

SUPERVISED DIMENSIONALITY-REDUCED MULTI-FIDELITY SURROGATE MODEL ASSISTED HIGH-DIMENSIONAL NATURAL LAMINAR FLOW OPTIMIZATION

Gan Yujie[1], Zhao Huan[1], Wang Shekun[1]

[1]School of Aeronautics, Northwestern Polytechnical University, Xi'an Shanxi, 710072, P. R. China

Abstract

The energy supplied by the aircraft power plant is primarily expended to overcome aerodynamic resistance during flight, and the total aircraft drag mainly consists of pressure drag and skin-friction drag, which accounted for a comparable proportion. When the aerodynamic layout of the aircraft is determined, there is very little room for the reduction of pressure drag, so the most focused and promising development direction for drag reduction in modern aircraft is the reduction of skin-friction drag. Therefore, laminar flow drag reduction technologies are now cutting-edge topics in aerodynamic design of aircraft. Among them, the more mature natural laminar flow (NLF) reduction technology has been successfully applied in engineering practice to reduce skin-friction drag by keeping a large laminar flow area in the attached surface layer. With the complexity of vehicle shapes, the refinement of designs, and the diversification of objectives, the design of natural laminar drag reduction is facing the problem 'curse of dimensionality'. To solve this situation, an efficient NLF optimization design method for high-dimensional problems is established based on the partial least squares dimensionality reduced multi-fidelity sparse polynomial chaos-kriging (PLS-DR-MFPCK) surrogate model. PLS-DR-MFPCK enables faster and more accurate prediction for high-dimensional problems by introducing supervised partial least squares (PLS) to train the surrogate model on a low-dimensional manifold. Such a method also combines the PLS and MF-PCK techniques into the jointly surrogate modelling process and adaptively establishes the accurate mapping from the high-dimensional inputs to the output of the system. This PLS-DR-MFPCK method fully considers the effect of inputs on outputs and adaptively trains these hyper-parameters in the surrogate modelling process. It is verified that the PLS-DR-MFPCK model improves modelling efficiency and global approximation accuracy for high-dimensional NLF optimization design problems. To verify the performance of the high-dimensional NLF optimization method based on proposed PLS-DR-MFPCK model, it is applied to the NLF optimization design of a transonic 36-dimensional RAE2822 airfoil. The result shows that the proposed high-dimensional NLF optimization design method significantly improves the optimization efficiency, effectiveness and meets the requirements of NLF optimization design. Compared to the initial profile, the optimized airfoil reduces drag by 55.54 counts (38.48%), while meeting the aerodynamic and geometric constraints. The optimized airfoil achieves an excellent balance between laminar flow area and shockwave strength, minimizing the total drag, which represents the fundamental design requirement of transonic NLF optimization. The NLF optimization framework based on the PLS-DR-MFPCK surrogate model excellently implements the NLF design.

Keywords: Multi-Fidelity Surrogate Model, Partial Least Squares, Natural Laminar Flow Optimization, Sparse Polynomial Chaos-Kriging

1. Introduction

With increasingly stringent requirements on aerodynamic efficiency, fuel economy, and environmental performance for future aircraft, aerodynamic shape optimization (ASO) is facing unprecedented challenges associated with higher-fidelity numerical simulations, refined geometric parameterization, and increasingly complicated multidisciplinary constraints [1-3]. In particular, as modern aircraft configurations become more sophisticated and design objectives become more diversified, natural laminar flow (NLF) drag-reduction technology has attracted extensive attention

https://doi.org/10.71945/icas2026_0502

because of its significant potential for reducing skin-friction drag and improving overall aerodynamic efficiency [4,5]. Since pressure drag reduction becomes progressively limited once the overall aerodynamic layout is determined, maintaining extended laminar flow regions over the aircraft surface has become one of the most promising approaches for further drag reduction in advanced aerodynamic design [6]. Consequently, efficient NLF optimization methods capable of handling complex aerodynamic responses and expensive computational fluid dynamics (CFD) analyses are urgently required.

Surrogate-based optimization (SBO) has become one of the most effective approaches for solving expensive aerodynamic optimization problems involving complex nonlinear design spaces and computationally intensive CFD simulations [3,5]. Among various surrogate modelling techniques, Kriging, polynomial chaos expansion (PCE), polynomial chaos-Kriging (PCK), and Gaussian process regression (GPR) have been widely applied in aerodynamic optimization because of their capability to approximate expensive aerodynamic responses with reduced computational cost [7]. In particular, the PCK surrogate model combines the global approximation characteristics of sparse polynomial chaos expansion with the local interpolation capability of Kriging, thereby demonstrating superior robustness and generalization ability for highly nonlinear aerodynamic responses compared with conventional Kriging or PCE methods [8-9].

To further reduce the dependence on expensive high-fidelity CFD samples, multi-fidelity surrogate modelling techniques have been extensively investigated in recent years [10-11]. These methods exploit the correlation between low-fidelity (LF) and high-fidelity (HF) data by constructing inexpensive LF surrogate models to capture the global trend of the target response and subsequently correcting the discrepancies using a limited number of HF samples. Representative methods, including Co-Kriging and multi-fidelity polynomial chaos expansion, have shown considerable potential in improving modelling efficiency for aerodynamic optimization applications [7,12]. However, the predictive performance and robustness of conventional multi-fidelity surrogate models often deteriorate significantly when confronted with highly nonlinear aerodynamic responses or high-dimensional design spaces [13]. In particular, as the number of design variables increases, the correlation matrices and polynomial candidate sets involved in multi-fidelity surrogate modelling grow rapidly, leading to excessive computational cost, deteriorated convergence, and reduced generalization capability.

With the increasing complexity of aerodynamic configurations and the refinement of shape parameterization strategies, modern NLF optimization problems commonly involve dozens or even hundreds of design variables [4-6,14]. Consequently, SBO algorithms suffer severely from the curse of dimensionality, where the required number of training samples grows exponentially with increasing dimensions [9]. For high-dimensional aerodynamic optimization, insufficient training samples frequently result in inaccurate surrogate models, poor global exploration capability, and excessive CFD evaluations during optimization [15-16]. Although gradient-enhanced surrogate models and high-dimensional model representation techniques have been proposed to alleviate these issues [17-20], their applicability is still limited for computationally expensive aerodynamic optimization involving strongly nonlinear responses and large-scale design spaces.

Another promising strategy for mitigating the curse of dimensionality is to construct surrogate models in a reduced latent space obtained through dimensionality-reduction techniques [21-25]. Existing approaches such as principal component analysis (PCA)-Kriging, active subspace (AS)-based surrogate modelling, and kernel principal component analysis (KPCA)-based methods have demonstrated the potential of dimensionality reduction in high-dimensional optimization [9]. Nevertheless, conventional unsupervised dimensionality-reduction methods usually neglect the correlation between input variables and aerodynamic responses, making it difficult to establish an accurate mapping from the reduced latent space to the target outputs. Moreover, linear dimensionality-reduction methods often fail to effectively capture the intrinsic nonlinear manifold structures embedded in complex aerodynamic datasets [26]. Therefore, supervised dimensionality-reduction techniques that preserve response-sensitive features are more promising for high-dimensional aerodynamic surrogate modelling.

Motivated by the above challenges, this work develops a supervised dimensionality-reduced multi-fidelity sparse polynomial chaos-Kriging surrogate model, namely the PLS-DR-MFPCK model, for

efficient high-dimensional NLF optimization. The proposed method integrates supervised partial least squares (PLS) dimensionality reduction and the multi-fidelity polynomial chaos-Kriging framework into a unified adaptive surrogate-modelling procedure. By projecting the original high-dimensional design variables onto response-sensitive latent spaces, the proposed method effectively alleviates the curse of dimensionality while preserving dominant aerodynamic characteristics. Simultaneously, the MF-PCK framework combines sparse polynomial chaos expansion and Gaussian-process correction to improve the robustness and prediction accuracy for highly nonlinear aerodynamic responses. An efficient NLF optimization framework based on the proposed PLS-DR-MFPCK model is subsequently established and applied to the transonic 36-dimensional RAE2822 airfoil optimization problem. Results demonstrate that the proposed method significantly improves surrogate modelling efficiency, optimization convergence, and aerodynamic drag-reduction performance while satisfying aerodynamic and geometric constraints. These results indicate that the proposed method provides a promising approach for high-dimensional aerodynamic optimization and efficient NLF design.

2. Supervised Partial Least Squares Dimensionality Reduced Multi-Fidelity Sparse Polynomial Chaos-kriging (PLS-DR-MFPCK) Surrogate Model

2.1 MF-PCK Surrogate Model

2.1.1 Low-Fidelity Polynomial Chaos-Kriging Approximation

The multi-fidelity polynomial chaos-Kriging (MF-PCK) framework first constructs a low-fidelity (LF) surrogate model to capture the global trend of the target response. Given a set of LF training samples $\boldsymbol{X}_l = \{\mathrm{X}^{(1)}, \mathrm{X}^{(2)}, \cdots, \mathrm{X}^{(N_l)}\}$ and the associated responses $\mathbf{Y}_l = (y_l(\mathrm{X}^{(1)}), y_l(\mathrm{X}^{(2)}), \cdots, y_l(\mathrm{X}^{(N_l)}))^{\mathrm{T}}$, the LF response is represented as a combination of sparse polynomial chaos expansion and Gaussian process correction:

$$y_l(\mathbf{X}) = \sum_{i=1}^{M} \beta_i \psi_i(\mathbf{X}) + z_l(\mathbf{X}) \tag{1}$$

Where $\psi_i(\mathbf{X})$ denotes the sparse polynomial basis selected using the Least Angle Regression (LAR) strategy. β_i represents the polynomial coefficients; $z_l(\mathbf{X})$ is a zero-mean Gaussian static stochastic process accounting for local discrepancies. The spatial correlation between any two samples $\mathbf{X}^{(i)}$ and $\mathbf{X}^{(j)}$ is characterized using a Gaussian kernel function:

$$R(\mathbf{X}^{(i)}, \mathbf{X}^{(j)}; \boldsymbol{\theta}_l) = \prod_{k=1}^{d} R_k(x_k^{(i)} - x_k^{(j)}, \theta_{lk}) = \prod_{k=1}^{d} \exp\left[-\theta_k (x_k^{(i)} - x_k^{(j)})^2\right] \tag{2}$$

where $\boldsymbol{\theta}_l = (\theta_{l1}, \theta_{l2}, \cdots, \theta_{ld})^{\mathrm{T}}$ are the correlation hyperparameters estimated through likelihood optimization:

$$\{\boldsymbol{\theta}_l\}_{ML} = \arg\min_{\theta_l} \left(\sigma_{z_l}^2 \left|\mathbf{R}_l\right|^{1/N_l}\right), \theta_{l_k} > 0 \tag{3}$$

Based on universal Kriging theory [13], the resulting low-fidelity PC-Kriging predictor can be expressed as:

$$\hat{y}_l(\mathbf{X}) = \sum_{i=1}^{M} \beta_i \psi_i(\mathbf{X}) + \mathbf{r}_l^{\mathrm{T}} \mathbf{R}_l^{-1} (\mathbf{Y}_l - \boldsymbol{\Psi}\boldsymbol{\beta}) \tag{4}$$

In the above equation, $\mathbf{r}_l$ is the correlation vector between the prediction point $\mathbf{X}$ and LF training samples; $\mathbf{R}_l$ is the sample correlation matrix; $\boldsymbol{\beta}$ is the polynomial coefficient, obtained by Generalized Least Squares (GLS) estimation.

2.1.2 Multi-Fidelity Correction Strategy

To improve prediction accuracy, the LF surrogate is further corrected using high-fidelity (HF) information through a hybrid multiplicative-additive correction formulation:

$$y_h(\mathbf{X}) = \alpha_0 \hat{y}_l(\mathbf{X}) + \underbrace{y_c(\mathbf{X}) + z_h(\mathbf{X})}_{C(\mathbf{X})} \tag{5}$$

where α_0 is a scaling coefficient, $y_c(\mathbf{X})$ denotes the polynomial chaos correction expansions, and $z_h(\mathbf{X})$ is an additional zero-mean static Gaussian stochastic process describing residual discrepancies between LF and HF responses. The model hyperparameters are $\boldsymbol{\theta}_h = (\theta_{h_1}, \theta_{h_2}, \cdots, \theta_{h_d})^{\mathrm{T}}$ .The correction component is represented using sparse polynomial chaos expansion:

$$y_c(\mathbf{X}) = \sum_{c_i \in A_c} \alpha_{c_i} \psi_{c_i}(\mathbf{X}) \tag{6}$$

where α_{c_i} is the corrected expanded polynomial term coefficients; A_C is the selected subset of important polynomial basis functions inherited from the LF surrogate model.

Combining the LF approximation and correction terms yields the final MF-PCK formulation:

$$\hat{y}_h(\mathbf{X}) = \alpha_0 \hat{y}_l(\mathbf{X}) + \sum_{c_i \in A_c} \alpha_{c_i} \psi_{c_i}(\mathbf{X}) + z_h(\mathbf{X}) \tag{7}$$

The optimal corrective expansion terms A_C is determined adaptively according to the leave-one-out cross-validation (LOOCV)-based adaptive basis function selection method[13], enabling efficient representation of nonlinear discrepancies across fidelity levels.

2.1.3 Multi-Fidelity PC-Kriging Model Construction

After determining the terms in Eq. (5), the MF-PCK model is trained within the universal Kriging framework. The regression coefficients are estimated using generalized least squares, while the covariance hyperparameters are optimized through likelihood maximization. The final predictor of the MF-PCK model is expressed as:

$$\hat{\boldsymbol{y}}_h(\mathbf{X}) = \alpha_0 \hat{\boldsymbol{y}}_l(\mathbf{X}) + \sum_{i=1}^{M_c} \alpha_{c_i} \psi_{c_i}(\mathbf{X}) + \mathbf{r}_h^{\mathrm{T}}(\mathbf{X}) \mathbf{R}_h^{-1} \left(\mathbf{Y}_h - \mathbf{F}\alpha \right) \tag{8}$$

where $\boldsymbol{\Gamma} = \mathbf{r}_h^{\mathrm{T}} \mathbf{R}_h^{-1} \mathbf{F} - \boldsymbol{\Lambda}^{\mathrm{T}}$; $\mathbf{F}$ denotes the measurement matrix; $\mathbf{R}_h$ is the HF correlation matrix; $\mathbf{r}_h$ is the correlation vector between the prediction point $\mathbf{X}$ and HF training samples; $\boldsymbol{\Lambda}$ denotes the trend vector. The mean-squared error of the MF-PCK model at the prediction point $\mathbf{X}$ is:

$$s_{y_h}^2(\mathbf{X}) = \sigma_{z_h}^2 \left(1.0 - \mathbf{r}_h^{\mathrm{T}} \mathbf{R}_h^{-1} \mathbf{r}_h + \boldsymbol{\Gamma} \left(\mathbf{F}^{\mathrm{T}} \mathbf{R}_h^{-1} \mathbf{F} \right)^{-1} \boldsymbol{\Gamma}^{\mathrm{T}} \right) \tag{9}$$

Compared with conventional Kriging-based multi-fidelity methods, the MF-PCK approach incorporates sparse polynomial chaos expansion into both trend approximation and discrepancy correction, thereby improving robustness and prediction capability for highly nonlinear aerodynamic responses.

2.2 Partial Least Squares Dimensionality Reduction

To alleviate the curse of dimensionality in high-dimensional surrogate modeling, the partial least squares (PLS) technique is introduced to identify latent directions that exhibit the strongest correlation between input variables and output responses.

Unlike conventional unsupervised dimensionality reduction methods, PLS extracts principal components t_l by maximizing the covariance between projected input variables $\mathbf{t}_l = \mathbf{X}^{(l-1)} \mathbf{w}^{(l)}$ and target responses $\mathbf{y}^{(l-1)}$ [27]. Consequently, the retained latent variables preserve the most response-relevant information in the original design space. For the *l*-th latent component, the projection direction is obtained by solving:

$$\mathbf{w}^{(l)} = \arg\max \ \mathrm{Cov}^2\left(\mathbf{X}^{(l-1)}\mathbf{w}^{(l)}, \mathbf{y}^{(l-1)}\right) \tag{10}$$

subject to the normalization constraint:

$$\| \mathbf{w}^{(l)} \| = 1 \tag{11}$$

The extracted latent variable is then used to update the residual matrices of both input and output spaces iteratively. The following relationship exists:

$$\mathbf{X}^{(l-1)} = \mathbf{t}_l \mathbf{p}^{(l)} + \mathbf{X}^{(l)}, \ \ \mathbf{y}^{(l-1)} = \mathbf{t}_l \mathbf{c}^{(l)} + \mathbf{y}^{(l)} \tag{12}$$

Where $\mathbf{p}^{(l)}$ (an $1\times d$ vector) and $\mathbf{c}^{(l)}$ (an $1\times m$ vector) are vectors of regression coefficients, also called loading vectors, and calculated by the following equations:

$$\mathbf{p}^{(l)} = \frac{\mathbf{X}^{(l-1)^T}\mathbf{t}_l}{\mathbf{t}_l^T\mathbf{t}_l}, \ \ \mathbf{c}^{(l)} = \frac{\mathbf{y}^{(l-1)^T}\mathbf{t}_l}{\mathbf{t}_l^T\mathbf{t}_l} \tag{13}$$

For more details about PLS, see [28-30].

Assuming that a total of h principal components are extracted, PLS can be understood as projecting the independent variables, $x_1,\ldots,x_d$ in a d -dimensional original space into an h -dimensional latent space to obtain h new variables-principal components. Usually, h is much less than d , $h \ll d$. The PLS model can be expressed in a compact matrix form using the score, loading, and weight matrices. Specifically, the score matrix $\mathbf{T} = [\mathbf{t}_1, \mathbf{t}_2, \ldots, \mathbf{t}_h] \in \mathbb{R}^{N\times h}$ (N denotes the number of samples) collects the latent score vectors, where each score vector is obtained as:

$$\mathbf{t}_i = \mathbf{X}^{(i-1)}\mathbf{w}^{(i)}, \quad i = 1,2,\ldots,h \tag{14}$$

With $\mathbf{W} = [\mathbf{w}^{(1)}, \mathbf{w}^{(2)}, \ldots, \mathbf{w}^{(h)}] \in \mathbb{R}^{d\times h}$ denoting the weight matrix that defines the directions of maximum covariance between the independent and dependent variables. The loading matrix $\mathbf{P} = [\mathbf{p}^{(1)}, \mathbf{p}^{(2)}, \ldots, \mathbf{p}^{(h)}] \in \mathbb{R}^{d\times h}$ characterizes the contribution of each original variable to the latent components. Consequently, the original input matrix can be approximated as

$$\mathbf{X} \approx \mathbf{T}\mathbf{P}^T \tag{15}$$

2.3 Supervised dimensionality reduced (DR) surrogate modelling by combining PLS and MF-PCK

Through supervised dimensionality reduction, the PLS technique effectively compresses the original high-dimensional design space while preserving dominant response-sensitive features, thereby significantly enhancing the training efficiency and predictive capability of the subsequent MF-PCK surrogate model in latent space. Since this is a multi-fidelity surrogate model, PLS dimensionality reduction is performed separately for the high-fidelity (HF) and low-fidelity (LF) datasets, resulting in projections onto two distinct latent spaces. The overall modeling procedure of the PLS-DR-MFPCK method is summarized as follows:

(1) Data standardization. The HF and LF samples are independently standardized, obtaining the normalized datasets $x_h^{(1)}, x_h^{(2)}, \cdots, x_h^{(N_h)}$ and $x_l^{(1)}, x_l^{(2)}, \cdots, x_l^{(N_l)}$

(2) PLS-based feature extraction and dimensionality reduction. PLS is applied independently to the HF and LF datasets to extract dominant features and reduce dimensionality. For HF samples, the resulting latent score matrix, loading matrix, and weight matrix are given by $\mathbf{T}_{train,h} = [\mathbf{t}_h^{(1)}, \mathbf{t}_h^{(2)}, \cdots, \mathbf{t}_h^{(n_h)}]$, $\mathbf{P}_h = [\mathbf{p}_h^{(1)}, \mathbf{p}_h^{(2)}, \cdots, \mathbf{p}_h^{(n_h)}]$ and $\mathbf{W}_h = [\mathbf{w}_h^{(1)}, \mathbf{w}_h^{(2)}, \cdots, \mathbf{w}_h^{(n_h)}]$ respectively. Similarly, for LF samples, the corresponding matrices are $\mathbf{T}_{train,l} = [\mathbf{t}_l^{(1)}, \mathbf{t}_l^{(2)}, \cdots, \mathbf{t}_l^{(n_l)}]$, $\mathbf{P}_l = [\mathbf{p}_l^{(1)}, \mathbf{p}_l^{(2)}, \cdots, \mathbf{p}_l^{(n_l)}]$ and $\mathbf{W}_l = [\mathbf{w}_l^{(1)}, \mathbf{w}_l^{(2)}, \cdots, \mathbf{w}_l^{(n_l)}]$. Here, n_h and n_l denote the numbers of retained principal components for HF and LF datasets, respectively, i.e., the dimensionalities of the reduced latent spaces.

(3) Construction of the MF-PCK surrogate model. The MF-PCK surrogate model is trained using the dimensionality-reduced datasets $(\mathbf{T}_{train,l},\ \mathbf{Y}_l)$ and $(\mathbf{T}_{train,h},\ \mathbf{Y}_h)$, during which the model hyperparameters are determined.

(4) Surrogate prediction. For a test sample $\boldsymbol{X}_{test}$, projections onto the LF and HF latent spaces are performed as:

$$\begin{aligned}\mathbf{T}_{test,l} &= \mathbf{X}_{test}\mathbf{W}_l(\mathbf{P}_l^T\mathbf{W}_l)^{-1}\\ \mathbf{T}_{test,h} &= \mathbf{X}_{test}\mathbf{W}_h(\mathbf{P}_h^T\mathbf{W}_h)^{-1}\end{aligned} \tag{16}$$

The MF-PCK surrogate model trained in the latent spaces is then employed to predict the corresponding responses on the test sample.

PLS-DR-MFPCK integrates PLS and MF-PCK into an adaptive joint modelling framework, which can be expressed as

$$\boldsymbol{y}_{\text{PLS-DR-MFPCK}} = \hat{\boldsymbol{y}}_{\text{MF-PCK}}(\mathbf{T}_l;\ \mathbf{T}_l = \text{PLS}(\mathbf{X}_l;\ n_l),\ \mathbf{T}_h;\ \mathbf{T}_h = \text{PLS}(\mathbf{X}_h;\ n_{\boldsymbol{h}}),\ \boldsymbol{\theta}_l,\ \boldsymbol{\theta}_h,\ \mathbf{Y}_l,\ \mathbf{Y}_h) \tag{17}$$

The effective dimensionality of the HF and LF samples after dimensionality reduction is determined by executing a nested optimization loop, as illustrated in Figure 1. In the outer loop, the PLS dimensionality-reduction models' parameters-the effective dimensions of the HF and LF feature spaces are optimized to minimize the prediction error until convergence. In the inner loop, the surrogate model hyperparameters are optimized by maximizing the likelihood function using the PLS-reduced samples as input until the optimal surrogate model is output alongside its corresponding predictions. This nested optimization strategy enables simultaneous optimization of both the MF-PCK surrogate and the PLS dimensionality-reduction model, thereby achieving joint training. To reduce CFD analysis time for high-fidelity samples, prediction errors are determined using Leave-One-Out Cross-Validation (LOOCV) error：

$$\varepsilon_{LOO} = \sum_{i=1}^{N_h}\left(\hat{y}_{MF-PCK}^{(-i)}\left(\mathbf{X}^{(i)}\right) - f\left(\mathbf{X}^{(i)}\right)\right)^2 \tag{18}$$

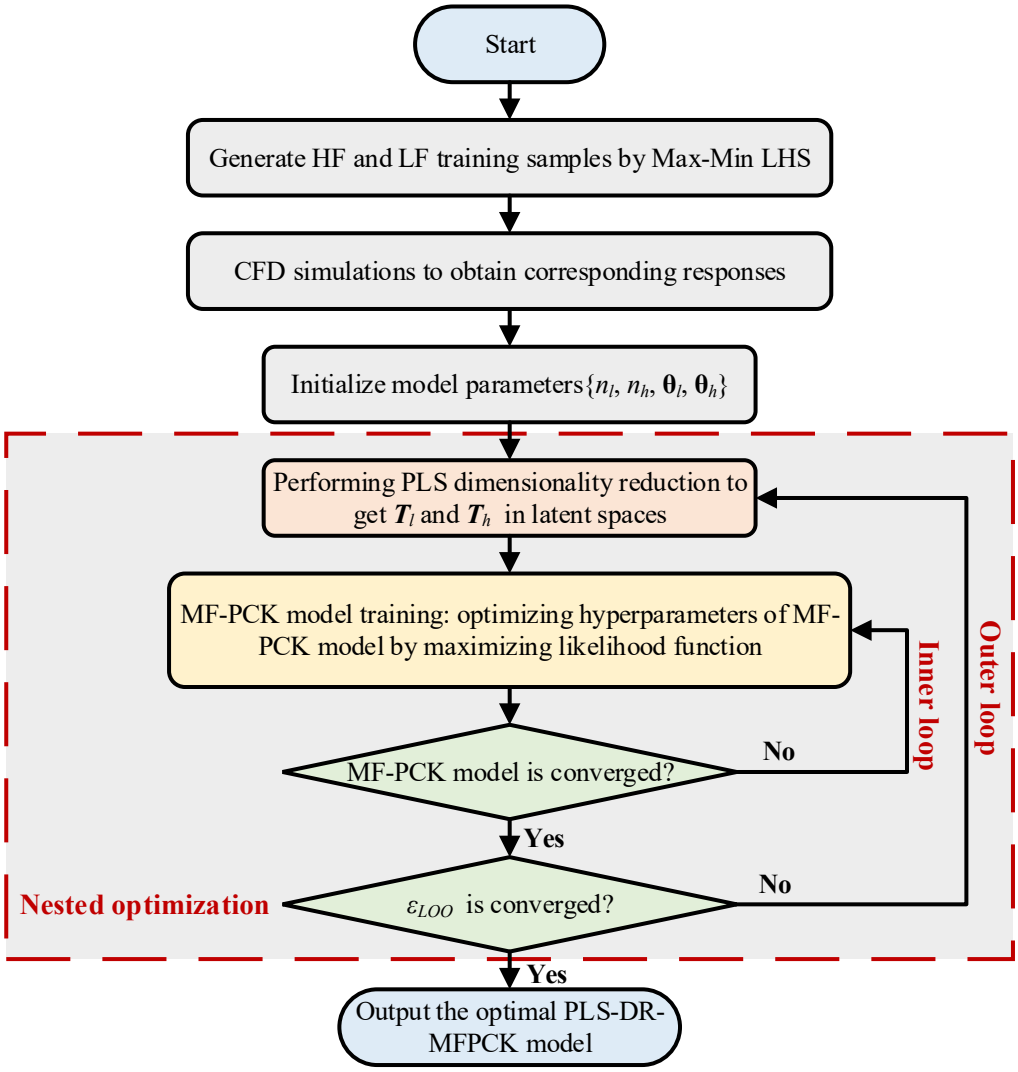


Figure 1 – The nested optimization flowchart of the PLS-DR-MFPCK modelling.

3. Natural laminar flow design frame

To enable efficient natural laminar flow (NLF) optimization, an NLF optimization framework is established based on the PLS-DR-MFPCK model, as illustrated in Figure 2. The overall procedure consists of the following steps:

(1) Defining the design variables, objective function and associated constraints of the optimization problem.

(2) Generating initial high-fidelity and low-fidelity samples within the design space using the Maximum-Minimum Latin Hypercube Sampling (LHS) strategy.

(3) Evaluating target responses at the sampled points by invoking the CFD solver.

(4) Constructing the initial PLS-DR-MFPCK surrogate model.

(5) Searching for the new candidate HF and LF points that minimize the objective function according to the VF-EI infill criterion (the responses are predicted by PLS-DR-MFPCK model), with the particle swarm optimization (PSO) algorithm.

(6) Evaluating the selected candidate HF and LF samples' corresponding target responses using the CFD solver.

(7) Evaluating whether the termination criteria are satisfied. If satisfied, terminating the optimization; otherwise, augmenting the training dataset with the newly selected HF and LF samples together with their corresponding CFD-evaluated responses, updating the surrogate model accordingly, and reinitiating the search for the optimal solution.

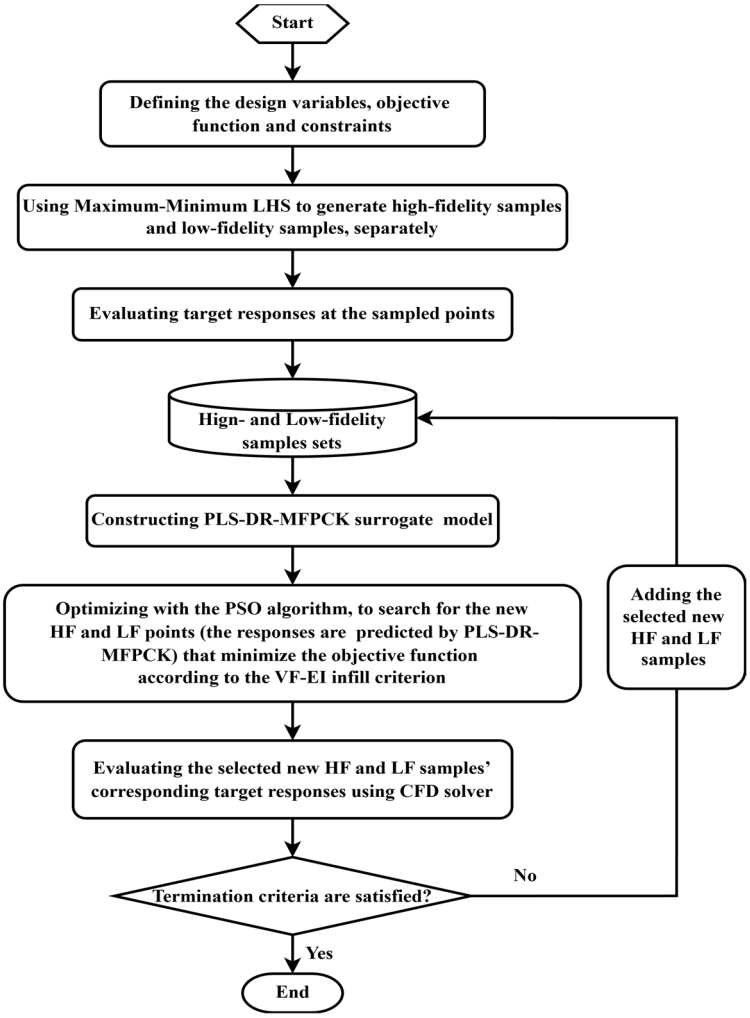
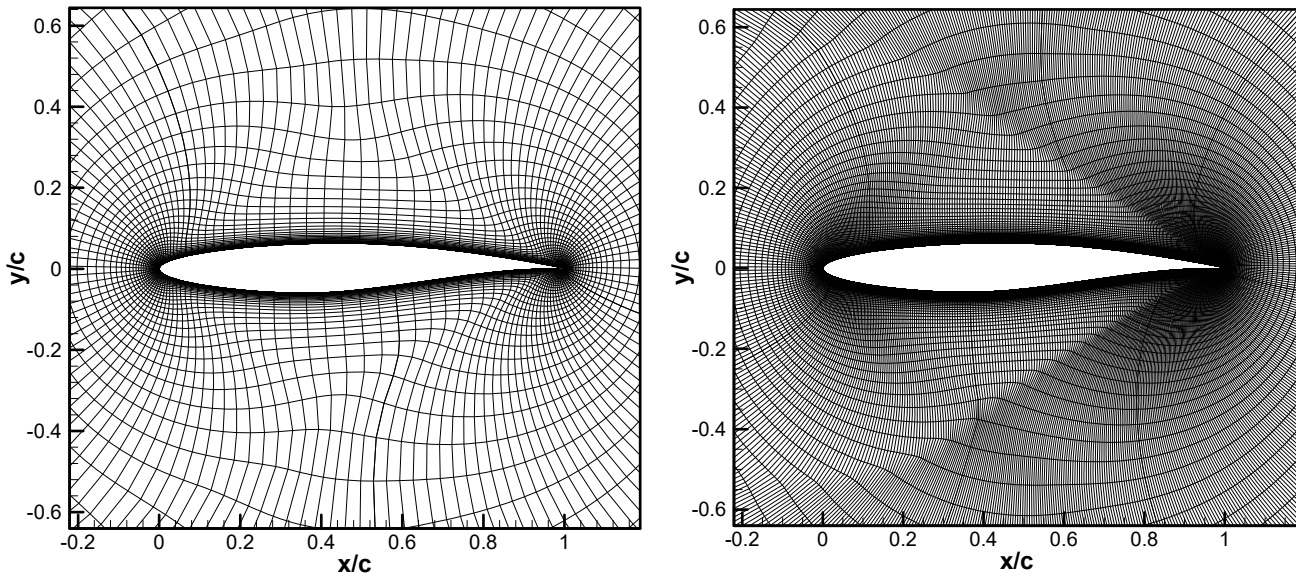


Figure 2 – The Natural Laminar Flow optimization Framework.

4. Natural laminar flow design of Transonic 36-dimensional RAE2822 airfoil

For this optimization case, the initial profile is the RAE2822 airfoil. Design cruise state is Ma = 0.734, C_L = 0.824, Re = 6.5 × 10^6. The HF and LF aerodynamic data are characterized using two sets of computational grids with different densities: the amount of HF grids' elements is 228,960 and the amount of LF grids' elements is 22,896, as shown in Figure 3. The geometry is parameterized by 36 class function/shape function transformation (CST) design variables (18 for the upper surface and 18 for the lower surface). The calculation expense of the HF computational grid is 8.314 times that of the LF computational grid. Table 1 lists the baseline aerodynamic coefficients computed on the two grids. Although both solutions are evaluated at the target lift coefficient, the LF grid gives a larger drag coefficient and a less negative pitching moment, indicating that direct LF prediction is insufficient and that multi-fidelity (MF) correction is necessary. Figure 4 presents the CST-based design space of the RAE2822 airfoil. The design space allows controlled changes in camber, thickness distribution and leading-edge radius while maintaining a realistic transonic airfoil topology. For this airfoil example, 100 HF samples and 500 LF samples are first generated globally using the maximum-minimum LHS method to construct the initial surrogate model.



(a) low fidelity (b) high fidelity

Figure 3 - Two sets of computational grids for RAE2822 airfoil

Table 1 - Calculation results of the baseline airfoil computed on HF and LF grids

Grid Type	C_L	C_D	C_M
High-Fidelity	0.824	1.44332×10^{-2}	-0.115516
Low-Fidelity	0.824	2.28621×10^{-2}	-9.34080×10^{-2}

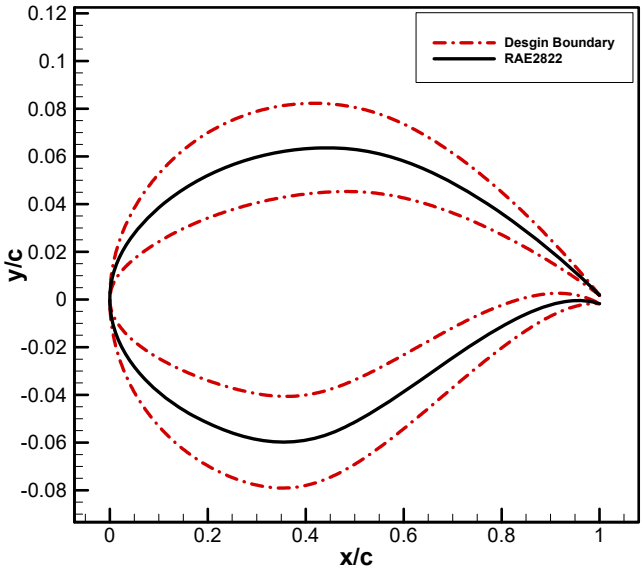


Figure 4 - The design space of the RAE2822 airfoil

Transonic natural laminar flow optimization of an airfoil profile requires a delicate trade-off between obtaining a larger laminar flow coverage area and mitigating shock-wave strength. This is because a long region of favorable pressure gradient helps to maintain the laminar flow region but leads to a steeper pressure recovery near the trailing edge, which tends to amplify shock waves' strength. Therefore, the optimization objective can be attributed to obtaining the minimum total drag coefficient of the airfoil under the fixed lift coefficient of 0.824. The constraints require that the maximum thickness of the airfoil is no less than its initial value, and that the pitching moment coefficient also remains greater than or equal to its initial value. The design optimization problem can be mathematically formulated as follows:

$$Min \quad C_D$$
$$s.t. \begin{cases} t_{max} \geq t_{max,\,initial} \\ C_L \geq C_{L,\,initial} \\ C_M \geq C_{M,\,initial} \end{cases} \tag{19}$$

PSO is adopted for global search. Both HF and LF samples are sequentially added using the VF-EI infill criterion to improve the prediction accuracy of the surrogate model in the region of interest.

Figure 5 compares the drag-convergence histories of the MF-PCK and PLS-DR-MFPCK assisted optimizations. With 250 additional HF samples, PLS-DR-MFPCK reduces C_D rapidly in the initial stage and finally reaches a considerably lower drag level. MF-PCK improves the design in several discrete steps but then stagnates, suggesting convergence to a local optimum. Referring to [19], Figure 6 shows the CPU-time breakdown in the first iteration. By introducing PLS-based dimension reduction, the proposed model substantially decreases surrogate-training cost relative to the original MF-PCK framework. Although the training time is still visible because separate response models are required for C_D and C_M, the reduced modeling expense improves the feasibility of high-dimensional transonic airfoil optimization. More importantly, the trained surrogate enables repeated prediction of numerous candidate geometries without repeatedly invoking expensive CFD analyses.

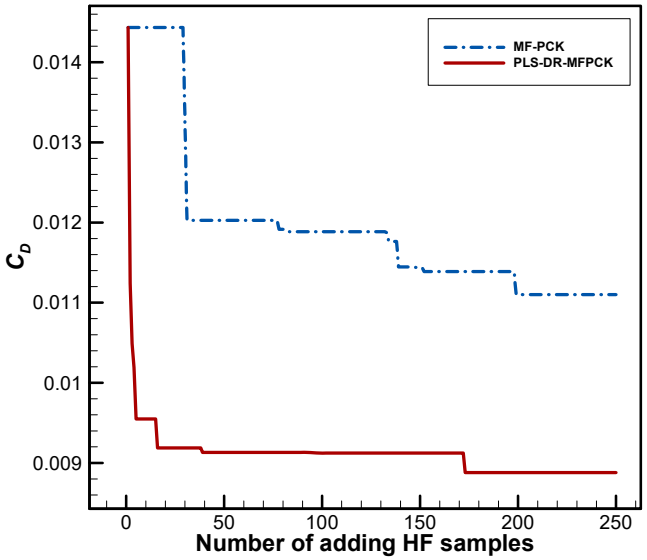


Figure 5 - C_D iteration history during the RAE2822 airfoil optimization

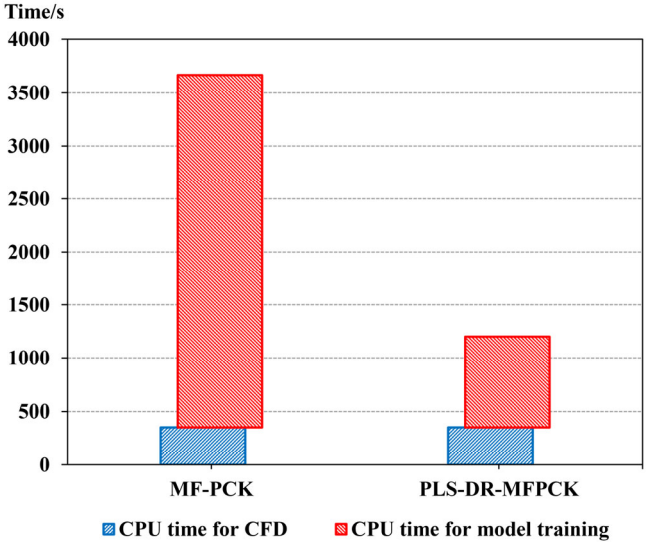


Figure 6 - Breakdown of the CPU time for RAE2822 airfoil optimization

Figure 7 compares the airfoil profiles before and after optimization. Both optimized geometries satisfy the thickness constraint. Relative to the baseline RAE2822, the optimized shapes slightly reduce the leading-edge radius, redistribute the upper-surface curvature and maintain adequate aft-body thickness. The PLS-DR-MFPCK result exhibits a smoother and more coordinated modification over the forward and mid-chord regions, whereas the MF-PCK result contains a less favorable local deformation. These geometric changes are closely related to weakening the transonic compression process while retaining the required lift and moment characteristics. Figure 8 gives the surface pressure-coefficient distributions. The baseline airfoil has an abrupt pressure recovery, corresponding to a pronounced shock. The optimized airfoils reshape the suction distribution and reduce the pressure jump. The PLS-DR-MFPCK design provides the smoothest pressure recovery and the weakest shock signature, which directly reduces wave-related pressure drag. Figure 9 shows the comparison of the pressure coefficient contours. The baseline case contains a broad and intense low-pressure region above the upper surface, while the optimized designs shrink this region and produce a more gradual downstream pressure field. Among them, the PLS-DR-MFPCK contour shows the most compact low-pressure zone and the mildest compression pattern, demonstrating better control of shock-induced pressure losses. Figure 10 presents the Mach-number contours. For

the baseline airfoil, a distinct local supersonic pocket develops on the upper surface and terminates in a relatively strong shock. After optimization, the high-Mach-number region becomes shorter and less intense, especially for the PLS-DR-MFPCK design. Therefore, the drag reduction is not achieved by suppressing acceleration completely, but by moderating shock strength and shock/boundary-layer interaction under the prescribed lift. Figure 11 compares the skin-friction coefficient distributions. The optimized cases exhibit earlier transition on the upper surface, with the MF-PCK design performing more remarkably. This is an acceptable trade-off in the present transonic NLF optimization, because pressure-gradient reshaping for shock mitigation may increase friction drag locally while significantly reducing pressure drag, enabling the PLS-DR-MFPCK design to achieve minimum total drag.

Table 2 summarizes the aerodynamic coefficients and geometric properties of the baseline and optimized airfoils. The MF-PCK design reduces C_D from 1.44332×10^{-2} to 1.11001×10^{-2}, corresponding to a reduction of 33.33 counts, or 23.09%. The PLS-DR-MFPCK design further reduces C_D to 8.87893×10^{-3}, giving a reduction of 55.54 counts, or 38.48%. This improvement mainly comes from the decrease in pressure drag, which is reduced by 60.54 counts, although the friction drag increases by 5.00 counts. Both optimized airfoils satisfy the constraints on pitching moment coefficient and thicknesses. Overall, the PLS-DR-MFPCK framework achieves a superior balance between laminar-flow-oriented pressure shaping, shock weakening and constraint satisfaction, consistent with the transonic natural laminar flow optimization design requirements. It also alleviates the local-optimum tendency observed in MF-PCK, demonstrating stronger robustness and effectiveness for high-dimensional transonic NLF airfoil optimization.

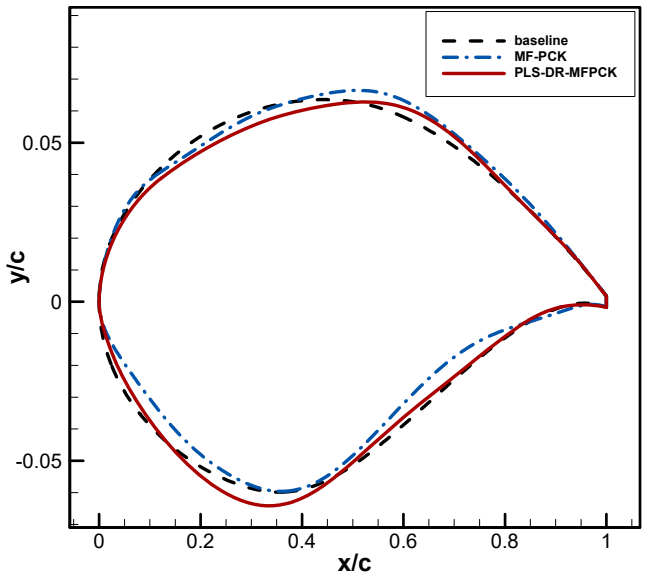


Figure 7 - Profiles of the airfoils before and after optimization

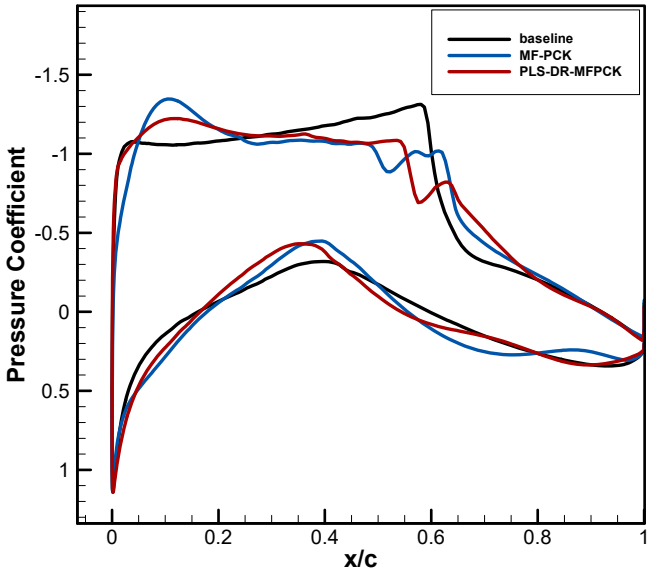
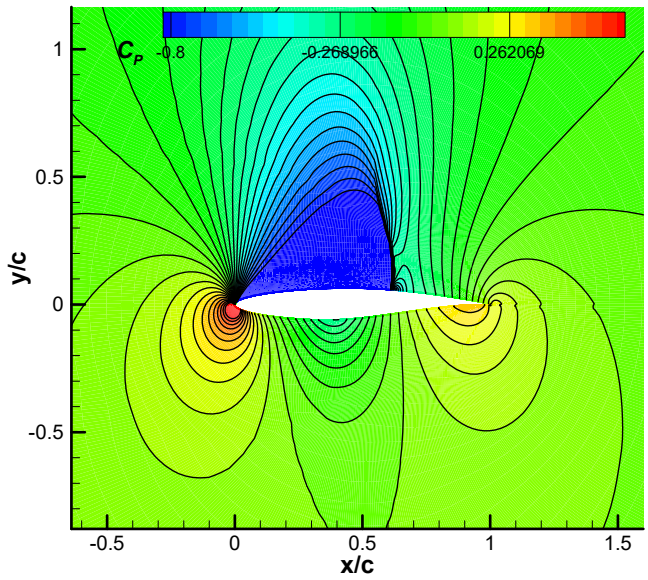
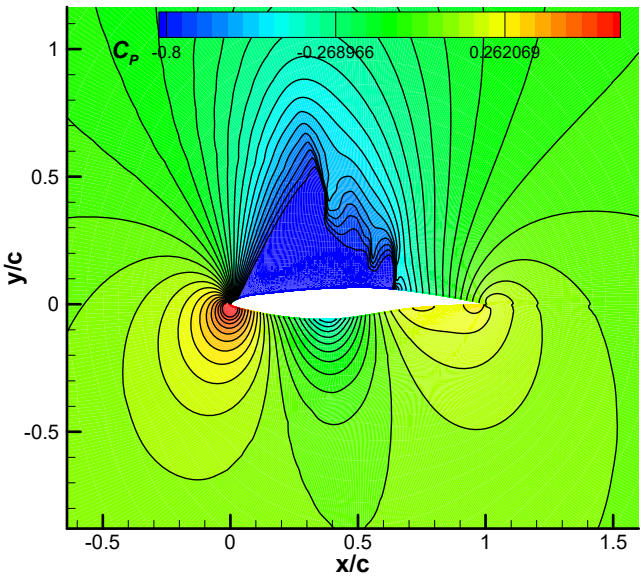


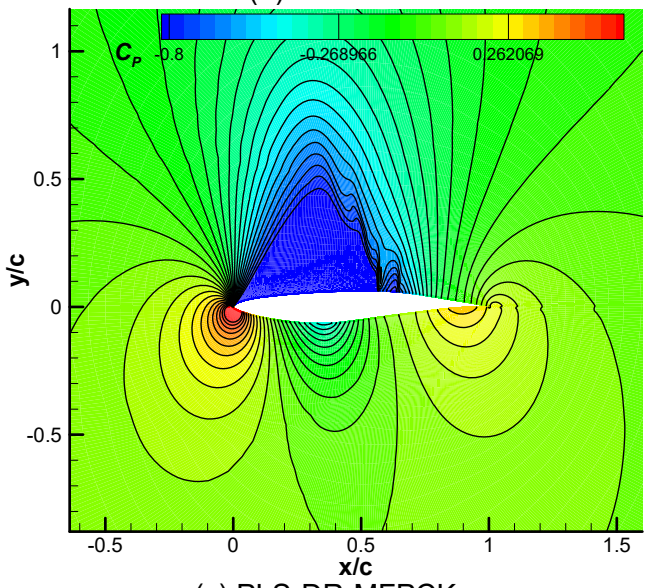
Figure 8 - Pressure coefficient distribution curves before and after optimization



(a) baseline

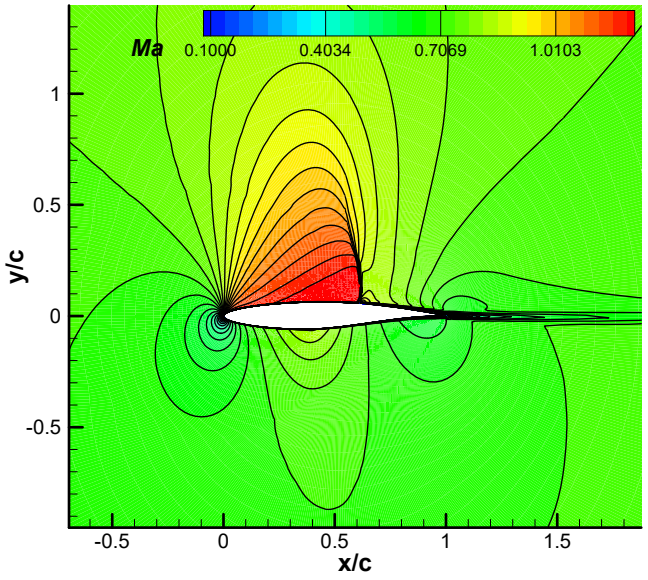


(b) MF-PCK

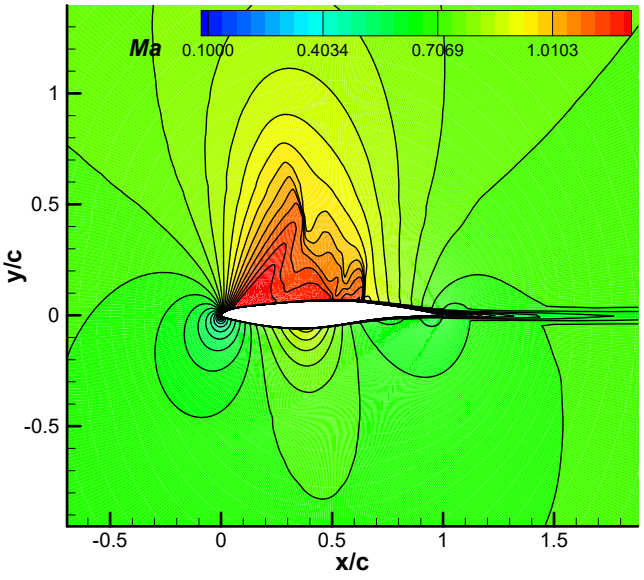


(c) PLS-DR-MFPCK

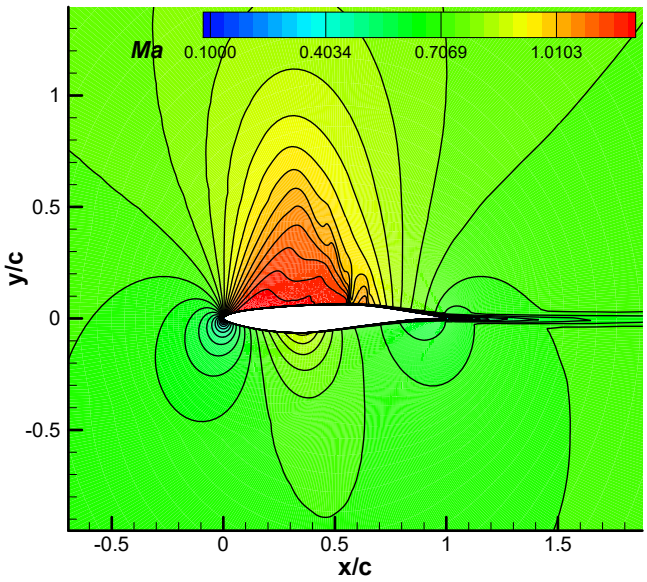
Figure 9 - Pressure coefficient contours before and after optimization



(a) baseline



(b) MF-PCK



(c) PLS-DR-MFPCK

Figure 10 - Mach number contours before and after optimization

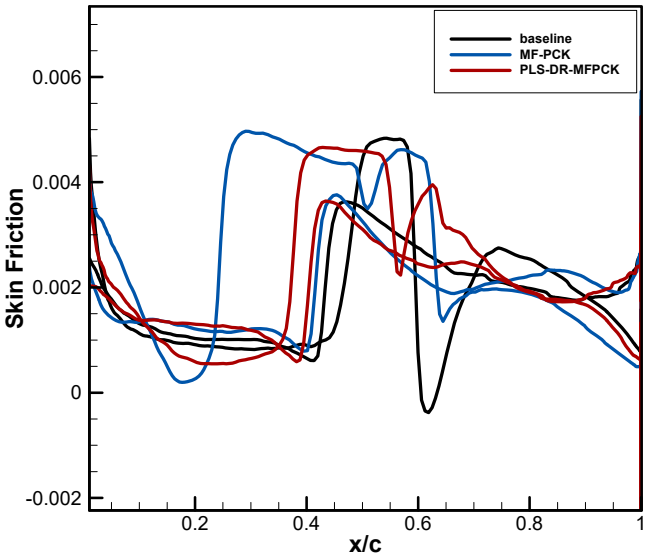


Figure 11 - Skin-friction drag coefficient curves before and after optimization

Table 2 - Comparison of the airfoils' aerodynamic data and geometric properties before and after optimization

Airfoil	C_L	C_D	C_{DP}	C_{DF}	C_M	t_{max}
baseline	0.824	1.44332×10^{-2}	1.07102×10^{-2}	3.72301×10^{-3}	-0.115516	0.1224296
MF-PCK	0.824	1.11001×10^{-2}	6.48960×10^{-3}	4.61050×10^{-3}	-0.110342	0.1224429
PLS-DR-MFPCK	0.824	8.87893×10^{-3}	4.65635×10^{-3}	4.22258×10^{-3}	-0.109631	0.1224301

5. Conclusion

This study develops a supervised dimensionality-reduced multi-fidelity surrogate-assisted framework for efficient high-dimensional natural laminar-flow optimization. The proposed PLS-DR-MFPCK model integrates partial least squares dimensionality reduction with multi-fidelity sparse polynomial chaos-Kriging in a unified adaptive modelling process. By projecting the original design variables into response-sensitive latent spaces, the method reduces the effective dimensionality of the problem while retaining the dominant aerodynamic features. Meanwhile, the multi-fidelity PCK formulation exploits inexpensive low-fidelity information and limited high-fidelity CFD samples to improve global approximation accuracy and reduce the cost of surrogate construction. The proposed framework

was applied to the transonic NLF optimization of a 36-dimensional RAE2822 airfoil at Ma = 0.734, C_L = 0.824 and Re = 6.5 × 10^6. Compared with the conventional MF-PCK-assisted optimization, the PLS-DR-MFPCK approach achieved faster drag convergence, lower surrogate-training cost and better avoidance of local optima. The optimized airfoil reduced the total drag coefficient from 1.44332 × 10^{-2} to 8.87893 × 10^{-3}, corresponding to a drag reduction of 55.54 counts, or 38.48%, while satisfying the prescribed lift, pitching-moment and maximum-thickness constraints. Flow-field analyses indicate that the improvement is mainly attributed to pressure-distribution reshaping, shock-wave weakening and mitigation of shock-induced pressure losses. Although the friction drag slightly increases, the substantial reduction in pressure drag leads to a clear decrease in total drag. These results demonstrate that the proposed PLS-DR-MFPCK-based optimization framework provides an accurate, efficient and robust approach for high-dimensional transonic NLF design, and shows strong potential for more complex aerodynamic optimization problems.

6. Contact Author Email Address

Zhao Huan, huanzhao@nwpu.edu.cn

7. Copyright Statement

The authors confirm that they, and/or their company or organization, hold copyright on all of the original material included in this paper. The authors also confirm that they have obtained permission, from the copyright holder of any third party material included in this paper, to publish it as part of their paper. The authors confirm that they give permission, or have obtained permission from the copyright holder of this paper, for the publication and distribution of this paper as part of the ICAS proceedings or as individual off-prints from the proceedings.

References

[1] Bons N P, He X, Mader C A, Martins J R. Multimodality in aerodynamic wing design optimization. *AIAA Journal*, Vol. 57, No. 3, pp 1004–1018, 2019.

[2] Poole D, Allen C, Rendall T. Global optimization of wing aerodynamic optimization case exhibiting multimodality. *Journal of Aircraft*, Vol. 55, No. 4, pp 1576–1591, 2018.

[3] Iuliano E, Quagliarella D, Donelli R S. Design of a supersonic natural laminar flow wing-body. *Journal of Aircraft*, Vol. 48, No. 4, pp 1147–1162, 2011.

[4] Yuan J, Sun J, Li L. Progress of supersonic aircraft laminar flow layout design and evaluation technologies. *ACTA AERONAUTICAET ASTRONAUTICA SINICA*, Vol. 43, No. 11, pp 526316 (in Chinese), 2022.

[5] Han Z-H, Chen J, Zhang K-S, Xu Z-M, Zhu Z, Song W-P. Aerodynamic shape optimization of natural-laminar-flow wing using surrogate-based approach. *AIAA Journal*, Vol. 56, No. 7, pp 2579–2593, 2018.

[6] Zhao R, Liu T, Wen C Y, et al. Theoretical modeling and optimization of porous coating for hypersonic laminar flow control. *AIAA Journal*, Vol. 56, No. 8, pp 2942–2946, 2018.

[7] Han Z-H, Xu C-Z, Qiao J-L, et al. Recent progress of efficient global aerodynamic shape optimization using surrogate-based approach. *Acta Aeronautica et Astronautica Sinica*, Vol. 41, No. 5, pp 623344, 2020 (in Chinese).

[8] Schöbi R, Sudret B, Wiart J. Polynomial-chaos-based Kriging. *International Journal for Uncertainty Quantification*, Vol. 5, No. 2, pp 171–193, 2015.

[9] Zhao H, Gong Z, Gan K, Gan Y, Xing H, Wang S. Supervised kernel principal component analysis-polynomial chaos-Kriging for high-dimensional surrogate modelling and optimization. *Knowledge-Based Systems*, Vol. 305, pp 112617, 2024.

[10] Peherstorfer B, Willcox K, Gunzburger M. Survey of multifidelity methods in uncertainty propagation, inference, and optimization. *SIAM Review*, Vol. 60, No. 3, pp 550–591, 2018.

[11] Fernández-Godino M G, Park C, Kim N H, Haftka R T. Issues in deciding whether to use multifidelity surrogates. *AIAA Journal*, Vol. 57, No. 5, pp 2039–2054, 2019.

[12] Schöbi R, Sudret B, Marelli S. Rare event estimation using polynomial-chaos-Kriging. *ASCE-ASME Journal of Risk and Uncertainty in Engineering Systems*, Vol. 3, No. 2, pp D4016002, 2017.

[13] Zhao H, Gao Z, Xu F, et al. Adaptive multi-fidelity sparse polynomial chaos-Kriging metamodeling for global approximation of aerodynamic data. *Structural and Multidisciplinary Optimization*, Vol. 64, No. 2, pp 829–858, 2021.

[14] Martins J R R A, Lambe A B. Multidisciplinary design optimization: a survey of architectures. *AIAA Journal*, Vol. 51, No. 9, pp 2049–2075, 2013.

[15] Wang C. Research on surrogate-based efficient aerodynamic optimization and many-objective problems. Ph.D. Dissertation, Aircraft Design Dept, Northwestern Polytechnical University, Xian, Dec, 2017.

[16] Zhao H, Gao Z, Xu F, Zhang Y, Huang J. An efficient adaptive forward–backward selection method for sparse polynomial chaos expansion. *Computer Methods in Applied Mechanics and Engineering*, Vol. 355, pp 456–491, 2019.

[17] Han Z-H, Görtz S, Zimmermann R. Improving variable-fidelity surrogate modeling via gradient-enhanced Kriging and a generalized hybrid bridge function. *Aerospace Science and Technology*, Vol. 25, No. 1, pp 177–189, 2013.

[18] Bouhlel M A, Martins J R R A. Gradient-enhanced Kriging for high-dimensional problems. *Engineering with Computers*, Vol. 35, No. 1, pp 157–173, 2019.

[19] Han Z-H, Zhang Y, Song C-X, et al. Weighted gradient-enhanced Kriging for high-dimensional surrogate modeling and design optimization. *AIAA Journal*, Vol. 55, No. 12, pp 4330–4346, 2017.

[20] Guo L, Narayan A, Zhou T. A gradient enhanced ℓ 1 -minimization for sparse approximation of polynomial chaos expansions. *Journal of Computational Physics*, Vol. 367, pp 49–64, 2018.

[21] Charles R C. Active subspaces: emerging ideas for dimension reduction in parameter studies. *Computing reviews*, Vol. 57, No. 8, pp 477–478, 2016.

[22] Li M, Wang R-Q, Jia G. Efficient dimension reduction and surrogate-based sensitivity analysis for expensive models with high-dimensional outputs. *Reliability Engineering & System Safety*, Vol. 195, No. 1, pp 106725, 2020.

[23] Constantine P G, Dow E, Wang Q. Active subspace methods in theory and practice: applications to Kriging surfaces. *SIAM Journal on Scientific Computing*, Vol. 36, No. 4, pp A1500–A1524, 2014.

[24] Zhou T, Peng Y. Kernel principal component analysis-based Gaussian process regression modelling for high-dimensional reliability analysis. *Computers & Structures*, Vol. 241, pp 106358, 2020.

[25] Lam R R, Zahm O, Marzouk Y M, et al. Multifidelity dimension reduction via active subspaces. *SIAM Journal on Scientific Computing*, Vol. 42, No. 2, pp A929–A956, 2020.

[26] Diez M, Campana E F, Stern F. Design-space dimensionality reduction in shape optimization by Karhunen–Lo`eve

expansion. *Comput. Meth. Appl. Mech. Eng*, Vol. 283, pp 1525–1544, 2015.

[27] Wold H. Partial least squares. *Encyclopedia of Statistical Sciences*, Vol. 6, pp 581-591, 1985.

[28] Helland I S. On the structure of partial least squares regression. *Communication in Statistics Simulation and Computation*, Vol. 17, No. 2, pp 581–607, 1988.

[29] Frank I E, Friedman J H. A statistical view of some chemometrics regression tools. *Technometrics*, Vol. 35, No. 2, pp 109–135, 1993.

[30] Alberto P R, González F G. Partial least squares regression on symmetric positive-definite matrices. *Revista Colombiana de Estadí-stica*, Vol. 36, No. 1, pp 177–192, 2012.