Smart Sensor Placement: A Correlation-Aware Attribution Framework (CAAF) for Real-world Data Modeling

Sze Chai Leung[1,*], Di Zhou[2,3], and H. Jane Bae[2]

[1]Department of Mechanical and Civil Engineering, California Institute of Technology, Pasadena, CA 91125, USA.
[2]Graduate Aerospace Laboratories, California Institute of Technology, Pasadena, CA 91125, USA.
[3]Department of Mechanical and Aerospace Engineering, University of Tennessee, Knoxville, TN 37996, USA.
[*]sleung@caltech.edu

ABSTRACT

Optimal sensor placement (OSP) is critical for efficient, accurate monitoring, control, and inference in complex real-world systems. We propose a machine-learning-based feature attribution framework to identify OSP for the prediction of quantities of interest. Feature attribution quantifies input contributions to a model's output; however, it struggles with highly correlated input data often encountered in real-world applications. To address this, we propose a Correlation-Aware Attribution Framework (CAAF), which introduces a clustering step before performing feature attribution to reduce redundancy and enhance generalizability. We first illustrate the core principles of the proposed framework through a series of validation cases, then demonstrate its effectiveness in real-world dynamical systems, such as structural health monitoring, airfoil lift prediction, and wall-normal velocity estimation for turbulent channel flow. The results show that the CAAF outperforms alternative approaches that typically struggle due to the presence of nonlinear dynamics, chaotic behavior, and multi-scale interactions, and enables the effective application of feature attribution for identifying OSP in real-world environments.

Introduction

Sensing involves the conversion of real-world physical phenomena into digital data representations. As sensing technologies and computational capabilities advance, there is a growing dependence on sensor measurements for diverse applications. However, since sensors typically provide localized observations, practical constraints on sensing resources impose fundamental limitations on data acquisition, preventing exhaustive spatial coverage or continuous temporal monitoring. Therefore, sensors must be strategically positioned to maximize information gain under these constraints. However, determining optimal sensor configurations for nonlinear chaotic systems, such as turbulent flows, geophysical processes, and biomedical applications, remains a problem of significant scientific interest and complexity due to the multi-scale and multi-process nature of these systems. Representative examples include early fire detection[1] and structural damage monitoring[2], where the goal is to achieve high detection accuracy while minimizing the number of deployed sensors. Similarly, environmental monitoring systems require carefully designed sensor networks to ensure accurate measurements of parameters like temperature and air quality[3,4]. In autonomous systems like ground and aerial vehicles, optimal sensing enables precise environmental perception, facilitating robust decision-making and control[5]. The integration of machine learning (ML) further amplifies the importance of sensing, as model performance depends heavily on input data quality. Well-designed sensor configurations reduce measurement noise and input correlations, thereby enhancing the reliability of ML-based predictions in control applications ranging from autonomous driving to industrial automation. Thus, systematic approaches to sensor placement remain essential for balancing resource constraints against information requirements across numerous domains.

The problem of finding optimal sensor placement (OSP) has been extensively studied, leading to the development of various methods to maximize coverage, minimize redundancy, and enhance prediction quality for specific applications. For instance, several approaches utilize analytical or algebraic formulations to determine sensor locations[6–8]. The proper orthogonal decomposition (POD) with QR-pivoting has been developed for OSP and demonstrated on applications such as sea surface temperature reconstruction[8] and greenhouse flow[9]. Optimal placements can also be found by maximizing the system's observability for state estimation[10]. The effective independence (EI) method[6] is proposed to identify the OSP for structural health monitoring (SHM) by maximizing the determinant or trace of the Fisher Information Matrix and applied to the design of space stations. Viewing the sensor placement problem from the perspective of optimization, some employed Bayesian inference to optimize sensor locations for reducing parameter estimation uncertainty[11,12]. Meanwhile, information-theoretic approaches

exist where mutual information is utilized to identify the best sensing configurations[12–15]. With the recent emphasis on ML and artificial intelligence, it is worth highlighting various ML-based sensor placement schemes that leverage data-driven techniques[3,16–18]. For example, an unsupervised OSP approach relying on an attention mechanism to prune insignificant channels for SHM was recently explored by Shi et al.[18]. Moreover, by formulating the sensor placement problem as a Markov decision process, Wang et al.[16] optimized the spatiotemporal modeling performance of the sensors through reinforcement learning.

Among these approaches, an ML-based method with significant potential for OSP is feature attribution (FA), which assigns importance scores to the input features of a model (e.g., words in a sentence, pixels in an image, etc.) based on their influence on the output, revealing the model's decision-making logic. Initially conceptualized for game theory by Shapley[19], a fair FA method should quantify the contribution of individual participants in a game based on four fundamental axioms: symmetry, efficiency, dummy, and additivity. The Shapley value, derived from these principles, was among the first to offer a systematic approach to estimating FA[20]. The GradientSHAP[21] and KernelSHAP[22,23] algorithms were proposed with an aim to estimate the Shapley value using different approaches.

More recently, FA has expanded to become a prevalent technique in ML research for quantifying the contributions of input features to a model's output[24–27]. Various algorithms have been developed with a focus on ML applications, including Saliency Map[28], DeepLift[29], and Integrated Gradients (IG)[24]. Regardless of the chosen algorithm, FA relies on an accurate surrogate model to capture the relationship between feature inputs and the inference target. By applying a FA algorithm to a model, scores are assigned to the input features based on their significance to the target predictions. It can be used not only to eliminate redundant input features, enhancing training efficiency and accuracy, but also to explain decisions made by large models. It has notable applications in visualizing attention regions in computer vision tasks[27] and revealing critical words or phrases in natural language processing models[25]. FA enables researchers to analyze and compare the behaviors of ML models with human reasoning patterns. While many existing OSP methods are application-specific, model-agnostic FA can be universally applied to tasks across different domains. It requires no coupling with or modification of the model architecture, allowing direct application to existing prediction models. Therefore, extending the application of FA to optimal sensor identification offers a promising research direction.

However, similar to many existing data-driven OSP methods, FA remains vulnerable to the challenges posed by highly correlated data commonly encountered in real-world scenarios. ML models often distribute weights across all contributing features, disregarding the inter-dependencies among input features. Consequently, FA algorithms applied to such models may assign high importance scores to correlated candidates, resulting in redundant sensor selections that convey overlapping information. This redundancy poses a significant limitation in OSP tasks, where the goal is to maximize information efficiency. For instance, when dealing with spatially correlated inputs, the selected sensors may cluster closely together, limiting the diversity of information captured. While such behavior is often acceptable or even desirable in the context of model interpretability, where highlighting groups of related features aids qualitative understanding, identifying optimal sensor locations necessitates scrutiny of individual elements rather than general regions. The ability to pinpoint the most informative sites is crucial.

To address this issue, we propose the Correlation-Aware Attribution Framework (CAAF) for identifying OSP. CAAF introduces a clustering step prior to model training and FA. This clustering process groups correlated candidate sensor locations together and selects a representative point from each cluster, thereby capturing the essential information within the group. As a result, highly correlated candidates are effectively filtered out before model construction to avoid redundancy. Following this step, only these cluster centers are retained as the final set of candidates for FA. Candidate clustering is performed using a correlation metric chosen based on the characteristics of the application. This preprocessing step enhances the robustness of the framework, ensuring that the selected sensors provide diverse and non-redundant information.

In summary, the CAAF identifies optimal sensor locations through five major steps.

1. Cluster the initial set of candidate sensors using a predefined correlation metric

2. Identify the cluster centers that serve as optimal representatives for their respective clusters

3. Develop a data-driven model to predict the target variable using inputs from the identified cluster-center sensors

4. Apply a FA algorithm to the model to evaluate the contributions of each cluster center's inputs, and rank the cluster centers based on their significance to the model's predictions

5. Select the desired number of cluster centers as the optimal sensor configuration based on the rankings

Overall, the CAAF is a novel ML-based approach to sensor selection that operates solely on candidate sensor measurements and ground truth data of the sensing objective. Its fully data-driven nature and remarkable adaptability offer a powerful new paradigm for tackling complex sensor placement problems across engineering applications.

Results

In this section, we present illustrative experiments and examples to intuitively demonstrate the working principles of CAAF. We then explore its applicability across various real-world scenarios and compare its performance with existing methods to highlight its advantages and limitations.

Demonstration of working principle

To establish the theoretical foundation of the CAAF methodology for OSP, we first present empirical evidence spanning from image classification tasks to synthetic validation problems. These results demonstrated the benefits of the added clustering step while validating CAAF's effectiveness on idealized test cases.

Correlation aware clustering: image clustering

A diverse collection of FA methods has been extensively tested and applied to explain deep ML models, with one of the most prominent applications being the interpretation of convolutional neural networks for image classification[25,27]. By visualizing the attention regions of the model that overlap with objects of interest, FA methods have demonstrated significant success in interpreting their decision-making processes.

However, this behavior may not be ideal for sensor placement. Figure 1 presents four selected samples from the ImageNet image classification dataset[30]. If FA is directly applied to an image classification model to determine optimal color-sensor locations, the resulting sensors tend to cluster within the object of interest, as these are the pixels the model primarily uses for classification (Figure 1(e–h)). Such a distribution is rarely optimal, as it lacks information diversity and may be particularly problematic for tasks that rely on global contextual information from the surrounding environment.

After implementing the clustering step prior to FA methods in image classification, the process systematically identifies all potential regions of interest along with their representative pixels as demonstrated in Figure 1(i–l). The results show distinct clusters corresponding to different objects (e.g., banana and starfish), with their respective cluster centers effectively representing multiple key regions of interest. The clustering correlation metric is defined as the Euclidean distance between pixels, integrating information from both the RGB color channels and the spatial positions of the pixels. This qualitative analysis illustrates how clustering can be used to efficiently pre-select salient image features and pinpoint the representative pixels within them. Importantly, these cluster centers form a reduced candidate set that encompasses information beyond the main object, thereby improving computational efficiency and sensor selection performance in subsequent FA analysis within CAAF.

Attribution based on correlation: correlated synthetic data

Correlated features cause a breakdown of the additivity axiom in Shapley attribution theory, which states that the total contribution of two features must equal the sum of their individual contributions, since part of their contribution may be shared. They complicate FA analysis as learned models may assign high importance scores to correlated inputs. This can occur because models, such as neural networks, can distribute weights arbitrarily among correlated features without losing predictive performance. During training, backpropagation optimizes the combined effect of features rather than individual contributions. If two features are perfectly correlated, the network may assign similar weights to both, as any linear combination of the two features yields the same output. This creates ambiguity in FA — while the model relies on the collective signal, FA methods may misleadingly assign importance to both features equally, even though only one is theoretically necessary. Although not visually shown here, we empirically demonstrate this phenomenon by training a neural network on two identical features, which indeed results in equal FA values for both features, revealing a key limitation in directly applying FA to interpret feature importance for correlated inputs.

We now further demonstrate how CAAF effectively addresses this challenge through clustering via an idealized sensor selection problem using synthetic data, with experimental details and results presented in Table 1. The task involves optimally selecting two sensors from three candidates, where Sensors 1 and 2 exhibit high mutual correlation, while Sensor 3 is statistically independent. Row "Corr(X_i, Y)" in the table lists the correlation coefficients between each sensor input and the target variable, providing a reference for their expected attribution-based contributions. If we naively apply FA without accounting for correlations, the method selects Sensors 1 and 2 due to their high individual FA scores in row "Naive FA". However, these sensors are highly correlated (with a correlation value of 0.9), rendering one redundant. In contrast, the CAAF first clusters the correlated candidates together, identifying Sensor 2 as the cluster center for the correlated group (Sensors 1 and 2), as shown in row "Cluster Label". By then performing FA to the cluster centers (Sensors 2 and 3), the framework correctly selects Sensors 2 and 3 as the optimal pair, aligning with intuitive expectations and avoiding redundancy.

Interestingly, when the input-output correlation coefficient "Corr(X_i, Y)" of a cluster center is increased, both the naive and clustered FA values associated with that sensor scale almost linearly with the correlation coefficient, as illustrated in Figure 2. However, after the removal of the correlated sensor, Sensor 1, the CAAF score or "Clustered FA" increases at a steeper rate, indicating a shift in relative importance among the remaining sensors. Namely, the difference between "Naive
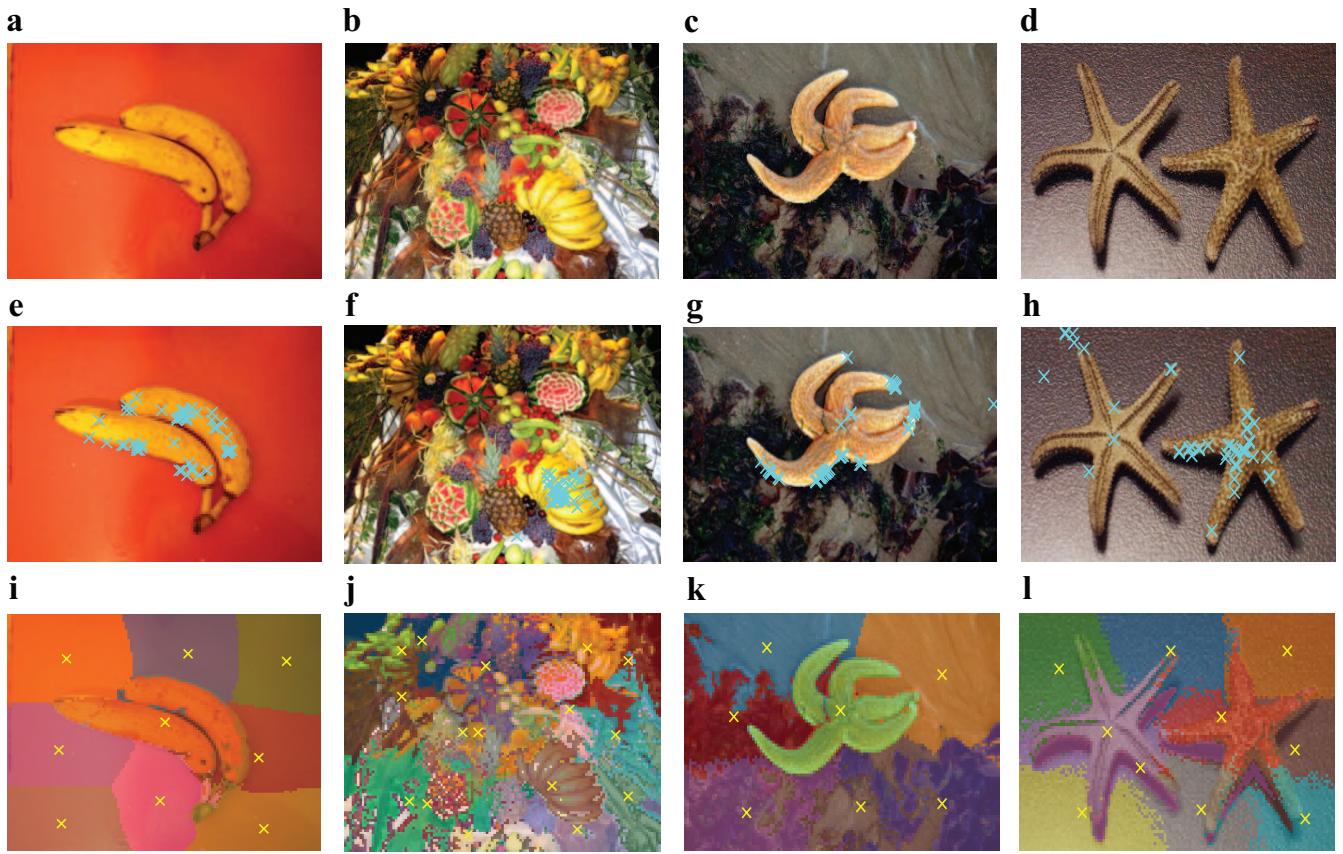
Figure 1. Clustering and naive FA applied to image classification. Clustering and naive FA outcome for 2 banana images (**a,e,i,b,f,j**) and 2 starfish images (**c,g,k,d,h,l**). The top row (**a–d**) shows the original ImageNet samples. The middle row (**e–h**) displays the naive FA results, where cyan crosses × indicate the 50 pixels with the highest FA scores. The bottom row (**i–l**) presents the clustering results, with each cluster region shaded in a distinct color and yellow crosses × marking the cluster centers.

FA" and "Clustered FA" increases as "Corr(X_i, Y)" increases, demonstrating that grouping correlated candidates and evaluating only the cluster centers amplifies the distinction between influential and redundant sensors. This effect arises because when naive FA is applied, duplicated information among correlated features dilutes the apparent importance of the uncorrelated ones. Consequently, incorporating clustering widens the FA gap between critical and non-critical candidates, enhancing the discriminative power of the selection process.

This behavior underscores a key advantage of CAAF. By condensing correlated sensors into representative cluster centers, the framework magnifies the relative importance of truly informative features, making optimal sensor selection more robust and reliable.

Real-world applications

Having illustrated the working principles of CAAF, we now evaluate its performance on real-world applications. This section demonstrates CAAF's applicability across three practical, control-oriented scenarios: identifying the most informative elements on a cantilever beam for SHM, predicting airfoil lift from surface pressure measurements, and inferring off-wall wall-normal velocity using wall pressure data in wall-bounded turbulence. These problems are chosen because efficiently determining the OSP is highly desirable for predicting the quantities of interest critical to downstream control applications, yet inherently nontrivial due to the complexity of these systems. These case studies validate the framework's effectiveness and showcase its utility in real-world settings.

Structural health monitoring
SHM ensures the safety and longevity of structures (bridges, buildings, etc.) by detecting damage and optimizing maintenance through sensor signals. Determining the best locations for a restricted number of sensors within a structure is a key step in acquiring precise, real-time measurements. Notably, in structural analysis, sensor positions must be carefully selected so

Table 1. Statistics for the synthetic data experiment.

Sensor Index	1	2	3
Corr(X_i, X_1)	1.0	0.9	0.0
Corr(X_i, Y)	0.65	0.89	0.32
Naive FA	29%	59%	12%
Cluster Label	0	**0**	**1**
Clustered FA		76%	24%

"Corr(X_i, X_1)" shows the correlation of each sensor's inputs with those of Sensor 1. "Corr(X_i, Y)" shows the level of correlation between the sensor inputs and the target variable. "Naive FA" indicates the FA percentages (or importance scores) of the individual candidate sensor inputs by directly performing FA on a model trained on all candidate features without precedent clustering. "Cluster Label" reveals the clusters formed using the Affinity Propagation clustering algorithm. The boldfaced labels indicate the cluster centers. "Clustered FA" prints the CAAF results, which are the FA percentages of the cluster centers.
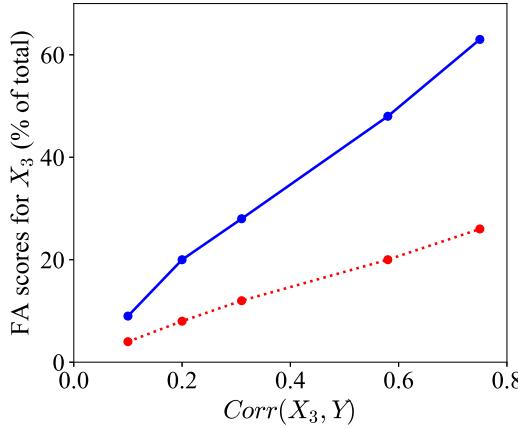

Figure 2. Variation of X_3's contribution fraction with correlation. As the correlation between X_3 and the target variable Y increases, the fractions of the total FA scores contributed by X_3 from the naive FA (- • -) and clustered FA (—•—) approaches are shown.

that the authentic mode shapes and target frequencies can be accurately recovered[31]. Additionally, SHM strategies rely on OSP techniques for various applications, including finite-element updating, state estimation, and damage localization[7,32]. Nonetheless, identifying OSP for SHM presents significant difficulties because it requires maximizing information gain about all potential states across the entire structure, while minimizing costs associated with the sensor equipment and computation.

OSP methods for SHM have been extensively studied, with common approaches including EI[6] and its variants (EI-mass[33], EI-dpr[34]), as well as energy-based methods like kinetic energy (KE) and weighted average KE[6,7,35]. In this study, we validate the data-driven CAAF method against these established analytical OSP approaches using a cantilever beam benchmark case and the standard SHM assessment metrics.

We apply CAAF to select 15 optimal sensor locations from 30 candidate nodes on a uniform cantilever beam, considering only the first three mode shapes and translational degrees of freedom. The CAAF uses the Affinity Propagation (AP) algorithm to perform clustering and identify cluster centers. As shown in Figure 3(a), AP groups the candidate sensor nodes into 19 clusters and captures the correlations among nodes located near the fixed joint. Next, the CAAF employs a multilayer perceptron (MLP) model trained to predict the modal coefficients of the first three modes from nodal deflection measurements at the candidate sensor locations. Figure 3(a) also displays the optimal sensor configurations from the CAAF alongside EI and KE methods. The spatial distributions reveal characteristic selection patterns unique to each methodology, with CAAF demonstrating a more balanced coverage of the structural domain.

The resulting sensor configurations from CAAF, EI, and KE methods are compared using three key performance metrics: root mean square error (RMS), condition number (CN), and determinant (DET). The RMS metric quantifies the orthogonality of the mode shape matrix, Φ, truncated to the selected sensor locations, while the CN and DET assess modal independence. The *Methods* section provides complete implementation details for these metrics. Optimal sensor configurations should minimize RMS and CN while maximizing DET, thereby ensuring maximal mode shape orthogonality and independence. High
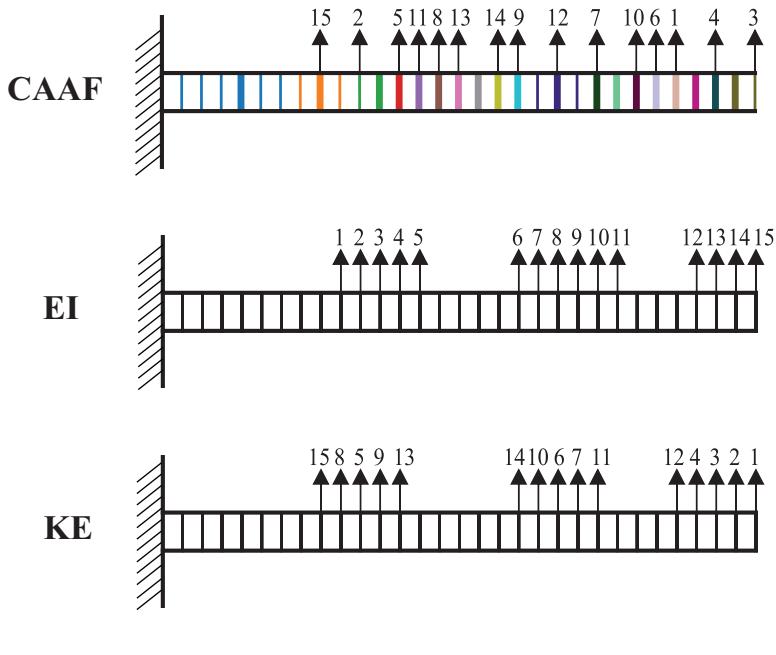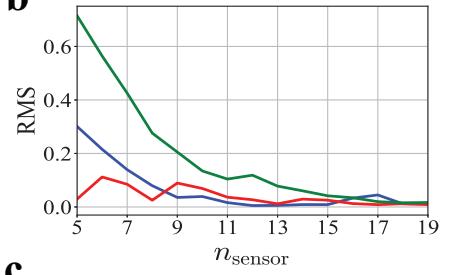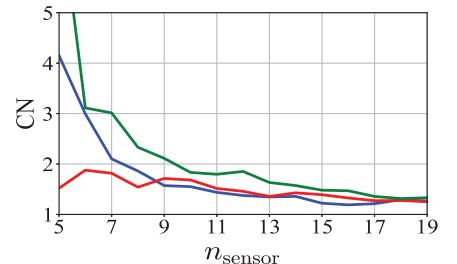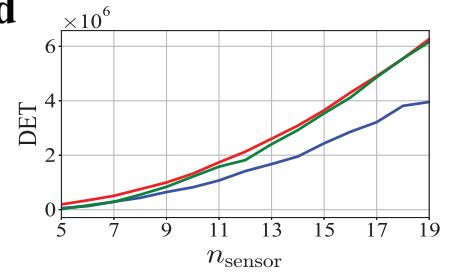
Figure 3. Locations and performance of SHM sensors. (**a**) 15 SHM sensors on the cantilever beam identified by the CAAF, EI, and KE methods, with 19 node clusters (colored bars) and their centers (thick colored bars). Performance variation measured using (**b**) RMS, (**c**) CN , and (**d**) DET versus the number of sensors (n_{sensor}) for configurations determined by CAAF (——), EI (——), and KE (——).

RMS/CN values or low DET values indicate suboptimal sensor placement.

As illustrated in Figure 3(b–d), CAAF achieves superior performance in both RMS and CN compared to EI and KE methods for high-density sensor configurations, despite the lower DET values. This performance pattern confirms that CAAF-based sensor placements yield better mode shape orthogonality and independence while maintaining competitive information content. These findings establish CAAF as a high-performing, data-driven OSP method applicable to traditionally analytical problems when implemented with proper model training and cluster selection.

Airfoil lift prediction

In active flow control applications, particularly gust mitigation for small aircraft and unmanned aerial vehicles, controller performance largely depends on accurate flow sensing. Turbulent winds and complex terrain generate unpredictable aerodynamic disturbances, exacerbated by the small size and low inertia of these vehicles[36–38]. Effective sensing and control strategies are essential to maintain stability, but they face significant challenges due to the stringent requirements for accurate perception and high bandwidth, low-latency response in highly unsteady real-world turbulence. OSP becomes critical in this context, as it directly affects the observability of key flow features and robustness of the control performance. Here, we focus on identifying OSP on an airfoil for surface pressure sensors to enable precise temporal predictions of the airfoil lift coefficients (C_L), which is crucial for mitigating gust-induced lift fluctuations[39–42].

Using large-eddy simulations (LES), we simulated flow over a NACA 0012 airfoil under both gusty and non-gusty conditions, collecting time-resolved airfoil surface pressure distributions and lift coefficients. Pressure data from the LES are averaged over the spanwise direction, as only the two-dimensional placement is investigated. The CAAF is then applied to determine optimal sensor locations, which involves clustering the candidates and training an MLP model with the LES data to estimate the instantaneous airfoil lift coefficient using pressure measurements at the candidate sensor locations. The predictive accuracy of the CAAF sensor configuration is benchmarked against those derived from naive FA, QR pivoting of POD modes,
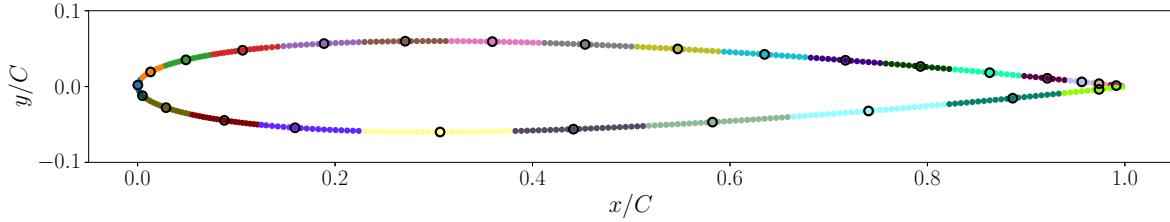
Figure 4. Clustering of airfoil surface pressure sensor candidates. Clustering results showing 27 identified clusters (colored dots) and their centers (◦).

Bayesian experimental design, and uniform sensor distributions methods.

Following our previous work[43], gusty inflow conditions are generated by placing a cylinder directly upstream of the airfoil. The Reynolds number Re_c based on chord length C and free-stream velocity U_∞ is 10^4. We simulate five distinct flow configurations, systematically varying cylinder geometry and angle of attack (AoA): (1) "None-5" (baseline, AoA = 5°), (2) "None-11" (baseline, AoA = 11°), (3) "Cylinder-5" (with cylinder, AoA = 5°), (4) "Cylinder-11" (with cylinder, AoA = 11°), and (5) "Cylinder2-0" (modified cylinder geometry, AoA = 0°). The complete geometric configuration is detailed in the *Methods* section. For our analysis, we use the "None-5", "None-11", and "Cylinder-5" cases for model training and sensor identification, reserving the "Cylinder-11" and "Cylinder2-0" cases exclusively for testing purposes.

We leverage the AP algorithm again to implement clustering for CAAF, with results visualized in Figure 4. It is observed that spatial correlation among pressure measurements leads to clustering based on proximity, with higher cluster density near the leading and trailing edges due to greater pressure variability.

The optimal 10-sensor configurations identified by the different OSP methods are displayed in Figure 5. Notably, the configurations derived from data-driven approaches exhibit distinct spatial distributions, reflecting their different underlying selection criteria. Nonetheless, some common features of the CAAF configurations can be observed in the others and explained by analyzing the physics of the flow around an airfoil. In addition to the large separation bubble that typically forms on the suction surface near the trailing edges of the airfoil, where most fluctuations occur due to re-circulation, informative regions also include the stagnation point just below the leading edge corresponding to the location of the first sensor on the CAAF configuration in Figure 5(a). The pressure around the stagnation point varies as the effective AoA deviates from the true AoA of the airfoil under the cylinder wake. Since the effective AoA is highly correlated with lift fluctuations, such an area is expected to play a key role in lift inference. Unique to the cases with gusty inflows, separation bubbles appear around the leading edge of the airfoil[43]. The flow characteristics within the separation region are significantly different, and the separation bubble size greatly impacts the airfoil lift coefficient. Considering this, the CAAF design recognizes it by prioritizing leading-edge sensor placement to capture these flow dynamics.

We evaluated the predictive performance of each optimal sensor configuration by training dedicated MLP models using only pressure measurements at the selected sensor locations. The prediction accuracy of the trained models acts as a direct metric for evaluating the optimality of each sensor configuration. The relative L_2 prediction error ε serves as our performance metric, defined as

$$\varepsilon = \frac{\left\|\widehat{C_L} - C_L\right\|_2}{\left\|C_L - \langle C_L\rangle\right\|_2} , \tag{1}$$

where $\widehat{\cdot}$ represents the predicted lift coefficient. Figure 6 shows the variation of ε with increasing number of sensors. The results demonstrate that CAAF-generated sensors consistently outperform other methods across most sensor counts, while matching their performance in remaining cases. This superior performance confirms CAAF's effectiveness for OSP in airfoil lift prediction tasks and suggests its broader applicability to similar target inference problems.

Turbulent wall-normal velocity estimation

Precise knowledge of the off-wall wall-normal velocity (v) in wall-bounded turbulence is highly valuable, as it enables active control of boundary layer flow structures via fluidic actuation to achieve drag reduction, potentially yielding substantial energy savings in various engineering applications[44,45]. Since making accurate non-intrusive measurements of the off-wall velocities is impractical, previous efforts[45–48] have focused on predicting the off-wall velocities using wall quantities such as the streamwise and spanwise velocity gradients at the wall, wall-shear stress, and wall pressure (p_w). However, prediction remains challenging due to the inherently nonlinear and chaotic nature of turbulent flows, particularly when relying on sparse measurements as required in real-world deployments, where most existing methods assume access to full-field inputs. There-
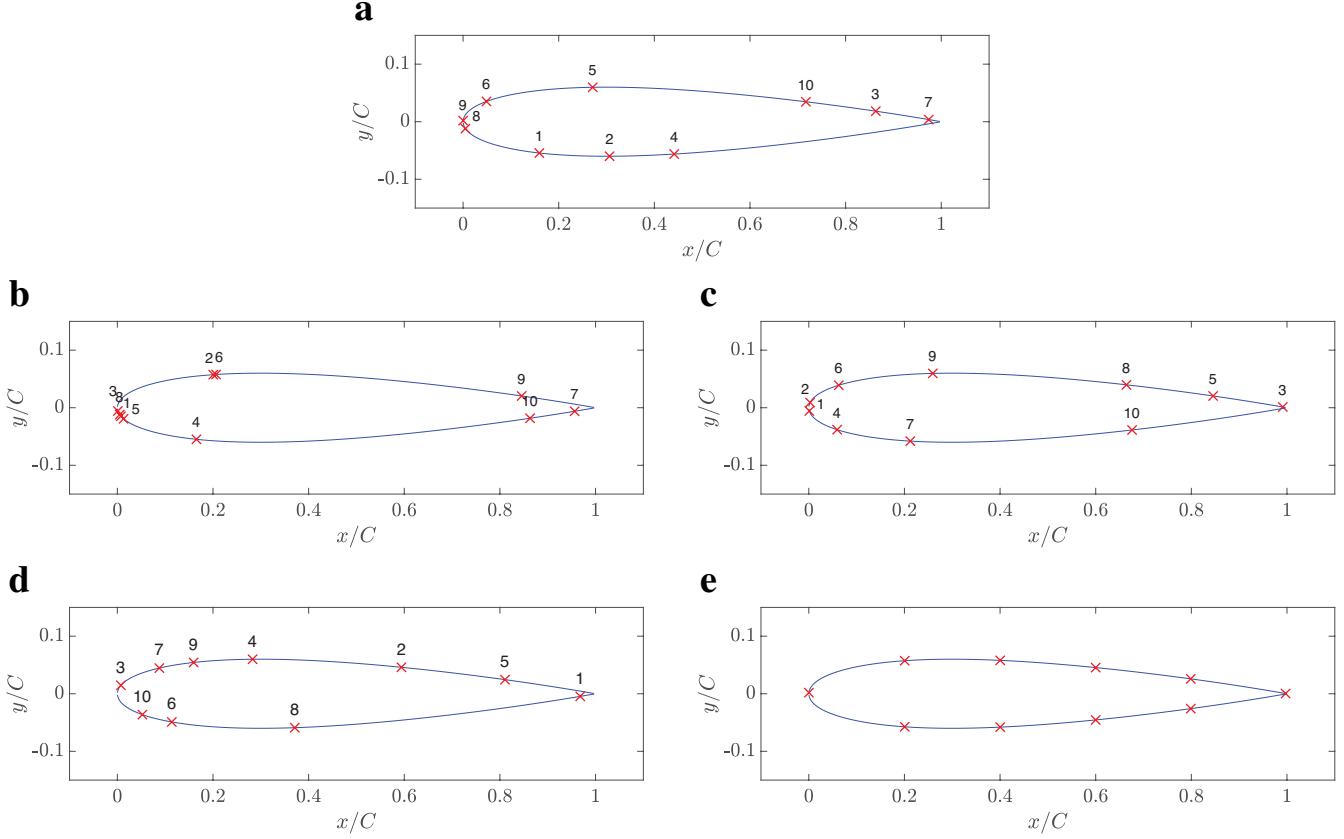
Figure 5. Optimal airfoil surface pressure sensor locations. The optimal 10-sensor configurations identified by (**a**) CAAF, (**b**) naive FA, (**c**) POD-based QR pivoting, (**d**) Bayesian experimental design, and (**e**) uniform distribution. Red crosses ✕ mark the locations of the sensors. Numerical labels indicate the sensors' importance ranking, reflecting their selection order, except for the POD-based QR pivoting sensors, which are not identified sequentially. For the uniform distribution case (**e**), sensors are not numbered as they are selected simultaneously in a single step.

fore, the objective of this application is to identify optimal wall sensing locations where wall pressure measurements provide the most informative input for estimating off-wall wall-normal velocity in turbulent channel flows.

Following previous studies[45,49], candidate locations of wall pressure probes for wall-normal velocity estimation are sampled from streamwise and spanwise positions within $-50 < \Delta x^+, \Delta z^+ < 50$, where the superscript $+$ denotes that the quantity is nondimensionalized by viscous units defined based on the kinematic viscosity and friction velocity. We employed exclusively the wall pressure as input, as it shows a high correlation with off-wall wall-normal velocity in the previous studies. Instantaneous wall-normal velocity at the wall-normal location $y^+ = 10$ and center of the sensing plane ($\Delta x^+, \Delta z^+ = 0$), denoted by v_{10}, is predicted by an MLP model using the wall pressure measurements (at $y^+ = 0$) as inputs at each time snapshot. The location $y^+ = 10$ is selected based on prior studies[44] indicating it as the optimal height for implementing effective opposition control for drag reduction. The ground truth data is obtained through a direct numerical simulation (DNS) of an incompressible turbulent channel flow at friction Reynolds number $Re_\tau = 186$. Figure 7 depicts the instantaneous fields of the wall pressure and wall-normal velocity fluctuation at $y^+ = 10$ obtained from the DNS.

As shown in Figure 8(a), the clustering analysis condenses the initial 361 probe locations to 58 representative candidates through AP. The clusters exhibit localized circular shapes consistent with the wall pressure spatial autocorrelation patterns reported in the literature[50,51]. By recognizing these features inherent to the problem, the clustering algorithm naturally groups points according to the underlying physical patterns of the flow, providing meaningful data representatives.

With the established cluster representatives and the trained MLP model for predicting v_{10}, CAAF is successfully applied. The optimal 10-sensor configuration is shown in Figure 8(b). To illustrate the physical basis for this sensor selection pattern, we present contours of the two-point cross-correlation coefficient between p_w and v_{10}, defined as

$$C_{v_{10}p_w}(\Delta x, \Delta z) = \frac{\langle v'_{10}(x,z,t) \, p'_w(x+\Delta x, z+\Delta z, t) \rangle}{\sigma_{v'_{10}}(x,z)\sigma_{p'_w}(x+\Delta x, z+\Delta z)}, \tag{2}$$
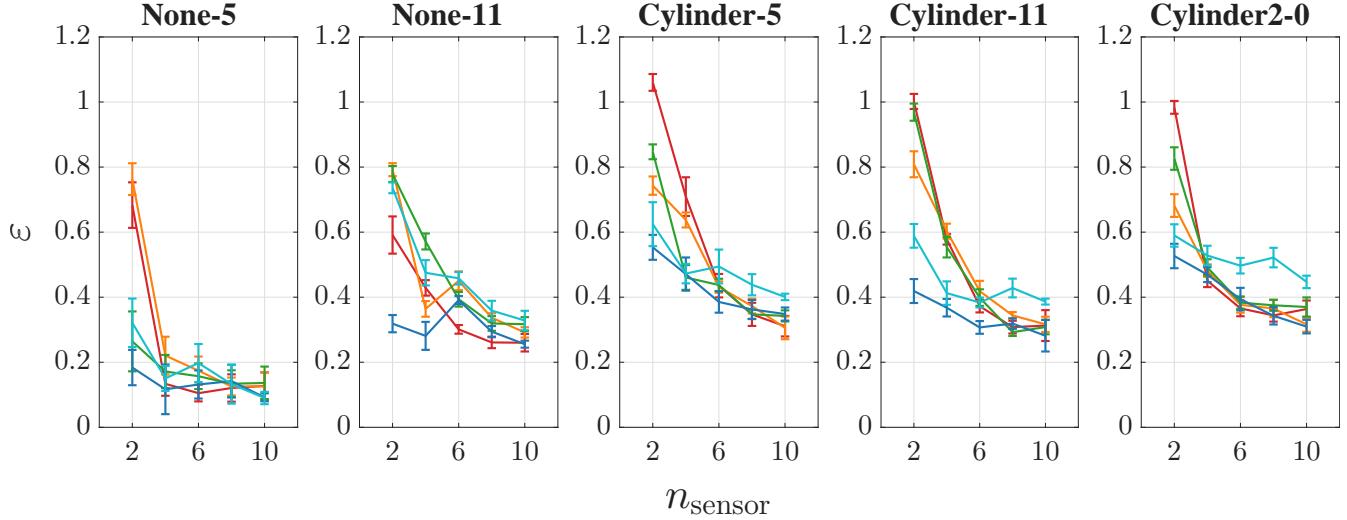
Figure 6. Prediction performance comparison between sensor configurations identified using different OSP methods. Lift prediction error for different number of desired sensors (n_{sensor}) using sensors identified by CAAF (——), naive FA (——), POD with QR pivoting (——), Bayesian experimental design (——), and uniform distribution (——). The error bars represent the standard deviation computed over a minimum of ten independent training runs.
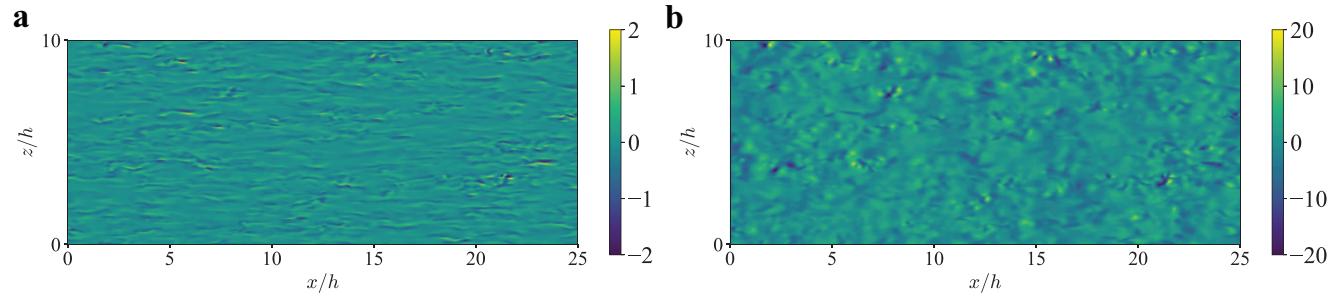

Figure 7. Instantaneous data visualization from DNS of channel flow. a Wall-normal velocity fluctuation v'^+ at $y^+ = 10$. **b** Wall pressure fluctuation $p_w'^+$.

where $(\cdot)'$ denotes fluctuations, $\langle\cdot\rangle$ is the time-averaging operator, and σ represents the standard deviation. As shown in Figure 8(b), the v_{10}-p_w cross-correlation reveals a distinct high-correlation region downstream of the prediction origin, which is largely symmetric. A secondary, slightly weaker high-correlation region appears upstream, centered around $\Delta x^+ \approx -25$ to -50. Similar spatial correlation structures have been reported in channel flows at $Re_\tau = 178$ and 578 by Park and Choi[45], who also visualized a saliency map from a convolutional neural network trained to predict v_{10} from p_w. Their results highlighted the same key regions of the pressure field contributing to the prediction, reinforcing the validity of the observed patterns. The CAAF-identified sensors exhibited physically interpretable placement, aligning with these flow features and supporting their relevance for wall-normal velocity prediction, as indicated by the cross-correlation statistics and prior literature.

As a benchmark configuration, we consider an array of 10 sensors uniformly positioned within the sensing plane, highlighted in Figure 8(b). Table 2 compares the wall-normal velocity time series predictions obtained using the CAAF-identified and the benchmark sensor configurations. The CAAF sensors achieve a L_2 prediction error of 0.683 and a correlation of 0.734 with the ground truth, outperforming the benchmark array which yields an error of 0.806 and a correlation of 0.600. Figure 9 contrasts the predicted wall-normal velocity sequences for a selected section of the testing dataset. Although tracking the reference signal with only 10 sensors is sub-optimal, the CAAF-identified sensors demonstrate more consistent tracking of the wall-normal velocity than the benchmark configuration, particularly during periods of large fluctuations.

These results demonstrate that CAAF effectively identifies optimal sensor locations for wall-normal velocity estimation. Notably, the CAAF-derived sensor configuration outperforms the uniform benchmark and achieves performance comparable to the full-field learning approach reported in the literature[45].
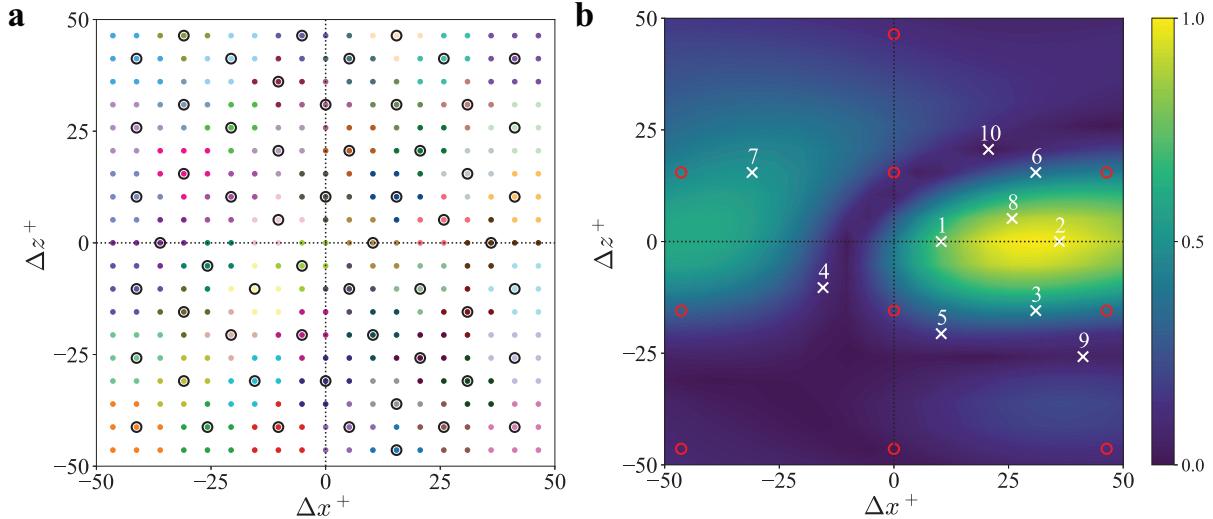
Figure 8. Clustering and optimal sensor configuration for wall-normal velocity estimation. a Initial candidate sensor locations (361 probes) and AP clustering results (reduced to 58 representative locations). Cluster members are shown as colored dots with cluster centers indicated by black open circles (○). **b** Normalized cross-correlation coefficient between p_w and v_{10} time series. The ten most optimal sensor locations identified by CAAF (white crosses) and the uniform benchmark 10-sensor configuration (red open circles ○). The numerical labels indicate the importance rankings of the CAAF-derived sensors.

Table 2. Wall-normal velocity prediction performance comparison between different sensor configurations.

| Sensor configuration	58 cluster centers	10 CAAF	10 uniform
$\mathrm{Corr}(v_{10}, \widehat{v_{10}})$	0.882	0.734	0.600
L_2 error	0.473	0.683	0.806

"$\mathrm{Corr}(v_{10}, \widehat{v_{10}})$" reports the correlation coefficient between the true and predicted wall-normal velocity, while "L_2 error" denotes the normalized L_2 error, computed as defined in Equation 1, but applied to the variable v_{10}.

Discussion

The validation cases presented in *Demonstration of working principle* illustrate that the additional clustering step not only reduces the problem's dimensionality to improve computational efficiency, but also enhances the effectiveness of FA for sensor placement in correlated datasets. Direct FA on the original candidate set tends to identify the global attribution maximum and its nearest neighbors, which often exhibit high correlation between sensor inputs and thus lead to redundant sensor selections. In contrast, by computing FA on cluster centers, the CAAF method prioritizes candidates with locally maximal FA values, collectively enabling a more comprehensive measurement strategy.

It is also worth noting that CAAF assumes no inherent restrictions on the positioning of the sensors. In addition, practical limitations, such as constraints imposed by geometry or the infeasibility of placing sensors at certain locations, can be readily incorporated into the sensor identification process. This can be achieved by excluding candidate sensor locations that are deemed unfeasible before applying the framework; for example, sensors at the trailing edge of the airfoil are often omitted due to sharpness and can thus be removed prior to the clustering step. The framework demonstrates broad applicability and strong reliability for real-world sensor placement problems. We showcase its effectiveness in predicting different quantities of interest, ranging from modal coefficients in solid structures to aerodynamic coefficients and velocity components in wall-bounded turbulent flows. Sensor configurations selected by CAAF yield performance comparable to and often superior to those obtained using analytical or empirical methods. As a purely data-driven approach, the framework is highly adaptable across diverse domains, owing to its decoupled clustering, model development, and FA stages.

However, this flexibility also introduces potential limitations. The quality of the resulting sensor configurations depends critically on the clustering and modeling processes, which may require significant tuning and domain knowledge depending on the application. Optimal performance necessitates clustering strategies informed by domain-specific correlation metrics and physical insights. For instance, in the aforementioned airfoil lift and wall-normal velocity applications, the clustering
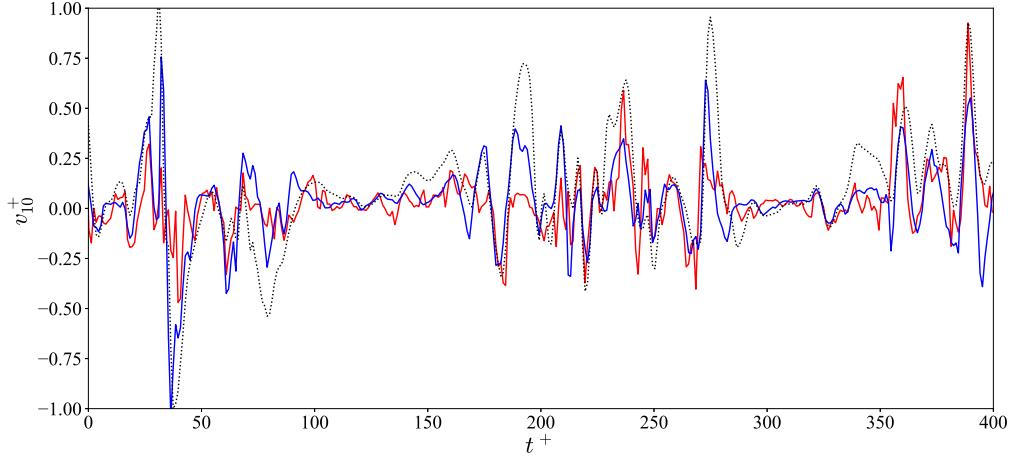
Figure 9. Wall-normal velocity (v_{10}^{+}) time-series predictions. CAAF-identified sensors (——) versus uniform benchmark sensor array (——) plotted against reference data (·····).

results agree with physical intuition, demonstrating their interpretability through domain-specific logic. Therefore, the choice of clustering and FA algorithms must also be carefully considered to ensure robust and interpretable results. Furthermore, the framework's efficacy relies on the accuracy of the surrogate model in capturing input-output relationships, which is influenced by data quality, model architecture, and hyperparameter selection. While the architecture and training process used in this study may not be fully optimized, the models still achieve sufficient accuracy to generate sensor configurations with satisfactory performance. Further optimization of the modeling approach could yield even higher-performing sensor configurations.

Despite its limitations, CAAF effectively determines sensing strategies for prediction tasks essential to the control of diverse systems. The data-driven nature of CAAF enables its application to a wide range of scenarios without requiring major modifications. By simply adjusting the set of candidate sensor locations, we can identify OSP in three-dimensional configurations and other more complex geometries beyond those examined in this study.

Methods

Although multiple clustering and FA algorithms are available, we employed the AP clustering and the IG attribution algorithm for their demonstrated efficiency and robustness in the investigated applications.

Affinity Propagation

AP is a clustering algorithm that identifies exemplars (representative data points) through iterative message passing between pairs. It utilizes two types of messages: responsibilities $r(i,k)$, which indicate how suitable point k is as an exemplar for point i compared to other candidates, and availabilities $a(i,k)$, reflecting how appropriate it is for i to choose k as its exemplar based on support from other points. The algorithm initializes availabilities to zero and iteratively updates responsibilities and availabilities until convergence, as described carefully in the original paper[52]. Exemplars emerge as points where $a(i,i)+r(i,i)$ is maximized. The algorithm requires no preset cluster count and handles non-metric or asymmetric similarities, efficiently converging to high-quality clusters. Since AP inherently identifies cluster centers through exemplars, it is particularly well-suited for the CAAF, which relies explicitly on the use of cluster centers. In contrast, when alternative clustering algorithms are employed, cluster centers can be determined through other methods, such as selecting the candidate with the highest average correlation to all other members within each cluster.

Integrated Gradients

For our task of finding optimal sensor locations, we chose to use the IG algorithm due to its two desirable properties: sensitivity and implementation invariance[24]. Sensitivity ensures that a feature's impact on the model's output is proportionally conveyed by the FA value; the implementation invariance provides robustness to variations in the model's architecture and implementation. The IG algorithm aggregates the model's gradients with respect to the input features along a linear path connecting a predefined baseline (often chosen to be the zero vector) with the input data, *i.e.*,

$$\text{IG}(x_i) := \int_0^1 \frac{\partial F(\alpha x)}{\partial \alpha x_i} x_i d\alpha \, , \tag{3}$$

where $\alpha \in [0,1]$ is a scaling parameter that interpolates between the baseline and the input, x represents the data point containing the inputs for all sensors $i = 1, \ldots, M$, where M is the number of candidate sensors, and $F(\cdot)$ is a data-driven model such that $y \approx F(x)$, with y denoting the target variable. In our implementation, we adopt the zero vector as the baseline. The magnitude of the FA scores, averaged over all data points, is computed for each candidate sensor and subsequently ranked to identify the most optimal sensors. While other FA algorithms, such as GradientSHAP and KernelSHAP, were also tested, they did not provide significant advantages over the IG method in terms of the performance of the resultant sensor configuration.

Demonstration of working principle

Correlation aware clustering: image clustering

The images used in this section are sourced from the ImageNet dataset[30] and are downsampled to half of their original resolution for the naive FA demonstration and to one quarter for the clustering demonstration.

Clustering is performed using the AP algorithm, with parameters set to a damping factor of 0.90, a maximum of 10,000 iterations, 5 convergence iterations, and a preference of -40. The affinity metric for the clustering algorithm is computed as the Euclidean distance between each pixel, incorporating both the RGB color channels R, G, B, and the x, y spatial coordinates, all scaled to the range [0, 1], given by

$$\text{Affinity} = (\Delta R^2 + \Delta G^2 + \Delta B^2 + \Delta x^2 + \Delta y^2)^{\frac{1}{2}}, \tag{4}$$

where Δ denotes the difference between two pixels.

The naive FA is conducted by applying the IG algorithm to a pre-trained EfficientNet model[53] for image classification.

Attribution based on correlation: correlated synthetic data

The dataset generation follows a structured procedure to create variables with the desired correlations. First, inputs for Sensor 1 are randomly generated as an independent reference dataset. Next, inputs for Sensor 2 and 3 are generated by combining the Sensor 1 inputs with a weighted random component to achieve specified correlation coefficients with the Sensor 1 data. The input datasets are standardized to have a zero mean and unit variance. A dataset for the target variable Y is then generated while ensuring specified correlations with the input datasets. The target dataset is constructed as a weighted linear combination of the input datasets, with weights determined by the desired correlation values. If necessary, an additional random component, orthogonal to all input datasets, is introduced to preserve unit variance while ensuring the specified correlations hold.

Clusters are produced using the AP algorithm with the affinity metric set to Pearson's correlation coefficient. The algorithm parameters are configured with a damping factor of 0.5, a maximum of 10,000 iterations, 10 convergence iterations, and a preference value of 0.7.

The IG algorithm for FA is performed on the MLP model trained with the datasets for the sensors and the target variable. The MLP architecture consists of a fully connected feedforward network with an input size equal to the number of sensors. It includes three hidden layers, each with 8 neurons, followed by batch normalization and LeakyReLU activation. The model was trained using the Adam optimizer with adaptive learning rate and L_2 regularization of 0.001 to prevent overfitting. The mean squared error loss function was used to measure the model's performance.

Real-world applications

FA values for candidate features are obtained by averaging the results across a minimum of 10,000 input data points, sampled from the training dataset.

Structural health monitoring

Inspired by Friswell and Castro-Triguero[32] and Liu et al.[7], the beam consists of 30 candidate nodes with only a translational degree of freedom, meaning the sensor inputs are the deflections at each node. It has a length L of 0.45 m, a width of 20 mm, and a thickness of 2 mm. The beam's elastic modulus is 32 GPa and density is 5219 kg/m^3. The first three vibration modes are considered. The mode shapes of the cantilever beam form a matrix $\Phi \in \mathbb{R}^{M \times N}$, where M is the number of candidate sensor nodes, N is the number of vibration modes. The mode shapes for a continuous cantilever beam are given by

$$X_N(x) = [\cosh(\beta_N x) - \cos(\beta_N x)] - \sigma_N [\sinh(\beta_N x) - \sin(\beta_N x)], \tag{5}$$

where

$$\sigma_N = \frac{\sinh(\beta_N L) - \sin(\beta_N L)}{\cosh(\beta_N L) + \cos(\beta_N L)}, \tag{6}$$

with $N = 1, 2, 3$ and $\beta_N L = n\pi$. The mode shapes are normalized by their 2-norm and mass. For every beam profile, the deflections at the nodes are a linear superposition of the three mode shapes, as described by $u_s = \Phi_s q$, where u_s is a vector

containing the beam deflections at each sensor node, $\Phi_s \in \mathbb{R}^{n_{\text{sensor}} \times N}$ is the matrix containing the mode shapes reduced to the sensor nodes with n_{sensor} denoting the desired number of optimal sensors, and q refers to the modal coefficient vector. A dataset of 50,000 beam profiles is generated by uniformly sampling each modal coefficient from the interval $[-1, 1]$.

The clustering implementation uses the same algorithm and parameters specified in Section *Attribution based on correlation: correlated synthetic data*, with the preference parameter adjusted to 0.991 to yield 19 clusters. To implement CAAF, we construct a data-driven model consisting of an MLP that takes deflection measurements at cluster center nodes as inputs and predicts modal coefficients as outputs. While maintaining the basic architecture described in Section *Attribution based on correlation: correlated synthetic data*, we modify the network by using ReLU activation functions in place of LeakyReLU and configuring each hidden layer with 12 neurons.

The main objective for finding OSP on a structure is to ensure that the mode shapes sampled at the sensor nodes are linearly independent and provide low low-variance estimation of the structural response. Success in achieving these goals is commonly measured using the metrics RMS of the off-diagonal entries of the mass-weighted modal assurance criterion (MMAC) matrix, CN of the mode shape matrix, and DET of the reduced Fisher information matrix. RMS of the MMAC measures the level of orthogonality between the reduced mode shapes, where

$$MMAC_{ij} = \frac{\left(\Phi_s^{(i)T}\mathbf{M}_s\Phi_s^{(j)}\right)^2}{\left(\Phi_s^{(i)T}\mathbf{M}_s\Phi_s^{(i)}\right)\left(\Phi_s^{(j)T}\mathbf{M}_s\Phi_s^{(j)}\right)} \tag{7}$$

for $i, j = 1, 2, \ldots, N$, the superscript ^{T} denotes transpose, and $\mathbf{M}_s \in \mathbb{R}^{n_{\text{sensor}} \times N}$ is the mass matrix corresponding to the final sensor set; a lower RMS indicates a better modal orthogonality. CN of the mode shape matrix Φ_s quantifies the linear independence between the mode shapes; a value closer to unity indicates better linear independence. DET of the reduced Fisher information matrix A_s, defined as $A_s = \Phi_s^T\Phi_s$, reveals the level of information contained in the reduced mode shapes and variance of the estimated response; an optimal sensor configuration should maximize the DET.

We compare CAAF to two different existing methods for OSP: EI and KE. The EI method[6] optimizes sensor placement by maximizing the linear independence of target mode shapes. Given a mode shape matrix $\Phi \in \mathbb{R}^{M \times N}$, the method evaluates the contribution of each sensor location through the Fisher Information Matrix $A = \Phi^T\Phi$. The EI vector $E_f \in \mathbb{R}^M$ is computed as the diagonal of $\Phi A^{-1}\Phi^T$, where each entry quantifies a sensor's ability to preserve mode shape independence. The sensor corresponding to the lowest E_f entry is iteratively eliminated until the desired number remains, ensuring minimal information loss. Our EI sensors match those from Friswell and Castro-Triguero[32], which validates our implementation. The KE method[6] identifies OSP by maximizing the captured vibrational energy of target modes. For a structure with a complete mass matrix $\mathbf{M} \in \mathbb{R}^{M \times M}$, the energy metric KE is computed as

$$KE = \Phi \otimes \mathbf{M}\Phi, \tag{8}$$

where $\otimes$ represents element-wise multiplication. Sensor locations are ranked by their row-summed modal energies, selecting positions with the highest dynamic activity.

Airfoil lift prediction
The flow simulation uses a finite-volume, unstructured-mesh LES solver[54]. The spatially-filtered incompressible Navier-Stokes equations are solved with second-order accuracy employing cell-based, low-dissipative, and energy-conservative spatial discretization and a fully-implicit, fractional-step time-advancement method with the Crank–Nicholson scheme. The Poisson equation for pressure is solved using the algebraic multigrid method[55]. The subgrid-scale stress is modeled using the dynamic Smagorinsky model[56,57].

Schematics of the simulation setup are shown in Figure 10. A NACA 0012 airfoil with chord length C is placed horizontally at the origin of the coordinate system. Gusty inflows are generated by placing a cylinder directly upstream of the airfoil. Such geometry has been extensively studied and characterized both experimentally and numerically by others[58–60]. Boundary conditions for both the airfoil surface and cylinder are defined as solid no-slip walls. The simulation domain is periodic in the spanwise direction with the domain size $L_z = 0.1C$. Uniform inflow of velocity U_∞ is set to come in through the left and bottom surfaces of the domain at $Re_c = 10^4$ with an angle (equal to the AoA of the airfoil) relative to the x-axis. Additionally, convective outflow boundary conditions are set for both the top and right surfaces of the domain. The cylinder size, D, and its distance from the airfoil, $L_{c,a}$, vary across different cases. A total of five different cases are simulated, named "None-5", "None-11", "Cylinder-5", "Cylinder-11", and "Cylinder2-0", respectively. Schematics of the simulated geometry for these cases are displayed in Figure 10. Table 3 summarizes the parameters associated with the aforementioned cases.

The computational meshes used in the simulations consist of approximately 20 million cells. The airfoil is enclosed by a C-type mesh, which features approximately 400 cells along the airfoil's circumference. Approximately 160 cells are used

Table 3. Parameters of the simulated cases.

Case	Disturbance	AoA	D/C	$L_{c,a}/D$
None-5	None	5°	N/A	N/A
None-11	None	11°	N/A	N/A
Cylinder-5	Cylinder wake	5°	0.35	3.8
Cylinder-11	Cylinder wake	11°	0.35	3.8
Cylinder2-0	Cylinder wake	0°	1.0	5.0

Type of upstream disturbance, the angle of attack (AoA) of the airfoil, the cylinder diameter (D/C), and the distance between the cylinder center and the airfoil leading edge ($L_{c,a}/D$) are listed for each case.

to discretize the surface of the cylinder in cases where it is present. The boundary layer on the airfoil is captured by a mesh with $\Delta y_1/C \approx 0.0015$, where Δy_1 is the height of the first wall-normal cell. The spanwise direction is uniformly discretized into 128 cells. To verify the LES results, we compare the results of pressure coefficient C_p and skin-friction coefficient C_f, defined as $C_p = P/\left(0.5\rho U_\infty^2\right)$, and $C_f = \tau_w/\left(0.5\rho U_\infty^2\right)$ respectively, where ρ is the fluid density, around the airfoil from the "None-5" case with those generated from DNS by Gao et al.[61] at the same Re_C. The present LES results agree well with the reference DNS results. Additionally, a grid convergence test performed by repeating the "None-11" case with a finer mesh of 40 million cells produces less than 3% difference in the time-averaged lift coefficients, suggesting that the current grid resolution is sufficiently high. The lift force, F_L, is obtained by numerically integrating the surface pressure and shear stress components normal to the freestream. The lift coefficient is then calculated using the lift force by

$$C_L = \frac{F_L}{\frac{1}{2}\rho U_\infty^2 L_z C}. \tag{9}$$

For additional details on the computational setup and validation results, readers are referred to our previous work[43].

With data for the pressure distribution and the lift coefficients generated from the simulations, optimal sensor configurations are identified using CAAF and other data-driven methods: naive FA, QR pivoting with POD modes and Bayesian experimental design. Baseline configurations given by uniformly distributed sensors are also tested. The simulation data serve as inputs to the data-driven methods for identifying optimal sensor locations. While QR pivoting with POD modes requires only pressure distribution data as input, Bayesian experimental design and FA methods require both pressure and lift data.

Neglecting the spanwise direction, the airfoil surface is spatially divided into a total of 376 candidate pressure sensor locations following the computational grid, symmetrically distributed about the chord, covering both the suction and pressure

None-5   None-11   Cylinder-5

Cylinder-11   Cylinder2-0
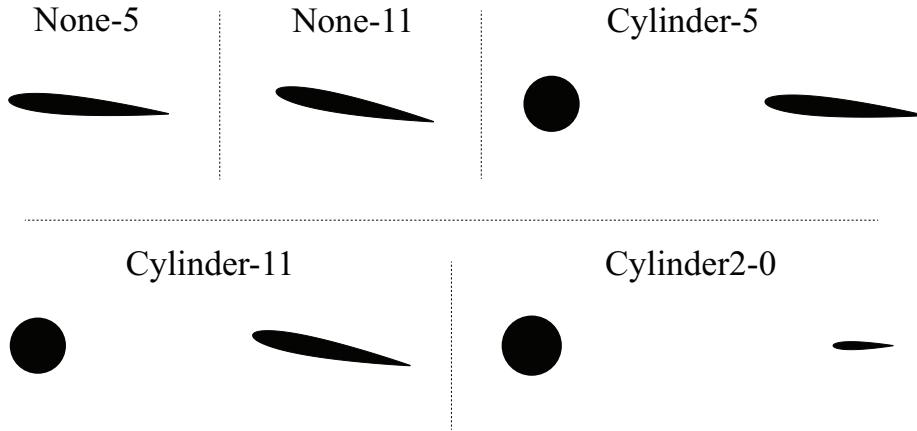
Figure 10. Schematics of the 5 simulated cases listed in Table 3.

sides. The pressure data matrix $P \in \mathbb{R}^{M \times N_t}$ is defined as

$$
P = \begin{bmatrix} P_1^{(1)} & P_2^{(1)} & \cdots & P_M^{(1)} \\ \vdots & \vdots & & \vdots \\ P_1^{(N_t)} & P_2^{(N_t)} & \cdots & P_M^{(N_t)} \end{bmatrix},
$$

(10)

where $P_i^{(j)}$ is the pressure measurement given by the i-th sensor at the j-th snapshot, with N_t time snapshots. The five different cases mentioned above are grouped into training and testing sets. The training set comprises of "None-5", "None-11", and "Cylinder-5" cases, whereas the testing set includes "Cylinder-11" and "Cylinder2-0" cases. Only the data from the training cases are used to determine the optimal sensor locations with $N_t > 30,000$. The implementation steps for QR pivoting of POD modes, Bayesian experimental design, and uniform distribution methods are described as follows.

Firstly, for the QR pivoting of POD modes method, a singular value decomposition is performed on the pressure matrix defined in Equation 10, yielding the predominant modes of the data as illustrated by

$$
P = \Psi \Sigma \Phi^T,
$$

(11)

where Ψ and Φ are the left and right singular vectors respectively, and Σ is the diagonal matrix of singular values. After that, the n_{sensor} most dominant left singular vectors $\Psi_{n_{\text{sensor}}}$ are extracted from the singular value decomposition. The final step involves computing the QR factorization with column pivoting of these vectors and selecting the top pivot locations.

The Bayesian experimental design method[11, 12, 14, 62] uses Bayesian inference to sequentially identify the design parameter s that provides the maximum information gain from prior to posterior, or equivalently the mutual information between the parameters r and the data h. First, a uniform prior probability distribution $Pr(r)$ is constructed. According to Bayes' rule, the posterior probability distribution $Pr(r \mid h, s)$ is proportional to the product of the prior distribution $Pr(r)$ and the likelihood $Pr(h \mid r, s)$. The information gain is measured by a utility function $\mathscr{U}(s)$, defined based on the concept of relative entropy, also known as the Kullback–Leibler (KL) divergence, given by

$$
\begin{aligned}
\mathscr{U}(s) &:= \mathbb{E}_{h|s}[KL(s,h)] = \int_{\mathscr{H}} KL(s,h) Pr(h \mid s) \, dh \\
&= \int_{\mathscr{H}} \int_{\mathscr{R}} Pr(r \mid h, s) \ln \frac{Pr(r \mid h, s)}{Pr(r)} Pr(h \mid s) \, dr \, dh \, .
\end{aligned}
$$

(12)

We evaluate the utility function by substituting the distribution of the lift coefficients $C_L = [C_L^{(1)}, \cdots, C_L^{(N_t)}]^T$ for h and the pressure distribution at a sensor location P_i for r to obtain the utility function $\mathscr{U}_i$, where the sensor location i is considered to be the design parameter s. To estimate the double integrals, we apply the Monte-Carlo sampling method to the $\mathscr{U}_i$ and acquire an estimator of the utility function

$$
\hat{\mathscr{U}}_i \approx \frac{1}{N_{mc}} \sum_{j=1}^{N_{mc}} \left\{ \ln \left[Pr \left(P_i^{(j)} \mid C_L^{(j)} \right) \right] - \ln \left[Pr \left(P_i^{(j)} \right) \right] \right\},
$$

(13)

where N_{mc} is the number of Monte-Carlo samples. The conditional probability distributions are generated by categorizing the lift coefficients C_L into bins based on their values. The pressure is assumed to follow Gaussian mixture distributions, with means and covariance matrices computed directly by fitting the simulated pressure data. Gaussian mixture distributions are necessary to incorporate data from multiple cases. We formed the optimal sensor configuration through a greedy algorithm that sequentially identifies the next optimal location by maximizing the marginal utility, which is expressed mathematically as $i^* = \text{argmax}_{1 \leq i \leq N_s} \mathscr{U}_i$. The utility is calculated for each candidate sensor location, followed by a grid search to determine the optimal location with the maximum utility. The sensor configuration is generated using 15 uniformly spaced bins spanning the range of C_L and $N = 20,000$ samples drawn from 15,000 time snapshots.

Finally, the baseline sensor configurations are obtained by uniformly distributing the sensors on the surface of the airfoil, resulting in an equally spaced sensor array. Moreover, to ensure the information from both the leading and trailing edges is captured by the baseline configurations, two sensors are fixed at the two edges before evenly placing the rest across the airfoil surface.

The clustering step in the CAAF is implemented by the AP algorithm with Pearson's correlation coefficient between the pressure inputs at different candidate sensor locations, P_i, as the affinity metric. The damping and preference parameters are set to 0.5 and the median of the input similarities, respectively. The lift predictions are made by training MLP models with the instantaneous pressure measurements from the identified sensors as input and the lift coefficients at the same time

step as output. The MLP model employs an architecture similar to that outlined in Section *Attribution based on correlation: correlated synthetic data*, with the addition of an extra hidden layer. The pressure and lift coefficient data are interpolated to a uniform time grid with an interval of $0.005C/U_\infty$. The training dataset is normalized and shuffled into batches of 64 data points. Normalization of the dataset is accomplished by scaling both the input pressure and output lift coefficients to a range between -1 and 1, *i.e.* $C_{L,scaled} = (C_L - \langle C_L \rangle)/(\max(C_L) - \min(C_L))$. Normalization is performed separately for the data of each case.

Turbulent wall-normal velocity estimation

In order to obtain high-fidelity data for wall-normal velocity and wall pressure, we perform a DNS of an incompressible turbulent channel flow at $Re_\tau \approx 186$. The simulations are performed by discretizing the incompressible Navier–Stokes equations with a staggered, second-order-accurate, central finite-difference method in space[63], and an explicit third-order-accurate Runge–Kutta method for time advancement[64]. The system of equations is solved via an operator splitting approach[65]. Periodic boundary conditions are imposed in the streamwise and spanwise directions, and the no-slip condition is applied at the walls. The code has been validated in previous studies of turbulent channel flows[66,67]. The streamwise, wall-normal and spanwise domain sizes are $L_{x,c}^+ \approx 5,300$, $L_{y,c}^+ \approx 372$ and $L_{z,c}^+ \approx 2,000$, respectively. The grid spacings in the streamwise and spanwise directions are uniform, with $\Delta x^+, \Delta z^+ \approx 5.2$; non-uniform meshes are used in the wall-normal direction, with the grid stretched towards the wall according to a hyperbolic tangent distribution with $\min(\Delta y^+) \approx 0.18$ and $\max(\Delta y^+) \approx 7.5$. The dataset comprises 900 temporal snapshots with $\Delta t^+ = 1.06$, each divided into 57 sensing sections, yielding over 50,000 data points scaled to the range [-1,1]. The testing dataset consists of all snapshots from one of the sensing sections.

We applied AP clustering to the candidate sensor locations based on streamwise velocity correlations, with damping and preference parameters set to $\lambda = 0.5$ and $p = 0.9$, respectively. All other hyperparameters maintain the same configuration detailed in Section *Attribution based on correlation: correlated synthetic data*. We then train an MLP model to predict v_{10} at the center of the sensing plane using p_w measurements from the 58 cluster centers identified by AP. The MLP architecture follows that described in Section *Attribution based on correlation: correlated synthetic data*, but with two hidden layers of 128 neurons each, totaling approximately 6,000 trainable parameters, and an L_2 regularization strength of 0.01. The model achieves a correlation of 0.88 between prediction and truth. This performance is comparable to the results reported by Park and Choi[45], who achieved a maximum L_2 prediction correlation of 0.96 using the full-state wall pressure as input.

Data and code availability

The source code used to conduct the experiments and analyze the data, as well as the datasets used in this study, are publicly available in the GitHub repository at https://github.com/Mickey-Leung/CAAF.git, with the exception of the channel flow velocity data for the wall-normal velocity estimation due to size limit but are available upon request.

References

1. Yu, L., Wang, N. & Meng, X. Real-time forest fire detection with wireless sensor networks. In *Proc. 2005 Int. Conf. Wirel. Commun. Netw. Mobile Comput.*, vol. 2, 1214–1217 (2005).

2. Tibaduiza Burgos, D. A., Gomez Vargas, R. C., Pedraza, C., Agis, D. & Pozo, F. Damage identification in structural health monitoring: A brief review from its implementation to the use of data-driven applications. *Sensors* **20**, 733 (2020).

3. Andersson, T. R. *et al.* Environmental sensor placement with convolutional gaussian neural processes. *Environ. Data Sci.* **2**, e32 (2023).

4. Castello, C. C., Fan, J., Davari, A. & Chen, R.-X. Optimal sensor placement strategy for environmental monitoring using wireless sensor networks. In *2010 42nd Southeastern Symp. Syst. Theory (SSST)*, 275–279 (IEEE, 2010).

5. Yeong, D. J., Velasco-Hernandez, G., Barry, J. & Walsh, J. Sensor and sensor fusion technology in autonomous vehicles: A review. *Sensors* **21**, 2140 (2021).

6. Kammer, D. C. Sensor placement for on-orbit modal identification and correlation of large space structures. *J. Guid. Control. Dyn.* **14**, 251–259 (1991).

7. Liu, K., Yan, R.-J. & Soares, C. G. Optimal sensor placement and assessment for modal identification. *Ocean. Eng.* **165**, 209–220 (2018).

8. Manohar, K., Brunton, B. W., Kutz, J. N. & Brunton, S. L. Data-driven sparse sensor placement for reconstruction: Demonstrating the benefits of exploiting known patterns. *IEEE Control. Syst. Mag.* **38**, 63–86 (2018).

9. Tabib, M., Skare, K., Bruaset, E. & Rasheed, A. Data-driven spatio-temporal modelling and optimal sensor placement for a digital twin set-up. *Eng. Proc.* **39**, 98 (2023).

10. Singh, A. K. & Hahn, J. Determining optimal sensor locations for state and parameter estimation for stable nonlinear systems. *Ind. & Eng. Chem. Res.* **44**, 5645–5659 (2005).

11. Verma, S., Papadimitriou, C., Lüthen, N., Arampatzis, G. & Koumoutsakos, P. Optimal sensor placement for artificial swimmers. *J. Fluid Mech.* **884**, A24 (2020).

12. Huan, X. & Marzouk, Y. M. Simulation-based optimal bayesian experimental design for nonlinear systems. *J. Comput. Phys.* **232**, 288–317 (2013).

13. Krause, A., Singh, A. & Guestrin, C. Near-optimal sensor placements in gaussian processes: Theory, efficient algorithms and empirical studies. *J. Mach. Learn. Res.* **9** (2008).

14. Lozano-Durán, A. & Arranz, G. Information-theoretic formulation of dynamical systems: causality, modeling, and control. *Phys. Rev. Res.* **4**, 023195 (2022).

15. Otto, S. E. & Rowley, C. W. Inadequacy of linear methods for minimal sensor placement and feature selection in nonlinear systems: A new approach using secants. *J. Nonlinear Sci.* **32**, 69 (2022).

16. Wang, Z., Li, H.-X. & Chen, C. Reinforcement learning-based optimal sensor placement for spatiotemporal modeling. *IEEE Trans. Cybern.* **50**, 2861–2871 (2019).

17. Zhong, Y., Fukami, K., An, B. & Taira, K. Sparse sensor reconstruction of vortex-impinged airfoil wake with machine learning. *Theor. Comput. Fluid Dyn.* 1–19 (2023).

18. Shi, S., Du, D., Mercan, O., Kalkan, E. & Wang, S. A novel data-driven sensor placement optimization method for unsupervised damage detection using noise-assisted neural networks with attention mechanism. *Mech. Syst. Signal Process.* **209**, 111075 (2024).

19. Shapley, L. S. A value for n-person games. *Annals Math. Stud.* 307–317 (1953).

20. Chen, H., Covert, I. C., Lundberg, S. M. & Lee, S.-I. Algorithms to estimate shapley value feature attributions. *Nat. Mach. Intell.* **5**, 590–601 (2023).

21. Ancona, M., Ceolini, E., Öztireli, C. & Gross, M. Towards better understanding of gradient-based attribution methods for deep neural networks. *arXiv preprint arXiv:1711.06104* (2017).

22. Lundberg, S. M. & Lee, S.-I. A unified approach to interpreting model predictions. *Adv. Neural Inf. Process. Syst.* **30** (2017).

23. Aas, K., Jullum, M. & Løland, A. Explaining individual predictions when features are dependent: More accurate approximations to shapley values. *Artif. Intell.* **298**, 103502 (2021).

24. Sundararajan, M., Taly, A. & Yan, Q. Axiomatic attribution for deep networks. In *Proc. Mach. Learn. Res.*, 3319–3328 (PMLR, 2017).

25. Sikdar, S., Bhattacharya, P. & Heese, K. Integrated directional gradients: Feature interaction attribution for neural nlp models. In *Proc. 59th Annu. Meet. Assoc. Comput. Linguist. 11th Int. Joint Conf. Nat. Lang. Process. (Vol. 1: Long Pap.)*, 865–878 (2021).

26. Linardatos, P., Papastefanopoulos, V. & Kotsiantis, S. Explainable ai: A review of machine learning interpretability methods. *Entropy* **23**, 18 (2020).

27. Wang, Z., Mardziel, P., Datta, A. & Fredrikson, M. Interpreting interpretations: Organizing attribution methods by criteria. In *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. Workshops*, 10–11 (2020).

28. Simonyan, K., Vedaldi, A. & Zisserman, A. Deep inside convolutional networks: Visualising image classification models and saliency maps. *arXiv preprint arXiv:1312.6034* (2013).

29. Shrikumar, A., Greenside, P. & Kundaje, A. Learning important features through propagating activation differences. In *Proc. Int. Conf. Mach. Learn.*, 3145–3153 (PMLR, 2017).

30. Deng, J. *et al.* Imagenet: A large-scale hierarchical image database. In *IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, 248–255 (IEEE, 2009).

31. Papadopoulos, M. & Garcia, E. Sensor placement methodologies for dynamic testing. *AIAA J.* **36**, 256–263 (1998).

32. Friswell, M. I. & Castro-Triguero, R. Clustering of sensor locations using the effective independence method. *AIAA J.* **53**, 1388–1391 (2015).

33. Garvey, S., Friswell, M. & Penny, J. Evaluation of a method for automatic selection of measurement locations based on subspace-matching. In *Proc. 14th Int. Modal Anal. Conf.*, vol. 2768, 1546 (1996).

34. Meo, M. & Zumpano, G. On the optimal sensor placement techniques for a bridge structure. *Eng. Struct.* **27**, 1488–1497 (2005).

35. Li, D., Li, H. & Fritzen, C. The connection between effective independence and modal kinetic energy methods for sensor placement. *J. Sound Vib.* **305**, 945–955 (2007).

36. Boettcher, F., Renner, C., Waldl, H.-P. & Peinke, J. On the statistics of wind gusts. *Boundary-Layer Meteorol.* **108**, 163–173 (2003).

37. Granlund, K., Monnier, B., Ol, M. & Williams, D. Airfoil longitudinal gust response in separated vs. attached flows. *Phys. Fluids* **26** (2014).

38. Galway, D., Etele, J. & Fusina, G. Modeling of the urban gust environment with application to autonomous flight. In *Proc. AIAA Atmos. Flight Mech. Conf. Exhibit*, 6565 (2008).

39. Renn, P. I & Gharib, M. Machine learning for flow-informed aerodynamic control in turbulent wind conditions. *Commun. Eng.* **1**, 45 (2022).

40. Beckers, D. & Eldredge, J. D. Deep reinforcement learning of airfoil pitch control in a highly disturbed environment using partial observations. *Phys. Rev. Fluids* **9**, 093902 (2024).

41. Nair, N. J. & Goza, A. Bio-inspired variable-stiffness flaps for hybrid flow control, tuned via reinforcement learning. *J. Fluid Mech.* **956**, R4 (2023).

42. Paris, R., Beneddine, S. & Dandois, J. Robust flow control and optimal sensor placement using deep reinforcement learning. *J. Fluid Mech.* **913**, A25 (2021).

43. Leung, S. C., Zhou, D. & Bae, H. J. Integrated gradients for optimal surface pressure sensor placement for lift prediction of an airfoil subject to gust. In *Proc. AIAA Aviat. Forum ASCEND 2024*, 4148 (2024).

44. Choi, H., Moin, P. & Kim, J. Active turbulence control for drag reduction in wall-bounded flows. *J. Fluid Mech.* **262**, 75–110 (1994).

45. Park, J. & Choi, H. Machine-learning-based feedback control for drag reduction in a turbulent channel flow. *J. Fluid Mech.* **904**, A24 (2020).

46. Guastoni, L. *et al.* Convolutional-network models to predict wall-bounded turbulence from wall quantities. *J. Fluid Mech.* **928**, A27 (2021).

47. Güemes, A., Discetti, S. & Ianiro, A. Sensing the turbulent large-scale motions with their wall signature. *Phys. Fluids* **31** (2019).

48. Milano, M. & Koumoutsakos, P. Neural network modeling for near wall turbulent flow. *J. Comput. Phys.* **182**, 1–26 (2002).

49. Yun, J. & Lee, J. Prediction of near wall velocity in turbulent channel flow using wall pressure based on artificial neural network. In *Proc. Korean Soc. Mech. Eng. Conf.*, 457–460 (KSME, 2017).

50. Choi, H. & Moin, P. On the space-time characteristics of wall-pressure fluctuations. *Phys. Fluids* **2**, 1450–1460 (1990).

51. Liu, Y., Zhou, D. & Wang, M. Wall-pressure fluctuations induced by a forward step in a high-reynolds-number turbulent boundary layer. In *Proc. AIAA/CEAS Aeroacoust. Conf. 2024*, 3392 (2024).

52. Frey, B. J. & Dueck, D. Clustering by passing messages between data points. *Science* **315**, 972–976 (2007).

53. Tan, M. & Le, Q. Efficientnetv2: Smaller models and faster training. In *International Conference on Machine Learning*, 10096–10106 (PMLR, 2021).

54. You, D., Ham, F. & Moin, P. Discrete conservation principles in large-eddy simulation with application to separation control over an airfoil. *Phys. Fluids* **20**, 101515 (2008).

55. Ruge, J. W. & Stüben, K. Algebraic multigrid (amg). In McCormick, S. F. (ed.) *Multigrid Methods*, vol. 3 of *Frontiers Appl. Math.*, 73–130 (SIAM, Philadelphia, PA, 1987).

56. Germano, M., Piomelli, U., Moin, P. & Cabot, W. H. A dynamic subgrid-scale eddy viscosity model. *Phys. Fluids* **3**, 1760–1765 (1991).

57. Lilly, D. K. A proposed modification of the germano subgrid-scale closure method. *Phys. Fluids* **4**, 633–635 (1992).

58. Lefebvre, J., Jones, A. R., Jarman, L. & Smith, M. J. Experimental and numerical investigation of airfoil performance in cylinder wake. In *Proc. 2018 Fluid Dyn. Conf.*, 3232 (2018).

59. Zhang, Z., Wang, Z. & Gursul, I. Aerodynamics of a wing in turbulent bluff body wakes. *J. Fluid Mech.* **937**, A37 (2022).

60. Jiang, Y., Mao, M.-L., Deng, X.-G. & Liu, H.-Y. Numerical investigation on body-wake flow interaction over rod–airfoil configuration. *J. Fluid Mech.* **779**, 1–35 (2015).

61. Gao, W., Zhang, W., Cheng, W. & Samtaney, R. Wall-modelled large-eddy simulation of turbulent flow past airfoils. *J. Fluid Mech.* **873**, 174–210 (2019).

62. Kraskov, A., Stögbauer, H. & Grassberger, P. Estimating mutual information. *Phys. Rev. E* **69**, 066138 (2004).

63. Orlandi, P. *Fluid flow phenomena: a numerical toolkit*, vol. 55 (Springer, 2000).

64. Wray, A. A. Minimal storage time advancement schemes for spectral methods. *NASA Tech. Rep.* **MS 202** (1990).

65. Chorin, A. J. Numerical solution of the navier-stokes equations. *Math. Comput.* **22**, 745–762 (1968).

66. Bae, H. J., Lozano-Duran, A., Bose, S. & Moin, P. Turbulence intensities in large-eddy simulation of wall-bounded flows. *Phys. Rev. Fluids* **3**, 014610 (2018).

67. Bae, H. J., Lozano-Durán, A., Bose, S. T. & Moin, P. Dynamic slip wall model for large-eddy simulation. *J. Fluid Mech.* **859**, 400–432 (2019).

Funding Declaration

Author contributions statement

H.J.B. conceived the experiments and provided guidance; S.C.L. developed the framework, conducted the experiments, analyzed the results, and wrote the manuscript; D.Z. assisted in conducting the experiments and the analysis. All authors reviewed the manuscript.

Competing interests

The authors declare no competing interests.