Multi-Fidelity Aerodynamic Surrogate Modeling for Conceptual Design Using Graph Neural Networks

Yiren Shen

Stanford University

Advisor: Prof. Juan J. Alonso

Collaborator: Dr. Jacob Needels

Advanced Modeling & Simulation (AMS) Seminar Series

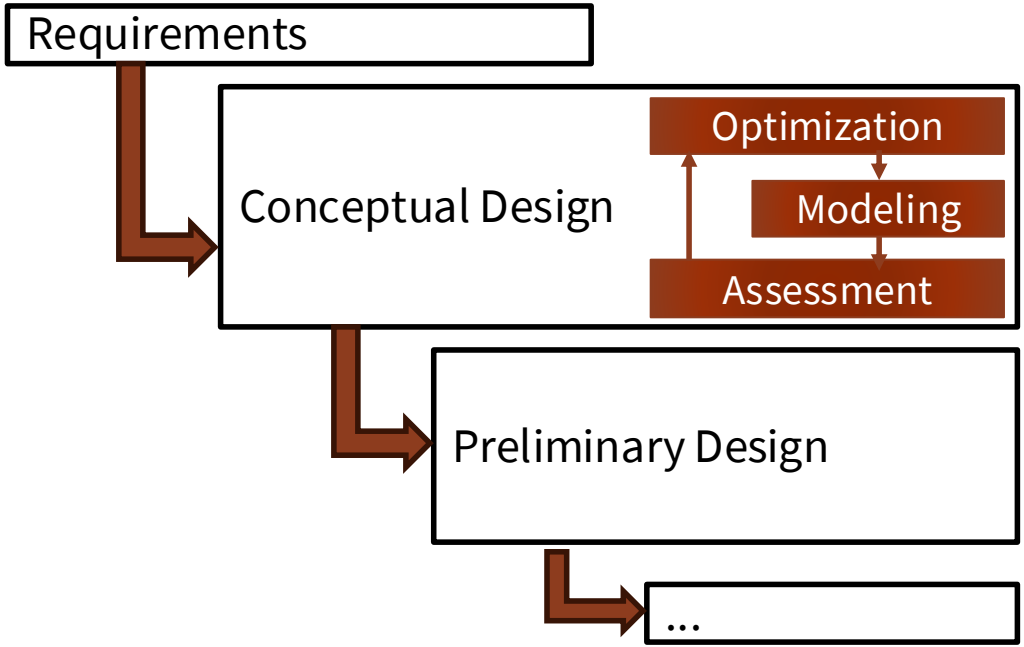
NASA Ames Research Center

July 9th, 2026





Aircraft Design Process (ADP)



- **“Selection”**
- configuration layout
- wing geometry

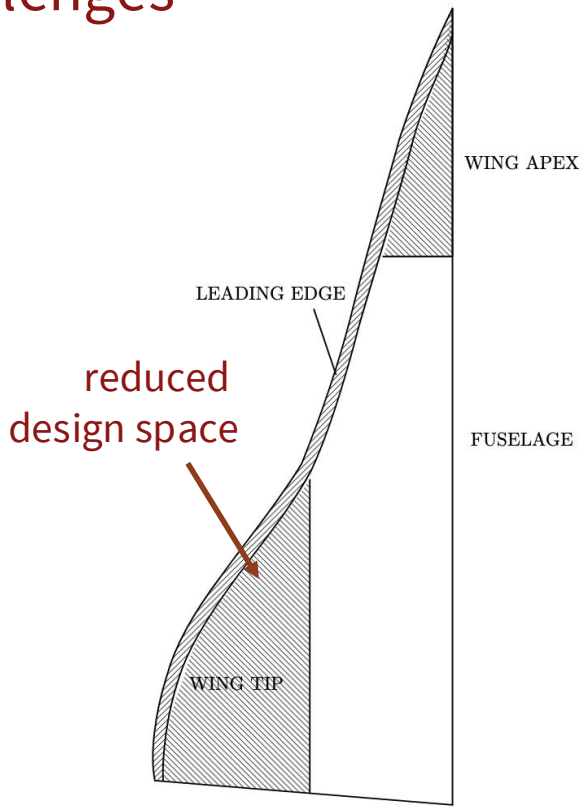
- **“Refinement”**

Flexibility and exploration → configuration commitment

aerospacedesignlab

Supersonic Transport (SST) Design Challenges

1. The “locked-in” effects:
   - Higher design confidence is needed in conceptual design

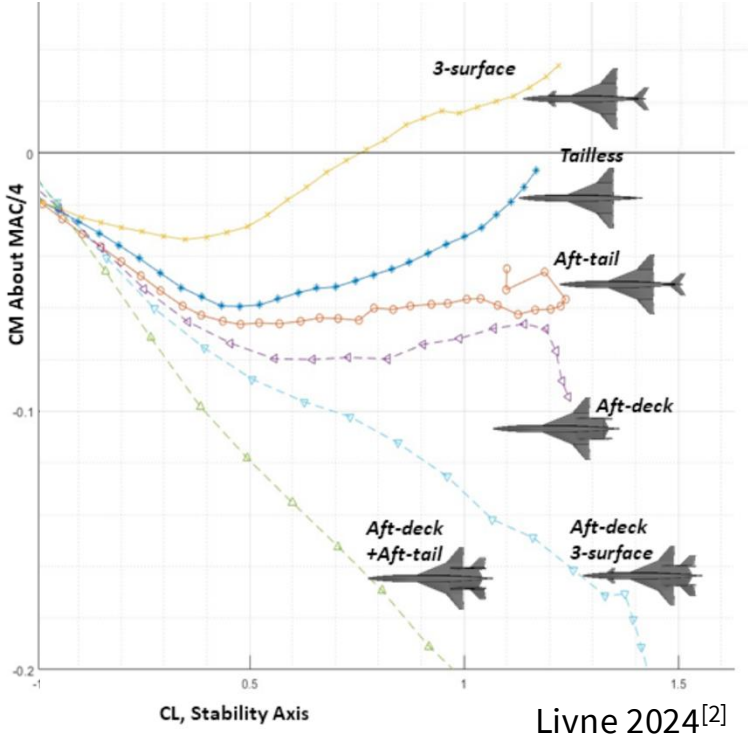


Wilde 1969 [1]

[1] Wilde and Cormery, “The aerodynamic derivation of the concorde wing,” 1969

Supersonic Transport (SST) Design Challenges

1. The “locked-in” effects:
   - Higher design confidence is needed in conceptual design
2. Conflicting objectives:
   - Supersonic vs. subsonic performance
   - Stability considerations

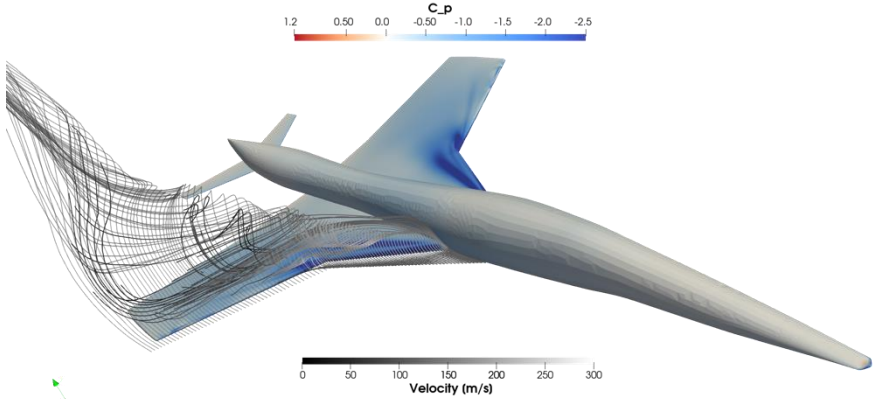


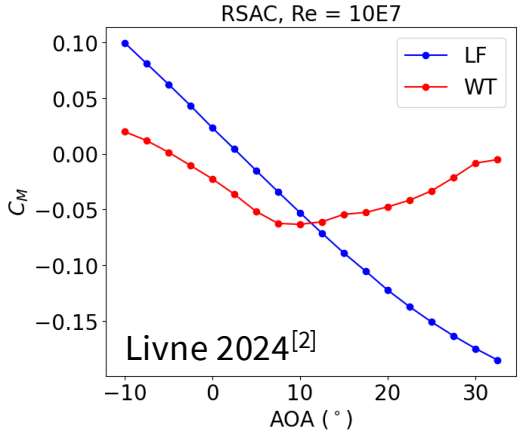
Livne 2024[2]

[2] Livne et al., “Compilation of Supersonic Configurations at Low Speeds (SCALOS) Final Report and Appendices A–F,” 2024

Supersonic Transport (SST) Design Challenges

1. The “locked-in” effects:
   - Higher design confidence is needed in conceptual design
2. Conflicting objectives:
   - Supersonic vs. subsonic performance
   - Stability considerations
3. Modeling limitations:
   - LF methods may produce biased predictions that mislead search direction
   - HF methods can be costly, fragile, and often infeasible





There is a fundamental fidelity-cost gap:

The **desired level of analysis fidelity** for conceptual design is often **unattainable** because of the **cost required** to generate it.

Two Paths to Close the Fidelity-cost Gap

Reduce the HF Cost

- Parallelization and GPU acceleration: $\mathcal{O}(\text{hrs}) \rightarrow \mathcal{O}(\text{min})$.
- Workflow automation: agentic DOE, case setup, and meshing.

Knowledge Transfer to LF Runtime

- Correct LF analysis using sparse HF observations.
- Built upon the current conceptual design workflow.

Multi-fidelity Methods and Surrogate Models

*"multi-fidelity methods use low-fidelity models to **reduce the runtime** ... but **recourse** to the high-fidelity model to preserve the accuracy of the outer-loop result..."* – Peherstorfer 2018 [3]

$$\hat{y}_{\mathrm{HF}} = \rho(z)\, y_{\mathrm{LF}}(z) + \delta(z) \quad [4]$$

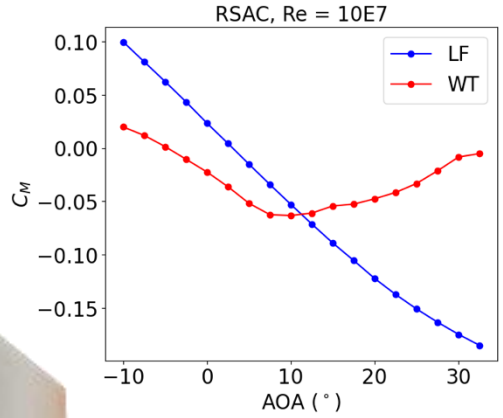
$$\hat{y}_{\mathrm{HF}} = \mathcal{F}\left(y_{\mathrm{LF}}, z \right) \quad [5]$$

Surrogate Models: bridging the fidelity-cost gap

[3] Peherstorfer et al., "Survey of multifidelity methods in uncertainty propagation, inference, and optimization,"2018
[4] Kennedy and O'Hagan, Bayesian Calibration of Computer Models, 2001
[5] Fernandez-Godino, Review of Multi-fidelity Models, 2023

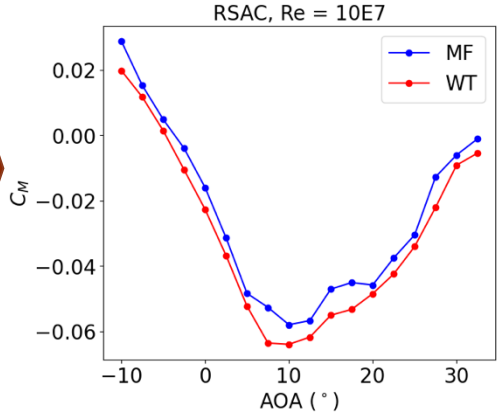
The Big Picture

LF Simulations



A
"Useful"
Multi-Fidelity
Surrogate
Model

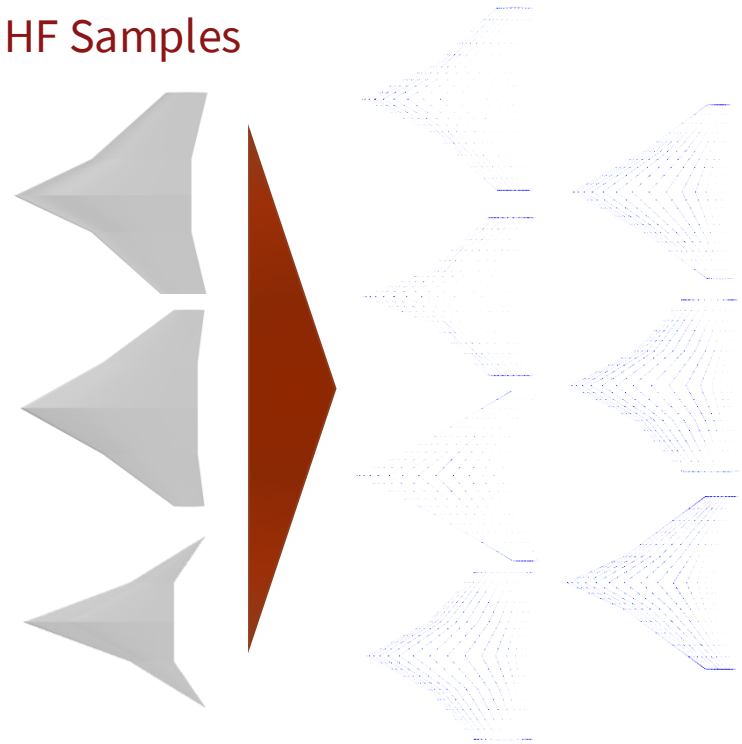
MF Results



What makes a surrogate model "useful" for conceptual design?

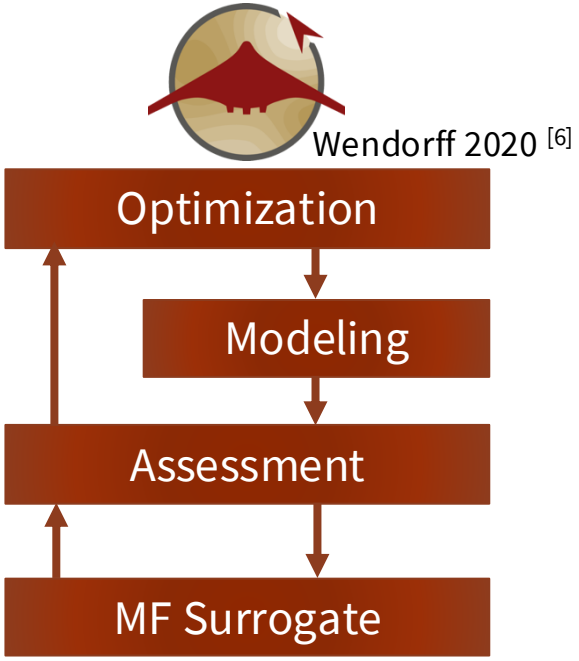
Useful Features of MF Models

1. Generalization across the design space
   - Learn from sparse HF samples



Useful Features of MF Models

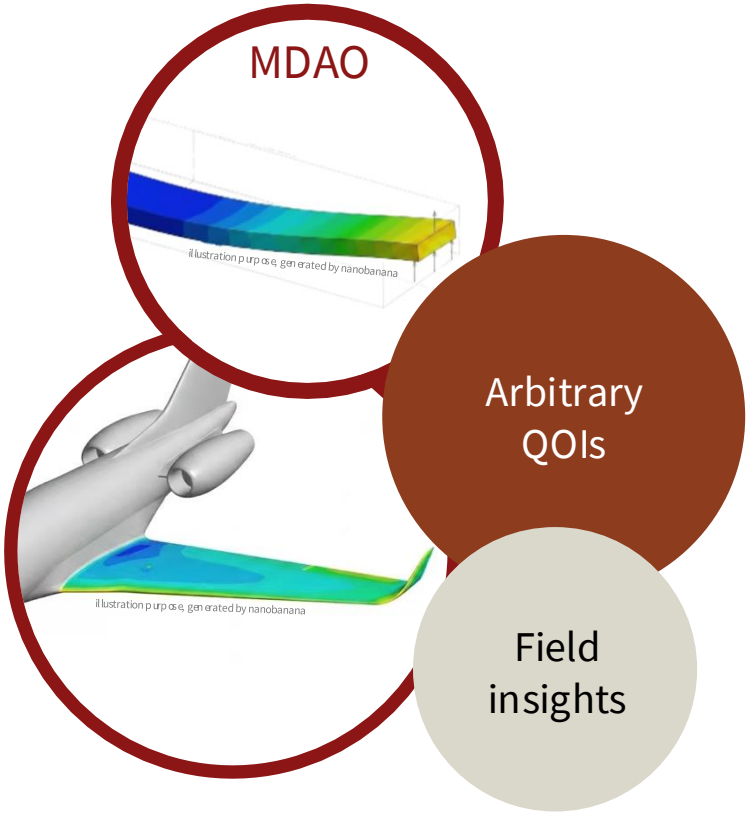
1. Generalization across the design space
   - Learn from sparse HF samples
2. Computational efficiency for online use
   - Design framework integration



[6] Wendorff et al., “SUAVE: An Aerospace Vehicle Environment for Designing Future Aircraft,” 2020

Useful Features of MF Models

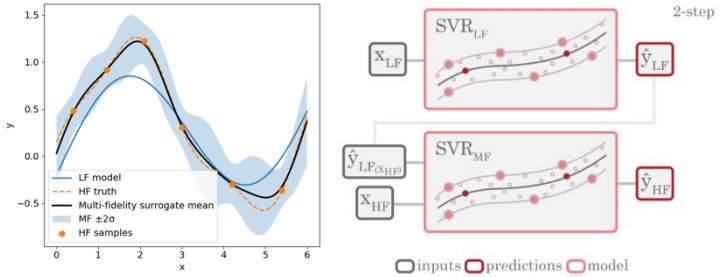
1. Generalization across the design space
   - Learn from sparse HF samples
2. Computational efficiency for online use
   - Design framework integration
3. Prediction of field quantities
   - High problem dimension



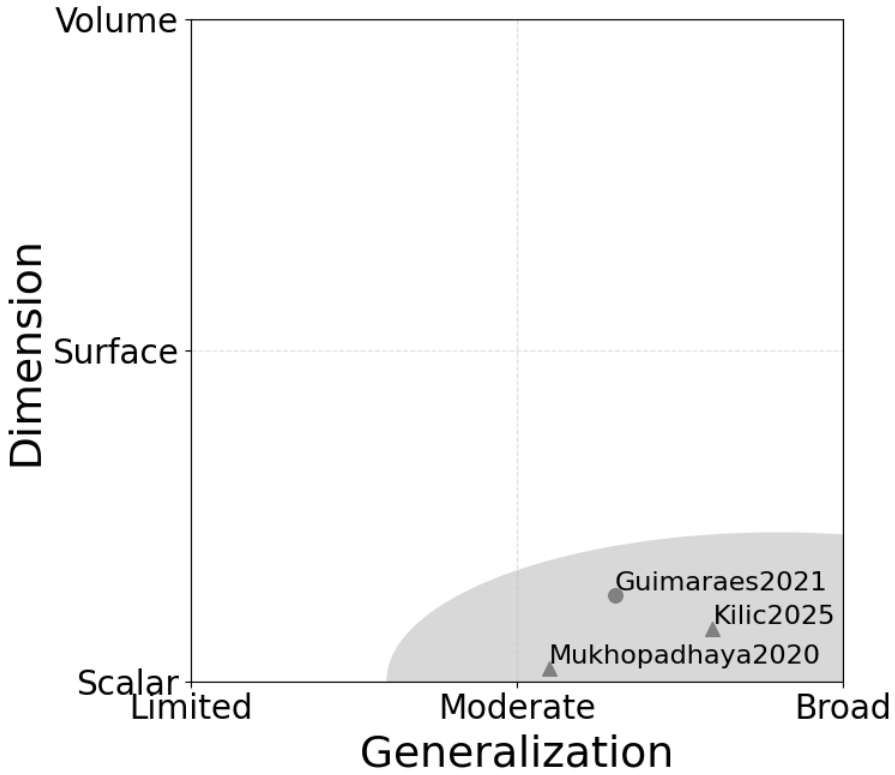
aerospacedesignlab

Literature Review

1. Scalar Surrogate Models
   - Gaussian-Process Regression
   - Support Vector Machine
   - Artificial Neural Network
   - …



Kilic 2025 [7]

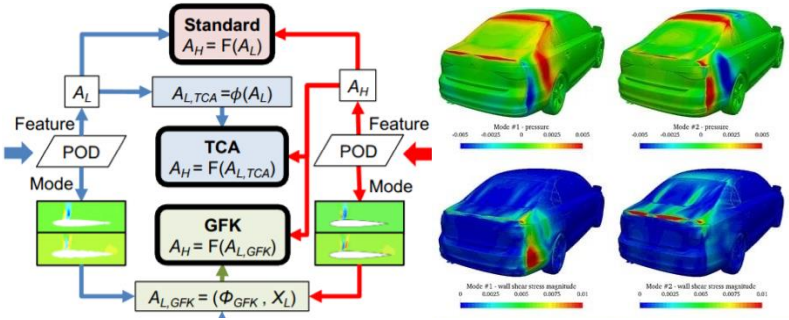


[7] Kilic and Nikbay, "A comparative study on multi- fidelity machine learning modeling for aerospace problems," 2025

aerospacedesignlab

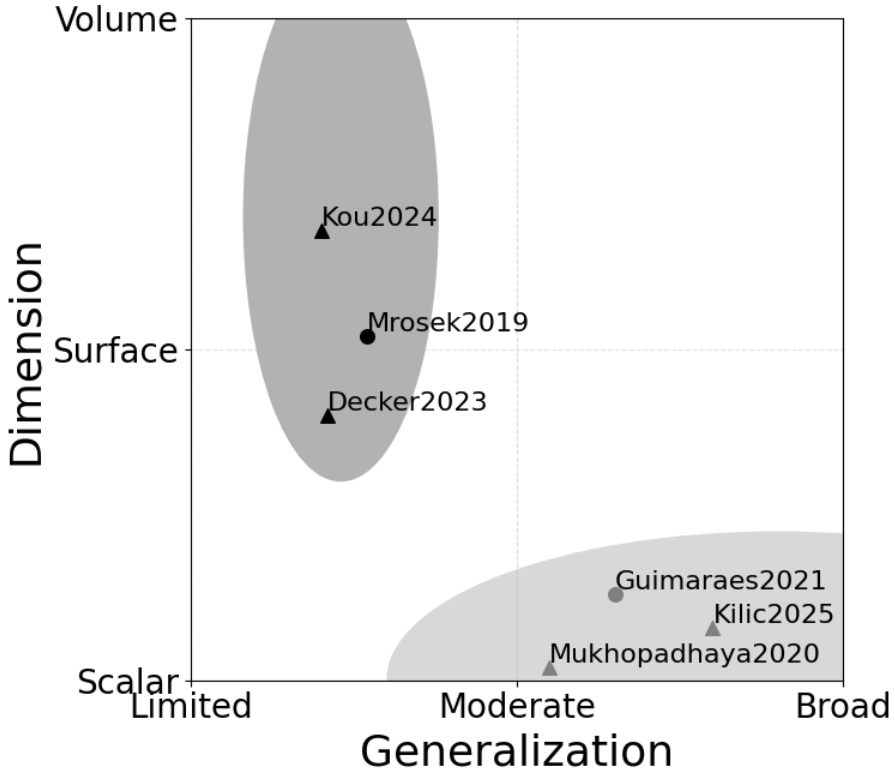
Literature Review

1. Scalar Surrogate Models
2. Reduced-Order Models
   - POD/POD-Kriging



Kou 2024 [8]

Mrosek 2019 [9]

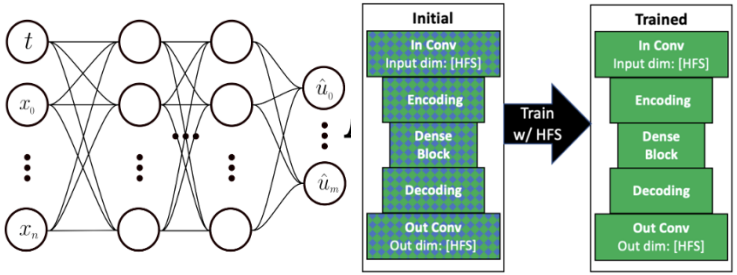


[8] Kou et al., "transfer learning for reduced-order modeling of transonic flows using multifidelity data," 2024
[9] Mrosek et al., "Reduced-order modeling of vehicle aerodynamics via proper orthogonal decomposition," 2019

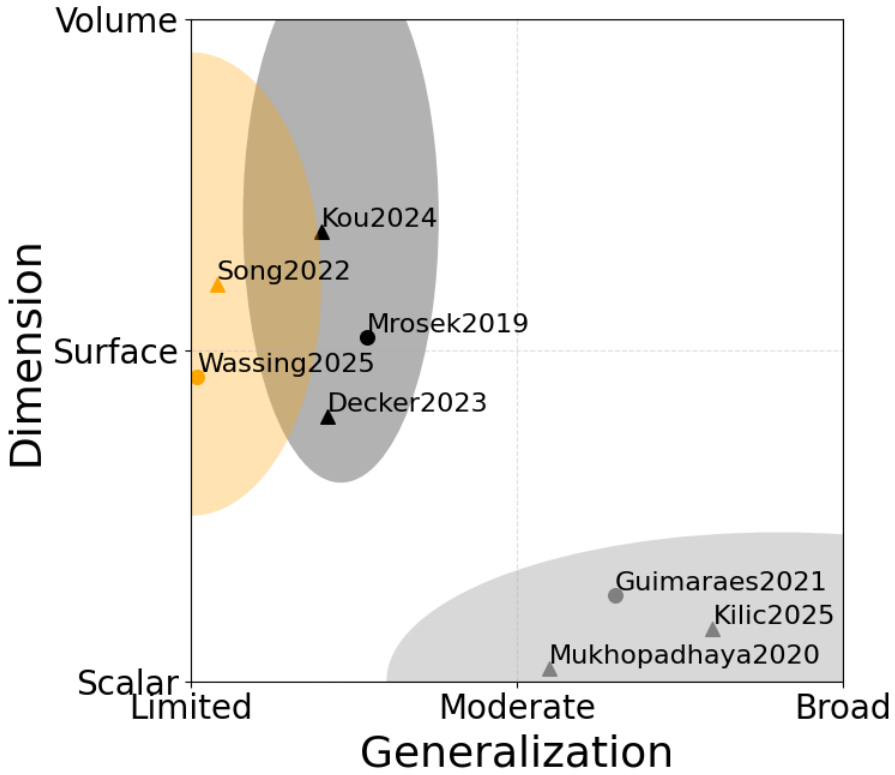
aerospacedesignlab

Literature Review

1. Scalar Surrogate Models
2. Reduced-Order Models
3. Image-based Neural Networks
   - › CNNs
   - › PINNs



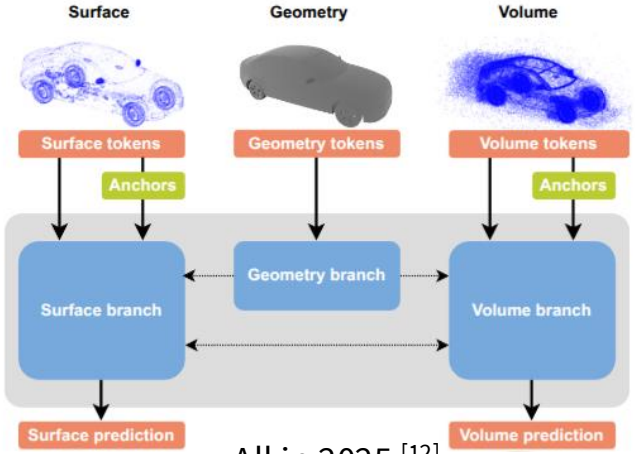
Wassing 2025 [10] Song 2022 [11]



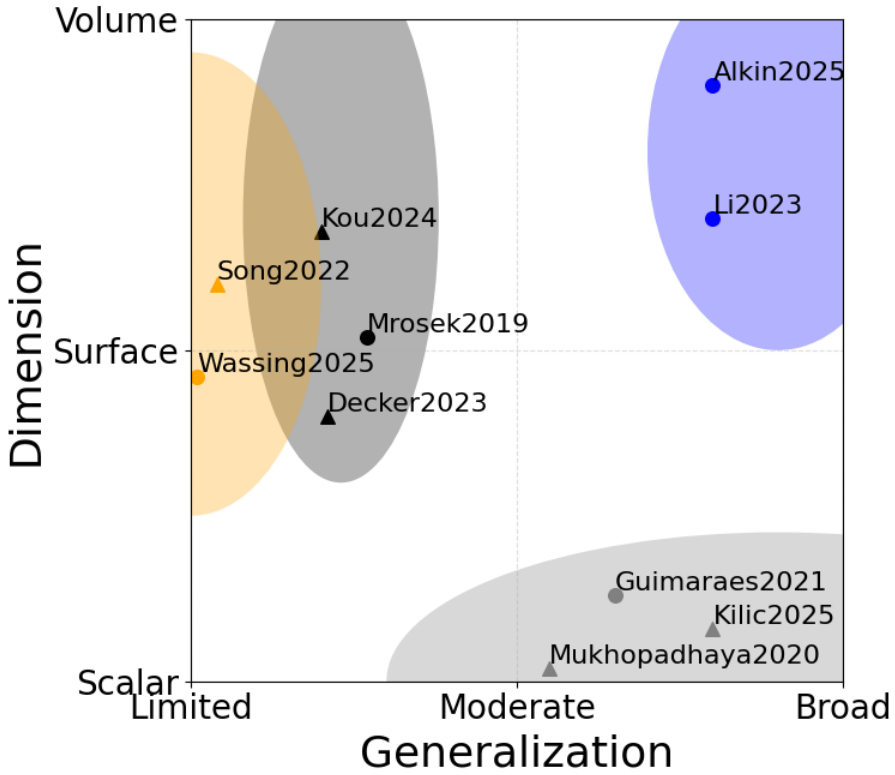
[10] Wassing et al., "Physics-Informed Neural Networks for Transonic Flows around an Airfoil," 2025
[11] Song and Tartakovsky, "Transfer learning on multifidelity data," 2022

Literature Review

1. Scalar Surrogate Models
2. Reduced-Order Models
3. Image-based Neural Networks
4. Operator Learning



Alkin 2025 [12]

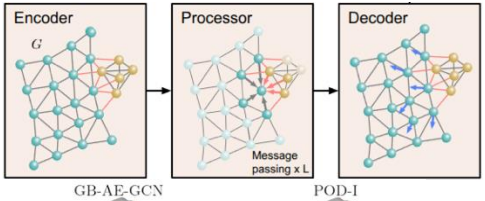


[12] Alkin et al., “AB-UPT: Scaling Neural CFD Surrogates for High-Fidelity Automotive Aerodynamics Simulations via Anchored-Branched Universal Physics Transformers,” 2025

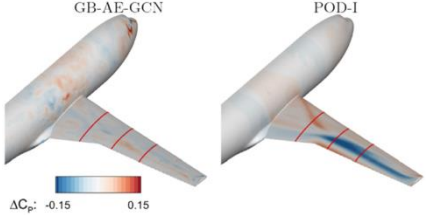
aerospacedesignlab

Literature Review

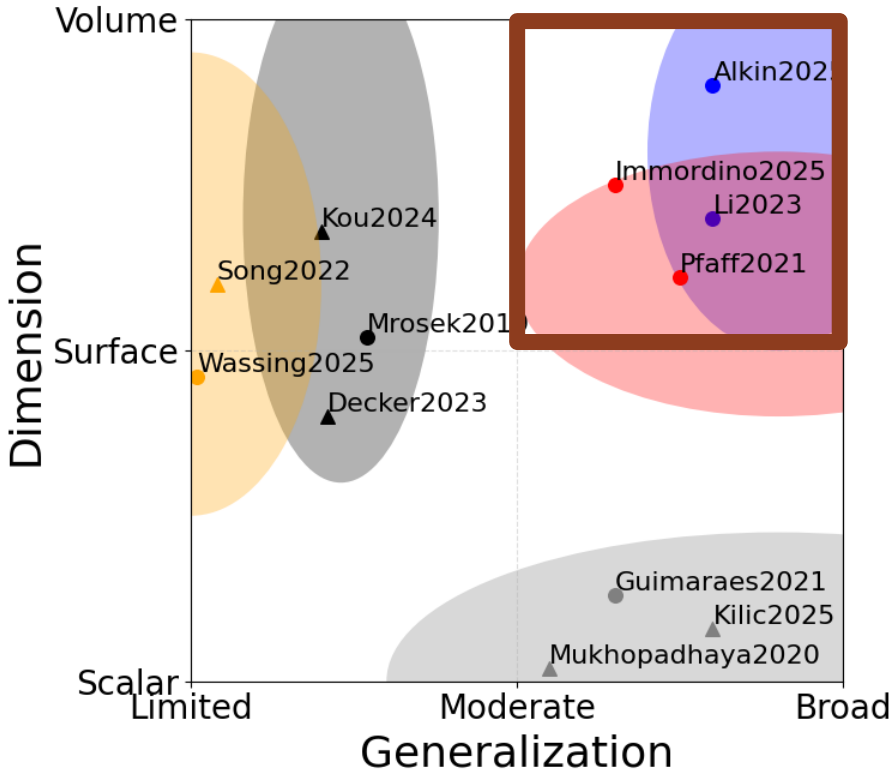
1. Scalar Surrogate Models
2. Reduced-Order Models
3. Image-based Neural Networks
4. Operator Learning
5. Geometry-Aware NNs
   - › Graph Neural Networks



Pfaff 2021 [13]



Immordino 2025 [14]



[13] Pfaff et al., “Learning Mesh-Based Simulation with Graph Networks,” 2021
[14] Immordino et al., “Predicting transonic flowfields in non–homogeneous unstr. grids using autoencoder GCN,”2025

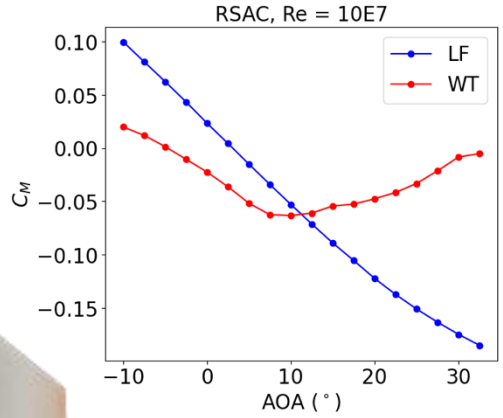
Research Objectives

Develop **multi-fidelity field data surrogate** modeling framework that brings **high-fidelity-informed corrections** to **low-fidelity design workflows**.

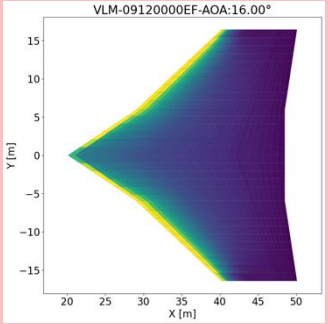
aerospacedesignlab

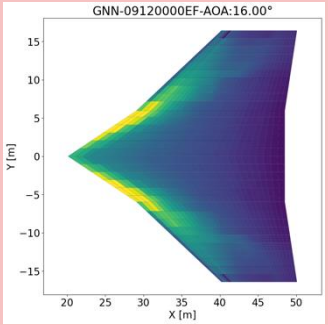
The Big Picture, Refined

LF Simulations

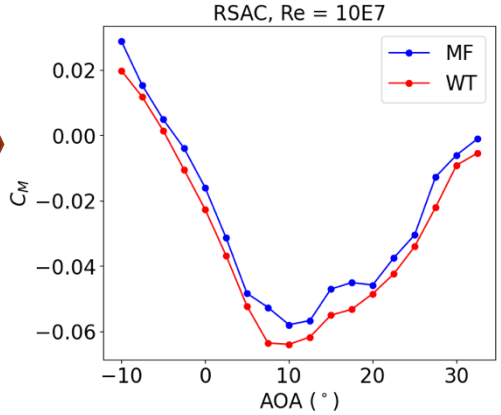


MF Surrogate





MF Analysis



MF Fields

Research Objectives

I
Multi-fidelity Field Surrogate Modeling

"How to build it?"

- Develop VortexNet to correct VLM pressure fields with CFD data.
- Deliver higher fidelity (w/ VLM) and geometric generalizability.

II
Dataset and Scaling

"What's the cost?"

- Quantify data–model–accuracy scaling for VortexNet training.
- Release an open-source multi-fidelity low-speed SST aerodynamics dataset.

III
Conceptual Design Integration

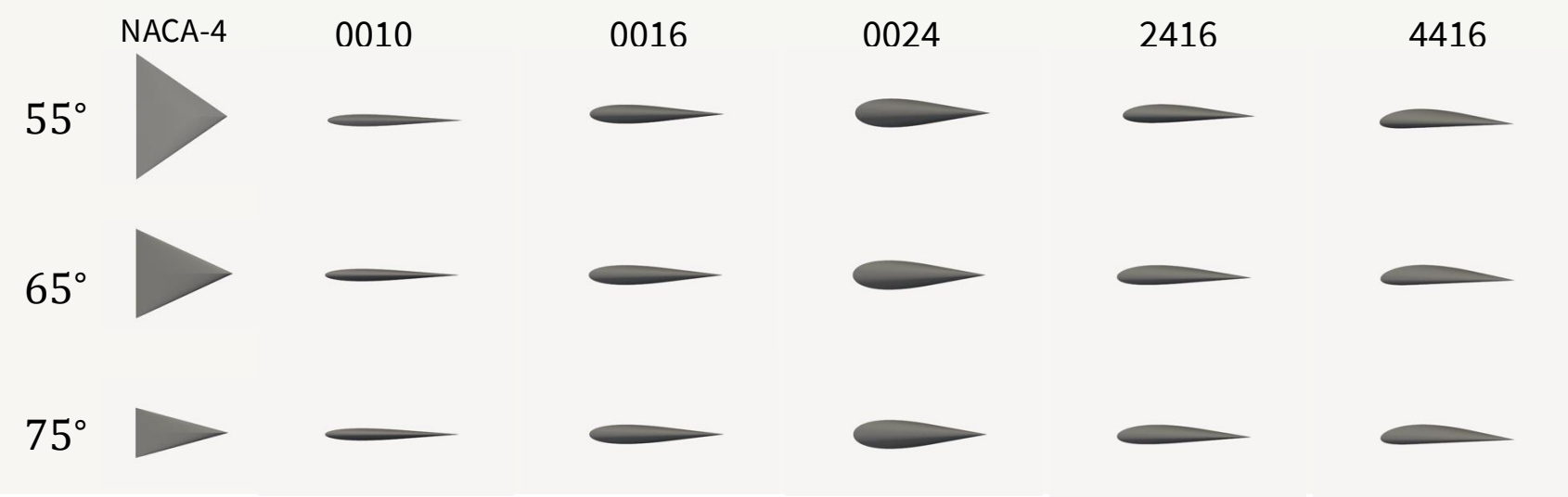
"How to use it?"

- Integrate VortexNet in conceptual design.
- Demonstrate fidelity gains through a shape optimization problem.

I
Multi-fidelity Field Surrogate Modeling

Problem Definition/Dataset

15 Delta wings:



40 free-stream conditions:

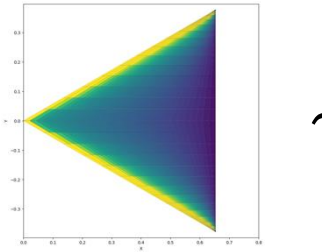
- Re_∞ : 6.5 × 10^6 to 1.0 × 10^7; AOA: 0° to 20°; Ma_∞ : 0.35 to 0.5

aerospacedesignlab

Reference State and Graph Definition

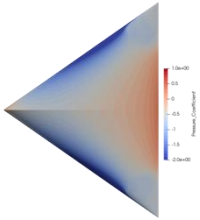
LF and HF fields approximate the same underlying physical state.

VLM ΔC_p^{LF} SUAVE



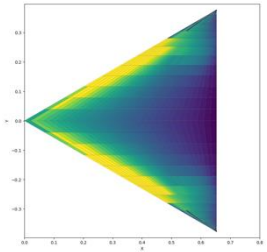
$\sim$

CFD C_p SU2 code

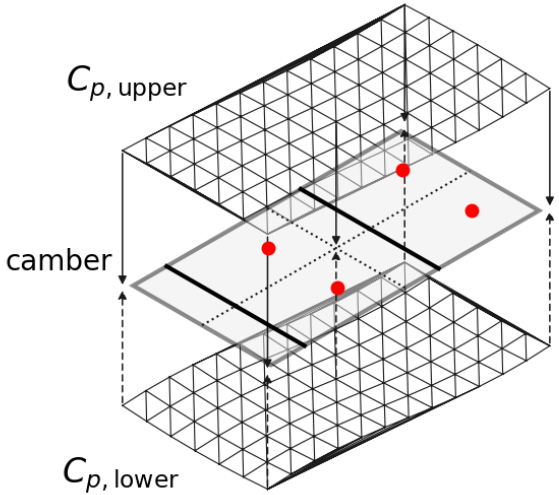


$\rightarrow$

CFD ΔC_p^{HF}

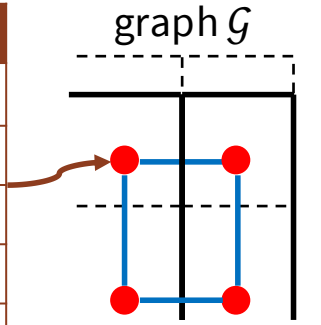


ΔC_p establishes the **reference state**, where multi-fidelity sources can be correlated.



Graph Conversion

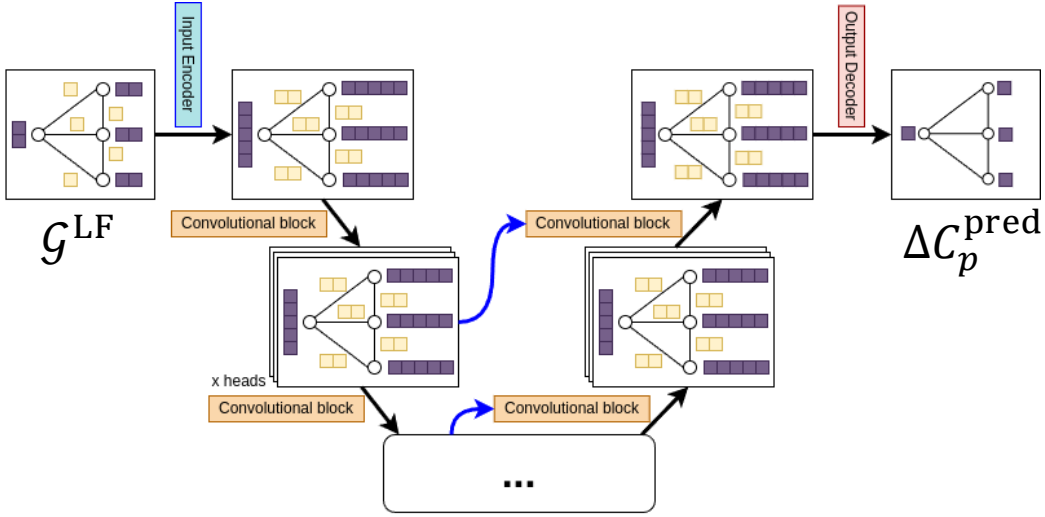
Nodal Feature [12x1]	
VLM ΔC_p	2x curvature [m^{-1}]
AOA [rad]	2x chord-wise slope
Ma_∞	control point x_c [m]
Re_∞	control point y_c [m]
thickness [m]	control point z_c [m]



VortexNet Architecture

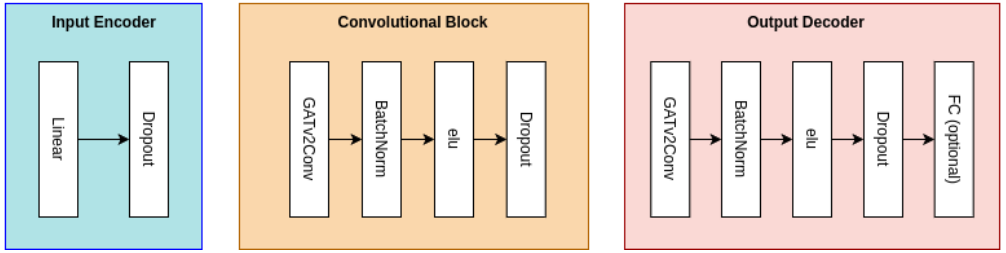
Key features:

- U-net style autoencoder.
  - ✓ Capture high-and-low frequency features.
- Multi-head message aggregation convolutional block using Graph Attention Network [15].



The network learns:

$$\Delta C_p^{\mathrm{HF}} \sim \Delta C_p^{\mathrm{pred}} = \mathcal{F}(\mathcal{G}^{\mathrm{LF}}, \omega_L)$$



[15] Veličković et al., "Graph Attention Networks," 2018

VortexNet Training and Loss Function

Multi-fold training:

- In each fold, 25% of the training set is reserved as cross-validation set.

Loss function:

$$\mathcal{L} = \text{SmoothL1Loss}\left(\Delta C_p^{\text{HF}}, \Delta C_p^{\text{pred}}\right) + \Lambda\, \text{PhysicalLoss}\left(\Delta C_p^{\text{pred}}\right)$$

Physical Loss improves model generalizability.

$$\text{PhysicalLoss} = \left\| \text{AIC}\, \Gamma^{\text{pred}}\left(\Delta C_p^{\text{pred}}\right) - \text{RHS} \right\|$$

Exponential decay weight Λ allows VortexNet to shift learning attention from global features (physics driven) to local features (data driven) towards the end:

$$\Lambda = \Lambda^{\max} \exp\left[\frac{\Lambda^{min}/\Lambda^{max}}{k-1}(\text{i} - 1)\right] \text{ for } i \in \{1, \ldots, K\}$$

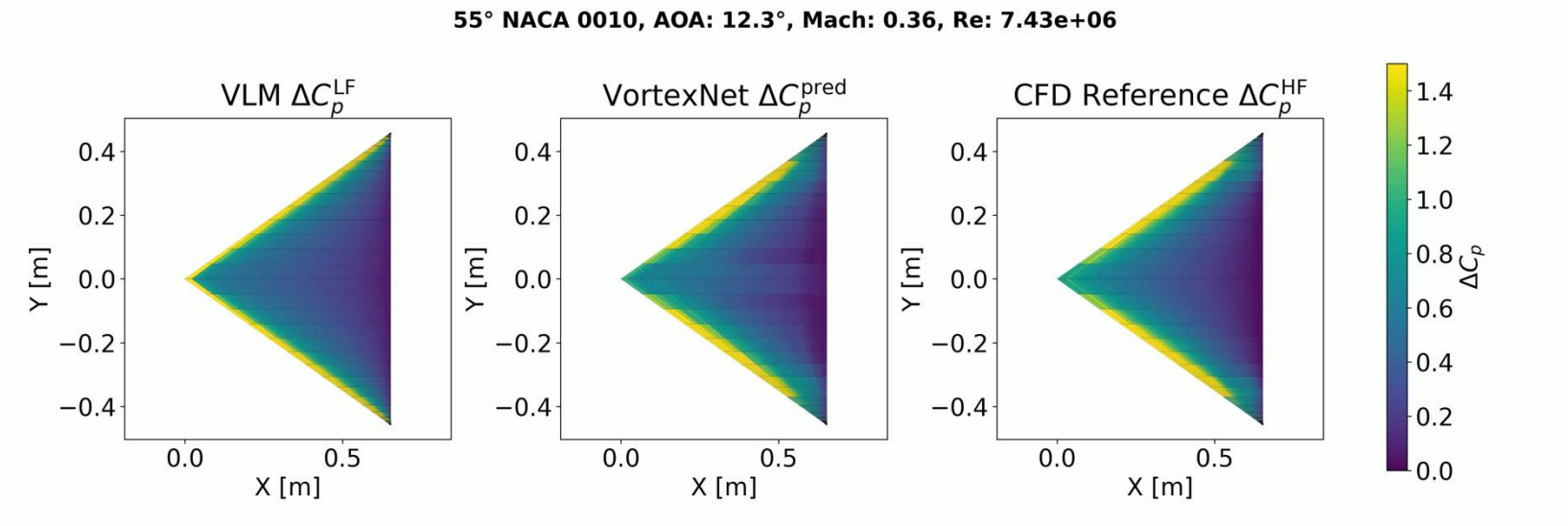
Results Organization

1. Performance evaluation:
   - Evaluate geometric generalizability through **interpolated geometries**

	0010	0013	0016	2416	3416	4416
55°	T		T	T		T
60°		H			H	
65°	T		T	T		T
70°		H			H	
75°	T		T	T		T

   - 2 Tests:
     - Pressure field prediction: qualitative and RMSE($\Delta C_p^{\text{pred}}, \Delta C_p^{\text{HF}}$)
     - $\mathbf{C_L}, \mathbf{C_M}$ prediction: NRMSE$\left(q^{\text{pred}}, q^{\text{CFD}}\right) = \text{RMSE}/\overline{q^{\text{CFD}}}$
2. Prediction mechanism explanation through information ablation

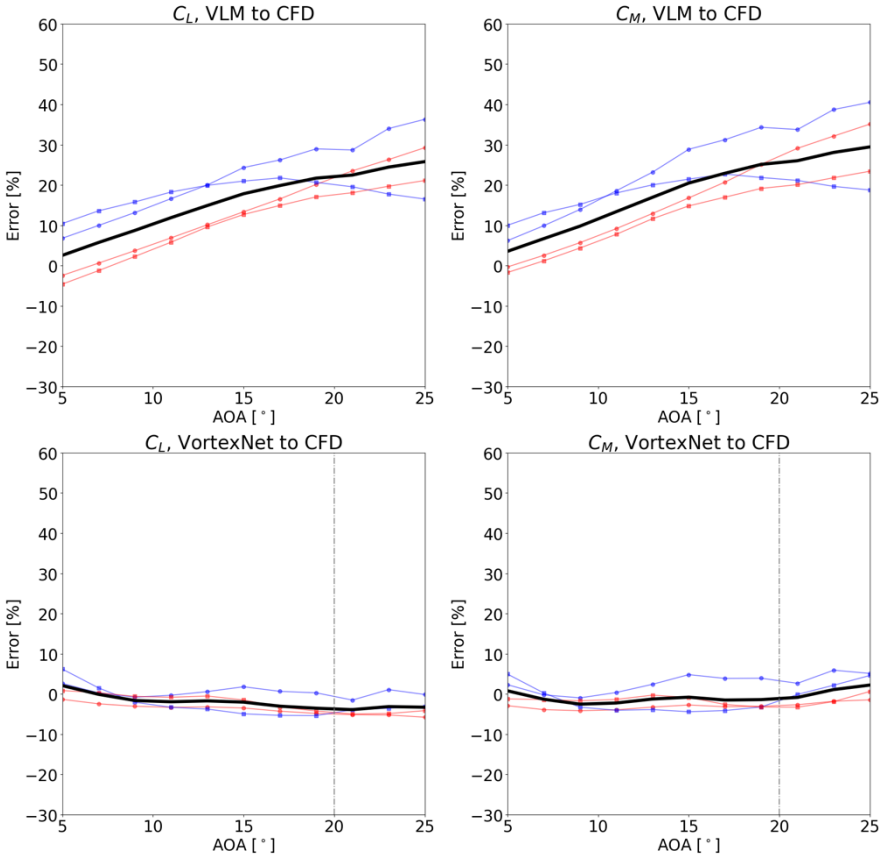
Test Geometries, Pressure Field Prediction

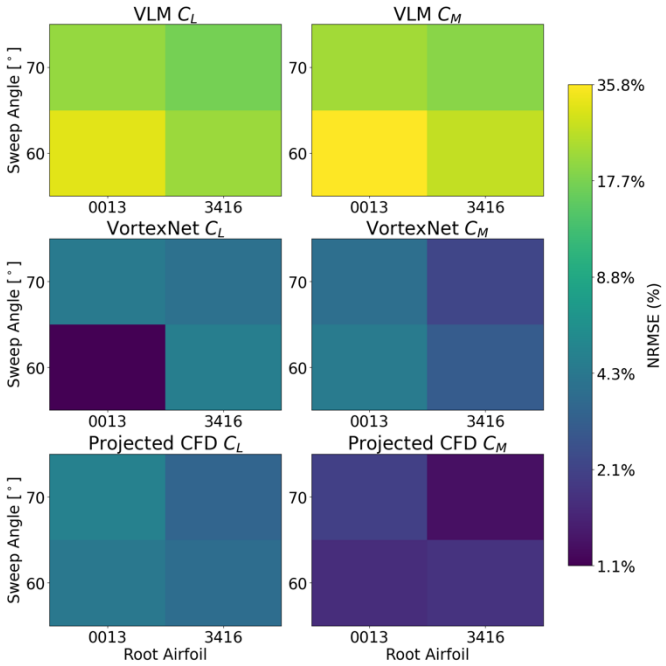


Sweep Angle [°]	Root Airfoil (NACA 4-digits)	
	0013	**3416**
60	0.1441	0.2695
70	0.2106	0.3293

aerospacedesignlab

Test Geometries, QOIs Prediction





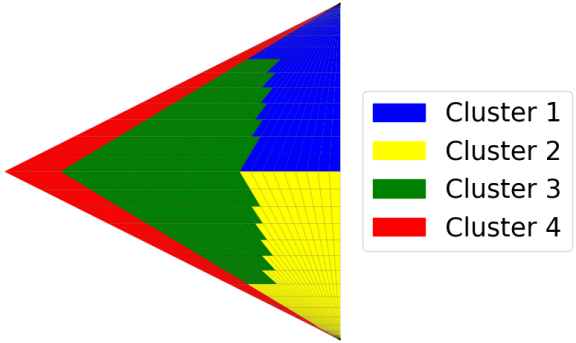
- 7x NRMSE reduction

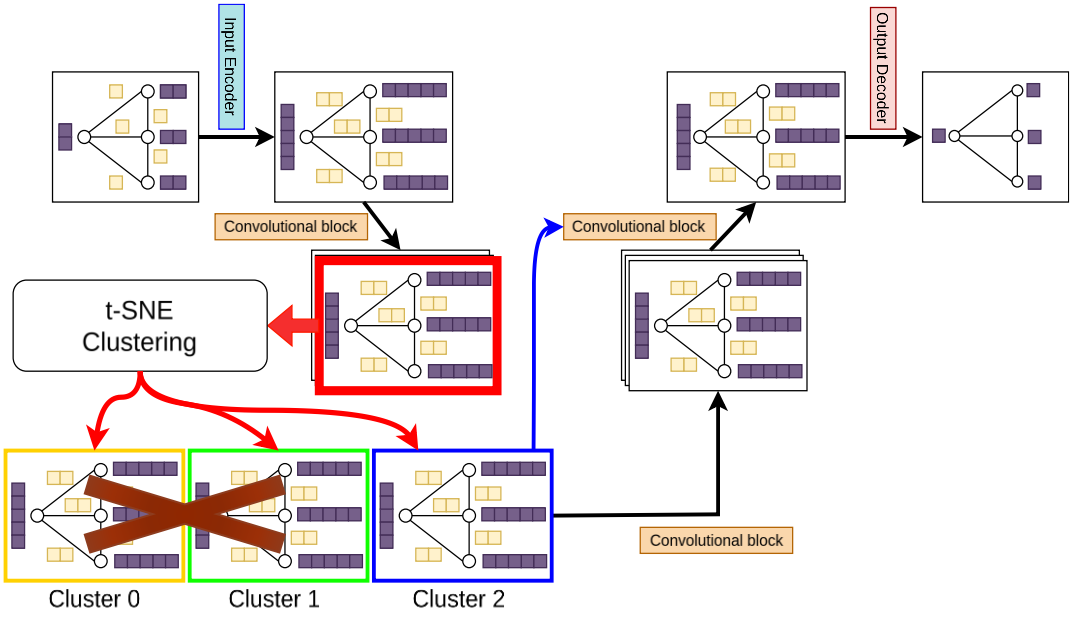
Mechanism Explanation: Method

Objective: understand how graph features are utilized across latent layers.

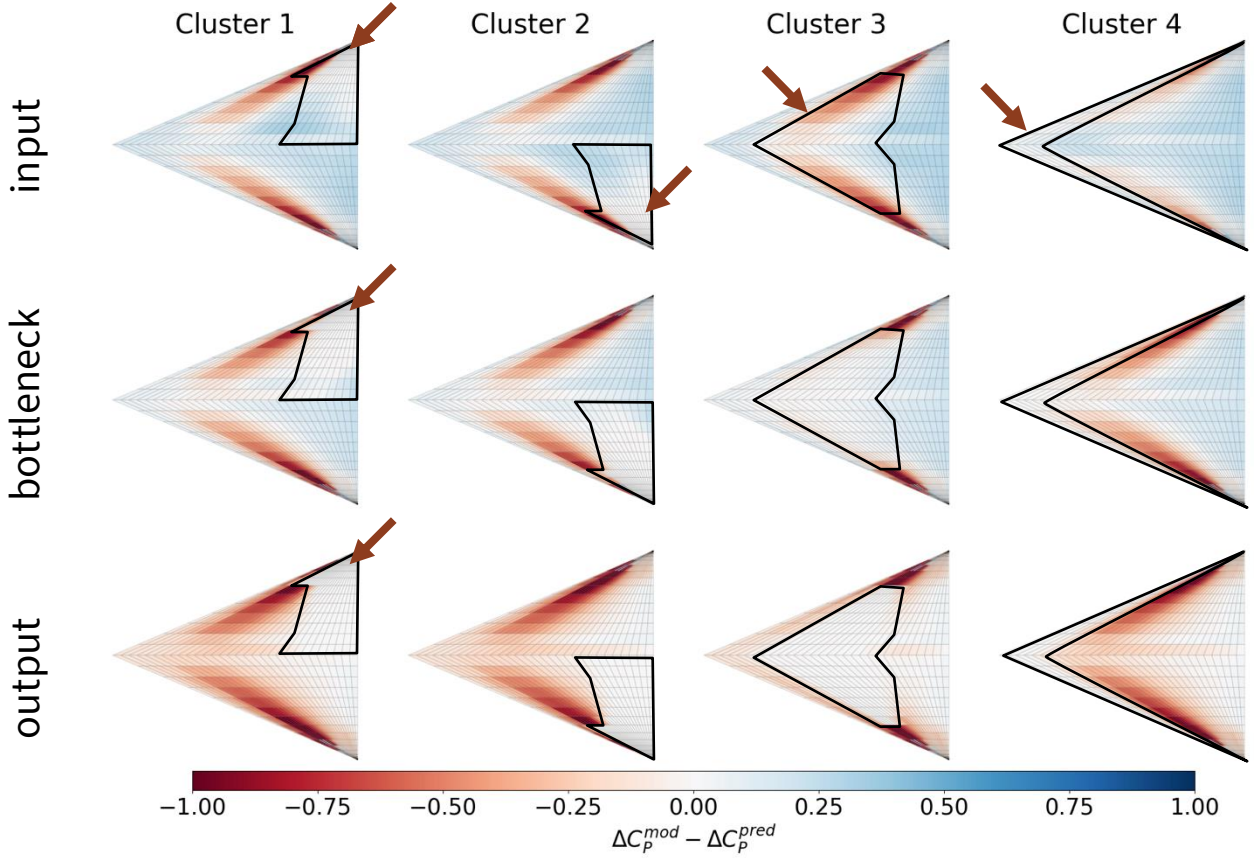
Method: ablation study on latent space information.

- Recursively allowing selected nodes to propagate while eliminating others.
- Repeated across all convolutional layers.





Results


Cluster 1
Cluster 2
Cluster 3
Cluster 4
input
bottleneck
output
−1.00
−0.75
−0.50
−0.25
0.00
0.25
0.50
0.75
1.00
$\Delta C_P^{mod} - \Delta C_P^{pred}$

Section Summary

VortexNet: a **first-of-its-kind field surrogate** bridging **high- and low-fidelity aerodynamics**.

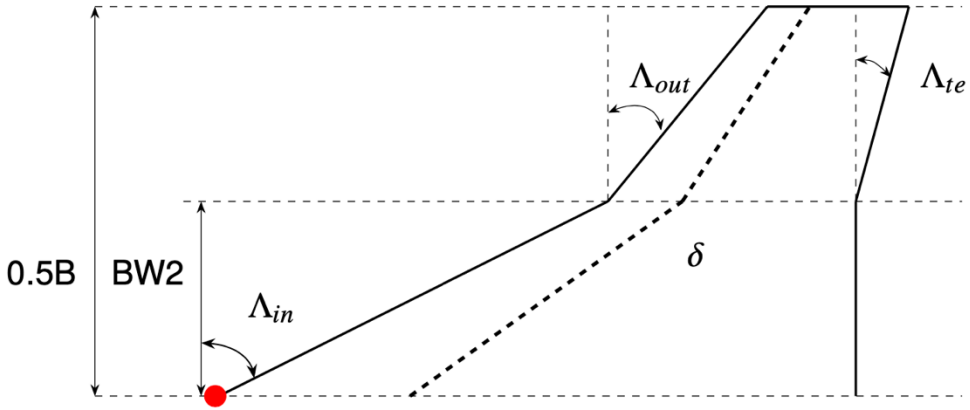
- Shows promising **field prediction quality** for conceptual design purpose.
- Achieves notable aerodynamic coefficients accuracy gains (**>5x**).
- Generalizable across unseen geometries and flow conditions.

Latent space study reveals that VortexNet achieves its prediction capability by blending **geometrical information** and **flow features** through **non-linear mapping**.

II – A
Multi-fidelity Dataset: “double-delta-aero”

Design Space Definition

- Design space: 6 design variables
- Design variable bounds based on the SCALOS [2] and the HSCT [17] program.
- A collaborative effort with colleagues from the University of Washington.



Design Variable	Lower	Upper
δ (°)	-8	7
Λ_{in} (°)	47	67
Λ_{out} (°)	35	60
Λ_{te} (°)	0	25
BW2	0.32	0.47
B (inches)	1100	1300

[17] Welge et al., "N+2 Supersonic Cencept Development and System Integration," 2010

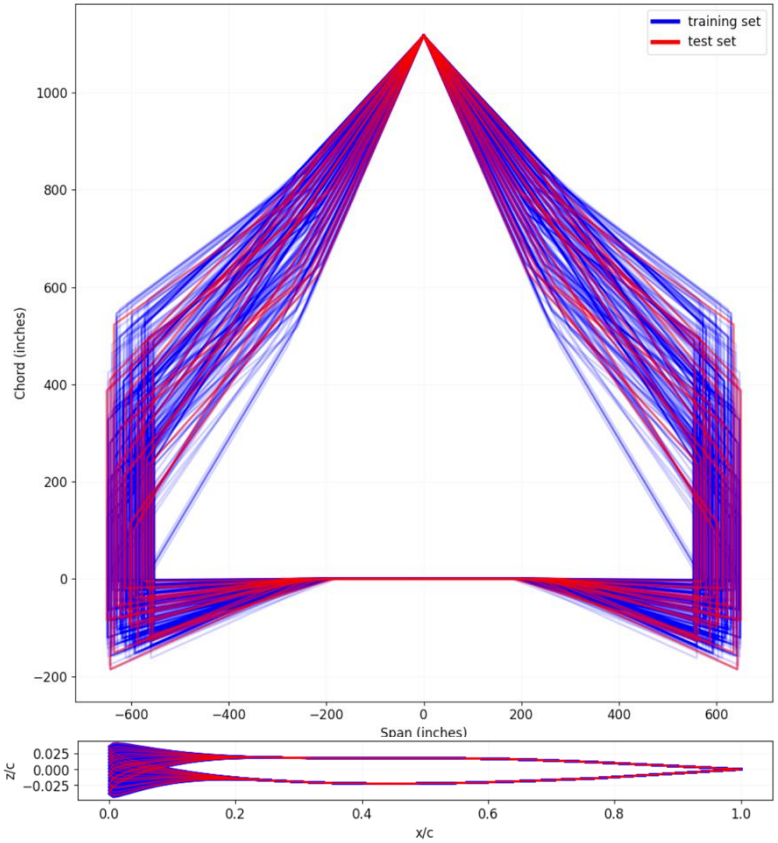
Design of Experiments

- DOE objectives:
  1. Uniform design-space coverage.
  2. Nested sampling points.
  3. Enable multiple model evaluation tasks.

- Saltelli's Sobol sequence sampling [18]

- 256 vehicles generated.

- Also provide a LHS holdout test set with 16 geometries.

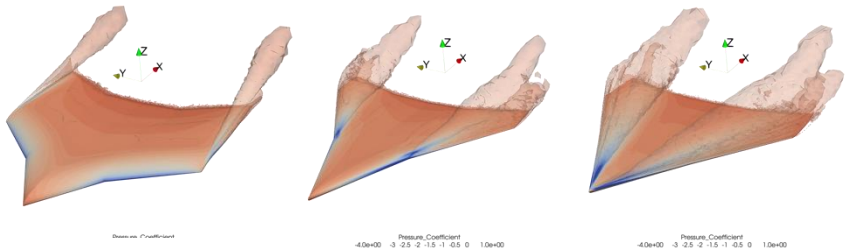


[18] Saltelli et al., "A quantitative model-independent method for global sensitivity analysis of model output,"1999

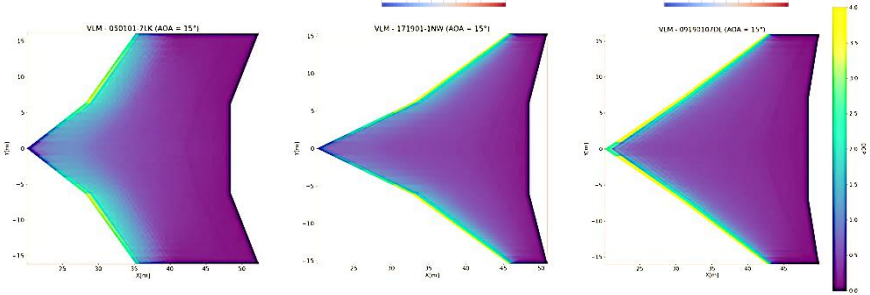
aerospacedesignlab

The “double-delta-aero” Dataset

CFD Solutions



VLM Solutions



QOI and config.

- Geometry: DVs and reference values.
- DOE configuration.
- **QOIs**: $C_L, C_D, C_{SF}, C_{M;x,y,z}$, and $C_{F;x,y,z}$.
- CFD convergence history.
- CFD and VLM run configurations.

Data Availability: HuggingFace Repository

- 2448 flow snapshots across 272 geometries.
- AOA from 11° to 19° (9 AOAs)
- Low-speed conditions: $Ma = 0.3$.
- 0.31M CPU-hours.
- Nested dataset:

Level	1	2	3	4	5	6
Geometry Count	8	16	32	64	128	256
Dataset Size (D)	72	144	288	576	1152	2304

II – B
VortexNet Scaling

I
Multi-fidelity Field Surrogate Modeling

II
Dataset & Scaling

III
Conceptual Design Integration

Empirical Scaling Law

- Empirical scaling laws describe **how model performance improves** with **increasing training/modeling scale**:

$$\epsilon = \left(\frac{N_C}{N}\right)^{\alpha_N} + \left(\frac{D_c}{D}\right)^{\alpha_D} + \epsilon_0 \qquad [19]$$

- N: model parameters; D: flow snapshots
- ϵ : mean squared error (MSE) on the **holdout test set**.

$$\epsilon = MSE(\Delta C_P^{\mathrm{HF}}, \Delta C_P^{\mathrm{pred}})$$

- **Objectives of the scaling experiment**:
  1. Empirically estimate the scaling exponents.
  2. How many geometries is sufficient to train the surrogate model?

[19] Kaplan et al., "Scaling Laws for Neural Language Models," 2020

Experiment Setups

Data size scaling (**6 dataset**):

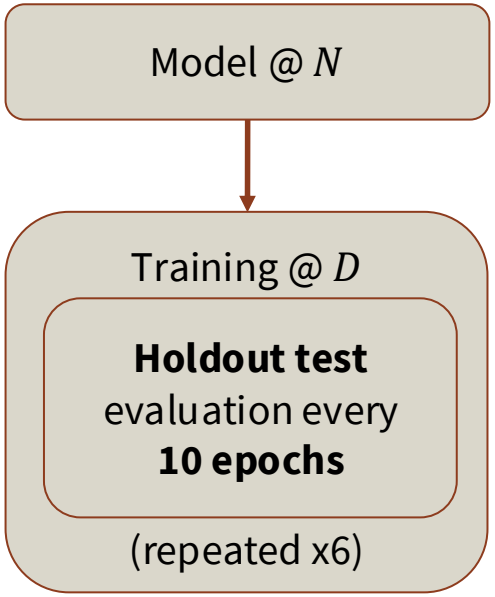
- Increased geometric sampling density

Level	1	2	3	4	5	6
Geometry Count	8	16	32	64	128	256
Dataset Size (D)	40	80	160	320	640	1280

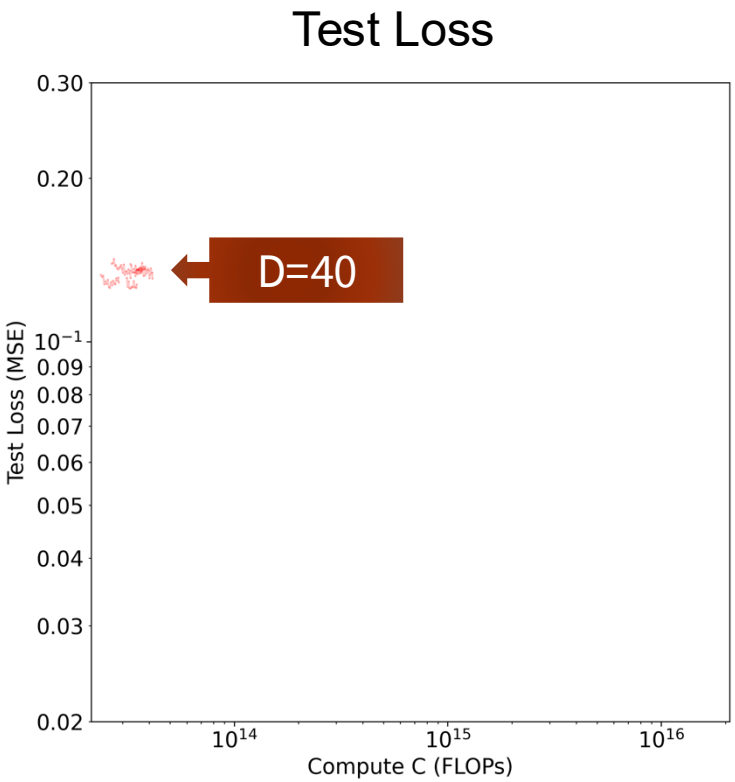
Model size scaling (**4 models**):

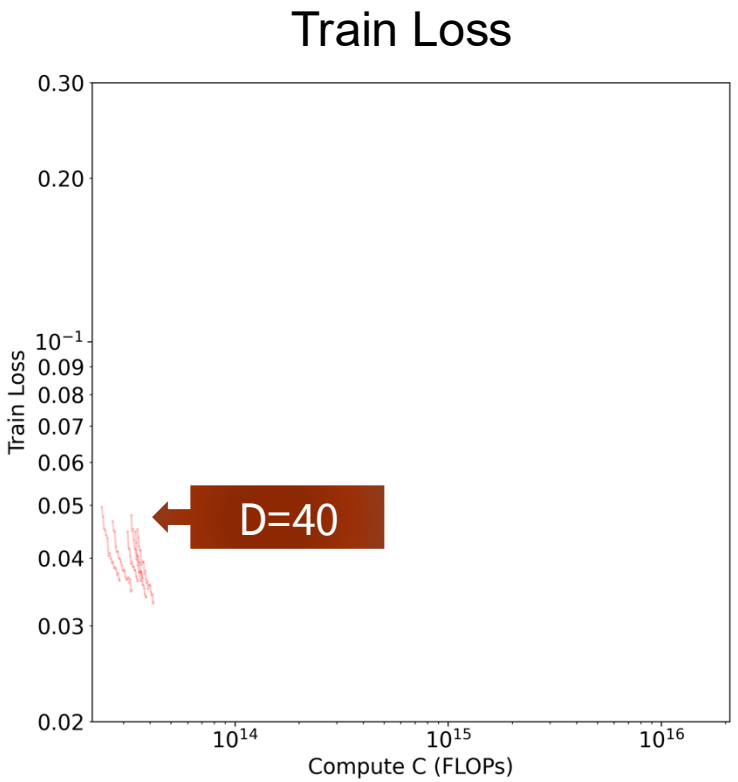
- Higher latent nodal feature dimension

Level	hidden_feature	heads	Model Size (N)
Mini	13	5	0.103×10^6
Small	27	5	0.538×10^6
Medium	37	5	1.144×10^6
Large	45	6	2.411×10^6

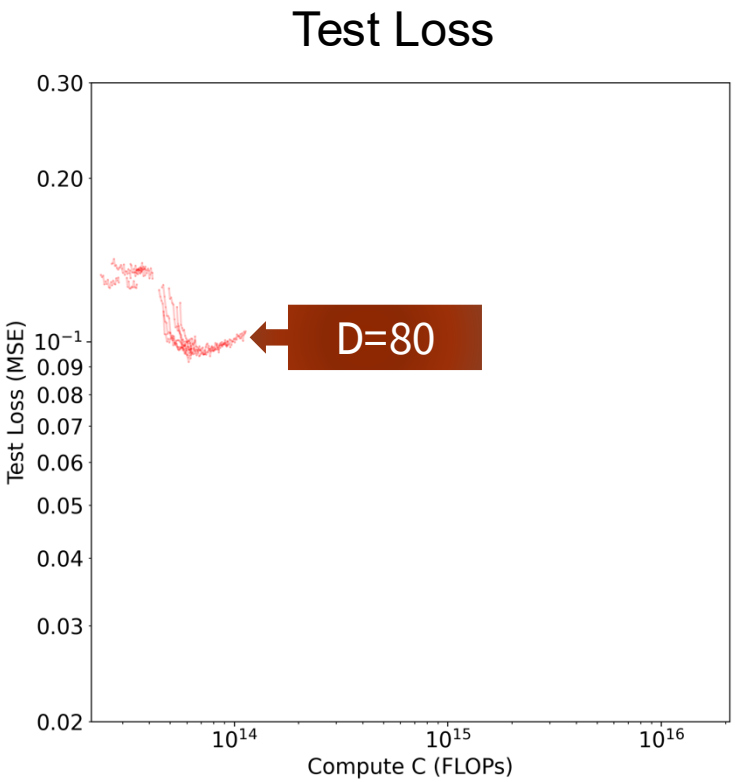


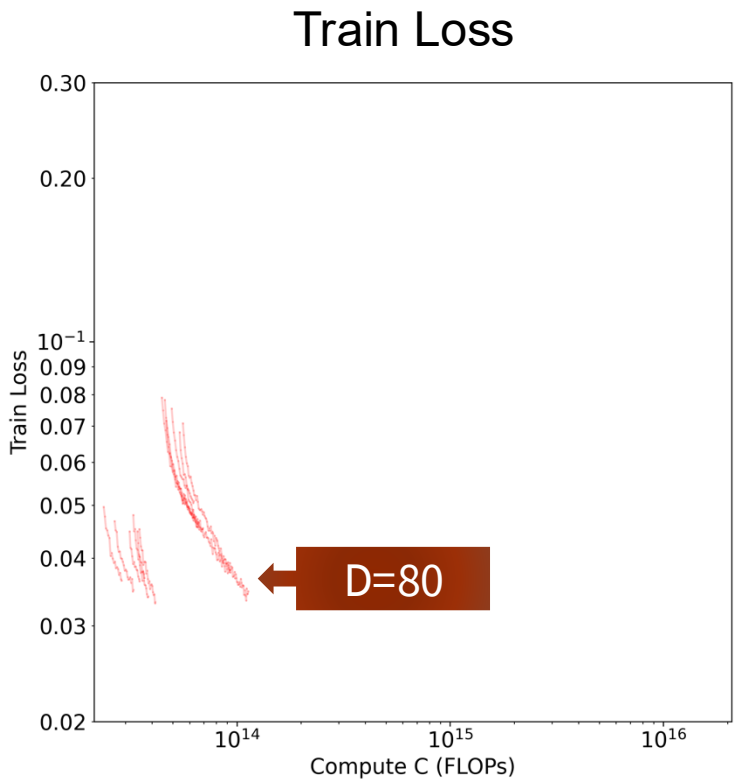
Test and Train Loss History





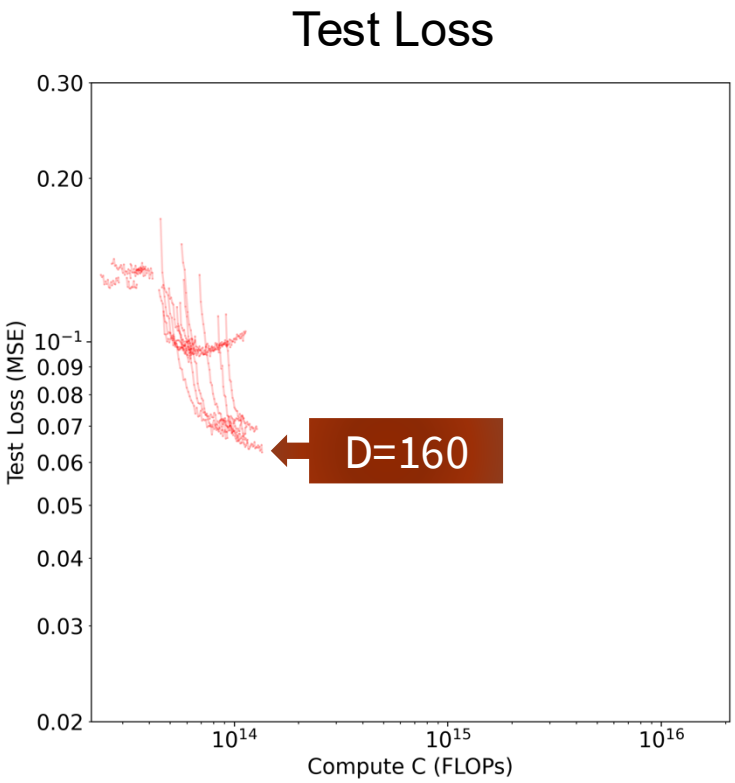
Test and Train Loss History

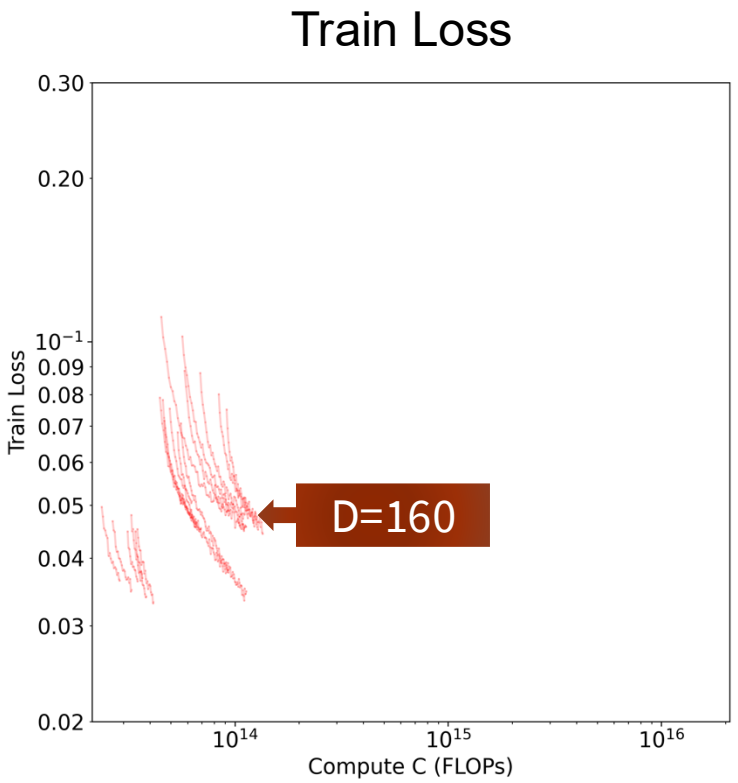




aerospacedesignlab

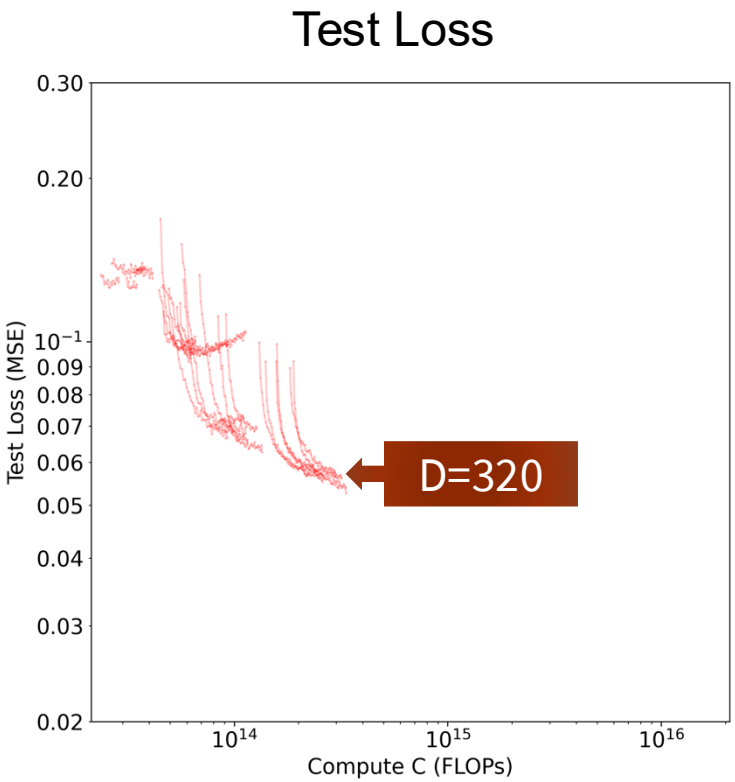
Test and Train Loss History

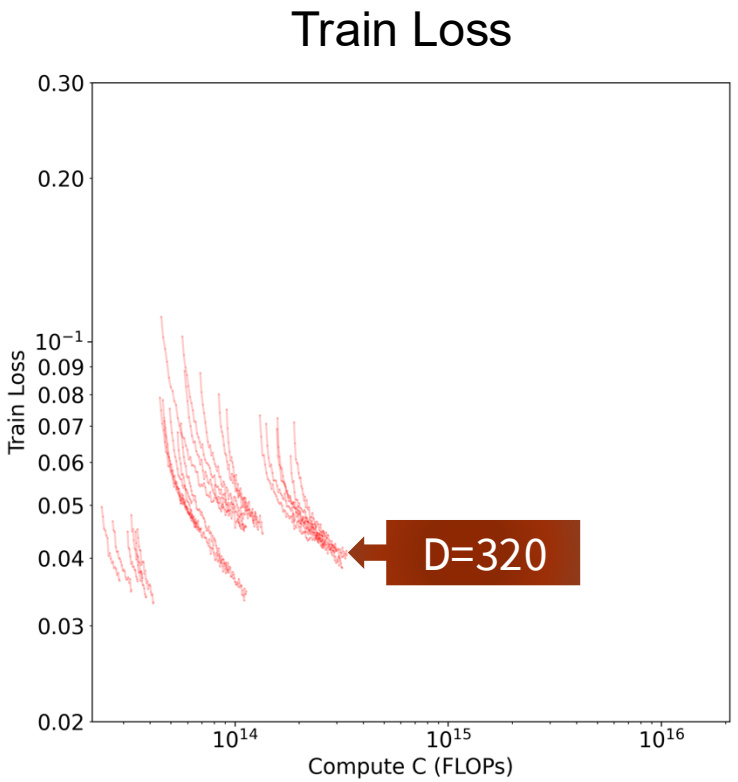




aerospacedesignlab

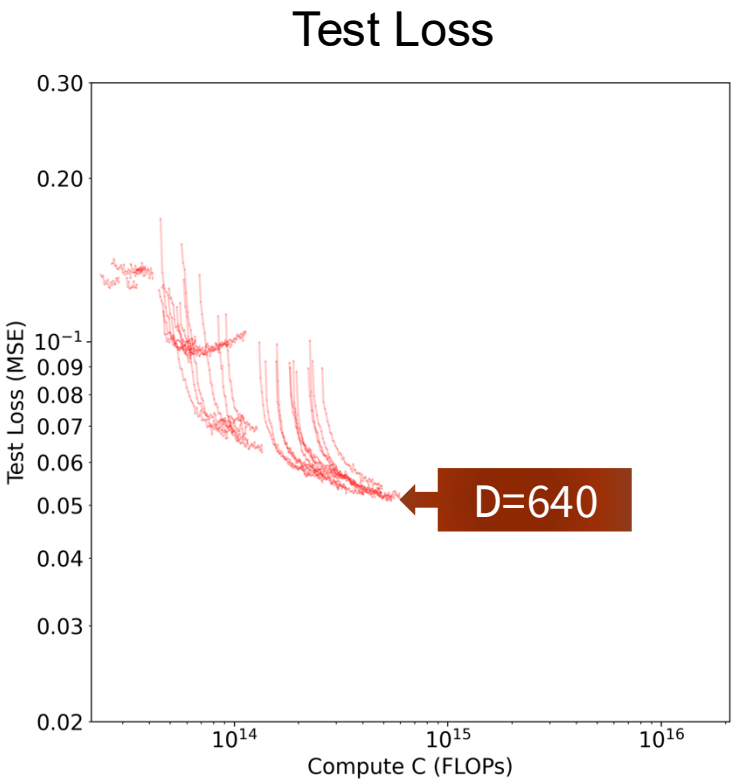
Test and Train Loss History

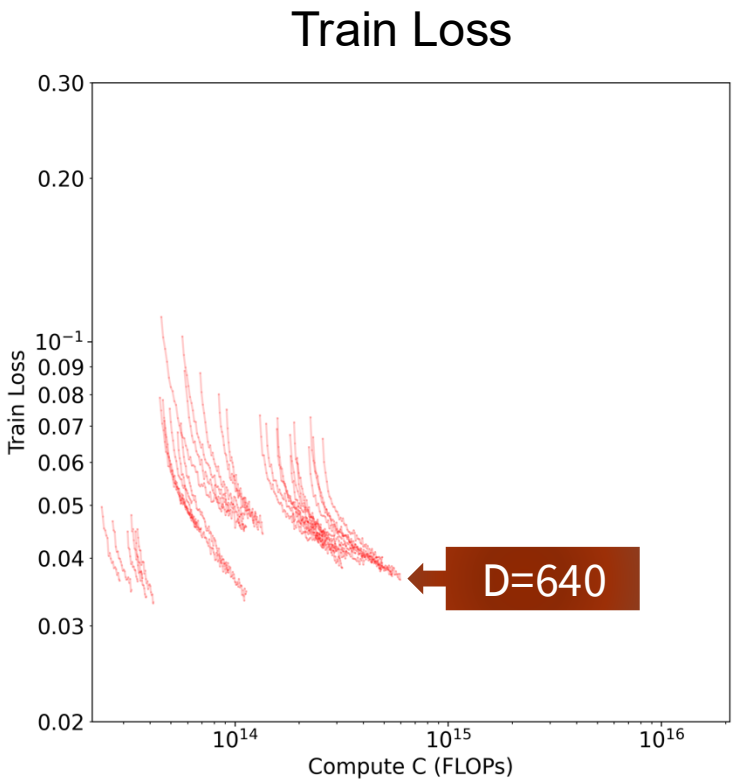




aerospacedesignlab

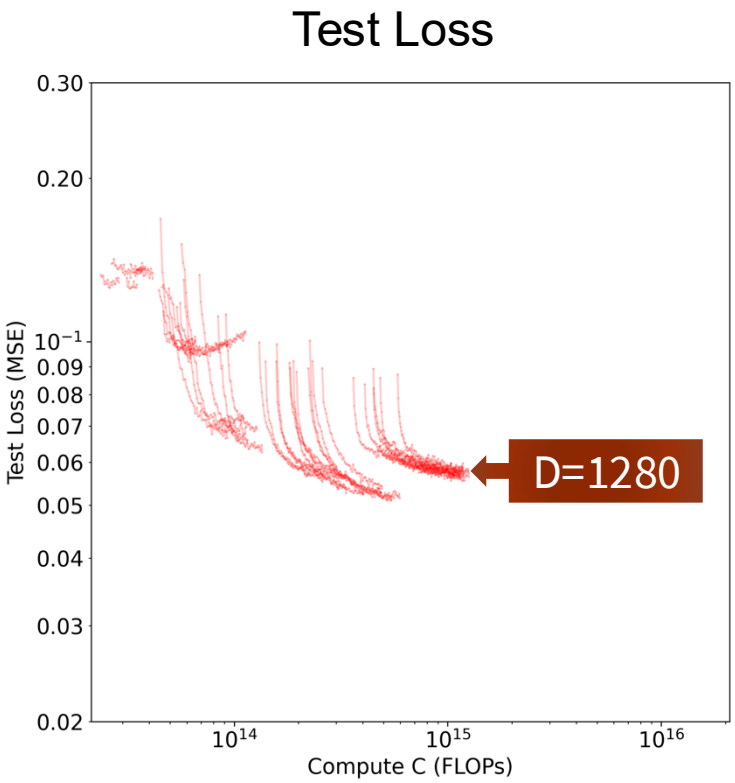
Test and Train Loss History

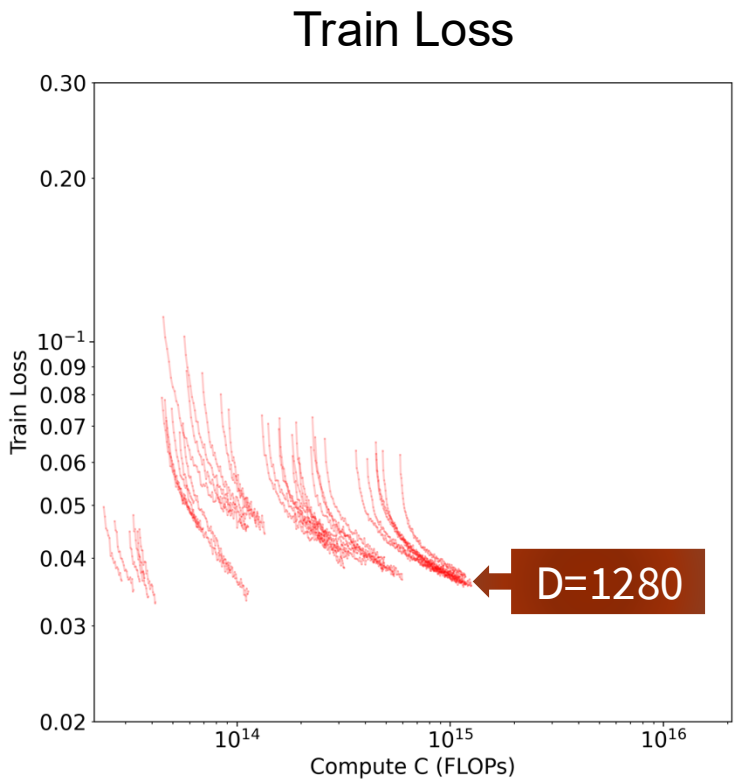




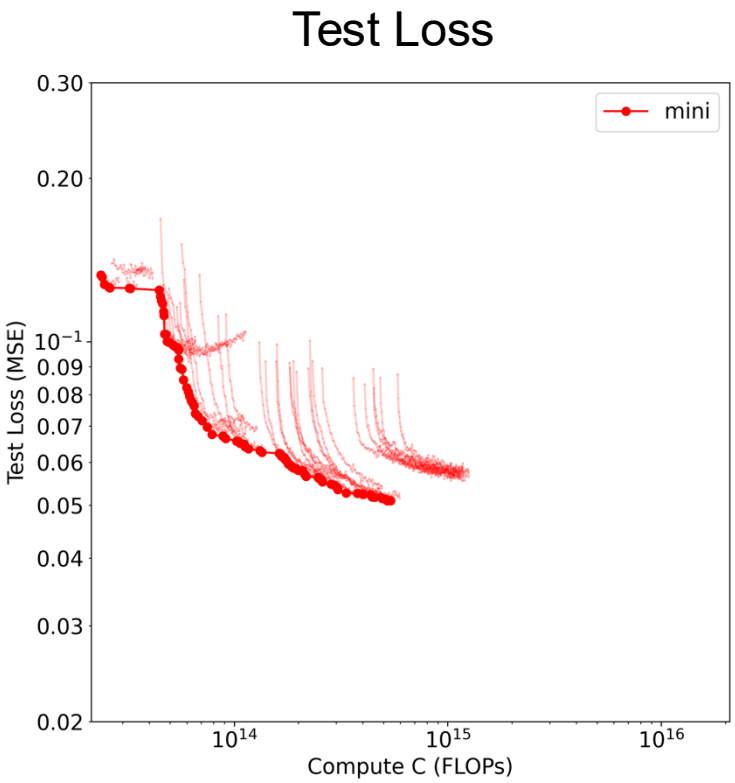
aerospacedesignlab

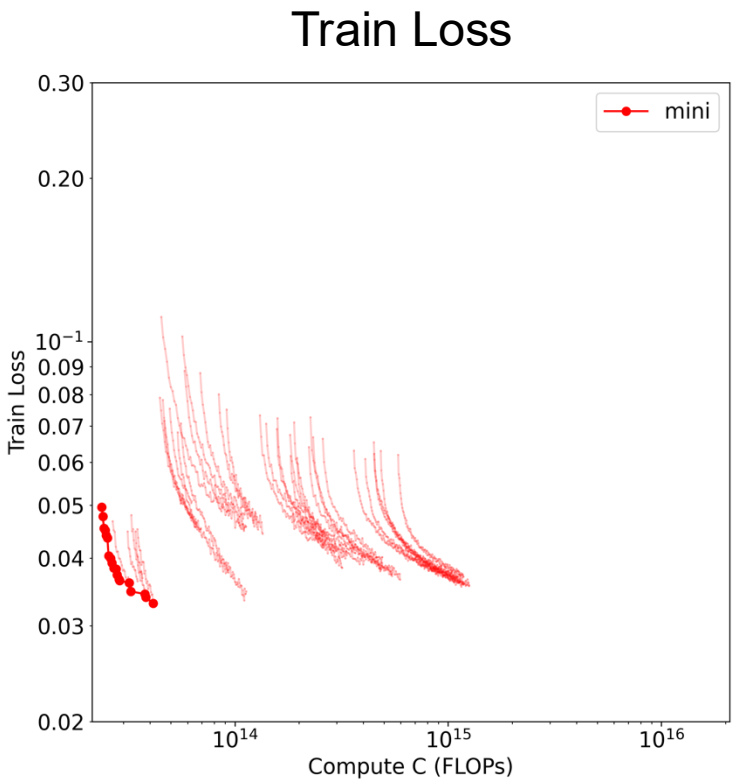
Test and Train Loss History



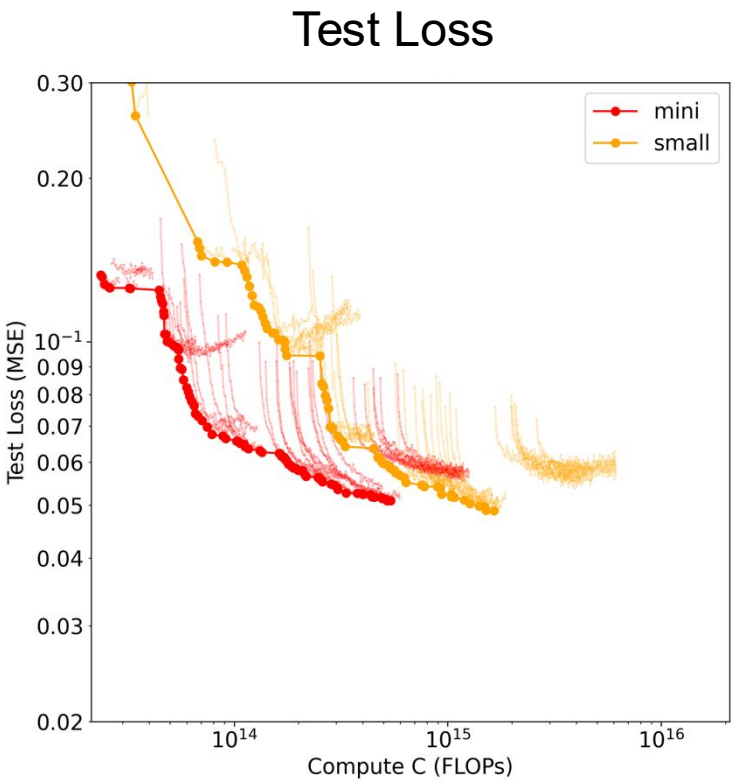


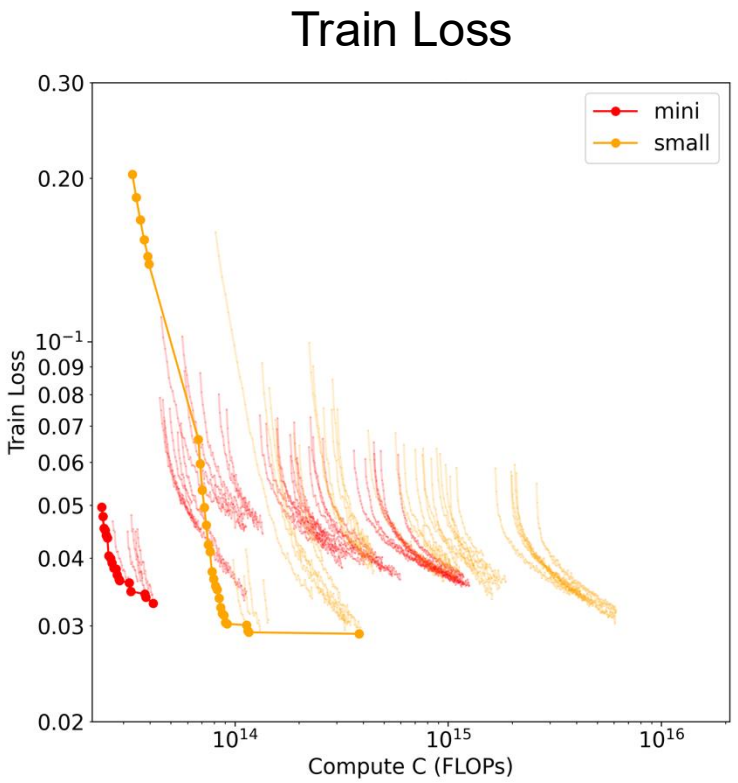
Test and Train Loss History



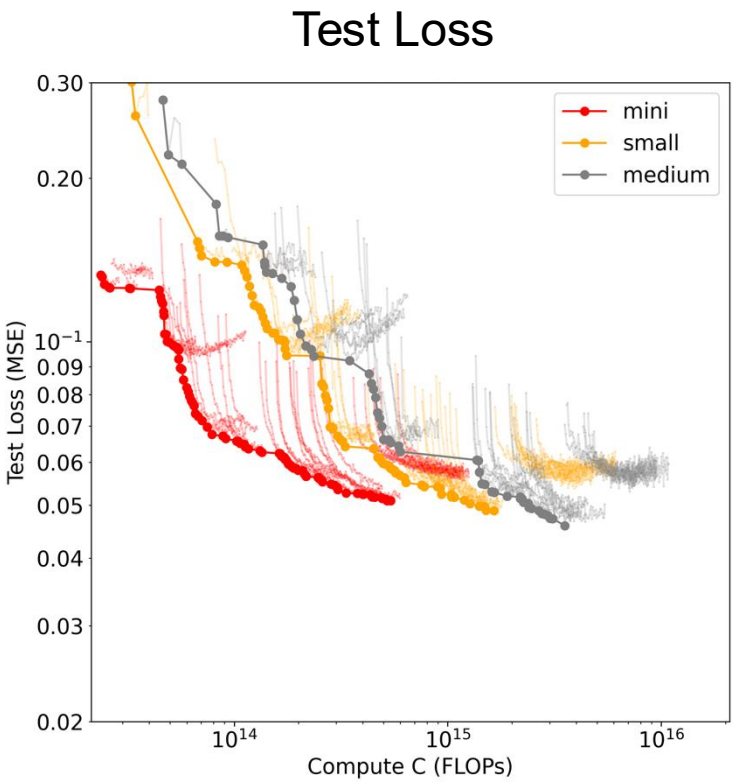


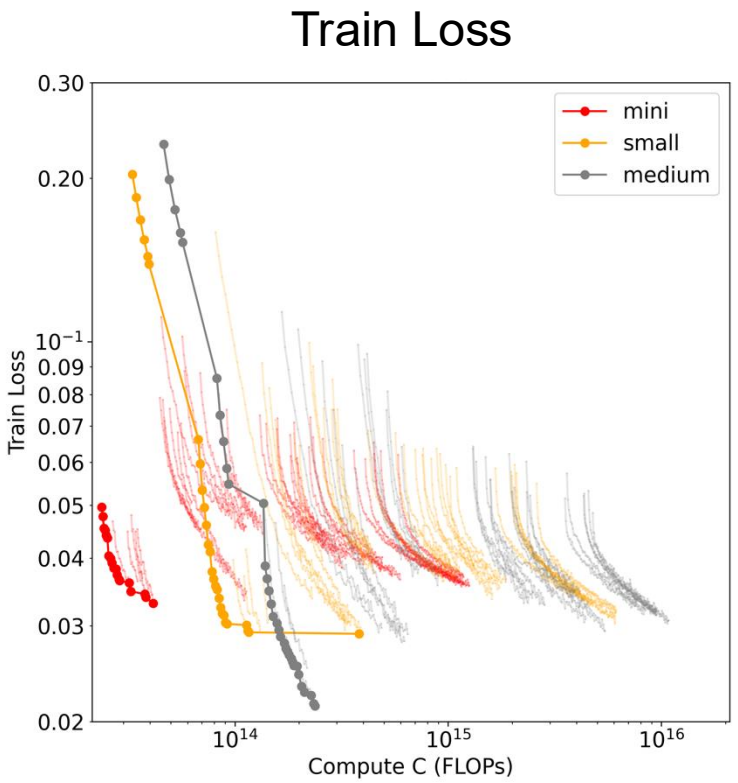
Test and Train Loss History



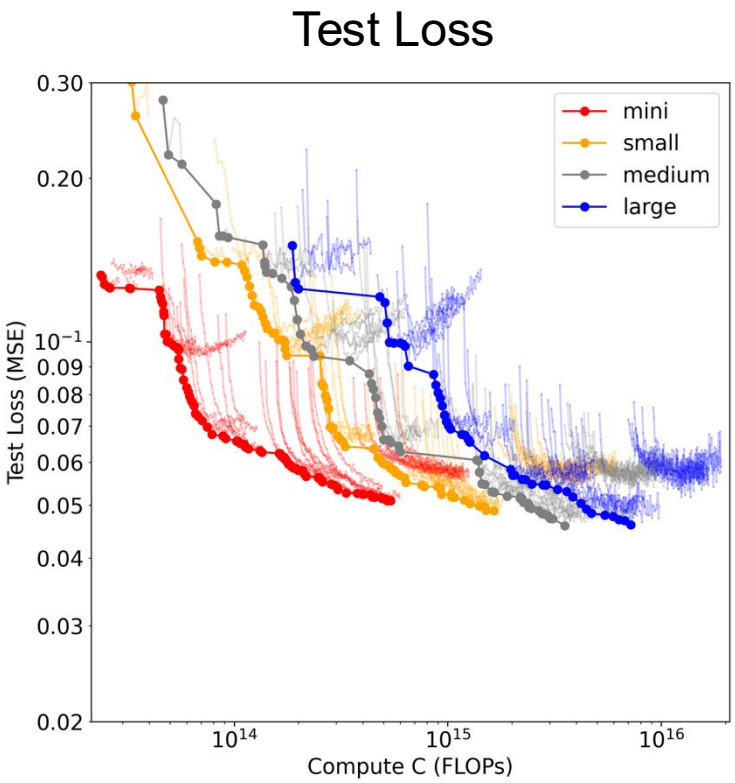


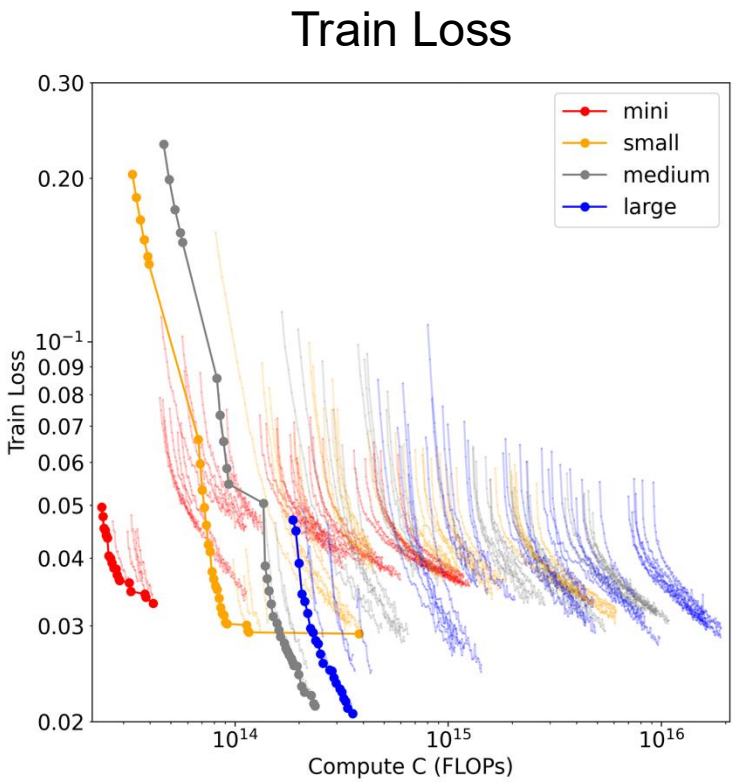
Test and Train Loss History





Test and Train Loss History

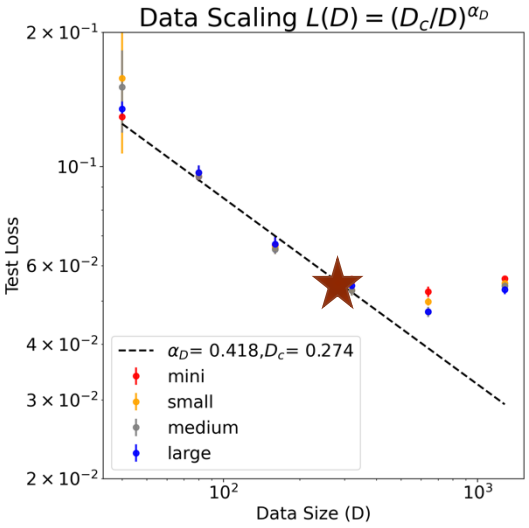




aerospacedesignlab

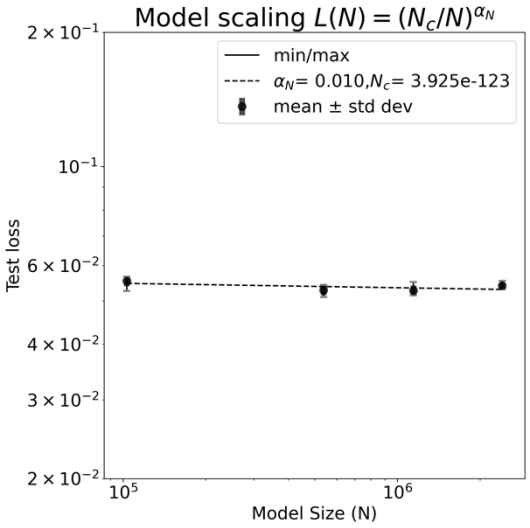
Data and Model Size Scaling (Best Test Loss)

Data Scaling (D)



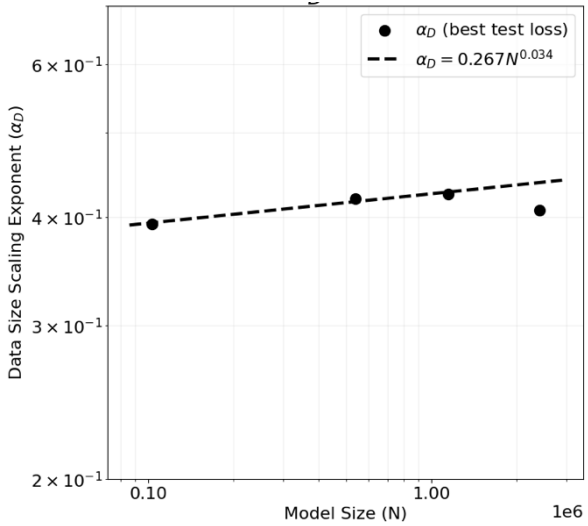
$$L(D) \sim \left(\frac{0.27}{D}\right)^{0.42}$$

Model Scaling (N)



$$L(N) \sim \epsilon$$

α_D vs. N



Larger models use data more efficiently

$$\text{Compute} \propto N\, D^{1.33}$$

Section Summary

1. Best Test Loss improves mainly with data, not model size:
   - More data is not always better; information quality matters.

2. Increasing GNN size alone does not guarantee lower test loss.
   - Further gains require better architecture and feature design.

3. Recommended design-space sampling density:
   - 8 samples per design variable.

III
Conceptual Design Integration

I
Multi-fidelity Field Surrogate Modeling

II
Dataset & Scaling

III
Conceptual Design Integration

The Optimization Problem Definition

- Showcases the fidelity gains of VortexNet in the conceptual design workflow.

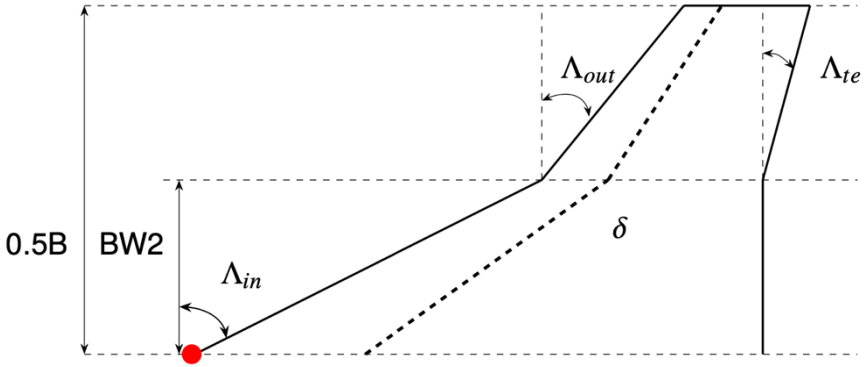
- **Objective**:

$$\textbf{minimize} \quad C_{D,\text{sup}}(\boldsymbol{x})$$

$$\textbf{s.t.}$$

$$C_L(\boldsymbol{x}) = C_L^*$$
$$|C_M(\boldsymbol{x})| \leq C_M^*$$
$$K_n(\boldsymbol{x}) \geq K_n^*$$

static margin $K_n = -\frac{C_{M,\alpha}}{C_{L,\alpha}}$ [20]



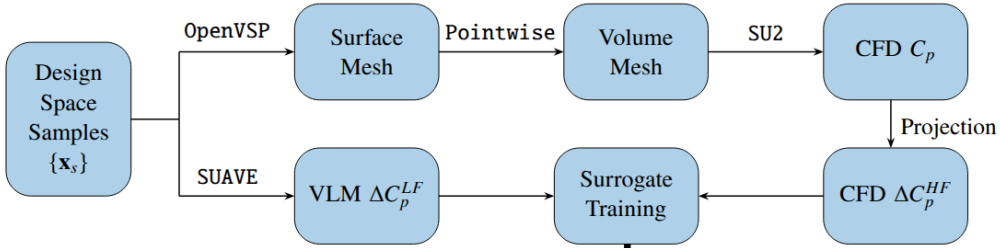
Conditions	Ma	C_L^*	C_M^*	K_n^*
supersonic	1.8	0.2	0.06	-
subsonic	0.3	0.7	0.06	5%

[20] Seraj and Martins, "Aerodynamic Shape Optimization of a Supersonic Transport Considering Low-Speed Stability," 2022

Aerodynamic Shape Optimization Workflow

Two-phase workflow:

- Offline: train VortexNet on paired low- and high-fidelity snapshots using the 32 geometries from the Lv.3 dataset.
- Online: optimizer-driven ASO with VortexNet augmentation.

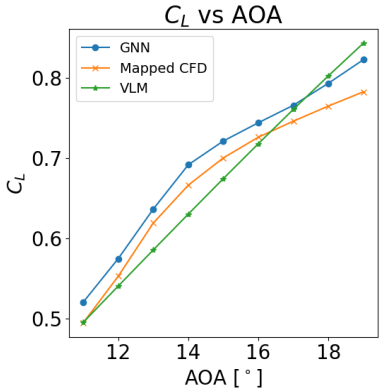


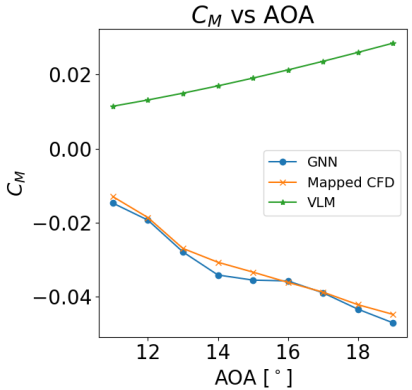
aerospacedesignlab

[Offline] Trained Surrogate Validation: Test Accuracy

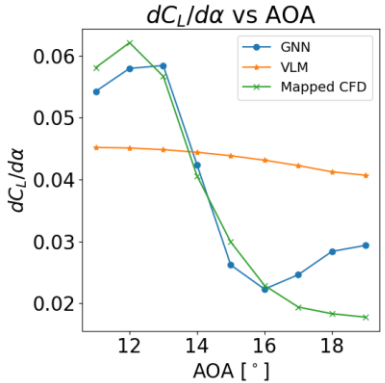
- How do the MF results compare with the LF and HF results?

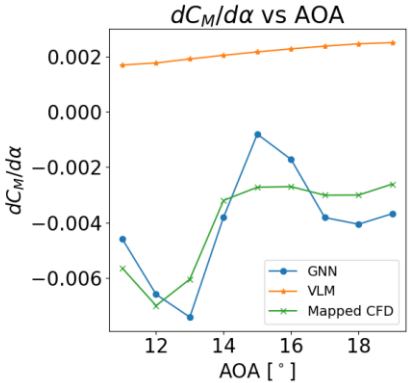
Configuration: 191301-3JT



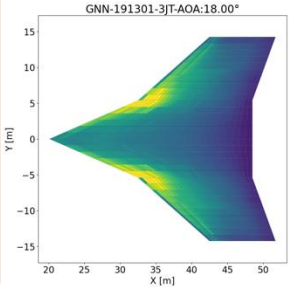


Configuration: 191301-3JT

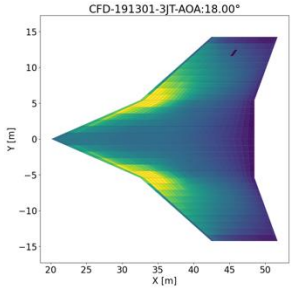




VortexNet



CFD

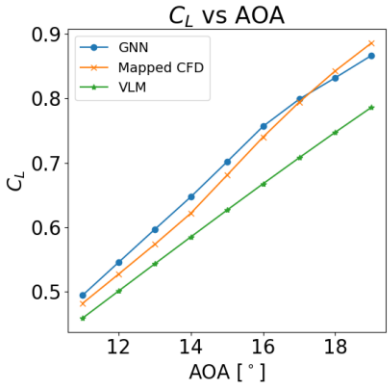


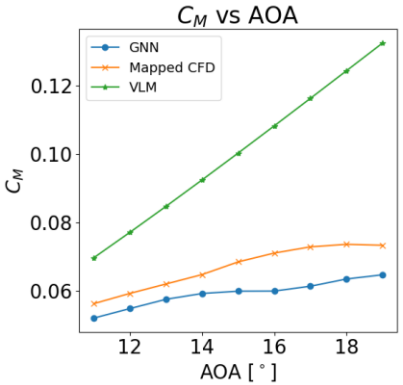
aerospacedesignlab

[Offline] Trained Surrogate Validation: Test Accuracy

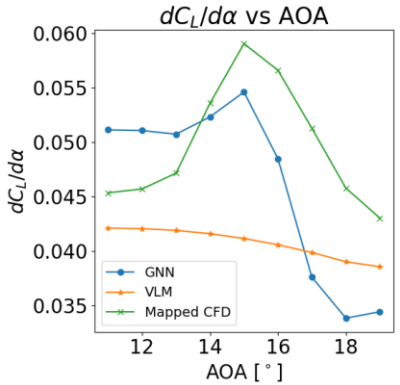
- How do the MF results compare with the LF and HF results?

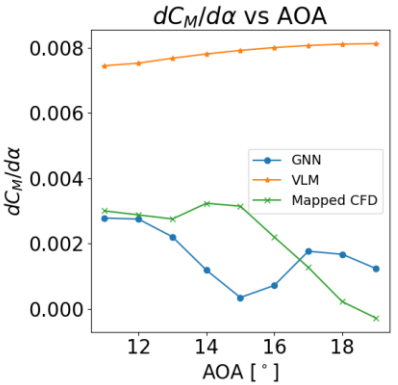
Configuration: 031600-5NR



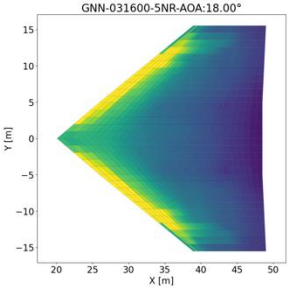


Configuration: 031600-5NR

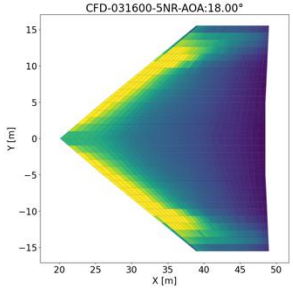




VortexNet



CFD

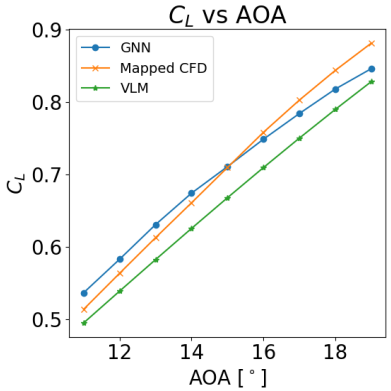


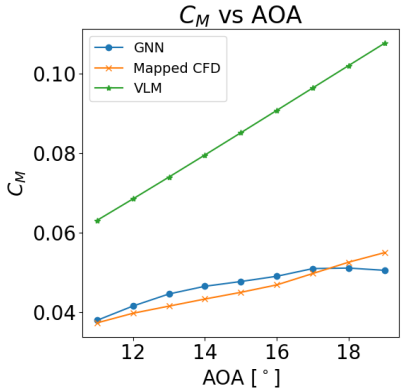
aerospacedesignlab

[Offline] Trained Surrogate Validation: Test Accuracy

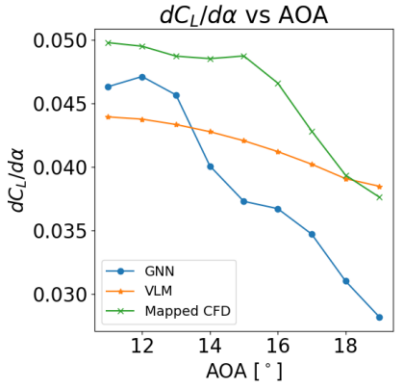
- How do the MF results compare with the LF and HF results?

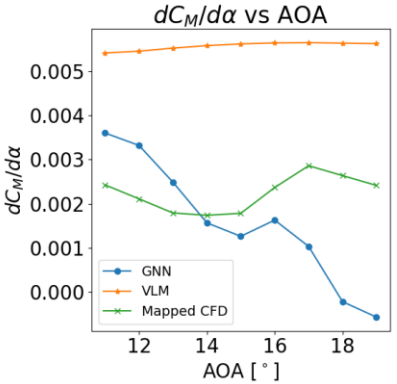
Configuration: 09230100EA



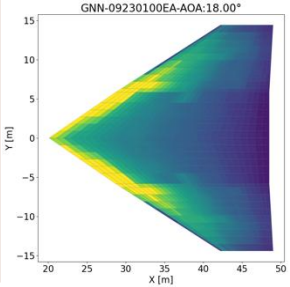


Configuration: 09230100EA

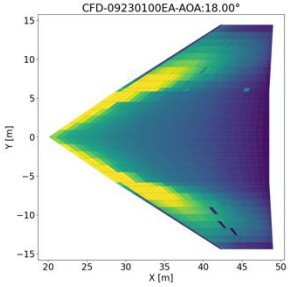




VortexNet



CFD

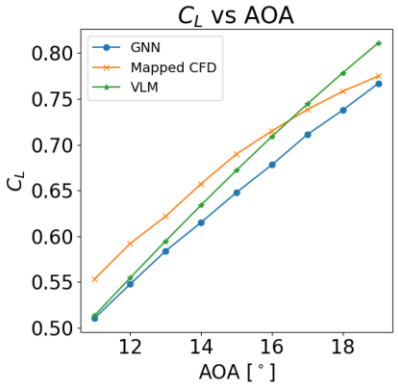


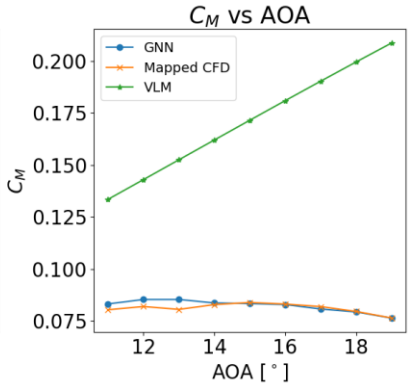
aerospacedesignlab

[Offline] Trained Surrogate Validation: Test Accuracy

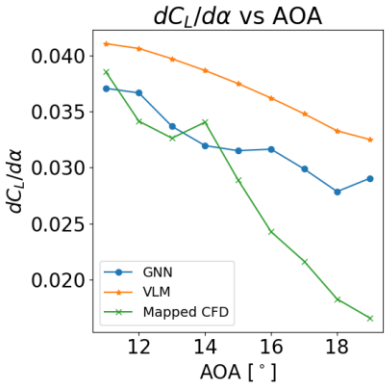
- How do the MF results compare with the LF and HF results?

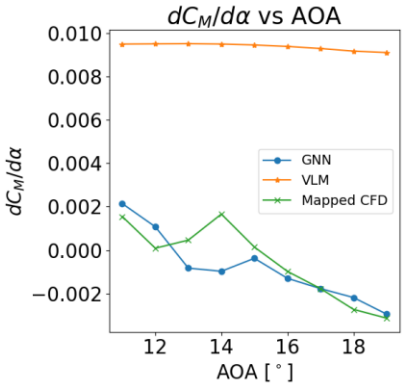
Configuration: 11050006QW



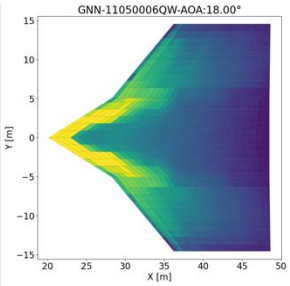


Configuration: 11050006QW

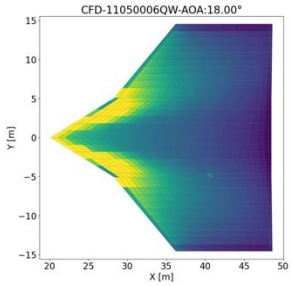




VortexNet



CFD



Trained Surrogate Validation: Test Accuracy Summary

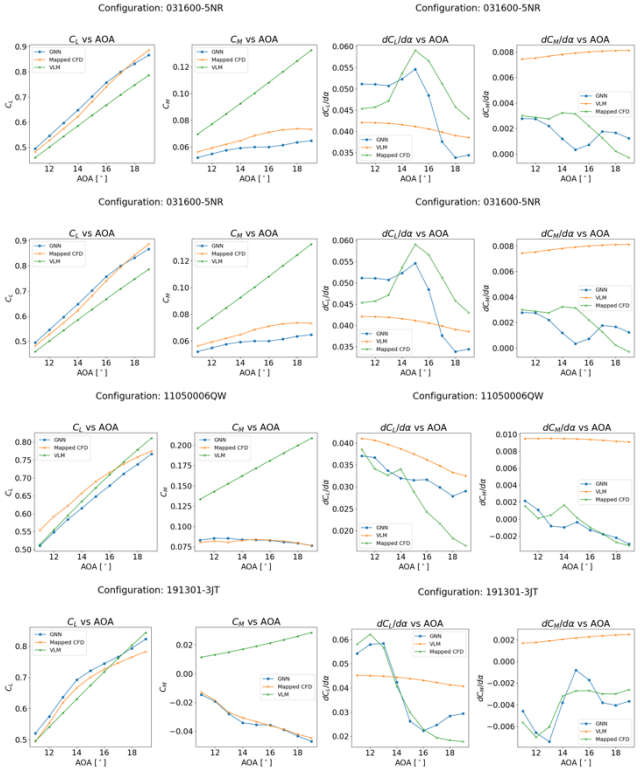
Compared with VLM, VortexNet provides **more accurate estimates** of C_L, C_M, $C_{L,\alpha}$, and $C_{M,\alpha}$.

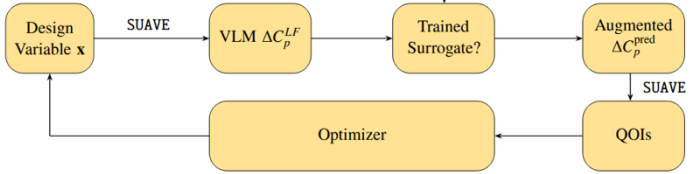
- Accurate K_n estimation!

Error comparison:

Methods	C_L NRMSE	C_M NRMSE
GNN-based	0.0535	0.2093
VLM-based	0.0739	1.6019

- $1.4 \times C_L$ error
- $7.7 \times C_M$ error





Online Phase

Optimizer:

- Particle Swarm Optimization (PSO) with 30 particles and 75 iterations.
- Constraints incorporated through a penalty method.

We ran two optimization cases: GNN-guided and VLM-based.

GNN-guided:

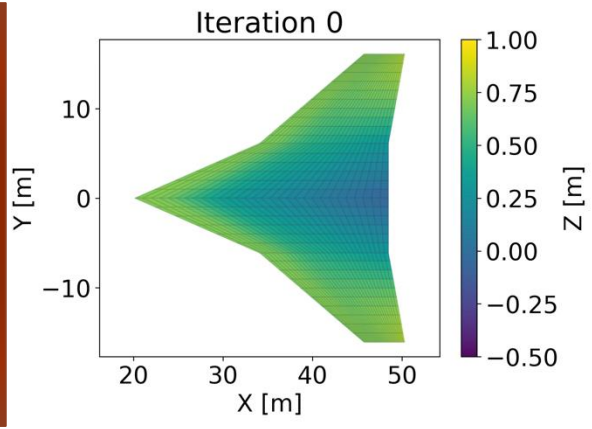
- **Subsonic QOIs** ($C_{L,\text{sub}}, C_{M,\text{sub}}, K_n$) are evaluated from VortexNet **augmented ΔC_p**.
- Supersonic QOIs ($C_{L,\text{sup}}, C_{M,\text{sup}}, C_{D,\text{sup}}$) are evaluated by **VLM**.

VLM-based:

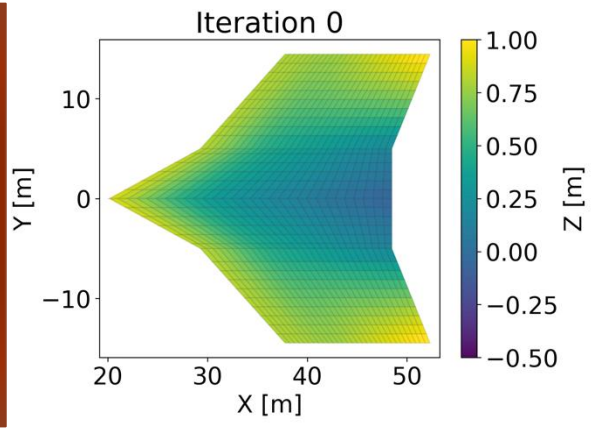
- **All** QOIs (supersonic and subsonic) are evaluated by **VLM only**.

Optimization Results

VLM-Based



GNN-Guided

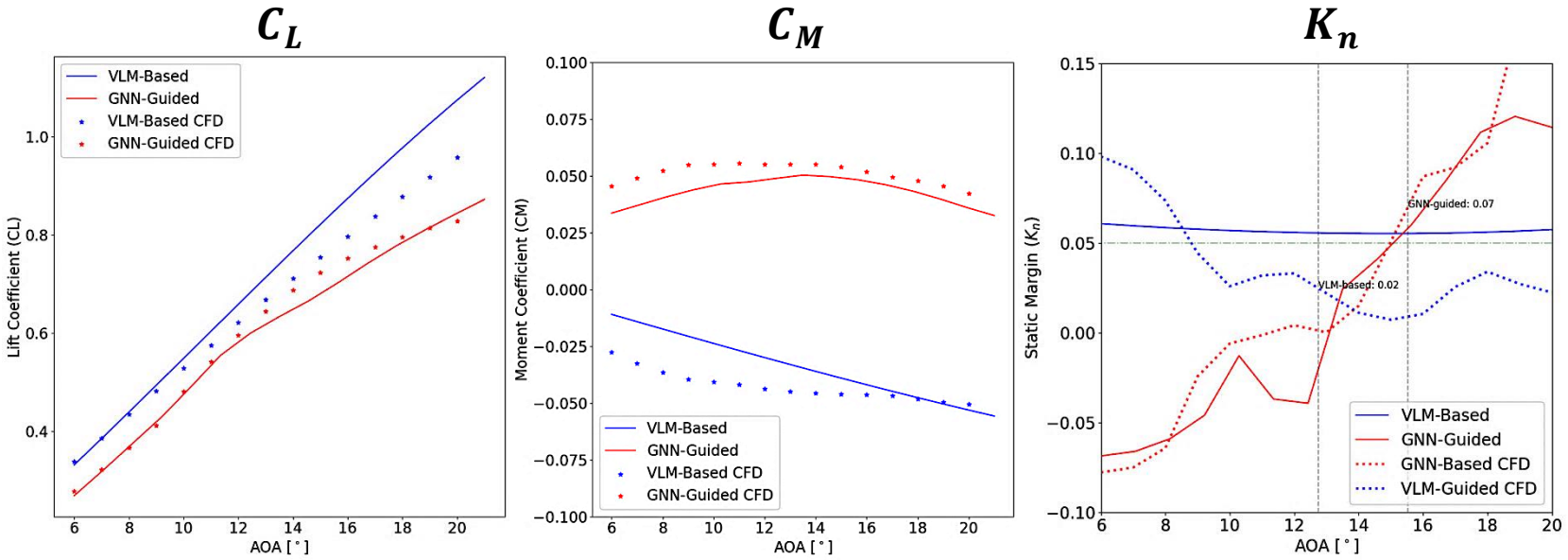


Metric	VLM-Based	GNN-Guided
Constraints		
Supersonic C_L	0.2000	0.1986
Supersonic C_M	-0.0291	-0.0059
Subsonic C_L	0.7002	0.6983
Subsonic C_M	-0.0323	0.0485
Subsonic K_n	0.0555	0.0576
Design Variables		
Subsonic AOA	12.7539°	15.5215°
Supersonic AOA	3.3393°	3.9636°
Objectives		
Subsonic C_D	0.1582	0.2046
Supersonic C_D	**0.0144**	**0.0174**

20.8% increase

Optimized Configurations Validation

CFD validation:



- At the corresponding operation conditions, VLM-Based design **failed** to meet the K_n requirement, while VortexNet-Guided **exceeded**.

Section Summary

Integrated VortexNet into the SUAVE conceptual design environment.

- "Drop-in" augmentation with minimal workflow disruption.

Two-point, stability-constrained optimization case shows the GNN-guided workflow produces a configuration that meets both **stability and aerodynamic targets**, while the **VLM-only workflow fails**.

- Increased fidelity accurately guides the search direction.
- Future work is required to incorporate modeling error into the optimization framework.

Summary and Future Directions

Summary

I: Multi-fidelity Field Surrogate Modeling

1. Developed VortexNet, a multi-fidelity field surrogate for conceptual design.
2. Delivers significant fidelity gains by embedding HF field information into LF analysis.

II: Dataset and Scaling

1. Identified VortexNet scaling laws: data matters more than model size; revealed GNN scaling limitations.
2. Created an open-source MF benchmark for low-speed SST aerodynamics.

III: Conceptual Design Integration

1. Integrated VortexNet into a conceptual design environment (SUAVE).
2. Showed that the added fidelity improves confidence in stability-constrained conceptual design.

GNNs can practically **bridge the fidelity-cost gap**, enabling higher **conceptual design confidence**, especially for concepts with limited historical data.

Future Directions

1. Improve model robustness and generalization
   - Strengthening feature encoding and graph construction to better handle unseen geometries.
2. Expand the framework beyond wing aerodynamics
   - Support full-vehicle conceptual design.
   - Incorporate additional field variables and additional fidelity levels.
3. Incorporate uncertainty quantification
   - Estimate prediction uncertainty, guide adaptive sampling, and better support engineering decision-making.

aerospacedesignlab

Funding and Computational Resources







Google Cloud





Stanford Data Science

This author was supported by Robert Bosch Research, the Air Force Office of Scientific Research under Grant FA9550-22-1-0004, and NASA, Award 80NSSC19K1661, under the Commercial Supersonics Technology (CST) program, Supersonic Configurations at Low Speeds (SCALOS), with Sarah Langston as the NASA technical grant monitor. Additionally, the author would like to thank Stanford Research Computing and Stanford Data Science for providing computational resources on the Sherlock and Marlowe clusters, as well as the Google Cloud Research Credit Program for granting access to GPU resources on the Google Cloud Platform.

THANK YOU!

GitHub: MF-VortexNet



double-delta-aero dataset

