Graph Neural Network-Guided Aerodynamic Shape Optimization for Conceptual Design of Supersonic Transport Wings

Yiren Shen* and Juan J. Alonso†
Stanford University, Stanford, CA 94305, USA

Low-speed stability constraints have historically been difficult to incorporate during the conceptual design phase due to the limited accuracy of low-fidelity simulation methods. This paper presents a graph neural network (GNN)-guided multi-fidelity optimization framework for the aerodynamic shape design of supersonic transport wings at the conceptual level. A GNN-based field prediction surrogate model is integrated into the SUAVE aircraft design environment, enabling real-time, high-fidelity-informed aerodynamic shape optimization across a defined design space. Using a two-point, stability-constrained shape optimization problem, we demonstrate that the GNN-guided workflow produces a design that satisfies both stability and aerodynamic performance requirements, while the traditional vortex lattice method (VLM)-based workflow results in a configuration that violates the intended design objectives. The proposed approach improves the fidelity of aerodynamic analysis and enhances confidence in conceptual design, increasing the likelihood that the resulting configuration will perform well in later design phases.

Nomenclature

C_D	=	drag coefficient	ΔC_p	=	local loading coefficient (DCP)
C_L	=	lift coefficient	K_n	=	static margin
C_M	=	pitching moment coefficient	$[x, y, z]$	=	Cartesian coordinate
$C_{L,\alpha}$	=	lift coefficient slope	$\mathbf{x}$	=	design variables
$C_{M,\alpha}$	=	pitching moment coefficient slope	y^+	=	dimensionless grid-to-wall spacing

I. Introduction

The demand for faster premium air travel in commercial passenger and cargo transportation called for renewed interests in supersonic transport (SST) [1]. Over the last two decades, significant research has focused on SST concept development, sonic-boom mitigation, and numerical methods [2–7]. These concepts tend toward a more slender aerodynamic shape, necessitating new trade-off studies among supersonic performance, sonic-boom mitigation, handling quality, and low-speed characteristics [8]. To address these challenges, the Supersonic Configurations at Low Speeds (SCALOS) project was initiated [9]. Throughout SCALOS, extensive efforts have produced aerodynamic and aeroelastic multidisciplinary design optimization (MDO) tools, flight-control co-design methodologies, wind-tunnel testing techniques, experimental datasets for diverse configurations, noise-prediction simulations, surrogate-modeling approaches, and the development of the reference supersonic airplane concept (RSAC) [5, 8, 10–20].

Among these contributions, the studies by Seraj and Martins [5, 13, 14] are of particular interest. Seraj and Martins validated the use of a Reynolds-averaged Navier-Stokes (RANS)-based computational fluid dynamics (CFD) solver to quantify lift (C_L), drag (C_D), and pitching-moment (C_M) coefficients for slender delta-wing configurations [14]. They then formulated an MDO problem in `MACH-Aero` [21] to perform free-form deformation (FFD)-based shape optimization of the wing under both subsonic (landing/take-off) and supersonic (cruise) conditions, subject to longitudinal stability constraints [13]. Their results reveals the trade-off between enforcing stability constraints and reducing cruise drag: enforcing a 5% static margin at low speed incurs a 5.8% drag penalty in supersonic cruise. Separately, Seraj and Martins

*Ph.D. Candidate, Department of Aeronautics and Astronautics, AIAA Student Member.
†Vance D. and Arlene C. Coffman Professor, Department of Aeronautics and Astronautics, AIAA Fellow.

embedded trim sizing into a drag-minimization optimization of the main wing and control surfaces, highlighting the necessity of incorporating stability constraints and component-to-component interactions in aerodynamic design [5].

These findings underscore the importance of including stability and control (S&C) constraints in the conceptual design of SCALOS-like configurations. During conceptual design, engineers select the overall aircraft configuration, sizing components, designing main wing, arranging systems, and estimating weights, to produce preliminary performance metrics that satisfy mission and business objectives, thereby laying the groundwork for a more detailed design, analysis, and optimization in the subsequent phases [22]. A broad design space, including different wing layouts, is surveyed using rapid analysis methods including comparative sizing based on existing configurations, empirical formulations, and low-fidelity tools such as the vortex lattice method (VLM), to eliminate concepts that do not meet program requirements [22, 23]. For SCALOS-like configurations, such incorporation of stability constraints is especially important, as indicated by Seraj and Martins, to reduce the risk of undesirable S&C behaviors carrying over into the later phases of vehicle development, thus mitigating the risk of program overhaul.

However, incorporating S&C constraints at conceptual design stage remains challenging. Typical conceptual-design suites like `SUAVE` [24] rely on low-fidelity methods such as VLM [25], which solve inviscid potential flow and cannot capture complex nonlinear phenomena such as vortex lift, vortex–boundary interactions, and flow separation, that critically influence stability [26]. Although theoretical and semi-empirical corrections exist [27], and multi-fidelity (MF) adjustments can incorporate experimental data to improve VLM predictions [28], these corrections are platform dependent and require extensive tuning across a large design space.

High-fidelity methods, such as CFD, can accurately model the fluid dynamics governing S&C behavior but remain impractical for broad conceptual surveys. Even with advances in high-performance computing (including GPU acceleration), large-scale CFD studies are computationally expensive. Moreover, as noted by Seraj and Martins [5], the mesh deformations induced by large shape morphing can cause numerical divergence in RANS simulations, further limiting the accessible design space that can be examined by CFD. Consequently, relying solely on high-fidelity CFD for aerodynamic shape optimization during conceptual design is undesirable.

To bridge the fidelity gap and increase design confidence during conceptual design, surrogate models are often employed [15, 29]. A multi-fidelity surrogate learns the discrepancy between different fidelity sources during offline training using a simplified mathematical framework, enabling rapid evaluations in the online phase of aerodynamic shape optimization with an accuracy comparable to the high-fidelity sources. However, most surrogate models are limited to predicting a fixed set of quantity of interestss (QOIs) defined *a priori* and cannot reproduce high-dimensional data such as full pressure-field distributions. In aircraft MDO, the ability to predict field data is highly desirable, as it allows arbitrary QOIs to be computed *a posteriori*. Furthermore, field data contain rich information for designers to analyze and are essential for multidisciplinary coupling.

To enable multi-fidelity field-prediction surrogate modeling, Shen et al. developed a graph-neural-network–based multifidelity surrogate named `VortexNet`, which maps low-fidelity VLM lattice-wise pressure loadings to a projected high-fidelity RANS CFD surface pressures, thereby correcting viscous, separation, and nonlinear vortex effects missing from the VLM model [16]. `VortexNet` represents panel-wise local pressure loadings, geometric information, and lattice connectivity as a graph, and uses message-passing layers to propagate local flow–geometry interactions and predict quasi-CFD–level loading coefficients. This surrogate demonstrates strong geometric generalizability across the design space and achieves significantly better prediction accuracy than VLM results, making it well suited for conceptual design [16, 17]. Moreover, `VortexNet` integrates seamlessly into the `SUAVE` multidisciplinary optimization workflow, enhancing design fidelity with negligible additional online cost.

Building on this integrated workflow, this paper presents a stability-constrained supersonic transport wing aerodynamic shape optimization study with a conceptual design focus. The remainder of the paper is organized as follows: Section II defines the optimization problem; Section III describes components used in the aerodynamic shape optimization framework; and Section IV reports the corresponding design results.

II. Optimization Problem

A. Vehicle Definition

The baseline vehicle in this study is a simplified version of the RSAC described by Wiersema et al. [20]. The main wing of the RSAC, designated as RW23, is a cranked delta wing with an inboard leading-edge sweep of 68° and an outboard sweep of 41°. Several modifications were applied to simplify the geometry. First, the canard and dorsal fin at the tail were removed. Second, the RW23 twist distribution was replaced: instead of the complex profile shown in

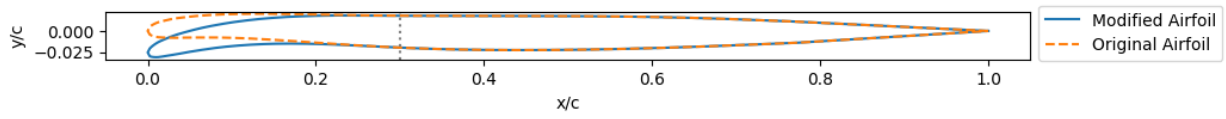


Fig. 1 Illustration of the RW23 root airfoil (original airfoil) and the resultant airfoil after a negative droop angle modification (modified airfoil).

Wiersema et al.'s Figure 2b, a constant 1.5° twist is applied to the inboard section, and a linear washout twist blend from 1.5° at section break to −1.5° at the tip is applied to the outboard section of the wing [20]. Third, the spanwise airfoil distribution was homogenized by applying a modified root airfoil shape across the entire span. Specifically, the first 30% of the chord of the root airfoil is rotated downward or upward to introduce a "droop" angle. To achieve this, the original camber line is interpolated with a cubic spline, then rotated by the prescribed droop angle and fitted with a Bézier curve to enforce first- and second-order continuity at the 30% break chord where the rotation is originated. Then the original thickness distribution, also interpolated via cubic splines, is added back to the modified camber line to reconstruct the upper and lower surfaces. Figure 1 shows the original RW23 root airfoil (dashed) and the modified shape with a −5° droop angle.

In addition to the "droop angle" design parameter, the aerodynamic shape optimization includes additional parameters that define the main wing shape. Figure 2 presents a schematic of half the wing platform, with design variables indicated by double arrows. Five parameters are shown:

- B: wing span, measured in inches;
- BW2: fractional location of the inboard–to–outboard break chord relative to the semi-span;
- SW1 and SW2: inboard and outboard leading-edge sweep angles (positive for aft-sweep), measured in degrees;
- SR2: outboard trailing-edge sweep angle (negative for aft-sweep), measured in degrees.

The inboard trailing edge remains at zero sweep. The red circle in Figure 2 marks the location of wing apex, which is fixed at $[x, y, z] = [20.12, 0.0, 0.66]$ meters from the nose datum.

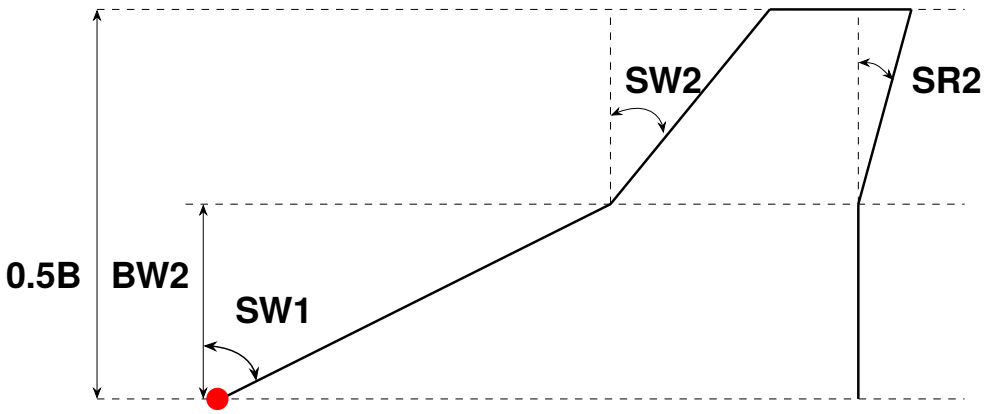


Fig. 2 Schematic drawing of the wing platform and its design variables.

Throughout aerodynamic shape optimization, these design variables $\mathbf{x} = [\text{Droop Angle}, \text{SW1}, \text{SW2}, \text{SR2}, \text{BW2}, \text{B}]$ are allowed to change and thus alter the configuration. As a result, the center of gravity (CG) of a vehicle needs to be recalculated. To compute CGs, we hold the fuselage, tail, engine, systems and vertical stabilizer locations and weights fixed, so any CG shift arises solely from changes to the wing platform. To model the effects of the change of wing platform to overall CG, we use the RSAC weight breakdown and compute moment balances about the nose datum, as shown in Table 1.

In Table 1, the specific weights for aerodynamic surfaces are taken from Welge et al. and Wiersema et al., with centroid locations as defined in `SUAVE` [2, 20, 24]. The remaining system weights and centroids follow the RSAC component-weight breakdown. For fuel and payload, we assume at approach and landing stage, the fuel mass onboard is 30% of the total capacity and is stored primarily in forward-body tanks. Payload is taken at 50% capacity, evenly distributed across the fuselage. The centroid location estimate is based on the fuel tanks location and evenly-distributed

Component Name	Weights [lbs]	Centroid Location-x from Datum [m]
Wing	10 lbs/ft^2	@0.4mean aerodynamic chord (MAC)
Horizontal Tail	5 lbs/ft^2	@0.3MAC
Vertical Tail	5 lbs/ft^2	@0.3MAC
Propulsion System	44540	42.40
Fuselage	29120	32.65
Systems and Fixed Equipments	43152	34.36
Fuel and Payload	47468	33.30

Table 1 Mass properties of system components used for center of gravity computation.

payload loading condition.

B. Optimization Problem

The optimization problem investigates how subsonic stability constraints affect the conceptual design of a wing platform optimized for minimum supersonic drag coefficient, $C_{D,\text{sup}}$. We adopt a two-point optimization case, following the example of Seraj and Martins, in which the wing is evaluated under both a supersonic cruise condition and a subsonic takeoff–landing condition [13]. The flight conditions, along with the target lift coefficients (C_L^*) and static margins (K_n^*), are listed in Table 2. The static margin is defined as:

$$K_n = -\frac{C_{M,\alpha}}{C_{L,\alpha}}, \tag{1}$$

where $C_{M,\alpha}$ and $C_{L,\alpha}$ are the moment and lift coefficient slope and can be approximated by finite difference by perturbing the angle-of-attack (AOA) by a small amount. In the current study, a perturbation of $O(10^{-2})$ degree is used for this finite difference approximation.

Since we optimize only the wing platform, imposing a trim condition on the wing results in a highly constrained problem with poor convergence. Instead, we bound C_M so that the wing's pitching moment remains small enough to be compensated by other trim surfaces, with minimal impact on C_L.

	Mach Number	Altitude [m]	Reynolds Number	C_L^*	C_M^*	K_n^*
Supersonic	1.8	16764	1.01×10^8	0.2	$-0.06 \leq C_M^* \leq 0.0$	-
Subsonic	0.3	500	8.04×10^7	0.7	$-0.06 \leq C_M^* \leq 0.0$	$\geq 5\%$

Table 2 Flight conditions at the supersonic and subsonic design points.

We then run the optimizer, as described in Section III.D, to minimize $C_{D,sup}$ subject to constraints on supersonic lift target ($C^*_{L,sup}$), supersonic moment bounds ($C^*_{M,sup}$), subsonic lift target ($C^*_{L,sub}$), subsonic moment bounds ($C^*_{M,sub}$), and static margin bounds (K_n^*). Eight design variables are optimized: six geometric parameters (Section II.A) and two angles of attack at subsonic and supersonic conditions AOA_{sub} and AOA_{sup}. Their bounds are given in Table 3.

III. Methodology

A. Flow Solvers

Two solvers, VLM and RANS CFD, are used to compute aerodynamics. Low-fidelity simulations are preformed with VLM. The specific VLM solver being used is shipped from `SUAVE`, which is a Python implementation of `VORLAX` [25] that approximates compressibility-corrected, inviscid, attached flows using a horseshoe-vortex panel method. Panels on an infinitesimally thin representation of the geometry carry bound and trailing vortices, whose strengths are solved via a linear system enforcing surface-tangency conditions. The resulting lattice vortex strength are converted into local loading coefficient (ΔC_p^{LF}), being the pressure difference coefficient (DCP) between the lower and upper wing

Design Variable	Upper	Lower	Scaling
Droop Angle	7.0°	−13.0°	8.0
B	1300 [inches]	1100 [inches]	1200
SW1	70°	60°	65
SW2	60°	38°	50
SR2	−5°	−30°	10
BW2	0.5	0.3	0.4
AOA_{sub}	21°	9°	15
AOA_{sup}	10°	0°	5

Table 3 List of design variables and their corresponding upper and lower bounds for the optimization problem.

surfaces. Aerodynamics QOIs, including lift (C_L), induced drag (C_D) and pitching moment (C_M) coefficients can be integrated across panel-wise ΔC_P^{LF}. Polhamus correction is applied to the leading edge of the wing. A more detailed formulation can be found at Shen and Alonso [17]. A 35×30 span-wise-by-chord-wise lattice discretization is used throughout this study. Even spatial distribution of the panels is used for both the span-wise and chord-wise directions.

High-fidelity simulations are performed with `SU2`'s [30] RANS solver using the Spalart–Allmaras turbulence model with rotation/curvature corrections [31]. The convective fluxes are discretized via the Jameson–Schmidt–Turkel (JST) scheme [32] with an adaptive Courant–Friedrichs–Lewy (CFL) number. The convergence criteria used is root-mean-square residual for the continuity equation drops below 10^{-6}, or a steady state value for C_D is found such that the change of C_D among the past 100 iterations (Cauchy convergence criterion) is below 5×10^{-5}. The geometry surface meshes are generated by `OpenVSP` [33], using a set of geometry parameters that balance the mesh sizes and accuracy. Specifically, a minimum edge length of 0.01 [m] and edge growth ratio of 1.2 are used. The volume meshes for CFD are then generated in `Pointwise` by importing the surface meshes generated. A script is used to maintain consistency in mesh quality among geometries. For boundary layer treatment, a 45-layers boundary layer extrusion with an height growth rate of 1.24 and a first layer height of 6×10^{-6} meters is used. These boundary layer parameters ensure consistent boundary layer resolution, such that $y^+ < 5$ over the majority of the wing. The far-field is defined as a sphere, with diameters roughly equals 40 MAC, being 630 [m]. Far-field boundary condition is applied to the far-field surface, and no-slip adiabatic walls boundary condition is applied on the vehicle. Once converged, the CFD pressure coefficients are projected onto the same 35×30 lattice used in VLM to produce the high-fidelity ΔC_P (DCP), ΔC_p^{HF}, for surrogate model training. This projection process is elaborated in Shen and Alonso [17].

B. Surrogate Modeling

The Graph Neural Network (GNN)-based MF surrogate model (`VortexNet`) developed by Shen et al. is used to map the VLM-computed DCP into its CFD counterpart [16]. This surrogate model is based on a Graph-U-Net structure, utilizing Graph Attention Network v2 (GATv2) layers [34] to perform message-passing rounds over a graph. Specifically, in `VortexNet`, the lattice grid used in VLM is treated as a graph: the control points at locations where the flow tangency condition is evaluated serve as graph nodes, and edge features capture spatial relationships via Euclidean distances among adjacent nodes to represent the spatial topology of the geometry under simulation [16]. To represent the flow and geometry state, each of these nodes is assigned a twelve-dimensional feature vector, including the DCP computed from VLM, the free-stream AOA, free-stream Mach number, free-stream Reynolds number, wing thickness evaluated at the control points, upper and lower surface curvatures, upper and lower surface chord-wise slopes, and the three-dimensional location $[x_c, y_c, z_c]$ of the control points. The addition of z_c expands the nodal feature dimension from eleven to twelve compared with the feature definition in Shen et al. [16], due to the richer camber profiles and twist distributions present in the current design space, where z_c improves the geometric representation.

Training of the surrogate model is conducted by first assembling a dataset containing diverse design samples within a prescribed design space. For each design sample at a prescribed flow condition, both VLM and CFD computations are performed to construct paired multi-fidelity result. Once the dataset is assembled, a customized k-fold cross-validation training scheme [16] is employed to train the surrogate model. A compound loss function is used, comprising a Smooth L1 loss, a sign-consistency penalty, and a physics-informed residual loss that quantifies the violation of the

surrogate model's field DCP prediction relative to the potential flow equation solved by VLM. This physical loss term is progressively annealed through the customized k-fold training scheme, shifting the training focus from enforcing potential flow's physical modeling, so that baseline pressure characteristics are accurately captured, to a more data-driven prediction that can represent nonlinear flow features.

Several hyper-parameters exist for the proposed surrogate model, and the tuning of these hyper-parameters is achieved through an automatic hyper-parameter tuning package named `Optuna` [35]. The resulting surrogate model demonstrates a strong capability to reconstruct nonlinear aerodynamic flow features while maintaining sufficient geometric generalizability, making it well-suited for augmenting the VLM aerodynamic computation in the conceptual design tasks discussed in Section II.

C. Multi-Fidelity Workflow

The overall workflow for `VortexNet` surrogate model training and GNN-guided aerodynamic shape optimization is shown in Figure 3. Two phases exist: in the offline phase, shown as blue-colored blocks in Figure 3, the surrogate model described in Section III.B is trained using a dataset of vehicle configurations ($\{\mathbf{x}_s\}$) sampled from a predefined design space. This design space is described in detail in Section IV.A, and the corresponding vehicles are constructed using `SUAVE`. Following vehicle definition, two fidelity levels branch out to conduct the low-fidelity (LF) and high-fidelity (HF) simulations, as discussed in Section III.A. The LF simulations use the VLM solver in `SUAVE`, while the HF simulations employ the `OpenVSP` API in `SUAVE` to build surface meshes, followed by volume mesh construction using `Pointwise`, and CFD computation using `SU2`. The final paired set of LF and HF field data is then used to train the `VortexNet` surrogate model.

Once trained, the surrogate model is ready to be deployed for online assessment, shown as gold-colored blocks in Figure 3. The thick black arrow connecting "Surrogate Training" and "Trained Surrogate" in Figure 3 highlights this transition. For online assessment, the design variable $\mathbf{x}$ is fed into `SUAVE`'s vehicle definition. Subsequently, an LF VLM simulation is conducted, which serves as the input to the surrogate model. When surrogate model augmentation is used, the surrogate model then enhances the VLM-predicted ΔC_p^{LF} to a quasi-HF field prediction ΔC_p^{pred}. This augmented DCP is then fed back into `SUAVE` for QOIs evaluation. When the surrogate model is not used, the VLM-predicted ΔC_p^{LF} is used for QOIs evaluation. It is worth noting that surrogate augmentation is applied only for the subsonic operating condition; all QOIs evaluations at the supersonic condition rely solely on VLM. These QOIs are then utilized by an optimizer, as discussed in Section III.D, to drive the convergence of the aerodynamic shape optimization.

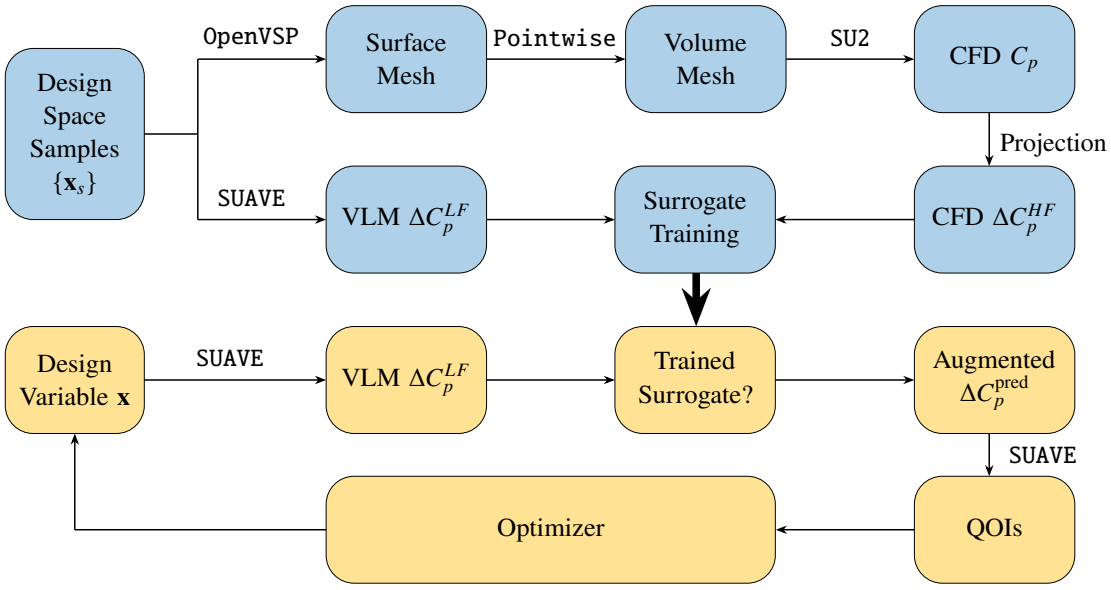


Fig. 3 Block diagram of the surrogate model training and the GNN-guided aerodynamic shape optimization workflow. Each block represents an intermediate stage in the design workflow, with "online" stages colored gold and "offline" stages colored blue, while the tools used for each specific process are annotated alongside the arrows.

D. Optimizer

A variety of solver choices can be used for the optimization problem discussed in Section II. In prior work, Seraj and Martins employed the gradient-based optimizer `SNOPT` [36], leveraging the discrete-adjoint capability of `ADflow` to obtain sensitivities of the free-form deformation (FFD) design variables [13]. However, the multifidelity workflow described in Section III.C is not adjoint-capable, so gradients must be approximated via finite differences. Constructing the Jacobian requires perturbing each of the eight design variables, as listed in Section II, and evaluating the objective and all constraints, rendering the cost of each optimization iteration relatively high. Moreover, the surrogate model (`VortexNet`) was trained only to predict the pressure field, without enforcing accurate design parameter to pressure field gradients. Previous validation shows normalized root mean square error (NRMSE) of 4.1% and 1.9% for C_L and C_M respectively [17]. Considering the small step size that will be used in $C_{M,\alpha}$ and $C_{L,\alpha}$ computation, the finite-difference gradients computed within these error bounds may be too noisy to guide a gradient-based optimizer reliably.

Hence, instead of relying on gradient information, zero-order algorithms are adopted for the optimization problem described in Section II. These methods explore the design space by tracking and updating a population of candidate solutions, eliminating the need for gradients computation. The stochastic nature, together with the board sampling space, avoids optimizer becoming trapped in local minima [37]. In this work, we select particle swarm optimization (PSO) [38], a stochastic, population-based method inspired by "swarm intelligence" [39]. In PSO, each candidate adjusts its position based on its current trajectory, its own best-found solution, and the swarm's global best, enabling efficient search without gradient evaluations. PSO has been successfully applied to aerodynamic shape optimization in previous studies [40], and its characteristics make it well suited to our problem's dimensionality and search-space requirements.

For a design point (particle) position $\mathbf{x_i} \in \mathbf{R}^n$, the following velocity v_i and position update is used:

$$\begin{aligned} v_i^{t+1} &= w v_i^t + c_1 r_1 (p_i^t - \mathbf{x_i}^t) + c_2 r_2 (g^t - \mathbf{x_i}^t), \\ \mathbf{x_i}^{t+1} &= \mathbf{x_i}^t + v_i^{t+1}, \end{aligned} \tag{2}$$

where $\mathbf{x_i}^t$ is the current particle position, $\mathbf{x_i}^{t+1}$ is the position after 1 step update, v_i^t is the current particle velocity, and v_i^{t+1} is the velocity after 1 step update; the parameters r_1 and r_2 are random numbers drawn from the uniform distribution $\mathcal{U}(0, 1)$; the coefficient w is the inertia coefficient, regulating how similar the current velocity to the velocity in the previous iteration; p_i is particle's best location, and g is the global best location; the coefficients c_1 and c_2 are the "memory" and "social" coefficients that pull particle towards its historical best and the global best identified. The initial coefficients are set to $w_{start} = 0.8$, $c_{1,start} = 1.5$ and $c_{2,start} = 0.2$ to allow each particle to explore the design space more independently. Throughout the optimization iterations, the following coefficient scheduling are used:

$$\begin{aligned} w &= w_{start} - \frac{t}{T}(w_{start} - w_{end}); \\ c_1 &= c_{1,start} - \frac{t}{T}(c_{1,start} - c_{1,end}); \\ c_2 &= c_{2,start} - \frac{t}{T}(c_{2,start} - c_{2,end}), \end{aligned} \tag{3}$$

where t is the current iteration index, and T is the prescribed total steps; w_{end}, $c_{1,end}$, and $c_{2,end}$ are 0.2, 0.5, 3.0 respectively. The linear scheduling of PSO coefficients encourages the balance between "exploration" and "exploitation" of the optimizer through iterations.

As PSO is designed for unconstrained optimization problems, constraints needs to be incorporated into the optimizer through penalty method. In addition to the supersonic drag coefficient $C_{D,sup}$, the total objective function that PSO optimizes is defined as:

$$f_{total} = f_{obj} + P_{total}, \tag{4}$$

where f_{obj} is a normalized supersonic drag count, defined as $f_{obj} = C_{D,sup}/0.05$; P_{total} is the total penalty term that captures the deviation of a current design point to the constraints bounds, and is computed as:

$$P_{total} = \sum_i w_t \cdot P_i \tag{5}$$

where w_t is an adaptive weight that decreases as the iterations progress, relaxing the constraints when the particles are sufficiently converged and enabling the optimizer to shift its focus from identifying feasible regions to exploiting the current configuration for supersonic drag reduction:

$$w_t = \begin{cases} w_0 & \text{if } \frac{t}{T} \leq 0.2; \\ w_0(1 - (\frac{t}{T} - 0.2))) & \text{if } \frac{t}{T} > 0.2, \end{cases} \tag{6}$$

where $w_0 = 5.0$.

Individual constraint penalties P_i are calculated as follows:

$$P_{C_L} = 1.6 w_t \cdot \left| \frac{C_L - C_L^*}{C_L^*} \right|; \tag{7}$$

$$P_{C_M} = 0.25 w_t \cdot \begin{cases} 1.5 \frac{C_M}{0.06} & \text{if } C_M \geq 0 \\ \frac{-0.06 - C_M}{0.06} & \text{if } C_M \leq -0.06 \\ 0 & \text{otherwise;} \end{cases} \tag{8}$$

$$P_{K_n} = 1.3 w_t \cdot \begin{cases} 2 \left| \frac{K_n - K_n^*}{K_n^*} \right| & \text{if } K_n < K_n^* \\ \left| \frac{K_n - K_n^*}{K_n^*} \right| & \text{otherwise,} \end{cases} \tag{9}$$

where, the superscript * denotes target (or reference) values. The scaling weights for each constraint are determined by trial and error to ensure that the relative penalty magnitudes remain comparable across all constraints.

Throughout the optimizations, particle may arrive at design points that produce unrealistic geometries, such that the tip chords of the wing are negative. In such cases, the particle is reset to a random location within the design space. In addition, if a particle's velocity would carry it beyond the design-space boundaries, boundary reflection is applied to reverse its velocity vector. In this work, the swarm comprises 30 particles and runs for 80 iterations.

IV. Results

A. Design Space Sampling

The design space used for the initial sampling is listed in Table 3. The term "initial design space sampling" refers to the process of generating samples of design variables drawn from the design space and constructing the corresponding vehicle configurations. These vehicle configurations are then used for offline surrogate model training, as discussed in Section III.C. Since the surrogate model is trained on these sampled configurations and inherently reflects the sampling bias in its predictions, the quality and methodology of the sampling process are critical. This results section is therefore provided to inform the reader of the sampling scheme, potential sampling bias, and the range of vehicle configurations used to train the surrogate model.

Within the design variable bounds, Saltelli's sampling [41] is used to draw samples from this six-dimensional design space. The sampling first construct Sobol-sequence matrices, each in dimension of 4-by-6, and then construct the swapped matrices by replacing columns between the Sobol-sequence matrices. This yields a total of $4 \times (2 + 6) = 32$ designs. The adoption of Saltelli's sampling enables estimation of first-order Sobol sensitivity indices, which results in a dataset that in the future can be used for global sensitivity analysis. In addition, the nested formulation for building Sobol-sequence matrices enables this dataset to be expanded to conduct future analyses without the need to resample the existing design points. The sampling scheme is designed to minimize the sampling clustering, and ensure a uniform coverage of the design space.

However, certain combinations of design variables produce nonphysical geometries. Specifically, negative tip chords may exist and must be excluded from the dataset. After removing these points, 27 valid design samples remain. Figure 4 shows these samples: off-diagonal panels display scatter plots of paired design parameters sampling locations, while diagonal panels show kernel-density estimated sampling density. The removal of nonphysical design points introduces bias in the sampling density. Specifically, at droop angles around 1° and spans near 1250 inches, lower sampling densities are observed. Additionally, the samples tend to be biased toward lower BW2 values. This bias may lead to a larger prediction errors for configurations in these regions, as the surrogate model has fewer training samples.

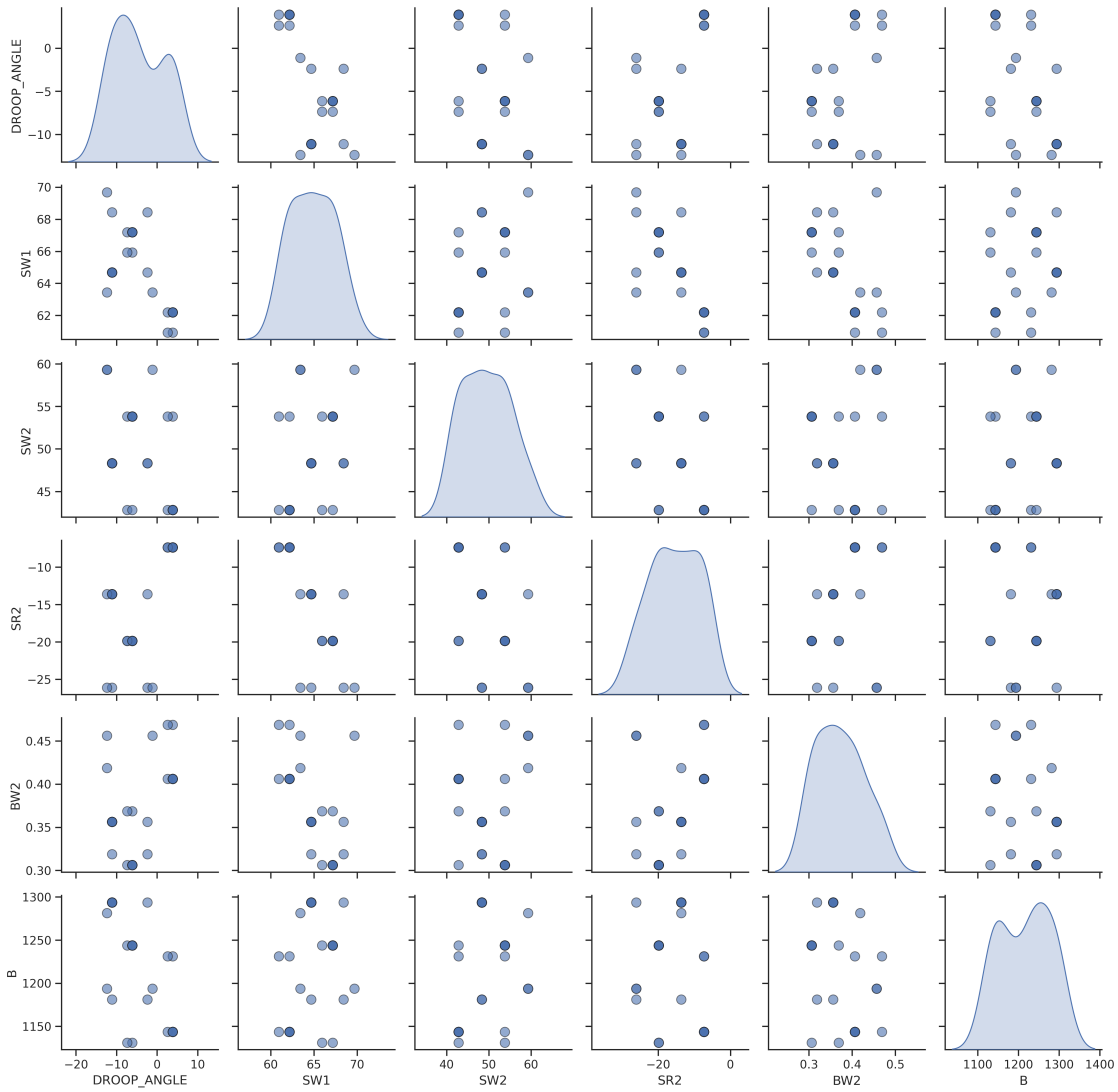


Fig. 4 Design space coverage and design variable distributions of the initial design space sampling dataset.

Among these 27 geometry samples, we further divide the dataset into two subsets. The "training geometries" set contains configurations used for surrogate model training, while the "test geometries" set is reserved for the geometric generalizability tests presented in Section IV.B and is not exposed to the model during training. The resulting configurations, along with their design variable values and assigned call signs, are listed in Table 4.

Call Sign	Droop Angle [°]	SW1 [°]	SW2 [°]	SR2 [°]	BW2	B [inches]
Training Geometries						
051001-11CR	-11.125	64.6875	48.3125	-13.59375	0.35625	1293.75
081001-11FM	-11.125	68.4375	48.3125	-13.59375	0.35625	1293.75
051001-11RS	-11.125	64.6875	48.3125	-13.59375	0.35625	1293.75
051001-11XH	-11.125	64.6875	48.3125	-26.09375	0.35625	1293.75
051000-11ZE	-11.125	64.6875	48.3125	-13.59375	0.31875	1293.75
081000-2BR	-2.375	68.4375	48.3125	-26.09375	0.31875	1181.25
102102-12BI	-12.375	69.6875	59.3125	-26.09375	0.45625	1193.75
032102-1SS	-1.125	63.4375	59.3125	-26.09375	0.45625	1193.75
032101-12UJ	-12.375	63.4375	59.3125	-13.59375	0.41875	1281.25
02050104LQ	3.875	62.1875	42.8125	-7.34375	0.40625	1143.75
02050103CN	2.625	62.1875	42.8125	-7.34375	0.40625	1143.75
01050104PW	3.875	60.9375	42.8125	-7.34375	0.40625	1143.75
02160104YE	3.875	62.1875	53.8125	-7.34375	0.40625	1143.75
02050204HR	3.875	62.1875	42.8125	-7.34375	0.46875	1143.75
02050104LY	3.875	62.1875	42.8125	-7.34375	0.40625	1231.25
01160203YK	2.625	60.9375	53.8125	-7.34375	0.46875	1231.25
071600-6LQ	-6.125	67.1875	53.8125	-19.84375	0.30625	1243.75
071600-7CQ	-7.375	67.1875	53.8125	-19.84375	0.30625	1243.75
061600-6JJ	-6.125	65.9375	53.8125	-19.84375	0.30625	1243.75
071600-6CD	-6.125	67.1875	53.8125	-19.84375	0.30625	1243.75
071601-6EI	-6.125	67.1875	53.8125	-19.84375	0.36875	1243.75
071600-6HX	-6.125	67.1875	53.8125	-19.84375	0.30625	1131.25
060501-7AY	-7.375	65.9375	42.8125	-19.84375	0.36875	1131.25
Test Geometries						
051001-2TZ	-2.375	64.6875	48.3125	-13.59375	0.35625	1293.75
051001-11LW	-11.125	64.6875	48.3125	-13.59375	0.35625	1181.25
070500-6IZ	-6.125	67.1875	42.8125	-19.84375	0.30625	1243.75
02050104VH	3.875	62.1875	42.8125	-7.34375	0.40625	1143.75

Table 4 Design parameters for configurations in the dataset used in the current study.

B. Surrogate Model Results

The goal of this section is to quantify the prediction error of the surrogate model and to establish its suitability for conceptual-level aerodynamic shape optimization by demonstrating the geometric generalizability of the proposed surrogate model on random configurations within the design space. Two tests are presented. The first test evaluates the model's prediction accuracy on vehicle configurations from the training geometries set but under randomly selected free-stream conditions that were not exposed to the surrogate model during training (the machine learning test set under the training geometries set). Specifically, a qualitative discussion of the DCP prediction quality is provided, along with a quantitative comparison between QOIs evaluated from the predicted DCP and those from VLM. The second test examines the surrogate model's QOIs evaluations on configurations from the test geometries set, in order to assess performance under both generalized geometries and varying free-stream condition inputs.

1. Training Geometries

We first evaluate the root mean square error (RMSE) of DCP predicted by the GNN-based surrogate model and compare the results against those from VLM, to assess the aerodynamic prediction improvement achieved by adopting the surrogate model. In the RMSE computation, the DCP from mapped CFD (ΔC_p^{HF}) is treated as the "ground truth," and the predicted DCP (ΔC_p^{pred}) or VLM's DCP evaluation are compared against it. The RMSEs for DCP prediction on the training set and the test set of configurations from the training geometries set are 0.182 and 0.186, respectively. The lack of significant difference between the training and test set RMSEs indicates that the surrogate model is well-trained and generalizes within the training geometry set. For comparison, the RMSEs of the DCP computed by VLM for the same training and test sets are 0.994 and 0.958, respectively. On average, a five-fold accuracy improvement is achieved by using the GNN-based surrogate model for DCP evaluation.

Figure 5 presents a qualitative comparison among DCPs obtained from VLM, the `VortexNet` surrogate model, and mapped CFD. The vehicle configurations are selected to provide a representative overview of DCPs across a broad design space. The specific AOAs at which these DCPs are evaluated belong to the test set of the surrogate model's training. The corresponding vehicle configuration call sign and the angle of attack of the free-stream condition are annotated at the top of each DCP figures. Four vehicle configurations, with AOAs ranging from 13° to 21°, are shown. The specific design variables for each configuration can be looked up from Table 4.

When the DCPs are compared column-wise (i.e., across different sources), it is evident that the surrogate model's predicted DCP more closely resembles the HF reference data than the LF VLM field distribution. Specifically, under conditions where vortex lift is present, such as 102102-12BI at 19° and 071600-6HX at 21°, the surrogate model successfully reconstructs the vortex lift suction effects, resulting in a pressure field that better matches the HF results. For conditions where vortex lift is less prominent, such as for 051001-11RS, the DCP distributions across VLM, `VortexNet`, and CFD appear similar. It is worth noting that the surrogate model does not introduce artificial vortex lift into the pressure field for such geometric and free-stream conditions. Overall, the GNN-guided surrogate model demonstrates strong generalizability and accurate field prediction across a wide range of geometric configurations and AOAs.

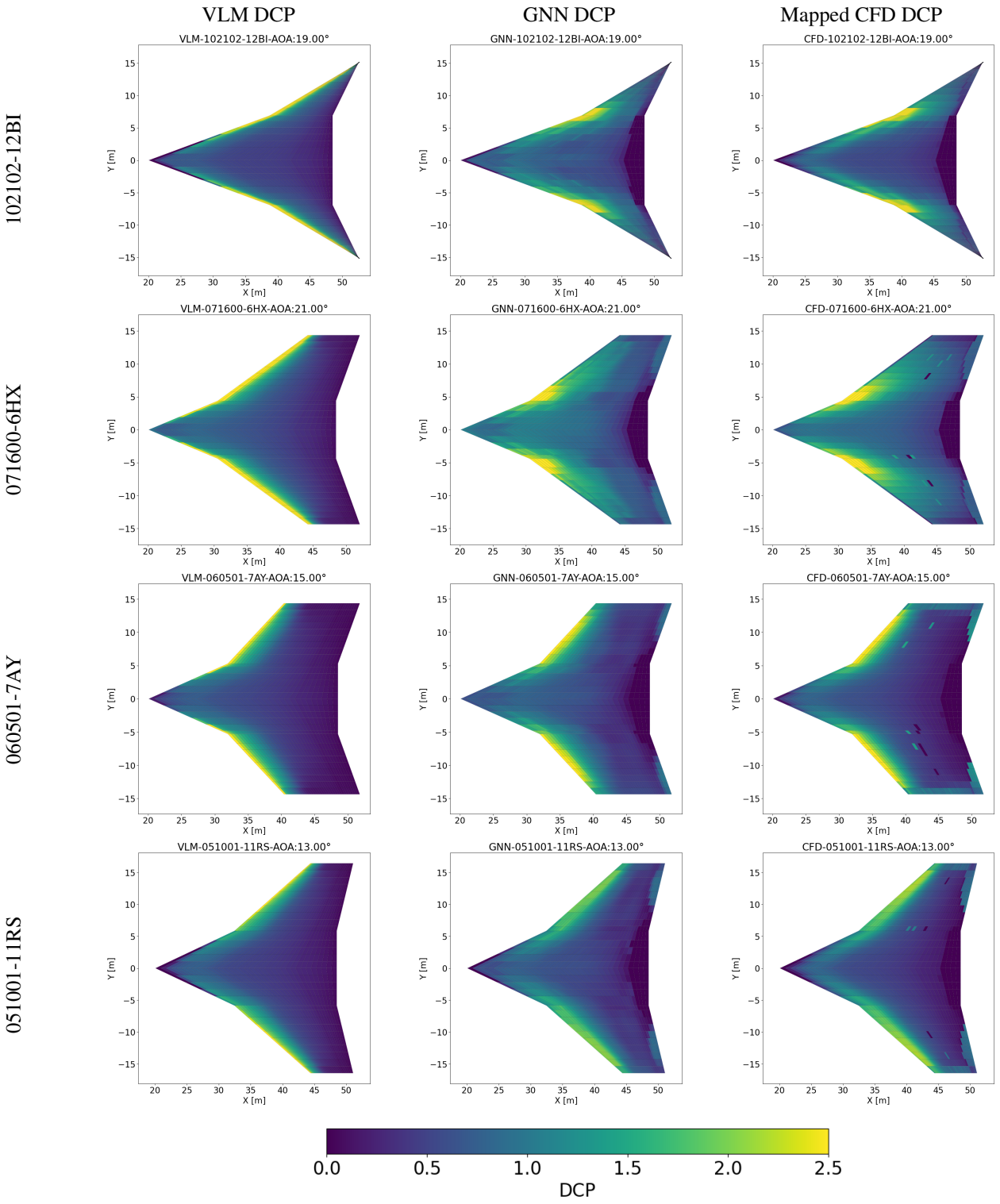


Fig. 5 Lattice panels and local loading coefficient (DCP) for various configurations and AOAs in the test set conditions from the training geometries. Left to right: VLM results, GNN results, and CFD surface pressure projected onto lattice panels.

Once the DCPs are obtained, the corresponding QOIs can be computed by feeding the DCPs back into `SUAVE`, as

outlined in Section III.C. Figure 6 presents the resulting QOIs as a function of AOA, ranging from 9° to 21° in 2° intervals, for the vehicle configurations shown in Figure 5. The vehicle call signs are indicated at the top of each plots. From left to right, each figure shows the variation of C_L, C_D, and C_M with respect to AOA. The QOIs predicted by the GNN-based surrogate model are plotted in blue, while results from VLM and CFD are shown in green and gold, respectively. It is worth noting that some of the free-stream conditions presented in these plots are included in the surrogate model's training set, allowing for a complete angle-of-attack sweep to be presented.

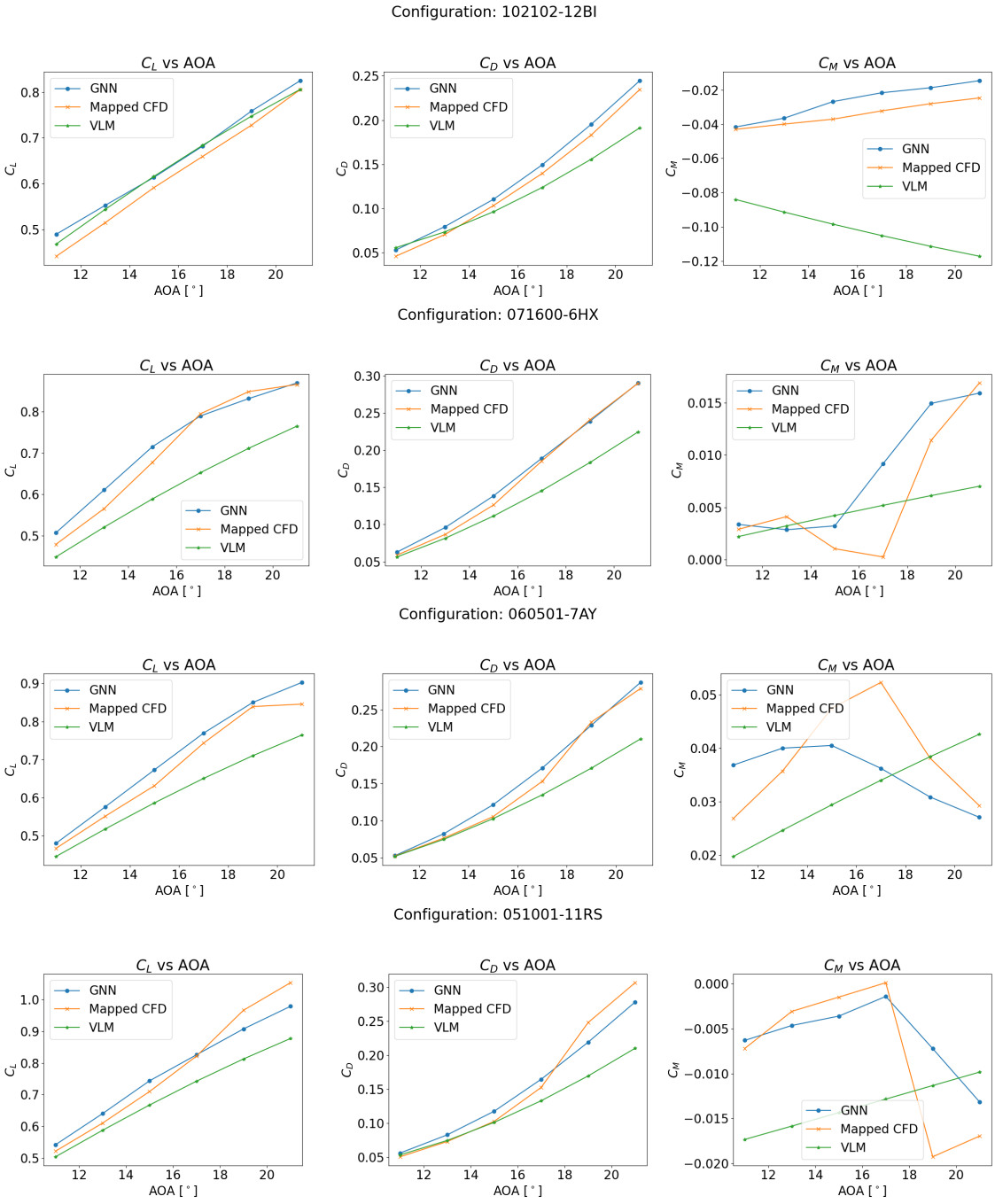


Fig. 6 QOIs sweep for selected configurations from the training geometries set. AOA from 9° to 21°.

Overall, the surrogate model-augmented QOIs estimations are significantly more accurate compared to those from VLM. Specifically, for C_L, the surrogate model not only reduces the prediction error in the absolute C_L values but

also captures the lift curve slope $C_{L,\alpha}$ accurately while the VLM-based computation fails to predict reliably. More importantly, the surrogate model provides a notably improved estimation of C_M: while VLM fails to capture the overall trend of C_M and cannot correctly identify the saddle point, which is critical for the static margin evaluation, while the surrogate model successfully reproduces these features. Considering both $C_{L,\alpha}$ and $C_{M,\alpha}$, although the surrogate model's estimation of $C_{M,\alpha}$ exhibits higher variance than that of $C_{L,\alpha}$, its adoption still substantially enhances design fidelity compared to relying solely on VLM for QOIs quantification.

One way to quantify the prediction accuracy across all AOAs for certain vehicle configuration is to use normalized root mean square error (NRMSE), defined as:

$$\text{NRMSE}(f^{\text{method}}) = \frac{1}{\overline{f^{\text{CFD}}}} \sqrt{\frac{1}{N} \sum_{i=1}^{N} (f_i^{\text{method}} - f_i^{\text{CFD}})^2}, \tag{10}$$

where f_i^{method} is a QOI evaluated by a specific method, f_i^{CFD} is the reference CFD QOI, $\overline{f^{\text{CFD}}}$ is the mean of the reference QOI values, and N is the total number of samples in that vehicle configuration set.

The NRMSE comparison results for C_L and C_M, computed using the surrogate model-augmented DCP and using VLM, across all test set free-stream conditions for configurations within the training geometries set, are presented in Table 5. A Z-score test with a two-standard-deviation threshold is applied to the GNN-based C_L and C_M NRMSE values to assess the consistency of the surrogate model's prediction accuracy across different geometries. It is observed that the C_L prediction for configuration 081001-11FM and the C_M prediction for configuration 032101-12UJ are considered as outliers. For the remaining configurations, the prediction quality is found to be relatively consistent. The deterioration in accuracy for these two outliers is likely due to their design points are near the boundaries of the design space. Both configurations exhibit relatively high droop angles and wider spans, which may introduce more complex flow physics that the surrogate model may not have encountered sufficiently during training. Nevertheless, the comparable NRMSE observed for the majority of vehicle configurations demonstrates that the surrogate model generalizes well across a variety of designs within the prescribed design space.

Call Sign	VLM C_L NRMSE	GNN C_L NRMSE	VLM C_M NRMSE	GNN C_M NRMSE
032101-12UJ	0.0280	0.0453	36.6143	5.2295
071600-6HX	0.1175	0.0406	0.8653	0.2025
102102-12BI	0.0338	0.0408	2.2498	0.3016
01050104PW	0.1648	0.0186	0.1790	0.0672
081000-2BR	0.1570	0.0481	1.5235	0.9771
02160104YE	0.2169	0.0669	0.4942	0.0942
051001-11RS	0.0516	0.0484	5.6214	0.8181
071600-7CQ	0.0467	0.0397	0.9940	0.1488
061600-6JJ	0.0992	0.0321	1.2457	0.2233
071600-6LQ	0.0545	0.0205	3.0687	0.1543
051000-11ZE	0.1331	0.0532	2.0196	1.3857
071601-6EI	0.0379	0.0376	21.8263	0.4049
060501-7AY	0.0714	0.0662	0.3804	0.1466
032102-1SS	0.1427	0.0381	1.1923	0.0616
071600-6CD	0.0548	0.0237	2.9414	0.1180
051001-11CR	0.1590	0.0608	0.4113	0.6245
02050104LQ	0.1936	0.0220	0.2742	0.1018
01160203YK	0.1112	0.0302	0.6517	0.0075
081001-11FM	0.1613	0.0886	0.0915	0.1973
02050104LY	0.1102	0.0017	0.3542	0.2016

Table 5 Comparison of VLM and GNN NRMSE for C_L and C_M across free-stream conditions in the machine learning test set for geometries within the training geometries set.

Finally, Table 6 summarizes the average NRMSE for each method, based on the data reported in Table 5. By incorporating the GNN-based surrogate model into the aerodynamic evaluation workflow, a 2.5× improvement in C_L prediction and an 8.2× improvement in C_M prediction are achieved.

Methods	C_L NRMSE	C_M NRMSE
GNN	0.0418	0.9022
VLM	0.1061	7.3888

Table 6 Averaged NRMSE for test set samples within the training geometries set, summarized by methods.

2. Test Geometries

The subsequent test examines the performance of the surrogate model on vehicle configurations that were withheld from training, as listed in Table 4. These four geometries are randomly selected and span a diverse range of shapes. The objective of this test is to evaluate the geometric generalizability of the surrogate model for accurate aerodynamic quantification under random geometries within a design space.

The resulting QOIs, evaluated using VLM, the GNN-based surrogate model, and the CFD reference, for geometries within the test geometries set are shown in Figure 7. All data points presented were not included in the surrogate model's training. Across all cases, the QOIs predicted by the surrogate model consistently exhibit higher accuracy than those computed using VLM. In particular, both C_L and $C_{L,\alpha}$ predictions from the surrogate model closely resemble the CFD results, although prediction accuracy varies across different configurations. Similarly, the surrogate model demonstrates improved accuracy in predicting C_M and $C_{M,\alpha}$ compared to VLM. The VLM-based predictions fail to capture the transition behavior of pitching moment at higher AOAs, due to limitations in its underlying physical model.

For some configurations, such as 070500-6IZ and 051001-2TZ, the VLM-predicted C_M trends are opposite to those obtained from CFD across all AOAs. This discrepancy arises because the DCP predicted by VLM is concentrated near the leading edge, due to the Polhamus suction correction used, resulting in inaccurate moment estimations for these geometries. In contrast, the surrogate model is able to capture the trend-wise evolution of C_M and $C_{M,\alpha}$, although the prediction accuracy for $C_{M,\alpha}$ in configuration 051001-2TZ is relatively lower than that of the other configurations. However, when a Z-score test with a two-standard-deviation threshold is applied to the configuration-wise NRMSE obtained from the GNN-guided QOIs estimations, no anomalies are identified. This result suggests that the surrogate model's prediction quality remains consistent across all configurations in this test set.

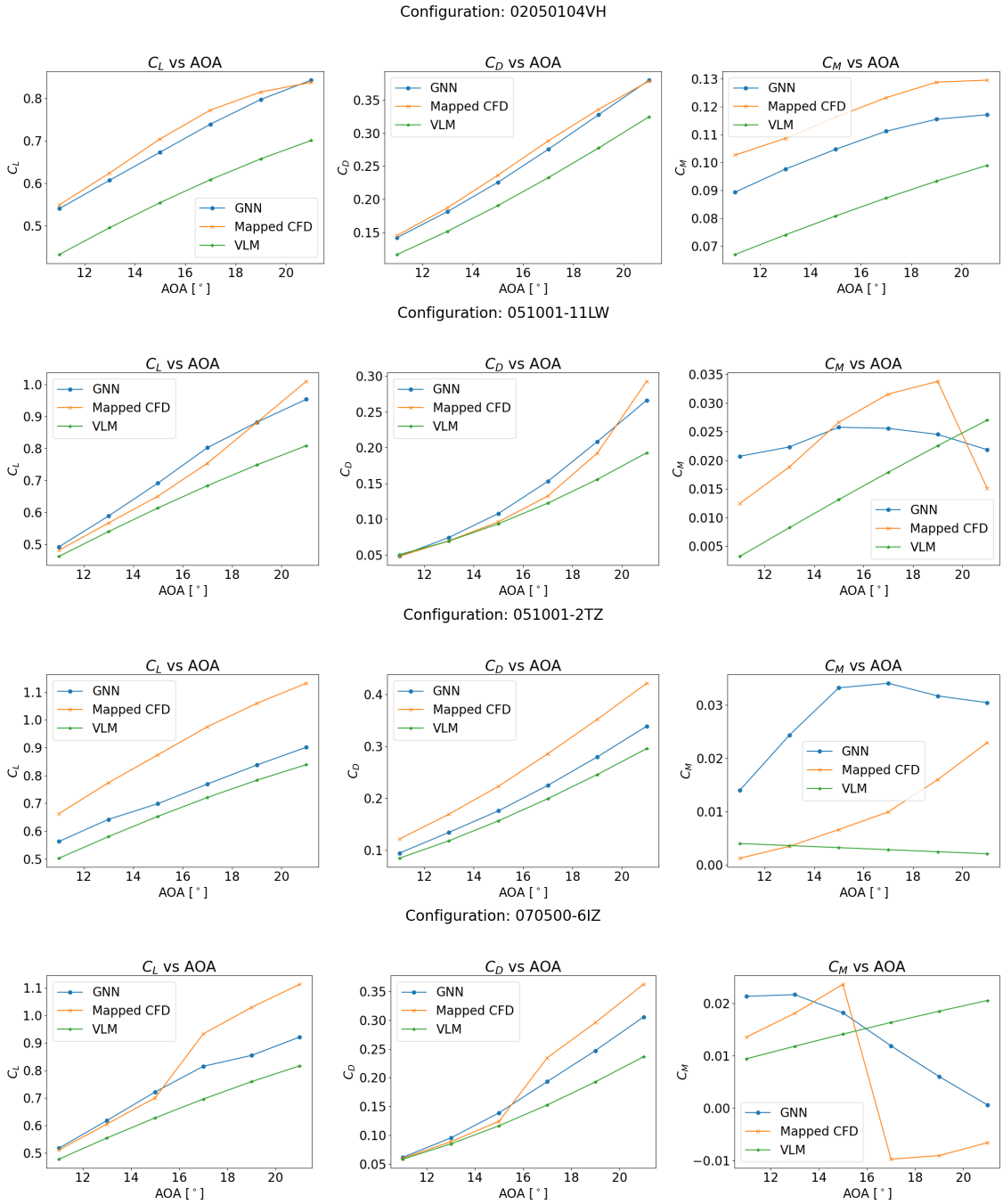


Fig. 7 QOIs sweep for configurations from the test geometries set. AOA from 9° to 21°.

Finally, we summarize the NRMSE values for C_L and C_M across all configurations and AOAs. The results are grouped by QOI evaluation method and are presented in Table 7. The GNN-based NRMSEs are found to be consistent with those reported in the previous test on configurations from the training geometries set, indicating similar prediction accuracy. When comparing the NRMSEs between the VLM and GNN-based methods, the incorporation of the GNN surrogate model in QOIs estimation yields a 2.0× improvement in C_L prediction and a 4.9× improvement in C_M prediction.

Methods	C_L NRMSE	C_M NRMSE
GNN-based	0.0925	0.9407
VLM-based	0.1805	4.6396

Table 7 Averaged NRMSE for all samples within the test geometries set, summarized by methods.

Based on both tests reported in this section, it is evident that the inclusion of the trained GNN-based surrogate model leads to significant improvements in aerodynamic estimation fidelity compared to VLM-based predictions. Furthermore, the GNN-based estimations of $C_{L,\alpha}$ and $C_{M,\alpha}$ closely match those obtained from mapped CFD results, enabling accurate quantification of static margin without requiring direct CFD evaluations within an aerodynamic shape optimization loop. Across all vehicle configurations, the GNN-based aerodynamic predictions exhibit relatively consistent accuracy. Therefore, we conclude that the trained surrogate model is likely to generalize well to other generated geometries within the current design space, that may encounter during the aerodynamic shape optimization.

C. Optimization Results

Two aerodynamic shape optimizations at the conceptual design level are conducted: one using the GNN-guided surrogate model for C_L, C_M, and K_n, and one using only VLM without surrogate augmentation. The optimization problem is defined in Section II, the workflow is described in Section III.C, and the optimizer is detailed in Section III.D. For both cases, the optimization uses the same bounds for the design variables, as listed in Table 3. While the initial design variable values for each particle may differ due to the random generation of initial particles, all other hyperparameters remain consistent across both runs. Therefore, the differences in the optimized configurations can be attributed to the improved accuracy of QOIs and static margin evaluation enabled by the inclusion of the GNN-based surrogate model within the aerodynamic shape optimization workflow.

The optimization history for the six geometric design variables is shown in Figure 8. The evolution of design variables from the GNN-guided workflow is shown in red, while that from the VLM-based workflow is shown in blue. In Figure 8, the solid line represents the design variables corresponding to the global best design, while the dashed line indicates the average values of the design variables across all particles at each iteration. The semi-transparent shaded region illustrates the spread of the particles at each iteration.

Across iterations, the VLM-based optimization converges to an optimal design faster than the GNN-based shape optimization. This is particularly evident from the noisy evolution of the design variable bounds in the GNN-based optimization between iterations 20 and 60. The spikes in the spread of particles are a result of the optimizer occasionally discovering combinations of design variables that produce a negative tip chord. When this occurs, the affected particle is reset to a random location within the initial design space, introducing additional challenges for convergence. Nevertheless, after 60 iterations, both optimization workflows yield sufficiently converged designs. The resultant wing shapes at the 1st, 20th, 49th, and final iterations obtained from both workflows, represented by the camber line plane, are shown in Figure 9 and Figure 10, respectively.

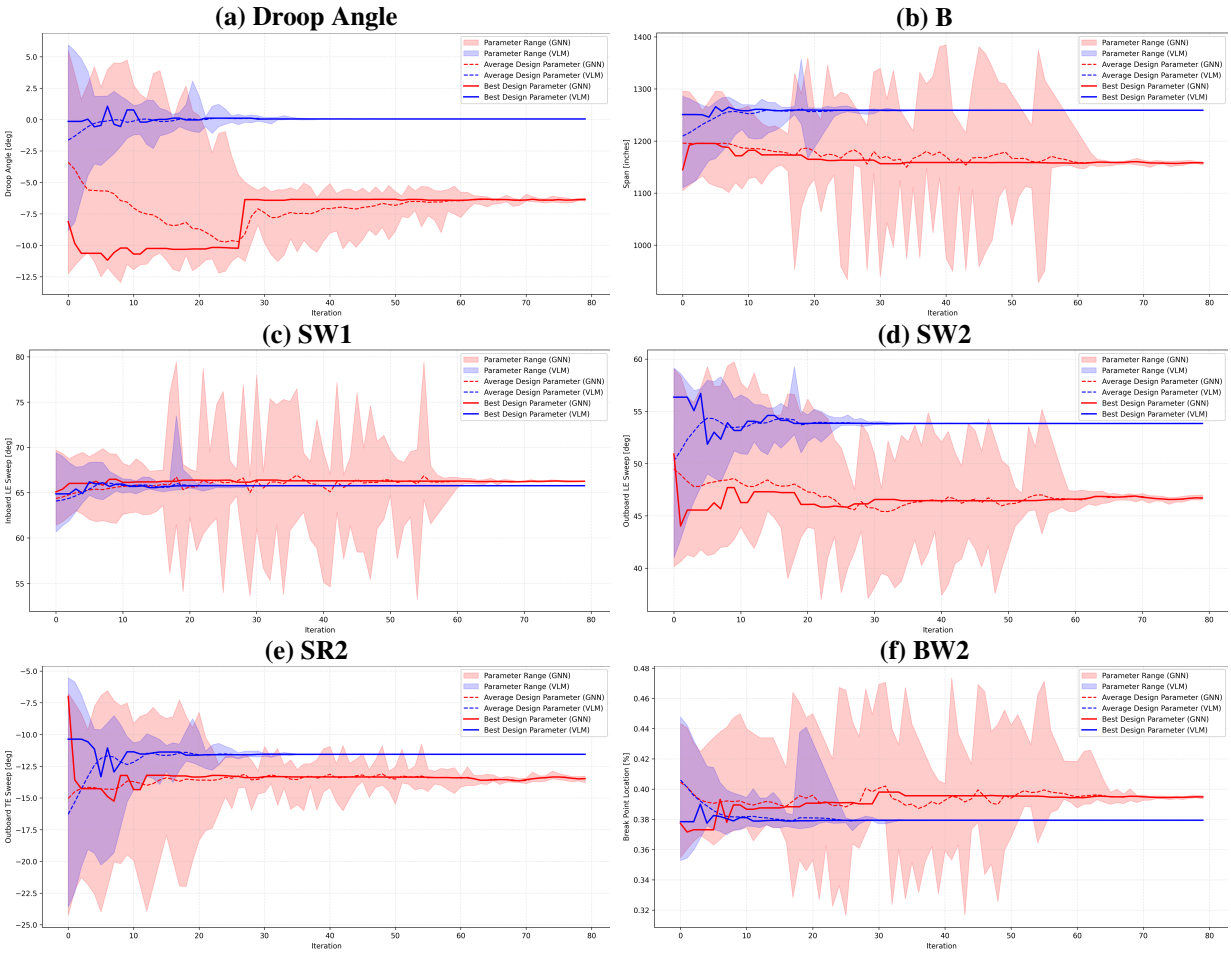


Fig. 8 Evolution of design variables during optimization.

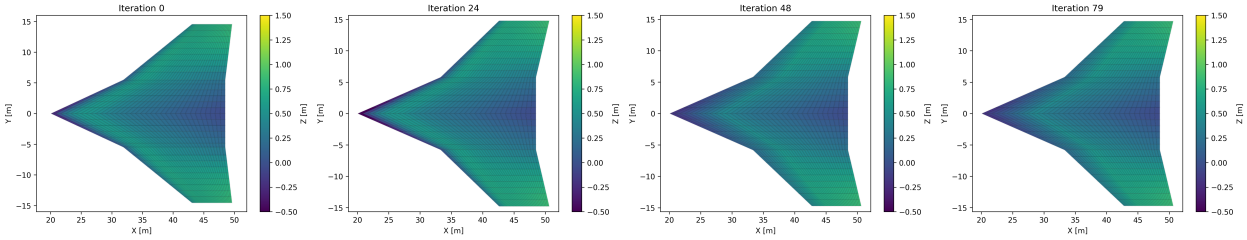


Fig. 9 Evolution of GNN-guided wing configurations, visualized by the camber plane in the x-y plane with the z-coordinate indicated by color.

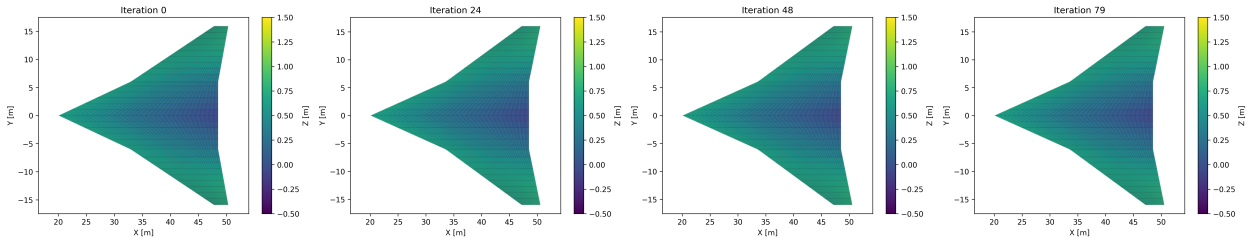


Fig. 10 Evolution of VLM-guided wing configurations, visualized by the camber plane in the x-y plane with the z-coordinate indicated by color.

The optimized designs from VLM-based and GNN-guided QOIs evaluation exhibit notable geometric differences. The VLM-based wing features a less cambered airfoil, with a droop angle of 0.0490°, whereas the GNN-guided wing exhibits significant camber, with a droop angle of −6.360°. The optimized span is also considerably wider in the VLM case, measuring 1259.1 inches compared to 1158.0 inches for the GNN-based configuration. The location of the inboard-outboard section break is more inboard in the VLM-based design, occurring at 0.379, compared to 0.395 in the GNN-guided design. The leading-edge sweep angles of the inboard and outboard sections for the VLM-based design are 65.8° and 53.8°, respectively, while the GNN-guided design adopts a less-swept outboard wing, with sweep angles of 66.3° for the inboard and 46.7° for the outboard sections. Finally, the trailing-edge sweep angle of the outboard section is 11.56° aft-sweep for the VLM-based design and 13.5° aft-sweep for the GNN-guided configuration. The estimated center of gravity locations of the full vehicle, however, are similar for both configurations with VLM-based configuration at 36.85 meters from datum and 36.84 meters for the GNN-guided configuration.

Table 8 summarizes the constraint satisfaction, the AOAs required for both designs at subsonic and supersonic conditions to meet those constraints, and the objective value of supersonic $C_{D,sup}$ and subsonic $C_{D,sub}$ (reported for reference). The subsonic QOIs for the VLM-based design are evaluated directly using VLM, while those for the GNN-guided design are obtained using the surrogate model-augmented workflow. The supersonic QOIs for both designs are evaluated using VLM. Both designs satisfy the specified constraints with comparable levels of constraint violation residuals.

To satisfy these constraints, the VLM-optimized vehicle operates at an AOA of 12.74° for the subsonic condition and 3.49° for the supersonic condition. In contrast, the GNN-guided vehicle operates at higher AOAs, specifically 15.79° and 5.56° for the subsonic and supersonic conditions, respectively. The resulting supersonic C_D, which is to be minimized in the aerodynamic shape optimization problem, is 0.0160 for the VLM-optimized vehicle and 0.0217 for the GNN-guided configuration. The inclusion of the GNN model for QOIs and constraints evaluation introduces a 35.6% increase in supersonic drag for the optimized configurations. On the other hand, the subsonic C_D evaluated for the GNN-guided configuration is 5.1% lower than that of the VLM-based design. This reduction is most likely due to a decreased wing loading, as the wing reference area of the GNN-guided configuration is 3.4% larger.

Metric	VLM-Based	GNN-Guided
Constraints		
Supersonic C_L	0.2000	0.2066
Supersonic C_M	-0.0301	-0.0584
Subsonic C_L	0.7000	0.6998
Subsonic C_M	-0.0308	-0.0158
Subsonic K_n	0.0500	0.0953
Design Variables		
Subsonic AOA	12.74°	15.79°
Supersonic AOA	3.49°	5.56°
Objectives		
Subsonic C_D	0.1749	0.1660
Supersonic C_D	**0.0160**	**0.0217**

Table 8 Comparison of performance metrics for VLM-based and GNN-guided optimized designs.

To better understand the underlying factors driving the optimization results, we conducted a detailed aerodynamic analysis of the optimized configurations. One approach to assess the aerodynamic behavior, as well as the limitations of the respective QOIs evaluation workflows, is to examine the AOA sweep. Figure 11 presents the AOA sweep of C_L, C_M, and K_n under subsonic conditions for the two optimized vehicle configurations. The blue curves represent the QOIs for the VLM-optimized vehicle, evaluated using the VLM-based workflow, while the red curves correspond to the GNN-guided optimized vehicle, with QOIs evaluated using the GNN-augmented workflow.

A comparison of the QOIs obtained from both workflows reveals three major differences. First, the C_L predicted by the GNN-guided workflow reflects the aerodynamic influence of the increased camber introduced by the 6.36° droop angle, resulting in a higher C_L at 0° AOA. Additionally, around 12° AOA, a noticeable reduction in $C_{L,\alpha}$ is observed. This reduction is attributed to a shift in the pressure distribution. Specifically, a weakening of the afterbody DCP and a strengthening of the leading-edge DCP. This transition also induces a mild pitch-up moment in the C_M curve between 12.5° and 13.9°, which diminishes at higher AOAs. In the static margin (K_n) analysis, the VLM-based workflow predicts a statically stable vehicle across all AOAs, whereas the GNN-guided workflow identifies a region of instability, with a negative K_n between AOA of 12.8° and 13.8°.

The aerodynamic behaviors evaluated using the GNN-guided workflow are consistent with low-speed wind tunnel experimental results for delta and cranked delta wings reported by Wentz and McMabon and Freeman [42, 43]. We believe that the aerodynamic evaluations produced by the GNN-guided workflow are of higher fidelity and are suitable for use in aerodynamic shape optimization.

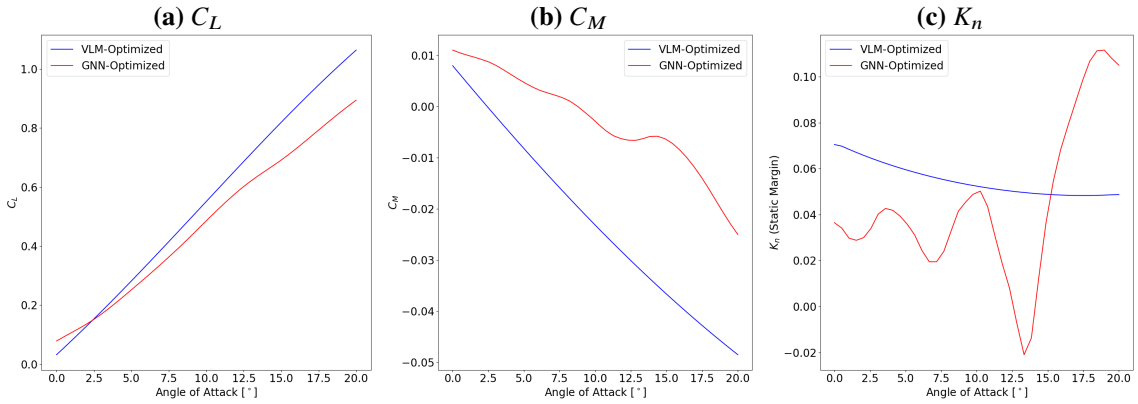


Fig. 11 Comparison of QOIs for (a) C_L, (b) C_M, and (c) static margin K_n, as functions of AOA for optimal configurations obtained from the VLM-based and the GNN-guided workflows.

To further understand the aerodynamic differences between the optimal configurations from both workflows, we examine the DCP at their corresponding subsonic AOAs that satisfy the design constraints (listed in Table 8). Figure 12(a) to Figure 12(c) show the comparison for the two optimal configurations, from the VLM-based and GNN-guided workflows, under two methods for DCP computation: ΔC_p^{LF} evaluated from VLM and ΔC_p^{pred} evaluated from GNN. As shown in Figure 12(a), the lift is concentrated at the leading edge of the vehicle. The resultant center of pressure is located 1.38 meters behind the center of gravity, suggesting that the vehicle is stable.

For the GNN-guided optimal configuration (Figure 12(b)), a less concentrated DCP at the leading edge is observed. In addition to the primary vortex contribution, a high-pressure region can be identified near the 0.3 chord position across the span, corresponding to the increased pressure effect on the lower surface induced by the dropped leading edge. The resulting center of pressure is located 0.6 meters behind its center of gravity. We further analyzed the DCP for the VLM-based optimal configuration using the GNN surrogate model augmentation and the result is shown in Figure 12(c). Due to the flat camber profile and higher wing sweep, a strong primary vortex forms and significantly increases lift near the break chord location. The resulting DCP distribution differs drastically from that shown in Figure 12(a), and the center of pressure shifts to 0.1 meters ahead of the center of gravity, indicating that this primary vortex has a destabilizing effect on the vehicle.

As a validation effort, we conducted two additional CFD simulations for the optimized configurations using the same CFD settings described in Section III.A. Since a direct comparison between the CFD-computed surface pressure coefficients and the DCP distributions shown in Figure 12(a)–(c) would be both cumbersome and potentially misleading, the CFD surface pressure coefficients are projected onto the lattice panels to generate CFD-based DCP ground truth. The resulting DCP fields for the VLM-based and GNN-guided configurations are shown in Figure 12(d) and (e), respectively.

Examining the ground truth result for the VLM-based configuration in Figure 12(d), the pressure distribution is qualitatively similar to that in Figure 12(c), where a strong primary vortex originating from the inboard section and extending outward is observed. However, the magnitude of the maximum pressure difference is under-predicted by the GNN model. This behavior was also observed in some test geometries and is likely due to the surrogate model having insufficient training data near these geometric configurations. Additionally, the use of an L1 loss function [16], which is less sensitive to local small deviations, may contribute to the under-prediction.

When comparing the GNN-guided DCP prediction (Figure 12(b)) to its CFD ground truth (Figure 12(e)), a similar error pattern is observed: the surrogate model tends to under-predict the peak DCP associated with the primary vortex. Furthermore, it appears to over-predict the influence of the leading edge droop, as indicated by the slightly increased pressure difference around the 0.3 chord location. This discrepancy in DCP magnitude may lead to inaccuracies in the magnitude estimation of C_M. We believe that these discrepancies can be mitigated by increasing the sampling density of training geometries and adopting a more sensitive data-driven loss function during surrogate model training.

We observe that increasing the leading-edge droop and reducing outboard sweep weakens the strength of the inboard leading-edge vortex, shifts the center of pressure rearward, and consequently improves stability. These design tendencies, driven by the optimizer, are aligned with findings observed by Seraj and Martins [13].

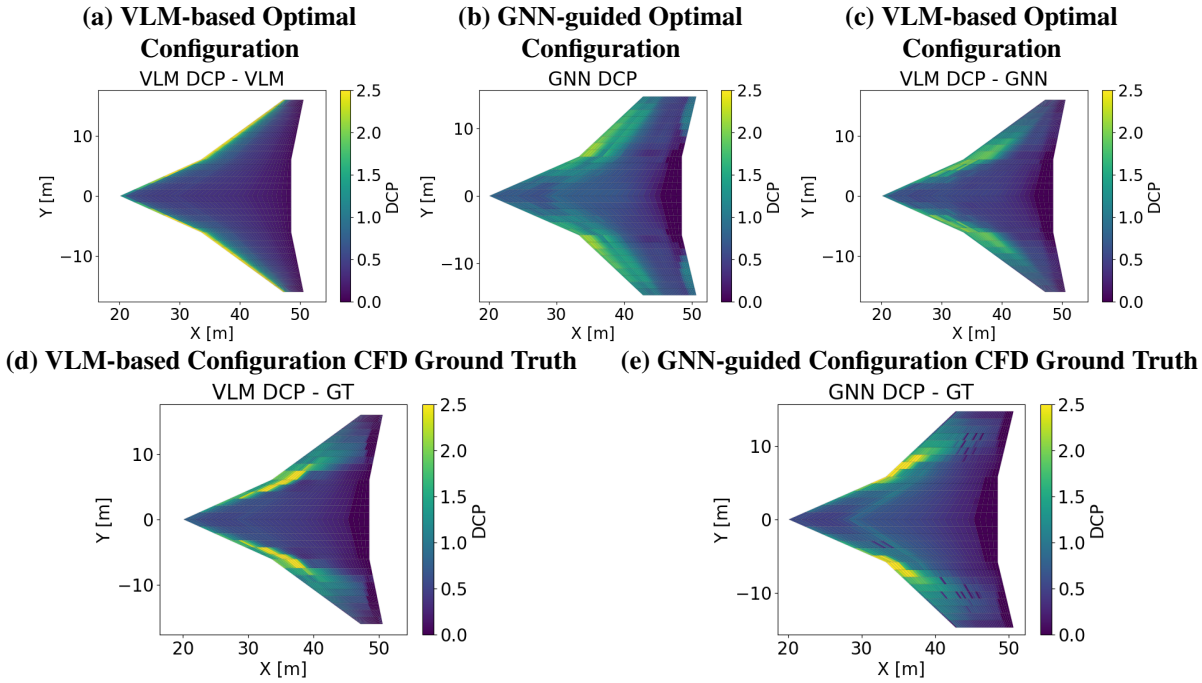


Fig. 12 Optimal wing design DCP distributions: (a) VLM-based configuration evaluated using VLM, (b) GNN-guided configuration evaluated using the GNN surrogate model, (c) VLM-based configuration evaluated using the GNN surrogate model, (d) VLM-based configuration evaluated using CFD, and (e) GNN-guided configuration evaluated using CFD.

Finally, we provide another set of AOA sweep results, similar to those reported in Figure 11, but with the QOIs evaluated using CFD. We ran CFD simulations for both the VLM-based and GNN-guided optimized configurations over AOA values ranging from 9° to 21°, with a one-degree interval. The same convergence criteria discussed in Section III.A are used. For the GNN-guided configuration, the CFD solver encountered convergence difficulties at AOA values greater than 19°. To improve the convergence, the fourth-order JST dissipation coefficient sensor was reduced from 0.02 to 0.018. As a result, the CFD results for the GNN-guided configuration at AOA values above 19° should not be interpreted as exhibiting the same level of convergence as the other data points. Nevertheless, as a preliminary validation effort, we plot the C_L, C_M, and K_n evolution along with the previously reported AOA sweep in Figure 11 in Figure 13. For the K_n plot, Equation 1 is used, and $C_{L,\alpha}$ and $C_{M,\alpha}$ are approximated using central difference.

Across QOIs, the GNN-augmented evaluations resemble the CFD results more closely than the VLM-based assessment. For C_M (Figure 13(b)), although the magnitude of the GNN-augmented evaluation differs from the CFD reference, the deviation remains approximately constant across AOAs. As a result, the $C_{M,\alpha}$ estimation remains credible. The origin of the C_M magnitude deviation is most likely due to the surrogate model's inaccuracy in predicting the maximum DCP, discretization error in the surface pressure projection step, and modeling error introduced by evaluating C_M using the camber plane profile in the current workflow rather than computing surface integration as in CFD.

For the static margin evaluation, as shown in Figure 13(c), the VLM-based workflow fails to capture K_n accurately and does not identify the unstable region. Under CFD evaluation, the VLM-based configuration failed to meet 5% K_n requirement at its 12.74° operation AOA. In contrast, the GNN-guided workflow successfully captures the K_n evolution within the operational range where AOA is less than 18°. At subsonic operation AOA of 15.79°, target K_n is satisfied. However, deviation is observed near AOA of 12°. The cause of this deviation requires further validation efforts in CFD and can be potentially reduced by an improvement of the GNN surrogate model training. Nevertheless, we believe that the improvement in K_n estimation, when applied to a stability-constrained aerodynamic shape optimization problem, can guide the optimizer toward configurations that better satisfy the design constraints and are more likely to align with the design intents aiding the subsequent stages of the design pipeline.

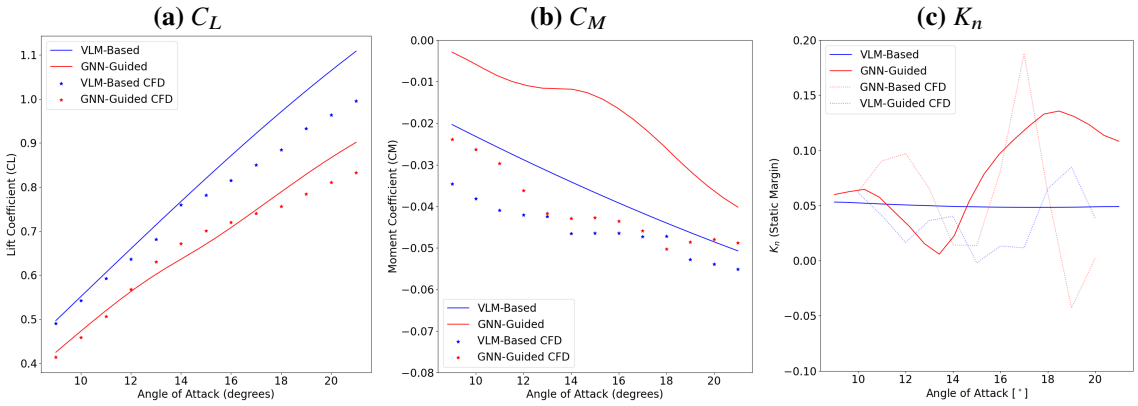


Fig. 13 Comparison of QOIs for (a) C_L, (b) C_M, and (c) static margin K_n, as functions of AOA for optimal configurations.

V. Conclusions and Future Work

This work presents a GNN-guided, multi-fidelity, conceptual-level aerodynamic shape optimization framework and the corresponding results from an optimization test case of a supersonic transport wing. By integrating a pre-trained field prediction surrogate model, `VortexNet`, into the design optimization workflow in `SUAVE` to augment VLM-computed local loading coefficients (DCP), we bridge the fidelity gap between low- and high-fidelity aerodynamic assessments. This integration enables accurate and fast online aerodynamic QOIs evaluations within the optimization loop.

A major focus of this paper is to elaborate the integration of the multi-fidelity aerodynamic shape optimization workflow, including the sampling technique used for training the surrogate model. Validation studies are conducted to evaluate the suitability of deploying the pre-trained surrogate model within an aerodynamic shape optimization loop. With the current workflow, arbitrary QOIs derived from the camber plane pressure field distribution can be evaluated, although in this study we focus on a subset of QOIs related to the longitudinal stability. A detailed comparison highlighting the fidelity gain achieved through the adoption of a GNN-guided workflow in quantifying lift (C_L), pitching moment (C_M) coefficients, and static margin (K_n) is presented. Compared to a traditional conceptual tool based solely on VLM, the augmented workflow achieves up to an 8× reduction in QOI estimation error. Furthermore, once trained on samples drawn from a pre-defined design space, the surrogate model maintains relatively consistent prediction accuracy for both DCP and QOIs across geometric configurations, demonstrating sufficient geometric generalizability for application to unseen, interpolated configurations within the same design space that may be encountered during optimization.

Two optimizations are performed for the same supersonic drag minimization problem under stability constraints and two-point free-stream conditions: one using the GNN-based surrogate model for QOI estimation and one without the GNN augmentation. When PSO is employed as the optimizer, the two cases converge to optima with significantly different geometries. Notably, the static margin constraint is not active in the optimization without GNN augmentation, as the VLM solver is unable to correctly evaluate $C_{L,\alpha}$ and $C_{M,\alpha}$, resulting in misleading and inaccurate static margin quantification. Such resulting optimal configuration, when examined using higher-fidelity methods, fails to satisfy the design constraints and yields suboptimal performance. In contrast, the GNN-guided workflow produces an optimal design that satisfies all constraints including the subsonic static margin requirements and exhibits geometric features and aerodynamic performance that are consistent with previous literature.

Overall, this study underscores the necessity of incorporating high-fidelity data for aerodynamic quantification in SCALOS-like configurations under low-speed stability constraints. The feasibility and effectiveness of embedding a GNN-based surrogate model for pressure field prediction and the resulting aerodynamic evaluations to guide aerodynamic shape optimization is demonstrated.

However, several limitations remain for future development. First, the accuracy of the proposed workflow strongly depends on the surrogate model's predictive performance, which tends to deteriorate when inadequate training samples are provided at near the boundaries of the design space. Additionally, the surrogate model may fail to accurately predict extrema in DCP. Criteria for evaluating surrogate model prediction quality should be developed to enable adaptive sampling or active learning during surrogate model's training. Second, the current conceptual design framework models

only the main wing, which may be insufficient for full-vehicle conceptual design. Future work should incorporate additional aerodynamic surfaces to address trim requirements and control surface sizing. Moreover, fuselage-to-wing aerodynamic interference is currently neglected. Third, the optimization problem considers only two operating conditions and does not reflect a full mission profile. Expanding the operational envelope would increase the utility of the proposed workflow for conceptual configuration refinement. Finally, the current validation effort is preliminary, and future work will focus on a more comprehensive validation of the proposed framework.

VI. Acknowledgment

This work was supported by 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. The authors extend their gratitude to Eli Livne and Sbaet Seraj for their valuable discussions on supersonic aircraft configurations and optimizations. The authors also wish to especially thank Kuang-Ying Ting and Chester Nelson for their insightful discussions on sharing the design intent, component weight balance, key design variables, and lessons learned from designing the reference supersonic airplane concept (RSAC). Additionally, the authors would like to thank Stanford Research Computing and Stanford Data Science for providing computational resources on the Sherlock and Marlowe clusters [44], as well as the Google Cloud Research Credit Program for granting access to GPU resources on the Google Cloud Platform.

References

[1] Bastedo, W., "Commerical Hypersonic Transportation Market Study," Contractor Report 20210014711, NASA Gleen Research Center, 2021.

[2] Welge, H. R., Bonet, J., Magee, T., Chen, D., Hollowell, S., Kutzmann, A., Mortlock, A., Stengle, J., Nelson, C., Adamson, E., Baughcum, S., Britt, R. T., Miller, G., and Tai, J., "N+2 Supersonic Cencept Development and System Integration," Contractor Report NASA/CR-2010-216842, NASA Langley Research Center, 2010.

[3] Morgenstern, J., Buonanno, M., Yao, J., Murugappan, M., Paliath, U., Cheung, L., Malcevic, I., Ramakrishnan, K., Pastouchenko, N., Wood, T., Martens, S., Viars, P., Tersmette, T., Lee, J., Simmons, R., Plybon, D., Alonso, J., Palacios, F., Lukaczyk, T., and Carrier, G., "Advanced concept studies for supersonic commercial transports entering service in the 2018-2020 period phase 2," Tech. Rep. June, NASA Gleen Research Center, 2015. URL https://ntrs.nasa.gov/citations/20150015837.

[4] Li, W., and Geiselhart, K., "Mtidisciplinary Design Optimization of Low-Boom Supersonic Aircraft with Mission Constraints," *AIAA Journal*, Vol. 59, 2020, pp. 1–15. https://doi.org/10.2514/1.J059237.

[5] Seraj, S., and Martins, J. R. R. A., "Minimum Trim Drag for a Three-Surface Supersonic Transport Aircraft," *Journal of Aircraft*, Vol. 0, No. 0, 0, pp. 1–12. https://doi.org/10.2514/1.C037899, URL https://doi.org/10.2514/1.C037899.

[6] Lukaczyk, T., Palacios, F., and Alonso, J., "Response Surface Methodologies for Low-Boom Supersonic Aircraft Design Using Equivalent Area Distributions," *12th AIAA Aviation Technology, Integration, and Operations (ATIO) Conference and 14th AIAA/ISSMO Multidisciplinary Analysis and Optimization Conference*, 2012. https://doi.org/10.2514/6.2012-5705, URL https://arc.aiaa.org/doi/abs/10.2514/6.2012-5705.

[7] Nöding, M., Schuermann, M., Bertsch, L., Koch, M., Plohr, M., Jaron, R., and Berton, J. J., "Simulation of landing and take-off noise for supersonic transport aircraft at a conceptual design fidelity level," *Aerospace*, Vol. 9, No. 1, 2022, pp. 1–23. https://doi.org/10.3390/aerospace9010009.

[8] Nelson, C. P., Ting, K. Y., Mavriplis, N., Soltani, R. M., and Livne, E., "Supersonic Configurations at Low Speeds (SCALOS): Project Background and Progress at University of Washington," *AIAA Science and Technology Forum and Exposition, AIAA SciTech Forum 2022*, 2022. https://doi.org/10.2514/6.2022-1803.

[9] Ting, K.-Y., Wiersema, K., Soltani, R., Nelson, C., and Livne, E., "Supersonic Configurations at Low Speeds (SCALOS) Lateral-Directional Aerodynamics: Configuration Variations and Control Surface Effects," *AIAA SCITECH 2025 Forum*, 2025. https://doi.org/10.2514/6.2025-2361, URL https://arc.aiaa.org/doi/abs/10.2514/6.2025-2361.

[10] Ting, K.-Y., Nelson, C. P., Wiersema, K., and Livne, E., "Supersonic Configurations at Low Speeds (SCALOS): Progress at the University of Washington," *AIAA Sci*, 2024. https://doi.org/10.2514/6.2024-0898.

[11] Mavriplis, N., Ting, K.-Y., Soltani, R. M., Nelson, C. P., and Livne, E., "Supersonic Configurations at Low Speeds (SCALOS): CFD Aided Wind Tunnel Data Corrections," *AIAA Scitech 2024 Forum*, 2023, pp. 1–28. https://doi.org/10.2514/6.2023-0231.

[12] Shanbhag, T. K., Zhou, B. Y., Molina, E., and Alonso, J., "A comparison of jet acoustic analysis methods," *AIAA Aviation 2021 Forum*, 2021, p. 2102.

[13] Seraj, S., and Martins, J. R., "Aerodynamic Shape Optimization of a Supersonic Transport Considering Low-Speed Stability," *AIAA SCITECH 2022 Forum*, 2022. https://doi.org/10.2514/6.2022-2177, URL https://arc.aiaa.org/doi/abs/10.2514/6.2022-2177.

[14] Seraj, S., and Martins, J. R. R. A., "Predicting the High-Angle-of-Attack Characteristics of a Delta Wing at Low Speed," *Journal of Aircraft*, Vol. 59, No. 4, 2022, pp. 1071–1081. https://doi.org/10.2514/1.C036618, URL https://doi.org/10.2514/1.C036618.

[15] Guimarães, T. A., Cesnik, C. E., and Kolmanovsky, I., "Low Speed Aerodynamic Modeling for Control-related Considerations in Supersonic Aircraft Design," *AIAA Aviation and Aeronautics Forum and Exposition, AIAA AVIATION Forum 2021*, 2021, pp. 1–18. https://doi.org/10.2514/6.2021-2531.

[16] Shen, Y., Needels, J. T., and Alonso, J. J., "VortexNet: A Graph Neural Network-Based Multi-Fidelity Surrogate Model for Field Predictions," *AIAA SCITECH 2025 Forum*, 2025. https://doi.org/10.2514/6.2025-0494, URL https://arc.aiaa.org/doi/abs/10.2514/6.2025-0494.

[17] Shen, Y., and Alonso, J. J., "Performance Evaluation of a Graph Neural Network-Augmented Multi-Fidelity Workflow for Predicting Aerodynamic Coefficients on Delta Wings at Low Speed," *AIAA SCITECH 2025 Forum*, 2025. https://doi.org/10.2514/6.2025-2360, URL https://arc.aiaa.org/doi/abs/10.2514/6.2025-2360.

[18] Magee, T. E., Hayes, P. J., and Khodadoust, A., "Dynamic Stability Characteristics for Commercial Supersonic Configurations at Low-Speed Flight Conditions – Part III," *AIAA SCITECH 2025 Forum*, 2025. https://doi.org/10.2514/6.2025-2363, URL https://arc.aiaa.org/doi/abs/10.2514/6.2025-2363.

[19] Ting, K.-Y., Nelson, C., Wiersema, K., and Livne, E., "Supersonic Configurations at Low Speeds (SCALOS): Wind Tunnel Tests at the University of Washington - an Overview and Some Key Insights," *AIAA SCITECH 2025 Forum*, 2025. https://doi.org/10.2514/6.2025-2359, URL https://arc.aiaa.org/doi/abs/10.2514/6.2025-2359.

[20] Wiersema, K., Ting, K.-Y., Nelson, C., and Livne, E., "Supersonic Configurations at Low Speeds(SCALOS): Correlation of CFD and Wind Tunnel Test Results," *AIAA SCITECH 2025 Forum*, 2025. https://doi.org/10.2514/6.2025-2362, URL https://arc.aiaa.org/doi/abs/10.2514/6.2025-2362.

[21] Mader, C. A., Kenway, G. K. W., Yildirim, A., and Martins, J. R. R. A., "ADflow—An open-source computational fluid dynamics solver for aerodynamic and multidisciplinary optimization," *Journal of Aerospace Information Systems*, 2020. https://doi.org/10.2514/1.I010796.

[22] Raymer, D. P., *Aircraft design: A conceptual approach*, 4th ed., AIAA education series, American Institute of Aeronautics and Astronautics, Reston, Va., 2006. URL http://www.loc.gov/catdir/toc/ecip068/2006004706.html.

[23] Hasan, Y. J., Flink, J., Freund, S., Klimmek, T., Kuchar, R., Liersch, C. M., Looye, G., Moerland, E., Pfeiffer, T., Schrader, M., et al., "Stability and control investigations in early stages of aircraft design," *2018 Applied Aerodynamics Conference*, 2018, p. 2996.

[24] MacDonald, T., Clarke, M., Botero, E. M., Vegh, J. M., and Alonso, J. J., "SUAVE: An Open-Source Environment Enabling Multi-Fidelity Vehicle Optimization," *18th AIAA/ISSMO Multidisciplinary Analysis and Optimization Conference*, 2017. https://doi.org/10.2514/6.2017-4437, URL https://arc.aiaa.org/doi/abs/10.2514/6.2017-4437.

[25] Miranda, L. R., "Extended applications of the vortex lattice method," *NASA. Langley Res. Center Vortex-Lattice Utilization*, 1976.

[26] Guimarães, T. A., and Cesnik, C. E., "The Impact of Vortex Breakdown on the Longitudinal Flight Behavior of the Scalos Aircraft," *AIAA Scitech 2024 Forum*, 2024, pp. 1–13. https://doi.org/10.2514/6.2024-0902.

[27] Polhamus, E. C., "A concept of the vortex lift of sharp-edge delta wings based on a leading-edge-suction analogy," *NASA. Langley Res. Center*, 1966. URL https://ntrs.nasa.gov/citations/19670003842.

[28] Guimarães, T. A., Cesnik, C. E., and Kolmanovsky, I. V., "An Integrated Low-Speed Aeroelastic-Flight-Dynamics Framework for Modeling Supersonic Aircraft," *AIAA SCITECH 2022 Forum*, 2022. https://doi.org/10.2514/6.2022-2175, URL https://arc.aiaa.org/doi/abs/10.2514/6.2022-2175.

[29] Mukhopadhaya, J., Mishra, A., Iaccarino, G., and Alonso, J., "Design Exploration under Epistemic Uncertainties," *Bulletin of the American Physical Society*, Vol. 63, 2018.

[30] Economon, T. D., Palacios, F., Copeland, S. R., Lukaczyk, T. W., and Alonso, J. J., "SU2: An Open-Source Suite for Multiphysics Simulation and Design," *AIAA Journal*, Vol. 54, No. 3, 2016, pp. 828–846. https://doi.org/10.2514/1.J053813, URL https://doi.org/10.2514/1.J053813.

[31] Shur, M. L., Strelets, M. K., Travin, A. K., and Spalart, P. R., "Turbulence Modeling in Rotating and Curved Channels: Assessing the Spalart-Shur Correction," *AIAA Journal*, Vol. 38, No. 5, 2000, pp. 784–792. https://doi.org/10.2514/2.1058.

[32] Jameson, A., Schmidt, W., and Turkel, E., "Numerical solution of the Euler equations by finite volume methods using Runge Kutta time stepping schemes," *14th Fluid and Plasma Dynamics Conference*, 1981. https://doi.org/10.2514/6.1981-1259.

[33] OpenVSP, "Open Vehicle Sketch Pad," https://openvsp.org/, 2025. Accessed: 2025-06-01.

[34] Brody, S., Alon, U., and Yahav, E., "How Attentive are Graph Attention Networks?" , 2022. URL https://arxiv.org/abs/2105.14491.

[35] Akiba, T., Sano, S., Yanase, T., Ohta, T., and Koyama, M., "Optuna: A Next-generation Hyperparameter Optimization Framework," *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, Association for Computing Machinery, New York, NY, USA, 2019, p. 2623–2631. https://doi.org/10.1145/3292500.3330701, URL https://doi.org/10.1145/3292500.3330701.

[36] Gill, P. E., Murray, W., and Saunders, M. A., "SNOPT: An SQP algorithm for large-scale constrained optimization," *SIAM review*, Vol. 47, No. 1, 2005, pp. 99–131.

[37] Kochenderfer, M. J., and Wheller, T. A., *Algorithms for Optimization, Second Edition*, The MIT Press, 2025.

[38] Kennedy, J., and Eberhart, R., "Particle swarm optimization," *Proceedings of ICNN'95 - International Conference on Neural Networks*, Vol. 4, 1995, pp. 1942–1948 vol.4. https://doi.org/10.1109/ICNN.1995.488968.

[39] Martins, J. R., and Ning, A., *Engineering Design Optimization*, Cambridge University Press, 2022.

[40] Lyu, Z., Xu, Z., and Martins, J. R. R. A., "Benchmarking Optimization Algorithms for Wing Aerodynamic Design Optimization," *Proceedings of the 8th International Conference on Computational Fluid Dynamics*, Chengdu, Sichuan, China, 2014. ICCFD8-2014-0203.

[41] Saltelli, A., Tarantola, S., and Chan, K.-S., "A quantitative model-independent method for global sensitivity analysis of model output," *Technometrics*, Vol. 41, No. 1, 1999, pp. 39–56.

[42] Wentz, W. H., and McMabon, M. C., "An Experimental Investigation of the Flow Fields about Delta and Double-Delta Wings at Low Speeds," *NASA CR 521*, 1966. URL https://ntrs.nasa.gov/api/citations/19660025758/downloads/19660025758.pdf.

[43] Freeman, D. C., "Low-Subsonic Aerodynamic Characteristics of a 60° Swept Delta-Wing Space Shuttle Orbiter," *NASA TM X-2762*, 1973. URL https://ntrs.nasa.gov/api/citations/19730018153/downloads/19730018153.pdf.

[44] Kapfer, C., Stine, K., Narasimhan, B., Mentzel, C., and Candes, E., "Marlowe: Stanford's GPU-based Computational Instrument," , Jan. 2025. https://doi.org/10.5281/zenodo.14751899, URL https://doi.org/10.5281/zenodo.14751899.