Multi-Fidelity Aerodynamic Optimization of a Low-Reynolds-Number Outlet-Guide-Vane Cascade

Proceedings of ASME Turbo Expo 2026
Turbomachinery Technical Conference and Exposition
GT2026
June 15–19, 2026, Milan, Italy

GT2026-176412

MULTI-FIDELITY AERODYNAMIC OPTIMIZATION OF A LOW-REYNOLDS-NUMBER OUTLET-GUIDE-VANE CASCADE

Shuai Li[1,*], Niklas Andersson[1]

[1]Chalmers University of Technology, Gothenburg, Sweden

ABSTRACT

This study presents an efficient aerodynamic optimization framework that integrates multi-fidelity surrogate modeling with advanced dimensionality reduction techniques. The study focuses on the shape optimization of a low-Reynolds-number Outlet-Guide-Vane (OGV) cascade, with the objective of minimizing pressure losses at the aerodynamic design point and two off-design conditions. Parameterization of the OGV cascade is enabled using DLR's BladeGen tool. To construct multi-fidelity kriging and neural network-based surrogate models, Reynolds-Averaged Navier-Stokes (RANS) Simulations and the MIT Inverse Solver for Euler Simulations (MISES) are respectively employed as high- and low-fidelity datasets. Optimization results demonstrate that both multi-fidelity kriging and neural network-based models yield Pareto fronts closely matching those generated by high-fidelity-only optimization, despite using just 22 high-fidelity simulations. The results also show that the multi-fidelity kriging model outperforms the multi-fidelity neural network model, likely due to the limited number of high-fidelity training samples, which necessitated the use of a small neural network architecture and, consequently, limited its ability to capture the underlying patterns in the data. Subsequent global sensitivity analysis (GSA) and Proper Orthogonal Decomposition (POD) indicate that the design space can be reduced from 15 to 5 dimensions. To further improve surrogate performance in high-dimensional design spaces, POD and Autoencoders are applied for dimensionality reduction (DR) of the blade design variable space. The results show that the Autoencoder achieves better dimensionality reduction than POD, with the reduced-space Pareto front more closely aligning with the full-space counterpart.

Keywords: Outlet Guide Vane, Aerodynamic Optimization, Kriging, Neural Network, Global Sensitivity Analysis

NOMENCLATURE

Abbreviations

ADP	Aerodynamic Design Point
AVDR	Axial Velocity Density Ratio
GENN	Gradient-Enhanced Neural Network
GSA	Global Sensitivity Analysis
HV	Hypervolume
IGD	Inverted Generational Distance
(K)PLS	(Kriging with) Partial Least Squares
KPLSK	Kriging combined with KPLS, namely, KPLS+K
MFKPLSK	Multi-Fidelity KPLSK
MISES	MIT Inverse Solver for Euler Simulations
MP	Measurement Plane
MSE	Mean Squared Error
NN	Neural Network
OGV	Outlet Guide Vane
OP	Operating Point
PCE	Polynomial Chaos Expansions
PF	Pareto Front
POD	Proper Orthogonal Decomposition
RANS	Reynolds-Averaged Navier-Stokes
SVD	Singular Value Decomposition

Roman letters

p	Static pressure [Pa]
p_t	Total pressure [Pa]
T	Static temperature [K]
T_t	Total temperature [K]
c	Chord length of the OGV blade [mm]
b	Pitch of the OGV blade [mm]
h	Height of the OGV blade [mm]
s	Root-mean-squared error of the kriging predictor [Pa]
X	Data matrix used for dimensionality reduction
z	Variables of reduced dimensions
$\tilde{X}$	Data matrix reconstructed from variables of reduced dimensions

Greek letters

β_s	Stagger angle [°]

*Corresponding author: shuai.li@chalmers.se
Documentation for `asmeconf.cls`: Version 1.45, March 16, 2026.

β_1	Inflow angle [°]
β_2	Outflow angle [°]
μ	Mean of a random variable
σ^2	Variance of a random variable

Dimensionless groups

ω	Pressure loss coefficient
Tu	Turbulence intensity
M	Mach number
Re	Reynolds number

Superscripts and subscripts

1,2	Values at MP1 and MP2
y	The vertical component of the Cartesian coordinate system

1. INTRODUCTION

The aviation sector remains a contributor to global greenhouse gas (GHG) emissions, with current aircraft engines producing substantial amounts of carbon dioxide [1–4]. As the industry moves toward environmentally friendly and climate-neutral aviation, continuous advances in engine technology are essential. Reducing the environmental footprint of air transport has therefore become a central objective for engine manufacturers worldwide. In line with the European Green Deal, which calls for ultra-efficient propulsion systems by 2035 and climate-neutral aviation by 2050 [5], there is a pressing need to improve every component that influences engine performance and efficiency. Among these components, the OGV plays a crucial aerodynamic role. Optimizing the OGV offers a promising route to improve engine efficiency, decrease fuel consumption, and consequently reduce GHG emissions [6]. In addition, the techniques developed for OGV optimization can be used to optimize other components of engines, contributing to the pursuit of sustainable aviation propulsion systems.

Aerodynamic shape optimization (ASO) provides a systematic means of improving the aerodynamic performance of engine components by exploring and refining their shape design parameters. Existing ASO strategies are commonly categorized as gradient-based or gradient-free approaches [7, 8]. Gradient-based methods (such as the adjoint method [9]) use the gradient of the objective function to efficiently guide the optimization process and often exhibit rapid convergence. However, their performance in nonconvex settings is sensitive to the initial guess and problem formulation, and convergence is generally restricted to local stationary points rather than global optima, which limits their robustness in complex optimization problems [10, 11]. In contrast, gradient-free methods (such as the genetic algorithm [12] and the particle swarm algorithm [13]) utilize stochastic or heuristic search strategies, providing enhanced robustness and global exploration capabilities, although with higher computational cost.

While gradient-free optimization methods are highly versatile, these methods often require a large number of function evaluations, which can be prohibitively expensive in computationally intensive applications. Surrogate models, as computationally inexpensive approximations of the true objective function, are commonly employed to reduce evaluation costs and guide the search toward promising regions. However, surrogate models built solely on high-fidelity data can still be costly, as each high-fidelity evaluation may demand substantial computational resources. Multi-fidelity surrogate models address this limitation by combining inexpensive low-fidelity data with a limited number of high-fidelity evaluations [14, 15]. By leveraging both sources of information, multi-fidelity approaches maintain sufficient accuracy while significantly reducing computational cost, enabling more efficient gradient-free optimization in complex and expensive problems.

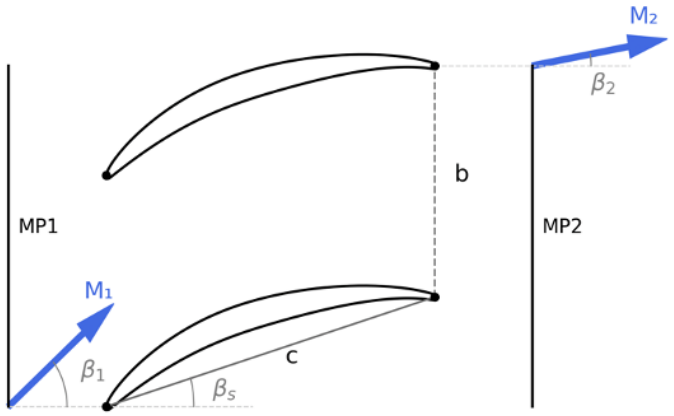


FIGURE 1: Schematic of the low–Reynolds-number OGV cascade showing the locations of Measurement Plane 1 (MP1) and Measurement Plane 2 (MP2).

High-dimensional design spaces pose a significant challenge in aerodynamic optimization [16], particularly when gradient-free methods and surrogate models are used. As the number of design variables increases, the required sampling effort grows rapidly. This is known as the curse of dimensionality [17, 18]. This not only increases computational cost but also degrades the accuracy and generalizability of surrogate models, which must approximate complex relationships in a large parameter space. Dimensionality-reduction techniques help alleviate these issues by identifying a lower-dimensional manifold [16, 19]. By reducing the effective dimensionality of the problem, these techniques improve surrogate-model fidelity, decrease sample requirements, and ultimately enable more efficient and robust optimization [19, 20].

This paper examines multi-fidelity surrogate models that integrate limited high-fidelity simulations with abundant low-fidelity simulations and evaluates dimensionality-reduction techniques for aerodynamic optimization of a low-Reynolds-number OGV cascade.

2. METHODS

The Non-dominated Sorting Genetic Algorithm II (NSGA-II) [21], implemented in *pymoo*, is used in the present study. NSGA-II is one of the most widely used evolutionary algorithms for solving multi-objective optimization problems. In the following, we introduce the database consisting of high-fidelity RANS simulations and low-fidelity MISES simulations, the kriging and neural network surrogate models, the multi-fidelity surrogate modeling strategy, and the dimensionality-reduction techniques. Here, RANS simulations are referred to as "high-fidelity" only

TABLE 1: OGV cascade design parameters.

Parameter	Value
Chord length c (mm)	70
Pitch to chord ratio b/c	0.577
Height to chord ratio h/c	2.4
Stagger angle β_s	16.04°
Inflow Mach number M_1	0.603
Inflow Reynolds number Re_1	150,000
Inflow angle β_1	43°
Inflow static pressure p_1 (Pa)	14408.0
Inflow total pressure $p_{t,1}$ (Pa)	18416.6
Inflow total temperature $T_{t,1}$ (K)	292.6
Turbulence intensity Tu	0.5%
Outflow static pressure p_2 (Pa)	16258.0
Outflow static temperature T_2 (K)	283.0
AVDR	1.03

in a relative sense, as they offer higher fidelity than MISES simulations. In the broader CFD context, however, the term "high-fidelity" typically denotes Direct Numerical Simulations (DNS) or Large-Eddy Simulations (LES).

2.1. The Database

The low–Reynolds-number OGV cascade is illustrated in Figure 1, with the design parameters listed in Table 1. The space of free variables is spanned by 15 design variables $x_i, i = 1, ..., 15$, provided by DLR's in-house tool BladeGen [22], which is used to generate the airfoil geometry. The geometric interpretations of the design variables are provided in Table 2. Specifically, variable x_1 controls the stagger angle, β_s, of the profile, as illustrated in Figure 2. Variables x_2 and x_3 define the leading-edge and trailing-edge angles, β_{LE} and β_{TE}, respectively (see Figure 2). Namely, $x_2 = \beta_{LE} + 90°$, and $x_3 = \beta_{TE} + 90°$. Variables x_4 and x_5 correspond to the leading-edge and trailing-edge wedge angles, $d\beta_{LE}$ and $d\beta_{TE}$. The radii of the leading- and trailing-edge circles, shown in red in Figure 2, are specified by x_6 and x_7, respectively. On the suction side, control point S_2 is defined by the distance x_8 between S_1 and S_2 along the tangent line $\overleftrightarrow{S_1S_2}$ of the leading-edge circle. Control point S_3 is specified by its Cartesian coordinates x_9 and x_{10}. Control point S_4 is determined by the distance x_{11} between S_4 and S_5 along the tangent line $\overleftrightarrow{S_4S_5}$ of the trailing-edge circle. Pressure-side control points are defined relative to the suction-side spline, typically by positive offsets. Control point P_2 is specified by the relative arc length along the suction side (x_{12}) and the normal distance from the suction side (x_{13}). Similarly, control points P_3 and P_4 depend on both the relative arc length and the distance to the suction side. However, for these points the arc length is fixed and only the distance is varied, resulting in design variables x_{14} for P_3 and x_{15} for P_4. For further details on these design variables, readers are referred to Ref. [22].

Geometries are randomly sampled from this 15-dimensional hypercube, with all sampled designs satisfying the geometric constraints. Namely, the points are randomly sampled to cover the design space of 15 dimensions. Using BladeGen [22], the 15

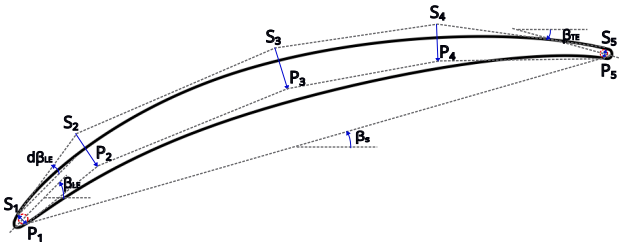


FIGURE 2: Illustration of the low–Reynolds-number OGV blade parameterization.

design variables are used to generate the blade geometries (see Figure 2), which are subsequently used in RANS and MISES simulations to achieve the pressure losses. The optimization problem aims to minimize the pressure losses at the ADP and off-design conditions, subject to constraints on the outflow angles at both the ADP and off-design conditions. Let $\mathbf{x}$ denote the vector of the 15 design variables. The optimization problem is formulated as follows:

$$\begin{aligned} \min_{\mathbf{x}} \quad & \big(f_1(\mathbf{x}),\, f_2(\mathbf{x})\big) \\ \text{s.t.} \quad & -0.2° \le \beta_{2,ADP}(\mathbf{x}) \le 1.4°, \\ & -1.4° \le \beta_{2,OP^-}(\mathbf{x}) \le 2.6°, \\ & -1.4° \le \beta_{2,OP^+}(\mathbf{x}) \le 2.6° \end{aligned} \tag{1}$$

where

$$f_1 := \omega_{\text{ADP}}, \tag{2}$$

$$f_2 := \frac{\omega_{\text{OP}^-} + \omega_{\text{OP}^+}}{2} \tag{3}$$

with ω_{ADP} denoting the loss coefficient at the design point, and ω_{OP^+} and ω_{OP^-} are the loss coefficients at off-design conditions OP^+ and OP^-, corresponding inlet-angle variations of ±5°. The definition of the loss coefficient ω is

$$\omega = \frac{\{\overline{p_{t,1}}\}_y - \{\overline{p_{t,2}}\}_y}{\{\overline{p_{t,1}}\}_y - \langle\overline{p_1}\rangle_y} \tag{4}$$

where $p_{t,1}$ and p_1 denote the total pressure and static pressure respectively in the measurement plane 20 mm upstream of the blade (i.e. MP1), and $p_{t,2}$ denotes the total pressure in the measurement plane 20 mm downstream of the blade (i.e. MP2). Here, the overbar denotes time-averaged quantities. The curly braces, $\{\cdot\}_y$, indicate mass-averaged values, while the angle brackets, $\langle\cdot\rangle_y$, denote surface-averaged values.

Table 3 summarizes the optimization constraints on the outflow angles at the ADP and off-design conditions. The outflow angle at the ADP, $\beta_{2,\text{ADP}}$, is bounded within a small admissible range, as reported in Table 3, $-0.2° \le \beta_{2,\text{ADP}} \le 1.4°$. The outflow angles at the OP^- and OP^+ are also bounded, but with a significantly wider range ($-1.4°$ to $2.6°$) than the ADP constraint. Although there are three constraints on the outflow angles, they act at different levels of strictness. $\beta_{2,\text{ADP}}$ is tightly constrained, but β_{2,OP^-} and β_{2,OP^+} have much broader bounds. Therefore, the problem is not over-constrained by the set of range of outflow angles β_{2,OP^-} and β_{2,OP^+}.

TABLE 2: Present OGV blade design variables.

Design variable	Minimum	Maximum	Interpretation
x_1	15.70	17.66	Stagger angle β_s (°)
x_2	120.42	125.68	Leading-edge angle $\beta_{LE} + 90°$ (°)
x_3	60.00	66.43	Trailing-edge angle $\beta_{TE} + 90°$ (°)
x_4	14.78	26.56	Leading-edge wedge angle $d\beta_{LE}$ (°)
x_5	0.55	4.35	Trailing-edge wedge angle $d\beta_{TE}$ (°)
x_6	5.00	6.81	Leading-edge radius r_{LE} (mm)
x_7	6.62	8.66	Trailing-edge radius r_{TE} (mm)
x_8	12.92	31.56	Distance between S_1 and S_2 (mm)
x_9	314.75	410.93	X-component of S_3 (mm)
x_{10}	191.53	205.08	Y-component of S_3 (mm)
x_{11}	818.78	856.70	Distance between S_4 and S_5 (mm)
x_{12}	136.03	227.50	Relative arc length along the suction side for P_2 (mm)
x_{13}	29.24	47.14	Normal distance of P_2 from the suction side (mm)
x_{14}	10.00	29.69	Normal distance of P_3 from the suction side (mm)
x_{15}	14.55	17.43	Normal distance of P_4 from the suction side (mm)

TABLE 3: Optimization constraints on the outflow angle.

Outflow angle (°)	Min.	Max.
$\beta_{2,ADP}$	-0.2	1.4
β_{2,OP^-}	-1.4	2.6
β_{2,OP^+}	-1.4	2.6

Two fidelity levels are employed in the CFD framework by using two different flow solvers. The high-fidelity RANS simulations are enabled by TRACE, a compressible RANS solver equipped with the Menter SST k-ω turbulence model [23] and γ-Re_θ transition model [24, 25]. The low-fidelity simulations are based on MISES, an inviscid blade-to-blade solver coupled with integral boundary-layer equations and an associated transition model [26, 27]. Both the RANS and MISES databases contain 3,918 designs. The objective functions take values in the ranges $0.032 \le f_1(\mathbf{x}) \le 0.054$ and $0.046 \le f_2(\mathbf{x}) \le 0.089$. The constraint variables lie within the following intervals: $-2.39° \le \beta_{2,ADP}(\mathbf{x}) \le 3.49°$, $-1.35° \le \beta_{2,OP^+}(\mathbf{x}) \le 5.92°$, and $-2.79° \le \beta_{2,OP^-}(\mathbf{x}) \le 3.27°$. The RANS database contains 1,516 feasible designs that satisfy the outflow-angle constraints in Table 3, while the MISES database includes 624 such designs. The database has also been used in a recent study [28]. Further details on the simulations can be found in Ref. [28].

2.2. Kriging Surrogate Model

In this study, we want to predict the two objective functions for any given design variables. Without loss of generality, assume that we wish to make a prediction at a given point $\mathbf{x}$. $Y(\mathbf{x})$ is a realization of a random variable that is normally distributed with mean μ and variance σ^2. Now consider two points $\mathbf{x}_i$ and $\mathbf{x}_j$. The correlation between the random variables is given by [29]

$$\mathrm{Corr}\big(Y(\mathbf{x}_i),\, Y(\mathbf{x}_j)\big) = \exp\left(-\sum_{\ell=1}^{d} \theta_\ell \, |x_{i\ell} - x_{j\ell}|^{p_\ell}\right), \qquad (5)$$

where θ_ℓ controls how quickly the correlation decreases as one moves along the ℓth coordinate direction. The parameter p_ℓ governs the smoothness of the function in that direction: values of p_ℓ close to 2 correspond to smoother functions, whereas values near 0 capture rough, non-differentiable behavior. For n points, the covariance function is given by [29]

$$\mathrm{cov}\big[Y(\mathbf{x}_i)\,,\, Y(\mathbf{x}_j)\big] = \sigma^2 R, \qquad (6)$$

where R is a $n \times n$ matrix with the (i, j) element given by Eq. (5). The values of μ, σ^2, θ_ℓ, and p_ℓ ($\ell = 1, \ldots, d$) are then estimated by maximizing the likelihood of the observed data. The kriging predictor is given by [29]

$$\hat{y}(\mathbf{x}^*) = \hat{\mu} + \mathbf{r}^T \mathbf{R}^{-1} (\mathbf{y} - \mathbf{I}\hat{\mu}), \qquad (7)$$

where $\mathbf{I}$ is a $n \times 1$ vector of ones, and $\mathbf{r}$ denotes the vector of correlations of $Y(\mathbf{x}^*)$ with $Y(\mathbf{x}_i)$, for $i = 1, \ldots, n$:

$$\mathbf{r} = \begin{pmatrix} \mathrm{Corr}(Y(\mathbf{x}^*),\, Y(\mathbf{x}_1)) \\ \vdots \\ \mathrm{Corr}(Y(\mathbf{x}^*),\, Y(\mathbf{x}_n)) \end{pmatrix}. \qquad (8)$$

The mean-squared error of the kriging predictor is [29]

$$s^2(\mathbf{x}^*) = \hat{\sigma}^2 \left[1 - \mathbf{r}^T \mathbf{R}^{-1} \mathbf{r} \; + \; \frac{\left(1 - \mathbf{r}^T \mathbf{R}^{-1} \mathbf{r}\right)^2}{\mathbf{I}^T \mathbf{R}^{-1} \mathbf{I}} \right]. \qquad (9)$$

For convenience, we often use the square root of the mean squared error, $s = \sqrt{s^2(\mathbf{x}^*)}$. This root mean squared error, also referred to as the "standard error", provides a natural measure of uncertainty in the predictions.

In this work, we use the KPLSK surrogate models implemented in the Surrogate Modeling Toolbox (SMT) library [30]. KPLSK is based on Kriging with Partial Least Squares (KPLS), a kriging variant that leverages the PLS method. By reducing the number of hyperparameters to estimate, KPLS offers faster training while maintaining good predictive accuracy.

2.3. Neural Network Surrogate Model

We focus on the Multilayer Perceptron (MLP), one of the earliest and most widely used neural network architectures. The MLP used in this study consists of an input layer with 15 components, an output layer with two components, and two hidden layers with six neurons each. The input vector, consisting of the 15 design variables, is successively transformed through linear operations followed by nonlinear activation functions within the hidden layers, ultimately producing the output vector—the objective functions. The fully connected neural network is implemented using a Gradient-Enhanced Neural Network (GENN) framework available in the SMT library [30], with the gradient-enhancement functionality disabled due to the absence of gradient information in the database. Consequently, the resulting GENN architecture is equivalent to a standard MLP.

In this study, the same network architecture is used across all datasets, consisting of two hidden layers with 16 neurons each. During training, a backtracking line search is employed to ensure stable optimization. The L2 regularization is applied to mitigate overfitting, and all inputs are normalized prior to training. The GENN surrogate was trained using full-batch Adam optimization for 1000 iterations.

2.4. Multi-Fidelity Surrogate Model

In multi-fidelity modeling, a common strategy is to couple low-fidelity predictions, which are typically inexpensive but approximate, with high-fidelity data, which are more accurate but much more costly to obtain. This fusion is often achieved through an autoregressive formulation of the form [14, 30]

$$y_{\text{high}}(\mathbf{x}) = \rho(\mathbf{x}) \cdot y_{\text{low}}(\mathbf{x}) + \delta(\mathbf{x}) \tag{10}$$

where $\rho(\mathbf{x})$ denotes a scaling factor that modulates the contribution of the low-fidelity model, and $\delta(\mathbf{x})$ represents a discrepancy function that captures systematic differences between the fidelities.

To more clearly illustrate the multi-fidelity method, we consider two analytical functions [30]:

$$f_{\text{LF}}(x) = \tfrac{1}{2}(6x-2)^2 \sin\big(2(6x-2)\big) + 10(x-0.5) - 5, \tag{11}$$

and

$$f_{\text{HF}}(x) = (6x-2)^2 \sin\big(2(6x-2)\big). \tag{12}$$

Six samples are drawn from the high-fidelity function $f_{\text{HF}}(x)$, while fourteen samples are taken from the low-fidelity function $f_{\text{LF}}(x)$. As shown in Figure 3, the high-fidelity samples form a subset of the low-fidelity samples. At identical values of x, a discrepancy $\delta(x)$ is observed between the low- and high-fidelity data. When $\rho(x)$ is assumed to be constant (e.g., $\rho = 1$), a surrogate model can be constructed for $\delta(x)$ along with a surrogate model for the low-fidelity data. The sum of these two surrogate models then yields the multi-fidelity surrogate constructed from both low- and high-fidelity samples. When $\rho(x)$ is not constant, it is treated as an unknown function and modeled jointly with $\delta(x)$ during the surrogate construction. In the example shown in Figure 3, the resulting multi-fidelity surrogate agrees closely with the high-fidelity function, demonstrating the effectiveness of the proposed multi-fidelity surrogate modeling approach.

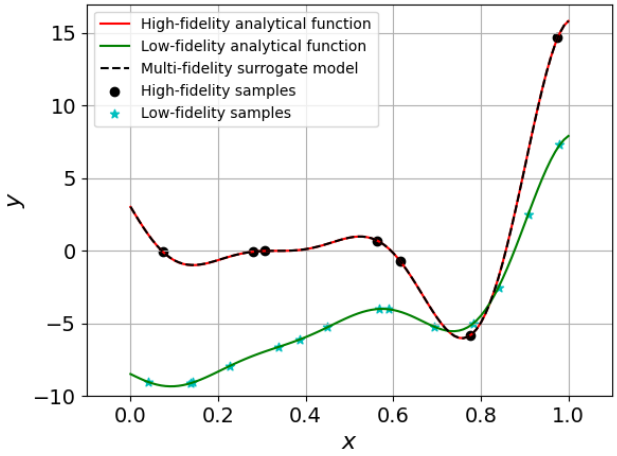


FIGURE 3: A one-dimensional example illustrating the construction of a multi-fidelity surrogate model [30].

Within the autoregressive kriging formulation of MFK-PLSK, a Gaussian process surrogate is first constructed for the low-fidelity response. The high-fidelity surrogate subsequently leverages this low-fidelity prediction as a conditioning input to model and correct the discrepancy. In the multi-fidelity NN surrogate model, the low-fidelity response and the discrepancy function $\delta(\mathbf{x})$ are both modeled using GENN and combined to construct the final multi-fidelity surrogate model.

In practice, various assumptions can be made about $\rho(\mathbf{x})$ depending on the complexity of the relationship between fidelities. A widely used simplification is to treat ρ as a constant. More flexible formulations allow $\rho(\mathbf{x})$ to vary over the design space, enabling the model to adapt to local differences between the low- and high-fidelity behaviors.

2.5. Dimensionality Reduction

Because high-dimensional design spaces greatly increase the complexity and sampling requirements of surrogate-based optimization, dimensionality-reduction techniques are desirable. While POD has been widely used in fluid dynamics since the 1980s and 1990s [31, 32], autoencoders emerged prominently only with the rise of deep learning in the 2010s [33]. The architecture of the autoencoder is depicted in Figure 4. The ability of autoencoders to learn nonlinear, task-specific latent spaces through neural networks represents a more modern approach. In this study, we evaluate and compare the effectiveness of POD and autoencoder-based methods for dimensionality reduction.

2.5.1. Proper Orthogonal Decomposition POD is a linear dimensionality-reduction procedure that begins by collecting the data matrix X and centering it to remove the mean. Singular value decomposition (SVD) is then performed to extract the dominant orthogonal modes of variation in the data. A reduced dimensionality k is selected so that the retained modes capture most of the total variance, typically 95–99%. The data are projected onto these modes to obtain reduced variables z, which provide a compact representation of the original system. Using z, the input

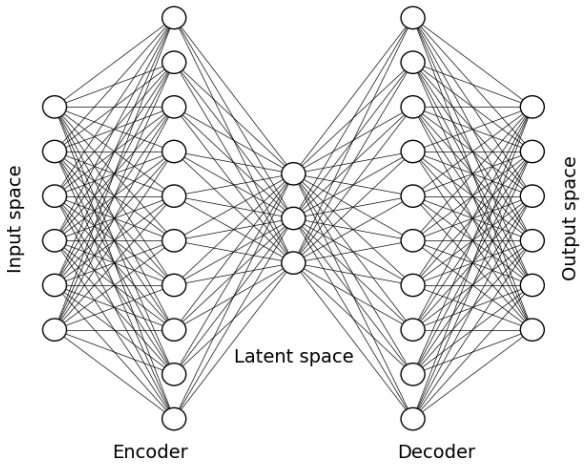


FIGURE 4: Neural-network architecture of the autoencoder, consisting of an encoder and a decoder.

space is reconstructed to form an approximation $\tilde{X}$, after which the function of interest is evaluated on this reconstruction. Finally, a surrogate model is built to relate the function evaluations $f(\tilde{X})$ to the reduced coordinates z, and this surrogate is used for efficient optimization.

2.5.2. Autoencoder Autoencoders provide a nonlinear alternative for dimensionality reduction, using a neural-network architecture composed of an encoder and a decoder. After collecting the data matrix X, the encoder maps the high-dimensional inputs into a lower-dimensional latent representation z. The decoder then reconstructs the input space from z, producing an approximation $\tilde{X}$ that captures the learned nonlinear structure of the data. The function of interest is evaluated on this reconstructed representation, similar to the POD approach. A surrogate model is subsequently constructed to relate the function values $f(\tilde{X})$ to the latent variables z, enabling optimization within the reduced nonlinear latent space.

An autoencoder with a single hidden layer in both the encoder and decoder can be described by the following equations. The encoding process is defined as [34]

$$z = \sigma\left(W_{\text{encoder}} X + b_{\text{encoder}}\right), \tag{13}$$

and the decoding process is given by [34]

$$\tilde{X} = \sigma\left(W_{\text{decoder}} z + b_{\text{decoder}}\right), \tag{14}$$

where σ is an activation function, and W and b are the weight matrices and bias vectors of the network, respectively. The encoder maps the input features into a compact representation, whereas the decoder attempts to reconstruct the original input from this representation. The objective of training the autoencoder is to minimize the reconstruction error, which quantifies the discrepancy between the input and its reconstruction. This reconstruction error is defined as [34]

$$L(X, \tilde{X}) = \left\| X - \tilde{X} \right\|_2^2, \tag{15}$$

where $\|\cdot\|_2$ is the Euclidean norm.

3. RESULTS AND DISCUSSION

3.1. RANS solutions of the baseline OGV configuration

The baseline OGV configuration is simulated using steady RANS equations with the SST k-ω turbulence model, both with and without the $\gamma - Re_\theta$ transition model, in the finite-volume solver Siemens STAR-CCM+ (version 2022.1). The compressible RANS equations are solved in a coupled manner. The numerical setup of the RANS simulations is summarized in Table 1. The residual convergence criterion for the RANS simulations is based on normalized residuals. Convergence is considered satisfied when the turbulent variables, i.e., turbulent kinetic energy (k) and specific dissipation rate (ω), decrease by six orders of magnitude (10^{-6}). In addition, the lift and drag histories are monitored and convergence is confirmed once these curves reach steady, flat behavior.

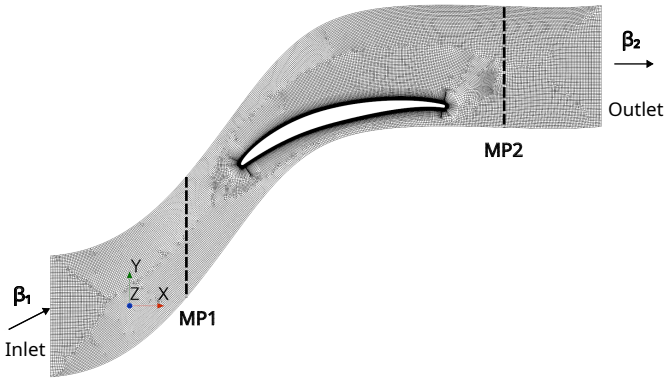


FIGURE 5: Computational domain and mesh for the baseline OGV configuration.

Figure 5 shows a two-dimensional S-shaped computational domain and the mesh for the baseline OGV configuration. The computational domain extends from $-0.06\ m$ to $0.13\ m$ in the x-direction, with the leading-edge point located at the origin of the coordinate system. The domain height is $0.04\ m$, corresponding to one pitch length. The unstructured mesh comprises a total of 50,287 cells and includes structured prism layers adjacent to the blade surface. The first near-wall cell height is chosen such that the dimensionless wall distance, y^+, remains below 1 over the blade surface.

During operation, the flow over a low-Reynolds-number OGV blade often leads to the formation of a laminar separation bubble and the transition from laminar to turbulent boundary layers [35]. Figure 6 presents the streamwise (X-direction) velocity component obtained from the RANS simulations performed with and without the $\gamma - Re_\theta$ transition model. The results show that the flow accelerates over the forward portion of the suction-side surface due to a favorable pressure gradient. While the RANS simulation without the transition model fails to predict the laminar separation bubble, the simulation including the transition model captures this feature. These results highlight the importance of incorporating transition modeling in RANS simulations of low-Reynolds-number OGV flows.

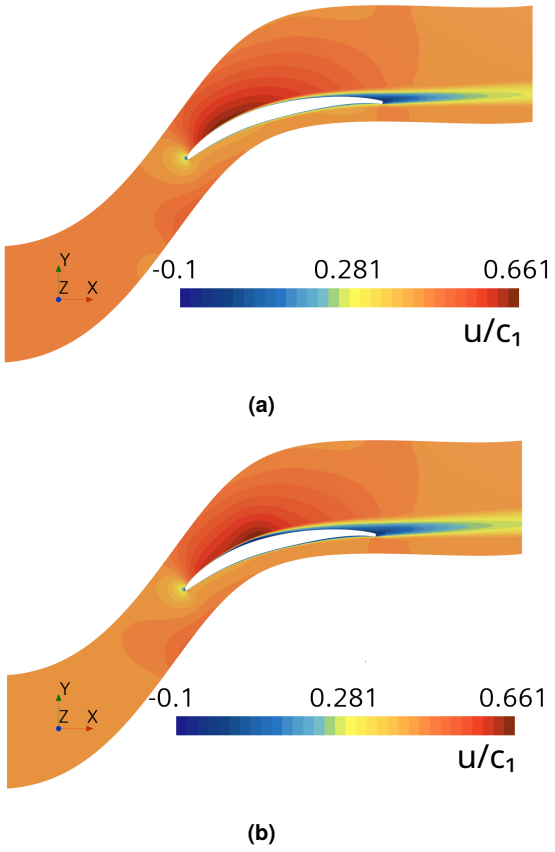


FIGURE 6: X-direction velocity component normalized by the sound speed at MP1: (a) RANS without the transition model, and (b) RANS with the transition model.

To further evaluate the predictive capability of the RANS simulations with and without the transition model, Figure 7 compares the isentropic Mach number distributions obtained from the experiments and both RANS simulations. The isentropic Mach number is defined as follows:

$$M_{is} = \sqrt{\frac{2}{\gamma - 1}\left[\left(\frac{p}{p_{t,1}}\right)^{-\frac{\gamma-1}{\gamma}} - 1\right]}, \tag{16}$$

where p is the static pressure and $\gamma = 1.4$ is the specific heat ratio. Results in Figure 7 show that, on the pressure side, both RANS simulations show very good agreement with the experimental data. On the suction side, where a laminar separation bubble and laminar-to-turbulent transition occur, the RANS simulation including the transition model provides better agreement with the measurements than the simulation without transition modeling. This further confirms the importance of incorporating a transition model in RANS simulations of low-Reynolds-number OGV flows.

3.2. Comparison Between Single- and Multi-Fidelity Surrogate-Based Optimization

The present surrogate models require nested sampling, that is, the high-fidelity samples must form a subset of the low-fidelity samples. From the 440 feasible and nested designs, we select 22

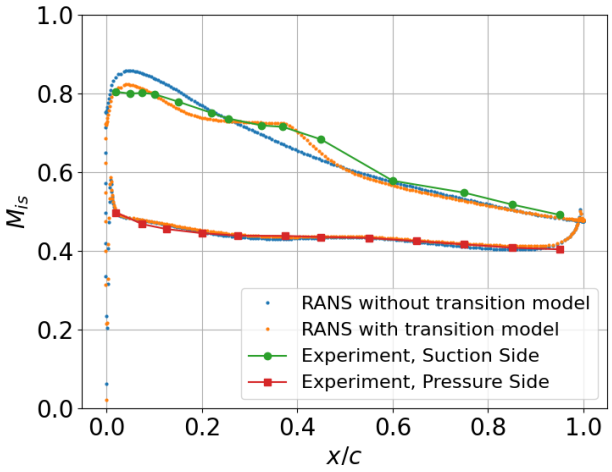


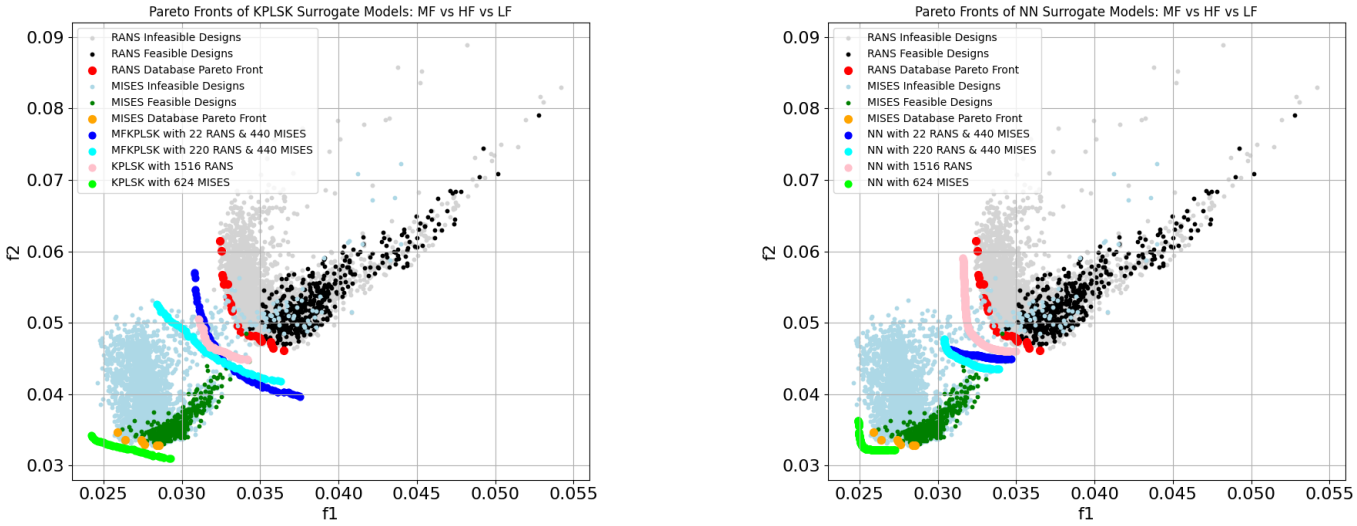
FIGURE 7: A comparison of isentropic Mach number distributions obtained from experiments and RANS simulations, with and without the transition model.

(5%) and 220 (50%) samples as high-fidelity data, together with the 440 low-fidelity samples, to construct the multi-fidelity surrogate models. The surrogate models were constructed using either KPLSK or GENN from the SMT library [30]. Since gradient information is not available in the database, the gradient-enhancement coefficient of GENN was set to zero, reducing GENN to a standard MLP.

Figure 8 compares the surrogate-based optimizations using multi-, high-, and low-fidelity models, constructed from feasible designs that satisfy the outflow-angle constraints and built with the KPLSK and NN surrogate models, respectively. The low-fidelity MISES optimization yields Pareto fronts with lower predicted pressure losses compared to the high-fidelity RANS optimization, which is expected given that MISES is based on inviscid Euler simulations. It is also observed that both KPLSK and NN surrogate models produce Pareto fronts with reduced (i.e., better optimized) pressure loss compared to the database Pareto fronts, as expected. The multi-fidelity KPLSK and NN models yield Pareto fronts close to those obtained from high-fidelity (RANS-only) optimization, demonstrating that a small number of high-fidelity samples (i.e., 22) can be sufficient for optimization. This approach reduces the computational cost of the optimization by requiring fewer expensive high-fidelity evaluations. However, the multi-fidelity NN performs worse than the multi-fidelity KPLSK, likely because the limited number of high-fidelity training samples necessitates a small neural network architecture, which cannot fully capture the underlying patterns in the data.

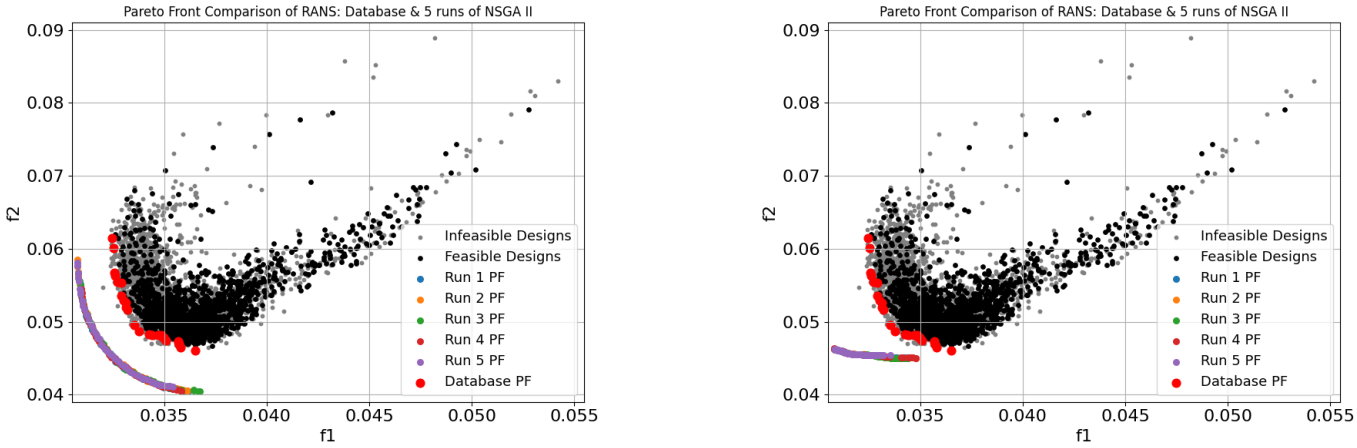
Figure 9 compares multiple runs of the NSGA-II optimization using multi-fidelity KPLSK and NN surrogate models constructed from 22 RANS and 440 MISES simulations. Despite the stochastic nature of the NSGA-II algorithm, the resulting Pareto fronts from the five independent runs show strong agreement, demonstrating the robustness and reliability of the reported Pareto fronts.

To evaluate the performance of the multi-fidelity surrogate-



(a) (b)

FIGURE 8: Comparison of Pareto fronts in the optimizations based on multi-, high-, and low-fidelity KPLSK and NN surrogate models: (a) KPLSK, and (b) NN.



(a) (b)

FIGURE 9: Comparison of different runs of the NSGA II optimizations based on multi-fidelity KPLSK and NN surrogate models built with 22 RANS and 440 MISES simulations: (a) Multi-fidelity KPLSK, and (b) multi-fidelity NN.

based optimizations, we compute the Inverted Generational Distance (IGD) [36] and the hypervolume (HV) of the Pareto fronts. The IGD quantifies the distance from each point on the high-fidelity Pareto front to its nearest point on the corresponding multi-fidelity Pareto front. The HV measures the area (or volume) dominated by a given set of solutions with respect to a reference point. For all four multi-fidelity surrogate models reported in Table 4, the same reference point is used, namely $(f_1, f_2) = (0.0413, 0.0628)$. Lower IGD values and higher hypervolume values indicate better optimization performance. As shown in Table 4, the multi-fidelity KPLSK surrogate models outperform the multi-fidelity NN models. Moreover, increasing the number of high-fidelity samples further improves optimization performance.

Figure 10 presents a comparison between the baseline OGV

TABLE 4: Optimization performance indicators IGD and Hypervolume for the multi-fidelity surrogate models.

Surrogate	(N_{HF}, N_{LF})	IGD	Hypervolume
KPLSK	(22, 440)	5.78×10^{-4}	2.16×10^{-4}
KPLSK	(220, 440)	9.39×10^{-4}	2.39×10^{-4}
NN	(22, 440)	4.57×10^{-3}	1.86×10^{-4}
NN	(220, 440)	4.29×10^{-3}	2.06×10^{-4}

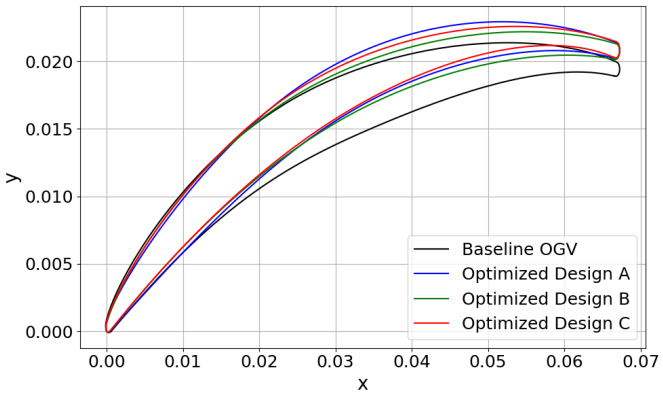


FIGURE 10: Comparison between the baseline OGV and three optimized designs. Designs A, B, and C correspond to the minimum f_1, minimum f_2, and 50% ranking of f_1, respectively, on the Pareto front of the MFKPLSK surrogate model built with 22 RANS and 440 MISES simulations.

and three optimized designs obtained from the Pareto front of the MFKPLSK surrogate model built with 22 RANS and 440 MISES simulations. Designs A, B, and C correspond to the minimum loss coefficient at the ADP, the minimum loss coefficient under off-design conditions, and the design ranked 50% in terms of the loss coefficient at the ADP, respectively. All three optimized designs exhibit a rearward shift in the profile, resulting from the larger stagger angles achieved through optimization. In addition, the optimized profiles are thinner, particularly for designs B and C. In contrast, the optimized designs closely match the baseline OGV at the leading edge and in the forward portion of the profile.

3.3. Global Sensitivity Analysis and Dimensionality Reduction

To identify the key design variables, Figure 11 presents the Sobol indices computed using the Gram–Schmidt Polynomial Chaos Expansions (PCE) [37, 38]. Two types of indices are shown: first-order Sobol indices, which measure the contribution of each input variable individually to the output variance, and total-order Sobol indices, which account for both the individual and interaction effects of each variable. Analysis of the first-order Sobol indices for both objective functions highlights five important design variables: x_1, x_4, x_6, x_8, and x_9. They correspond to the re-staggering (stagger angle), leading-edge wedge angle, leading-edge radius, and suction-side control points S_2 and S_3, respectively, as shown in Table 2 and Figure 2. These key design variables suggest that the leading-edge and suction-side parameters play a more important role in affecting the pressure losses at both the ADP and off-design conditions. Based on this global sensitivity analysis, the design space is reduced from 15 to 5 dimensions.

Figure 12 presents the explained variance of the POD modes computed from the high-fidelity database. The first five modes capture almost all of the total energy, also confirming the validity of reducing the dimensions from 15 to 5.

It is important to note that Sobol screening and POD should not be interpreted as mutually reinforcing methods. Sobol screening identifies the input variables and their interactions that contribute most significantly to the variance of the objective functions, whereas POD extracts dominant directions in the input space based solely on input variance, without consideration of the objectives. Because high variability in the input space does not necessarily correspond to high variability in the outputs, and vice versa, these two methods cannot be used to validate or reinforce each other.

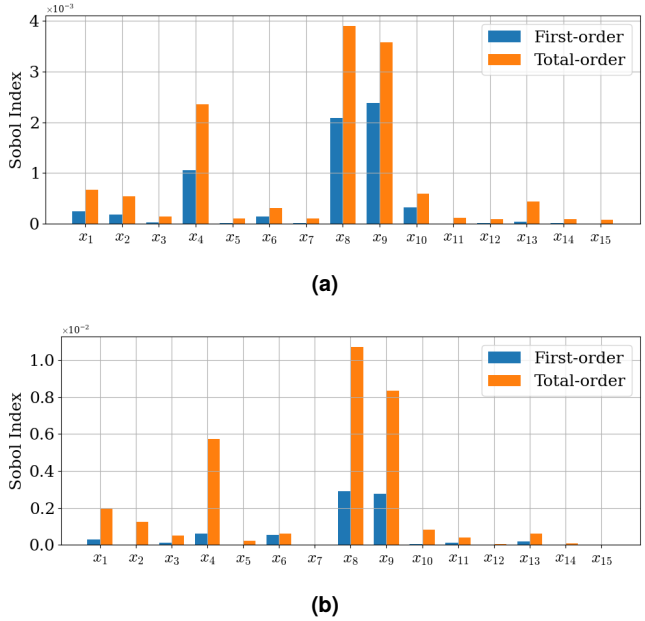


FIGURE 11: Sobol indices obtained from the global sensitivity analysis for the objective functions: (a) f_1 and (b) f_2.

To assess the effectiveness of POD and autoencoders for dimensionality reduction in the aerodynamic optimization of the OGV blade, we applied both techniques within the surrogate-based optimization framework. Figures 13 and 14 compare POD and autoencoder–based dimensionality reduction for the high-fidelity KPLSK and NN surrogate optimizations, respectively. The results show that the autoencoder achieves better dimensionality reduction: its reduced-space Pareto front aligns more closely with the full-space Pareto front than that obtained using POD.

Table 5 reports the reconstruction mean squared error (MSE)

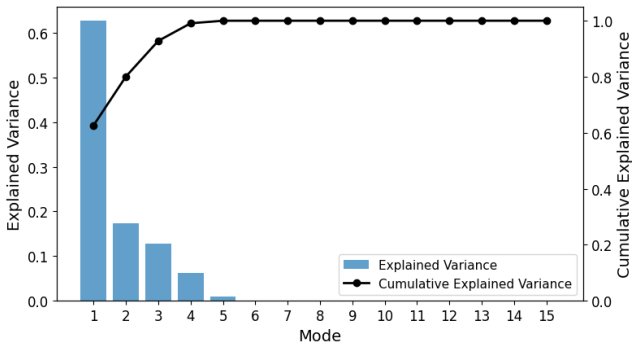


FIGURE 12: Explained variance of the POD modes computed from the high-fidelity database.

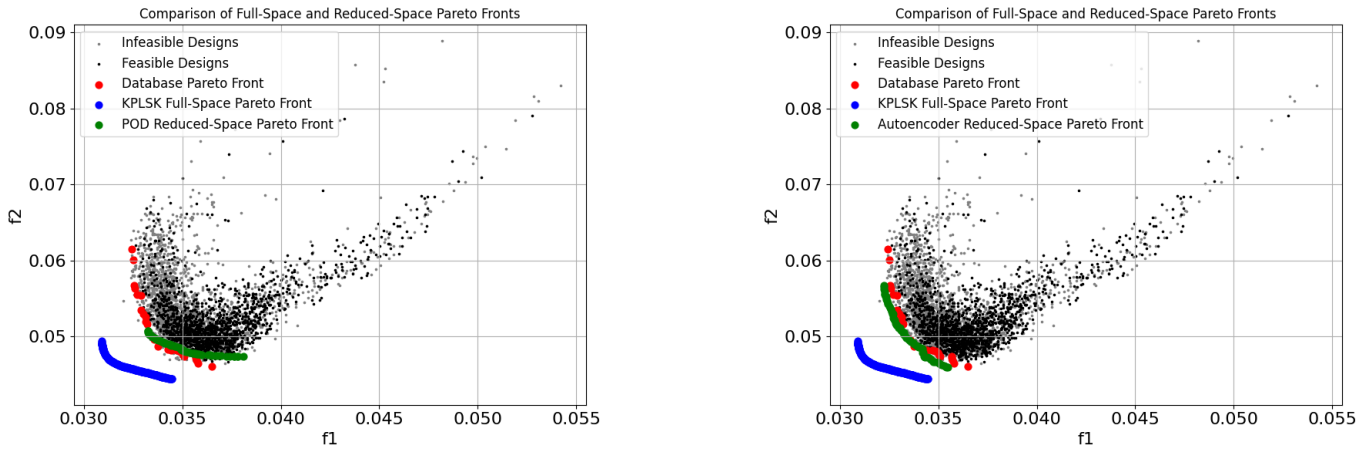


FIGURE 13: Comparison of dimensionality reduction using (a) POD and (b) Autoencoder in KPLSK surrogate-based optimization using 1,516 RANS samples.

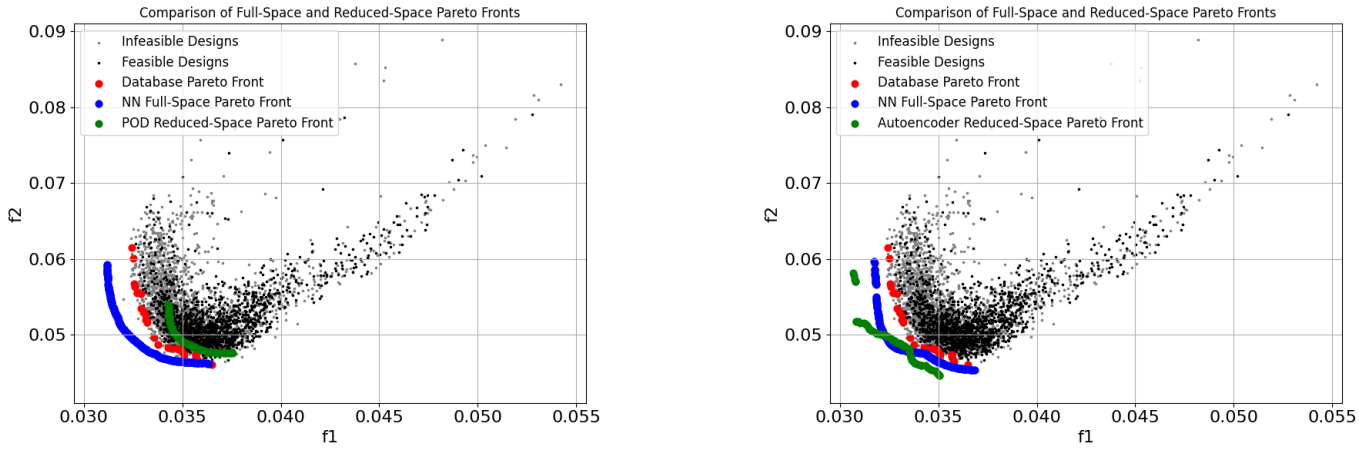


FIGURE 14: Comparison of dimensionality reduction using (a) POD and (b) Autoencoder in NN surrogate-based optimization using 1,516 RANS samples.

of the POD and autoencoder used for dimensionality reduction in high-fidelity KPLSK- and NN-based surrogate optimizations. Given that both f_1 and f_2 are on the order of 10^{-2}, the reconstruction MSEs for both POD and the autoencoder are reasonably low, indicating that the dimensionality-reduction techniques perform well in the KPLSK- and NN-surrogate-based optimizations. Since KPLSK already incorporates a dimensionality-reduction step to facilitate initial hyperparameter estimation, we further investigate whether the use of POD or autoencoders becomes redundant when applied with KPLSK. To this end, we also perform surrogate-based optimizations using standard kriging (KRG), as summarized in Table 5. The results show that both KPLSK and KRG surrogates yield identical reconstruction MSE values when combined with either POD or autoencoder dimensionality-reduction techniques, indicating that the use of POD or autoencoders with KPLSK is not redundant. Furthermore, the Pareto fronts obtained from the standard kriging-based optimizations are identical to those generated using KPLSK-based optimizations (see Figure 13) and are therefore not shown here.

TABLE 5: Reconstruction MSE of the POD and autoencoder used for dimensionality reduction in high-fidelity KPLSK, NN, and standard kriging surrogate-based optimizations.

Surrogate	DR technique	MSE of f_1	MSE of f_2
KPLSK	POD	7.62×10^{-6}	3.05×10^{-5}
KPLSK	Autoencoder	5.69×10^{-6}	2.22×10^{-5}
NN	POD	7.18×10^{-6}	2.98×10^{-5}
NN	Autoencoder	5.73×10^{-6}	2.13×10^{-5}
KRG	POD	7.62×10^{-6}	3.05×10^{-5}
KRG	Autoencoder	5.69×10^{-6}	2.22×10^{-5}

4. CONCLUSION

This study presents an efficient aerodynamic optimization framework that integrates multi-fidelity surrogate modeling with advanced dimensionality reduction techniques. The optimization is applied to a low-Reynolds-number OGV cascade, with the objective of minimizing pressure losses at the aerodynamic design point and two off-design conditions, within the context of the ongoing Horizon Europe Sci-Fi-Turbo project. Parameterization of the OGV cascade is enabled using the BladeGen tool developed at the DLR Institute of Propulsion Technology [22]. To construct multi-fidelity kriging and neural network-based surrogate models, RANS and MISES simulations are respectively employed as high- and low-fidelity datasets. RANS simulations of the baseline configuration show that, although the simulation without the transition model fails to predict the laminar separation bubble on the suction side, the simulation with the transition model captures this feature, highlighting the importance of incorporating transition modeling in RANS simulations of low-Reynolds-number OGV flows. Regarding the optimization results, the IGD and hypervolume performance indicators show that both the multi-fidelity KPLSK- and neural-network-based optimizations produce high-quality Pareto fronts, with the multi-fidelity KPLSK approach performing better. Beside, optimization results demonstrate that both multi-fidelity KPLSK and neural network-based models yield Pareto fronts closely matching those generated by high-fidelity-only optimization, despite using just 22 high-fidelity simulations. This highlights the potential of multi-fidelity surrogate modeling to significantly reduce computational cost while maintaining accuracy. The results also show that the multi-fidelity KPLSK model outperforms the multi-fidelity neural network model, likely due to the limited number of high-fidelity training samples, which necessitated the use of a small neural network architecture and, consequently, limited its ability to capture the underlying patterns in the data. An examination of the blade profiles of three designs optimized using multi-fidelity KPLSK surrogate-based optimization shows that the profiles closely follow the baseline OGV near the leading edge and in the forward section, while exhibiting a rearward shift and reduced thickness in the aft portion to reduce pressure losses. Subsequent analysis of the first-order Sobol indices for both objective functions highlights five important design variables: the re-staggering (stagger angle), leading-edge wedge angle, leading-edge radius, and two suction-side control points. A POD analysis of the high-fidelity database indicates that the design space can be reduced from 15 to 5 dimensions. To further improve surrogate performance in high-dimensional design spaces, POD and Autoencoders are applied for dimensionality reduction of the blade design variable space. The reasonably low reconstruction MSEs for both POD and the autoencoder indicate that both POD and the autoencoder perform well in the KPLSK- and NN-surrogate-based optimizations. The results show that the Autoencoder achieves better dimensionality reduction than POD, with the reduced-space Pareto front more closely aligning with the full-space counterpart. Overall, this study demonstrates a robust and computationally efficient framework for the multi-fidelity aerodynamic optimization of turbomachinery components, supporting the development of next-generation, low-emission aero engines.

ACKNOWLEDGMENTS

This research was funded by the Horizon Europe project Scale-resolving Simulations for Innovations in Turbomachinery Design (Sci-Fi-Turbo) under Grant Agreement No. 101138080. The authors thank the project coordinator, DLR Institute of Propulsion Technology, for providing the database. The authors also thank Sorbonne University and Cenaero in the Sci-Fi-Turbo consortium, who contributed to helpful discussions on the multi-fidelity surrogate models during the project meetings.

REFERENCES

[1] Grewe, V., Gangoli Rao, A., Grönstedt, T., Xisto, C., Linke, F., Melkert, J., Middel, J., Ohlenforst, B., Blakey, S., Christie, S. et al. "Evaluating the climate impact of aviation emission scenarios towards the Paris agreement including COVID-19 effects." *Nature communications* Vol. 12 No. 1 (2021): p. 3841.

[2] Boucher, O., Borella, A., Gasser, T. and Hauglustaine, D. "On the contribution of global aviation to the CO2 radiative forcing of climate." *Atmospheric Environment* Vol. 267 (2021): p. 118762.

[3] Teoh, R., Engberg, Z., Shapiro, M., Dray, L. and Stettler, M. E. J. "The high-resolution Global Aviation emissions Inventory based on ADS-B (GAIA) for 2019–2021." *Atmospheric Chemistry and Physics* Vol. 24 No. 1 (2024): pp. 725–744.

[4] Lee, D. S., Fahey, D. W., Skowron, A., Allen, M. R., Burkhardt, U., Chen, Q., Doherty, S. J., Freeman, S., Forster, P. M., Fuglestvedt, Jan. et al. "The contribution of global aviation to anthropogenic climate forcing for 2000 to 2018." *Atmospheric environment* Vol. 244 (2021): p. 117834.

[5] Schafer, A. W. and Barrett, R. H. "Pathways towards 90% decarbonization of aviation by 2050." *Nat. Clim. Chang* Vol. 12 (2022): pp. 895–896.

[6] Clemen, C. "Aero-mechanical optimisation of a structural fan outlet guide vane." *Structural and multidisciplinary optimization* Vol. 44 No. 1 (2011): pp. 125–136.

[7] Li, J., Du, X. and Martins, J. R. R. A. "Machine learning in aerodynamic shape optimization." *Progress in Aerospace Sciences* Vol. 134 (2022): p. 100849.

[8] Abergo, L., Morelli, M. and Guardone, A. "Aerodynamic shape optimization based on discrete adjoint and RBF." *Journal of Computational Physics* Vol. 477 (2023): p. 111951.

[9] Jameson, A. "Aerodynamic design via control theory." *Journal of scientific computing* Vol. 3 No. 3 (1988): pp. 233–260.

[10] Nocedal, J. "Numerical optimization." *Springer Ser. Oper. Res. Financ. Eng./Springer* (2006).

[11] Bonnans, J. F., Gilbert, J. C., Lemaréchal, C. and Sagastizábal, C. A. *Numerical optimization: theoretical and practical aspects*. Springer (2006).

[12] Holland, J. H. "Genetic algorithms and the optimal allocation of trials." *SIAM journal on computing* Vol. 2 No. 2 (1973): pp. 88–105.

[13] Kennedy, J. and Eberhart, R. "Particle swarm optimization." *Proceedings of ICNN'95-international conference on neural networks*, Vol. 4: pp. 1942–1948. 1995. ieee.

[14] Fernández-Godino, M. G. "Review of multi-fidelity models." *arXiv preprint arXiv:1609.07196* (2016).

[15] Schouler, M., Belme, A. and Cinnella, P. "Comparison of multi-fidelity surrogate models for multi-objective aerodynamic optimization in turbomachinery under extreme cost imbalance." *Advanced Modeling and Simulation in Engineering Sciences* Vol. 12 No. 1 (2025): p. 35.

[16] Pang, B., Zhang, Y., Li, J., Wang, X., Chang, M. and Bai, J. "Data-driven surrogate model for aerodynamic design using separable shape tensor method." *Chinese Journal of Aeronautics* Vol. 37 No. 9 (2024): pp. 41–58.

[17] He, Y. and Luo, J. "Efficient hierarchical kriging modeling method for high-dimension multi-fidelity problems." *Chinese Journal of Mechanical Engineering* Vol. 37 No. 1 (2024): p. 151.

[18] Saves, P. "High-dimensional multidisciplinary design optimization for aircraft eco-design." Ph.D. Thesis, ISAE-SUPAERO. 2024.

[19] Hou, C. K. J. and Behdinan, K. "Dimensionality reduction in surrogate modeling: A review of combined methods." *Data science and engineering* Vol. 7 No. 4 (2022): pp. 402–427.

[20] Ji, B., Huang, J., Lu, X., Wu, Y. and Liu, J. "An improved approach for reducing the dimensionality of wing aerodynamic optimization considering longitudinal stability." *Aerospace* Vol. 11 No. 1 (2024): p. 80.

[21] Deb, K., Pratap, A., Agarwal, S. and Meyarivan, T. "A fast and elitist multiobjective genetic algorithm: NSGA-II." *IEEE transactions on evolutionary computation* Vol. 6 No. 2 (2002): pp. 182–197.

[22] Voß, C., Siggel, M., Backhaus, J., Goinis, G. and Pahs, A. "A differentiated geometry blade parameterization methodology for gas turbines." *Computers & Fluids* Vol. 292 (2025): p. 106588.

[23] Menter, F. R. "Two-equation eddy-viscosity turbulence models for engineering applications." *AIAA journal* Vol. 32 No. 8 (1994): pp. 1598–1605.

[24] Langtry, R. B. "A correlation-based transition model using local variables for unstructured parallelized CFD codes." *Stuttgart: University of Stuttgart* (2006).

[25] Langtry, R. B. and Menter, F. R. "Correlation-based transition modeling for unstructured parallelized computational fluid dynamics codes." *AIAA journal* Vol. 47 No. 12 (2009): pp. 2894–2906.

[26] Drela, M. and Giles, M. B. "Viscous-inviscid analysis of transonic and low Reynolds number airfoils." *AIAA journal* Vol. 25 No. 10 (1987): pp. 1347–1355.

[27] Youngren, H. and Drela, M. "Viscous/inviscid method for preliminary design of transonic cascades." *27th Joint Propulsion Conference*: p. 2364. 1991.

[28] Goinis, G., Satcunanathan, S. and Bergmann, M. "Examining the Potential of High-Order SRS to Support RANS-Based Compressor Airfoil Optimization." *Turbo Expo*, Vol. 88872: p. V011T32A025. 2025. American Society of Mechanical Engineers.

[29] Jones, D. R. "A taxonomy of global optimization methods based on response surfaces." *Journal of global optimization* Vol. 21 No. 4 (2001): pp. 345–383.

[30] Saves, P., Lafage, R., Bartoli, N., Diouane, Y., Bussemaker, J., Lefebvre, T., Hwang, J. T., Morlier, J. and Martins, J. R. R. A. "SMT 2.0: A Surrogate Modeling Toolbox with a focus on hierarchical and mixed variables Gaussian processes." *Advances in Engineering Software* Vol. 188 (2024): p. 103571.

[31] Sirovich, L. "Turbulence and the dynamics of coherent structures. I. Coherent structures." *Quarterly of applied mathematics* Vol. 45 No. 3 (1987): pp. 561–571.

[32] Holmes, P. *Turbulence, coherent structures, dynamical systems and symmetry*. Cambridge university press (2012).

[33] Vincent, P., Larochelle, H., Lajoie, I., Bengio, Y., Manzagol, P.-A. and Bottou, L. "Stacked denoising autoencoders: Learning useful representations in a deep network with a local denoising criterion." *Journal of machine learning research* Vol. 11 No. 12 (2010).

[34] Alaghbari, K. A., Lim, H.-S., Saad, M. H. M. and Yong, Y. S. "Deep autoencoder-based integrated model for anomaly detection and efficient feature extraction in iot networks." *IoT* Vol. 4 No. 3 (2023): pp. 345–365.

[35] Bergmann, M., Morsbach, C., Möller, F. M., Klose, B. F., Hergt, A. and Goinis, G. "A Comparative Study of Varying Incidence Angle Effects on a Low-Reynolds-Number Compressor Cascade Based on Experiments and Low-Fidelity and High-Fidelity Numerical Simulations." *International Journal of Turbomachinery, Propulsion and Power* Vol. 10 No. 4 (2025): p. 42.

[36] Coello Coello, C. A. and Reyes Sierra, M. "A study of the parallelization of a coevolutionary multi-objective evolutionary algorithm." *Mexican international conference on artificial intelligence*: pp. 688–697. 2004. Springer.

[37] Xiu, D. and Karniadakis, G. E. "The Wiener–Askey polynomial chaos for stochastic differential equations." *SIAM journal on scientific computing* Vol. 24 No. 2 (2002): pp. 619–644.

[38] Salehi, S., Raisee, M., Cervantes, M. J. and Nourbakhsh, A. "Efficient uncertainty quantification of stochastic CFD problems using sparse polynomial chaos and compressed sensing." *Computers & fluids* Vol. 154 (2017): pp. 296–321.