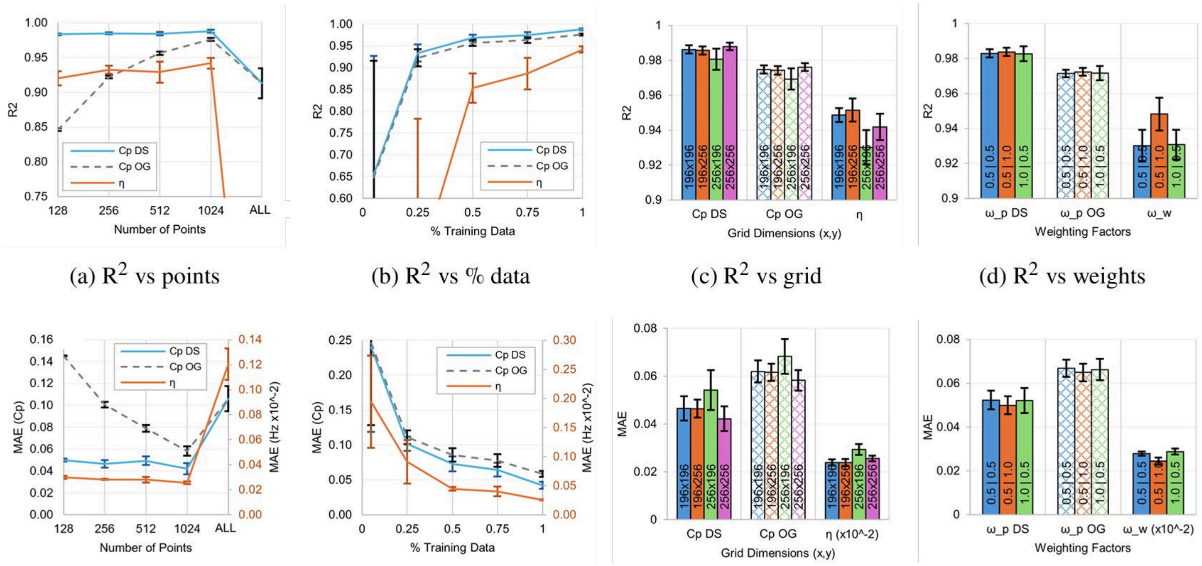
[17] O. Ronneberger, P. Fischer, and T. Brox. U-Net: Convolutional networks for biomedical image segmentation. In Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015, Lecture Notes in Computer Science, vol. 9351, pages 234–241. Springer, Cham, 2015. doi: 10.1007/978-3-319-24574-4_28.

*Corresponding Author, K. Odendaal (k.odendaal@marin.nl); G. Drakoulas (george.drakoulas@damen.com)

7. Appendix

7.1. Hyperparameter optimisation

A targeted sensitivity study was conducted to evaluate the effects of point sampling density, training data volume, wave-grid resolution, and loss weighting. Figure A.1 summarises these ablations with mean performance and 95% confidence intervals over the test set. Three response types are considered: pressure prediction accuracy on the downsampled point cloud used during training, pressure accuracy after re-interpolation to the full-resolution hull surface, and free-surface wave elevation accuracy. Since the regDGCNN backbone operates on a fixed number of sampled points, training is necessarily performed on a reduced representation of the hull. Adequate sampling density is therefore essential to retain sufficient geometric information for accurate reconstruction on the full domain.



(a) R^2 vs points (b) R^2 vs % data (c) R^2 vs grid (d) R^2 vs weights

(e) MAE vs points (f) MAE vs % data (g) MAE vs grid (h) MAE vs weights

Figure A.1. Learning-test curve summaries across ablation studies. (a-d) vary the number of points and training-data fraction; (e-h) vary grid resolution and loss-weight factors. Error bars denote 95% confidence intervals across 3 random seed variations. DS denotes Cp evaluated on down-sampled geometry points, while OG denotes Cp evaluated on the original full geometry, reconstructed via interpolation from the reduced point set, thereby capturing reconstruction error. The curve shows wave elevation prediction performance.

A consistent performance gap is observed between pressure accuracy on the sampled point set and on the reconstructed full-resolution geometry. This gap decreases monotonically with increasing point count and becomes negligible at approximately , indicating that this sampling density preserves the relevant geometric features. Both and MAE follow the same trend, with performance degrading as point count decreases and reconstruction error increases. With respect to training data volume, performance saturates at roughly 50% of the available dataset, corresponding to approximately 80 hull-speed samples, demonstrating strong data efficiency. Training on the full dataset using a single batch leads to degraded performance, which is attributed to optimisation limitations rather than model capacity.

Wave-grid resolution exhibits comparatively weak sensitivity within the tested range. However, changes to wave-related hyperparameters and loss weighting affect both wave and pressure accuracy, reflecting the shared latent representation in the multi-head architecture. In particular, loss weighting plays a critical role, as imbalanced task objectives can bias the learned representation toward one output at the expense of the other. These results emphasize the importance of careful hyperparameter selection in jointly trained geometric surrogate models.

*Corresponding Author, K. Odendaal (k.odendaal@marin.nl); G. Drakoulas (george.drakoulas@damen.com)

Simulation of multi-component flash boiling sprays with data-driven vapor-liquid equilibrium calculations

[1]Y. Li*; [2]Z. Lu; [2]H. Guo; [1]X. Ma; [3]Z. Han

[1] *School of Vehicle and Mobility, Tsinghua University, Beijing, China;*

[2]*School of Power and Energy, Northwestern Polytechnical University, Xi'an 710129, China*

[3]*Institute of aeroengines, Tsinghua University, Beijing, China*

Abstract

A data-driven model to calculate the vapor liquid equilibrium (VLE) was proposed to accelerate the CFD simulation of multi-component (*n*-hexane/ethanol) liquid sprays under flash-boiling conditions. This method significantly improves computational efficiency by using a deep feedforward neural network (DFNN) to replace the complex iterative solution of multi-component VLE. The DFNN takes instantaneous pressure, temperature, and system composition as input and predicts the corresponding phase equilibrium state. A parametric study was conducted to optimize the neural network's hyperparameters, including the activation function, number of hidden layers, and neurons per hidden layer. The rate of phase change is then calculated as a linear relaxation toward phase equilibrium, guiding subsequent computational steps in the CFD solver. A series of cases were performed to test the proposed methodology, involving the injection of a superheated *n*-hexane/ethanol blend into a gaseous nitrogen environment. The simulation results and computational cost were examined. It is found that the DFNN model, while accurately representing the non-ideal non-equilibrium phase change of a multi-component injection flow, speeds up the VLE solution by four orders of magnitude, leading to a reduction >60% in overall flow simulation time. This model shows promise for injection flow simulations, especially for systems with a large number of compositions, such as sustainable aviation fuels.

Keywords: Multi-component liquid; Neural network; Spray; In-nozzle flow;

Introduction

With recent advancements in high-performance computing, there has been a significant enhancement in high-fidelity numerical simulations for cutting-edge propulsion and power systems. This has sparked a growing interest in computational fluid dynamics (CFD) research on fuel injection flows inside and near nozzles. A comprehensive understanding of the fluid dynamics within and immediately outside injectors is paramount for accurately representing subsequent phenomena such as atomization, fuel-air mixing, and combustion processes.

The homogeneous mixture model (HMM) is an important method for modeling fuel injection flow, known for its effectiveness in dealing with the small-scale flow channel inside the injectors [1]. The homogeneous relaxation model (HRM) [2] is frequently coupled with HMM to describe phase transitions. This model takes non-equilibrium effects into account by conceptualizing phase change as a linear approach to equilibrium. There is abundant evidence in the literature that validates HRM's ability to accurately predict transitions from liquid to gas in internal flow scenarios.

The HRM has a defect in that it is only feasible for single-component liquids, while most fuels are multi-component. It is also unable to describe the mixing-driven vaporization outside the nozzle. To overcome these drawbacks, a non-equilibrium phase transition model, named the unified model, was introduced [3, 4]. This model incorporates VLE for mixtures of any composition and takes into account the timescales of different phase change mechanisms. The unified model, while a valuable extension to HRM, still has certain limitations. It relies on Raoult's law to calculate VLE, so it cannot account for non-ideal mixing, which is important for components with different molecular structures. This issue could be addressed by including activity coefficients in the VLE solver to consider intermolecular interactions, but that would enhance algorithmic complexity and memory requirements. Moreover, the computational efficiency of the unified model decreases significantly as the number of components in the system increases.

Deep learning neural network is a thriving subdomain within the vast landscape of artificial intelligence with proficiency in learning and approximating complex non-linear functions, as well as handling the interpolation of unseen data encountered beyond the training phase. The capabilities of DNN offer a potential solution to address the constraints of the unified model. Therefore, in this study, a framework has been developed to swiftly and accurately

*Corresponding Author, Dr Yanfei Li, liyf2018@tsinghua.edu.cn

predict VLE using DNNs. This enables a significant improvement over the traditional iteration-based approach. The framework has been seamlessly incorporated into the CFD solver, CONVERGE version 3.0 [5], using UDFs.

To compare the performance of the data-driven model with the iteration-based unified model, simulations have been carried out involving the injection of a superheated liquid ethanol-water mixture into gaseous nitrogen using both the new model and the unified model.

Numerical Method

The process of developing the deep-learning-based approach involves three main steps. First, a database of VLE was created to account for non-ideal mixing effects in a multi-species system across a wide range of thermodynamic conditions. Next, a deep feedforward neural network (DFNN) was fine-tuned with optimal hyperparameters to enable accurate and efficient VLE predictions. Finally, the DFNN was incorporated into the flow solver to replace the iterative VLE calculator, and the rate of phase transition was calculated using the predicted phase equilibrium. This phase change modelling approach is referred to as the DFNN model from here on. It is noted that in this work, deep neural network is essentially a complicated curve fit to data generated by solution of vapor-liquid equilibrium model. The details of these three steps are outlined below.

<u>Traditional VLE calculation and validation</u>

The VLE algorithm based on fugacity has been implemented using MATLAB. An iteration convergence threshold of 1.0e-5 was used to ensure that iterative convergence was achieved. To validate the solution, Figure 1 compares the bubble and dew points obtained from experiments and our in-house solver. The n-hexane/ethanol blends were used as they are typical mixtures that show significant deviations from ideal states. The predicted results by the VLE solver matched well with the experimental data for both the bubble and dew points, representing the non-ideal mixing behaviours and the azeotrope concentration accurately.

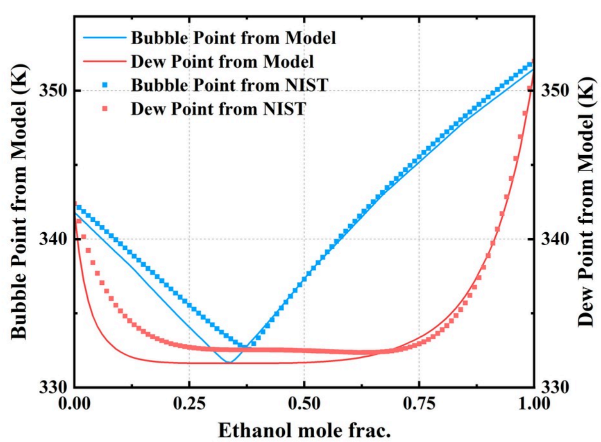


Figure 1 Bubble and dew points calculated by traditional VLE solver compared with experimental measurement.

<u>Neural network based VLE solver development and validation</u>

In order to assess the accuracy of the DFNN's predictions, Figure 2 displays regression plots that compare the output variables with the predicted values by DFNN on the y-axis and the expected values on the x-axis. Each data point in the validation dataset is represented by a scatter on the plot, and the R2-scores for each variable on the validation dataset are also included. Overall, the predictions made by DFNN closely match the expected results, with R^2-scores exceeding 0.999 for each output variable.

To assess the computational efficiency of the DFNN, Figure 6 provides a comparison of the time taken to generate output between the VLE solver and the DFNN model. The indicated times correspond to generating up to 500 random data points. The calculations were performed on a single processor, with the only difference being the method of

*Corresponding Author, Dr Yanfei Li, liyf2018@tsinghua.edu.cn

Accelerating Numerical Simulation in CFD by Reduced Order Surrogate Modelling and Scientific Machine Learning

Gianluigi Rozza*

SISSA – International School for Advanced Studies, mathLab, Trieste, Italy

Abstract

Computational Fluid Dynamics (CFD) is a cornerstone for the analysis of complex flow phenomena. However, the computational cost of high-fidelity simulations often restricts their use in many-query scenarios such as design optimization, uncertainty quantification, and parametric analysis.

Reduced Order Models (ROMs) alleviate this limitation by constructing low-dimensional surrogates that retain the dominant flow dynamics at a significantly reduced computational cost. This work presents complementary data-driven strategies combining reduced-order modeling and scientific machine learning to improve the accuracy, robustness, and generalization capabilities of ROMs for CFD applications.

The first contribution focuses on non-intrusive ROMs through a space-dependent aggregation framework where multiple surrogate models are locally combined using spatially varying convex weights learned from data, improving predictive accuracy in challenging benchmark problems such as transonic airfoil flows. The second contribution addresses intrusive ROMs through parametric closure modeling using deep operator networks to learn reduced correction operators that compensate unresolved dynamics and enhance generalization across the parameter space. Finally, a hybrid reduced-order modeling framework is introduced for high-Reynolds-number fluid–structure interaction problems in an Arbitrary Lagrangian–Eulerian setting, combining POD–Galerkin projection with machine-learning-based turbulence modeling and reduced interpolation techniques for mesh motion.

Overall, these approaches advance hybrid reduced-order modeling by integrating surrogate aggregation, operator learning, and data-driven turbulence closure within a flexible and scalable framework for fast and reliable CFD predictions.

Keywords: Reduced Order Modeling, Computational Fluid Dynamics, Scientific Machine Learning, Digital Twins, Operator Learning