

A Resources Availability and Licensing

DrivAerStar is distributed under the CC BY-NC-SA 4.0 license to facilitate academic dissemination while protecting intellectual property rights. Under this license, users are permitted to copy, distribute, display, and perform the work, as well as create derivative works, provided that they give appropriate credit to the original authors, provide a link to the license, and indicate if changes were made. **Commercial use of the work is strictly prohibited without explicit written permission from the copyright holders.**

All associated resources, including source code, datasets, and documentation, are publicly accessible to ensure transparency and enable reproducibility. Users are encouraged to contribute improvements through established channels, enhancing the collective value of these open-source resources within the academic community.

B Broader Impact

DrivAerStar addresses critical needs across multiple research and industrial domains. In academic research, the high-fidelity dataset enables development of novel geometric deep learning models for complex fluid dynamics, advancing understanding of how neural networks capture physical phenomena, and supporting investigations into physics-informed machine learning approaches. The dataset expands available benchmarks for graph neural networks and point cloud methods, establishing new evaluation standards for fluid dynamics applications.

For industrial applications, DrivAerStar enables development of fast surrogate models that can replace computationally expensive CFD simulations in design optimization workflows. This capability supports more efficient automotive development cycles and extends to aerospace applications where virtual testing reduces reliance on expensive physical prototypes. The dataset addresses growing industrial demand for understanding deep learning generalization across large solution domains, promoting integration of high-fidelity physics-based simulations with artificial intelligence.

Beyond immediate applications, DrivAerStar establishes a foundation for data-driven design methodologies that can transform engineering practices across disciplines facing computational constraints. By democratizing access to industrial-quality aerodynamic data, the dataset empowers researchers worldwide to develop next-generation computational methods for more efficient and sustainable engineering designs.

Table A1: **Parametric deformation ranges for automotive geometry generation via lattice-based morphing.** Twenty parameters control geometric variations across vehicle components, with ranges established from production vehicle design standards and aerodynamic optimization requirements.

Number	Parameter	Minimum	Maximum	Unit
1	Global Scaling	80%	120%	-
2	Vehicle Width	-0.1	0.1	m
3	Vehicle Length	-0.1	0.1	m
4	Ramp Angle	-10	10	degree
5	Front Bumper Length	-0.1	0.1	m
6	Windscreen X	-0.05	0.05	m
7	Windscreen Z	-0.05	0.05	m
8	Side Mirrors X	-0.05	0.05	m
9	Side Mirrors Z	-0.05	0.05	m
10	Rear Window X	-0.05	0.05	m
11	Rear Window Z	-0.05	0.05	m
12	Trunk Lid Angle	-10	10	degree
13	Trunk Lid X	-0.05	0.05	m
14	Trunk Lid Z	-0.05	0.05	m
15	Diffuser Angle	-10	10	degree
16	Greenhouse Angle	-10	10	degree
17	Front Hood Angle	-10	10	degree
18	Intake Hood Angle	-10	10	degree
19	Tire Diameter	-0.033	0.033	m
20	Tire Width	-0.015	0.015	m

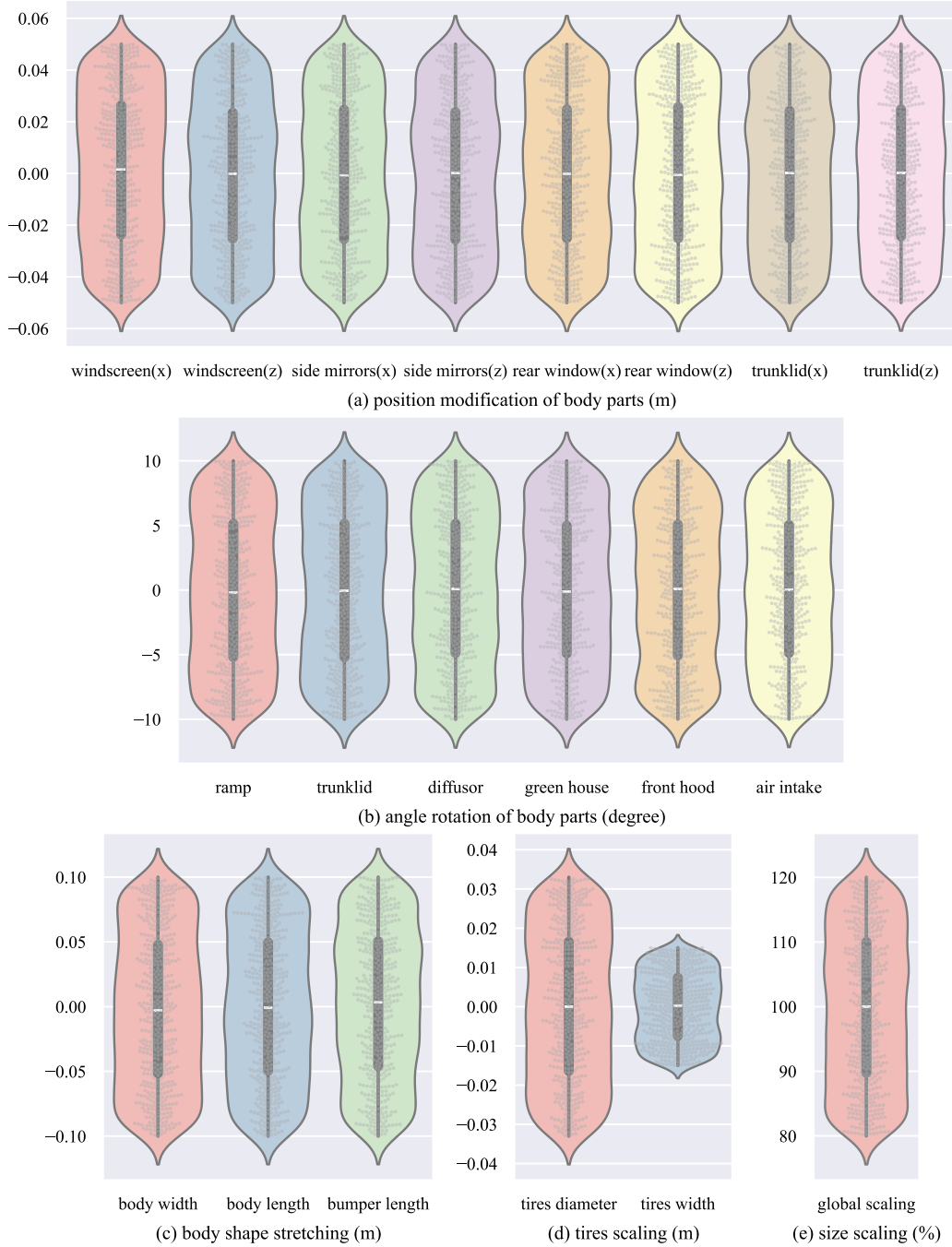


Figure A1: **Parametric deformation diversity in DrivAerStar.** Five deformation categories generate geometric variations through Latin Hypercube Sampling: (a) Linear displacement of four vehicle components within $\pm 0.05\text{m}$ in x and z directions. (b) Angular rotation of six body elements ranging from -10° to 10° . (c) Stretching and compression of vehicle body and front bumper between $\pm 0.1\text{m}$. (d) Tire deformation with diameter variations of $\pm 0.033\text{m}$ and width variations of $\pm 0.015\text{m}$. (e) Global scaling from 80% to 120% of baseline dimensions.

C Dataset Generation, Simulation, and Validation

C.1 Geometric Morphing via Reference Vehicle Architecture and Parametric Design

We develop a parametric geometry generation framework using lattice deformation to create diverse automotive geometries from the open-source DrivAer vehicle model. Our approach employs FFD

techniques implemented through Python scripts on the Blender[®] platform, establishing a $32 \times 8 \times 8$ three-dimensional lattice structure that enables fine-grained control over local body regions while maintaining geometric integrity.

The implementation begins with bounding box analysis to determine spatial distribution of geometric features, establishing mapping relationships between local and global coordinate systems. The parametric framework divides the vehicle into nine key regions including front bumper, trunk, and roof, with regional deformation control achieved through predefined index grouping. We define 20 key parameters covering body morphology, dimensional scaling, and wheel configurations (Table A1). Parameter ranges reference production vehicle design standards and incorporate CFD simulation experience to set boundary conditions that ensure engineering feasibility, with diffuser inclination constrained within $\pm 10^\circ$ to balance downforce and drag requirements.

LHS generates 1000 parameter samples across engineering-reasonable ranges for statistical representativeness, producing thousands of unique geometric configurations. Each parameter corresponds to specific physical modifications: dimensional stretches (front bumper length $\pm 0.1\text{m}$), positional adjustments (trunk inclination $\pm 0.05\text{m}$), size scaling (vehicle width $\pm 0.1\text{m}$), and angular rotations (windscreen inclination $\pm 0.05\text{m}$). Parameter distributions are detailed in Figure A1, with each configuration generating unique variants through nine independent deformation functions.

For cooling system integration—absent in previous datasets—we bind front grille, radiator, engine bay, and gearbox components to unified lattice deformers using LHS methodology. This coordinated deformation prevents spatial interpenetration while maintaining positional coherence among complex internal components during geometric transformations.

Precise wheel alignment employs a dedicated four-wheel positioning algorithm with predefined feature points including 3D coordinate systems for all four wheels. The algorithm dynamically calculates axle positions by tracking spatial transformations of these feature points during deformation, ensuring geometric matching between axles and deformed body while maintaining lattice deformation controllability and avoiding component distortion issues common in traditional morphing techniques.

This lattice deformation approach reduces geometric generation time by over 90% compared to traditional CAD parametric methods while preserving industrial-grade surface quality. The systematic parameterization via FFD maintains geometric continuity, topological consistency, and engineering feasibility while establishing a deformable configuration space rooted in industrial design practices. As illustrated in Figures A2 to A4, the method preserves aerodynamic baseline characteristics of the reference vehicle while enabling parametric integration of engineering details such as cooling systems and chassis components, generating complex geometric interactions including length-width coupling variations that provide foundation for neural networks to capture nonlinear flow features.

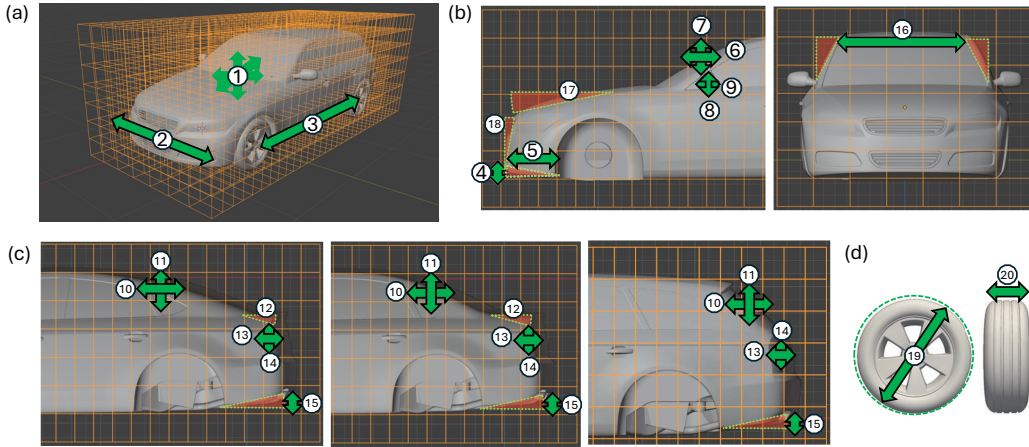


Figure A2: **Lattice deformation framework for DrivAer vehicle configurations.** (a) Deformation lattice structure with vehicle scaling, width, and length adjustment parameters. (b) Front vehicle section parameters controlling bumper and forward geometry modifications. (c) Rear section parameter sets for three vehicle types: fastback (left), notchback (center), and estateback (right) configurations. (d) Wheel geometry parameters for tire diameter and width variations.

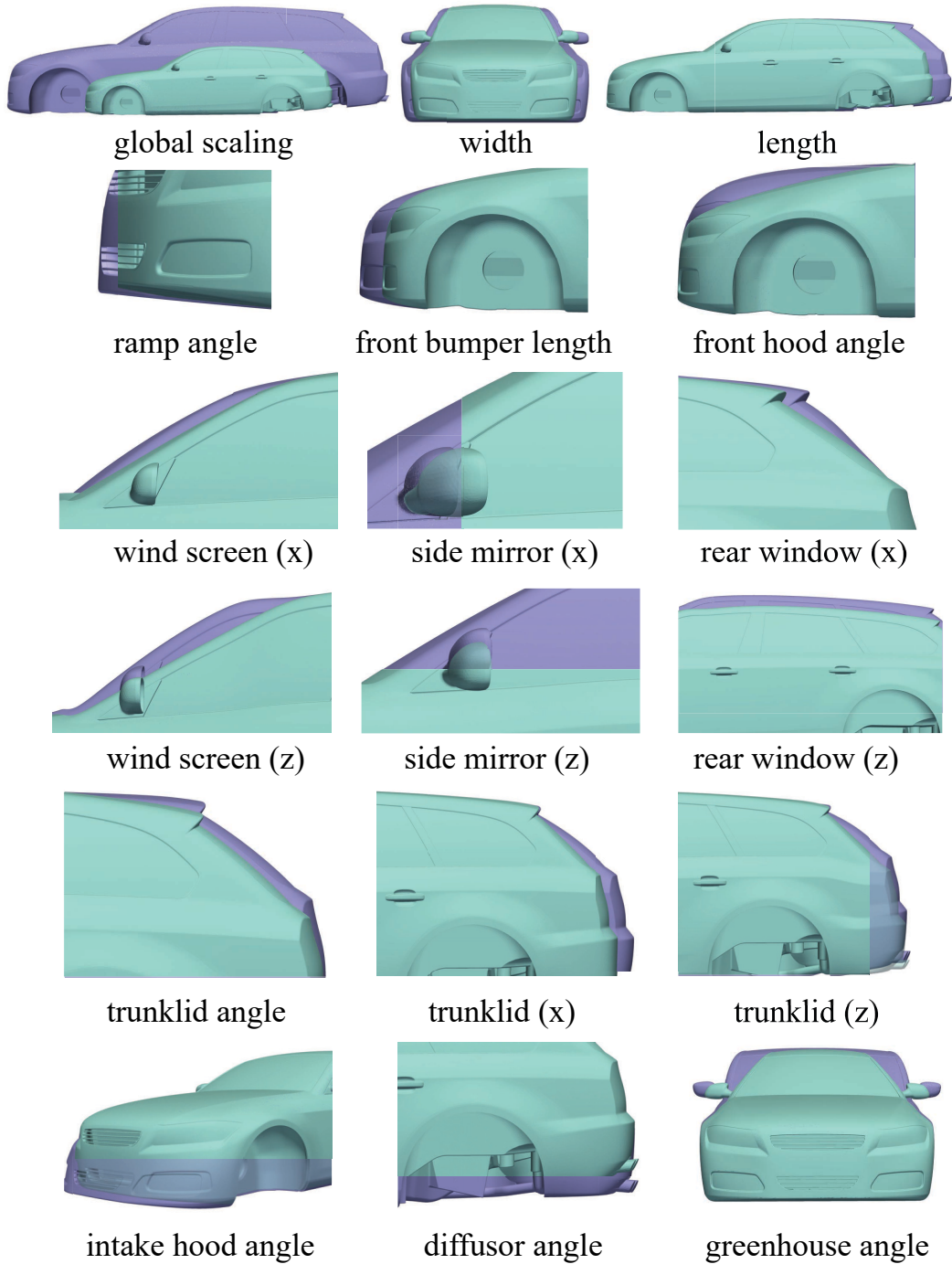


Figure A3: **Individual parameter deformation visualization.** Lattice deformation results for single-parameter variations across minimum to maximum ranges. Blue models represent minimum parameter values, purple models represent maximum values. Wheel size parameters are excluded to emphasize vehicle body geometry modifications.

Future research will integrate RANS simulation results to establish high-dimensional mapping models between morphological parameters and aerodynamic coefficients, supporting development of intelligent aerodynamic optimization frameworks.

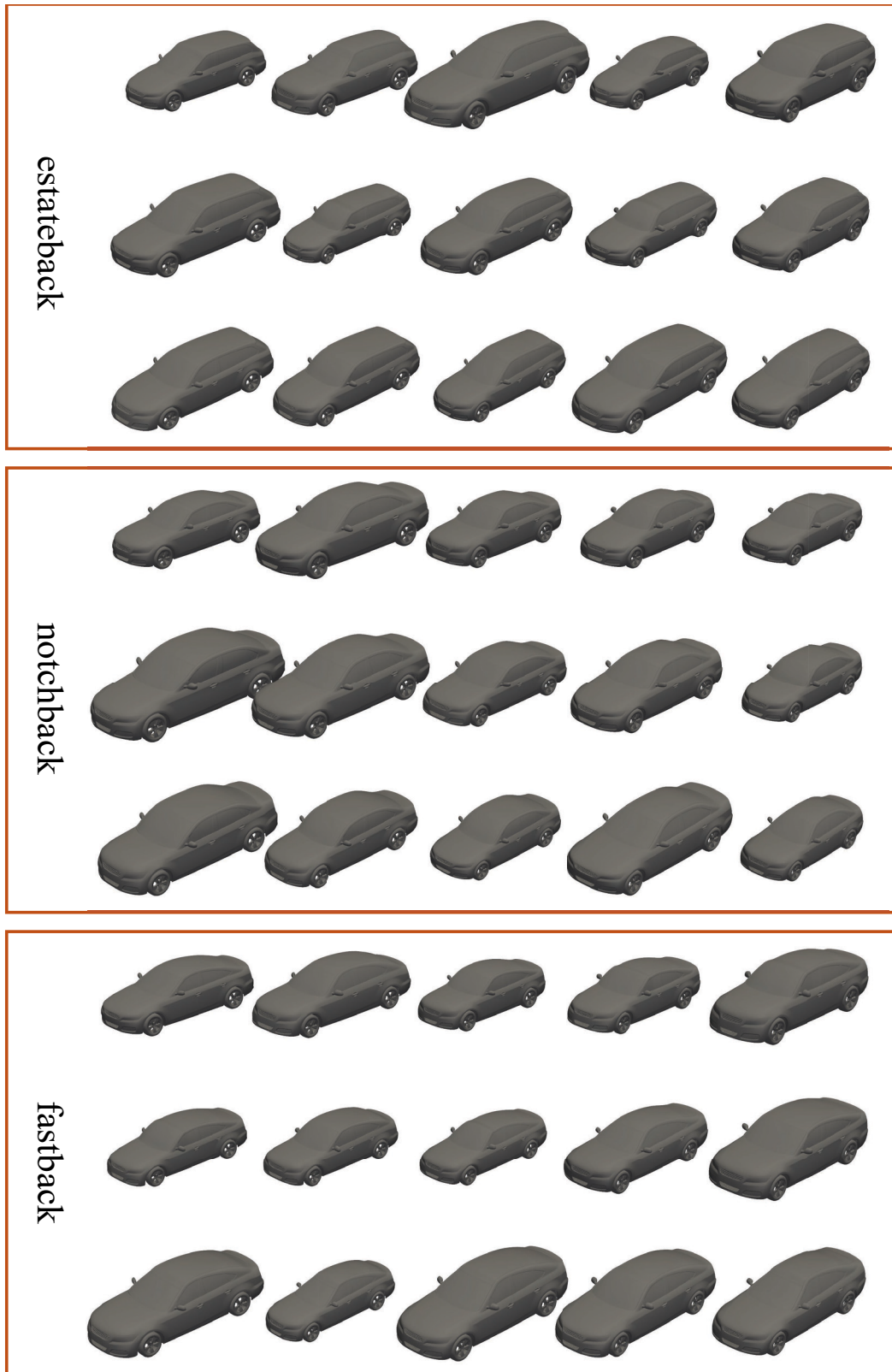


Figure A4: **Multi-parameter deformation examples by vehicle type.** Deformation variations for three rear configurations: estateback (top), notchback (middle), and fastback (bottom).

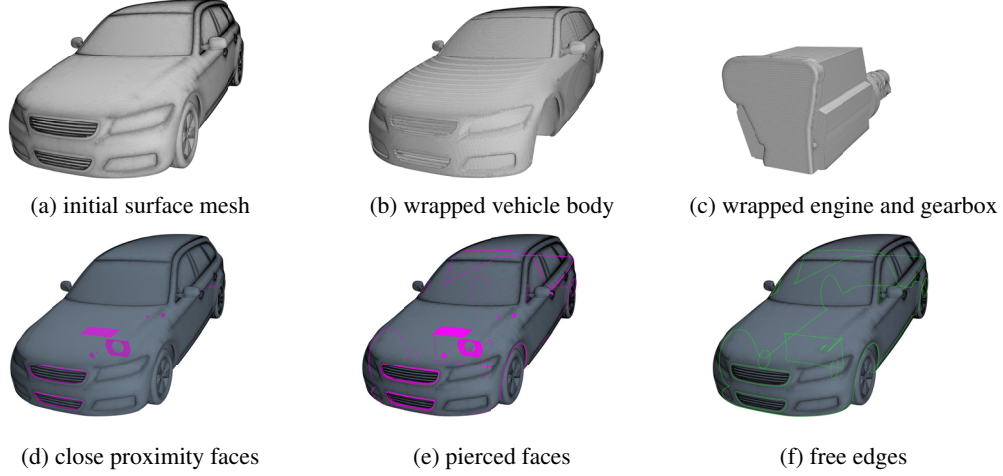


Figure A5: **Surface wrapping process using STAR-CCM+®**. Surface wrapping workflow applied to vehicle components: (a) Initial generated surface mesh with critical defects. (b) Wrapped vehicle body mesh. (c) Wrapped engine and gearbox components. The process addresses three fundamental surface errors: (d) close proximity faces, (e) pierced faces, and (f) free edges, with defective faces highlighted in purple and problematic edges in green.

C.2 Surface Wrapping Implementation Details

Generated vehicle geometries exhibit critical surface defects that prevent direct CFD computation, requiring mandatory surface wrapping preprocessing for all components. As illustrated in **Figure A5**, we employ STAR-CCM+®'s surface wrapping functionality to independently process each component (body, drivetrain, wheels) of the initial surface mesh, addressing three fundamental mesh errors: (i) *Close Proximity* faces with minimum distances below 0.1 mm, (ii) *Pierced Faces* that create flow leakage paths, and (iii) *Free Edges* that compromise watertight geometric continuity.

Aerodynamically critical regions, including grilles and wheel arches, receive enhanced treatment through local refinement with 10% size reduction and curvature adaptation using a minimum spacing of 0.5 mm. Post-wrapping validation ensures aspect ratios remain below 20 and skewness angles exceed 10° across all mesh components. This optimized workflow achieves 98.6% defect reduction while maintaining 0.28 mm geometric fidelity, enabling reliable CFD simulations with aerodynamic feature resolution below 1 mm accuracy.

C.3 Regional Mesh

Mesh Generation Framework The parallel mesh generation employs a global base size $L_0 = 0.24$ m with surface growth rate 1.3, utilizing hexahedral elements enhanced by triangular surface reconstruction for geometric fidelity. Dual refinement strategies ensure critical feature capture: curvature refinement with 0.01 m deviation distance and 36 points per circle, plus proximity refinement for chassis, engine bay, and wheel components.

Volume growth adopts *very slow* mode with maximum element size constrained to 0.48 m ($200\%L_0$) and minimum surface size to 0.024 m ($10\%L_0$). Three-tiered volume controls implement progressive refinement: Block 2 at 0.06 m ($25\%L_0$), Block 3 matching L_0 , and Block 4 at 0.024 m ($10\%L_0$).

Surface and Boundary Layer Treatment Vehicle body surfaces receive enhanced treatment with target surface size 0.0144 m ($6\%L_0$) and minimum size 0.0096 m ($4\%L_0$). Prism layers follow geometric progression with first-layer height $h_1 = 0.005$ mm, expansion ratio $r = 1.2$, and 8-layer configuration yielding total thickness $H = h_1 \cdot (r^8 - 1)/(r - 1) \approx 0.114$ mm. External boundaries (inlet/outlet/side/top) utilize coarser 0.48 m elements without prism layer refinement.

Quality Control and Validation Mesh quality maintains triangular surface elements above a 0.05 minimum threshold with boundary layer resolution enforcing $y^+ > 30$ using blended wall functions.

Mesh independence verification employs three resolutions—coarse (0.48 m), medium (0.24 m), and fine (0.12 m)—validated through drag coefficient and flow field analysis.

C.4 Navier–Stokes Equations and Turbulence Model

Reynolds-Averaged Navier-Stokes Approach Automotive external aerodynamics simulation requires appropriate turbulent flow treatment. While Direct Numerical Simulation (DNS) captures all turbulence scales, it demands prohibitive computational resources for practical automotive applications. This study employs the RANS approach, modeling time-averaged flow fields at significantly reduced computational cost compared to DNS or Large Eddy Simulation (LES) alternatives.

The RANS equations derive from Reynolds decomposition, where each flow variable decomposes into mean and fluctuating components: $\phi = \bar{\phi} + \phi'$. For steady-state incompressible flow, the governing equations become:

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \bar{\mathbf{v}}) = 0, \quad (\text{A1})$$

$$\frac{\partial}{\partial t}(\rho \bar{\mathbf{v}}) + \nabla \cdot (\rho \bar{\mathbf{v}} \otimes \bar{\mathbf{v}}) = -\nabla \bar{p} + \nabla \cdot (\bar{\mathbf{T}} + \mathbf{T}_{\text{RANS}}), \quad (\text{A2})$$

where ρ , $\bar{\mathbf{v}}$, \bar{p} , and $\bar{\mathbf{T}}$ represent air density, mean velocity vector, mean pressure, and mean viscous stress tensor, respectively. The Reynolds stress tensor \mathbf{T}_{RANS} requires turbulence modeling closure:

$$\mathbf{T}_{\text{RANS}} = -\rho \begin{pmatrix} \overline{u'u'} & \overline{u'v'} & \overline{u'w'} \\ \overline{u'v'} & \overline{v'v'} & \overline{v'w'} \\ \overline{u'w'} & \overline{v'w'} & \overline{w'w'} \end{pmatrix}, \quad (\text{A3})$$

where u' , v' , w' represent velocity fluctuations in the x -, y -, z -directions, respectively.

SST k - ω Turbulence Model This study employs the SST k - ω turbulence model (Menter, 1994), solving transport equations for turbulent kinetic energy k and specific dissipation rate ω . This model provides superior performance for boundary layers under adverse pressure gradients and applies throughout the boundary layer without wall-distance computation, yielding a reliable approximation for automotive aerodynamics.

The transport equations are formulated as:

$$\frac{\partial}{\partial t}(\rho k) + \nabla \cdot (\rho k \bar{\mathbf{v}}) = \nabla \cdot [(\mu + \sigma_k \mu_t) \nabla k] + P_k - \rho \beta^* f_{\beta^*} (\omega k - \omega_0 k_0), \quad (\text{A4})$$

$$\frac{\partial}{\partial t}(\rho \omega) + \nabla \cdot (\rho \omega \bar{\mathbf{v}}) = \nabla \cdot [(\mu + \sigma_\omega \mu_t) \nabla \omega] + P_\omega - \rho \beta f_\beta (\omega^2 - \omega_0^2), \quad (\text{A5})$$

$$\mu_t = \min \left(\frac{1}{\omega}, \frac{1}{F_2 S} \right), \quad (\text{A6})$$

$$\phi = F_1 \phi_1 + (1 - F_1) \phi_2, \quad (\phi = \sigma_k, \sigma_\omega, \beta), \quad (\text{A7})$$

where μ and μ_t denote dynamic and turbulent eddy viscosity, f_{β^*} and f_β represent free-shear and vortex-stretching modification factors, k_0 and ω_0 are ambient turbulence values (Spalart and Rumsey, 2007), and F_1 , F_2 are blending functions connecting inner and outer boundary layer regions.

Model parameters are: $\sigma_{k1} = 0.85$, $\sigma_{k2} = 1.0$, $\sigma_{\omega1} = 0.5$, $\sigma_{\omega2} = 0.856$, $\beta_1 = 0.075$, $\beta_2 = 0.0828$. Following the Boussinesq hypothesis, the Reynolds stress tensor is modeled as:

$$\mathbf{T}_{\text{RANS}} = \mu_t \left(\frac{\partial \bar{u}_i}{\partial x_j} + \frac{\partial \bar{u}_j}{\partial x_i} \right) - \frac{2}{3} \rho k \delta_{ij}, \quad (\text{A8})$$

providing complete system closure.

Porous Media Treatment Vehicle cooling components (radiators, condensers) are modeled as porous media due to their complex internal geometries that are too fine for individual meshing. The void fraction ε_g represents the geometry at coarse mesh scales, while momentum exchange is captured through body force terms.

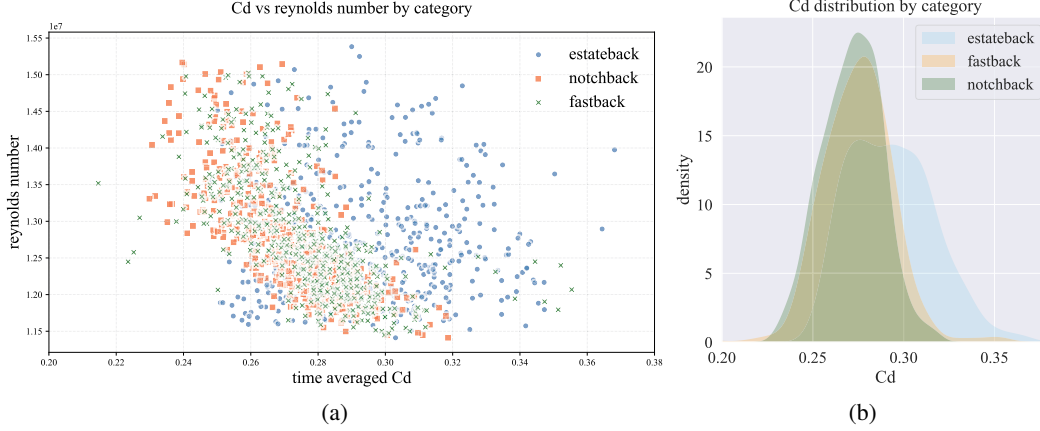


Figure A6: **Aerodynamic performance comparison of three rear configurations.** (a) Drag coefficient (C_D) versus Reynolds number (Re) with 95% confidence intervals (shaded regions). (b) Probability density distributions of C_D values highlighting modal characteristics. Estateback (blue), notchback (orange), and fastback (green) configurations are consistently color-coded across subfigures.

The governing equations for porous media become:

$$\frac{\partial(\varepsilon_g \rho)}{\partial t} + \nabla \cdot (\rho \bar{\mathbf{v}}) = 0, \quad (\text{A9})$$

$$\frac{\partial(\rho \bar{\mathbf{v}})}{\partial t} + \nabla \cdot (\rho \bar{\mathbf{v}} \otimes \bar{\mathbf{v}}) = -\nabla p + \nabla \cdot (\bar{\mathbf{T}} + \mathbf{T}_{\text{RANS}}) + \mathbf{F}_p, \quad (\text{A10})$$

where the flow resistance \mathbf{F}_p is given by:

$$\mathbf{F}_p = -(\mathbf{P}_v + \mathbf{P}_i |\bar{\mathbf{v}}|) \cdot \bar{\mathbf{v}}, \quad (\text{A11})$$

with \mathbf{P}_v and \mathbf{P}_i representing viscous and inertial resistance tensors. The superficial velocity $\bar{\mathbf{v}}$ relates to the actual pore velocity \mathbf{v}_g through $\bar{\mathbf{v}} = \varepsilon_g \mathbf{v}_g$.

Further implementation details are available in the [STAR-CCM+® Documentation](#).

C.5 Computational Infrastructure

Dataset generation utilized a high-performance computing cluster with 100 nodes, each equipped with Intel® Xeon® Gold 6148 processors, consuming approximately 1,080,000 core-hours. Computational efficiency was optimized through the AMG linear solver (Stüben, 2001), which accelerates convergence while reducing memory requirements.

C.6 Simulation Result

The aerodynamic behavior of three rear configurations (estateback, notchback, fastback) is quantified in Figure A6, with comprehensive flow field visualizations including pressure distributions (Figure A8), wall shear stress patterns (Figure A9), and velocity profiles (Figure A10).

Reynolds Number Dependencies Drag characteristics exhibit distinct Reynolds number trends across configurations. The notchback configuration demonstrates consistent C_D reduction with increasing Re throughout the observed range, while the fastback shows progressive C_D decrease across its operational regime. In contrast, the estateback maintains relatively uniform C_D values without significant monotonic variation, though with broader dispersion compared to other configurations.

Statistical Characteristics The estateback exhibits bimodal density peaks at $C_D = 0.31$ and 0.35 , contrasting with unimodal distributions of notchback ($\mu = 0.28$, $\text{IQR} = 0.04$) and fastback ($\mu = 0.33$). The notchback's narrow interquartile range confirms superior aerodynamic consistency, while the fastback's right-skewed distribution reflects intermittent high-drag flow states.

C.7 Computational Validation and Quality Assurance

Automated Quality Control A comprehensive quality assessment framework with specific acceptance criteria ensures dataset integrity. Automated filtering verifies that physical quantities remain within engineering ranges: pressure values (volumetric and surface) and wall shear stress do not exceed $\pm 2 \times 10^4$ Pa. Simulations exhibiting numerical anomalies are automatically excluded from the dataset.

Convergence Criteria All cases undergo minimum 1,000 iteration steps with convergence determined by C_D stabilization within $\pm 1\%$ over the final 100 iterations. A hybrid convergence metric considers both relative C_D change and asymptotic behavior of solution residuals across continuity, turbulence, and momentum equations. Cases with unresolved oscillatory behavior or non-physical trends are identified through spectral analysis and excluded.

Mesh and Solution Quality Mesh quality verification ensures appropriate wall treatment application, with all simulations maintaining wall y^+ values outside the buffer layer for accurate near-wall flow resolution using blended wall functions. Meshes with significant skewness or aspect ratio defects undergo refinement or exclusion. Final quality control discards simulations with the lowest 5% composite quality scores, retaining only the top 60% of simulations based on computational quality metrics for subsequent analysis.

C.8 Differences Between Simulation and Experiments

Geometric Configuration Differences The primary geometric differences between simulation and experimental setups concentrate on wheel mounting systems and radiator representation. The front wheels utilize MacPherson strut-type mounting assemblies, while the rear wheels employ solid axle configurations. The experimental radiator geometry is unavailable in the CAD model and is instead represented through porous media modeling with pressure drop characteristics derived from hexagonal aluminum honeycomb and perforated sheet specifications.

The experimental model features five-spoke wheels with detailed tire tread patterns, symmetric mirror assemblies, comprehensive chassis geometry, and three rear configurations: fastback, notchback, and estateback. A precisely machined floor section maintains 4 mm wheel clearances, while detailed engine bay components include integrated cooling systems and spotlights. Wheel assembly tolerances are maintained within ± 0.5 mm.

Physical Modeling Validation Table 2 presents drag coefficient comparison between experimental measurements and DrivAerStar simulation results. All three rear configuration coefficients have been calibrated to approximate open-road operating conditions. The primary error sources include: flow regime complexities around the four-wheel assembly, pressure drop modeling through radiator porosity representation, and inherent turbulence model limitations in capturing separation and reattachment phenomena.

The validation demonstrates acceptable agreement between computational and experimental results, with discrepancies primarily attributed to geometric simplifications in the cooling system representation and the challenge of accurately modeling complex wheel-ground interactions in the computational domain.

D Benchmark Setup and Evaluation

D.1 Evaluation Metrics

Relative L_2 Loss (ϵ) The relative L_2 error serves as the primary evaluation metric for flow field regression across all experimental configurations. To ensure comparability between different physical variables, each flow field variable ϕ (velocity, pressure, wall shear stress) undergoes standardization using dataset-wide statistics:

$$\hat{\phi} = \frac{\phi - \mu_\phi}{\sigma_\phi}, \quad (\text{A12})$$

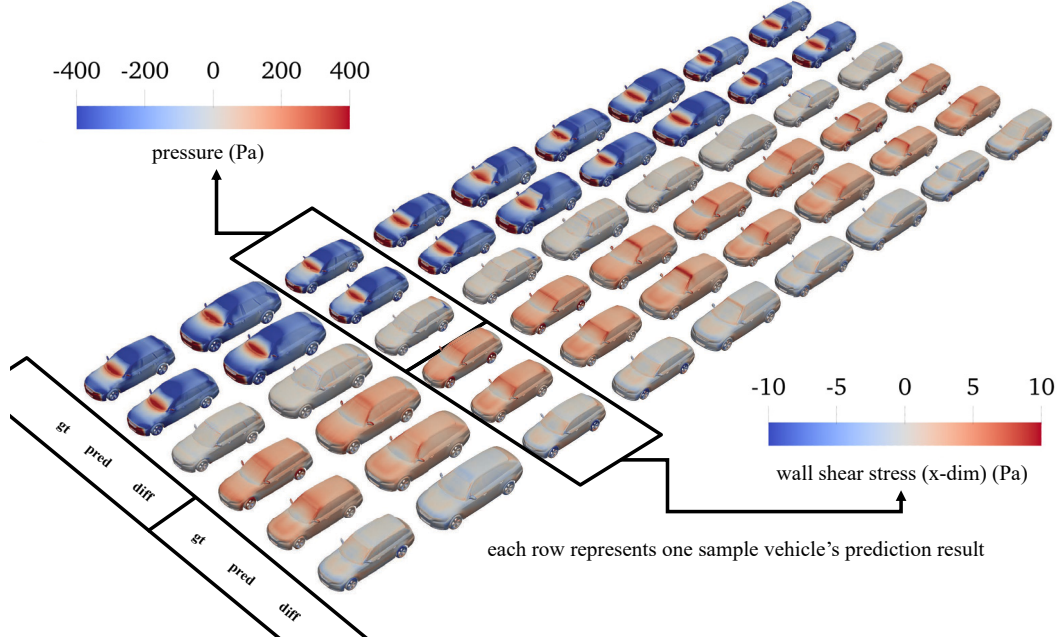


Figure A7: **Benchmark results visualization.** Nine testing cases from the single-vehicle setup demonstrate model performance across unique geometric configurations. Each row presents pressure fields (left) and wall shear stress results (right), with ground truth CFD data, machine learning predictions, and difference maps arranged from left to right within each prediction task.

where μ_ϕ and σ_ϕ represent the mean and standard deviation computed over the entire dataset. The normalized relative L_2 error is then defined as:

$$\epsilon = \frac{\|\hat{\phi}_{\text{pred}} - \hat{\phi}_{\text{true}}\|_2}{\|\hat{\phi}_{\text{true}}\|_2}, \quad (\text{A13})$$

where $\hat{\phi}_{\text{pred}}$ and $\hat{\phi}_{\text{true}}$ denote standardized predicted and ground-truth flow fields, respectively. This normalization eliminates scale discrepancies and aligns with established practices in industrial CFD model evaluation.

Drag Coefficient Counts Error Automotive industry practice quantifies drag coefficient accuracy using “counts error,” where 1 count equals 0.001 in C_D units. For example, if the simulated C_D is 0.300 and the predicted value is 0.320, the count error is 20 counts. This metric provides intuitive error quantification directly relevant to vehicle development targets.

D.2 Benchmark Result

Table A2 presents a systematic performance comparison between Transolver, GNOT, and PointNet frameworks across aerodynamic prediction tasks, revealing critical insights into their respective capabilities.

Single-Vehicle Specialization Transolver achieves superior performance in estateback-specific training, recording the lowest validation losses at 1,200 samples: pressure (0.1996), WSS (0.7953), and C_D (0.0266). The framework demonstrates strong scalability, with C_D prediction error decreasing 15.4% when expanding training data from 400 to 1,200 samples. GNOT shows substantial improvement in high-data regimes, with 1,200-sample C_D validation loss (0.0357) approaching Transolver’s performance, indicating enhanced learning capacity with sufficient data. PointNet exhibits inconsistent convergence behavior, showing pressure validation loss degradation (3.1% increase from 0.2991 to 0.2984) despite tripling training data. Single-vehicle benchmark results are visualized in **Figure A7**.

Multi-Vehicle Generalization Transolver maintains robust performance across vehicle configurations, showing minimal WSS validation loss degradation (2.9%) between single-vehicle (0.7953)

Table A2: **Training and validation loss comparison between Transolver and GNOT.** Loss curves for pressure, wall shear stress (WSS), and drag coefficient (C_D) prediction tasks. Training configuration: 500 samples per vehicle type (estateback, fastback, notchback) for 1,500 total samples. All loss values are normalized by initial conditions for comparative analysis.

(a) Transolver							
Category	Configuration	Training Loss			Validation Loss		
		Pressure	WSS	C_D	Pressure	WSS	C_D
Single-Vehicle	400×Estateback	0.2112	0.7841	0.0305	0.2173	0.7889	0.0314
	800×Estateback	0.2050	0.7793	0.0296	0.2086	0.7837	0.0283
	1200×Estateback	0.1969	0.7881	0.0315	0.1996	0.7953	0.0266
Multi-Vehicle	133×3 Vehicles	0.2076	0.8050	0.0421	0.2194	0.8049	0.0375
	266×3 Vehicles	0.1967	0.8130	0.0335	0.2075	0.8170	0.0335
	400×3 Vehicles	0.1957	0.7973	0.0333	0.2074	0.7871	0.0286

(b) GNOT							
Category	Configuration	Training Loss			Validation Loss		
		Pressure	WSS	C_D	Pressure	WSS	C_D
Single-Vehicle	400×Estateback	0.3250	0.9619	0.1144	0.3250	0.9619	0.1144
	800×Estateback	0.2185	0.7991	0.0463	0.2254	0.7986	0.0516
	1200×Estateback	0.1960	0.8135	0.0032	0.2087	0.8217	0.0357
Multi-Vehicle	133×3 Vehicles	0.2069	0.7802	0.0301	0.2122	0.7795	0.0344
	266×3 Vehicles	0.1968	0.7909	0.0284	0.2002	0.7920	0.0309
	400×3 Vehicles	0.1954	0.8171	0.0282	0.1964	0.8124	0.0315

(c) PointNet							
Category	Configuration	Training Loss			Validation Loss		
		Pressure	WSS	C_D	Pressure	WSS	C_D
Single-Vehicle	400×Estateback	0.2805	0.6460	0.0627	0.2991	0.6547	0.0502
	800×Estateback	0.2549	0.6755	0.0376	0.2787	0.6769	0.0499
	1200×Estateback	0.2717	0.6157	0.0571	0.2984	0.6273	0.0436
Multi-Vehicle	133×3 Vehicles	0.2971	0.6716	0.0722	0.3201	0.6767	0.0536
	266×3 Vehicles	0.2731	0.6875	0.0585	0.3007	0.7041	0.0653
	400×3 Vehicles	0.2872	0.6316	0.0589	0.3030	0.6560	0.0757

and 400×3 multi-vehicle (0.7871) scenarios. GNOT demonstrates superior cross-vehicle generalization, with 400×3 multi-vehicle C_D validation loss (0.0315) outperforming single-vehicle baseline (0.0357) by 11.8%, suggesting inherent architectural advantages for heterogeneous geometries. PointNet struggles with multi-task learning, exhibiting 41.7% higher average WSS validation loss compared to Transolver in equivalent configurations.

Task-Specific Analysis Wall shear stress prediction presents the greatest challenge across all frameworks, with the highest relative errors observed (Transolver: 0.7953, GNOT: 0.8124, PointNet: 0.6560 at 400×3 validation). Drag coefficient predictions reveal framework-dependent characteristics: Transolver achieves ultra-low errors (0.0266) through specialized training, while GNOT’s physics-informed architecture enables superior error consistency ($\pm 6.2\%$ variation) across vehicle configurations.

Sample Efficiency Analysis Transolver reaches 90% peak performance with 800 samples, showing marginal improvement ($\pm 2\%$) at 1,200 samples. GNOT requires 1,200 samples for comparable accuracy, while PointNet exhibits negative scaling beyond 800 samples, with 1,200-sample training yielding 4.3% higher pressure loss than the 800-sample case, indicating architectural limitations in leveraging large datasets.

D.3 Scaling Analysis with Training Sample Size

We conducted comprehensive scaling experiments using Transolver in multi-vehicle configurations to assess training dataset size impact on drag coefficient prediction. The analysis utilized 12,000 total samples partitioned into training subsets of varying sizes, with 150 independently generated samples for testing. This revised data splitting strategy ensures robust evaluation, with minor discrepancies from initial results attributed to refined subset partitioning. Validation loss and relative improvement (calculated as $\text{Improvement} = \frac{\text{Loss}_{400} - \text{Loss}_N}{\text{Loss}_{400}} \times 100\%$ against the 400-sample baseline) are presented in Table 3.

Results demonstrate consistent C_D validation loss reduction with increased training data: doubling samples from 400 to 800 reduces loss by 10.67%, scaling to 1,200 samples achieves 23.73% reduction, and the full 12,000-sample dataset delivers 29.07% total improvement. This trend confirms Transolver’s effective data scaling properties, where larger training subsets directly translate to enhanced C_D prediction accuracy. The substantial accuracy gains achieved with the 12,000-sample dataset provide meaningful improvements for practical CFD applications, where marginal C_D prediction error reductions drive significant advances in vehicle aerodynamic optimization workflows. Raw experimental data, including detailed loss curves and test sample distributions, are provided in supplementary materials to ensure reproducibility.

E Data Sources of Automotive CFD Datasets

The comparative metrics presented in Table 1 were compiled from multiple sources. For the DrivAerNet++ dataset (Elrefaie et al., 2024), the wall y^+ range and the average C_D precision were derived from the published simulation cases, while the wind tunnel error was extracted from Table 8 of the original paper. For the DrivAerML dataset (Ashton et al., 2024c), the wall y^+ range was not reported in the available literature; the average C_D precision was estimated from Figure 13(a) of the original paper, and the wind tunnel error was obtained from Table 1 in its appendix.

F Experimental Setting and Details

Implementation Details To ensure experimental reproducibility, all random number generators were initialized with a fixed seed of 42. Models were trained and evaluated on the DrivAerStar dataset under two configurations: 1 Vehicle (single configuration) and 3 Vehicle (multi-configuration). To assess data efficiency, the training set size was varied across 400, 800, and 1,200 samples, while the validation and test sets were fixed at 150 samples each. All models were trained from scratch for up to 500 epochs to ensure convergence. Training was conducted using the Adam optimizer with an initial learning rate of 0.001, adjusted by a cosine annealing scheduler, and a batch size of 4 and 1 for training and validation, respectively.

Network Architecture: Transolver The Transolver architecture (Wu et al., 2024) consists of 4 encoding layers, each implementing multi-head attention with 4 heads. The model uses a hidden dimension of 64 and maintains an MLP ratio of 1 in its feed-forward networks. The network processes 7-dimensional spatial input data and generates 4-dimensional outputs. The architecture incorporates 5 downsampling operations, utilizes a reference size of 8, and processes 16 data slices in parallel. During training, gradient clipping with a maximum norm of 0.1 was applied to stabilize optimization, while a weight decay of 0.0001 was used for regularization.

Network Architecture: PointNet The PointNet architecture (Qi et al., 2017) processes irregular mesh data through a transformation network (T-Net) and hierarchical feature extraction. The input transformation network generates a 7×7 transformation matrix using three 1D convolutional layers with channel dimensions of 64, 128, and 256. The main network employs sequential 1D convolutional layers with batch normalization and ReLU activation, progressively increasing feature dimensions from 64 to 256 channels. A residual connection bridges the initial 64-channel features to the final representation, while point-wise fully connected layers map the 256-dimensional features to 4-dimensional outputs.

Network Architecture: GNOT The GNOT architecture (Hao et al., 2023) processes irregular mesh data using neural operator principles. The input transformation network processes 7-dimensional features through three 1D convolutional layers (64, 128, and 256 channels), followed by max pooling and fully connected layers to generate transformation matrices. The main network consists of sequential 1D convolutional layers with batch normalization and ReLU activation, progressively increasing feature dimensions from 64 to 128 and finally to 256 channels. A residual connection via 1×1 convolution bridges the initial 64-channel features to the final 256-channel representation. Point-wise fully connected layers conclude the architecture by mapping 256-dimensional features to the 4-dimensional output space.

G Limitations and Future Work

Geometric Representation While our dataset employs 20 deformation parameters, this parameterization remains insufficient to fully characterize complete vehicle geometries. The simplified geometric representation limits the model’s ability to capture fine-scale aerodynamic features that significantly impact flow behavior in real-world applications.

Simulation Fidelity Although our simulations incorporate many industrial practices, they cannot capture all real-world complexities, including dynamic effects, environmental conditions, and manufacturing tolerances. The steady-state RANS approach, while computationally efficient, may not fully resolve transient flow phenomena critical for certain aerodynamic assessments.

Model Performance Validation Our experiments confirm that data-driven scaling laws remain effective, with model performance improving consistently as training samples increase. However, the observed improvements may plateau beyond current dataset sizes, requiring investigation of alternative approaches for continued accuracy gains.

Future Directions Future work will focus on dataset enhancement through expanded geometric parameterization and foundation model development for industrial applications. We plan to explore physics-informed learning methods and data assimilation techniques incorporating wind tunnel experimental data to improve model performance in limited-data scenarios. Additionally, investigation of multi-fidelity approaches combining RANS, LES, and experimental data may enhance prediction accuracy while maintaining computational efficiency for practical automotive design workflows.

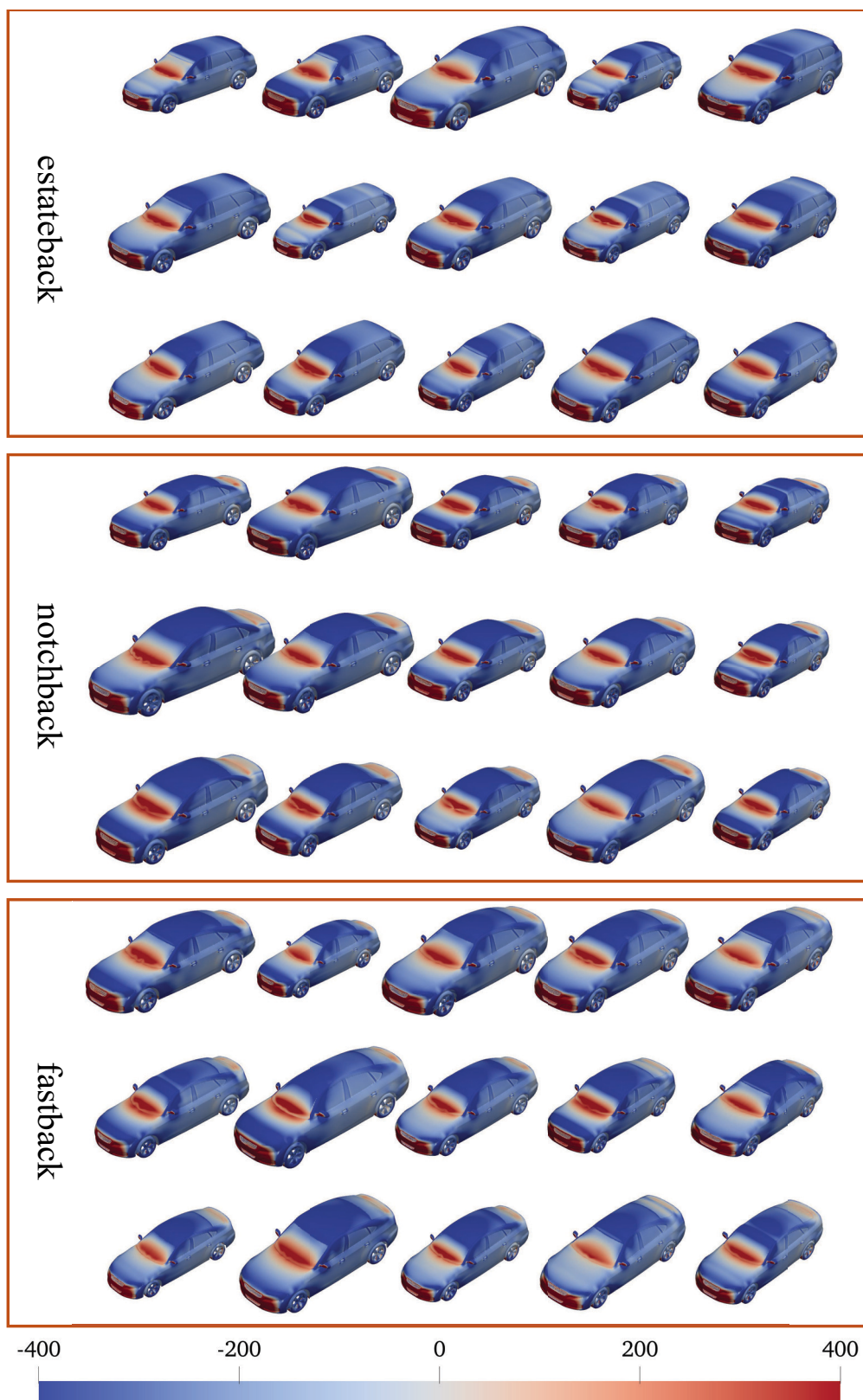


Figure A8: DrivAerStar simulation results: pressure field distributions (Pa).

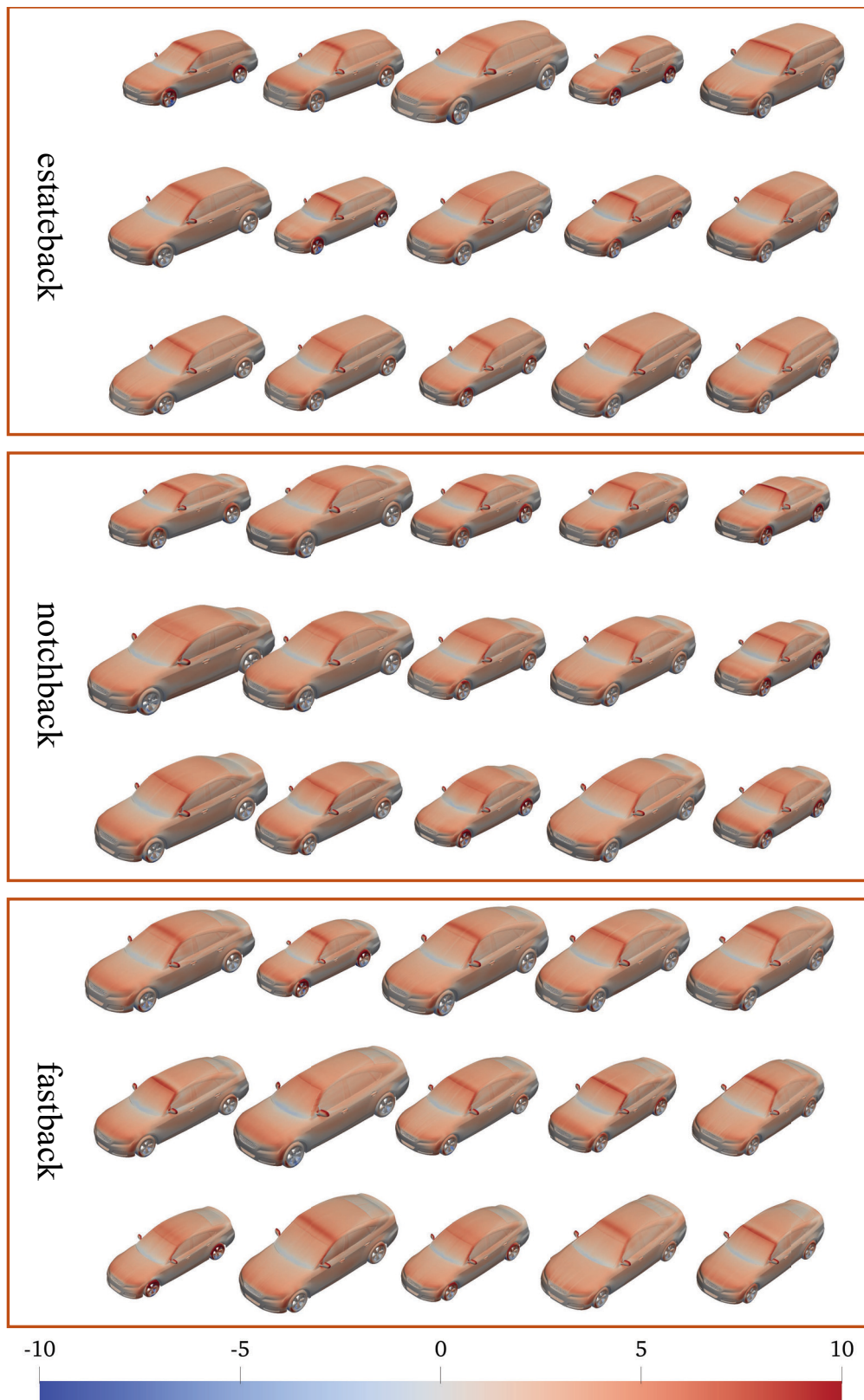


Figure A9: DrivAerStar simulation results: wall shear stress distributions (Pa).

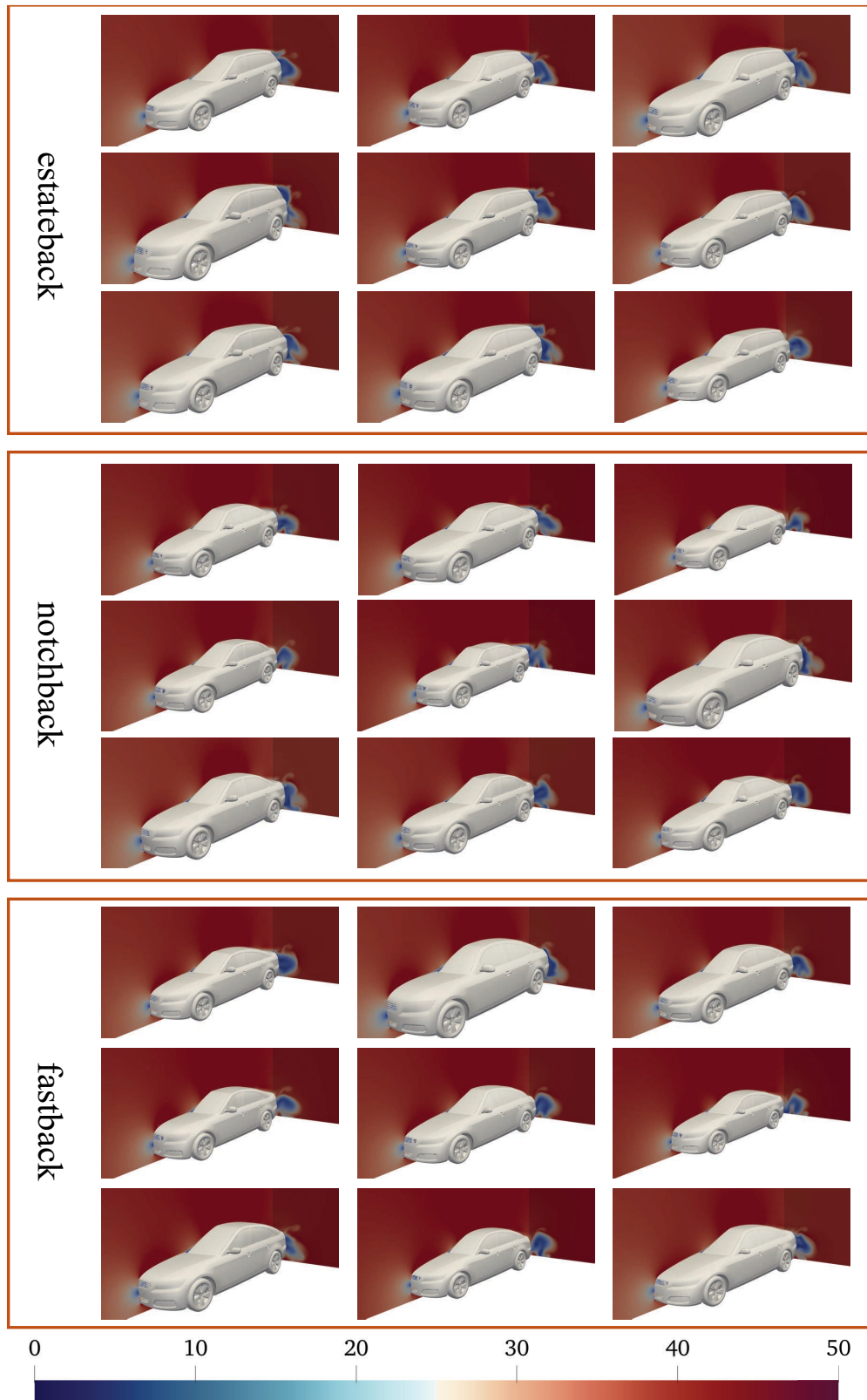


Figure A10: DrivAerStar simulation results: velocity field distributions (m/s)