Proceedings of the Seventh International Conference on
Railway Technology: Research, Development and Maintenance
Edited by: J. Pombo
Civil-Comp Conferences, Volume 15, Paper 5.6
Civil-Comp Press, Edinburgh, United Kingdom, 2026
ISSN: 2753-3239, doi: 10.4203/ccc.15.5.6

Optimisation of Sparse Sensor Placement and Real-Time Pressure Field Prediction for High-Speed Trains

W. Qiu[1,2], Z. Sun[1], D. Guo[1] and G. Yang[1]

[1]Institute of Mechanics, Chinese Academy of Sciences, Beijing, China
[2]School of Engineering Science, University of Chinese Academy of Sciences, Beijing, China

Abstract

This paper presents a rapid prediction framework for the real-time reconstruction of surface pressure fields on complex three-dimensional high-speed trains using extremely sparse sensor data. The methodology integrates generalised proper orthogonal decomposition, compressive sensing, and an improved particle swarm optimisation algorithm. By extracting the dominant spatial modes of the flow field, the number of significant coefficients required for reconstruction is reduced to fewer than ten, fundamentally minimising the required physical sensors. The optimisation algorithm successfully configures a minimal sensor array of seven pressure measurement points on the leading car. The framework accurately reconstructs the surface pressure field, achieving an average computation time of 0.0033 seconds per case and an average relative reconstruction error of three per cent. Evaluations across varied operating conditions confirm high fidelity and aerodynamic rationality. By bypassing the need to resolve the entire volumetric flow field, the proposed approach delivers a transformative reduction in computational effort compared to traditional computational fluid dynamics, providing a robust foundation for real-time aerodynamic load calculation and operational safety monitoring.

Keywords: high-speed trains, surface pressure field, compressive sensing, proper orthogonal decomposition, particle swarm optimisation, sparse sensor placement, flow field reconstruction.

1 Introduction

The CR450 electric multiple unit (EMU) completed its prototype roll-out and joint tests in China in July 2025, with an operating speed of 400 km/h. At this speed, aerodynamic load characteristics have become the core factor dominating the research and development, operational safety, and energy efficiency performance of high-speed trains, with aerodynamic drag accounting for more than 90% of the total drag. To accurately obtain the aerodynamic load performance of high-speed trains, academia and industry mainly rely on three major research methods: numerical simulation, model testing, and full-scale vehicle testing. These methods each have advantages and disadvantages, and together they constitute the research technology system for high-speed train aerodynamic loads.

The numerical simulation method uses computers to build appropriate models to solve the Navier-Stokes equation, thereby accurately simulating the flow field. Its advantage is that it can perform flexible calculations for high-speed trains of different shapes and environmental conditions, yielding a series of physical quantities with precise spatial and temporal distributions. Its disadvantage is that it is highly dependent on computing power, and the calculation results are significantly affected by grid quality, calculation models, and empirical parameters.

Model testing is based on the principle of similarity, using scaled-down models to simulate the airflow around high-speed trains in a controlled test environment. Its advantage is that the test results are more reliable than numerical simulations. Its disadvantages are the high costs of model production and test equipment, the spatially sparse measurement results, and the time-consuming and labour-intensive process.

Full-scale vehicle testing involves testing the actual train in a real operating environment. Its measurement results are the most reliable, but the testing cycle is long and the cost is extremely high. To obtain surface pressure information, 240 pressure monitoring points need to be arranged on the lead car alone [1]. Furthermore, the acquired data are limited to sparse measurement locations, making it impossible to obtain complete full-field information.

A flow field reconstruction technique proposed in recent years can reconstruct a high-resolution flow field from sparse, low-resolution input data, thereby enabling precise monitoring of the train surface pressure distribution during actual operation. In particular, obtaining a high-precision pressure field on the surface of a high-speed train is of great significance for accurate experimental measurements under complex operating conditions [2]. Meanwhile, as an auxiliary measurement technology, it can be utilised in the experimental verification stage of various flow control technologies [3][4].

Inoba et al. [5] proposed a method to predict the surface pressure distribution of a simple car model under the action of wind direction through data-driven optimisation of sparse pressure sensors. However, the obtained points are still arranged symmetrically around the model, which may not be the most suitable under crosswind conditions. In addition, using the least squares method to reconstruct surface pressure is not applicable to high-speed trains with a high slenderness ratio and high operating speed. Therefore, there is an urgent need to develop a novel high-precision method for predicting the surface pressure field from sparse measurement data.

To solve the difficulty of predicting the full-car surface pressure field using sparse measurement data, this paper proposes a rapid prediction framework for the surface pressure field suitable for high-speed trains based on the compressive sensing (CS) method. This framework obtains a sparse representation of the complex flow field by performing generalised proper orthogonal decomposition (GPOD) on the training data, and then solves for a small number of large coefficients through compressive sensing, thereby achieving an accurate reconstruction of the pressure field. At the same time, an improved particle swarm optimisation (IPSO) algorithm is adopted to optimise the layout scheme of the sparse measurement points, reducing redundant sensors that rely on manual experience and significantly lowering the measurement cost of real-train tests.

This paper is organized as follows: Section 2 introduces the main methods and theoretical basis used in the rapid prediction framework. Section 3 compares and analyses the pressure field reconstruction results of the high-speed train. Section 4 summarises the main contributions of this paper.

2 Methods

2.1 Computational Domain and Grid Configuration

In Figure 1, the 1:1 scale simplified three-car CRH380A model utilised for the numerical simulation in this study is shown. It consists of three carriages: a leading car, a middle car, and a trailing car, each with a length of 25 m. The train model has a height of H = 3.55 m, a maximum width of W = 3.38 m, and a maximum cross-sectional area of S = 11.497 m^2. The streamlined nose cones of both the leading and trailing cars have a length of 12 m. As this study solely aims to verify the feasibility of applying the rapid prediction framework to objects with complex three-dimensional external structures, such as high-speed trains, the simulation model omits detailed appendages including bogies, air-conditioning compartments, and pantographs.

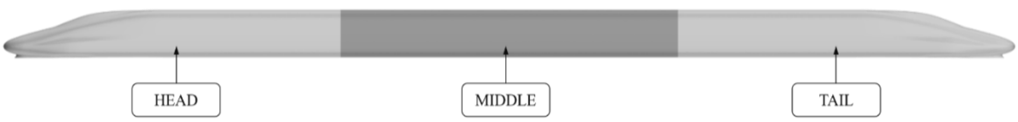


Figure 1: Simplified computational model.

In Figure 2, a schematic of the computational domain is shown. To ensure the flow field fully develops and to minimise boundary effects, the longitudinal (X-axis) distance from the front of the leading car to the inlet boundary is approximately 40 m, and the distance from the rear of the trailing car to the outlet is roughly 65 m. The total domain dimensions are 180 m in length, 100 m in width, and 40 m in height. To accurately capture the leeward flow under crosswind conditions, the train is positioned asymmetrically in the lateral (Y-axis) direction, with upstream and downstream distances set to 40 m and 60 m, respectively.

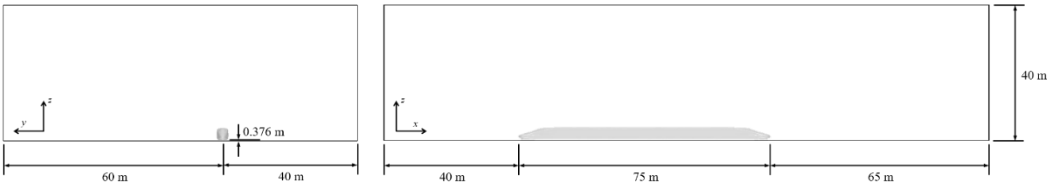


Figure 2: Computational domain.

Velocity inlet conditions are applied to the front and left boundaries of the domain, with the longitudinal velocity matching the train's operating speed and the lateral velocity equalling the crosswind speed. Conversely, zero-pressure outlet conditions are assigned to the rear and right boundaries. A no-slip wall condition is imposed on both the train and ground surfaces. To accurately simulate the relative ground effect, the train remains stationary while the ground surface translates at the longitudinal running speed. Finally, a symmetry boundary condition is applied to the top of the domain.

In Figure 3 and Figure 4, the computational mesh generated using the commercial software STAR-CCM+ is shown. The computational domain is divided into five rectangular refinement zones to capture complex flow features, with base mesh sizes of 0.04 m, 0.08 m, 0.16 m, 0.32 m, and 0.64 m, respectively. To accurately resolve the boundary layer, 16 prism layers are extruded from the train surface with a growth ratio of 1.2. The first layer thickness is set to 4×10^{-4} m, ensuring the dimensionless wall distance satisfies $30 < y^{+} < 100$. This study utilises a high-y^{+} wall treatment, employing wall functions to approximate near-wall flow and ensuring the first grid cell lies within the logarithmic region of the boundary layer. The total mesh comprises approximately 55 million cells.

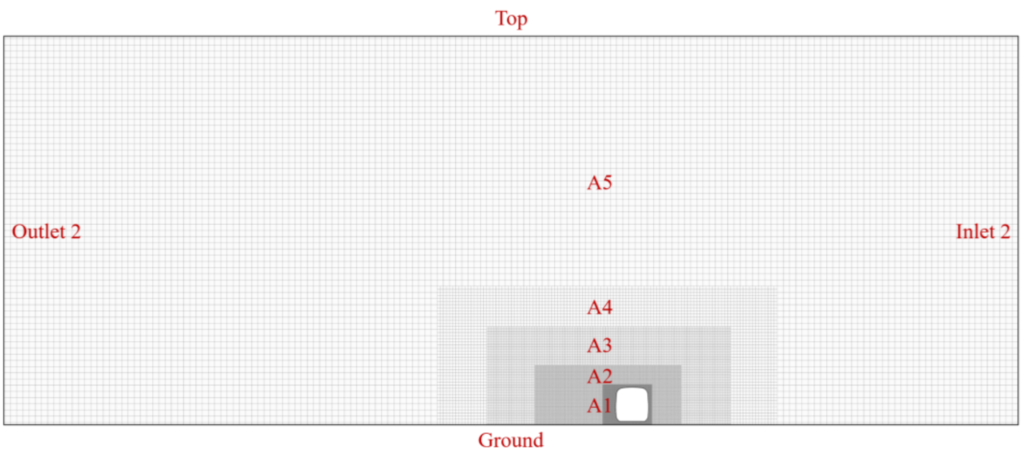


Figure 3: Side view of the computational mesh.

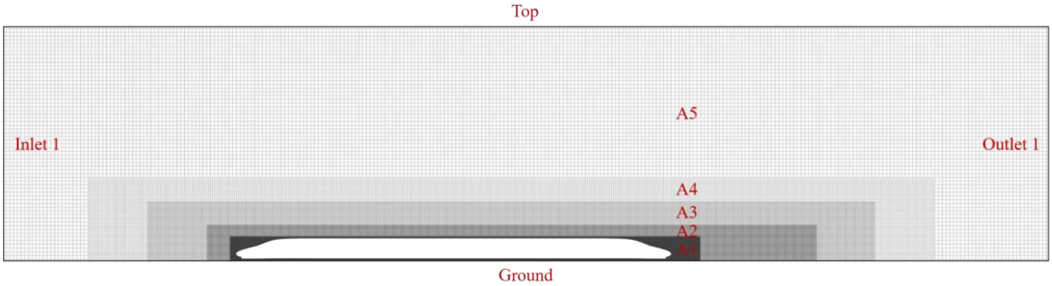


Figure 4: Front view of the computational mesh.

2.2 Numerical Algorithm and Validation

As the maximum operating speed in this study is 350 km/h, yielding a Mach number below 0.3, the flow field is treated as incompressible. The fluid motion is governed by the three-dimensional incompressible Navier-Stokes (N-S) equations. To close the governing equations, the improved delayed detached eddy simulation (IDDES) method, based on the shear-stress transport (SST) model, is employed [6]. Within this hybrid framework, the Reynolds-averaged Navier-Stokes (RANS) approach is applied in the near-wall region to model small-scale turbulent fluctuations, whereas large eddy simulation (LES) is utilised away from the wall to resolve the motion of large separation vortices. Furthermore, the IDDES formulation inherently mitigates grid-induced separations and accurately accounts for wall shear stress.

The aerodynamic drag coefficient (C_d) is defined as follows:

$$C_d = \frac{2F_x}{\rho U^2 S}, \tag{1}$$

where F_x is the streamwise aerodynamic force, computed by integrating both normal pressure and shear stresses over the train surface. The parameter ρ represents the air density (1.225 kg/m^3), U is the incoming flow velocity, and S denotes the maximum cross-sectional reference area (11.497 m^2).

To validate the mesh strategy and numerical solver, simulations are compared against wind tunnel tests of a 1:8 scaled simplified CRH380A train model. Both methods evaluate an open-air operating speed of 300 km/h. A comparison between the numerical results and the experimental data is given in Table 1. Minor geometric discrepancies between the numerical and physical models account for the slight deviations observed. Nevertheless, the numerical predictions align well with the experimental measurements, yielding an overall drag coefficient error of approximately 2.52%. This confirms that the adopted numerical methodology provides accurate aerodynamic data for high-speed trains.

	C_d of leading car	C_d of middle car	C_d of trailing car	C_d of total
Wind tunnel experiment	0.1248	0.0819	0.1194	0.3261
Computational fluid dynamics	0.1206	0.0827	0.1147	0.3180
Error	3.37%	1.00%	3.94%	2.52%

Table 1: Comparison between wind tunnel experimental results and Computational fluid dynamics results.

2.3 Dataset Construction

Following the validation of the numerical methodology, a series of simulations was conducted to generate the pressure field dataset. The operating conditions encompassed running speeds from 270 km/h to 350 km/h and crosswind speeds from 0 m/s to 15 m/s. The resulting dataset was randomly partitioned into a training set and a test set at a 4:1 ratio. The training set is utilised to extract spatial modes via GPOD and to evaluate fitness during the IPSO process. Subsequently, the test set is employed to assess the reconstruction accuracy of the proposed framework.

To maintain computational efficiency, the reconstruction process for the case study exclusively utilises data from the leading car surface, comprising 56,564 spatial measurement points. This targeted approach facilitates the investigation of various control parameters on the CS reconstruction without incurring prohibitive computational costs. For data structuring, the surface pressure field for each operating condition is flattened into a one-dimensional column vector. Because the spatial coordinates of the mesh nodes remain identical across all cases, these individual vectors from the training set are concatenated horizontally. This process yields a two-dimensional snapshot matrix, structured as spatial dimensions by operating conditions, which serves as the foundational input for the dimensionality reduction framework.

2.4 Compressed sensing

Compressed sensing, initially proposed in 2006 by E.J. Candes, J. Romberg, T. Tao, and D.L. Donoho et al. [7][8], has demonstrated that when an original signal possesses sparsity or compressibility, it is possible to reconstruct the signal from randomly subsampled data, well below the sampling rate required by the Nyquist-Shannon sampling theorem.

In mathematics, a compressible signal $\mathrm{x} \in \mathbb{R}^{m}$ can be represented as a sparse vector $\mathrm{s} \in \mathbb{R}^{t}$ in a new sparse transform basis $\boldsymbol{\Psi} \in \mathbb{R}^{m\times t}$, taking the form of:

$$\mathrm{x} = \boldsymbol{\Psi}\mathrm{s}. \tag{2}$$

A measurement matrix $\boldsymbol{\Phi}$ can be constructed based on the sensor positions by initializing a d-by-m zero matrix, where each row corresponds to a sensor. The original signal is then sampled into a set of measurements $\boldsymbol{y}$ with length d through the measurement matrix $\boldsymbol{\Phi}$, satisfying the relationship:

$$y = \boldsymbol{\Phi}x = \boldsymbol{\Phi\Psi}s = \boldsymbol{\Theta}s. \tag{3}$$

When $d < m$, the system in Equation (3) is underdetermined, thereby admitting infinitely many solutions. Solutions derived via the least squares method are typically non-sparse. Consequently, under limited measurements, this conventional approach fails to capture all essential characteristic values, yielding poor reconstruction accuracy. However, by leveraging the inherent high sparsity of natural signals, it is possible to identify the sparsest coefficient vector $\boldsymbol{s}$ that remains consistent with the measurement vector $\boldsymbol{y}$. The original high-dimensional signal is subsequently recovered through an inverse sparse transform. In Figure 5, the mathematical framework of compressive sensing [9] is shown.

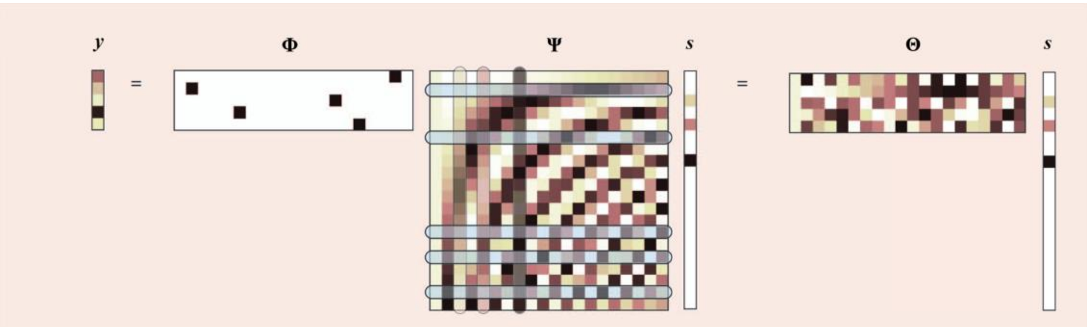


Figure 5: Framework of the compressive sensing method.

The specific equation for solving s is as follows:

$$s = \underset{s'}{arg\,min} \left\|s'\right\|_0, \ s.t. \ \ y = \boldsymbol{\Theta}s' = \boldsymbol{\Phi\Psi}s'. \tag{4}$$

However, the optimization problem associated with the l_0 norm is highly intractable, requiring a search over all possible sparse vector combinations, which is an NP-hard problem. A significant innovation in compressed sensing technology is the allowance to relax this problem to a convex optimization problem with the l_1 norm minimization.

$$s = arg\min_{s'} \|s'\|_1, \; s.t. \;\; y = \mathbf{\Theta}s' = \mathbf{\Phi\Psi}s'. \tag{5}$$

Compressive sensing substantially reduces the required number of measurement sensors from m to t ($t << m$). This reduction significantly lowers the deployment costs associated with model tests and full-scale vehicle tests, while simultaneously mitigating the risk of sensor failure. To ensure high fidelity in the reconstructed signal, the judicious selection of the measurement matrix $\mathbf{\Phi}$ and the sparsity matrix $\mathbf{\Psi}$ is of paramount importance. In the present study, the sparsity matrix is constructed from the dominant spatial modes extracted via GPOD. Concurrently, the measurement matrix is optimised using the IPSO algorithm, and the signal reconstruction is executed employing the basis pursuit (BP) algorithm.

2.5 Improved Particle Swarm Optimisation

The particle swarm optimisation (PSO) algorithm is a population-based stochastic optimisation metaheuristic, initially introduced by Kennedy and Eberhart in 1995 [10]. Rather than relying on gradient information, the algorithm navigates the search space through collective information sharing and individual cooperation among a swarm of candidate solutions, termed particles, to locate the global optimum. Building upon this foundation, various dynamic theoretical analyses [11], adaptive parameter tuning [12], topological structure enhancements, and hybridisations with other evolutionary algorithms have yielded numerous algorithmic variations.

Given its minimal adjustable parameters, robust global search capabilities, and relative ease of implementation, this study proposes an improved particle swarm optimisation (IPSO) method to optimise the layout of the pressure sensors. To enhance the convergence rate and search efficiency, the IPSO incorporates advanced search strategies and a novel fitness function. The variables involved in the IPSO algorithm and their corresponding meanings are given in Table 2.

Control Parameter	Definition
i, j, k, h	Variable used for iteration
m	Total number of surface control points
n	Number of operating conditions
n_{train}	Number of the training set
num	Number of the particles
d	Number of measurement points currently used
$h_{\max}$	Maximum number of iterations
$d_{\max}$	Maximum number of measurement points
$v_{\max,\text{initial}}$	Initial maximum speed of the particles
$v_{\max,\text{final}}$	Final maximum speed of the particles
w_{initial}	Initial inertia weight factor
w_{final}	Final inertia weight factor
η_1, η_2	Individual and social learning factors
r_1, r_2	Random numbers uniformly distributed between [0, 1]
ρ_{train}	Proportion of training conditions sampled when calculating the fitness function
$\varepsilon_{\text{limit}}$	Required convergence error
$\boldsymbol{x}_i=[x_{i,1},x_{i,2},\ldots,x_{i,d}]$	Position state of the i-th particle
$\boldsymbol{v}_i=[v_{i,1},v_{i,2},\ldots,v_{i,d}]$	Velocity state of the i-th particle
$\boldsymbol{p}_{\text{rec}}$	Reconstructed pressure
$\boldsymbol{p}_{\text{cfd}}$	Pressure obtained from CFD
e_k	Relative error of the k-th control point
ε_j	Relative reconstruction error of the j-th operating condition
c	Scaling factor
f_i	Fitness state of the i-th particle

Table 2: Variable definitions in the IPSO algorithm.

The fundamental PSO process entails iteratively updating the local optimal position and fitness of each particle alongside the global optimal position of the entire swarm. This iterative procedure concludes upon satisfying a predefined convergence criterion, thereby yielding the optimal solution. The velocity state of each particle is updated as follows:

$$v_i^h = w \times v_i^{h-1} + \eta_1 \times r_1 \times \left(\text{xbest\_local}_i^{h-1} - x_i^{h-1}\right) + \eta_2 \times r_2 \times \left(\text{xbest\_global}_i^{h-1} - x_i^{h-1}\right), \tag{6}$$

where xbest_local denotes the local optimal position and xbest_global represents the global optimal position.

The proposed IPSO framework advances the standard PSO by introducing a linearly decaying inertial weight factor and maximum particle velocity as a function of the iteration number. This modification ensures enhanced global exploration during the initial stages and promotes rapid convergence in the later stages:

$$\begin{cases} w^{h} = w_{\text{initial}} - \frac{h}{h_{max}(w_{\text{initial}} - w_{\text{final}})} \\ h_{\text{max,initial}}\frac{h}{h_{max}\left(v_{\text{max,initial}} - v_{\text{max,final}}\right)}\{ \\ v_{max} \end{cases} \tag{7}$$

Furthermore, a kinematic collision model is integrated. When a particle intersects the search domain boundary, its velocity is reversed and attenuated, thereby preventing the non-physical clustering of particles at the domain extremities:

$$v_i^h = \left\{ v_i^h, \quad v_i^h \in \left[-v_{max}^{h_{max}^h}, \right| -0.1 v_i^h, \quad v_i^h \notin \left[-v_{max}^{h_{max}^h}. \right] \{ \right] \tag{8}$$

The surface pressure field of high-speed trains exhibits pronounced spatial heterogeneity. Extensive areas, such as the straight carriage sections, are characterised by low pressure gradients, whereas localised zones, including the nose tip and leeward regions, demonstrate significantly elevated gradients. Although these high-gradient regions occupy a minimal surface area, they decisively influence the overall aerodynamic performance of the vehicle. Consequently, traditional global error metrics, such as the mean squared error (MSE) and the mean absolute error (MAE), exhibit distinct limitations when evaluating the reconstruction accuracy. These metrics fail to account for spatial gradient distributions, causing significant reconstruction errors in high-gradient zones to be numerically diluted by the minor errors in expansive low-gradient regions through area-weighted averaging. Furthermore, these functions are highly sensitive to the original data scale; thus, a numerically small global error may still correspond to marked deviations from the true pressure field in critical high-gradient regions.

To address this limitation, a statistically robust fitness function is implemented. For the j-th operating condition, the relative reconstruction error, ε_j, is defined as the ratio of the number of control points where the relative discrepancy between the reconstructed and the computational fluid dynamics (CFD) computed pressure coefficients exceeds 5% to the total number of surface control points. To mitigate overfitting, a random subset of the training conditions, denoted by the proportion ρ_{train}, is selected for each fitness evaluation. A scaling factor, c, is incorporated to ensure the aggregate fitness function converges towards the target error exclusively when the majority of the training conditions satisfy the predefined threshold. This formulation directs the IPSO algorithm to minimise the relative reconstruction error across the entire geometry, yielding highly generalisable optimisation results. The specific mathematical formulations are as follows:

$$c_j = \begin{cases} 1, & \varepsilon_j \leq \varepsilon_{limit} \\ 0, & \varepsilon_j > \varepsilon_{limit} \end{cases} \tag{9}$$

$$c = \frac{1}{n_{\text{train}} \times \rho_{\text{train}} + 1} \left(1 + \sum_{j=1}^{n_{\text{train}} \times \rho_{\text{train}}} c_j\right), \tag{10}$$

$$f_i = \frac{1}{c \times n_{\text{train}}} \sum_{j=1}^{n_{\text{train}}} \varepsilon_j. \tag{11}$$

3 Results

In this section, the proposed framework is applied to reconstruct the sparse surface pressure of the aforementioned train model under varying crosswind conditions. To validate the feasibility of this methodology, the analysis is strictly confined to the surface pressure field reconstruction of the leading car.

The first 10 spatial modes obtained from the GPOD process serve as the sparsity matrix, while the set of IPSO results, comprising 7 pressure sensors, constitutes the measurement matrix. In Figure 6, the specific measurement positions are shown. The 7 measurement points derived from the IPSO optimisation are entirely concentrated on the windward side of the leading car, predominantly situated in the middle and rear bottom regions.

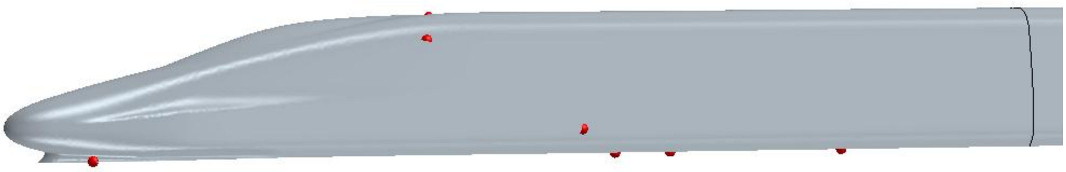


Figure 6: Layout of pressure sensors.

A random subset of operating conditions from the test set was selected to validate the proposed framework. In Figure 7, a visual comparison of the reconstructed surface pressure field at a running speed of 330 km/h and a crosswind speed of 5 m/s is shown. The entire reconstruction process requires merely 0.002 seconds, yielding a relative reconstruction error of 0.06%. This indicates that within the leading car region, only 0.06% of the surface area exhibits a relative error in the pressure coefficient exceeding 5%. Furthermore, the relative error contours reveal that regions of elevated reconstruction error are confined to minute areas at the nose tip and the rear of the leeward side.

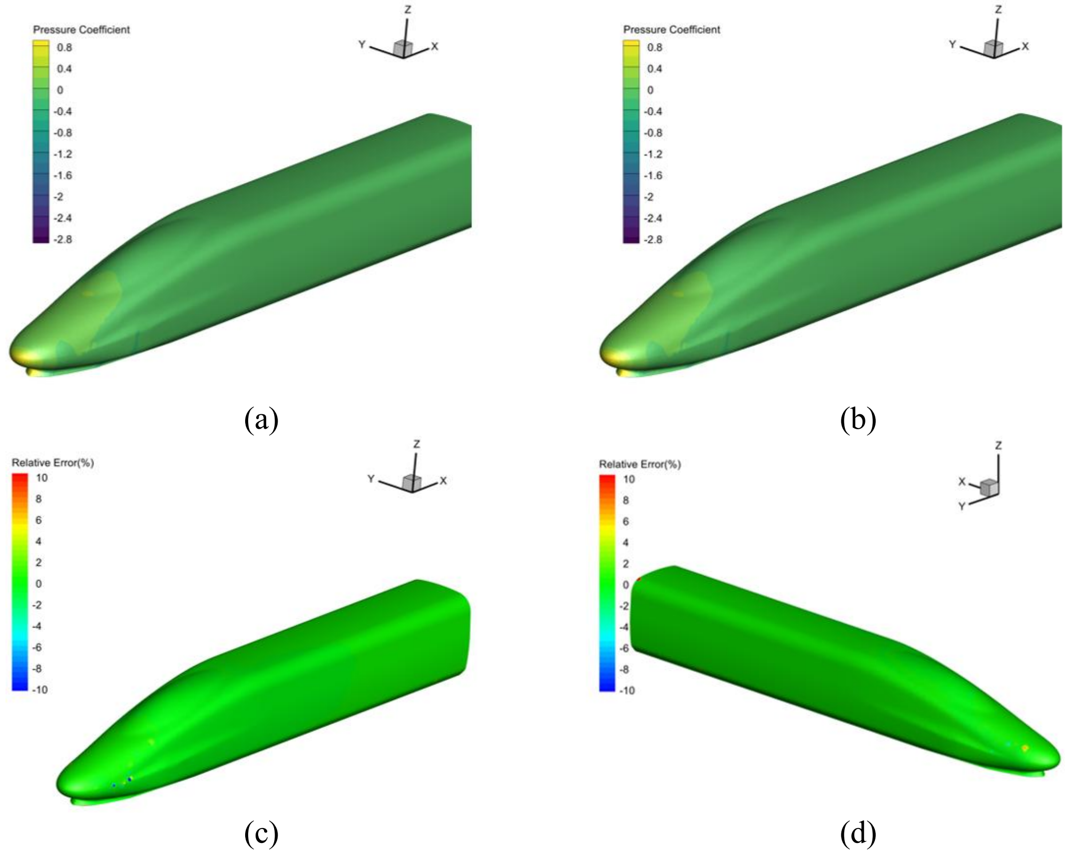
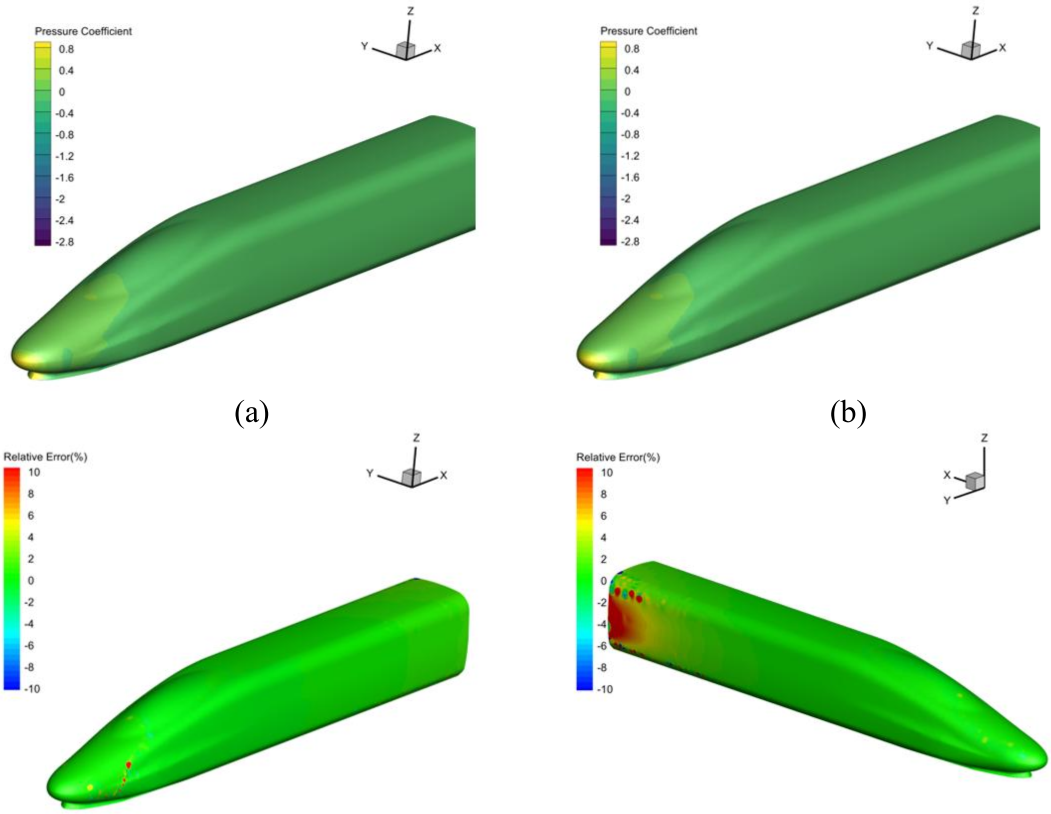


Figure 7: Comparison of the surface pressure fields at a running speed of 330 km/h and a crosswind speed of 5 m/s: (a) CFD-computed pressure coefficient contours; (b) reconstructed pressure coefficient contours; (c) relative error distribution on the windward side; (d) relative error distribution on the leeward side.

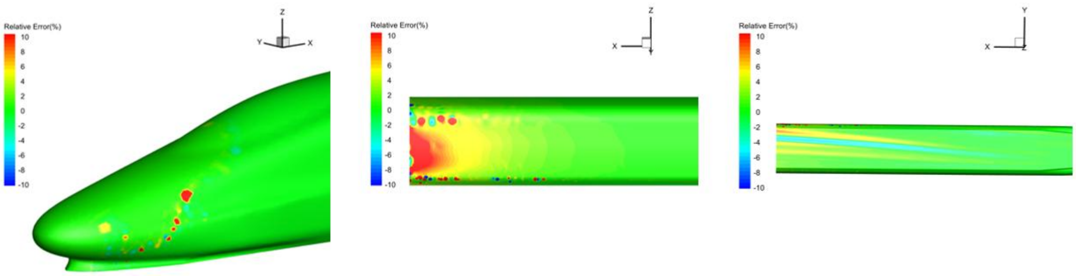
In Figure 8, the visualisation of the reconstructed surface pressure field at a running speed of 330 km/h and a crosswind speed of 10 m/s is shown. The crosswind velocity substantially impacts the reconstruction accuracy. As the crosswind speed increases from 5 m/s to 10 m/s, the error regions expand on the windward side of the nose tip and the leeward side near the tail of the leading car. Furthermore, a stripe-like error region, tapering towards the front, emerges at the rear of the bottom area. In Figure 9, an enlarged view of these high-error regions is shown. Aerodynamically, when the train operates in open air, the incoming flow undergoes compression at the windward nose tip, causing a sharp velocity reduction and generating substantial aerodynamic drag. Under crosswind conditions, the flow on the leeward side experiences separation, initiating periodic vortex shedding. Concurrently, due to the relative ground effect, the flow traversing from the leading to the trailing car encounters the continuous development of the ground boundary layer and stagnation effects at the underbody. This induces energy loss within the airflow, where the synergy of viscous forces and adverse pressure gradients decelerates the vertical velocity component to zero. The flow subsequently detaches from the ground, forming separated vortices. These compounded aerodynamic phenomena generate severe local pressure gradients, directly contributing to the significant deviations observed during the pressure field reconstruction.



(a) (b)

(c) (d)

Figure 8: Comparison of the surface pressure fields at a running speed of 330 km/h and a crosswind speed of 10 m/s: (a) CFD-computed pressure coefficient contours; (b) reconstructed pressure coefficient contours; (c) relative error distribution on the windward side; (d) relative error distribution on the leeward side.



(a) (b) (c)

Figure 9: Enlarged views of regions exhibiting large relative reconstruction errors: (a) windward side of the nose tip; (b) rear of the leeward side; (c) rear of the underbody.

To further quantify the impact of crosswind velocity, a subset of training and test conditions was selected for CS reconstruction, maintaining identical measurement point locations to evaluate error magnitudes. In Figure 10, the error comparison results for these conditions are shown; notably, these results remain independent of the measurement point ordering. The average relative reconstruction error across these conditions is 1.37%, with all instances exceeding a 1% error occurring exclusively under crosswind conditions. The maximum relative reconstruction error of 7.13% manifests at a running speed of 300 km/h and a crosswind speed of 15 m/s. This specific operating condition resides at the boundary of the numerical simulation design space. Because there is insufficient surrounding data to constrain and correct these boundary conditions, an elevated reconstruction error is theoretically anticipated. Consequently, future numerical design spaces should explicitly account for boundary sparsity by extending the coverage of simulation conditions marginally beyond the strict actual application range.

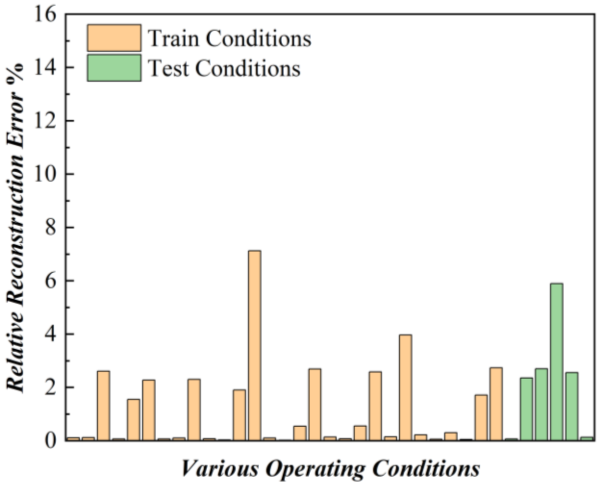


Figure 10: Comparison of relative reconstruction errors under various operating conditions.

Computationally, these surface pressure reconstructions for the leading car required an average of 0.0033 seconds per case on a single CPU core. In stark contrast, a single computational fluid dynamics (CFD) simulation under identical hardware conditions demanded an average of 800 hours. The immense computational burden of CFD fundamentally necessitates supercomputing platforms and parallel processing architectures. By bypassing the need to resolve the entire volumetric flow field, the proposed framework's ability to directly predict surface control points delivers a transformative reduction in computational effort and a robust increase in efficiency.

4 Conclusions and Contributions

To enable the accurate and real-time reconstruction of surface pressure fields on complex three-dimensional geometries using extremely sparse sensor data, this study proposes a rapid prediction framework integrating generalised proper orthogonal decomposition (GPOD), compressive sensing (CS), and an improved particle swarm optimisation (IPSO) algorithm. Applying this methodology, the surface pressure field of a high-speed train under complex operating conditions is reconstructed with high fidelity. The main conclusions are as follows:

1. The GPOD method extracts the dominant spatial modes of the flow field, effectively reducing the number of significant coefficients required for the CS reconstruction to fewer than 10. This dimensionality reduction fundamentally minimises the necessary number of physical pressure sensors.

2. The proposed IPSO algorithm optimises the sensor layout to capture essential flow features using a minimal sensor array. Maintaining high reconstruction accuracy, a configuration of merely 7 pressure measurement points is established. Employing the basis pursuit (BP) algorithm, the surface pressure field of the leading car is accurately reconstructed, achieving an average computation time of 0.0033 seconds and an average relative reconstruction error of 3%. Evaluations across various operating conditions indicate an increase in reconstruction error at the boundaries of the design space.

The proposed framework extends beyond high-speed trains and is applicable to other complex aerodynamic and hydrodynamic bodies, such as aircraft and submarines. By integrating the real-time surface pressure distributions to compute aerodynamic loads, this methodology provides a robust foundation for operational safety monitoring, highlighting its practical engineering value. Furthermore, the framework can be adapted to reconstruct other three-dimensional physical fields, including velocity and temperature. However, the current model is specifically trained for the CRH380A high-speed train; applying it to fundamentally different geometric configurations necessitates the construction of separate databases. Future studies will focus on enhancing the generalisation capability of the model across diverse vehicle shapes and validating its predictive performance using data from full-scale vehicle tests.

Acknowledgements

This work was supported by the National Key R&D Program of China (Grant No. 2025ZD0122004). The numerical calculations in this study were carried out on the Sugon Supercomputer.

References

[1] T. Liu, Z. Chen, Z. Guo, S. Krajnović, "Reasonable pressure tap layout to measure the aerodynamic forces of a train at different yaw angles", Measurement, 166, 108255, 2020.

[2] H. Gao, T. Liu, X. Chen, H. Huo, Z. Chen, J. Zhang, B.C. Khoo, "An accurate and efficient methodology on wind spectra relative to moving trains: field measurements of wind characteristics in complex terrains", Stochastic Environmental Research and Risk Assessment, 1-31, 2025.

[3] J. Zhang, A. Xu, F. Huang, Y. Bai, T. Liu, "A novel vortex control method for improving anti-overturning performance of a high-speed train with leeward airbag structures under crosswinds", Physics of Fluids, 36(6), 2024.

[4] J. Zhang, F. Huang, Y. Yu, S. Han, Y. Ding, G. Gao, "A novel wake flow control method for drag reduction of a high-speed train with vortex generators installing on streamlined tail nose", Physics of Fluids, 35(10), 2023.

[5] R. Inoba, K. Uchida, Y. Iwasaki, T. Nagata, Y. Ozawa, Y. Saito, T. Nonomura, K. Asai, "Optimization of sparse sensor placement for estimation of wind direction and surface pressure distribution using time-averaged pressure-sensitive paint data on automobile model", Journal of Wind Engineering and Industrial Aerodynamics, 227, 105043, 2022.

[6] M.S. Gritskevich, A.V. Garbaruk, J. Schütze, F.R. Menter, "Development of DDES and IDDES formulations for the k-ω shear stress transport model", Flow, Turbulence and Combustion, 88, 431-449, 2012.

[7] D.L. Donoho, "Compressed sensing", IEEE Transactions on Information Theory, 52, 1289-1306, 2006.

[8] E.J. Candes, T. Tao, "Near-optimal signal recovery from random projections: Universal encoding strategies?", IEEE Transactions on Information Theory, 52, 5406-5425, 2006.

[9] K. Manohar, B.W. Brunton, J.N. Kutz, S.L. Brunton, "Data-driven sparse sensor placement for reconstruction: Demonstrating the benefits of exploiting known patterns", IEEE Control Systems Magazine, 38(3), 63-86, 2018.

[10] J. Kennedy, R. Eberhart, "Particle swarm optimization", in "Proceedings of ICNN'95 - International Conference on Neural Networks", Volume 4, IEEE, 1942-1948, 1995.

[11] M. Clerc, J. Kennedy, "The particle swarm-explosion, stability, and convergence in a multidimensional complex space", IEEE Transactions on Evolutionary Computation, 6, 58-73, 2002.

[12] P.N. Suganthan, "Particle swarm optimiser with neighbourhood operator", in "Proceedings of the 1999 Congress on Evolutionary Computation-CEC99", Volume 3, IEEE, 1958-1962, 1999.