



Multi-fidelity load reconstruction in drone composite wings using in-service strain measurements

Ilias N. Giannakeas^{*}, Yuhang Pan, Sadjad Naderi[✉], Zahra Sharif Khodaei, Ferri M.H. Aliabadi

Department of Aeronautics, Imperial College London South Kensington Campus, Exhibition Road, SW7 2AZ, London, UK

ARTICLE INFO

Keywords:

Digital twin
Load reconstruction
Multi-fidelity modelling
Wind tunnel testing

ABSTRACT

Accurate reconstruction of aerodynamic loads from sparse strain measurements is a longstanding challenge in structural health monitoring (SHM), particularly for lightweight composite wings in unmanned aerial vehicles (UAVs) where space and weight constraints limit instrumentation. This work presents a data-efficient pipeline for UAV wing load reconstruction under sparse sensing, explicitly designed to operate with limited experimental data. The proposed framework uses a two-stage approach consisting of multi-objective optimization for sensor placement and a hybrid load reconstruction strategy that combines physics-based (low fidelity) with data-driven (high-fidelity) observations. In the first stage, the optimal strain gauge locations are identified numerically through a Pareto-based optimization that balances the number of sensors against reconstruction accuracy. Then, in the second stage, a multi-fidelity approach is developed for the reconstruction of the aerodynamic loads where a calibration matrix (CM), derived from finite element simulations, is combined with an artificial neural network (ANN) trained on limited experimental data to account for modelling discrepancies and improve prediction fidelity. The framework is validated both numerically and experimentally, using wind tunnel experiments on a composite UAV wing. The results show that the multi-fidelity framework improves load reconstruction relative to the CM-only baseline and is particularly effective when experimental data are limited. With only three wind-tunnel runs used for training, the multi-fidelity model reduced the mean MAE from 8.72 to 4.58 for lift and from 8.94 to 4.08 for drag. For lift, the corresponding mean absolute percentage error (MAPE) decreased from 49.23% to 18.44%. When ten training runs were used, the mean MAE was further reduced from 8.73 to 3.85 for lift and from 8.44 to 3.91 for drag, while the lift MAPE decreased from 47.51% to 12.92%. The proposed framework is also compared against a fully data-driven ANN trained on the same experimental datasets. Across the full data-availability study, the multi-fidelity approach reached accurate and stable predictions with fewer training runs than the ANN, which exhibited higher variability under data-scarce conditions. These results highlight the potential of the proposed framework as a promising basis for SHM and digital twin (DT) applications in aerospace structures.

1. Introduction

Structural health monitoring (SHM) of composite aerospace structures plays a critical role in ensuring mission reliability, structural integrity, and aerodynamic performance throughout service life. Recent advances in sensing technologies and computational capabilities have enabled increasingly sophisticated monitoring strategies, supporting the transition from periodic inspection toward continuous, data-informed assessment and digital twin-based operation of aerospace structures [1]. This is particularly relevant for composite materials, which, despite

their high strength-to-weight ratios, are susceptible to internal damage mechanisms such as delaminations and fibre breakage that are difficult to detect using conventional inspection techniques [2]. Within this context, integrating SHM with reliable load reconstruction capabilities is essential for linking measured structural responses to operational loads, enabling fatigue assessment, remaining useful life estimation, and performance-informed decision-making.

Aeronautical structures experience complex and highly variable loading throughout their operational life. Accurately capturing the aerodynamic load spectra encountered in service not only provides insight into historical usage but also supports improved structural

^{*} Corresponding author.

E-mail addresses: i.giannakeas19@imperial.ac.uk (I.N. Giannakeas), y.pan21@imperial.ac.uk (Y. Pan), sadjad.naderi@imperial.ac.uk (S. Naderi), z.sharif-khodaei@imperial.ac.uk (Z.S. Khodaei), m.h.aliabadi@imperial.ac.uk (F.M.H. Aliabadi).

<https://doi.org/10.1016/j.ast.2026.113171>

Received 22 February 2026; Received in revised form 11 June 2026; Accepted 7 July 2026

Available online 8 July 2026

1270-9638/© 2026 Published by Elsevier Masson SAS.

Nomenclature

$\mathbf{A}$	Calibration matrix for load estimation
$\mathbf{A}_{RF}$	Calibration matrix for estimation of reactions
$\mathbf{B}_{\mathcal{D}}$	Bias tensor
$\mathbf{F}_L$	Load vector
$\mathbf{F}_{RF}$	Vector of reaction forces
RF_y	Drag force
RF_z	Lift force
$\mathbf{W}_{\mathcal{D}}$	Weight tensor
$\mathbf{c}$	Boolean sensor indicator vector
n_L	Number of load components / unit load cases
n_s	Number of strain sensors
$\mathbf{x}_{RF}$	Location of force transducer
δ	Discrepancy term
ϵ^m	Measured Strains
ϵ_{ij}	Strain component
λ	Regularization parameter

design, certification, and maintenance planning [3,4]. SHM systems can support this need by enabling continuous monitoring capabilities for load estimation, early detection of performance degradation, and the development of digital twins that reflect the true operational state of the structure [5]. In composite wing structures, reconstructing aerodynamic lift and drag forces from strain-based measurements offers a practical pathway for monitoring stresses at critical locations and assessing fatigue damage accumulation under realistic operating conditions [6,7].

Common sensing technologies employed to acquire these strain histories include conventional strain gauges (SGs), piezoelectric sensors, and fibre-optic-based systems, such as Fiber Bragg Grating (FBG) sensors and distributed fibre optic (DFOS) sensors [8]. FBGs can be surface mounted or embedded into composites and allow multiplexing of many sensing points along a single fibre, offering great performance against electromagnetic interference [9]. Novel, miniaturized acquisition systems have made possible the integration of FBGs into space- and weight-limited structures [10]. Currently, integration of SGs is the standard practice in industry and has also been used extensively in wing structures [11,12]. However, in unmanned aerial vehicles (UAVs), constraints on space, weight, and power significantly limit the number and type of sensors that can be deployed. It is therefore typical that load reconstruction must be achieved from sparse strain measurements, often under conditions where experimental training data are limited.

Inferring loads, stresses, or displacements from strain measurements is an inherently ill-posed inverse problem [13]. In the specific case of reconstructing displacement or deformation fields, this class of methods is often referred to as ‘shape sensing’ [14]. A wide range of methodologies has been proposed to address this challenge. Data-driven approaches employ machine learning models such as Artificial Neural Networks (ANN) or Gaussian Process regression trained on numerical or experimental datasets to establish strain-to-load mappings [15,16]. While such methods can achieve high accuracy, their performance depends strongly on the availability of sufficiently rich and representative training datasets, and their generalization capability may degrade when data are sparse, or operating conditions deviate from those observed during training. More recently, physics-informed learning approaches have also been proposed for shape sensing, including models that incorporate strain–displacement relations, constitutive equations and equilibrium constraints within neural-network training [17].

In contrast, physics-based approaches exploit mechanical models to reconstruct loads or deformations without reliance on extensive training data. Typical examples include: i) the inverse finite element method (iFEM) which reconstructs full-field displacements by minimizing the least-squares error between measured and numerical strains, without

requiring prior knowledge of material properties or loading conditions [12,18], ii) strain integration approaches such as Ko’s displacement theory [19] and iii) the Calibration Matrix (CM) method in which strain responses to unit load cases computed via finite element (FE) models are used to infer unknown loads through inverse mapping. Owing to their simplicity and interpretability, CM-based approaches have seen widespread adoption across multiple engineering domains [10,20,21]. In [22] and [14], demonstrated the effectiveness of the CM approaches for load identification while their integration with FE models allowed full-field displacement, strain or stress reconstruction. In such two-step approaches, the loads are first identified from measured strains and are then applied to a refined FE model to recover the structural field quantities. These methods are highly valuable when the objective is full-field reconstruction and when a sufficiently representative structural model and sensor set are available.

Despite their respective advantages, both data-driven and physics-based approaches face limitations in practical UAV applications. Physics-based methods are sensitive to modelling assumptions, parameter uncertainties, sensor noise, and sparse sensor placement, which can lead to degraded reconstruction accuracy. Conversely, purely data-driven approaches often require extensive experimental datasets to achieve robust performance, which are costly and difficult to obtain.

To address these challenges, recent research has explored two complementary directions. The first concerns optimal sensor placement, recognising that reconstruction accuracy depends on both the number and spatial distribution of sensors. Approaches based on statistical, modal, and optimisation techniques such as Fisher Information Matrix (FIM), Effective Independence (EI), Modal Assurance Criterion (MAC), and genetic algorithms (GA) have been widely studied [23–25]. Additionally, data-driven sensor selection strategies have been recently proposed [26], leveraging machine-learned feature importance to identify informative sensing locations with reduced computational cost. These approaches aim to identify the monitoring locations that would obtain the maximum value out of the available sensor budget.

The second direction focuses on hybrid load reconstruction frameworks that combine physics-based models with data-driven corrections. Such approaches aim to retain the interpretability and generality of physics-informed methods while exploiting data-driven components to compensate for modelling discrepancies and unmodelled dynamics. For example, Zhou et al. [27] proposed a hybrid framework in which convolutional neural networks are coupled with reduced-order inverse models to reconstruct the full-field aerodynamic loads and corresponding structural deformations. Duan et al. [28] introduced a reduced-basis reconstruction enhanced by ANN-based corrections to achieve fast and robust aerodynamic force estimation from sparse measurements. These studies demonstrate the promise of hybrid methodologies for load reconstruction under practical constraints.

Despite these advances, a fundamental disconnect remains between optimal sensing and robust load reconstruction in operational settings. Sensor placement studies typically optimize information content or reconstruction accuracy in isolation, without explicitly considering how the selected sensor network performs within hybrid or data-driven reconstruction frameworks under realistic data constraints. Conversely, hybrid reconstruction methods often assume fixed sensor configurations and focus on algorithmic performance, with limited attention to how sensor placement, data availability, and experimental uncertainty jointly affect reconstruction fidelity during operation. As a result, existing studies provide only partial insights into how to achieve accurate and reliable aerodynamic load reconstruction from sparse strain measurements. A study that jointly optimises sensor placement and assesses data efficiency and load-reconstruction fidelity (explicitly accounting for limited experimental data, measurement noise, and modelling discrepancies) remains lacking, particularly for composite UAV wings operating under practical constraints.

To address this gap, this work proposes a data-efficient multi-fidelity framework for aerodynamic load reconstruction in composite UAV

wings that integrates numerical modelling, sensor placement optimization, and multi-fidelity reconstruction within a unified pipeline. Multi-fidelity is defined here as additive discrepancy correction of a physics-based prior. First, high-fidelity FE simulations are used to construct the CM that forms the mapping between unitary loads and strain response and to support a multi-objective optimization that identifies optimal strain gauge locations by balancing reconstruction accuracy against sensor count. Then, a multi-fidelity reconstruction strategy is introduced, in which a physics-based CM formulation provides a baseline load estimate, and an ANN trained on a limited set of wind tunnel measurements learns a discrepancy correction to improve prediction fidelity. This formulation enables accurate load reconstruction under sparse sensing while reducing dependence on large experimental datasets. In the present study, the primary quantity of interest is the reconstruction of the global aerodynamic resultant forces, rather than the recovery of the complete force distribution or full-field displacement/stress response.

The proposed methodology is validated both numerically and experimentally using wind tunnel tests on composite UAV wings. The results are compared against a CM-only method as well as a fully data-driven method. The main contributions of this work are as follows:

1. The development of an end-to-end, experimentally validated pipeline that jointly addresses sensor placement and aerodynamic load reconstruction under sparse sensing constraints in composite UAV wings.
2. Wind tunnel validation on a composite UAV wing, providing practical insight into the interaction between sensor placement, modelling discrepancies, and reconstruction fidelity under realistic operational conditions.
3. A systematic comparison using experimental observations under typical UAV operational conditions of physics-based (CM-only), fully data-driven, and hybrid multi-fidelity reconstruction, performed on the same sensor network and aerodynamic test campaign.
4. A quantitative assessment of data efficiency, indicating that multi-fidelity reconstruction can reach comparable accuracy with fewer experimental training runs within the tested operating envelope.

The remainder of this paper is structured as follows. Section 2 presents the proposed methodology, introducing the CM formulation for aerodynamic load reconstruction, the multi-objective optimization strategy for sensor placement, and the hierarchical multi-fidelity framework for discrepancy learning. Section 3 describes the wind tunnel experimental campaign, including the test configuration, instrumentation, and data acquisition procedures used to generate experimental datasets. Section 4 reports the numerical and experimental results, covering the identification of optimal sensor layouts, analysis of

strain measurements, and a comprehensive comparison of physics-based, data-driven, and multi-fidelity reconstruction strategies under varying data availability scenarios. Finally, Section 5 summarizes the main findings and discusses their implications for UAV structural health monitoring and digital twin applications.

2. Methodology

The proposed methodology is formulated as an integrated, multi-fidelity pipeline that combines numerical modelling, sparse sensing design, and experimental validation to enable accurate aerodynamic load reconstruction in composite UAV wings. The framework is developed and demonstrated using the ASTERO fixed-wing UAV by Nordic Wing [29], which provides a representative lightweight composite wing platform. The overall structure and information flow is illustrated schematically in Fig. 1, while the individual methodological components are detailed in the following subsections.

The methodology builds on a physics-informed backbone based on a FE model of the wing. This numerical model serves as a low-cost surrogate for simulating structural strain responses under unitary aerodynamic loading conditions and forms the basis for constructing CM mappings between applied loads, measured strains, and reaction forces. These mappings provide an interpretable, physics-based baseline for load reconstruction.

To explicitly address the sparse sensing constraints inherent to UAV platforms, a multi-objective sensor placement optimization is incorporated into the framework. Using information derived from the FE-based CM formulation, the optimization identifies strain gauge configurations that balance reconstruction accuracy against the number of sensors. Importantly, this step is not treated as a standalone design exercise, but as an integral component of the overall framework that provides actionable design feedback for both instrumentation strategy and subsequent operational load monitoring. In this way, the sensor layout is co-designed with the reconstruction methodology, ensuring consistency between numerical modelling, experimental validation, and in-service application (Fig. 1).

Recognising that numerical models alone cannot fully capture real structural behaviour, the framework adopts a multi-fidelity reconstruction strategy. In this formulation, the CM-based reconstruction is treated as a low-fidelity, physics-driven estimate, while an ANN trained on a limited set of experimental measurements learns a discrepancy correction that accounts for modelling errors, sensor placement imperfections, and experimental variability. This multi-fidelity approach preserves the interpretability and robustness of physics-based methods while providing adaptability when experimental data are limited.

Experimental wind tunnel tests conducted at Imperial College London on an instrumented wing-set from the ASTERO UAV provide the

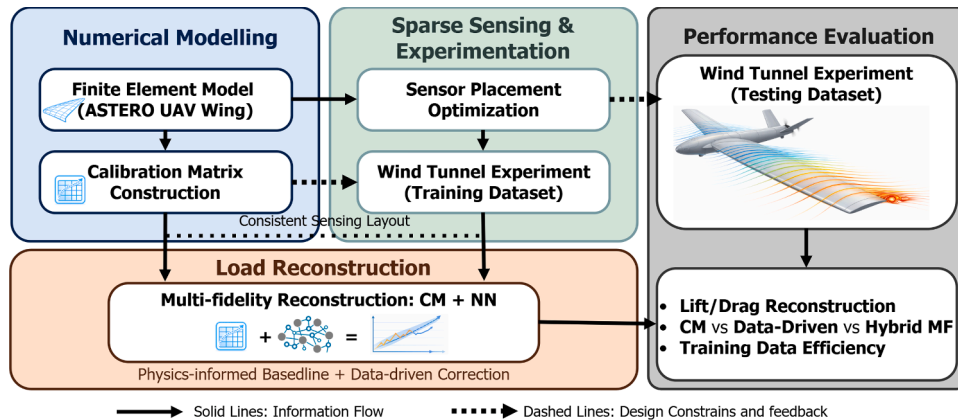


Fig. 1. Schematic overview of the proposed multi-fidelity load reconstruction framework integrating numerical modelling, optimization, and experimental validation.

data required for training and validation. These experiments enable systematic evaluation of the proposed framework under controlled yet representative aerodynamic conditions, allowing direct comparison between CM-only, purely data-driven, and hybrid multi-fidelity reconstruction strategies.

2.1. Calibration matrix (CM) method for aerodynamic load reconstruction

The CM method forms the foundation of the proposed multi-fidelity framework, as it embeds the underlying physics of the structural response. Load reconstruction relies on a suitable parameterization of the external load field acting on the structure, enabling the mapping of strain measurements to global load components. It is assumed that the structure operates within its elastic range, allowing the deformation state to be represented as a linear superposition of individual unitary loads. This assumption is generally valid in practical applications, since aerospace structures are designed to remain within their elastic regime under normal operating conditions.

The aerodynamic pressure acting on the wing surface can be parameterized as a set of concentrated forces defined as $\mathbf{F}_L = [F_{L,1}, F_{L,2}, \dots, F_{L,n_L}]^T$, where n_L is the number of individual concentrated forces. A common choice is to assume that the concentrated forces are applied at each nodal location of the underlying FE mesh that is constructed for the spatial discretization of the problem [10,30].

The application of parameterized aerodynamic load, $\mathbf{F}_L$, leads to the deformation of the structure and a strain field $\boldsymbol{\varepsilon}$. Let the structure be instrumented with n_s strain sensors, located at coordinates $\mathbf{x}_s$. At each time instant, the measured strains can be arranged in the vector $\boldsymbol{\varepsilon}^m = [\varepsilon_1^m, \varepsilon_2^m, \dots, \varepsilon_{n_s}^m]$. Using the linear-elastic assumption, the i^{th} strain ε_i^m can be described as a superposition of the strain contributions ε_{ij}^m from each $F_{L,j}$ load component as:

$$\varepsilon_i^m = \sum_{j=1}^{n_L} \varepsilon_{ij}^m = \sum_{j=1}^{n_L} a_{ij} F_{L,j} \quad (1)$$

where a_{ij} is an influence value and describes the effect of the j^{th} unit load on the i^{th} strain [30]. Repeating Eq. (1) for each element in $\boldsymbol{\varepsilon}^m$, the influence values a_{ij} can be arranged in an $[n_s \times n_L]$ matrix leading to:

$$\boldsymbol{\varepsilon}^m = \mathbf{A} \mathbf{F}_L \quad (2)$$

where $\mathbf{A}$ is termed the calibration matrix (CM). Each a_{ij} element of $\mathbf{A}$ can be viewed as the ε_i^m strain value that would be observed if a single unitary load $F_{L,j}$ was applied [31], and transforms the applied loads to sensed strains at the sensor locations. Thus, the elements of $\mathbf{A}$ are computed column by column by applying iteratively a unitary load at each load location. This process is done numerically using a FE model.

Assume that at a time instant, the installed sensor system measured the strain vector $\boldsymbol{\varepsilon}^m$. The error between the measured and the estimated strains under the application of the parameterized load can be computed as the norm $e = \|\boldsymbol{\varepsilon}^m - \mathbf{A} \mathbf{F}_L\|^2$. Using the error e for a given vector $\boldsymbol{\varepsilon}^m$ of measured strains, it is possible to estimate $\mathbf{F}_L$ by solving the inverse problem through a least-squares minimization of e (i.e. $\partial e / \partial \mathbf{F}_L = 0$) leading to [30]:

$$\hat{\mathbf{F}}_L = (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T \boldsymbol{\varepsilon}^m \quad (3)$$

where $\hat{\mathbf{F}}_L$ is the estimation of the applied load.

Generally, $n_s \neq n_L$ leading to a non-square matrix $\mathbf{A}$. The large condition number of $\mathbf{A}$ can lead to significant errors when solving the inverse problem described in Eq. (3). At the same time, measurement noise will be amplified and propagated to the final estimations. It has been shown that the reconstruction accuracy does not strictly increase as

the number of sensing points (n_s) increases [13]. To manage the reconstruction error in the identified load vector, the Tikhonov regularization technique is a common choice, which is suitable in removing additive Gaussian noise [13,32]. The load identification problem can be recast into an optimization problem $\min_{\mathbf{F}_L} \|\mathbf{A} \mathbf{F}_L - \boldsymbol{\varepsilon}^m\|_2^2 + \lambda^2 \|\mathbf{F}_L\|_2^2$ and the identified load can be derived through [33]:

$$\hat{\mathbf{F}}_L = (\mathbf{A}^T \mathbf{A} + \lambda^2 \mathbf{I})^{-1} \mathbf{A}^T \boldsymbol{\varepsilon}^m = \mathbf{A}^* \boldsymbol{\varepsilon}^m \quad (4)$$

where λ is the regularization parameter. Selection of the appropriate λ value is crucial to avoid over-smoothing the results or introducing redundant errors. The selection of parameter λ can be optimized to reduce measurement noise, leading to improved predictive accuracy [13,30,34,35]. A common approach is the implementation of the Generalized Cross Validation (GCV) approach which is also adopted here. For the implementation steps of GCV the interested reader is directed to [13] and the references therein. The regularization addresses the numerical instability of the inverse problem for a given CM by reducing noise amplification and improving the conditioning of the reconstruction. However, it does not correct systematic errors in the CM itself, nor discrepancies arising from modelling assumptions, boundary-condition differences, or sensor positioning errors. Therefore, regularization is necessary for stable CM-based reconstruction but is not sufficient to account for model-experiment mismatch. The proposed multi-fidelity framework is applied on top of the regularized CM estimate to account for such errors.

In Eq. (2), matrix $\mathbf{A}$ maps the applied forces $\mathbf{F}_L$ to the sensed strains $\boldsymbol{\varepsilon}^m$, formulating a forward problem. We can work in an analogous manner and create a mapping between the applied loads and specific reaction forces. Let $\mathbf{F}_{RF} \in \mathbb{R}^{n_{RF}}$ be the vector containing the reactions at a given control location $\mathbf{x}_{RF}$. Matrix $\mathbf{A}_{RF}$ can be defined to map $\mathbf{F}_L$ to the vector of reaction forces $\mathbf{F}_{RF}$ as:

$$\mathbf{F}_{RF} = \mathbf{A}_{RF} \mathbf{F}_L \quad (5)$$

where the elements in $\mathbf{A}_{RF}$ are defined and computed in an analogous manner to the influence values of $\mathbf{A}$. Therefore, given a measurement $\boldsymbol{\varepsilon}^m$ from the sensing system, the reaction forces can be estimated using Eq. (4) and 5 as:

$$\hat{\mathbf{F}}_{RF} = \mathbf{A}_{RF} \mathbf{A}^* \boldsymbol{\varepsilon}^m \quad (6)$$

Essentially, Eq. (6) combines a direct (Eq. (5)) with an inverse problem (Eq. (4)) to link the measured strains with the reactions observed. Similar formulations can be implemented to obtain mappings for various quantities of interest such as loads, stresses and strain fields [31]. For the UAV study conducted, this formulation aims to predict the total lift and the drag acting on the drone during flight. The focus on resultant lift and drag forces is motivated by their practical relevance for UAV operational monitoring. These quantities provide compact indicators of aerodynamic loading experienced by the structure and can be used to support load-history reconstruction, fatigue assessment, flight-envelope monitoring, and digital-twin updating.

2.2. Selection of the optimal sensing positions

The accuracy of load reconstruction depends on both the number and placement of strain gauges. Weight and space limitations in UAVs restrict the number of sensors that can be installed. Here, sensor placement optimization is not treated as a standalone objective, but as a supporting design step, aiming to identify the most informative locations, to ensure that the sparse sensor network maximizes the value of each sensor and enables robust load reconstruction within the proposed multi-fidelity framework.

Within this context, sensor placement is formulated as a constrained, multi-objective combinatorial optimization problem and performed prior to the experimental campaign using numerical predictions only. To

ensure tractability and reflect realistic installation constraints, candidate sensor positions are predefined as a finite set of feasible discrete locations on the wing surface, accounting for wiring routes and accessibility. Based on the internal layout of the wing structure, regions near the leading and trailing edges cannot be instrumented due to limited access during manufacturing and integration. Although the strain gauges used in the subsequent experimental campaign are installed on the exterior surface of the wing, the proposed methodology is formulated with broader applicability in mind, including scenarios where internal access for sensor integration is possible during manufacturing, representing a realistic and generalisable sensing design space for UAV wing applications.

Based on these considerations, candidate locations are restricted to three spanwise lines: one located near the leading-edge region, one aligned with the internal stiffener, and one near the trailing-edge region. Along each line, candidate sensor locations are defined at regular intervals of 100 mm, resulting in a total of $n_s = 50$ potential sensor positions across the top and bottom wing surfaces (Fig. 2A and Fig. 2B). This discretisation reflects common practice in both strain gauge and fibre-optic based monitoring systems, where longitudinal strain (ϵ_{xx}) is measured at regular spanwise intervals using distributed arrays [12].

The optimization addresses two competing objectives: (i) minimizing the reconstruction error of the aerodynamic reaction forces, and (ii) minimizing the number of installed sensors. This trade-off naturally leads to a multi-objective optimization problem, for which no single solution simultaneously optimizes both objectives. Instead, the outcome is a Pareto front comprising non-dominated solutions, where improvements in one objective can only be achieved at the expense of the other [37].

The reconstruction error objectives are evaluated over a numerical validation dataset composed of quasi-static distributed pressure load cases. These load cases are generated using the Gaussian pressure profiles described in Section 2.3. Therefore, the optimization is not performed for a single load case, but for a collection of idealized distributed loading scenarios. Let $\mathbf{c} = [s_1, s_2, \dots, s_{50}]$ be a Boolean vector indicating a selected sensor combination, where $s_i = 1$ or $s_i = 0$ indicates that a sensor has been selected or not, respectively. For each candidate sensor configuration, matrix $\mathbf{A}$ is reduced by retaining only the rows of active sensor locations while matrix $\mathbf{A}_{RF}$ is kept unchanged, since it maps the reconstructed load vector to the resultant reaction forces. Using the reduced matrix $\mathbf{A}$, the reaction forces are then reconstructed for each distributed pressure load case. The first objective can be expressed as:

$$f_1 = \sum RMSE(\mathbf{F}_{RF,i}, \hat{\mathbf{F}}_{RF,i}) \quad (7)$$

and it describes the summation of the root mean square error (RMSE) between the components in the true ($\mathbf{F}_{RF}$) and the estimated reaction forces ($\hat{\mathbf{F}}_{RF}$) computed using Eq. (6). The second objective can be written as:

$$f_2 = \sum_{i=1}^{n_s} s_i = \|\mathbf{c}\| \quad (8)$$

representing the total number of selected sensors. To constrain the search space and reflect practical deployment limitations, lower and upper bounds are imposed on the number of sensors. Specifically, a minimum of two sensors is required to ensure a solvable reconstruction problem, while the maximum number is limited to twenty four due to data acquisition hardware constraints. The resulting optimization problem can be written as:

$$\begin{aligned} & \min f_1 \\ & \min f_2 \\ & \text{s.t.} : 2 \leq f_2 \leq 24 \end{aligned} \quad (9)$$

The combinatorial nature of this formulation results in an exponentially growing search space, rendering exhaustive evaluation computationally infeasible. To efficiently explore candidate configurations, a GA is employed. The GA operates on a population of candidate solutions encoded as binary vector $\mathbf{c}$ (Fig. 3), which are evaluated using the two objective functions. Through successive generations involving selection, crossover, and mutation, the algorithm promotes high-performing configurations while maintaining population diversity and avoiding premature convergence. The optimisation process ultimately yields a set of Pareto-optimal sensor configurations that capture the trade-off between reconstruction accuracy and sensor economy.

2.3. Numerical model and construction of calibration matrices

The practical construction of matrices $\mathbf{A}$ and $\mathbf{A}_{RF}$, which are central to the identification of $\mathbf{F}_L$ and $\mathbf{F}_{RF}$, requires a computational representation of the wing. To this end, a FE model of the ASTERO wing is developed in ABAQUS [36] to simulate its structural response under a range of loading conditions.

Sensor Number

s_1	s_2	s_3	$\dots$	s_{50}	
0	0	1	$\dots$	1	c_1
1	1	0	$\dots$	1	c_2
0	0	1	$\dots$	0	c_3
1	0	1	$\dots$	1	c_4
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$
1	1	1	$\dots$	1	c_n

Sensor Combination

Fig. 3. Illustration of binary vectors $\mathbf{c}$ for the identification of active sensor locations.

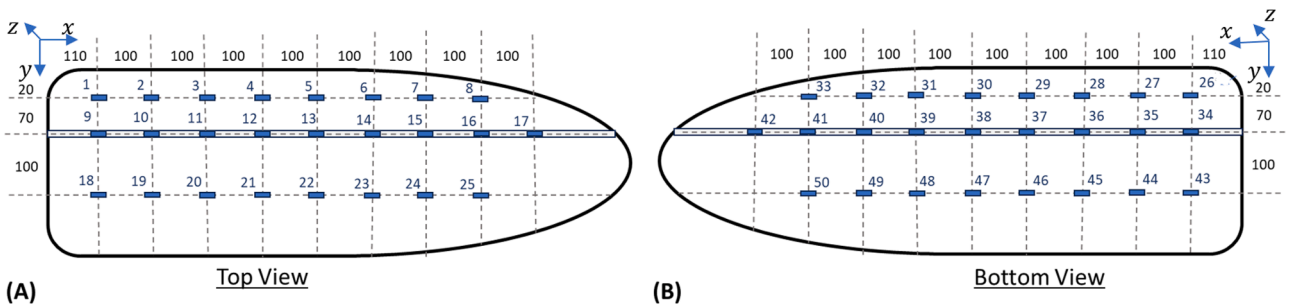


Fig. 2. A) Candidate locations on the top surface of the wing and B) candidate locations on the bottom surface of the wing.

The structure of each wing consists of a top and bottom skin and an internal stiffener (Fig. 4A). Linear quadrilateral shell elements with reduced integration (S4R) are used for spatial discretization. Each wing (left and right) is analysed individually, acting as a cantilever with fixed boundary conditions applied at the root. This modelling choice is made to ensure that the boundary conditions are well defined and directly controllable during the experimental campaign in the wind tunnel. The resulting FE mesh consists of 5538 elements and 5385 nodes. The discretization size is selected to ensure convergence of the structural response while further mesh refinement was found to have negligible influence on reconstruction accuracy while increasing the computational cost associated with matrix construction.

The wings of ASTERO are made of a proprietary composite structure with thickness of approximately 1 mm. Due to confidentiality constraints, the properties of the composite material used, and internal structure are not disclosed. Based on the information available, the global structural response of the wing is expected to be approximately quasi-isotropic. Therefore, in the present study, the wing is represented using an equivalent isotropic material model. This modelling choice is intended to capture the global elastic behaviour required for CM construction, rather than the detailed ply-level response.

The equivalent Young's modulus, E , and density, ρ , are identified experimentally by collecting vibration measurements from four ASTERO wings (2 left and 2 right wings) in the laboratory (Fig. 4B). The inclusion of responses from multiple wing sets ensures that the material and geometrical variability from the manufacturing process is captured in the estimations. Each wing is clamped at its root and excited to record tip accelerations [37]. A calibration study is then performed to adjust E and ρ such that the first two natural frequencies of the FE model matched those obtained experimentally (Fig. 4C), providing good agreement between the numerical and physical global dynamic response. The estimated values are reported in Table 1 which are used for all numerical simulations performed in this study. It is noted that the present framework targets quasi-steady aerodynamic load reconstruction under stable wind tunnel conditions. Dynamic aeroelastic phenomena, such as flutter, are outside the scope of the current study. Remaining discrepancies between the equivalent FE model and the experimental response are addressed through the multi-fidelity discrepancy-learning correction introduced in Section 2.4.

For the construction of $\mathbf{A}$ and $\mathbf{A}_{RF}$, unitary forces are applied sequentially at each nodal location of the FE mesh on the surface of the wing. Aerodynamic loading is assumed to be adequately represented as a superposition of lift (F_z) and drag (F_y) components, while longitudinal contributions (F_x) are neglected due to their negligible influence within the operating conditions considered. A load case is constructed where a unitary lift and drag forces are applied individually at each surface node. In total $n_L = 2 \times 5295 = 10590$ unit load cases are constructed where the nodes at the root boundary have been excluded.

For each unit load case, the resulting strain field is computed and the longitudinal strain component ϵ_{xx} is extracted at the 50 candidate strain-gauge locations (see Fig. 2). The strain responses associated with a single unit load form the entries of the corresponding column of matrix $\mathbf{A}$. In

Table 1

Equivalent material properties used for the simulation of the wing response.

Young's Modulus E [MPa]	Density ρ [kg/m ³]	Thickness t [mm]	Poisson's Ratio
16000	2000	1	0.3

parallel, the global reaction forces at the wing root are recorded to populate matrix $\mathbf{A}_{RF}$. Specifically, the reaction forces RF_z (total lift reaction) and RF_y (total drag reaction) are extracted with respect to a reference point located at $[-105, 75, -221.67]$ mm. This reference point corresponds to the monitoring position of the force transducer that is used in the subsequent wind tunnel experiments. The adopted force directions, coordinate axes, and reference point are illustrated schematically in Fig. 5A.

Using the solutions from the simulated load cases, matrices $\mathbf{A}$ and $\mathbf{A}_{RF}$ are constructed. If strain gauges are available at all candidate sensor locations, the dimension of these matrices would be 50×10590 and 2×10590 , for $\mathbf{A}$ and $\mathbf{A}_{RF}$, respectively. The construction of $\mathbf{A}$ and $\mathbf{A}_{RF}$ enables the estimation of lift and drag reaction forces (RF_z and RF_y , respectively) for each wing, given a vector of strain measurements $\mathbf{e}^m$ at any time instant.

For validation of the CM method, a numerical dataset of simulated load scenarios is constructed. In each scenario, a distributed pressure field is applied normal to the wing surface, following the Gaussian profile illustrated in Fig. 5B and centered at $\mathbf{x}_c = (x_c, y_c)^T$ as:

$$q(\mathbf{x}) = A_q \exp\left(-\frac{\|\mathbf{x} - \mathbf{x}_c\|^2}{2\sigma_q^2}\right) \quad (10)$$

where $A_q \sim U(1.125 \cdot 10^{-3}, 3.75 \cdot 10^{-3})$ N/mm² is the amplitude of the pressure load and $\sigma_q = 30$ mm is the standard deviation. In total, 400 samples are generated and simulated using the FE model, by randomly varying $\mathbf{x}_c$ over the wing surface. During each simulation run, the longitudinal strains ϵ_{xx} are computed at each of the 50 candidate sensor locations, along with the reaction forces at the reference point. This numerical validation dataset is used to evaluate the CM performance and to guide the sensor placement optimization in the following sections and it is distinct from the unit nodal lift and drag load cases used to construct $\mathbf{A}$ and $\mathbf{A}_{RF}$.

The reaction forces are reconstructed based on the strain measurements using Eq. (6) and matrices $\mathbf{A}$ and $\mathbf{A}_{RF}$. It is noted that for this validation step, all 50 locations are assumed active. The predicted versus true values of RF_y and RF_z are reported in Fig. 6. The results indicate that using the CM method, it is possible to reconstruct accurately the reaction forces with RMSE errors of 0.13 and 1.63 for RF_y and RF_z , respectively. The high accuracy in the reconstruction is expected as matrices $\mathbf{A}$ and $\mathbf{A}_{RF}$ and the validation dataset are both generated from exactly the same system (the FE model in this case).

In the following sections, the validation dataset will be used as a benchmarking tool to evaluate the reconstruction accuracy of different sensor configurations during the sensor placement optimization.



Fig. 4. A) Numerical discretization of the wing geometry using shell elements, B) experimental configuration for the extraction of acceleration time histories and C) comparison of the first 2 natural frequencies between the numerical model with the equivalent material properties and the actual wings.

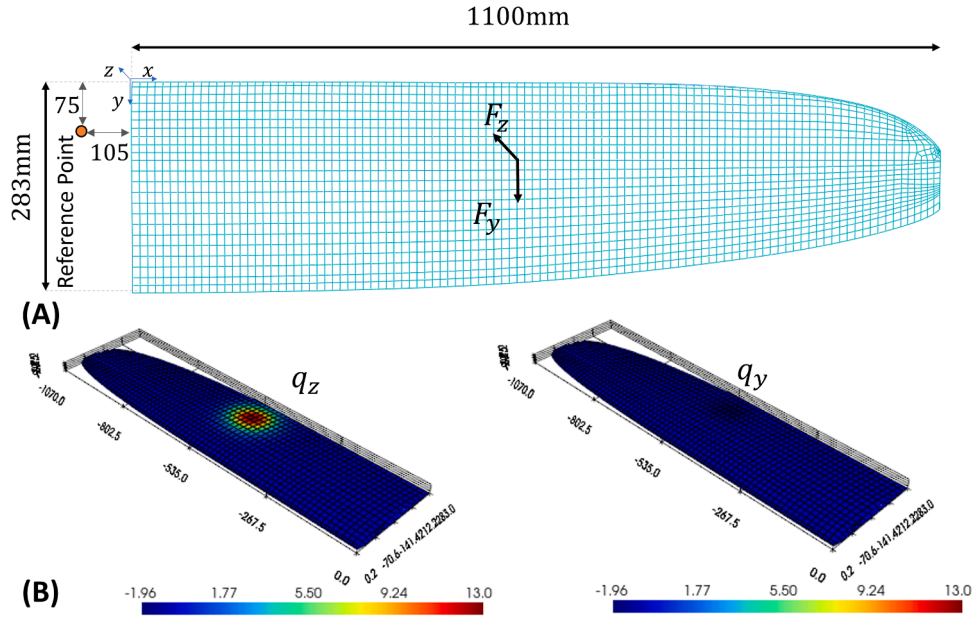


Fig. 5. A) Top down view of the meshed wing. An example of unitary lift and drag forces is depicted along with the position of the reference point. B) Example of a test load with the application of Gaussian pressure on the wing surface.

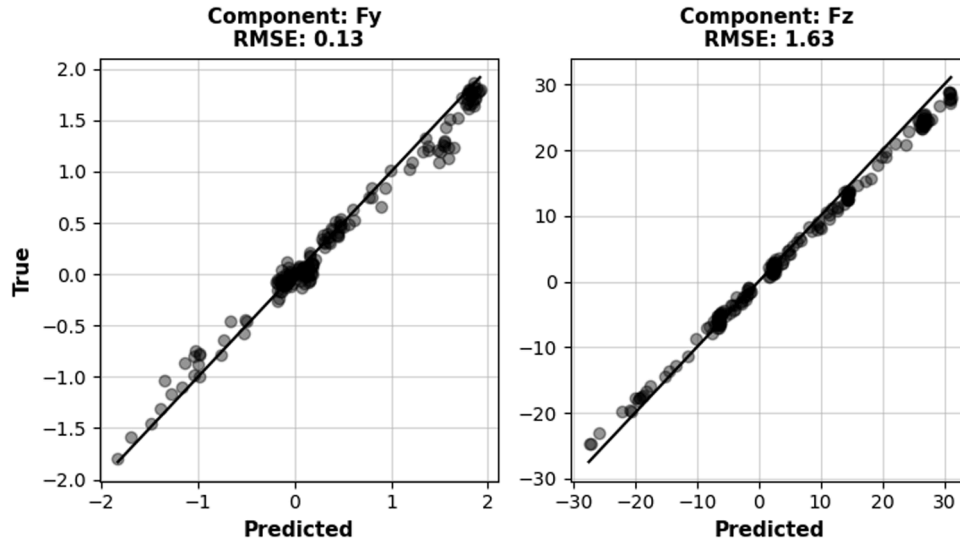


Fig. 6. Evaluation of the reconstruction accuracy of the CM method using the numerical dataset generated.

2.4. Multi-fidelity load reconstruction through discrepancy learning

While FE models are valuable for establishing the CM, their predictive accuracy is affected by several sources of uncertainty. Material property variability, geometric tolerances, and modelling simplifications can all reduce the accuracy of the simulated response. In addition, the inverse problem formulation in Eq. (4) is ill-conditioned and susceptible to numerical instabilities, meaning that even small discrepancies in the strain data may lead to significant errors in the reconstructed loads. On the experimental side, uncertainties arise from sensor misplacement/misalignment, environmental and operational conditions, and intrinsic measurement errors of the strain gauges. These modelling and experimental uncertainties can significantly impair reconstruction accuracy.

To mitigate these effects, a hierarchical multi-fidelity framework is adopted, in which the FE-based CM method is explicitly treated as a low-fidelity, physics-informed prior, and a discrepancy term is introduced to

learn systematic deviations from experimental observations using limited data [38]. This formulation effectively decomposes the reconstruction into two components: (i) a physics-based estimate derived from the FE model and inverse mapping, and (ii) a data-driven correction that learns systematic biases and unmodeled dynamics. Such an additive structure is particularly well suited to SHM applications, as it preserves the interpretability of the physics-based framework while exploiting the adaptability of machine learning to compensate for systematic uncertainties.

The hierarchical approach can be expressed as:

$$\mathbf{F}_{RF}^{HF}(\mathbf{e}^m) = \mathbf{F}_{RF}^{LH}(\mathbf{e}^m) + \delta(\mathbf{e}^m) \quad (11)$$

where, $\mathbf{F}_{RF}^{LH}(\mathbf{e}^m) = \hat{\mathbf{F}}_{RF}(\mathbf{e}^m)$ is the low-fidelity estimate based on Eq. (6), $\delta(\mathbf{e}^m)$ is the discrepancy term and $\mathbf{F}_{RF}^{HF}(\mathbf{e}^m)$ is the high-fidelity estimation.

In practice, the identification of $\delta(\mathbf{e}^m)$ is not straightforward and requires field observations to characterize the mismatch between model-

based and real responses [39]. Here, $\delta(\epsilon^m)$ is approximated by training an ANN on experimental data to learn this mismatch:

$$\delta(\epsilon^m) = \mathbf{F}_{RF}^{\text{Measured}} - \mathbf{F}_{RF}^{\text{LH}}(\epsilon^m) \approx \text{ANN}(\epsilon^m; \mathbf{W}_{\mathcal{D}}, \mathbf{B}_{\mathcal{D}}) \quad (12)$$

where, $\mathbf{W}_{\mathcal{D}}$ and $\mathbf{B}_{\mathcal{D}}$ are the weight and bias matrices learned on a training dataset $\mathcal{D}$ of strain-load pairs. Substituting Eq. (12) into Eq. (11) yields:

$$\mathbf{F}_{RF}^{\text{HF}}(\epsilon^m) \approx \hat{\mathbf{F}}_{RF}(\epsilon^m) + \text{ANN}(\epsilon^m; \mathbf{W}_{\mathcal{D}}, \mathbf{B}_{\mathcal{D}}) \quad (13)$$

Because the ANN learns only the residual discrepancy rather than the full strain-load mapping, substantially fewer training runs are required compared to fully data-driven approaches.

This hierarchical approximation retains the CM prediction as a physics-based prior and learns only the residual discrepancy between the CM estimate and the measurements with a data-driven correction, providing a robust framework that accounts for both numerical inaccuracies and experimental uncertainties. Both the numerical simulations and experimental measurements that are used in the above equations implement the sensor network that is suggested following the optimization process defined in Section 2.2, ensuring consistency between the two stages. Although an ANN is used in this work, the formulation is not restricted to neural networks. Other surrogate models, such as Gaussian processes [39,40], Polynomial Chaos Expansion [41,42], and statistical methods [43] have been implemented in similar applications. In this study, a feed-forward ANN was selected because it provides a simple and computationally efficient way to approximate the residual mapping from sparse strain measurements to force-discrepancy terms, while allowing rapid inference once trained.

Further calibration or updating of the FE model could also improve the fidelity of the numerical simulator and, consequently, the initial CM-based estimates by recreating the $\mathbf{A}$ and $\mathbf{A}_{RF}$ matrices. Such updating may be performed using preliminary laboratory tests, as adopted here through matching of the first two frequencies, or using online strain measurements during operation [44]. Even an updated FE model may not fully capture all sources of experimental discrepancy, particularly when these arise from effects that are difficult to parameterize explicitly, such as local sensor positioning and bonding errors, boundary-condition tolerances, aerodynamic contributions from the complete wind tunnel fixture, and manufacturing variability in the composite wing. For this reason, the proposed residual-learning step is not intended to replace FE calibration, but to complement it by correcting the remaining systematic mismatch at the level of the reconstructed resultant forces.

3. Experimental test campaign

3.1. Wind tunnel facility and flow conditions

An experimental test campaign is conducted to generate the data required for training the discrepancy term and for validating the pro-

posed multi-fidelity load reconstruction framework under controlled yet representative operating conditions. All experiments are performed in the low-speed wind tunnel at Imperial College London, which features a test section of 1.5m $\times$ 3.0m (Fig. 7A) and provides uniform flow velocities at a controlled ambient temperature of 18°C. The total blockage ratio of the experimental configuration was approximately 2.1%, remaining below the commonly accepted 5% threshold for low-speed wind tunnel testing, thereby minimising blockage-related interference effects.

The test conditions are selected to reflect the operational envelope of the ASTERO UAV while ensuring stable and interpretable aerodynamic loading. The free-stream velocity was limited to 24m/s, and the angle of attack (AoA) is varied between -15° and $+20^\circ$. These limits are chosen to avoid damaging the load cell that is used for measuring the acting forces, while sufficiently covering the expected operational range of the ASTERO UAV.

3.2. Mounting configuration and boundary conditions

To reproduce controlled and repeatable boundary conditions, each ASTERO wing is mounted onto custom fixtures machined to match the wing profiles (Fig. 7B), enabling precise positioning and load transfer while ensuring geometric compatibility with the wing structure. An elliptical nose profile is also added to emulate a simplified fuselage geometry during testing. This configuration allows each wing to be treated as a cantilever, enabling precise control of the boundary conditions and ensuring consistency with the numerical modelling assumptions.

The fixtures are bolted to a 15 mm thick aluminium flange creating a rigid assembly. Holes are drilled through the flange along a circular pattern allowing the assembly to rotate. The rotation axis was positioned close to the aerodynamic centre of the wing section (approximately one-third of the chord length from the leading edge), such that pitch adjustments produced realistic variations in aerodynamic loading. The holes are spaced to allow changing the AoA in the range from -15° to $+20^\circ$ in 5° increments. The range is selected to cover the expected operational profile of ASTERO during flight.

3.3. Instrumentation and data acquisition

Aerodynamic forces are measured using an ATI Delta 330-30 triaxial force transducer [45] mounted beneath the aluminium flange. The transducer records total lift and drag force components with sampling frequency of 1000 Hz and provides the reference measurements used for training the discrepancy model and validating the reconstructed loads. To minimise boundary-layer interference and blockage effects, the model is mounted on an aluminium column, raising the wing centreline to 0.75 m above the tunnel floor. All wiring is routed along the tunnel floor and passed outside the test section through a sealed opening to avoid flow disturbance.

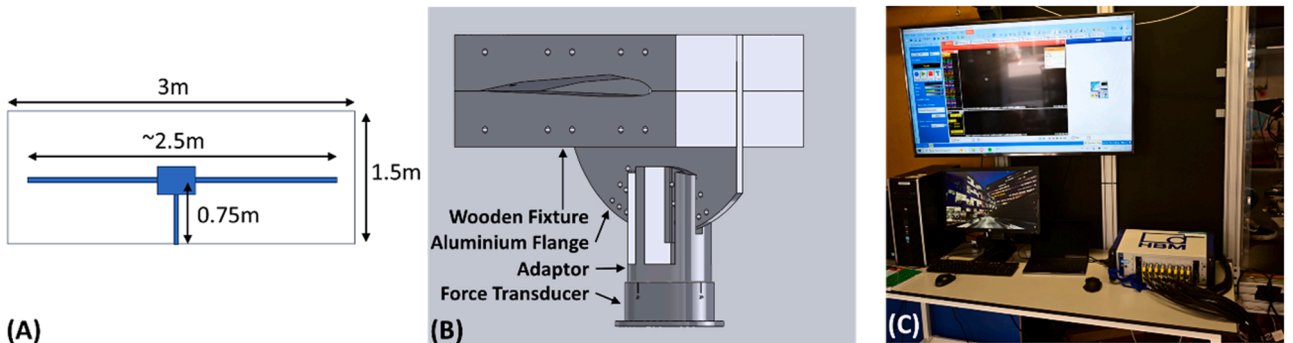


Fig. 7. A) Cross section of wind tunnel with positioning of the model aircraft, B) Fixture configuration for the installation of the ASTERO wings C) Data acquisition set-up with HBM GEN4tB.

Strain measurements are acquired using an HBM GEN4tB high-speed data acquisition (DaQ) system [46], which provides a total of 24 channels for simultaneous measurement (Fig. 7C). The maximum number of channels in the DaQ dictates the maximum number of sensors in the optimization step. The system is configured to sample each channel at 1000 Hz, fully synchronised with the force transducer, enabling direct time-aligned correlation between measured aerodynamic loads and structural strain responses. The strain gauges are installed according to the optimised sensor configurations identified in Section 2.2, ensuring consistency between the numerical optimisation, multi-fidelity reconstruction, and experimental validation stages.

4. Results and discussion

4.1. Numerical identification of optimum sensor placement

The sensor placement optimization problem is solved using the CM constructed in Section 2.3. The optimization is formulated as a multi-objective combinatorial problem and implemented using the NSGA-II algorithm available in the pymoo package [47]. The GA is initialized with a random population of 100 Boolean vectors $\mathbf{c}$, each representing a candidate sensor configuration. Binary random sampling is used for initialization, while half-uniform crossover (crossover probability 0.9) and bit-flip mutation (per-variable mutation probability equal to $1/n_s = 0.02$) are employed to generate offspring solutions. Duplicate individuals are eliminated to maintain population diversity.

For each candidate configuration the Boolean sensor availability vector $\mathbf{c}$ indicates which sensors are available and thus the dimensions of matrix $\mathbf{A}$ change and can be written as $\sum \mathbf{c} \times 10590$, where $2 \leq \sum \mathbf{c} \leq 24$, given the constraints on the allowable number of sensors defined in Section 2.2. The reconstruction error and sensor-count objectives are then evaluated using Eqs. 7 and 8, respectively. The numerical validation dataset of distributed pressure load cases constructed in Section 2.3 is used at each iteration to assess the reconstruction accuracy associated with each candidate sensor configuration. The algorithm is executed for 40 generations, resulting in a total of 4000 function evaluations.

The optimization yields a set of Pareto-optimal sensor configurations that represent the trade-off between reconstruction accuracy and the number of installed sensors. The resulting objective space and the Pareto front are shown in Fig. 8. Each red marker corresponds to a sensor configuration evaluated during the optimization, while the blue markers identify the non-dominated solutions forming the Pareto front. As expected, configurations with fewer sensors generally lead to higher

reconstruction errors. A clear region of diminishing returns is observed between approximately 10 and 15 sensors, beyond which further increasing the total sensor count yields only marginal improvements in reconstruction accuracy. This behaviour provides a strong numerical justification for sparse instrumentation in the ASTERO wing.

To further interpret the optimization outcomes, the sensor configurations on the Pareto front are visualized in Fig. 9A using the binary selection vector $\mathbf{c}$. Distinct spatial patterns emerge where sensor locations aligned with the internal stiffener are consistently selected on both the upper and lower wing surfaces, while positions near the wing root and trailing edge also appear frequently. In contrast, sensor locations close to the leading edge are rarely selected, indicating limited sensitivity to the reaction forces of interest.

To extract additional trends and insights, the subset of solutions yielding RMSE error below 20 are grouped, and the 20 most frequently selected sensor locations were identified. Fig. 9B shows the most frequently selected sensor locations within this subset, highlighting a small number of structurally informative regions that dominate the optimal solutions. These results indicate that the sensor placement problem is highly structured, with optimal configurations concentrating sensors along key load-bearing features rather than distributing them uniformly.

While the mathematically optimal configurations identified by the optimization often exhibit irregular and asymmetric layouts, such solutions are undesirable in practice due to increased wiring complexity and integration challenges. Guided by the observed selection patterns and constrained by the 24-channel limit of the available data acquisition system, a final sensor layout comprising 12 uniaxial strain gauges per wing is selected. This choice both aligns with the observed region of diminishing returns, where adding further sensing locations would be expected to yield only marginal improvements in reconstruction accuracy as well as aims to fully utilize all available channels in the DaQ system used during the subsequent wind tunnel testing. The chosen configuration includes sensors 9, 11, 13, 15, 18, 20, 34, 36, 38, 40, 43, and 45 (see Fig. 2 for numbering convention). The selected locations are divided into 6 at the top surface (Fig. 9B, indicated with green circles) and 6 at the bottom surface of the wing, preserving symmetry between the upper and lower surfaces and arranging the sensors along two straight spanwise lines to facilitate easy installation while remaining consistent with the optimization results. Therefore, the sensor configuration considers both the insights gained from the optimization study as well as practical implementation.

Since the selected configuration does not match exactly the configuration on the Pareto front, the reconstruction performance of the selected layout is compared against the best-performing 12-sensor configuration from the Pareto front. The comparison summarized in Table 2, shows that the proposed layout achieves comparable reconstruction accuracy for both lift (RF_z) and drag (RF_y) forces, despite not being the mathematically optimal solution. This confirms that near-optimal numerical performance can be achieved using structured and practically deployable sensor layouts, providing a robust basis for the subsequent experimental campaign and multi-fidelity reconstruction analyses.

4.2. Wind tunnel demonstration under representative operational conditions

Using the selected sensor configuration, a set of composite wings are instrumented with standard foil strain gauges [48] featuring a gauge length of 15 mm and a nominal resistance of 350 Ω . A total of 12 gauges are installed on each wing, with six mounted on the upper surface and six on the lower surface (Fig. 10).

To minimize disruption of the aerodynamic surface, electrical tracks are screen-printed on flexible Kapton film using silver ink. The gauge leads are bonded to these tracks via copper pads and conductive adhesive, providing a low-profile and reliable connection. The flexible films

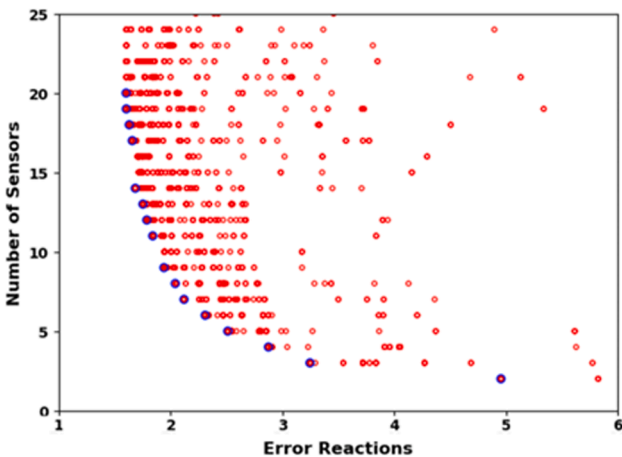


Fig. 8. Objective Space for all samples during the GA optimization. Red circles indicate a sample solution while blue circles indicate non-dominated solutions on the Pareto front.

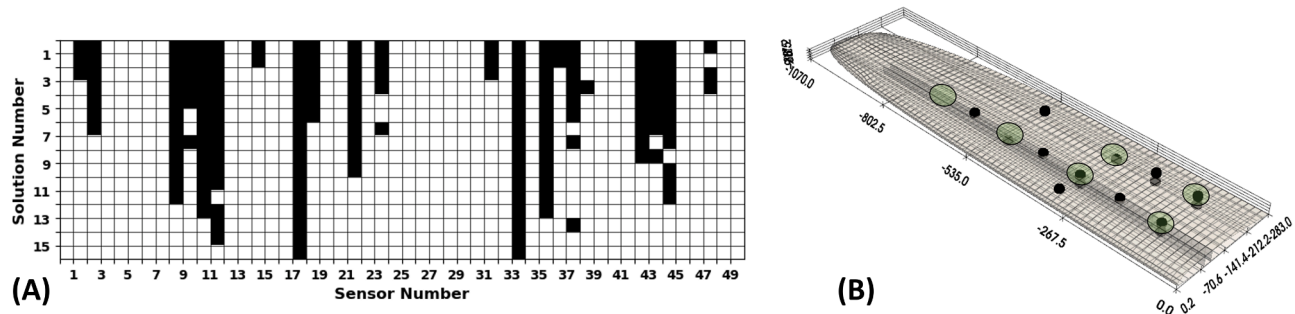


Fig. 9. A) Visualization of the Pareto front sensor combinations. Black boxes indicate active sensors. B) Visualization of the most commonly picked sensor locations for RMSE less than 20. Green circles indicate the final selected pattern (symmetrical on both sides).

Table 2
Comparison of the predictive accuracy between the 12 optimal sensor locations and the configuration selected.

Configuration	RF_y [N]	Error in RF_y [%]	RF_z [N]	Error in RF_z [%]
Selected 12 Sensors	29.1	5.2	241.8	3.4
Optimal 12	32.2	4.9	253.6	1.3
True Values	30.7	-	250.4	-

carrying the printed tracks were subsequently bonded to the wing surface using heat blankets to ensure uniform adhesion. The bonding procedure consisted of ramping the temperature to 90°C, holding it steady for 10 minutes. The printed tracks, developed and supplied by the Danish Technological Institute, exhibited a resistance of less than 0.2 Ω , thereby reducing the need for conventional wiring and limiting

aerodynamic interference.

A preliminary screening run is conducted to characterize the aerodynamic response of the wings and to define a safe and representative operating envelope for the experimental campaign. During this initial assessment, the wind velocity is set to 12m/s and the AoA is varied from 10° to 25° in 5° increments (Fig. 11). Surface-mounted tufts are used to visualize flow behaviour and their motion is recorded during testing using a camera installed inside the wind tunnel. While lift and drag forces remain small and stable up to 15° angle of attack, significant flow separation and associated structural vibrations are observed at 20° and 25°. Similar vibrations were also recorded for AoA below - 5°. To avoid excessive structural excitation and potential damage to the load cell, the combinations of wind velocity and AoA tested in the main campaign are restricted accordingly and summarized in Table 3. A total of 16 combinations are tested.

For each operating point, the wind tunnel turbine is gradually ram-

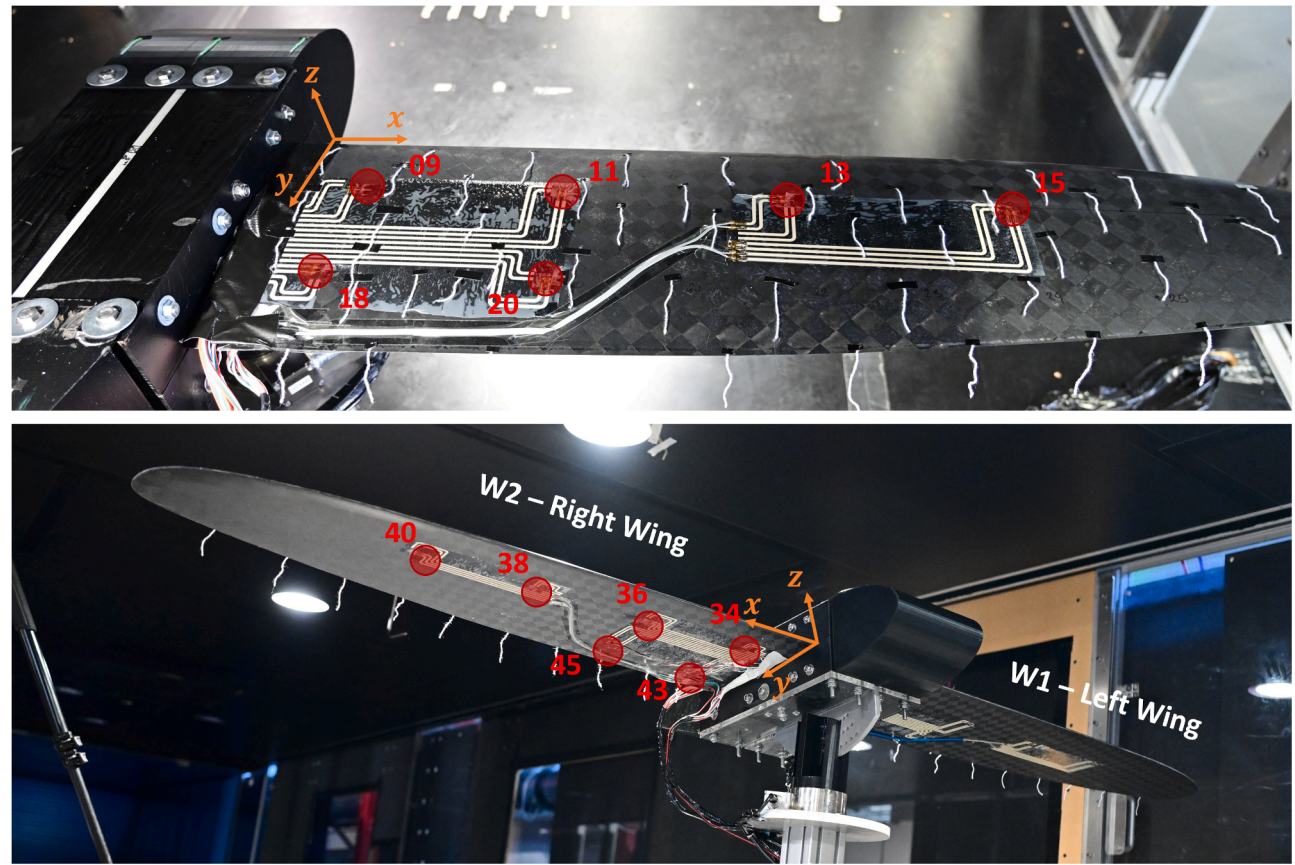


Fig. 10. Sensorized composite wings and installation of the model drone inside Imperial's wind tunnel. Each wing is sensorized with 12 strain gauges and printed tracks are used to minimize the disturbance of the aerodynamic surface.

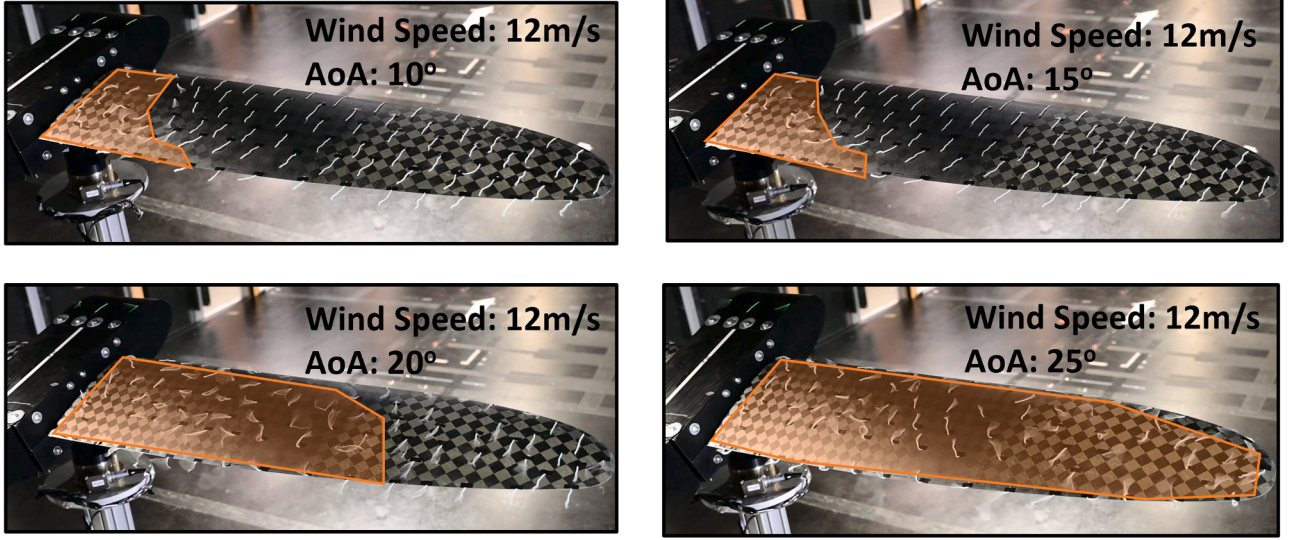


Fig. 11. Airflow separation at different AoA. Orange regions indicate the observed separation area.

Table 3

Wind velocity and Angle of Attack combinations tested during the wind tunnel experiments.

		Wind Velocity [m /s]				
		12	16	20	22	24
Angle of Attack [°]	-5	✓	✓	✓		
	0	✓				
	5	✓	✓	✓	✓	✓
	10	✓	✓	✓	✓	
	15	✓	✓	✓		

ped to the target velocity, after which the flow is held steady for at least 30 seconds to ensure stable aerodynamic conditions and allow the structural response to settle. Strain and reaction forces are recorded simultaneously throughout this period, yielding a set of synchronised strain–reaction force dataset $\mathcal{S}$ that forms the experimental dataset used for training and validation of the proposed hierarchical reconstruction framework.

Because the force transducer is mounted beneath the rotating flange, the measured reaction forces are expressed in the global tunnel coordinate system. These measurements are transformed into aerodynamic lift and drag components using a standard coordinate rotation:

$$\mathbf{F}_{RF} = \begin{bmatrix} RF_y \\ RF_z \end{bmatrix} = \begin{bmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{bmatrix} \begin{bmatrix} RF'_y \\ RF'_z \end{bmatrix} \quad (14)$$

where α is the AoA and RF'_y , RF'_z are the reaction forces measured at the transducer coordinate system (Fig. 12). This transformation is applied consistently across all test cases to compute the aerodynamic lift and drag forces. It is noted that the recorded values correspond to the combined forces acting on both wings.

Fig. 13 (W1: left wing, W2: right wing) presents strain time histories from all instrumented locations on both wings for a single wind tunnel run (16 m/s wind velocity and AoA 5°), illustrating the symmetry, repeatability, and residual discrepancies observed in the experimental data. As expected, gauges mounted on the upper surfaces record compressive strains, while those on the lower surfaces record tensile strains, consistent with the bending response of the wing under aerodynamic loading. The measurements from the two wings show good overall agreement in both magnitude and temporal evolution, indicating a consistent structural response under identical aerodynamic loading. Across all sensor locations, the strain time histories exhibit similar

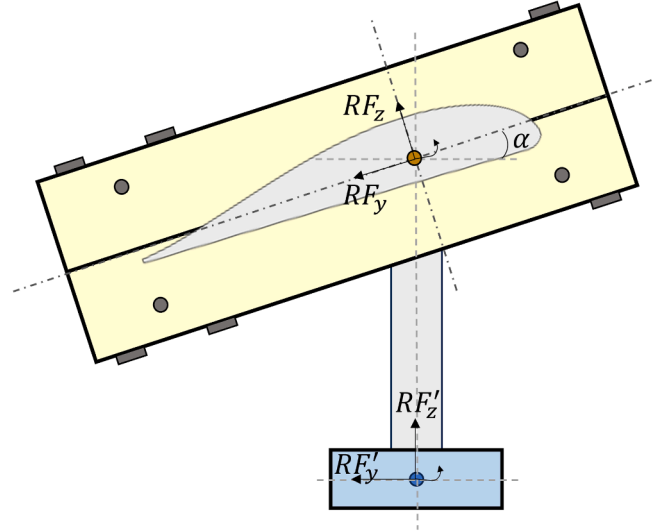


Fig. 12. Schematic illustration of applied and sensed reaction forces.

transient behaviour during the ramp-up and ramp-down phases, as well as stable plateaus under steady flow conditions. Gauges mounted on the upper and lower surfaces record compressive and tensile strains, respectively, consistent with the expected bending response of the wing.

Minor discrepancies are observed between corresponding sensors on the two wings, primarily in the form of small amplitude offsets and local variations in strain magnitude. These differences are attributed to a combination of material variability inherent to composite structures, slight deviations in sensor positioning and bonding, and tolerances in boundary conditions during installation. Importantly, the discrepancies are largely systematic rather than stochastic, and no significant phase lag or drift is observed between the responses of the two wings. At the same time, the presence of small but persistent wing-to-wing differences highlights the limitations of purely physics-based models and motivates the use of the proposed multi-fidelity reconstruction framework to account for such discrepancies in subsequent load estimation analyses.

4.3. Hierarchical reconstruction of aerodynamic loads

The strain time histories in Fig. 13 exhibit small but persistent

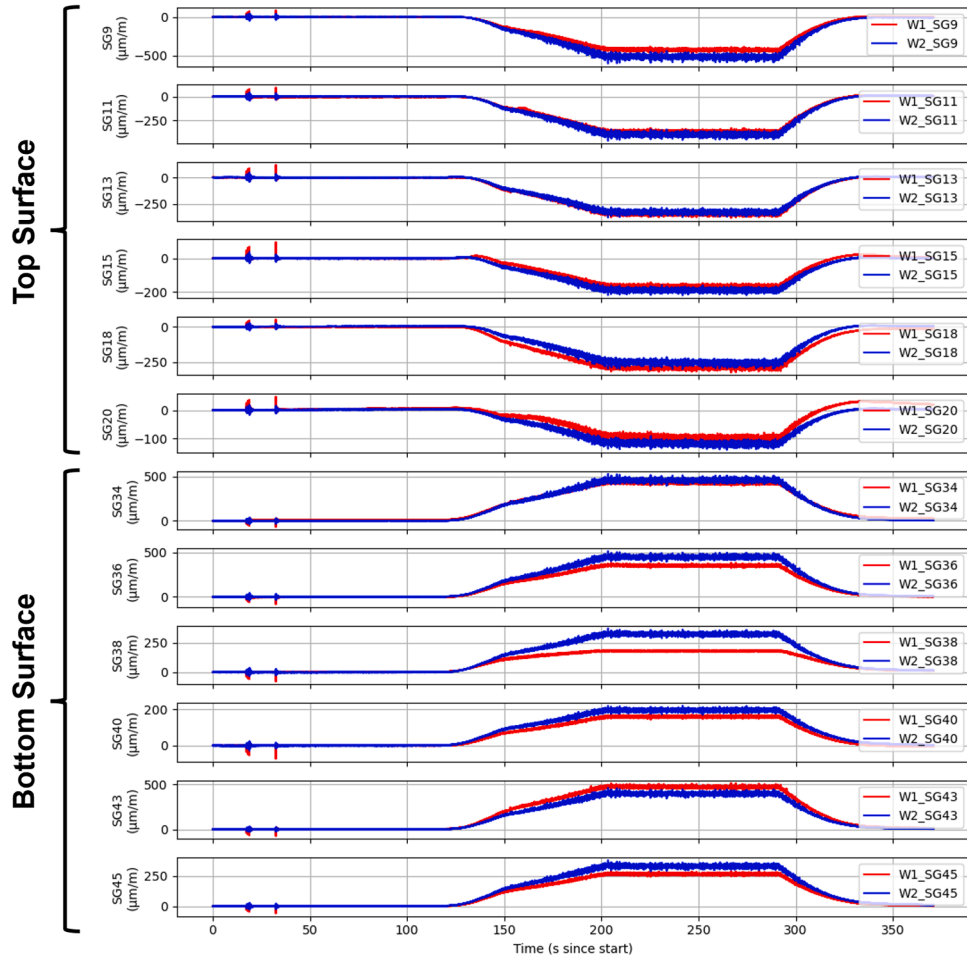


Fig. 13. Strain gauge measurements of both wings for 16 m/s wind velocity and AoA 5°.

differences between nominally identical sensor locations on the left and right wings. While these deviations are expected in composite structures, they can propagate through the inverse mapping and reduce load reconstruction accuracy. The observed mismatch arises from a combination of factors, including manufacturing and material variability, slight differences in sensor positioning and bonding, boundary condition tolerances during installation, and minor geometric misalignment between the wings. To account for these effects, the discrepancy term introduced in Section 2.3 is employed to capture the residual error associated with both measurement noise and modelling inaccuracies inherent to the CM-based reconstruction. The discrepancy term is approximated by training an ANN on the error between the estimated reaction forces from the CM method and the true experimental response using the data collected from the wind tunnel campaign.

For reconstruction, each wing is treated independently and the total aerodynamic forces are obtained by summing the contributions estimated from the left and right wings. To assess and compare the proposed approach three reconstruction strategies are considered:

- **Case 1: CM Method (Baseline)** – The reaction forces are reconstructed directly using the CM formulation described in Eq. (6). This approach serves as the baseline for evaluating reconstruction accuracy.
- **Case 2: ANN Only (Data-driven)** – A neural network model is trained to estimate the reaction forces solely based on the measured strain data. This purely data-driven case relies entirely on the correlations learned from the available training dataset without incorporating any physics-based information.

- **Case 3: Proposed Multi-Fidelity Method (Hybrid)** – The proposed multi-fidelity framework augments the physics-based CM reconstruction with an ANN trained on the discrepancy correction learned from the experimental data.

Prior to aerodynamic load reconstruction, the ANN-based models must be trained using the experimental wind tunnel dataset. Each wind tunnel run listed in Table 3 consists of synchronised strain and reaction-force time histories acquired at 1000 Hz. To focus on steady aerodynamic conditions, the ramp-up and ramp-down portions of each run are removed prior to training, as these transient intervals are not representative of the steady response targeted in this work.

Each retained time instant provides one paired sample, $(\epsilon^m(t), \mathbf{F}_{RF}^{measured}(t))$, where $\epsilon^m(t)$ contains the 12 measured strain values from the selected sensor network and $\mathbf{F}_{RF}^{measured}(t)$ contains the corresponding measured reaction-force components. A training set containing N_{train} wind-tunnel runs therefore consists of all paired strain-force samples extracted from those complete runs, while the remaining runs are used only for testing.

The same experimental dataset is used for both Cases 2 and 3. In Case 2, the ANN is trained to directly map the measured strain vector $\epsilon^m(t)$ to the measured reaction forces $\mathbf{F}_{RF}^{measured}(t)$. In Case 3, instead of predicting the aerodynamic reaction forces directly, the network is trained to estimate the discrepancy between the low-fidelity CM-based reconstruction and the experimentally measured ground truth. Following the formulation introduced in Section 2.3, the CM-based estimate is treated as the low-fidelity prediction:

$$\mathbf{F}_{RF}^{LH}(\epsilon^m) \approx \hat{\mathbf{F}}_{RF}(\epsilon^m) \quad (15)$$

where $\hat{\mathbf{F}}_{RF}(\epsilon^m)$ is obtained from Eq. (6). The discrepancy term is defined as the residual between the measured reaction forces and the low-fidelity estimate and is trained to approximate this residual by minimizing the following loss function of the training dataset $\mathcal{D}$:

$$\mathcal{L} = \|\mathbf{F}_{RF}^{Measured} - (\hat{\mathbf{F}}_{RF}(\epsilon^m) + \text{ANN}(\epsilon^m; \mathbf{W}_{\mathcal{D}}, \mathbf{B}_{\mathcal{D}}))\|_1 \quad (16)$$

During inference, the ANN-predicted discrepancy is added to the CM-estimated reactions to obtain the final corrected load predictions (Eq. (13)). This formulation allows the multi-fidelity model to retain the physical interpretability of the CM method while leveraging the data-driven component to adaptively compensate for systematic modelling errors, measurement noise and experimental uncertainties.

First, the hyperparameter tuning study is presented for Case 3. The experimental dataset contains 16 wind tunnel runs spanning the combinations of wind speed and angle of attack (Table 3). The hyperparameter tuning study is conducted to assess the influence of the ANN architecture on reconstruction performance. The ANN consists of an input layer, an output layer and (N_h) fully connected hidden layers, each with the same width (w_h) and ReLU activation functions. The tested architectural configurations are compared in Table 4, where the performance of each configuration is assessed using the mean and median Mean Absolute Error (MAE) over 10 repeated random run-based splits.

The effect of the ANN hyperparameters is evaluated using a repeated random run-based train/test splitting procedure. Splitting is performed at the level of complete wind tunnel runs, rather than individual time samples, to avoid leakage between highly correlated measurements from the same operating condition. For each configuration, N_{train} wind tunnel runs are randomly selected for training, while the remaining $16 - N_{train}$ runs are held out for testing. For the hyperparameter tuning study, $N_{train} = 10$ corresponding to a 62.5/37.5 split, with 10 runs used for training and 6 runs used for testing. The ANN is trained using the Adam optimiser with a learning rate of $5 \cdot 10^{-4}$ for a maximum of 80 epochs. To reduce the risk of overfitting, early stopping is applied with a patience of 20 epochs. For each hyperparameter combination, the training process is repeated 10 times using different random train/test splits, and the ANN weights are re-initialized at the beginning of each repeat.

The results in Table 4 shows that increasing the ANN size does not

Table 4
Hyperparameter tuning for the discrepancy-correction ANN. Reported values correspond to the mean $\pm$ standard deviation and median MAE over 10 repeated random run-based train/test splits.

N_h	w_h	Total Parameters	RF_y MAE mean $\pm$ std	RF_y MAE median	RF_z MAE mean $\pm$ std	RF_z MAE median
2	20	1182	2.15 $\pm$ 0.43	2.23	5.44 $\pm$ 2.48	4.5
2	40	3962	2.10 $\pm$ 0.36	2.21	3.96 $\pm$ 1.64	3.53
2	80	14322	2.05 $\pm$ 0.29	2.14	3.27 $\pm$ 1.38	3.06
2	100	21902	2.18 $\pm$ 0.42	2.19	3.23 $\pm$ 1.04	3.10
2	200	83802	2.06 $\pm$ 0.32	2.05	3.80 $\pm$ 2.35	3.57
3	20	1602	2.27 $\pm$ 0.33	2.26	4.99 $\pm$ 2.40	4.10
3	40	5602	2.11 $\pm$ 0.34	2.21	3.66 $\pm$ 2.13	3.23
3	80	20802	2.05 $\pm$ 0.28	2.08	3.90 $\pm$ 2.02	2.96
3	100	32002	1.91 $\pm$ 0.37	1.96	4.40 $\pm$ 4.92	2.83
3	200	124002	1.94 $\pm$ 0.24	2.00	5.42 $\pm$ 4.89	3.42

necessarily improve reconstruction accuracy. Although wider networks reduce the error in some cases, the improvement is accompanied by a substantial increase in trainable parameters. For scarce experimental datasets, excessively large ANN architectures are more prone to overfitting because their high number of trainable parameters can fit split-specific noise or experimental artefacts rather than generalizable relationships. The 3×40 ($N_h \times w_h$) configuration is selected as a balanced architecture. It improves upon the smaller models for both reaction forces, while achieving comparable accuracy to wider networks with considerably fewer parameters.

Having selected the 3×40 architecture, the effect of experimental data availability is investigated next for the ANN-based reconstruction strategies. The same architecture is also used for the ANN of Case 2 to ensure a consistent comparison. For this purpose, the number of wind tunnel runs used for training, N_{train} is varied from 2 to 14. For each value of N_{train} , Cases 2 and 3 are trained and tested 10 times, where each time a random run-based train/test split is used. This procedure enables the reconstruction performance and robustness of the two ANN-based approaches to be assessed under progressively increasing levels of experimental data availability, while reducing dependence on any single split. For consistency, at each iteration, Case 1 is evaluated on the same held-out test runs, although it does not require experimental training.

Fig. 14 compares the reconstructed and experimentally measured lift and drag forces for a representative random iteration when only $N_{train} = 3$ wind tunnel runs are used during training. For completeness and for better comparison, the performance of each method is expressed through MAE, RMSE and Mean Absolute Percentage Error (MAPE). The high MAPE value should be interpreted with caution as values close to zero are included. Under this highly data-scarce regime, Case 1 exhibits the lowest accuracy. The drag component shows limited agreement with the experimental measurements. For the lift component, Case 1 performs better than for drag and the predictions follow the general increasing trend of the measured force. However, there is systematic deviation from the one-to-one line, indicating that the CM alone cannot fully reproduce the experimental response. In contrast, both ANN-based approaches improve reconstruction accuracy in this data-scarce case. For drag, in this representative random split, Case 2 gives the lowest error, although Cases 2 and 3 perform similarly. For lift, Case 3 performs best, indicating that the CM-based estimate provides a useful physics-informed prior under limited training data. Both ANN-based methods also reduce the lift MAPE compared with the CM baseline.

Although the results reported in Fig. 14 provide useful indications of the performance of each case, the values are influenced by the specific runs selected for training and testing. Table 5 provides the comparison across the 10 repetitions when $N_{train} = 3$. For RF_y , both ANN-based approaches improve substantially over the CM baseline, reducing the MAE from 8.94 in Case 1 to 4.21 and 4.08 in Cases 2 and 3, respectively. Similar trends are observed for RMSE, which decreases from 9.69 in Case 1 to 5.07 and 4.95 for Cases 2 and 3. The MAPE for RF_y is also reduced, although these values should be interpreted cautiously because the measured drag force includes values close to zero.

For RF_z , the advantage of the multi-fidelity formulation is clearer under this data-scarce condition. Case 3 reduces the mean MAE from 8.72 in Case 1 to 4.58, and the RMSE from 11.78 to 6.03. The corresponding MAPE decreases from 49.23% to 18.44%, indicating a substantial improvement in relative prediction accuracy. In contrast, Case 2 shows high variability for RF_z , with a mean MAE of 18.27 and RMSE of 24.27, although its median errors are lower than the mean values. This indicates strong sensitivity to the selected training and test runs and suggests reduced generalisation of the purely data-driven model under highly limited experimental data.

Next, the case with $N_{train} = 10$ is considered. Fig. 15 reports the reconstructed reaction forces for a representative random split with $N_{train} = 10$. As expected, the increased availability of experimental data improves the performance of both ANN-based approaches relative to the CM baseline. the representative split, both Cases 2 and 3 substantially

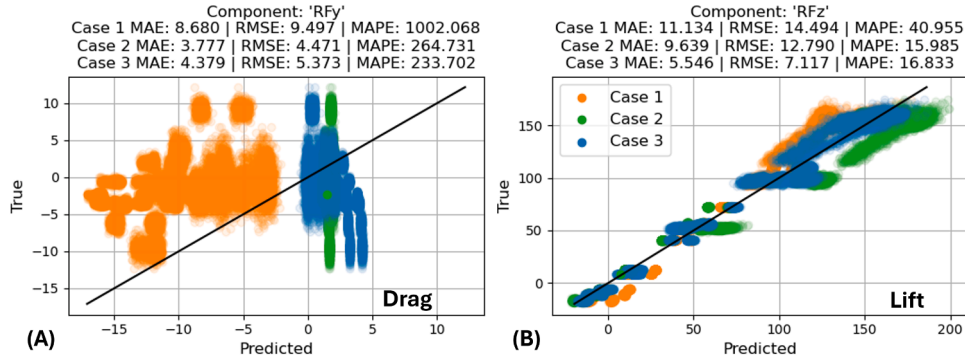


Fig. 14. Comparison of predictive accuracy when 3 experimental runs are used for training: (A) reconstructed drag component (RF_y) and (B) reconstructed lift component (RF_z).

Table 5

Repeated training summary for $N_{train} = 3$.

Case	Metric	RF_y mean $\pm$ std	RF_y median	RF_z mean $\pm$ std	RF_z median
Case 1	MAE	8.94 ± 0.46	8.75	8.72 ± 0.25	8.82
	RMSE	9.69 ± 0.41	9.56	11.78 ± 0.34	11.91
	MAPE	988.04 ± 43.44	995.63	49.23 ± 9.10	48.42
Case 2	MAE	4.21 ± 0.71	3.92	18.27 ± 15.38	11.00
	RMSE	5.07 ± 0.90	4.56	24.27 ± 19.99	14.63
	MAPE	324.30 ± 151.67	272.56	110.09 ± 177.46	20.13
Case 3	MAE	4.08 ± 0.45	3.89	4.58 ± 1.47	4.12
	RMSE	4.95 ± 0.54	4.71	6.03 ± 1.55	5.43
	MAPE	262.87 ± 39.11	268.25	18.44 ± 9.20	17.81

reduce the drag error compared with Case 1, with Case 3 giving the lowest MAE/RMSE values. For lift, both ANN-based models also improve over Case 1, with Case 3 providing the lowest MAE and MAPE.

The results across all 10 repetitions with $N_{train} = 10$ in Table 6 confirm the performance improvement after the available data are increased. For RF_y , Cases 2 and 3 both outperform the CM baseline, reducing the mean MAE from 8.44 in Case 1 to 4.34 and 3.91, respectively. This corresponds to reductions of approximately 49% and 54%. A similar improvement is observed in RMSE, which decreases from 9.12 in Case 1 to 5.43 and 4.86 for Cases 2 and 3. The MAPE for RF_y is also reduced, although the percentage errors remain high because of near-zero drag values.

For RF_z , Case 2 now improves over the CM baseline, reducing the mean MAE from 8.73 to 7.39 and the RMSE from 11.74 to 9.63. However, Case 3 achieves the best overall lift reconstruction, with a mean MAE of 3.85 and RMSE of 5.13, corresponding to reductions of

approximately 56% relative to Case 1 for both metrics. The improvement is also reflected in the relative percentage error. The mean MAPE decreases from 47.51% in Case 1 to 13.33% in Case 2 and 12.92% in Case 3.

By comparing the results in Table 5 and Table 6, it is observed that the purely data-driven (Case 2) and multi-fidelity (Case 3) approaches both improve when a rich training dataset is available. This convergence is expected, as sufficiently rich training data allow ANNs to learn the nonlinear mappings between measured strains and aerodynamic loads with high accuracy through interpolation within the observed operating envelope. Under such conditions, the reliance on prior physical information becomes less critical. This observation naturally raises a key question for practical SHM applications: *What is the added value of a multi-fidelity approach when purely data-driven models can eventually achieve similar accuracy?*

To investigate the influence of data availability on reconstruction

Table 6

Repeated training summary for $N_{train} = 10$.

Case	Metric	RF_y mean $\pm$ std	RF_y median	RF_z mean $\pm$ std	RF_z median
Case 1	MAE	8.44 ± 0.59	8.40	8.73 ± 1.00	8.83
	RMSE	9.12 ± 0.62	9.23	11.74 ± 1.04	11.76
	MAPE	1045.65 ± 228.35	1098.17	47.51 ± 19.89	46.13
Case 2	MAE	4.34 ± 1.05	4.42	7.39 ± 3.19	7.17
	RMSE	5.43 ± 1.08	5.38	9.63 ± 3.90	9.79
	MAPE	426.31 ± 332.94	278.56	13.33 ± 3.21	12.91
Case 3	MAE	3.91 ± 0.68	3.70	3.85 ± 2.11	3.44
	RMSE	4.86 ± 0.80	4.61	5.13 ± 2.87	4.42
	MAPE	283.77 ± 167.50	217.72	12.92 ± 7.18	11.59

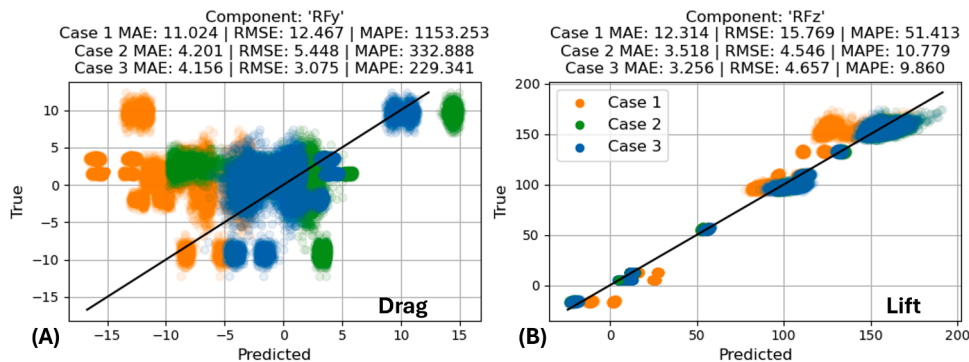


Fig. 15. Comparison of predictive accuracy when 10 experimental runs are used for training: (A) reconstructed drag component (RF_y) and (B) reconstructed lift component (RF_z).

accuracy, Fig. 16 reports the distribution of RF_z MAE values obtained from repeated random run-based train/test splits as N_{train} increases from 2 to 14. Case 1 shows an almost constant level of performance, as expected for the CM-based method, with some increased variability at higher N_{train} values because fewer runs remain available for testing and the results become more sensitive to the specific held-out operating conditions.

The results show that both ANN-based approaches improve as more wind tunnel runs are included in training. However, the purely data-driven Case 2 approach exhibits substantially larger variability under low-data conditions, with several long upper tails. This indicates that when the experimental dataset is sparse, the direct strain-to-load mapping may fail to generalise reliably to unseen aerodynamic states.

In contrast, the multi-fidelity approach (Case 3) shows lower median errors and a narrower error distribution over most of the data-scarce regime. This indicates that the CM-based estimate provides a useful physics-informed prior, reducing the learning effort on the ANN and improving robustness when experimental data are limited. As N_{train} increases, the performance of Case 2 and Case 3 gradually converges, with both approaches reaching lower and more stable MAE values.

While Case 2 can achieve high performance given sufficient experimental data, the multi-fidelity approach (Case 3), grounded by a physics-based prior, reaches higher accuracy significantly faster and with fewer training runs. After being trained on only $N_{train} = 3$ wind tunnel runs, Case 3 achieves an RF_z MAE of 4.58 ± 1.47 , with a median value of 4.12 and clearly outperforms the baseline set by Case 1. This is already within the same order of accuracy as the results obtained with substantially larger training sets, where the median MAE for $N_{train} = 10 - 14$ ranges approximately between 2.65 and 3.57. The main accuracy improvement for Case 3 is achieved early, while additional training runs further reduce the error and stabilize the predictions Case 2 requires approximately $N_{train} = 6$ training runs before its mean and median errors consistently fall below those of Case 1, and around $N_{train} = 8$ training runs before its predictions become more stable and consistently concentrated in a lower-error range. Even when Case 2 becomes competitive at higher training sizes, its low-data variability remains larger than Case 3. The full summary statistics for each run are included in the Appendix, Tables A1-A13.

In practical terms, these results highlight the data efficiency advantage of the multi-fidelity approach which may reduce the experimental burden associated with wind tunnel testing. By achieving reliable reconstruction with fewer training runs, the method may lessen the need to densely sample the full operational envelope, translating in savings in

experimental effort and costs.

4.4. Limitations of the proposed framework and future work

The results presented demonstrate the potential of the proposed multi-fidelity framework for aerodynamic load reconstruction under sparse sensing and limited experimental data. However, several limitations should be noted. First, the FE model used to construct the CM is an equivalent isotropic representation of a proprietary ASTERO wing. This model is intended to capture the global elastic response relevant to lift and drag reconstruction, rather than the detailed ply-level orthotropic response or local strain field of the laminate. The calibration against the first two natural frequencies provides a validation of the global stiffness-mass behaviour, but it does not guarantee reproduction of the full spatial strain field across the wing surface under all possible loading conditions. At the same time, the sensitivity of the CM to uncertainties in material properties, geometric tolerances and FE modelling assumptions has not been quantified.

Furthermore, the CM is constructed using linear static unit load simulations. The present formulation therefore targets quasi-steady aerodynamic load reconstruction within the stable operating conditions tested in the wind tunnel. Dynamic aeroelastic phenomena, such as flutter or strongly coupled fluid-structure instabilities, are outside the scope of the current study. This choice is consistent with the experimental campaign, where operating points associated with significant flow separation and structural vibrations were excluded. In total, the experimental validation is based on 16 wind tunnel operating conditions. Although repeated random run-based train/test splitting and retraining were used to assess data efficiency and reduce dependence on a particular random split, the dataset remains limited in terms of distinct aerodynamic states. Additional states need to be assessed including different pitch, roll and aileron positions. The reported results should therefore be interpreted as evidence of performance within the tested operating envelope, rather than as proof of generalisation to arbitrary in-flight manoeuvres or untested aeroelastic regimes.

An additional limitation concerns the measured drag component. The drag force recorded by the wind tunnel transducer may include contributions from the mounting fixture and attached nose, which are captured in the measured reaction force but may not produce corresponding strain changes in the wing structure used to construct the CM. Consequently, future studies using a single-wing configuration in the wind tunnel, or improved aerodynamic isolation of the mounting system, could mitigate this effect and improve the observability of the drag

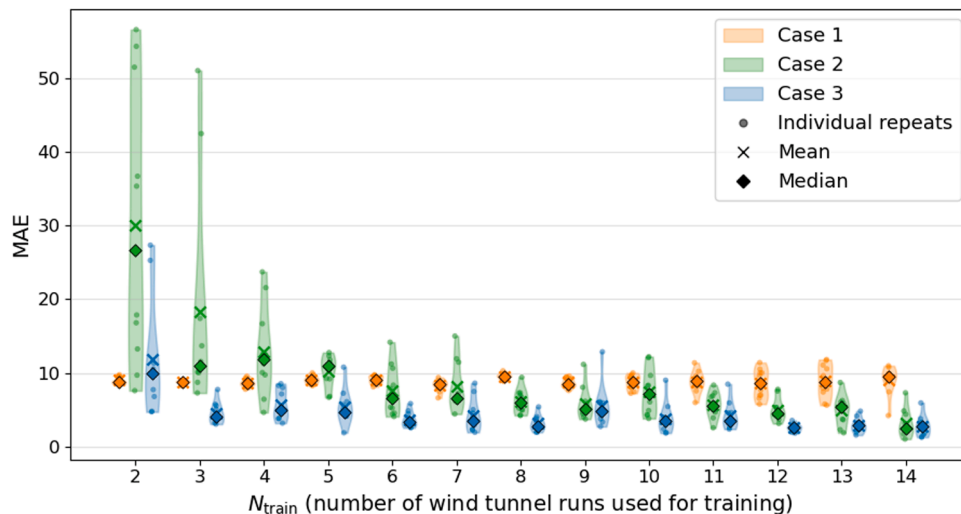


Fig. 16. Distribution of RF_z MAE as a function of the number of wind-tunnel runs used for training, N_{train} . Violin plots show the error distribution across repeated random run-based splits for Cases 1, 2 and 3.

force from the wing strain measurements.

Finally, the ANN component is used to learn residual discrepancies between the CM-based estimate and the measured response. This reduces the learning burden compared with a fully data-driven strain-to-load mapping, but it does not eliminate the need for representative experimental data. Although the results indicate that the proposed framework can significantly reduce data requirements, future work should extend the validation to larger experimental datasets, additional flight-like loading conditions, uncertainty propagation in the CM and a formal sensitivity analysis, and online adaptation using in-flight measurements.

5. Conclusions

This study presented a multi-fidelity framework for aerodynamic load reconstruction in composite UAV wings using sparse strain measurements. The proposed approach integrates a physics-based calibration matrix derived from FE simulations with a data-driven ANN trained to learn residual discrepancies between numerical predictions and experimental observations. A multi-objective sensor placement strategy was employed to identify strain-gauge configurations that balance reconstruction accuracy against practical instrumentation constraints.

Wind tunnel experiments conducted on the composite wings provided experimental validation of the framework under controlled but representative operating conditions. The results demonstrated that the multi-fidelity reconstruction strategy substantially improves load estimation accuracy compared to a CM-only approach. Even with as few as three experimental training runs, the proposed method reduced the mean MAE from 8.94 to 4.08 for the drag component and from 8.72 to 4.58 for the lift component, corresponding to reductions of approximately 54% and 47%, respectively. For lift, the MAPE was also reduced from 49.23% to 18.44%, indicating improved relative prediction accuracy under data-scarce conditions. Adding more training data further improves the reconstruction performance. When trained with 10 experimental runs, the multi-fidelity approach continued to outperform the CM-only baseline. In particular, the lift MAPE decreased from 47.51% to 12.92%.

The repeated random run-based train/test analysis further showed that the main advantage of the multi-fidelity approach lies in its data efficiency. The purely data-driven approach can achieve competitive performance compared to the proposed method when sufficient experimental data are available, but it exhibits larger variability when trained with only a small number of wind-tunnel runs. In contrast, the multi-fidelity formulation reaches low and stable errors earlier. This behaviour confirms that embedding physics-based constraints within the learning process can improve data efficiency and prediction stability within the tested operating envelope. Therefore, the proposed multi-fidelity approach provides promising load-reconstruction accuracy under the data-scarce conditions examined, a critical advantage for UAV

SHM systems where extensive experimental data collection is impractical.

Beyond improved reconstruction accuracy, the framework offers additional practical benefits. The CM preserves physical interpretability, enabling insight into load-strain relationships, while the discrepancy-learning formulation limits the learning task to systematic modelling and measurement errors. Together with the sensor placement optimization, this holistic methodology supports deployable, resource-efficient SHM solutions for lightweight composite structures. The results demonstrate that combining physics-informed modelling with targeted data-driven correction provides a promising pathway toward higher-fidelity load reconstruction in UAV applications, subject to further validation under broader flight-like conditions. Future work will focus on extending the framework toward online adaptation and in-flight learning, enabling real-time digital twin implementations for operational monitoring and decision support.

Declaration of generative AI and AI-assisted technologies in the writing process

During the preparation of this manuscript, the authors used ChatGPT (version 5.1) for language editing and to improve grammatical and syntactical clarity. Following its use, the authors carefully reviewed and edited the content as necessary and take full responsibility for the final published version.

CRediT authorship contribution statement

Ilias N. Giannakeas: Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Yuhang Pan:** Writing – review & editing, Investigation, Data curation. **Sadjad Naderi:** Writing – review & editing, Data curation. **Zahra Sharif Khodaei:** Writing – review & editing, Supervision, Investigation, Funding acquisition. **Ferri M.H. Aliabadi:** Writing – review & editing, Supervision, Investigation, Funding acquisition.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This project has received funding from the European Climate, Infrastructure and Environment Executive Agency (CINEA) under Grant agreement No. 101096073, and from United Kingdom Research and Innovation (UKRI) under Grant agreement No. 10065739.

Appendix. Repeated training summaries for N_{train} in the range from 2 to 14

MAPE values should be interpreted cautiously because measured values approach zero.

Table A1

Repeated training summary for $N_{train} = 2$.

Case	Metric	RFy mean $\pm$ std	RFy median	RFz mean $\pm$ std	RFz median
Case 1	MAE	8.73 $\pm$ 0.18	8.74	9.01 $\pm$ 0.45	8.82
	RMSE	9.51 $\pm$ 0.19	9.52	12.07 $\pm$ 0.37	11.93
	MAPE	989.99 $\pm$ 92.84	1032.93	49.95 $\pm$ 7.38	52.94
Case 2	MAE	4.17 $\pm$ 1.09	3.88	29.95 $\pm$ 19.27	26.59
	RMSE	5.03 $\pm$ 1.47	4.65	36.57 $\pm$ 23.29	30.92
	MAPE	381.58 $\pm$ 269.48	254.00	130.18 $\pm$ 121.19	89.83

(continued on next page)

Table A1 (continued)

Case	Metric	RFy mean $\pm$ std	RFy median	RFz mean $\pm$ std	RFz median
Case 3	MAE	4.51 $\pm$ 0.75	4.37	11.85 $\pm$ 7.96	9.99
	RMSE	5.65 $\pm$ 0.95	5.40	14.61 $\pm$ 8.90	12.79
	MAPE	357.50 $\pm$ 199.79	277.53	52.97 $\pm$ 44.65	39.02

Table A2Repeated training summary for $N_{train} = 3$.

Case	Metric	RFy mean $\pm$ std	RFy median	RFz mean $\pm$ std	RFz median
Case 1	MAE	8.94 $\pm$ 0.46	8.75	8.72 $\pm$ 0.25	8.82
	RMSE	9.69 $\pm$ 0.41	9.56	11.78 $\pm$ 0.34	11.91
	MAPE	988.04 $\pm$ 43.44	995.63	49.23 $\pm$ 9.10	48.42
Case 2	MAE	4.21 $\pm$ 0.71	3.92	18.27 $\pm$ 15.38	11.00
	RMSE	5.07 $\pm$ 0.90	4.56	24.27 $\pm$ 19.99	14.63
	MAPE	324.30 $\pm$ 151.67	272.56	110.09 $\pm$ 177.46	20.13
Case 3	MAE	4.08 $\pm$ 0.45	3.89	4.58 $\pm$ 1.47	4.12
	RMSE	4.95 $\pm$ 0.54	4.71	6.03 $\pm$ 1.55	5.43
	MAPE	262.87 $\pm$ 39.11	268.25	18.44 $\pm$ 9.20	17.81

Table A3Repeated training summary for $N_{train} = 4$.

Case	Metric	RFy mean $\pm$ std	RFy median	RFz mean $\pm$ std	RFz median
Case 1	MAE	8.94 $\pm$ 0.18	8.96	8.69 $\pm$ 0.53	8.59
	RMSE	9.73 $\pm$ 0.18	9.75	11.86 $\pm$ 0.58	11.93
	MAPE	973.92 $\pm$ 116.09	965.21	50.24 $\pm$ 6.93	48.79
Case 2	MAE	4.08 $\pm$ 0.79	4.32	12.91 $\pm$ 6.12	11.89
	RMSE	5.01 $\pm$ 0.97	5.31	17.03 $\pm$ 8.41	15.62
	MAPE	330.21 $\pm$ 157.42	282.90	34.39 $\pm$ 32.29	21.23
Case 3	MAE	3.97 $\pm$ 0.46	4.17	5.75 $\pm$ 2.04	4.98
	RMSE	4.90 $\pm$ 0.51	5.05	7.23 $\pm$ 2.36	6.76
	MAPE	286.32 $\pm$ 94.73	260.00	28.07 $\pm$ 18.93	20.04

Table A4Repeated training summary for $N_{train} = 5$.

Case	Metric	RFy mean $\pm$ std	RFy median	RFz mean $\pm$ std	RFz median
Case 1	MAE	8.83 $\pm$ 0.49	8.89	9.06 $\pm$ 0.57	9.09
	RMSE	9.59 $\pm$ 0.49	9.65	12.07 $\pm$ 0.53	12.06
	MAPE	1020.76 $\pm$ 98.35	1008.36	53.51 $\pm$ 9.38	55.68
Case 2	MAE	4.49 $\pm$ 0.90	4.58	10.20 $\pm$ 2.15	10.91
	RMSE	5.69 $\pm$ 0.93	5.88	13.72 $\pm$ 3.13	14.43
	MAPE	375.65 $\pm$ 211.41	297.03	21.38 $\pm$ 12.85	18.01
Case 3	MAE	3.95 $\pm$ 0.38	3.92	5.38 $\pm$ 2.38	4.74
	RMSE	5.04 $\pm$ 0.35	5.00	6.87 $\pm$ 2.66	6.40
	MAPE	247.28 $\pm$ 56.84	226.17	24.76 $\pm$ 23.85	17.50

Table A5Repeated training summary for $N_{train} = 6$.

Case	Metric	RFy mean $\pm$ std	RFy median	RFz mean $\pm$ std	RFz median
Case 1	MAE	8.89 $\pm$ 0.51	8.75	9.11 $\pm$ 0.55	9.12
	RMSE	9.68 $\pm$ 0.45	9.58	12.16 $\pm$ 0.56	12.23
	MAPE	965.28 $\pm$ 152.78	961.39	56.09 $\pm$ 10.79	58.86
Case 2	MAE	4.63 $\pm$ 0.49	4.64	7.59 $\pm$ 3.47	6.62
	RMSE	5.66 $\pm$ 0.66	5.60	10.60 $\pm$ 4.84	8.99
	MAPE	362.54 $\pm$ 159.82	306.00	16.91 $\pm$ 5.99	17.12
Case 3	MAE	4.06 $\pm$ 0.36	4.02	3.71 $\pm$ 1.07	3.39
	RMSE	4.99 $\pm$ 0.34	5.00	4.90 $\pm$ 1.29	4.43
	MAPE	229.28 $\pm$ 39.33	222.61	18.45 $\pm$ 8.32	16.70

Table A6Repeated training summary for $N_{train} = 7$.

Case	Metric	RFy mean $\pm$ std	RFy median	RFz mean $\pm$ std	RFz median
Case 1	MAE	9.08 $\pm$ 0.51	8.94	8.30 $\pm$ 0.79	8.53
	RMSE	9.73 $\pm$ 0.46	9.69	11.32 $\pm$ 1.05	11.76
	MAPE	1040.08 $\pm$ 133.07	1044.11	52.98 $\pm$ 14.08	47.76
Case 2	MAE	3.93 $\pm$ 0.95	3.87	8.23 $\pm$ 3.35	6.62
	RMSE	5.01 $\pm$ 1.02	4.79	10.51 $\pm$ 3.53	8.90
	MAPE	288.30 $\pm$ 198.20	242.59	20.75 $\pm$ 20.65	13.97
Case 3	MAE	4.04 $\pm$ 0.65	3.91	4.23 $\pm$ 2.26	3.56
	RMSE	4.86 $\pm$ 0.66	4.74	5.60 $\pm$ 2.95	4.59
	MAPE	241.79 $\pm$ 54.05	218.59	23.56 $\pm$ 21.58	16.75

Table A7Repeated training summary for $N_{train} = 8$.

Case	Metric	RFy mean $\pm$ std	RFy median	RFz mean $\pm$ std	RFz median
Case 1	MAE	9.17 $\pm$ 0.70	9.15	9.51 $\pm$ 0.45	9.57
	RMSE	9.93 $\pm$ 0.66	9.97	12.45 $\pm$ 0.48	12.56
	MAPE	970.66 $\pm$ 179.53	939.00	65.50 $\pm$ 10.90	69.31
Case 2	MAE	4.82 $\pm$ 0.59	4.85	6.15 $\pm$ 1.67	5.97
	RMSE	5.88 $\pm$ 0.59	6.06	8.37 $\pm$ 2.06	7.97
	MAPE	316.64 $\pm$ 190.32	261.60	17.87 $\pm$ 8.69	14.95
Case 3	MAE	4.50 $\pm$ 0.58	4.55	3.17 $\pm$ 1.05	2.79
	RMSE	5.44 $\pm$ 0.62	5.45	4.27 $\pm$ 1.36	3.88
	MAPE	232.24 $\pm$ 29.33	228.69	13.83 $\pm$ 3.95	13.08

Table A8Repeated training summary for $N_{train} = 9$.

Case	Metric	RFy mean $\pm$ std	RFy median	RFz mean $\pm$ std	RFz median
Case 1	MAE	9.05 $\pm$ 0.84	9.19	8.72 $\pm$ 0.65	8.45
	RMSE	9.79 $\pm$ 0.75	9.78	11.86 $\pm$ 0.68	11.84
	MAPE	934.15 $\pm$ 269.97	931.20	55.59 $\pm$ 22.55	56.29
Case 2	MAE	3.91 $\pm$ 0.48	3.96	5.84 $\pm$ 2.24	5.17
	RMSE	4.91 $\pm$ 0.67	5.04	7.80 $\pm$ 2.60	7.30
	MAPE	246.71 $\pm$ 67.96	264.11	16.52 $\pm$ 11.82	12.22
Case 3	MAE	4.03 $\pm$ 0.90	4.08	5.49 $\pm$ 2.84	4.87
	RMSE	4.96 $\pm$ 1.04	4.91	7.14 $\pm$ 3.05	6.10
	MAPE	213.57 $\pm$ 36.03	220.02	31.17 $\pm$ 38.68	17.43

Table A9Repeated training summary for $N_{train} = 10$.

Case	Metric	RFy mean $\pm$ std	RFy median	RFz mean $\pm$ std	RFz median
Case 1	MAE	8.44 $\pm$ 0.59	8.40	8.73 $\pm$ 1.00	8.83
	RMSE	9.12 $\pm$ 0.62	9.23	11.74 $\pm$ 1.04	11.76
	MAPE	1045.65 $\pm$ 228.35	1098.17	47.51 $\pm$ 19.89	46.13
Case 2	MAE	4.34 $\pm$ 1.05	4.42	7.39 $\pm$ 3.19	7.17
	RMSE	5.43 $\pm$ 1.08	5.38	9.63 $\pm$ 3.90	9.79
	MAPE	426.31 $\pm$ 332.94	278.56	13.33 $\pm$ 3.21	12.91
Case 3	MAE	3.91 $\pm$ 0.68	3.70	3.85 $\pm$ 2.11	3.44
	RMSE	4.86 $\pm$ 0.80	4.61	5.13 $\pm$ 2.87	4.42
	MAPE	283.77 $\pm$ 167.50	217.72	12.92 $\pm$ 7.18	11.59

Table A10Repeated training summary for $N_{train} = 11$.

Case	Metric	RFy mean $\pm$ std	RFy median	RFz mean $\pm$ std	RFz median
Case 1	MAE	9.11 $\pm$ 0.86	9.00	8.83 $\pm$ 1.48	8.90
	RMSE	9.75 $\pm$ 0.78	9.68	11.71 $\pm$ 1.64	11.82
	MAPE	1016.14 $\pm$ 261.27	935.75	59.49 $\pm$ 23.02	55.29
Case 2	MAE	4.32 $\pm$ 1.27	4.28	5.67 $\pm$ 1.73	5.61
	RMSE	5.39 $\pm$ 1.54	5.77	7.55 $\pm$ 2.19	7.43

(continued on next page)

Table A10 (continued)

Case	Metric	RFy mean $\pm$ std	RFy median	RFz mean $\pm$ std	RFz median
Case 3	MAPE	265.83 $\pm$ 86.68	244.89	14.94 $\pm$ 8.04	12.79
	MAE	4.05 $\pm$ 0.75	4.04	4.19 $\pm$ 1.85	3.57
	RMSE	4.86 $\pm$ 0.87	4.77	5.34 $\pm$ 2.15	4.53
	MAPE	228.04 $\pm$ 41.51	244.18	21.41 $\pm$ 16.19	13.96

Table A11

Repeated training summary for $N_{train} = 12$.

Case	Metric	RFy mean $\pm$ std	RFy median	RFz mean $\pm$ std	RFz median
Case 1	MAE	9.05 $\pm$ 1.09	8.79	8.55 $\pm$ 1.97	8.65
	RMSE	9.66 $\pm$ 0.99	9.47	11.33 $\pm$ 1.95	11.52
	MAPE	994.73 $\pm$ 170.32	1011.99	57.77 $\pm$ 24.02	56.77
Case 2	MAE	4.39 $\pm$ 1.44	4.09	5.03 $\pm$ 1.55	4.58
	RMSE	5.70 $\pm$ 1.57	5.42	6.72 $\pm$ 1.93	5.99
	MAPE	330.18 $\pm$ 340.90	216.41	12.84 $\pm$ 4.22	12.42
Case 3	MAE	4.07 $\pm$ 0.55	3.78	2.55 $\pm$ 0.52	2.65
	RMSE	4.91 $\pm$ 0.66	4.73	3.41 $\pm$ 0.74	3.41
	MAPE	215.93 $\pm$ 32.62	218.67	11.34 $\pm$ 5.83	11.03

Table A12

Repeated training summary for $N_{train} = 13$.

Case	Metric	RFy mean $\pm$ std	RFy median	RFz mean $\pm$ std	RFz median
Case 1	MAE	8.73 $\pm$ 0.86	8.61	8.93 $\pm$ 2.33	8.78
	RMSE	9.45 $\pm$ 0.84	9.43	12.03 $\pm$ 2.25	12.11
	MAPE	823.68 $\pm$ 176.72	746.70	56.31 $\pm$ 33.66	59.43
Case 2	MAE	4.16 $\pm$ 1.00	4.15	4.93 $\pm$ 2.05	5.43
	RMSE	5.49 $\pm$ 1.17	5.90	6.57 $\pm$ 2.46	7.12
	MAPE	227.34 $\pm$ 51.38	227.46	11.51 $\pm$ 3.06	11.44
Case 3	MAE	4.15 $\pm$ 0.46	4.03	2.94 $\pm$ 0.99	2.85
	RMSE	5.13 $\pm$ 0.62	5.07	3.95 $\pm$ 1.20	3.98
	MAPE	209.55 $\pm$ 35.78	199.78	11.77 $\pm$ 8.57	11.02

Table A13

Repeated training summary for $N_{train} = 14$.

Case	Metric	RFy mean $\pm$ std	RFy median	RFz mean $\pm$ std	RFz median
Case 1	MAE	9.42 $\pm$ 1.01	9.38	9.07 $\pm$ 1.96	9.56
	RMSE	10.12 $\pm$ 0.99	10.17	11.82 $\pm$ 2.17	12.62
	MAPE	1099.00 $\pm$ 312.54	1064.20	69.75 $\pm$ 32.85	74.09
Case 2	MAE	4.16 $\pm$ 1.38	3.91	3.17 $\pm$ 1.86	2.44
	RMSE	5.31 $\pm$ 1.92	5.25	4.33 $\pm$ 2.67	3.35
	MAPE	256.79 $\pm$ 56.64	245.27	9.76 $\pm$ 4.76	8.10
Case 3	MAE	3.99 $\pm$ 1.14	3.92	2.75 $\pm$ 1.41	2.70
	RMSE	4.87 $\pm$ 1.47	4.93	3.60 $\pm$ 1.74	3.67
	MAPE	229.18 $\pm$ 69.59	205.21	14.09 $\pm$ 8.04	13.17

Data availability

Data will be made available on request.

References

- [1] G. Sala, L. Di Landro, A. Airolidi, P. Bettini, Fibre optics health monitoring for aeronautical applications, *Meccanica* 50 (10) (2015) 2547–2567.
- [2] M.F. Aliabadi, Z.S. Khodaei, Structural Health Monitoring for Advanced Composite Structures, 8, World Scientific, 2017.
- [3] C. Sun, et al., A multi-model architecture based on deep learning for aircraft load prediction, *Commun. Eng.* 2 (1) (2023) 47.
- [4] C.Y. Park, J.H. Kim, S.-M. Jun, A structural health monitoring project for a composite unmanned aerial vehicle wing: Overview and evaluation tests, *Struct. Control Health Monit.* 19 (7) (2012) 567–579.
- [5] M. Xiong, H. Wang, Digital twin applications in aviation industry: a review, *Int. J. Adv. Manuf. Technol.* 121 (9) (2022) 5677–5692.
- [6] X. Zhou, S. He, L. Dong, S.N. Atluri, Real-time prediction of probabilistic crack growth with a helicopter component digital twin, *Aiaa J.* 60 (4) (2022) 2555–2567.
- [7] J. He, Y. Zhou, X. Guan, W. Zhang, Y. Wang, W. Zhang, An integrated health monitoring method for structural fatigue life evaluation using limited sensor data, *Materials* 9 (11) (2016) 894.
- [8] G. Souza, J. Tarpani, Distributed fiber optics sensing applied to laminated composites: embedding process, strain field monitoring with OBR and fracture mechanisms, *J. Nondestruct. Eval.* 39 (4) (2020) 77.
- [9] K. Loupos, A. Amditis, Structural health monitoring fiber optic sensors. *Fiber Optic Sensors: Current Status and Future Possibilities*, Springer, 2016, pp. 185–206.
- [10] L. Colombo, C. Sbarufatti, W. Zielinski, K. Dragan, M. Giglio, Numerical and experimental flight verifications of a calibration matrix approach for load monitoring and temperature reconstruction and compensation, *Aerosp. Sci. Technol.* 118 (2021) 107074.
- [11] W. Akl, S. Poh, A. Baz, Wireless and distributed sensing of the shape of morphing structures, presented at the. *Smart Structures and Materials 2006: Sensors and*

- Smart Structures Technologies for Civil, Mechanical, and Aerospace Systems, SPIE, 2006, pp. 267–275.
- [12] M. Gherlone, P. Cerracchio, M. Mattone, Shape sensing methods: Review and experimental comparison on a wing-shaped plate, *Prog. Aerosp. Sci.* 99 (2018) 14–26.
 - [13] R. Liu, E. Dobriban, Z. Hou, K. Qian, Dynamic load identification for mechanical systems: a review, *Arch. Comput. Methods Eng.* 29 (2) (2022) 831–863.
 - [14] M. Esposito, M. Gherlone, P. Marzocca, External loads identification and shape sensing on an aluminum wing box: an integrated approach, *Aerosp. Sci. Technol.* 114 (2021) 106743.
 - [15] M.J. Candon, O. Levinski, A. Altaf, R. Carrese, P. Marzocca, Aircraft transonic buffet load prediction using artificial neural networks, presented at the, in: *AIAA Science and Technology Forum and Exposition 2019*, 2019, p. 0763.
 - [16] D. Poloni, D. Oboe, C. Sbarufatti, M. Giglio, Towards a stochastic inverse finite element method: a Gaussian process strain extrapolation, *Mech. Syst. Signal Process.* 189 (2023) 110056.
 - [17] Y. Qiu, P.K. Arunachala, C. Linder, Sensenet: A physics-informed deep learning model for shape sensing, *J. Eng. Mech.* 149 (3) (2023) 04023002.
 - [18] A. Tessler, J.L. Spangler, A least-squares variational method for full-field reconstruction of elastic deformations in shear-deformable plates and shells, *Comput. Methods Appl. Mech. Eng.* 194 (2–5) (2005) 327–339.
 - [19] W. L. Ko, W. L. Richards, and V. T. Fleischer, “Applications of Ko displacement theory to the deformed shape predictions of the doubly-tapered Ikhan Wing,” 2009.
 - [20] S. Bionda, E. DI GALLEONARDO, G. Cazzulani, F. Braghin, Development of a new algorithm for wheel-rail contact force estimation using reduced measurement sets, presented at the, in: *Proceedings of the 14th International Conference on Vehicle System Dynamics, Identification and Anomalies*, 2014, pp. 1–12.
 - [21] H. Abdel-Jaber, B. Glisic, Monitoring of prestressing forces in prestressed concrete structures—an overview, *Struct. Control Health Monit.* 26 (8) (2019) e2374.
 - [22] M. Esposito, M. Mattone, M. Gherlone, Experimental shape sensing and load identification on a stiffened panel: a comparative study, *Sensors* 22 (3) (2022) 1064.
 - [23] J. Kyung, H.-C. Eun, An optimal strain gauge layout design for the measurement of truss structures, *Sensors* 23 (5) (2023) 2738.
 - [24] H. Peng, B. Wang, Y. Ning, S. Cao, M. Liu, Strain Gauge Location Optimization for Operational Load Monitoring of an Aircraft Wing Using an Improved Correlation Measure, *Appl. Sci.* 14 (19) (2024) 9078.
 - [25] V. Mallardo, M. Aliabadi, Z.S. Khodaei, Optimal sensor positioning for impact localization in smart composite panels, *J. Intell. Mater. Syst. Struct.* 24 (5) (2013) 559–573.
 - [26] R. Semaan, Optimal sensor placement using machine learning, *Comput. Fluids* 159 (2017) 167–176.
 - [27] X. Zhou, M. Dziendzikowski, K. Dragan, L. Dong, M. Giglio, C. Sbarufatti, In-service Load Monitoring for an UAV Digital Twin, *E-J. Nondestruct. Test.* 29 (7) (2024) 1–8.
 - [28] J. Duan, Q. Wang, J.S. Hesthaven, Machine-learning-enhanced real-time aerodynamic forces prediction based on sparse pressure sensor inputs, *AIAA J.* 62 (7) (2024) 2601–2621.
 - [29] “World leading fixed wing drone - Nordic wing.” Accessed: Jan. 08, 2026. [Online]. Available: <https://nordic-wing.com/>.
 - [30] A. Airolidi, L. Marelli, P. Bettini, G. Sala, A. Apicella, Strain field reconstruction on composite spars based on the identification of equivalent load conditions, presented at the, *Sensors and Smart Structures Technologies for Civil, Mechanical, and Aerospace Systems 2017*, SPIE, 2017, pp. 207–226.
 - [31] L. Colombo, C. Sbarufatti, W. Zielinski, K. Dragan, M. Giglio, Numerical and experimental flight verifications of a calibration matrix approach for load monitoring and temperature reconstruction and compensation, *Aerosp. Sci. Technol.* 118 (2021) 107074.
 - [32] J. B. Bell, “Solutions of Ill-posed problems,” 1978.
 - [33] H. Sun, D. Feng, Y. Liu, M.Q. Feng, Statistical regularization for identification of structural parameters and external loadings using state space models, *Comput. Civ. Infrastruct. Eng.* 30 (11) (2015) 843–858.
 - [34] B. Miao, F. Zhou, C. Jiang, X. Chen, S. Yang, A comparative study of regularization method in structure load identification, *Shock Vib* 2018 (1) (2018) 9204865.
 - [35] M. Aucejo, O. De Smet, Bayesian source identification using local priors, *Mech. Syst. Signal Process.* 66 (2016) 120–136.
 - [36] G. Abaqus, Abaqus 6.11, Dassault Syst. Simulia Corp. Provid. RI USA 3 (2011) 73.
 - [37] “Model: 352A24,” PCB Piezotronics. Accessed: Sep. 16, 2025. [Online]. Available: <https://www.pcb.com/products?m=352a24>.
 - [38] M.C. Kennedy, A. O’Hagan, Bayesian calibration of computer models, *J. R. Stat. Soc. Ser. B Stat. Methodol.* 63 (3) (2001) 425–464.
 - [39] I.N. Giannakeas, Z.S. Khodaei, M. Aliabadi, Digital clone testing platform for the assessment of SHM systems under uncertainty, *Mech. Syst. Signal Process.* 163 (2022) 108150.
 - [40] I.N. Giannakeas, F. Mazaheri, O. Bacarreza, Z.S. Khodaei, F.M. Aliabadi, Probabilistic residual strength assessment of smart composite aircraft panels using guided waves, *Reliab. Eng. Syst. Saf.* 237 (2023) 109338.
 - [41] D.N. John, M. Schick, V. Heuveline, Learning model discrepancy of an electric motor with Bayesian inference, presented at the, *Proceedings* (2019) 11.
 - [42] D. Rossat, et al., Bayesian inference with correction of model bias for thermo-hydro-mechanical models of large concrete structures, *Eng. Struct.* 278 (2023) 115433.
 - [43] O.A. Vanli, S. Jung, Statistical updating of finite element model with Lamb wave sensing data for damage detection problems, *Mech. Syst. Signal Process.* 42 (1–2) (2014) 137–151.
 - [44] T. Marwala, I. Boulkaibet, S. Adhikari, Probabilistic Finite Element Model Updating Using Bayesian Statistics: Applications to Aeronautical and Mechanical Engineering, John Wiley & Sons, 2016.
 - [45] “ATI industrial automation: F/T sensor delta,” ATI Industrial Automation. Accessed: Sep. 18, 2025. [Online]. Available: https://www.ati-ia.com/products/f/t/ft_models.aspx?id=delta.
 - [46] Hottinger Brüel & Kjær (HBK), “GEN series mainframe data acquisition systems.” [Online]. Available: https://www.hbkworld.com/en/products/instruments/daq-systems/genhs/mainframe#1ref_hbm.com.
 - [47] J. Blank, K. Deb, Pymoo: Multi-objective optimization in python, *IEEE Access* 8 (2020) 89497–89509.
 - [48] Tokyo Measuring Instruments Laboratory. Accessed: Sep. 19, 2025. [Online]. Available: https://tml.jp/e/product/strain_gauge.