

FEDERAL UNIVERSITY OF ABC  
GRADUATE COURSE IN

Allan Moreira de Carvalho

REDUCED ORDER MODELS FOR DATA-DRIVEN FLOW FIELD FLOW  
RECONSTRUCTION USING MACHINE LEARNING

Santo André, SP – Brazil  
July 2025



Allan Moreira de Carvalho

REDUCED ORDER MODELS FOR DATA-DRIVEN FLOW FIELD FLOW  
RECONSTRUCTION USING MACHINE LEARNING

presented to the Graduate course of Federal University of ABC,  
as a partial requirement for obtaining a Master's degree in .

Advisor: Prof. Dr. Daniel Jonas Dezan  
Coadvisor: Prof. Dr. Wallace Gusmão Ferreira

Santo André, SP – Brazil  
July 2025

Allan Moreira de Carvalho,

Reduced Order Models for Data-Driven Flow Field Flow Reconstruction  
using Machine Learning/ Allan Moreira de Carvalho. – 2025.

?? p.

Advisor: Prof. Dr. Daniel Jonas Dezan

Coadvisor: Prof. Dr. Wallace Gusmão Ferreira

() – Federal University of ABC, Graduate Course in , Santo André, 2025.

1. K1. 2. K2. 3. K3. 4. K4. 5. K5. I. Jonas Dezan, Daniel II.  
Gusmão Ferreira, Wallace III. Graduate Course in , 2025. IV. Title

This exemplar was revised and altered from the original version according to the observations done by the defense board at the defense day, under responsibility of author and advisor's consent.

Santo André, \_\_\_\_\_ , 20 \_\_\_\_\_

Author's signature: \_\_\_\_\_

Advisor's signature: \_\_\_\_\_



---

SIGN SHEET

---

Signatures of examining board members, which examined and approved the defense of candidate Allan Moreira de Carvalho in July 23<sup>rd</sup>, 2025:

---

Profa. Dra. Marcia Maria Penteado Marchesini - Examinadora Externa

---

Prof. Dr. Wilson Carlos da Silva Júnior - Examinador Externo

---

Profa. Dra. Franciane Freitas Silveira - Examinadora Interna Suplente

---

Prof. Dr. Rômulo Gonçalves Lins - Presidente





Dedication....



## **Acknowledgement**

Obrigado a todos



## Epigraph

*“Learn from the mistakes of those who followed your advice”  
(Unknown author)*



## Abstract

The aerothermodynamic design of high-performance propulsion systems, such as rocket nozzles, presents a critical computational challenge. High-fidelity (HF) simulations, typically based on Reynolds-Averaged Navier-Stokes (RANS) equations coupled with conjugate heat transfer (CHT) models, provide the necessary accuracy to predict wall heat flux and temperature but are prohibitively expensive for many-query applications like design optimization and uncertainty quantification (UQ). Conversely, low-fidelity (LF) models, such as quasi-one-dimensional Euler solvers, are computationally cheap but lack the physical fidelity to predict critical viscous and thermal phenomena. This dissertation addresses this dilemma by developing and validating a data-driven, multi-fidelity framework for the rapid and accurate reconstruction of two-dimensional compressible flow fields in a water-cooled, convergent-divergent nozzle. The methodology leverages a small database of paired LF and HF simulations to train a computationally inexpensive surrogate model. Proper Orthogonal Decomposition (POD) is employed for dimensionality reduction, demonstrating exceptional efficiency by compressing the high-dimensional flow state into a latent space of just 10 modes with negligible information loss. A central contribution of this work is a rigorous comparison of two surrogate modeling techniques for mapping the LF latent space to the HF latent space: an Artificial Neural Network (ANN) and a Kriging (Gaussian Process Regression) interpolator. The results demonstrate the framework's success, with the Kriging-based Reduced Order Model (ROM) achieving a computational speedup of over 38,000 times compared to the HF simulation while maintaining remarkable accuracy, evidenced by a normalized root mean squared error of just 0.01% for the temperature field and 0.04% for the wall heat flux. A key finding is the superior performance of the Kriging model over the ANN, highlighting that for engineering problems characterized by smooth, low-dimensional mappings and limited data, established statistical methods can offer greater accuracy and efficiency than more complex deep learning architectures. This work provides a robust framework that breaks the trade-off between computational cost and accuracy, enabling previously intractable many-query analyses essential for modern aerospace design.

**Keywords:** Reduced Order Model; Data-Driven; Machine Learning; Fluid Flow Reconstruction





## Resumo

O projeto aerotermodinâmico de sistemas de propulsão de alto desempenho, como bocais de foguetes, apresenta um desafio computacional crítico. Simulações de alta fidelidade (HF), tipicamente baseadas nas equações de Navier-Stokes Médias por Reynolds (RANS) acopladas a modelos de transferência de calor conjugada (CHT), fornecem a precisão necessária para prever o fluxo de calor na parede e a temperatura, mas são proibitivamente caras para aplicações de muitas consultas, como otimização de projeto e quantificação de incerteza (UQ). Por outro lado, modelos de baixa fidelidade (LF), como solucionadores de Euler quase unidimensionais, são computacionalmente baratos, mas carecem da fidelidade física para prever fenômenos viscosos e térmicos críticos. Esta dissertação aborda esse dilema desenvolvendo e validando uma estrutura multi-fidelidade orientada por dados para a reconstrução rápida e precisa de campos de fluxo compressíveis bidimensionais em um bocal convergente-divergente resfriado a água. A metodologia aproveita um pequeno banco de dados de simulações LF e HF pareadas para treinar um modelo substituto computacionalmente barato. A Decomposição Ortogonal Própria (POD) é empregada para redução de dimensionalidade, demonstrando eficiência excepcional ao comprimir o estado de fluxo de alta dimensão em um espaço latente de apenas 10 modos com perda de informação desprezível. Uma contribuição central deste trabalho é uma comparação rigorosa de duas técnicas de modelagem substituta para mapear o espaço latente LF para o espaço latente HF: uma Rede Neural Artificial (ANN) e um interpolador Kriging (Regressão por Processo Gaussiano). Os resultados demonstram o sucesso da estrutura, com o Modelo de Ordem Reduzida (ROM) baseado em Kriging alcançando um aumento de velocidade computacional de mais de 38.000 vezes em comparação com a simulação HF, mantendo uma precisão notável, evidenciada por um erro quadrático médio normalizado de apenas 0,01% para o campo de temperatura e 0,04% para o fluxo de calor na parede. Uma descoberta chave é o desempenho superior do modelo Kriging em relação à ANN, destacando que para problemas de engenharia caracterizados por mapeamentos suaves e de baixa dimensão e dados limitados, métodos estatísticos estabelecidos podem oferecer maior precisão e eficiência do que arquiteturas de aprendizado profundo mais complexas. Este trabalho fornece uma estrutura robusta que quebra a relação de compromisso entre custo computacional e precisão, permitindo análises de muitas consultas anteriormente intratáveis essenciais para o projeto aeroespacial moderno.

**Palavras-chave:** Modelo de Ordem Reduzida; Orientado a Dados; Aprendizado de Máquina; Reconstrução de Fluxo de Fluidos;



## List of Figures



## List of Tables



## Contents





## Introduction

### 1.1 The Convergent-Divergent Nozzle: A Cornerstone of Aerospace Propulsion

The convergent-divergent nozzle, often referred to as the De Laval nozzle, represents a foundational technology in the field of high-speed aerospace propulsion (??). Its primary function is to accelerate a hot, pressurized gas to supersonic velocities, thereby generating thrust. This is achieved by converting the thermal and pressure energy of the working fluid, typically the product of combustion, into kinetic energy (?). The nozzle's characteristic shape, with a converging section, a minimum-area throat, and a diverging section, is meticulously designed to manage the compressible flow physics, as illustrated in Figure ??.

In operation, the gas accelerates to sonic velocity (Mach 1) at the throat—a condition known as choked flow—and subsequently expands and accelerates to supersonic speeds in the divergent section (??). The ultimate thrust produced is governed by the momentum of the exiting gas and any pressure differential at the nozzle exit, as described by the fundamental thrust equation, which underscores the critical importance of maximizing the exhaust velocity for efficient propulsion (?).

Figure 1.1: Schematic representation of a convergent-divergent (De Laval) nozzle. The top panel shows the physical contour, identifying the subsonic converging section, the throat where the flow becomes sonic ( $M = 1$ ), and the supersonic diverging section. The bottom panel illustrates the qualitative behavior of key flow properties along the nozzle axis ( $x$ ): pressure ( $p$ ) and temperature ( $T$ ) decrease, while velocity ( $v$ ) and Mach number ( $M$ ) increase.

### 1.2 The Critical Challenge of Aerothermodynamic Design

The operational environment within a convergent-divergent nozzle is exceptionally severe, characterized by extreme temperatures and pressures that impose extreme thermo-mechanical stresses on the structure (??). Consequently, a paramount challenge in nozzle design is the management of heat transfer from the hot gas to the nozzle walls. This aerothermodynamic problem is critical for ensuring the structural integrity and operational reliability of the entire propulsion system (?). Uncontrolled wall temperatures can lead to material degradation, thermal fatigue, and catastrophic failure (?).

To mitigate these risks, modern high-performance nozzles often incorporate active

cooling systems, such as circulating water through channels within the nozzle walls. This transforms the design problem into one of conjugate heat transfer (CHT), where the heat transfer within the fluid (convection) is intricately coupled with the heat transfer within the solid structure (conduction), as depicted in Figure ?? (?). Accurately predicting the temperature distribution and heat flux at the fluid-solid interface is therefore not merely an academic exercise but a mission-critical design requirement.

Figure 1.2: Conceptual diagram of the conjugate heat transfer (CHT) problem in a cooled nozzle wall. The figure shows the coupled physics: convective heat transfer ( $q_{conv}$ ) from the hot gas flow to the inner wall surface, heat conduction ( $q_{cond}$ ) through the solid wall material, and convective heat transfer ( $q_{cool}$ ) from the channel walls to the internal coolant flow. The temperature at the fluid-solid interface is an outcome of this coupled system.

### 1.3 The Computational Dilemma: Accuracy vs. Cost

The modeling of such complex CHT phenomena necessitates the use of high-fidelity (HF) computational fluid dynamics (CFD) simulations. These simulations, which typically solve the Reynolds-Averaged Navier-Stokes (RANS) equations, are capable of capturing the essential physics of turbulence, viscosity, and heat transfer that govern the system's behavior (??). However, the computational cost associated with these HF simulations is substantial. A single simulation can require hours or days of high-performance computing resources, rendering them impractical for the many-evaluations scenarios that define modern engineering design (???). Processes such as design optimization, uncertainty quantification (UQ), and real-time control demand thousands of model evaluations, a task for which HF simulations are prohibitively expensive (?).

On the other end of the spectrum, low-fidelity (LF) models, such as simplified quasi-one-dimensional (quasi-1D) Euler solvers, offer computational speeds that are orders of magnitude faster. While these models can capture the bulk flow dynamics, their underlying physical assumptions—typically inviscid flow—prevent them from predicting viscous-driven phenomena like boundary layer development and, critically, wall heat transfer (?). This creates a significant computational dilemma, illustrated in Figure ??: the models that are fast enough are not accurate enough for the problem at hand, and the models that are accurate enough are far too slow for practical design cycles.

Figure 1.3: The trade-off between model fidelity and computational cost. Low-fidelity (LF) models are computationally inexpensive but offer lower predictive accuracy. High-fidelity (HF) models provide high accuracy at a prohibitive computational cost. The goal of this research is to develop a Reduced Order Model (ROM) that achieves HF accuracy at a computational cost comparable to the LF model, effectively breaking the trade-off.

## 1.4 A Data-Driven Paradigm: Reduced Order Models

This research addresses the computational dilemma by proposing a data-driven, multi-fidelity framework that synergistically combines the strengths of both HF and LF models. The core of this approach is the development of a Reduced Order Model (ROM), a computationally inexpensive surrogate that learns the complex relationship between the outputs of the LF and HF simulations (???). This surrogate model functions as a sophisticated "black-box" approximation, trained on a pre-computed database of paired LF and HF results. The general workflow is depicted in Figure ???. Once trained, the ROM can rapidly generate predictions with accuracy approaching that of the HF model, but at a computational cost comparable to the LF model (???). This paradigm effectively breaks the trade-off between accuracy and cost, enabling the rapid exploration of the design space required for advanced engineering analyses.

Figure 1.4: The data-driven Reduced Order Model (ROM) methodology, divided into two stages. (a) The offline training stage is computationally expensive and performed once. It involves generating paired high-fidelity (HF) and low-fidelity (LF) simulation snapshots, using Proper Orthogonal Decomposition (POD) to find low-dimensional bases, and training a surrogate model to map the LF coefficients to the HF coefficients. (b) The online prediction stage is computationally cheap and performed in real-time. It involves running a new fast LF simulation, projecting its result onto the LF basis, using the trained surrogate to predict the HF coefficients, and reconstructing the full HF flow field.

## 1.5 Thesis Statement and Dissertation Structure

This dissertation details the development, implementation, and validation of a data-driven, multi-fidelity framework for the rapid and accurate reconstruction of high-fidelity, two-dimensional compressible flow fields in water-cooled, convergent-divergent nozzles. The framework utilizes Proper Orthogonal Decomposition for dimensionality reduction and critically compares the efficacy of Artificial Neural Networks and Kriging interpolation as surrogate models to bridge the fidelity gap.

The subsequent chapters are structured as follows. Chapter 2 establishes the theoretical foundations of compressible flow, nozzle aerothermodynamics, and turbulence modeling. Chapter 3 details the multi-fidelity computational framework, including the setup, verification, and validation of the high- and low-fidelity models. Chapter 4 presents the core data-driven methodology, explaining the principles of Reduced Order Modeling, Proper Orthogonal Decomposition, and the surrogate models employed. Chapter 5 provides a comprehensive analysis of the framework's performance, comparing the accuracy and computational efficiency of the different approaches. Finally, Chapter 6 synthesizes the findings, draws conclusions, and outlines promising directions for future research.



# Theoretical Foundations of Compressible Flow and Nozzle Aerothermodynamics

## 2.1 Governing Equations of Fluid Motion

### 2.1.1 The Euler Equations

The foundation for modeling inviscid, compressible fluid flow is the set of Euler equations, which represent the conservation of mass, momentum, and energy. For a quasi-one-dimensional flow through a duct with a varying cross-sectional area  $S(x)$ , the equations can be expressed in a conservative, integral form. The low-fidelity model in this work employs an in-house finite volume solver for the conservative variables. The governing law can be written in vector form as:

$$\frac{\partial \mathbf{U}}{\partial t} + \frac{\partial \mathbf{F}}{\partial x} = \mathbf{Q} \quad (2.1)$$

where  $\mathbf{U}$  is the vector of conserved variables,  $\mathbf{F}$  is the flux vector, and  $\mathbf{Q}$  is the source term vector, defined as:

$$\mathbf{U} = \begin{pmatrix} \rho S \\ \rho u S \\ E_{tot} S \end{pmatrix}, \quad \mathbf{F} = \begin{pmatrix} \rho u S \\ (\rho u^2 + p) S \\ (E_{tot} + p) u S \end{pmatrix}, \quad \mathbf{Q} = \begin{pmatrix} 0 \\ p \frac{dS}{dx} \\ 0 \end{pmatrix}$$

Here,  $\rho$  is the fluid density,  $u$  is the velocity in the axial direction  $x$ ,  $p$  is the static pressure, and  $E_{tot}$  is the total energy per unit volume, given by  $E_{tot} = \rho(e + u^2/2)$ , where  $e$  is the specific internal energy. For a calorically perfect gas, this is related to pressure and temperature ( $T$ ) by the ideal gas law,  $p = \rho R T$ , and the relation  $e = c_v T = p/((\gamma - 1)\rho)$ , where  $\gamma$  is the ratio of specific heats. The source term  $\mathbf{Q}$  arises from the dimensional reduction to a single spatial coordinate, accounting for the pressure forces exerted by the nozzle walls on the fluid as the area  $S(x)$  changes (?). This formulation provides the mathematical basis for the computationally efficient low-fidelity model. The numerical implementation uses a finite volume method with appropriate flux splitting schemes (e.g., Steger-Warming) to handle the hyperbolic nature of the equations and accurately capture discontinuities like shock waves. Boundary conditions typically involve specifying total pressure and total temperature at the inlet and static pressure at the outlet.

### 2.1.2 The Navier-Stokes Equations

To capture the effects of viscosity and heat conduction, which are neglected in the Euler equations but are crucial for predicting wall friction and heat transfer, the more comprehensive Navier-Stokes equations must be employed.[1] These equations extend the Euler equations by incorporating stress tensor and heat flux terms, which account for the dissipative effects within the fluid (?). They form the physical and mathematical basis for high-fidelity CFD simulations.

In their integral, conservative form for a fixed control volume  $\Omega$  with boundary  $\partial\Omega$ , the compressible Navier-Stokes equations can be written as [2, 3]:

$$\frac{\partial}{\partial t} \int_{\Omega} \mathbf{U} d\Omega + \oint_{\partial\Omega} \mathbf{F}_c \cdot \mathbf{n} dS = \oint_{\partial\Omega} \mathbf{F}_v \cdot \mathbf{n} dS \quad (2.2)$$

where  $\mathbf{U}$  is the vector of conserved state variables, and the total flux is split into a convective (inviscid) part  $\mathbf{F}_c$  and a viscous part  $\mathbf{F}_v$ . For a 2D flow, these are defined as [4]:

$$\mathbf{U} = \begin{pmatrix} \rho \\ \rho u \\ \rho v \\ \rho E \end{pmatrix}, \quad \mathbf{F}_c = \begin{pmatrix} \rho u & \rho v \\ \rho u^2 + p & \rho uv \\ \rho uv & \rho v^2 + p \\ (\rho E + p)u & (\rho E + p)v \end{pmatrix}, \quad \mathbf{F}_v = \begin{pmatrix} 0 & 0 \\ \tau_{xx} & \tau_{xy} \\ \tau_{yx} & \tau_{yy} \\ u\tau_{xx} + v\tau_{xy} + q_x & u\tau_{yx} + v\tau_{yy} + q_y \end{pmatrix}$$

Here,  $u$  and  $v$  are the velocity components,  $E$  is the total energy per unit mass,  $\boldsymbol{\tau}$  is the viscous stress tensor, and  $\mathbf{q}$  is the heat flux vector, given by Fourier's law  $\mathbf{q} = -k\nabla T$ . The viscous stress tensor components  $\tau_{ij}$  are functions of the fluid viscosity and the velocity gradients. This complete formulation allows for the modeling of boundary layers and heat transfer, which are essential for the high-fidelity analysis in this work.[1, 4]

### 2.1.3 Reynolds-Averaged Navier-Stokes (RANS)

For the turbulent flows typically encountered in propulsion systems, directly solving the Navier-Stokes equations (DNS) is computationally infeasible for engineering applications. The Reynolds-Averaged Navier-Stokes (RANS) approach offers a practical alternative by decomposing flow variables into a mean and a fluctuating component. This time-averaging process introduces new terms, known as Reynolds stresses, into the momentum equations, creating a closure problem that requires additional modeling (??). The final form of the RANS momentum equation for an incompressible fluid can be written in tensor notation as:

$$\frac{\partial(\rho \bar{u}_i)}{\partial t} + \frac{\partial(\rho \bar{u}_i \bar{u}_j)}{\partial x_j} = -\frac{\partial \bar{p}}{\partial x_i} + \frac{\partial}{\partial x_j} \left[ \mu \left( \frac{\partial \bar{u}_i}{\partial x_j} + \frac{\partial \bar{u}_j}{\partial x_i} \right) - \overline{\rho u'_i u'_j} \right] + F_i \quad (2.3)$$

where the overbar denotes a time-averaged quantity and the prime denotes a fluctuating component. The new term,  $-\overline{\rho u'_i u'_j}$ , is the \*\*Reynolds stress tensor\*\*. It represents the

net transport of momentum due to turbulent fluctuations and acts as an additional stress on the mean flow. Since this term contains unknown correlations of fluctuating velocities, it must be modeled to close the system of equations. A turbulence model provides this closure by relating the Reynolds stresses to known mean flow quantities. A detailed derivation of the RANS equations is provided in Appendix A.

## 2.2 Physics of Convergent-Divergent Nozzle Flow

The operation of a De Laval nozzle is governed by the principles of isentropic compressible flow theory. As a pressurized gas enters the converging section at subsonic speed, its velocity increases while its pressure and temperature decrease. For a sufficiently low back pressure, the flow accelerates until it reaches sonic velocity ( $\text{Mach} = 1$ ) precisely at the throat, the point of minimum cross-sectional area. This condition is known as choked flow (??). Beyond the throat, in the diverging section, the flow continues to expand and accelerate into the supersonic regime. Under certain operating conditions, particularly when the nozzle exit pressure does not match the ambient pressure, discontinuities in the flow known as shock waves can form within or outside the nozzle (??). These shocks can significantly impact nozzle performance. A key capability of the simplified quasi-1D Euler model used in this work is its ability to accurately capture the position of such shock waves, a critical feature for validating its representation of the bulk flow physics (?).

## 2.3 Conjugate Heat Transfer (CHT)

The aerothermodynamic analysis of a cooled nozzle wall is fundamentally a conjugate heat transfer (CHT) problem. CHT refers to the coupled simulation of two distinct physical domains: the fluid domain, governed by the Navier-Stokes equations, and the solid domain, governed by the heat conduction equation (?). Within the solid domain, the governing equation for heat transfer is the heat diffusion equation. For a steady-state problem without internal heat generation, as is the case for the nozzle wall itself, this equation simplifies to:

$$\nabla \cdot (k_s \nabla T_s) = 0 \quad (2.4)$$

where  $k_s$  is the thermal conductivity of the solid material and  $T_s$  is the temperature field within the solid. This equation is solved numerically by the CHT solver (in this work, SU2) for the solid region.

At the interface between the fluid and the solid, the temperatures and heat fluxes must be matched, creating a tightly coupled system. A typical numerical solution strategy, as employed in this research, involves an iterative process. The flow field is solved to

determine the convective heat transfer to the wall. This heat flux then serves as a boundary condition for solving the heat equation within the solid wall material. The resulting wall temperature is then passed back as a boundary condition for the fluid simulation. This iterative exchange continues until the temperature and heat flux at the fluid-solid interface converge to a steady-state solution (?).

## 2.4 Turbulence Modeling: The Shear Stress Transport (SST) Model

The choice of turbulence model is critical for the accuracy of the high-fidelity simulation. This work employs the Menter’s Shear Stress Transport (SST)  $k - \omega$  model, a robust and widely used two-equation eddy-viscosity model (??). The SST model is formulated as a hybrid model that leverages the strengths of two traditional models: the  $k - \omega$  model and the  $k - \epsilon$  model (?). The  $k - \omega$  model is known for its superior accuracy in resolving the flow within the viscous sublayer and logarithmic region of the boundary layer, which is the region immediately adjacent to the wall (?). Since the primary quantity of interest is wall heat transfer—a phenomenon dominated by near-wall physics—the accuracy of the model in this region is paramount. However, the standard  $k - \omega$  model can be overly sensitive to the freestream turbulence conditions far from the wall.

Conversely, the  $k - \epsilon$  model is more robust and less sensitive in these freestream regions (?). The SST model ingeniously combines these two approaches using a blending function,  $F_1$ , as illustrated in Figure ??. This function activates the  $k - \omega$  model in the near-wall region ( $F_1 = 1$ ) to capitalize on its accuracy and smoothly transitions to the more stable  $k - \epsilon$  model in the freestream ( $F_1 = 0$ ), away from the boundary layer. The transport equations for the turbulent kinetic energy,  $k$ , and the specific dissipation rate,  $\omega$ , are:

$$\frac{\partial(\rho k)}{\partial t} + \frac{\partial(\rho u_j k)}{\partial x_j} = P_k - \beta^* \rho \omega k + \frac{\partial}{\partial x_j} \left[ (\mu + \sigma_k \mu_t) \frac{\partial k}{\partial x_j} \right] \quad (2.5)$$

$$\frac{\partial(\rho \omega)}{\partial t} + \frac{\partial(\rho u_j \omega)}{\partial x_j} = \frac{\gamma}{\nu_t} P_k - \beta \rho \omega^2 + \frac{\partial}{\partial x_j} \left[ (\mu + \sigma_\omega \mu_t) \frac{\partial \omega}{\partial x_j} \right] + 2(1 - F_1) \rho \sigma_{\omega,2} \frac{1}{\omega} \frac{\partial k}{\partial x_j} \frac{\partial \omega}{\partial x_j} \quad (2.6)$$

The final term in the  $\omega$  equation is the cross-diffusion term, which is active only in the outer region ( $F_1 \approx 0$ ) and facilitates the transition to the  $k - \epsilon$  formulation. All model constants ( $\beta^*$ ,  $\sigma_k$ , etc.) are blended using the same function  $F_1$ . This hybrid formulation makes the SST model particularly well-suited for predicting flows with adverse pressure gradients and boundary layer separation, both of which can occur within a nozzle. Therefore, its use represents a methodical choice to maximize the “fidelity” of the high-fidelity simulation for the specific engineering goal of accurately predicting wall heat transfer.



Figure 2.1: Illustrative diagram of the SST turbulence model's blending function ( $F_1$ ) across a boundary layer. The function activates the accurate  $k - \omega$  model ( $F_1 = 1$ ) in the near-wall region and transitions to the robust  $k - \epsilon$  model ( $F_1 = 0$ ) in the outer part of the boundary layer and the freestream, combining the strengths of both approaches.



## A Multi-Fidelity Computational Framework

### 3.1 High-Fidelity Simulation: The 2D RANS-CHT Approach

#### 3.1.1 The SU2 CFD Suite

The high-fidelity simulations were conducted using the SU2 software suite, an open-source platform developed for multiphysics analysis and design optimization (???). SU2 (Stanford University Unstructured) is a collection of C++ and Python-based tools designed to solve partial differential equations on unstructured meshes using state-of-the-art numerical methods (?). Its robust capabilities for handling compressible flows, conjugate heat transfer, and a wide range of turbulence models make it a suitable choice for this research (??).

#### 3.1.2 Model Setup

The high-fidelity model was configured to solve the conjugate heat transfer problem in its entirety. Within the fluid domain, the Reynolds-Averaged Navier-Stokes (RANS) equations were solved using a finite volume method. The Menter's SST turbulence model was employed to close the equations, as detailed in the previous chapter (?). In the solid domain, representing the nozzle wall, the energy equation was solved to model heat conduction. To achieve a steady-state solution, the simulation utilized an implicit Euler integration method coupled with time marching until convergence was reached at the fluid-solid interface (?).

#### 3.1.3 Verification: Grid Independence Study

A critical step in any CFD analysis is verification, which ensures that the numerical solution is not an artifact of the computational grid. A grid independence study was performed using the Grid Convergence Index (GCI) method proposed by Roache (1994) to quantify the discretization error (?). Three structured grids of increasing refinement were generated, as detailed in Table ??.

Table 3.1: Numerical domains for the grid independence study. The table details the number of nodes in the axial ( $N_x$ ) and radial ( $N_y$ ) directions, the total number of cells ( $N_{cells}$ ), the grid refinement ratio ( $r$ ), and the approximate near-wall grid spacing in wall units ( $Y^+$ ) (?).

	$N_x$	$N_y$	$N_{cells}$	$r$	$Y^+$
<b>Grid 1</b>	270	440	118091	1.3	$\approx 1$
<b>Grid 2</b>	210	330	68761	1.3	$\approx 2$
<b>Grid 3</b>	160	250	39591		$\approx 3$

The GCI was calculated for key output quantities: average heat flux along the wall ( $q$ ), pressure along the centerline ( $p$ ), and temperature at the wall ( $T$ ). The results, summarized in Table ??, show that the solutions exhibit monotonic and asymptotic convergence, as indicated by the positive order of convergence ( $p$ ) and small GCI values. An asymptotic ratio ( $GCI_{asymptotic}$ ) close to 1 signifies that the solution is in the asymptotic range of convergence and is effectively grid-independent. Based on this rigorous analysis, Grid 2 was selected for all subsequent high-fidelity simulations, as it provided a reliable solution with a manageable cell count (?). This verification process is fundamental to the entire study, as it establishes a credible and converged high-fidelity baseline against which the reduced-order models are ultimately judged.

Table 3.2: Results for the grid independence study for average heat flux ( $q$ ), pressure ( $p$ ), and wall temperature ( $T$ ). The analysis confirms asymptotic convergence, justifying the selection of Grid 2 (?).

	$q$ [ $W/m^2$ ]	$N_{cells}$	$r$	GCI	$GCI_{asymptotic}$	$p$
<b>Grid 1</b>	$519 \times 10^3$	118091	1.3	6.95%		
<b>Grid 2</b>	$502 \times 10^3$	68761	1.3	11.40%	1.034	1.71
<b>Grid 3</b>	$519 \times 10^3$	39591				
	$p$ [kPa]	$N_{cells}$	$r$	GCI	$GCI_{asymptotic}$	$p$
<b>Grid 1</b>	$737 \times 10^3$	118091	1.3	0.23%		
<b>Grid 2</b>	$736 \times 10^3$	68761	1.3	0.30%	1.001	1.02
<b>Grid 3</b>	$736 \times 10^3$	39591				
	$T$ [K]	$N_{cells}$	$r$	GCI	$GCI_{asymptotic}$	$p$
<b>Grid 1</b>	518	118091	1.3	0.17%		
<b>Grid 2</b>	520	68761	1.3	0.47%	0.998	3.73
<b>Grid 3</b>	523	39591				

### 3.1.4 Comparison with Experimental Data

Following verification, the high-fidelity model was validated by comparing its predictions against experimental data for test case no. 313 from Back et al. (1964) (?). This validation process faced certain challenges due to incomplete information in the reference study, such as the exact composition of the nozzle wall, which was assumed to be AISI 302 stainless steel. Furthermore, the boundary condition for the outer cooling wall was simplified to a constant temperature of 300 K. Despite these limitations, the validation showed promising results, as seen in Figure ??. The predicted pressure distribution along the nozzle centerline demonstrated good agreement with the experimental measurements. While the predicted wall temperature showed a slight deviation, likely attributable to the simplified boundary conditions and material property assumptions, the overall agreement was sufficient to establish confidence in the high-fidelity model's ability to capture the essential physics of the problem (?).

Figure 3.1: Validation of the high-fidelity (HF) model. Comparison of numerical predictions from the SU2 RANS-CHT simulation against experimental data from Back et al. (1964). (a) Pressure distribution along the nozzle centerline. (b) Temperature distribution along the nozzle wall. The good agreement validates the HF model’s predictive capability.

### 3.2 Low-Fidelity Simulation: The Quasi-1D Euler Approach

The low-fidelity model is an in-house, quasi-one-dimensional Euler solver designed for maximum computational efficiency (?). It solves the conservation laws for an inviscid, ideal gas on a regular 1D mesh. The numerical scheme employs the finite volume method with Steger-Warming for spatial discretization and the AUSM method for flux evaluation. A fourth-order Runge-Kutta integrator is used for time marching to a steady-state solution (?). To verify this simplified model, its results were compared against published numerical data from Arina (2004) for a standard nozzle test case. The comparison, shown in Figure ??, showed excellent agreement in the pressure and density profiles along the nozzle, including the accurate capture of a shock wave’s position and strength. This verification confirmed that despite its inherent simplifications, the low-fidelity model provides a reliable representation of the core, one-dimensional flow dynamics (?).

Figure 3.2: Verification of the low-fidelity (LF) quasi-1D Euler solver. Comparison of the in-house solver’s predictions against reference numerical data from Arina (2004) for a standard nozzle test case with a shock wave. (a) Pressure profile. (b) Density profile. The excellent agreement, including the shock position and strength, verifies the correctness of the LF model implementation.

### 3.3 Parametric Data Generation for Model Training

To train the data-driven surrogate models, a comprehensive dataset mapping the low-fidelity inputs to the high-fidelity outputs was required. This dataset was generated by running 181 paired simulations, where each pair consisted of one LF and one HF simulation for the same set of boundary conditions (?). The key parameter varied across these simulations was the inlet total temperature ( $T_{0in}$ ), which was sampled across a range of 285 K to 1115 K.

To ensure efficient and uniform exploration of this parameter space, Latin Hypercube Sampling (LHS) was employed. LHS is a statistical method for Design of Experiments (DoE) that generates a more representative sample of the parameter space compared to simple random sampling, making it a standard and effective practice for generating training data for surrogate models (?). The resulting dataset of 181 snapshot pairs was then partitioned into three subsets for the machine learning workflow: 145 pairs for training the models, 18 for validation during the training process to prevent overfitting, and a final

18 pairs held out as a blind test set to evaluate the final performance of the trained models (?). A summary of the dataset is provided in Table ??.

Table 3.3: Summary of the dataset generation and partitioning for machine learning model development.

Purpose	# of Samples	Sampling Method	Parameter Varied	Range
Training	145	LHS	Inlet Total Temp. ( $T_{0in}$ )	285 K - 1115 K
Validation	18	LHS	Inlet Total Temp. ( $T_{0in}$ )	285 K - 1115 K
Testing	18	LHS	Inlet Total Temp. ( $T_{0in}$ )	285 K - 1115 K
<b>Total</b>	<b>181</b>			





## Data-Driven Flow Reconstruction via Reduced Order Modeling

### 4.1 The Philosophy of Reduced Order Modeling (ROM)

Reduced Order Modeling is a paradigm designed to address the computational bottleneck inherent in high-fidelity simulations. ROMs create computationally inexpensive, low-dimensional surrogate models that approximate the behavior of a full-order, complex system (???). These models are particularly valuable in many-query contexts like optimization, control, and uncertainty quantification, where repeated evaluations of the full model are infeasible (??). ROMs can be broadly categorized into two classes: intrusive and non-intrusive. Intrusive methods, such as Galerkin projection, require modification of the governing equation solvers. In contrast, non-intrusive methods are purely data-driven, treating the full-order model as a black box and learning the input-output relationship from simulation data (??). The framework developed in this dissertation falls into the non-intrusive category.

### 4.2 Dimensionality Reduction with Proper Orthogonal Decomposition (POD)

#### 4.2.1 Theoretical Underpinnings of POD

A foundational step in building an effective ROM is dimensionality reduction. Proper Orthogonal Decomposition (POD) is a powerful mathematical technique for extracting a low-dimensional basis that optimally captures the variance within a high-dimensional dataset (???). Mathematically, POD is equivalent to Principal Component Analysis (PCA) and is implemented via the Singular Value Decomposition (SVD) of the data matrix (??). The method decomposes a complex, spatiotemporal field into a set of orthogonal spatial modes (the POD modes) and corresponding temporal coefficients. These modes are hierarchical; the first mode captures the most energy (or variance) in the data, the second captures the most of the remaining energy, and so on. This property allows for an extremely efficient, low-rank approximation of the original data by retaining only the first few most energetic modes (??). The core of the method is the SVD of a snapshot matrix  $\mathbf{X}$ , which is detailed in the following section. A full mathematical formulation of POD is

provided in Appendix B.

### 4.2.2 The Snapshot Method

The specific implementation used in this study is the snapshot method of POD. The process begins by assembling the data from the 181 high-fidelity simulations into a snapshot matrix,  $\mathbf{X}_h \in \mathbb{R}^{m \times n}$ , where each of the  $n = 181$  columns is a vector containing the flattened field data (pressure, temperature, Mach number, and wall heat flux) from a single simulation, and  $m$  is the total number of degrees of freedom (grid points  $\times$  variables) (?). A similar matrix,  $\mathbf{X}_l$ , is created for the low-fidelity data.

The SVD is then performed on these matrices. For a generic snapshot matrix  $\mathbf{X}$ , the SVD is given by:

$$\mathbf{X} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T \quad (4.1)$$

where  $\mathbf{U} \in \mathbb{R}^{m \times m}$  is an orthogonal matrix whose columns are the left singular vectors (the POD spatial modes),  $\mathbf{\Sigma} \in \mathbb{R}^{m \times n}$  is a diagonal matrix of singular values  $\sigma_i$ , and  $\mathbf{V} \in \mathbb{R}^{n \times n}$  is an orthogonal matrix whose columns are the right singular vectors (related to the temporal coefficients).

By projecting the high-dimensional snapshot data onto a small number of these modes (a truncated basis  $\mathbf{\Psi}_l = [\mathbf{u}_1, \dots, \mathbf{u}_l]$ ), one obtains a set of low-dimensional coefficients that serve as a compact representation of the full system state. The approximation of a snapshot  $\mathbf{x}_j$  and the calculation of its coefficient vector  $\mathbf{a}_j$  are given by:

$$\mathbf{x}_j \approx \mathbf{\Psi}_l \mathbf{a}_j \quad (4.2)$$

$$\mathbf{a}_j = \mathbf{\Psi}_l^T \mathbf{x}_j \quad (4.3)$$

### 4.2.3 Application and Efficacy

The application of POD to the nozzle flow dataset yielded a remarkable degree of dimensionality reduction. The high-fidelity state, originally described by 208,110 degrees of freedom (the number of grid points times the number of variables), was accurately represented using only 10 POD modes. The effectiveness of this compression is visualized in Figure ???. The energy spectrum shows a rapid decay, indicating that the vast majority of the system's variance is captured by the first few modes. The projection error associated with this truncation was exceptionally low, with an average global relative error in the temperature field of just 0.0006% and a maximum error for a sample case of only 0.0133% (?). This demonstrates the extraordinary efficiency of POD in compressing the vast

amount of information contained in the CFD solutions into a highly compact, latent variable representation with negligible loss of information.

Figure 4.1: Efficacy of Proper Orthogonal Decomposition (POD). (a) Normalized energy spectrum, showing the rapid decay of singular values, which indicates that the first few modes capture most of the system’s energy. (b) Visualization of the first three dominant spatial POD modes for the temperature field, representing the fundamental patterns of temperature variation across the dataset.

This powerful compression is the key enabler for the entire surrogate modeling framework. The original problem—mapping a 633-dimensional low-fidelity state vector to a 208,110-dimensional high-fidelity state vector—is a high-dimensional regression task that would be intractable for any machine learning model without an enormous dataset. POD transforms this problem. It acts as an expert feature extractor, identifying the fundamental building blocks (the modes) of the flow. The complex task of learning the spatially distributed physics is thus converted into a much simpler one: learning the mapping between the 10 coefficients of the low-fidelity model and the 10 coefficients of the high-fidelity model. In this synergistic partnership, POD performs the heavy lifting of identifying the relevant physical structures, while the surrogate model learns the relatively simple correlation between their amplitudes. This synergy is the cornerstone of the framework’s success.

### 4.3 Surrogate Modeling for Latent Space Mapping

Once the dimensionality of the problem has been reduced via POD, a surrogate model is required to learn the nonlinear mapping from the low-fidelity latent space to the high-fidelity latent space. This study compares two distinct approaches for this task: an Artificial Neural Network and a Kriging interpolator.

#### 4.3.1 The Artificial Neural Network (ANN) Approximator

Artificial Neural Networks are powerful computational models inspired by biological neural systems, renowned for their ability to act as universal function approximators (?). The specific architecture used here is a dense Multilayer Perceptron (MLP), which consists of an input layer, multiple hidden layers of interconnected “neurons,” and an output layer (?). Each connection has an associated weight, and each neuron applies a nonlinear activation function to its input. The network “learns” by adjusting these weights through a process called back-propagation, which uses gradient descent to minimize a loss function between the network’s predictions and the true data (?).

The overall mapping from a new vector of low-fidelity coefficients,  $\mathbf{a}_l^*$ , to the predicted high-fidelity coefficients,  $\hat{\mathbf{a}}_h^*$ , can be expressed as a nested function representing the entire

network:

$$\hat{\mathbf{a}}_h^* = \mathcal{N}(\mathbf{a}_l^*; \boldsymbol{\theta}_{NN}) = f^{(L)} \left( \mathbf{W}^{(L)} f^{(L-1)} \left( \dots f^{(1)} \left( \mathbf{W}^{(1)} \mathbf{a}_l^* + \mathbf{b}^{(1)} \right) \dots \right) + \mathbf{b}^{(L)} \right) \quad (4.4)$$

where  $L$  is the number of layers,  $f^{(k)}$  is the activation function of layer  $k$ , and  $\boldsymbol{\theta}_{NN} = \{\mathbf{W}^{(k)}, \mathbf{b}^{(k)}\}_{k=1}^L$  is the set of all trained weights and biases of the network.

The implementation in this work utilized the TensorFlow library to construct a dense MLP with the architecture shown in Figure ??(a): an input layer with 10 neurons (for the 10 LF POD coefficients), four hidden layers each containing 10 neurons, and an output layer with 10 neurons (for the 10 HF POD coefficients). The hyperbolic tangent ('tanh') was used as the activation function, and the ADAM optimizer was employed for efficient training. The training history, shown in Figure ??(b), showed a rapid decrease in both the training and validation loss functions, indicating effective learning of the mapping between the latent spaces without significant overfitting (?).

Figure 4.2: Artificial Neural Network (ANN) model details. (a) Schematic of the Multilayer Perceptron (MLP) architecture used, showing the input layer (10 LF coefficients), four hidden layers, and the output layer (10 HF coefficients). (b) Learning curve from the training process, plotting the Mean Squared Error (MSE) versus training epochs for both the training and validation datasets. The convergence of both curves indicates successful training.

### 4.3.2 The Kriging Interpolator

Kriging, also known as Gaussian Process Regression, is a sophisticated statistical interpolation method that has its roots in geostatistics (??). It frames the interpolation problem from a Bayesian perspective, modeling the unknown function as a realization of a Gaussian process. A prediction at a new point is calculated as a weighted average of the known training data points. Crucially, these weights are not based on simple distance but are determined by a statistical model of spatial correlation (a covariance function or variogram) that is learned from the data itself (??).

Since the mapping is from a vector of low-fidelity coefficients  $\mathbf{a}_l \in \mathbb{R}^{10}$  to a vector of high-fidelity coefficients  $\mathbf{a}_h \in \mathbb{R}^{10}$ , an independent Kriging model is trained for each of the 10 components of the output. The prediction for the  $j$ -th component of the high-fidelity coefficient vector,  $\hat{a}_{h,j}^*$ , given a new low-fidelity coefficient vector  $\mathbf{a}_l^*$ , is given by the Ordinary Kriging predictor:

$$\hat{a}_{h,j}^*(\mathbf{a}_l^*) = \mathcal{K}_j(\mathbf{a}_l^*; \boldsymbol{\theta}_{Kriging}) = \hat{\mu}_j + \mathbf{k}_{*,j}^T \mathbf{K}_j^{-1} (\mathbf{y}_j - \hat{\mu}_j \mathbf{1}) \quad (4.5)$$

Here,  $\mathbf{y}_j$  is the vector of the  $j$ -th high-fidelity coefficient from all training snapshots,  $\mathbf{K}_j$  is the covariance matrix of the low-fidelity training inputs computed using a covariance function (e.g., exponential), and  $\mathbf{k}_{*,j}$  is the vector of covariances between the new input  $\mathbf{a}_l^*$

and the training inputs. The model parameters,  $\boldsymbol{\theta}_{Kriging}$ , are the hyperparameters of the chosen covariance function (such as correlation lengths and variance), which are optimized during the training phase. A key advantage of Kriging is that, in addition to providing a prediction, it also provides a measure of the prediction uncertainty (the variance), which can be invaluable for engineering analysis (??). A detailed mathematical formulation of the Kriging predictor and variance is provided in Appendix C.

The implementation in this study employed ordinary Kriging with an exponential covariance model, assuming no underlying mean bias in the data. The model was trained using the same 145 snapshots from the dataset to learn the correlation structure of the latent space mapping (?).



## Analysis of Model Performance and Computational Efficiency

### 5.1 Flow Field Reconstruction Accuracy

The ultimate test of the Reduced Order Models is their ability to accurately reconstruct the full high-fidelity flow field from a new, unseen low-fidelity input. The reconstruction process follows a four-step algorithm, as depicted in Figure ??: (1) run a new LF simulation; (2) project the LF results into the pre-computed low-dimensional basis; (3) use the trained surrogate model (NN or Kriging) to map the LF coefficients to predicted HF coefficients; and (4) reconstruct the full HF flow field from the predicted coefficients and the HF basis vectors (?).

Figure 5.1: Flowchart of the online prediction and reconstruction stage. A new input parameter prompts a fast low-fidelity (LF) simulation. The resulting LF snapshot is projected onto its POD basis to obtain low-dimensional coefficients. The trained surrogate model (ANN or Kriging) maps these to predicted high-fidelity (HF) coefficients, which are then used with the HF POD basis to reconstruct the full HF flow field.

#### 5.1.1 Neural Network Reconstruction

The Neural Network-based ROM demonstrated high accuracy. For a representative test case with an inlet total temperature of 935 K, the reconstructed temperature field was qualitatively very similar to the ground truth full-order model (FOM) result, as shown in Figure ?. The mean relative error for the temperature field reconstruction was only 0.19% (?). More importantly for the engineering application, the predicted distributions of wall heat flux and wall temperature were almost indistinguishable from the high-fidelity simulation results, confirming the model's ability to accurately predict the critical quantities of interest (?).

Figure 5.2: Performance of the Neural Network (ANN) based ROM for a test case. (a) Temperature field from the high-fidelity (FOM) simulation. (b) Reconstructed temperature field from the ANN-ROM. (c) Pointwise absolute error between the FOM and ROM. (d) Comparison of wall heat flux. (e) Comparison of wall temperature. The model shows high accuracy in both the 2D field and the 1D wall quantities.

### 5.1.2 Kriging Reconstruction

The Kriging-based ROM performed even better, achieving a remarkable level of accuracy. For the same test case, the normalized root mean squared error for the temperature field reconstruction was a mere 0.001% (?). Visually, as seen in Figure ??, the reconstructed temperature field, as well as the predicted wall heat flux and wall temperature distributions, were identical to the high-fidelity results for all practical purposes. This indicates an exceptionally precise mapping learned by the Kriging model (?).

Figure 5.3: Performance of the Kriging-based ROM for a test case. (a) Temperature field from the high-fidelity (FOM) simulation. (b) Reconstructed temperature field from the Kriging-ROM. (c) Pointwise absolute error, which is visibly negligible. (d) Comparison of wall heat flux. (e) Comparison of wall temperature. The Kriging model achieves near-perfect reconstruction of both the 2D field and the 1D wall quantities.

## 5.2 A Comparative Assessment of Surrogate Modeling Techniques

A central contribution of this research is the direct, quantitative comparison between the modern deep learning approach (ANN) and the established statistical method (Kriging). The performance of both models, along with the baseline simulations, is synthesized in Table ??.

Table 5.1: Computational cost and accuracy comparison. The table shows the CPU time for training and a single evaluation, the speedup relative to the high-fidelity model, and the Normalized Root Mean Squared Error for reconstructed Temperature (NRMSE-T) and wall heat flux (NRMSE-q) for the test dataset (?).

	CPU time	speedup	NRMSE-T	NRMSE-q
<b>high-fidelity model</b>	1,532.42 seconds	-	-	-
<b>low-fidelity model</b>	1.02 seconds	1,502.37 x	-	-
<b>NN training</b>	231.14 seconds	-	-	-
<b>Kriging training</b>	1.63 seconds	-	-	-
<b>single NN evaluation</b>	0.23 seconds	6,662.70 x	1.38%	1.55%
<b>single Kriging evaluation</b>	0.04 seconds	38,310.50 x	0.01%	0.04%

The data in Table ?? shows that while both ROMs provide massive computational gains, the Kriging model is superior across every metric. It is more accurate (NRMSE for temperature is 0.01% vs. 1.38% for the NN), faster to evaluate (0.04s vs. 0.23s), and orders of magnitude cheaper to train (1.63s vs. 231.14s).

This counter-intuitive result, where a "classic" statistical method outperforms a more "modern" neural network, warrants deeper analysis. The reason for Kriging's superior



performance likely lies in the specific nature of this engineering problem: a relatively small training dataset (145 points) is used to learn a smooth, low-dimensional mapping (10 input variables to 10 output variables). Artificial Neural Networks are highly flexible models with a large number of free parameters (weights and biases). While this flexibility is a great strength for complex, high-dimensional problems with vast amounts of data, it can be a weakness with smaller datasets, where the network may struggle to generalize perfectly. Kriging, in contrast, is a method specifically designed for interpolation from sparse data (??). Its underlying statistical assumptions—that the function is a realization of a Gaussian process with a certain spatial correlation—provide a strong “inductive bias”. This bias constrains the model, guiding it to find the best linear unbiased predictor based on the correlations observed in the limited training data (??). For this particular problem, Kriging’s statistically-grounded approach proved to be a more efficient and robust strategy for interpolation than the more flexible but less constrained neural network. This finding serves as an important reminder for the engineering community: the application of complex deep learning models is not always the optimal path, and mature statistical techniques should be carefully considered, as they may be more appropriate and effective for the problem at hand.

### 5.3 Evaluation of Computational Gains and Engineering Impact

The practical implications of this research are profound. The Kriging-based ROM achieves a speedup of over 38,000 times compared to the high-fidelity simulation, reducing a computation that takes over 25 minutes to just 0.04 seconds, with negligible loss in accuracy (?). This dramatic acceleration transforms the landscape of what is possible in nozzle design and analysis.

Previously infeasible studies are now brought within reach. For instance, a comprehensive multi-objective design optimization, which might require tens of thousands of function evaluations to explore trade-offs between performance and structural integrity, can now be performed in hours instead of years (??). Full-scale uncertainty quantification, propagating uncertainties in manufacturing tolerances or operational conditions through the model to predict the reliability of the design, becomes a tractable problem (??). Furthermore, the near-instantaneous prediction speed opens the door to applications in inverse design (determining the nozzle shape required to produce a desired flow field) and even real-time control systems for advanced propulsion concepts (??).



## Synthesis, Conclusions, and Future Directions

### 6.1 Synthesis of Findings

This dissertation has successfully detailed the development and validation of a multi-fidelity, data-driven framework capable of accelerating the aerothermodynamic analysis of convergent-divergent nozzles by orders of magnitude. The key achievements of this work can be synthesized as follows:

- A robust multi-fidelity simulation workflow was established, coupling a high-fidelity 2D RANS-CHT solver (SU2) with a computationally inexpensive quasi-1D Euler solver. Both models were rigorously verified and validated.
- Proper Orthogonal Decomposition was shown to be exceptionally effective at dimensionality reduction, compressing the high-dimensional state of the flow field into a compact latent space of only 10 variables with minimal information loss.
- Two distinct machine learning-based surrogate models—an Artificial Neural Network and a Kriging interpolator—were successfully implemented and trained to map the low-fidelity latent space to the high-fidelity latent space.
- Quantitative analysis demonstrated that the resulting Reduced Order Models can reconstruct the high-fidelity flow field and critical wall heat transfer quantities with high accuracy, achieving a computational speedup of up to 38,000 times over the full-order model.

### 6.2 Concluding Remarks

The central conclusion of this work is twofold. First, it confirms that data-driven, non-intrusive Reduced Order Models provide a powerful and effective strategy for bridging the fidelity gap in the computational analysis of complex engineering systems like rocket nozzles. They successfully leverage the speed of simple models and the accuracy of complex ones, breaking the traditional compromise between computational cost and predictive fidelity. Second, and perhaps more significantly, this study demonstrates that for engineering problems characterized by smooth, low-dimensional mappings and limited data availability, established statistical methods like Kriging can outperform more complex, modern deep

learning architectures in terms of accuracy, training cost, and evaluation speed. This highlights the critical importance of selecting a modeling technique that is appropriate for the specific structure and scale of the problem at hand.

### 6.3 Future Directions and Broader Impact

While this study has established a successful framework, several promising avenues for future research exist that could extend its capabilities and impact.

- **Geometric Variations:** The current model is parameterized only by an inlet boundary condition. A significant extension would be to incorporate variations in the nozzle geometry itself, such as the contour shape or throat radius. This would require a more sophisticated parameterization of the geometry (e.g., using B-splines or Class Shape Function Transformations) and a substantially larger training dataset to cover the expanded design space (??).
- **Extension to 3D:** Applying this methodology to full three-dimensional nozzle simulations is the next logical step toward industrial relevance. This presents significant challenges, as the computational cost of the HF model and the sheer size of the data snapshots increase by orders of magnitude, demanding more advanced data handling and potentially more sophisticated ROM techniques (?).
- **Advanced ROM Techniques:** Future work could explore more advanced data-driven methods. For non-linear dimensionality reduction, deep autoencoders could be investigated as an alternative to the linear basis of POD (??). For unsteady flows, where transient behavior is important, methods like Dynamic Mode Decomposition (DMD) could be employed to capture the temporal dynamics of the system (??).
- **Integration into Optimization Loops:** The validated Kriging-based ROM is an ideal candidate for integration into a formal design optimization framework. It could be coupled with a multi-objective genetic algorithm, such as NSGA-II, to perform rapid optimization of the nozzle shape or its cooling channel design to maximize performance while respecting thermal constraints (?).
- **Uncertainty Quantification:** The Kriging model's inherent ability to provide a prediction variance should be fully exploited. This would enable a comprehensive Uncertainty Quantification (UQ) analysis, allowing designers to understand how uncertainties in input parameters (e.g., material properties, operating conditions) propagate to uncertainties in key performance metrics, leading to more robust designs (???).

The impact of this research contributes to the ongoing shift in engineering design towards a "digital twin" paradigm. By creating fast and accurate virtual representations of physical assets, these data-driven models enable continuous monitoring, real-time optimization, and predictive maintenance, paving the way for the next generation of intelligent and efficient aerospace systems (??).



## Mathematical Formulation of Reynolds-Averaged Navier-Stokes (RANS) Equations

The derivation of the RANS equations begins with the instantaneous Navier-Stokes equations for an incompressible fluid. In tensor notation, the momentum equation is [1, 2]:

$$\frac{\partial u_i}{\partial t} + \frac{\partial(u_i u_j)}{\partial x_j} = -\frac{1}{\rho} \frac{\partial p}{\partial x_i} + \nu \frac{\partial^2 u_i}{\partial x_j \partial x_j} \quad (\text{A.1})$$

where  $u_i$  are the velocity components,  $p$  is the pressure,  $\rho$  is the density, and  $\nu$  is the kinematic viscosity.

The core idea is the Reynolds decomposition, which separates any flow variable into a time-averaged (mean) component and a fluctuating component.[2, 3] For velocity and pressure, this is written as:

$$u_i(\mathbf{x}, t) = \bar{u}_i(\mathbf{x}) + u'_i(\mathbf{x}, t) \quad (\text{A.2})$$

$$p(\mathbf{x}, t) = \bar{p}(\mathbf{x}) + p'(\mathbf{x}, t) \quad (\text{A.3})$$

By definition, the time average of a fluctuating component is zero, i.e.,  $\overline{u'_i} = 0$ .

Substituting this decomposition into the momentum equation and then applying the time-averaging operator to each term yields:

$$\frac{\partial(\bar{u}_i + \overline{u'_i})}{\partial t} + \frac{\partial((\bar{u}_i + \overline{u'_i})(\bar{u}_j + \overline{u'_j}))}{\partial x_j} = -\frac{1}{\rho} \frac{\partial(\bar{p} + \overline{p'})}{\partial x_i} + \nu \frac{\partial^2(\bar{u}_i + \overline{u'_i})}{\partial x_j \partial x_j} \quad (\text{A.4})$$

Using the properties of the averaging operator (e.g.,  $\bar{\bar{u}_i} = \bar{u}_i$ ,  $\overline{\bar{u}_i} = 0$ , and  $\overline{\bar{u}_i \bar{u}_j} = \bar{u}_i \bar{u}_j = 0$ ), the equation simplifies. The nonlinear convective term expands to:

$$\overline{(\bar{u}_i + \overline{u'_i})(\bar{u}_j + \overline{u'_j})} = \bar{u}_i \bar{u}_j + \overline{\bar{u}_i u'_j} + \overline{u'_i \bar{u}_j} + \overline{u'_i u'_j} = \bar{u}_i \bar{u}_j + \overline{u'_i u'_j} \quad (\text{A.5})$$

The term  $\overline{u'_i u'_j}$  is the time average of the product of fluctuating velocities and is generally non-zero.

The resulting time-averaged momentum equation, or RANS equation, is [3, 4]:

$$\frac{\partial \bar{u}_i}{\partial t} + \bar{u}_j \frac{\partial \bar{u}_i}{\partial x_j} = -\frac{1}{\rho} \frac{\partial \bar{p}}{\partial x_i} + \frac{\partial}{\partial x_j} \left( \nu \frac{\partial \bar{u}_i}{\partial x_j} - \overline{u'_i u'_j} \right) \quad (\text{A.6})$$

This equation resembles the original momentum equation for the mean flow, but with

an additional term,  $-\overline{u'_i u'_j}$ . This term is moved to the right-hand side and interpreted as a stress. The term  $\overline{\rho u'_i u'_j}$  is known as the \*\*Reynolds stress tensor\*\*, representing the transport of momentum by turbulent fluctuations.[2, 3] The RANS equations are not a closed system due to this new unknown term, which gives rise to the "closure problem" of turbulence. Turbulence models, such as the SST model used in this work, are required to model the Reynolds stresses in terms of the mean flow variables to close the system.



## Mathematical Formulation of Proper Orthogonal Decomposition (POD)

Proper Orthogonal Decomposition (POD) is a technique to find an optimal low-dimensional basis for a given set of data. The formulation used in this work is the method of snapshots.[5, 6]

Let the data consist of  $n$  snapshots, where each snapshot  $y_j \in \mathbb{R}^m$  is a vector representing the state of the system at a specific time or parameter value. Here,  $m$  is the number of spatial degrees of freedom, and  $n$  is the number of snapshots. These snapshots are assembled into a matrix:

$$Y = [y_1, y_2, \dots, y_n] \in \mathbb{R}^{m \times n} \quad (\text{B.1})$$

The goal of POD is to find an orthonormal basis  $\{\psi_i\}_{i=1}^\ell$  of rank  $\ell \ll m$  that is optimal in the sense that it maximizes the captured energy (variance) of the snapshots. This is equivalent to minimizing the mean squared projection error [7, 8]:

$$\min_{\{\psi_i\}_{i=1}^\ell} \sum_{j=1}^n \left\| y_j - \sum_{i=1}^\ell \langle y_j, \psi_i \rangle \psi_i \right\|^2 \quad (\text{B.2})$$

This optimization problem is solved by finding the eigenvectors of the spatial correlation matrix  $K = YY^T \in \mathbb{R}^{m \times m}$ . The POD basis vectors  $\psi_i$  are the eigenvectors of  $K$  corresponding to the largest eigenvalues  $\lambda_i$ .

For typical CFD problems,  $m \gg n$ , making the eigenvalue problem for the  $m \times m$  matrix  $K$  computationally intractable. The \*\*method of snapshots\*\* circumvents this by solving the eigenvalue problem for the much smaller temporal correlation matrix  $C = Y^T Y \in \mathbb{R}^{n \times n}$ . [5, 6]

The connection between these two problems is established through Singular Value Decomposition (SVD). The SVD of the snapshot matrix  $Y$  is given by [7, 9]:

$$Y = U \Sigma V^T \quad (\text{B.3})$$

where  $U \in \mathbb{R}^{m \times m}$  and  $V \in \mathbb{R}^{n \times n}$  are orthogonal matrices, and  $\Sigma \in \mathbb{R}^{m \times n}$  is a diagonal matrix of singular values  $\sigma_i \geq 0$ .

The key relationships are:

- $YY^T = U \Sigma \Sigma^T U^T$

- $Y^T Y = V \Sigma^T \Sigma V^T$

From this, we can see:

- The \*\*POD modes\*\* (spatial basis vectors)  $\{\psi_i\}$  are the columns of  $U$ . They are the eigenvectors of the spatial correlation matrix  $Y Y^T$ . [7, 9]
- The eigenvalues of  $Y Y^T$  are the squares of the singular values,  $\lambda_i = \sigma_i^2$ .
- The eigenvectors of the temporal correlation matrix  $Y^T Y$  are the columns of  $V$ .

Therefore, we can compute the POD modes efficiently from the eigenvectors of the smaller matrix  $Y^T Y$ .

Once the truncated POD basis  $\Psi_\ell = [\psi_1, \dots, \psi_\ell] \in \mathbb{R}^{m \times \ell}$  is obtained, any snapshot can be approximated by its projection onto this basis:

$$y_j \approx \Psi_\ell a_j \tag{B.4}$$

where  $a_j \in \mathbb{R}^\ell$  is the vector of temporal coefficients for snapshot  $y_j$ , computed by projection:

$$a_j = \Psi_\ell^T y_j \tag{B.5}$$

In matrix form, the matrix of all temporal coefficients is  $A = \Psi_\ell^T Y = \Sigma_\ell V_\ell^T$ , where  $\Sigma_\ell$  and  $V_\ell$  are the truncated versions of  $\Sigma$  and  $V$ . [10]



## Mathematical Formulation of Kriging

Kriging, or Gaussian Process Regression (GPR), is a statistical interpolation method that provides a prediction and a measure of uncertainty for that prediction.[11, 12] It models the unknown function of interest,  $f(x)$ , as a realization of a Gaussian Process (GP) [13]:

$$f(x) \sim \mathcal{GP}(m(x), k(x, x')) \quad (\text{C.1})$$

where  $m(x)$  is the mean function and  $k(x, x')$  is the covariance function (or kernel), which defines the correlation between function values at different points.

In this work, Ordinary Kriging is used, which assumes a constant but unknown mean,  $m(x) = \mu$ . [11, 14] Given a set of  $n$  observations  $y = [f(x_1), \dots, f(x_n)]^T$  at input locations  $X = \{x_1, \dots, x_n\}$ , we want to predict the function value  $f_*$  at a new point  $x_*$ .

Under the GP assumption, the joint distribution of the observed values  $y$  and the value to be predicted  $f_*$  is a multivariate Gaussian:

$$\begin{pmatrix} y \\ f_* \end{pmatrix} \sim \mathcal{N} \left( \begin{pmatrix} \mu \mathbf{1} \\ \mu \end{pmatrix}, \begin{pmatrix} K & k_* \\ k_*^T & k_{**} \end{pmatrix} \right) \quad (\text{C.2})$$

where  $\mathbf{1}$  is a vector of ones,  $K$  is the  $n \times n$  covariance matrix of the training points with entries  $K_{ij} = k(x_i, x_j)$ ,  $k_*$  is the  $n \times 1$  vector of covariances between the training points and the new point,  $k_* = [k(x_1, x_*), \dots, k(x_n, x_*)]^T$ , and  $k_{**} = k(x_*, x_*)$  is the variance at the new point.

The Kriging predictor  $\hat{f}(x_*)$  is the mean of the conditional (posterior) distribution  $p(f_* | X, y, x_*)$ . For Ordinary Kriging, the Best Linear Unbiased Predictor (BLUP) is given by [15, 16]:

$$\hat{f}(x_*) = \hat{\mu} + k_*^T K^{-1} (y - \hat{\mu} \mathbf{1}) \quad (\text{C.3})$$

where  $\hat{\mu}$  is the estimated constant mean, often calculated as:

$$\hat{\mu} = \frac{\mathbf{1}^T K^{-1} y}{\mathbf{1}^T K^{-1} \mathbf{1}} \quad (\text{C.4})$$

A key advantage of Kriging is its ability to provide a measure of prediction uncertainty. The variance of the posterior distribution, known as the Kriging variance or mean squared

error (MSE), is given by [17]:

$$\sigma^2(x_*) = k_{**} - k_*^T K^{-1} k_* + \frac{(1 - \mathbf{1}^T K^{-1} k_*)^2}{\mathbf{1}^T K^{-1} \mathbf{1}} \quad (\text{C.5})$$

The covariance function  $k(x, x')$  depends on hyperparameters (e.g., correlation length) which are typically estimated from the training data by maximizing the log-marginal likelihood.[13, 16]