

Faster by Design: Interactive Aerodynamics via Neural Surrogates Trained on Expert-Validated CFD

Nicholas Thumiger¹, Andrea Bartezzaghi¹, Mattia Rigotti¹, Cezary Skura¹, Thomas Frick¹,
Elisa Seriola², Fabrizio Arbucci², A. Cristiano I. Malossi¹

¹IBM Research

²Dallara Automobili

Abstract

Computational Fluid Dynamics (CFD) is central to race-car aerodynamic development, yet its cost — tens of thousands of core-hours per high-fidelity evaluation — severely limits the design space exploration feasible within realistic budgets. AI-based surrogate models promise to alleviate this bottleneck, but progress has been constrained by the limited complexity of public datasets, which are dominated by smoothed passenger-car shapes that fail to exercise surrogates on the thin, complex, highly loaded components governing motor-sport performance. This work presents three primary contributions. First, we introduce a high-fidelity RANS dataset built on a parametric LMP2-class CAD model and spanning six operating conditions (map points) covering straight-line and cornering regimes, generated and validated by aerodynamics experts at Dallara to preserve features relevant to industrial motorsport. Second, we present the Gauge-Invariant Spectral Transformer (GIST), a graph-based neural operator whose spectral embeddings encode mesh connectivity to enhance predictions on tightly packed, complex geometries. GIST guarantees discretization invariance and scales linearly with mesh size, achieving state-of-the-art accuracy on both public benchmarks and the proposed race-car dataset. Third, we demonstrate that GIST achieves a level of predictive accuracy suitable for early-stage aerodynamic design, providing a first validation of the concept of interactive design-space exploration — where engineers query a surrogate in place of the CFD solver — within industrial motorsport workflows.

Introduction

Computational Fluid Dynamics (CFD) sits at the core of modern race-car aerodynamic development, driving the iterative loop through which teams extract aero metrics such as drag and downforce under tight regulatory and budget constraints. Yet, CFD is computationally expensive: a single high-fidelity steady-state evaluation of a full car typically consumes thousands of core-hours, and meaningful design exploration requires sweeping both geometric parameters — such as wing profiles, endplates, underfloor geometry — and operating conditions, i.e., the map points defining heave, pitch, yaw, roll and steer. The combinatorial cost of this exploration is a fundamental bottleneck of motorsport aerodynamic design, directly limiting the number of concepts feasible to evaluate within time and resource budget.

AI-based surrogate models offer the opportunity to accelerate design cycles by replacing the numerical solver with a learned approximator that predicts flow fields and integrated loads at a fraction of the cost of a CFD run (Azizzadenesheli et al. 2024). Within this space, data-driven approaches, which learn directly from precomputed simulation data rather than embedding the governing equations into the training objective, have emerged as the most practical route for the high-dimensional geometrically complex configurations encountered in industrial aerodynamics. Neural operators in particular have substantially raised the ceiling of surrogate accuracy on PDE-governed problems (Kovachki et al. 2023).

However, in the automotive domain, progress has been constrained by the limited complexity of publicly available CFD datasets. Existing benchmarks are dominated by smoothed passenger-car shapes or academic bluff bodies, none of which contain the thin, complex, highly loaded components (e.g. wings, endplates, strakes, splitters, diffusers) that dominate race-car aerodynamics. Surrogates trained or validated on such data are therefore not exercised in the regimes that matter for motorsport. Moreover, numerical solutions available in the public domain are often not validated by aerodynamics experts. Additionally, existing datasets are typically generated at a single fixed operating condition, whereas industrial aerodynamic development requires sweeping map points — combinations of heave, pitch, yaw, roll, and steer — to characterize vehicle behaviour across the full range of track conditions.

This work presents the following complementary contributions: (i) a high-fidelity CFD dataset spanning six map points across straight-line and cornering conditions, built with the purpose to expose surrogate models to race-car representative geometry and flow features, and developed and validated by aerodynamics domain experts at Dallara, (ii) GIST, a graph-based neural operator model designed to operate on the graphs that arise from such complex geometries, and to scale to the mesh sizes such problems entail, and (iii) empirical evidence that GIST’s accuracy on the race-car dataset is sufficient for early-stage aerodynamic design, pointing toward its adoption within industrial workflows.

We firstly present the dataset, then we briefly introduce the method, and finally we present results obtained on the race-car dataset and an example of how the surrogate model can aid optimization of geometrical design.

Methodology

Dataset

The dataset is generated from a simplified parametric CAD model of a Le Mans Prototype 2 (LMP2) race car, designed in-house by aerodynamics domain experts at Dallara to retain the features that govern vehicle-scale aerodynamics: for example, the rotating wheels encapsulated into the wheelhouse create complex wakes, which interact with the complete underfloor, from front splitter to rear diffuser, in ground effect (Venturi behaviour). To our knowledge, no public CFD dataset on cars combines this level of aerodynamic detail with expert-driven generation and validation.

Each geometric configuration of the parametric CAD model is evaluated at six map points, reproducing both straight-line and cornering track conditions through different combinations of heave, pitch, yaw, roll, and steer. To our knowledge, no existing public automotive CFD dataset covers multiple operating conditions in this way: standard benchmarks fix a single flow condition, which precludes training surrogates capable of predicting aerodynamic behaviour across the map-point space central to race-car development. The car surface is decomposed into 20 part identifiers (PIDs), each corresponding to a distinct aerodynamic component or to special surfaces (to be treated accordingly). Individual PIDs can be treated separately, enabling component-level analysis, or merged into a single watertight surface mesh for full-car surrogate training and inference. The geometric parameters are chosen specifically because they induce large, non-local changes in the flow: e.g., a modification on the top hole gurney of the front wheelhouse reshapes the wheel wake, feeding the underfloor differently, thus producing strong load variations on components that were not themselves modified. The dataset is therefore designed to stress a surrogate model’s ability to capture both the direct effect (loads and surface flow features on the parametrized component) and the indirect effect (load redistribution across the remainder of the car), as typical of the elliptic field defined by the incompressible Navier-Stokes equations. For race-car aerodynamics, where global aerodynamic balance is at least as important as local peak loads, this distinction is essential.

The geometry also includes a radiator modeled as a porous medium with a single inlet and outlet (marked as different PIDs), where a pressure drop across the flow direction is imposed following Darcy’s law. This captures the effect of the cooling system on the external aerodynamics, a coupling that is routinely neglected in public datasets yet materially affects global downforce and drag.

CFD simulations are performed as incompressible steady-state RANS (Pope 2000):

$$\frac{\partial \bar{u}_i}{\partial x_i} = 0$$

$$\bar{u}_j \frac{\partial \bar{u}_i}{\partial x_j} = -\frac{1}{\rho} \frac{\partial \bar{p}}{\partial x_i} + \nu \frac{\partial^2 \bar{u}_i}{\partial x_j \partial x_j} - \frac{\partial}{\partial x_j} (\overline{u'_i u'_j}),$$

where the Reynolds stress tensor ($\overline{u'_i u'_j}$) is modeled according to the Boussinesq hypothesis, which relates the stress to the

Table 1: LMP2 race car dataset sample characteristics.

Average points per sample	~ 4.8M
Average cells per sample	~ 2.4M
Average sample size	360 MB
Map points per configuration	6
Operating conditions	heave, pitch, yaw, roll, steer
Total samples	625

mean strain rate:

$$-\overline{u'_i u'_j} = \nu_t \left(\frac{\partial \bar{u}_i}{\partial x_j} + \frac{\partial \bar{u}_j}{\partial x_i} \right) - \frac{2}{3} k \delta_{ij}.$$

The $k - \omega$ SST turbulence model (Menter 1992), properly customized by Dallara experts for correlation purposes, is chosen to model the kinematic turbulent viscosity ν_t .

The mesh has been created according to the best practices of Dallara experts, preserving geometric features with high-quality elements, realizing a fine modeling of the boundary layer and of the volumetric regions where strong gradients of the solution happen. Second order spatial discretization on the velocity-pressure coupled solver guarantees high accuracy, with convergence evaluated on both residuals and aerodynamics metrics. Target fields stored on the car surface for each sample include mean pressure and mean wall shear stress along the three directional components. Together, these provide sufficient information for both load integration and surface-flow diagnostics.

Every configuration has been reviewed and aero-verified by Dallara aerodynamic engineers against expected load trends, providing an expert-in-the-loop validation step grounded in motorsport domain expertise. A summary of the key dataset parameters is provided in Table 1. Extensions to the dataset are currently in progress, including additional map points and a substantial refinement of the geometry.

Neural operators

Evolution of Neural Operators Two foundational frameworks established the operator-learning paradigm: FNO (Li et al. 2020), which applies spectral convolutions in Fourier space, and DeepONet (Lu et al. 2021), which approximates operators via branch-trunk network pairs. Early data-driven CFD methods built on these ideas: graph network simulators such as MeshGraphNet (Pfaff et al. 2021) demonstrated that learned message passing on unstructured meshes could reproduce complex fluid dynamics, while the Geometry-Informed Neural Operator (GINO, Li et al. 2023) combined graph encoders with FNO-based architectures to handle 3D car shapes. As the field transitioned toward transformer-based architectures, standard attention mechanisms presented a severe quadratic scalability problem when applied to the massive, high-resolution meshes required for vehicle aerodynamics. Transolver mitigates this scalability bottleneck by projecting spatial coordinates onto a reduced set of context tokens. However, this efficiency inherently sacrifices the faithful encoding of the high-fidelity mesh structure, limiting the model’s ability to generalize reliably across varying grid resolutions (Wu

et al. 2024). The Geometry Aware Operator Transformer (GAOT, Wen et al. 2025) similarly combines an attentional graph neural operator encoder with a vision transformer processor and geometry-aware embeddings, achieving strong performance across diverse PDE benchmarks. Like Transolver, however, GAOT relies on Euclidean proximity for graph construction, making it susceptible to unphysical cross-surface interactions on thin-walled geometries.

GIST To resolve the scalability limitations of existing architectures without compromising geometric fidelity, we introduce the Gauge-Invariant Spectral Transformer (GIST), detailed in Rigotti, Thumiger, and Frick (2026).

A surface mesh discretizing the car’s CAD geometry is a graph $G = (V, \mathcal{E})$, where vertices V are the mesh nodes and edges \mathcal{E} encode their connectivity. Treating meshes as graphs rather than unstructured point clouds is critical: it preserves topological information that Euclidean proximity alone cannot capture, such as the fact that two vertices on opposite sides of a thin wing are geometrically close but topologically distant. A neural operator is discretization-invariant if its predictions are consistent across different mesh resolutions and samplings of the same physical surface, a property that most architectures violate in practice because their node features shift non-trivially as the mesh changes. Spectral methods offer a principled route through the normalized graph Laplacian $L = I - D^{-1}A$ (where A is the adjacency matrix of G and D is the diagonal degree matrix), whose spectrum converges to that of the continuous Laplace-Beltrami operator as mesh resolution increases. However, spectral methods face two compounding obstacles in this setting. First, computing the eigenvectors of L via full eigendecomposition costs $\mathcal{O}(N^3)$ for a mesh with N nodes, which is entirely intractable at the $\sim 4.8\text{M}$ -node scale of our dataset. Second, eigenvectors of L suffer from gauge ambiguity: any eigenvector u_k may be arbitrarily sign-flipped or rotated within degenerate eigenspaces, making raw spectral coordinates solver-dependent and therefore unreliable as node features across different meshes.

GIST resolves both problems simultaneously. It avoids eigendecomposition by computing spectral embeddings via random projection (Chen et al. 2019), and it derives attention weights exclusively from gauge-invariant quantities. Given a spectral filter f and a random matrix $R \in \mathbb{R}^{N \times r}$ with $\mathbb{E}[RR^\top] = I$, each node i is embedded as the i -th row of $\tilde{\Phi} = f(P) \cdot R \in \mathbb{R}^{N \times r}$, where $P = D^{-1}A$ is the random-walk matrix of G and r is the embedding dimension. Attention weights are computed from inner products $\langle \tilde{\phi}_i, \tilde{\phi}_j \rangle$, which are unbiased estimators of the gauge-invariant kernel

$$K(i, j) = [f(P)f(P)^\top]_{ij} = \sum_{\ell} f(\lambda_{\ell})^2 (u_{\ell})_i (u_{\ell})_j,$$

where λ_{ℓ} and u_{ℓ} are the eigenvalues and eigenvectors of L , respectively. Since $K(i, j)$ depends only on the operator P and not on any particular eigenvector basis, it is immune to sign flips and eigenspace rotations.

As a direct consequence, GIST admits a formal bound on its discretization mismatch error. For two discretizations G_n and $G_{n'}$ of the same compact m -dimensional Riemannian manifold with $n \leq n'$ nodes, the attention kernel values at

corresponding mesh points satisfy:

$$|\langle \tilde{\phi}_i^n, \tilde{\phi}_j^n \rangle - \langle \tilde{\phi}_i^{n'}, \tilde{\phi}_j^{n'} \rangle| = \mathcal{O}\left(n^{-\frac{1}{m+4}}\right) + \mathcal{O}\left(r^{-\frac{1}{2}}\right).$$

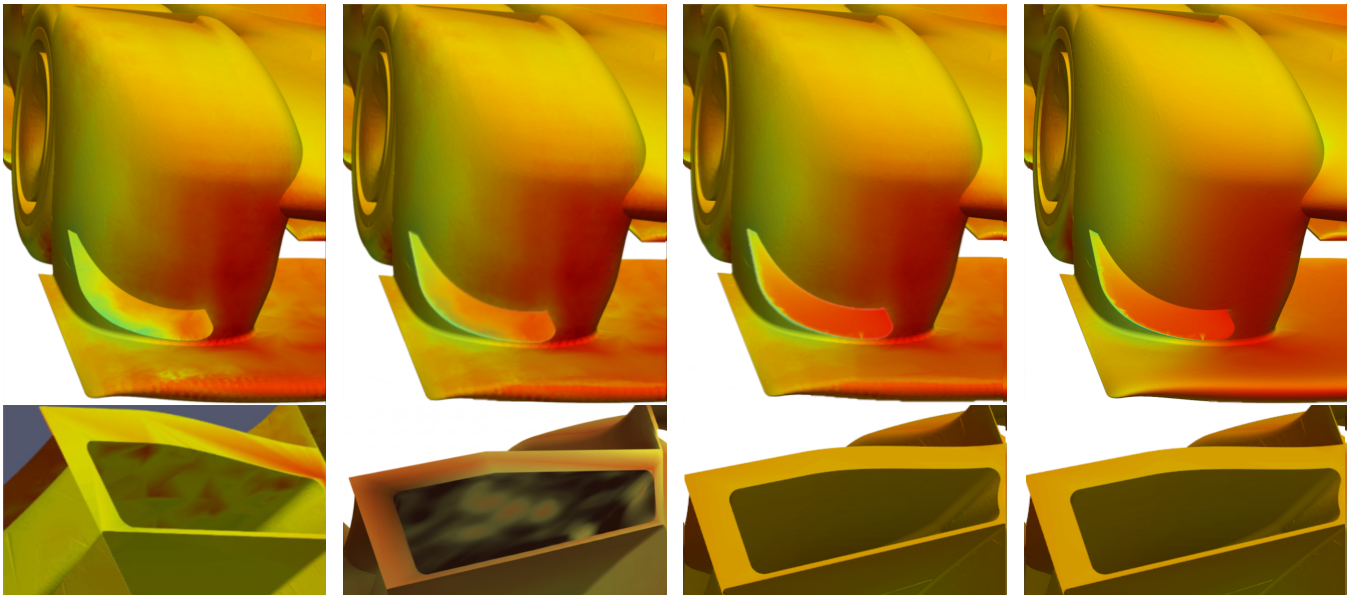
The first term reflects spectral convergence to the Laplace-Beltrami operator and vanishes as the mesh is refined; the second is a statistical error independently controlled by r , the size of the random projection space. This quantified rate is the key theoretical guarantee absent from prior approaches: GIST does not merely claim mesh-invariance heuristically but bounds the error explicitly. Finally, computing $\tilde{\Phi}$ via iterative power methods avoids eigendecomposition and costs $\mathcal{O}(N \cdot r)$, yielding linear scaling with mesh size that makes GIST practical on the $\sim 4.8\text{M}$ -point surface meshes in our dataset. For a more detailed description of the method and the full mathematical formulation we refer the reader to Rigotti, Thumiger, and Frick (2026).

Results

The scarcity of publicly available, complex CFD datasets has hindered the community from identifying core limitations in existing neural operator approaches. A fundamental issue is the common treatment of meshes as simple point clouds, which neglects underlying mesh connectivity. Recent architectures, such as those by (Alkin et al. 2025) and (Wu et al. 2024), rely on the assumption that Euclidean proximity and local metadata (e.g., surface normals) provide a sufficient representation of geometry. However, for complex race-car geometries featuring thin components, two vertices may be spatially proximate yet topologically distant. A primary example is a pair of vertices located on opposite sides of a thin spoiler or flick.

Figure 1 illustrates this phenomenon. The Fourier Neural Operator (FNO) based model utilizes a signed distance function as input to the FNO layer and a graph neural operator (GNO) as a decoder back onto the mesh. This approach introduces a regular error pattern reflecting the underlying FNO grid (visible on the front flick) and produces unphysical predictions on the surface due to interference from invalid points in the regular grid. Similarly, while GAOT and Transolver-based methods produce smoother field predictions, they suffer from poor accuracy on thin surfaces, likely due to unphysical interactions between vertices on opposing sides. By constructing vertex embeddings from the spectral structure of the mesh graph via the gauge-invariant kernel $K(i, j)$, GIST ensures that points close in Euclidean space maintain a large embedding difference if they are topologically distant, precisely the failure mode exposed by thin race-car components such as spoilers and flicks. Consequently, GIST retains the flexibility of a point-cloud representation while rigorously treating the surface as a manifold mesh.

As a result of these considerations and the favorable scalability properties of GIST, significant performance gains are obtained over other methods. To evaluate said performance, we assess both point-wise field accuracy and the resulting integrated aerodynamic loads when inference is run on every point in the geometry. The primary objective of the model is to predict the surface pressure p and wall shear stress τ



(a) FNO-Based Method

(b) Transolver-Based Method

(c) GIST

(d) Ground Truth

Figure 1: Qualitative comparison of the pressure field predicted by different AI-surrogate models on the front area (top row) and on the rear area (bottom row) of the LMP2 race car. The pressure calculated by the GIST model is the closest to the reference CFD ground truth.

at every vertex. We quantify the error using Mean Squared Error (MSE) and the Coefficient of Determination (R^2):

$$\text{MSE} = \frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2, \quad R^2 = 1 - \frac{\sum_i (y_i - \hat{y}_i)^2}{\sum_i (y_i - \bar{y})^2}$$

where N is the number of vertices of the surface mesh, y_i and \hat{y}_i represent the ground truth (CFD) and predicted values at vertex i , respectively, and \bar{y} is the mean of the ground truth values.

Furthermore, the surrogate model must accurately reproduce the integrated loads used in race car development. We compute six aerodynamic coefficients by integrating force and torque contributions over the surface mesh. Without going into details, assuming the surface mesh being discretized in surface mesh elements, the force acting on element k is approximated as:

$$F_k = (\bar{p}_k n_k + \bar{\tau}_{k,\text{wall}}) A_k$$

where \bar{p}_k and $\bar{\tau}_{k,\text{wall}}$ are representative static pressure and wall shear stress vector, respectively, computed from the point-wise predicted values on the vertices of the element k , n_k is the outward-facing normal and A_k is the area. To resolve these forces into the aerodynamic frame, we utilize the wind-aligned coordinate system. The aerodynamic coefficients for drag ($C_x S$), side force ($C_y S$), and downforce ($C_z S$) are calculated by normalizing the integrated contributions:

$$C_{\{x,y,z\}} S = \frac{\sum_k (F_k \cdot e_{\{x,y,z\},\text{wind}})}{0.5 \rho_\infty V_\infty^2},$$

where $\rho_\infty = 1.225 \text{ kg/m}^3$ and $V_\infty = 50 \text{ m/s}$. Torque coeffi-

Table 2: Performance comparison on the Dallara LMP2 dataset. Metrics include MSE (units of 10^{-2}) and R^2 for surface pressure, and Mean Percentage Error (%) for integrated drag (C_d) and downforce (C_l) coefficients. Higher is better for R^2 ; lower is better for all other metrics.

Model	Pressure (p)		Integral Coeffs (% error)	
	MSE \downarrow	R^2 \uparrow	C_d \downarrow	C_l \downarrow
FNO-based	6.82	0.968	4.0%	3.6%
Transolver	4.12	0.980	1.2%	1.2%
GAOT	3.55	0.983	0.9%	1.0%
GIST (ours)	1.88	0.987	0.6%	0.7%

icients ($C_{mx} S, C_{my} S, C_{mz} S$) are similarly integrated as:

$$C_{m\{x,y,z\}} S = \frac{\sum_k (M_k \cdot e_{\{x,y,z\},\text{wind}})}{0.5 \rho_\infty V_\infty^2 L_{ref}}$$

where L_{ref} is the characteristic length and $M_k = (\bar{r}_k - r_0) \times F_k$ is the torque of element k relative to the origin r_0 . For this paper, we present for simplicity the drag and downforce aerodynamic coefficients.

Model results

The results in Table 2 demonstrate GIST's superiority over existing state-of-the-art methods. Baselines such as GAOT and GINO require significant downsampling of the input data via random subsampling to accommodate GPU memory constraints during training. While such downsampling may be negligible for simpler datasets like ShapeNet or DrivAerNet,

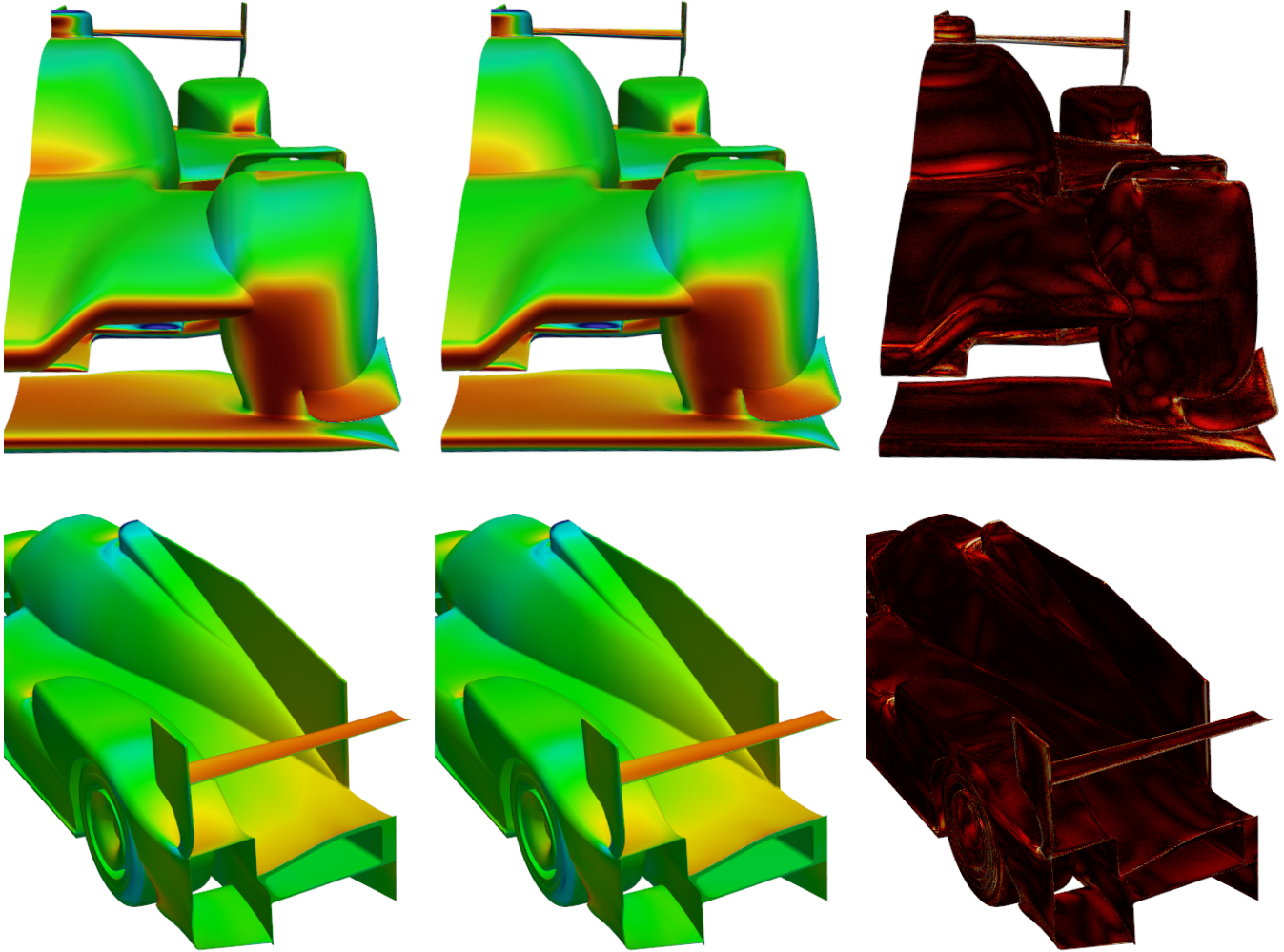


Figure 2: Pressure field distribution on the complete LMP2 race car. Top row: front view. Bottom row: back view. Left: CFD ground truth. Center: GIST prediction. Right: delta.

it has a measurable impact on performance for complex race-car geometries characterized by sharp features and intricate perturbations. In contrast, GIST’s ability to treat the mesh as a connected graph preserves these critical features while simultaneously improving inference efficiency at test time.

From a qualitative standpoint, the model accurately predicts general pressure trends even within high-complexity regions. Consistent with many data-driven methods, GIST exhibits a tendency to smooth high-frequency features, such as steep pressure gradients or localized flow structures resulting from indirect geometric coupling.

For instance, on the upper surface of the splitter—where a vortex is generated via aerodynamic interaction—the model correctly identifies the vortical structure but under-predicts the sharpest pressure gradients at the wingtip. Nevertheless, the surrogate’s ability to capture the global solution profile and the presence of critical aerodynamic features provides

sufficient fidelity for rapid design exploration and iterative optimization.

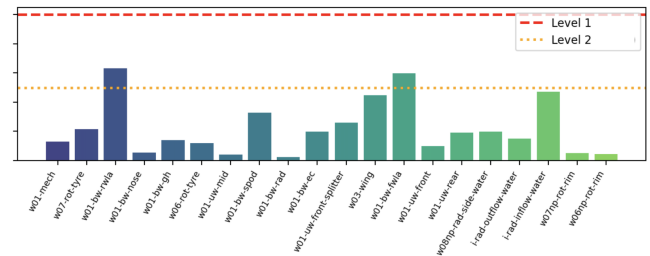


Figure 3: Absolute drag coefficient error calculated as the difference between the GIST model prediction and the CFD ground truth on each of the twenty car PIDs. The usability threshold is met for every part, while the CFD replacement threshold is met by 18/20 parts.

To evaluate the model’s practical utility, Dallara established specific absolute error thresholds in advance for integrating AI into their CFD simulation pipeline. Figure 3 illustrates the average absolute error for the predicted drag coefficient across each component of the vehicle. The red line denotes the “usability threshold”, indicating the accuracy level required for designers to reliably use the surrogate for rapid concept exploration and configuration testing. The yellow line represents a more stringent “CFD-replacement threshold”, at which the surrogate could potentially substitute for high-fidelity simulations in specific design phases. While the specific numerical values are redacted for confidentiality, the results confirm that GIST achieves a performance level suitable for immediate productive use within Dallara’s industrial workflow. The objective is not to replace classical numerical solvers entirely, but to significantly accelerate iterative design cycles.

Study on the optimization of rear diffuser

A key use case for a trained surrogate is rapid parameter sweeps that would be highly expensive with a full CFD campaign. To illustrate this, we study the effect of rear diffuser angle on the two primary design objectives: drag (C_d) and downforce (C_l). Figure 4 shows both coefficients predicted by GIST alongside CFD ground-truth evaluations across the full range of diffuser angles, demonstrating strong agreement between the surrogate and the solver.

The two objectives exhibit competing trends. Drag increases monotonically with angle across the entire range. Downforce magnitude, by contrast, grows steeply for negative and small positive angles but saturates beyond approximately $1.5\text{--}2^\circ$: increasing the angle further delivers diminishing additional downforce. This saturation creates a natural efficiency knee, a region where the aerodynamic gain per unit drag penalty is maximized. A designer can formalize this as optimizing the efficiency $E = |C_l|/C_d$, or equivalently, maximizing $|C_l|$ subject to a drag budget imposed by circuit requirements or regulations. Both formulations identify an optimal angle in the $1\text{--}2^\circ$ range, beyond which added downforce no longer justifies the drag cost.

Crucially, GIST produces this full sweep in seconds on a single GPU, compared to the tens of thousands of core-hours an equivalent CFD campaign would require. This acceleration transforms diffuser angle selection from an offline batch study into an interactive design exploration, enabling engineers to respond to regulation changes or new circuit requirements in near real time.

Next steps

The Dallara dataset enables research trajectories impossible with current public benchmarks. Our focus is moving from static prediction to a dynamic, physics-aware design tool by leveraging the unique global and local boundary conditions inherent in our data. Unlike standard datasets that focus on fixed operating conditions, our work utilizes “map points”—simulations covering varying heave, pitch, yaw, roll, and steer at multiple wind speeds. This allows for training models conditioned on diverse flow regimes, predicting aerodynamic

performance during complex transient maneuvers like high-speed cornering or heavy braking.

Our most immediate and achievable objective is the integration of internal cooling aerodynamics. While academic datasets typically treat car bodies as hollow shells, each model in our dataset includes a fully functional internal radiator modeled via Darcy-flow equations. Because these porous medium labels are already present in our data, we can uniquely train the surrogate to account for “cooling drag” and its impact on the performance. This provides a real-time decision-making tool for the trade-off between thermal management and aerodynamic efficiency—a feature set currently unsupported by any existing AI models in the literature.

Principled uncertainty quantification is another promising direction. Equipping GIST with epistemic uncertainty estimates (Frick et al. 2024) would identify regions of the design space where predictions are less reliable, while distribution-free conformal prediction intervals (Deutschmann, Rigotti, and Rodriguez Martinez 2024) could provide statistically guaranteed coverage on integrated loads, giving engineers a rigorous criterion for when to defer to the full CFD solver.

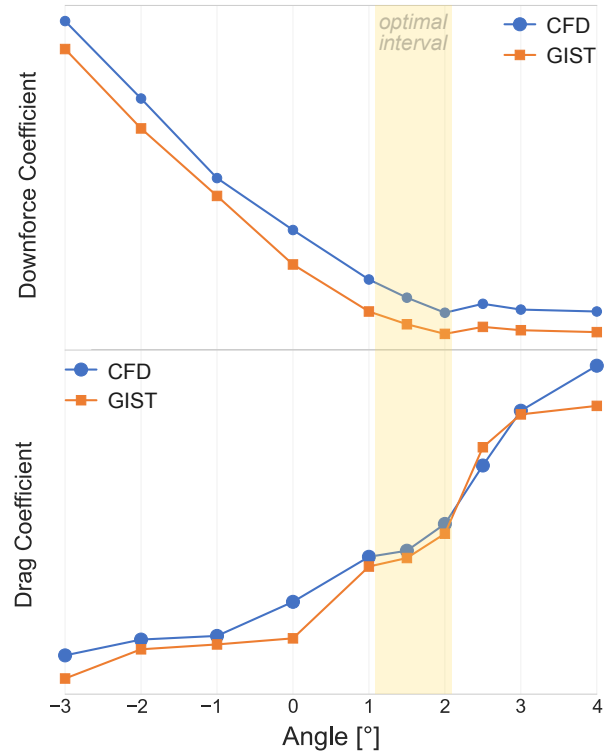


Figure 4: Downforce (top) and drag (bottom) coefficients as a function of rear diffuser angle. GIST predictions (orange) closely track CFD ground truth (blue) across the full sweep range. The shaded region highlights the optimal interval ($1\text{--}2^\circ$), where the downforce-to-drag ratio is maximized before downforce saturates. Note that downforce is a negative force, thus lower values mean more downforce. Values at 4° are out of domain for the AI model: no training samples were produced near that value.

Embedding interpretability into the GIST transformer, for instance through concept-based attention mechanisms (Rigotti et al. 2022), could further ground predictions in human-readable aerodynamic concepts and surface non-obvious cross-component couplings, deepening engineer trust and guiding targeted design exploration.

Finally, we are leveraging the differentiability of the GIST architecture to develop an Inverse Modeling framework. By propagating gradients from target performance coefficients back to input geometries, the model can automatically suggest optimal tweaks to components, from the diffuser angle to more complex and nonlinear areas. To ensure industrial reliability, we aim to bridge the “sim-to-real” gap using a multi-fidelity approach, fine-tuning the RANS-trained model with sparse, high-accuracy wind tunnel and track data. This calibration allows the surrogate to correct for inherent RANS limitations, such as flow separation inaccuracies, aligning AI predictions with the physical reality measured on the ground.

Conclusions

This work addresses the computational bottleneck of motorsport aerodynamic development by introducing a high-fidelity, expert-validated simplified-LMP2 dataset and the Gauge-Invariant Spectral Transformer (GIST). By explicitly encoding mesh connectivity through spectral embeddings, GIST overcomes the limitations of point-cloud-based surrogates, accurately capturing the complex flow physics of thin, highly loaded race-car components. Our results demonstrate that GIST not only achieves state-of-the-art performance on public benchmarks but also meets the rigorous accuracy thresholds required for industrial deployment at Dallara. Ultimately, this work provides a scalable, discretization-invariant foundation for the next generation of aerodynamic design tools, transitioning from passive field prediction toward dynamic, physics-aware optimization and real-time “sim-to-real” calibration.

References

Alkin, B.; Bleeker, M.; Kurle, R.; Kronlachner, T.; Sonnleitner, R.; Dorfer, M.; and Brandstetter, J. 2025. AB-UPT: Scaling neural CFD surrogates for high-fidelity automotive aerodynamics simulations via anchored-branched universal physics transformers. *arXiv preprint arXiv:2502.09692*.

Azizzadenesheli, K.; Kovachki, N.; Li, Z.; Liu-Schiaffini, M.; Kossaifi, J.; and Anandkumar, A. 2024. Neural operators for accelerating scientific simulations and design. *Nature Reviews Physics*, 6: 320–328.

Chen, H.; Sultan, S. F.; Tian, Y.; Chen, M.; and Skiena, S. 2019. Fast and Accurate Network Embeddings via Very Sparse Random Projection. In *Proceedings of the 28th ACM International Conference on Information and Knowledge Management*.

Deutschmann, N.; Rigotti, M.; and Rodriguez Martinez, M. 2024. Adaptive Conformal Regression with Split-Jackknife+ Scores. *Transactions on Machine Learning Research*.

Frick, T.; Antognini, D.; Giurghi, I.; Grewe, B.; Malossi, C.; Zhu, R. J.; and Rigotti, M. 2024. MC Layer Normalization for calibrated uncertainty in Deep Learning. *Transactions on Machine Learning Research*.

Kovachki, N.; Li, Z.; Liu, B.; Azizzadenesheli, K.; Bhattacharya, K.; Stuart, A.; and Anandkumar, A. 2023. Neural operator: Learning maps between function spaces with applications to PDEs. *Journal of Machine Learning Research*, 24(89): 1–97.

Li, Z.; Kovachki, N.; Azizzadenesheli, K.; Liu, B.; Bhattacharya, K.; Stuart, A.; and Anandkumar, A. 2020. Fourier neural operator for parametric partial differential equations. *arXiv preprint arXiv:2010.08895*.

Li, Z.; Kovachki, N.; Choy, C.; Li, B.; Kossaifi, J.; Otta, S.; Nabian, M. A.; Stadler, M.; Hundt, C.; Azizzadenesheli, K.; et al. 2023. Geometry-informed neural operator for large-scale 3d pdes. *Advances in Neural Information Processing Systems*, 36: 35836–35854.

Lu, L.; Jin, P.; Pang, G.; Zhang, Z.; and Karniadakis, G. E. 2021. Learning nonlinear operators via DeepONet based on the universal approximation theorem of operators. *Nature Machine Intelligence*, 3: 218–229.

Menter, F. R. 1992. Improved two-equation k-omega turbulence models for aerodynamic flows. Technical report, NASA Technical Memorandum.

Pfaff, T.; Fortunato, M.; Sanchez-Gonzalez, A.; and Battaglia, P. W. 2021. Learning mesh-based simulation with graph networks. In *International Conference on Learning Representations*.

Pope, S. B. 2000. *Turbulent Flows*. Cambridge University Press.

Rigotti, M.; Miksovic, C.; Giurghi, I.; Gschwind, T.; and Scotton, P. 2022. Attention-based Interpretability with Concept Transformers. In *International Conference on Learning Representations*.

Rigotti, M.; Thumiger, N.; and Frick, T. 2026. GIST: Gauge-Invariant Spectral Transformers for Scalable Graph Neural Operators. *arXiv preprint arXiv:2603.16849*.

Wen, S.; Kumbhat, A.; Lingsch, L.; Mousavi, S.; Zhao, Y.; Chandrashekar, P.; and Mishra, S. 2025. GAOT: Geometry Aware Operator Transformer. *arXiv preprint arXiv:2505.18781*.

Wu, H.; Luo, H.; Wang, H.; Wang, J.; and Long, M. 2024. Transolver: A fast transformer solver for pdes on general geometries. *arXiv preprint arXiv:2402.02366*.