ISTANBUL TECHNICAL UNIVERSITY ★ GRADUATE SCHOOL

VISION TRANSFORMER-BASED PHYSICS INFORMED CFD PREDICTION OF AXIAL FANS WITH SELF-SUPERVISED CONTRASTIVE LEARNING FOR ENHANCED GEOMETRIC SENSITIVITY

M.Sc. THESIS

Mehmet Fatih REYHAN

Department of Aeronautics and Astronautics Engineering

Aeronautics and Astronautics Engineering Programme

JUNE 2025

ISTANBUL TECHNICAL UNIVERSITY ★ GRADUATE SCHOOL

VISION TRANSFORMER-BASED PHYSICS INFORMED CFD PREDICTION OF AXIAL FANS WITH SELF-SUPERVISED CONTRASTIVE LEARNING FOR ENHANCED GEOMETRIC SENSITIVITY

M.Sc. THESIS

Mehmet Fatih REYHAN
(511221122)

Department of Aeronautics and Astronautics Engineering

Aeronautics and Astronautics Engineering Programme

Thesis Advisor: Assoc. Prof. Dr. Baha ZAFER

JUNE 2025

İSTANBUL TEKNİK ÜNİVERSİTESİ ★ LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ

EKSENEL FANLARDA GEOMETRİK FARKLILIKLARIN YAKALANMASINA YÖNELİK KONTRASTİF ÖĞRENME DESTEKLİ, FİZİK BİLİNÇLİ VISION TRANSFORMER TABANLI CFD TAHMİNİ

YÜKSEK LİSANS TEZİ

Mehmet Fatih REYHAN
(511221122)

Uçak ve Uzay Mühendisliği Anabilim Dalı

Uçak ve Uzay Mühendisliği Programı

Tez Danışmanı: Doç. Dr. Baha ZAFER

HAZİRAN 2025

Mehmet Fatih REYHAN, a M.Sc. student of ITU Graduate School student ID 511221122 successfully defended the thesis entitled “VISION TRANSFORMER-BASED PHYSICS INFORMED CFD PREDICTION OF AXIAL FANS WITH SELF-SUPERVISED CONTRASTIVE LEARNING FOR ENHANCED GEOMETRIC SENSITIVITY”, which he/she prepared after fulfilling the requirements specified in the associated legislations, before the jury whose signatures are below.

Thesis Advisor : **Assoc. Prof. Dr. Baha ZAFER** ....
Istanbul Technical University

Jury Members : **Dr. Saliha Banu YILMAZ** ....
Istanbul Technical University

Assoc. Prof. Dr. Ümit Güneş ....
Yıldız Technical University

....

Date of Submission : 30 May 2025
Date of Defense : 27 June 2025

To my family,

FOREWORD

I would like to express my deepest gratitude to my mentor Yavuz Selim Mehel for providing the vision that shaped this work, to my dear friend Atalay Çelik for always answering my questions and sharing his insight in the field of deep learning, to my wife for her unwavering patience and encouragement, to my grandmother for her heartfelt prayers that accompanied me throughout this journey and to my advisor Assoc. Prof. Baha Zafer for his valuable support.

June 2025 Mehmet Fatih REYHAN

TABLE OF CONTENTS

ABBREVIATIONS

CFD : Computational Fluid Dynamics
CNN : Convolutional Neural Network
DNN : Deep Neural Network
FiLM : Feature-wise Linear Modulation
IGV : Inlet Guide Vane
LSTM : Long Short-Term Memory
MLP : Multi-Layer Perceptron
NGO : Northern Goshawk Optimization
NT-Xent : Normalized Temperature-Scaled Cross-Entropy
OGV : Outlet Guide Vane
RPM : Revolutions Per Minute
TBNN : Tensor Basis Neural Network
ViT : Vision Transformer

SYMBOLS

τ : NT-Xent Loss Temperature
$\mathbf{Q}$: Query Tensor
$\mathbf{K}$: Key Tensor
$\mathbf{V}$: Value Tensor
$\mathbf{B}$: Batch Size
$\mathbf{r}$: Radial Coordinate
θ : Angular Coordinate
$\mathbf{z}$: Axial Coordinate
$\mathbf{M}$: Mask Tensor, Mach Number
γ : Scaling Parameter, Specific Heat Ratio
β : Shifting Parameter
$\mathbf{z}_j$: Latent Vector
$\mathscr{L}$: Loss
P_t : Absolute Total Pressure
P_s : Static Pressure
T_t : Absolute Total Temperature
T_s : Static Temperature
C_d : Discharge Coefficient

LIST OF TABLES

Page

LIST OF FIGURES

VISION TRANSFORMER-BASED PHYSICS INFORMED CFD PREDICTION OF AXIAL FANS WITH SELF-SUPERVISED CONTRASTIVE LEARNING FOR ENHANCED GEOMETRIC SENSITIVITY

SUMMARY

In recent years, deep learning has emerged as a transformative approach in aerodynamic modeling, offering efficient alternatives to Computational Fluid Dynamics methods, particularly for design-intensive applications in turbomachinery. This thesis proposes a physics-informed deep learning framework for predicting aerodynamic performance parameters across multi-stage fan geometries, leveraging the Vision Transformer (ViT) architecture for its capacity to model complex spatial patterns and geometric variability.

The proposed architecture is composed of an encoder-decoder structure specifically designed for turbomachinery geometries. The encoder utilizes a ViT backbone, where input geometries are divided into non-overlapping patches and embedded into high-dimensional feature vectors via convolutional patch embedding. Positional encoding is added to preserve spatial structure, followed by multiple self-attention layers that extract long-range geometric dependencies. To account for varying stage configurations, attention masking and a learnable inlet station token are integrated into the encoder pipeline.

The decoder begins with an attention-based upsampler that increases the spatial resolution of the latent representation, followed by a boundary conditioning module that modulates the geometric features using Feature-wise Linear Modulation (FiLM) for inlet conditions and cross-attention for RPM and outlet pressure. The final stage of the decoder, referred to as the projector, consists of multi-layer perceptrons that reduce the feature dimensionality and map the encoded information to the target output space. The architecture is carefully designed to ensure that both spatial and physical structures are retained throughout the transformation.

To enhance geometric sensitivity and avoid overfitting to global input–output mappings, the encoder is first pretrained using contrastive learning with the NT-Xent loss. This step enables the model to distinguish subtle geometric variations, which are critical in turbomachinery design. The decoder is then fine-tuned using supervised learning, guided by a combination of a percentage-based loss and two physics-based losses: a continuity loss enforcing mass conservation, and an isentropic loss ensuring thermodynamic consistency between predicted flow variables.

The model is evaluated across a comprehensive CFD dataset covering multiple speedlines and six distinct fan designs, including variations in metal angle and hub radius. The results demonstrate that the model successfully captures both radial and streamwise flow behaviors, performs reliably across off-design operating conditions, and accurately simulates performance trends in response to design modifications.

Notably, even with subtle geometric changes, the model reflects physically consistent shifts in aerodynamic quantities.

This work highlights the efficacy of integrating self-supervised learning with physics-based constraints in achieving generalizable, computationally efficient aerodynamic models. The approach offers significant promise as a surrogate model for accelerating turbomachinery design cycles and supporting early-stage performance evaluations under varying geometric and operational scenarios.

EKSENEL FANLARDA GEOMETRİK FARKLILIKLARIN YAKALANMASINA YÖNELİK KONTRASTİF ÖĞRENME DESTEKLİ, FİZİK BİLİNÇLİ VISION TRANSFORMER TABANLI CFD TAHMİNİ

ÖZET

Bu tez çalışmasında, çok kademeli fan geometrilerinin aerodinamik performansını hızlı ve fiziksel olarak tutarlı bir şekilde tahmin etmek amacıyla fizik bilinciyle yapılandırılmış bir derin öğrenme modeli geliştirilmiştir. Hesaplamalı Akışkanlar Dinamiği yöntemleri, özellikle yüksek çözünürlüklü ve doğruluğu yüksek çözümler sağlamak açısından güçlü olsa da, her tasarım iterasyonunda baştan çözüm alınmasının getirdiği yüksek hesaplama maliyeti, bu yöntemlerin pratik uygulamalarda sınırlı kalmasına neden olmaktadır. Özellikle turbomakinelerde farklı devirlerde ve koşullarda (stall/choke) hesaplamaların gerekliliği yüksek bir hesaplama maliyeti oluşturmaktadır. Bu nedenle, hem doğruluğu hem de hesaplama verimliliği hedefleyen yapay zeka tabanlı tahmin sistemlerine olan ihtiyaç giderek artmaktadır.

Literatürde son yıllarda turbomakine alanında derin öğrenme tabanlı birçok çalışma yapılmıştır. Bu çalışmalar genellikle belirli bir geometrik konfigürasyon veya sınırlı bir operasyonel rejim için özel olarak geliştirilmiş modeller sunmaktadır. Kimi çalışmalar doğrudan akış alanı tahmini üzerine odaklanırken, kimileri ise basınç kaybı ya da deviasyon gibi türetilmiş parametrelerin tahminine yönelmiştir. Bununla birlikte, bu modellerin büyük çoğunluğu yalnızca sınırlı sayıda girişe ve dar bir tasarım aralığına sahip olup, farklı sınır koşullarına ve tasarımsal varyasyonlara karşı genelleme yapma kapasitesine sahip değildir. Literatürde, özellikle tasarım iterasyonlarını yönlendirecek; farklı geometrilerle, çoklu sınır koşulları altında çalışabilen ve fiziksel anlamda genelleştirilebilir bir model eksikliği göze çarpmaktadır. Bu çalışma, söz konusu boşluğu doldurmayı hedeflemekte ve fiziksel tutarlılığı sağlayabilen, geometrideki değişimlere karşı hassas ve sınır koşullarına göre doğru sonuçlar verebilen genel amaçlı bir tahmin modeli geliştirmeyi amaçlamaktadır.

Modelin eğitimi ve değerlendirmesi için kullanılan veri seti, NUMECA yazılımı ile daha önce TEI bünyesinde çözülmüş toplam 1600 adet CFD sonucundan oluşturulmuştur. Veri seti, farklı RPM değerlerinde, iki ve üç kademeli fan konfigürasyonlarında, değişen çıkış basınçları ve farklı geometrilerdeki eksenel fan çözümlerini içermektedir. Bazı konfigürasyonlarda IGV ve OGV bulunurken, bazılarında bu elemanlar yer almamaktadır. Bu çeşitlilik sayesinde hem tasarım noktası hem de tasarım noktası dışındaki koşullarda modelin genelleme yeteneği sınanabilmiştir.

CFD verileri işlenirken, her rotor ve stator kanatçıklarının önünde ve arkasında tanımlanan akış düzlemlerinden, radyal yönde 15 adet streamtube oluşturulmuştur. Bu streamtube düzlemleri, radyal yönde eşit kütlesel debi taşıyacak şekilde oluşturulmuş

ve her bir kesitten toplam ve statik basınç, sıcaklık, Mach sayısı gibi fiziksel parametreler çıkarılmıştır. Bu yapılandırma, modelin hem spanwise hem de streamwise yönde anlamlı çıkışlar üretmesini mümkün kılmıştır.

Geometri verileri, endüstride yaygın olarak kullanılan geomTurbo formatında elde edilmiştir. Bu format, her rotor veya stator kanadının basınç ve emme yüzeylerine ait noktasal koordinatlarını içermektedir. Başlangıçta Kartezyen koordinat sisteminde tanımlanmış olan bu veriler, modelin turbomakinalardaki simetrik yapıyı daha etkin öğrenebilmesi amacıyla kutupsal koordinat sistemine dönüştürülmüştür. Elde edilen koordinatlar, her bir kanat yüzeyi için radyal, açısal ve eksensel bileşenleri içerecek şekilde yapılandırılmıştır. Böylece her bir geometri örneği, suction ve pressure yüzeyleri için toplam 6 kanallı bir tensör olarak temsil edilmiştir.

Modelin girişinde kullanılacak geometrik tensörlerin boyutu sabit tutulabilmesi için, maksimum üç kademe ve bir IGV (yani yedi rotor/stator sırası) içerecek şekilde yapılandırılmıştır. Veri setinde yer alan bazı örneklerin yalnızca iki ya da üç kademeye sahip olması nedeniyle, eksik kalan kısımlar sıfırlarla doldurulmuş, yani padding işlemi uygulanmıştır. Ancak verilerin bu yapay kısımlarının, modelin öğrenme sürecini olumsuz etkilememesi ve dikkat mekanizması içinde hesaba katılmaması için, attention mask adı verilen özel bir maskeleme stratejisi kullanılmıştır. Bu strateji sayesinde, sadece gerçek (yani fiziksel olarak mevcut) koordinatlar birbirleriyle bilgi alışverişi yapabilirken; doldurulmuş (sıfırlanmış) kısımlar tamamen izole edilmiştir. Böylece model, sadece geçerli geometrik içerik üzerinden öğrenme gerçekleştirir ve sahte girişlerin eğitimi bozmasının önüne geçer.

Modelin mimarisi olarak iki ana yapıdan oluşmaktadır: encoder ve decoder. Encoder tarafında, Vision Transformer (ViT) mimarisinden esinlenilmiş özel bir yapı kullanılmaktadır. Giriş geometrisi, önce 3 kanat kesidini ve her kanat yüzeyi için 225 noktayı kapsayacak şekilde parçalara (patch) bölünmektedir. Her bir parça, 3x225 boyutundaki bir bölgeyi temsil eder ve bu bölgeden 192 boyutlu bir öznitelik vektörü çıkarılmaktadır. Tüm geometri, bu şekilde 35 perçaya bölünmekte (5 span x 7 sıra) ve bu parçalar ardışık bir sırayla Transformer katmanlarına verilmektedir. Bu katmanlar içerisinde hem dikkat mekanizmaları, hem normalizasyon katmanları hem de atlama bağlantıları (skip-connection) bulunmaktadır.

Encoder'ın iç yapısında, her bir 3x225'lik tensörü temsil eden parçaların vektörlerei çok başlı dikkat (multi-head self-attention) ve beslemeli ileri ağlar (feed-forward networks) içeren katmanlar boyunca işlenmektedir. Bu yapılar, her bir geometrik parçanın diğer tüm parçalarla olan ilişkisini modelleyerek, uzun menzilli ve çapraz etkileşimleri yakalayabilmektedir. Attention mask, bu katmanlar içinde doğrudan dikkat skorlarına uygulanmakta; geçersiz (padding ile doldurulmuş) parçaların dikkate alınmaması sağlanmaktadır. Bu, modelin sadece fiziksel olarak mevcut geometri bölümleri arasında ilişki kurmasına olanak tanır.

Buna ek olarak encoder mimarisine, standart Vision Transformer mimarisinden farklı olarak öğrenilebilir bir giriş istasyonu (learnable inlet station) adlı bir vektör eklenmiştir. Bu vektör, giriş koşullarına karşılık gelen sabit bir istasyonu temsil eder ve eğitim sürecinde öğrenilir. Encoder tarafından çıkarılan tüm parçalarla birlikte bu özel vektör de işlenerek, modelin başlangıç akış koşullarını yapısal olarak anlamlandırmasına katkı sağlanır. Böylece, geometrik temsile fiziksel anlam katılmış olur.

Decoder tarafında üç önemli yapı yer almaktadır. İlk olarak, dikkat tabanlı bir upsampler modülü ile 40 parçadan (5 span x 8 istasyon) oluşan encoder çıktısı, 120 parçaya çıkarılarak 15 span × 8 istasyonluk bir yapı oluşturulmaktadır. Daha sonra, Boundary Conditioning modülü devreye girmekte; burada giriş koşulları (örneğin RPM, giriş toplam basınç ve sıcaklık) FiLM (Feature-wise Linear Modulation) yöntemi ile modele işlenmekte, çıkış basıncı gibi downstream koşullar ise cross-attention ile geometrik özelliklere bağlanmaktadır. Bu sayede model, sınır koşullarına duyarlı tahminler üretmekte ve speedline üzerindeki konumunu doğru yakalamayı (stall veya choke rejiminde olup olmadığını) öğrenmektedir. Son aşamada, beslemeli ileri ağlardan (feed-forward networks) oluşan projector modülü, 192 boyutlu vektörleri fiziksel çıktılara dönüştürmektedir. Elde edilen tensör, 7 farklı büyüklüğü temsil eden ve 15 span × 8 istasyonluk bir çıktıya dönüştürülmektedir.

Modelin eğitimi iki aşamalı olarak gerçekleştirilmiştir. İlk olarak encoder, kontrastif öğrenme yöntemi ile eğitilmiş; benzer geometriler yakın, farklı olanlar uzak temsillerle kodlanacak şekilde NT-Xent kaybı kullanılmıştır. Bu sayede model, minimal geometrik farkları bile ayrıştırabilir hâle gelmiştir. Ardından encoder dondurularak, decoder kısmı CFD verileri üzerinden eğitilmiştir. Eğitim sürecinde yalnızca yüzde bazlı hata dikkate alınmamış, aynı zamanda iki fizik temelli kayıp fonksiyonu entegre edilmiştir. Bunlardan ilki, süreklilik (continuity) ilkesine dayanan ve tüm istasyonlarda kütle korunumunu sağlayan bir hata fonksiyonudur. İkincisi ise izantropik akış koşullarından türetilmiş ve statik parametrelerden hesaplanan teorik toplam basınç ve sıcaklık ile modelin tahmin ettiği değerler karşılaştırılarak oluşturulmuştur. Modelin toplam kaybı, bu üç farklı kaybın toplamından oluşmuş ve eğitim, toplam hata oranı yüzde birin altına düştüğünde durdurulmuştur. Yüzde kayıba ek olarak eklenen bu iki fiziksel kayıp sayesinde modelin genelleştirme kabiliyeti artmıştır.

Modelin değerlendirme sürecinde hem doğrudan tahmin edilen hem de hesaplanan büyüklükler üzerinden analiz yapılmıştır. Yüksek, orta ve düşük hızlarda yapılan testlerde modelin, spanwise ve streamwise yönde hem doğruluğu hem de fiziksel eğilimleri başarıyla yakaladığı gösterilmiştir. Ayrıca model, düşük veri içeren stall ve choke bölgelerinde bile CFD ile tutarlı eğilimler göstermiştir. Bununla birlikte, speedline üzerindeki nokta tahminlerinde model, çalışma rejimini RPM ve çıkış basıncı değerine göre doğru bir şekilde konumlandırabilmiş ve beklenen eğilimleri yansıtmıştır.

Son olarak, modelin farklı geometrik varyasyonlara karşı duyarlılığı incelenmiştir. Küçük değişiklikler içeren ancak fiziksel anlamı büyük olan modifikasyonlarda, örneğin metal açı değişimi ve hub yarıçapındaki küçülme gibi durumlarda modelin çıktıları CFD ile aynı yönde eğilim göstermiştir. Bu da modelin yalnızca veri ezberleyen bir yapı değil, fiziksel neden-sonuç ilişkilerini öğrenebilen ve tasarım sürecinde karar destek aracı olarak kullanılabilecek bir yapı olduğunu ortaya koymaktadır.

Bu tez kapsamında geliştirilen fizik bilincine sahip, kontrastif ön-eğitimle desteklenmiş ViT tabanlı derin öğrenme modeli, turbomakine aerodinamiğinde hem doğruluk hem de hesaplama verimliliği açısından umut verici bir alternatif sunmaktadır.

Modelin gelecekte daha fazla veri setiyle eğitilmesi, fiziksel kısıtların çeşitlendirilmesi, endüstriyel uygulamalar için çok daha yetkin bir tahmin aracı ortaya çıkaracaktır.

1. INTRODUCTION

Deep learning has been utilized in a wide range of industries in recent years, from industrial automation and autonomous systems to healthcare and finance. This is also accurate in the field of aerodynamics, where complex fluid phenomena are increasingly being modeled using data-driven methodologies. Deep learning techniques have become an attractive alternative for conventional computational fluid dynamics (CFD) methods in the context of turbomachinery, where fluid dynamics are essential in determining performance and efficiency. These techniques have the potential to accelerate simulations, lower computational costs, and increase the accuracy of aerodynamic predictions, especially in scenarios with complex flow patterns and geometric variability.

CFD simulations have always been essential for simulating the aerodynamic performance of turbomachinery parts. Although high-fidelity methods such as Large Eddy Simulation and Direct Numerical Simulation offer comprehensive insights into flow dynamics, their high computational cost makes them unsuitable for industrial applications. Because of its relatively low computational cost, Reynolds-Averaged Navier-Stokes modeling is widely used in the aerospace industry. Computational power is a crucial limiting factor for achieving optimal designs, though numerous solutions are needed for design iterations during the design process. This challenge is particularly evident in turbomachinery applications, where performance and stability are ensured by multiple CFD simulations for both design and off-design conditions, as well as stall and surge evaluations. This challenge has motivated the integration of deep learning models to predict aerodynamical parameters while significantly reducing computational time.

1.1 Literature Review

Deep learning models in literature have demonstrated considerable potential in approximating flow fields, enabling faster predictions without sacrificing accuracy. For instance, Aulich et. al. demonstrates that deep neural networks (DNNs) can be effectively adapted for predicting complete flow solutions in turbomachinery applications [2]. By using xyz-coordinates from CFD mesh data along with rotational speed and boundary conditions, the DNN successfully predicts velocities, pressure, and density within the flow field. The model is trained using a small dataset of twenty random 3D fan stage designs and shows robust generalization capabilities, accurately predicting flow characteristics for both random and Pareto optimal designs, demonstrating its potential for rapid aerodynamic performance assessments. A recent study by Feng et. al. introduced a physics-informed deep learning cascade loss model aimed at enhancing the accuracy of aerodynamic loss predictions in turbomachinery [3]. The model effectively bridges the gap between empirical models and purely data-driven approaches by embedding physical mechanisms, such as velocity distribution within the boundary layer, as intermediate variables. The integration of physics-informed constraints significantly improves prediction accuracy, particularly at high Mach numbers and high incidence angles, achieving a 57.4% error reduction compared to traditional empirical models. Moreover, Hipple et. al. show an alternative approach involving the use of machine learning to predict compressor stall events in gas turbine systems [4]. Utilizing a Long Short-Term Memory (LSTM) neural network, the model analyzes multivariate time series data to identify patterns indicative of impending stall. Training on real-world compressor data demonstrated that the model could accurately forecast stall occurrences 5–20 ms before they happen, providing a valuable time window for preventive interventions. Li et. al. explore the application of deep learning techniques for instability prediction in axial flow compressors, particularly under distorted inflow conditions [5]. Utilizing a Long Short-Term Memory (LSTM) neural network optimized with the Northern Goshawk Optimization (NGO) algorithm, the model predicts stall and surge events based on dynamic pressure data collected from the compressor casing. Experimental validation shows that the NGO-LSTM model can predict stall signals at least 1 second in advance under both uniform and distorted inflow conditions, offering significant improvements in prediction accuracy compared to traditional empirical models. Furthermore, a data-driven method has been introduced by Liu et. al. to predict flow fields in compressor cascades by employing deep learning in combination with ℓ_1 regularization [6]. The approach focuses on calibrating inlet boundary conditions and turbulence model parameters using sparse experimental data to improve prediction accuracy. The results show a substantial reduction in prediction errors—approximately 70% compared to traditional methods—highlighting the effectiveness of integrating data-driven techniques with physical calibration in turbomachinery analysis. Bruni et. al. present a deep learning framework named C(NN)FD for predicting flow fields and aerodynamic performance in multi-stage axial compressors [7]. By employing a physics-based dimensionality reduction technique, the model transforms the regression problem from an unstructured to a structured one, significantly reducing computational complexity. The architecture is shown to achieve accuracy comparable to CFD benchmarks while providing real-time predictions, making it suitable for industrial applications involving gas turbine compressors. Aulich et. al. proposed a model leveraging transformer architectures to approximate 3D flow solutions for turbomachinery compressor designs [8]. By training on CFD-generated data, the model can rapidly predict flow fields, including velocity and pressure distributions, across different compressor configurations. The integration of self-attention mechanisms allows the model to effectively handle complex geometrical variations and operating conditions, significantly reducing computational time compared to traditional CFD simulations. Ji and Du introduce a tensor basis neural network (TBNN) model designed to improve turbulence modeling in transonic axial compressors [9]. By integrating the non-linear eddy viscosity approach within the TBNN framework, the model effectively captures anisotropic turbulence characteristics, addressing the limitations of traditional linear eddy viscosity models. The proposed $k-\omega$-SST-TBNN model, trained on high-fidelity large eddy simulation (LES) data, demonstrates significant improvements in predicting complex flow features such as tip-gap vortices and radial flow parameter distribution in

NASA Rotor 37, achieving higher accuracy and robustness compared to conventional turbulence models. Finally, Chen and Thuerey showed considerable potential in predicting compressible turbulent flows over aerofoils, generating Reynolds-averaged Navier-Stokes (RANS) solutions with high accuracy [10]. The use of structured grid encodings and body-fitted coordinate mappings enables the model to efficiently capture the flow field. The proposed model achieves great performance, accurately predicting pressure coefficients, skin friction coefficients, and wake total pressure profiles, thereby demonstrating its capability for rapid aerodynamic design optimization.

There are also studies that employ relatively new architectures, such as transformers, for aerodynamic modeling, leveraging their ability to capture long-range dependencies. One such example is the AeroDiT model, which uses diffusion transformers to accurately predict flow fields around airfoils at high Reynolds numbers [11]. By integrating scalable diffusion models with transformer architectures, AeroDiT effectively balances computational efficiency and accuracy, enabling real-time aerodynamic predictions. The model demonstrates significant improvements over conventional methods, achieving low relative L2 errors for pressure and velocity components while maintaining rapid inference speeds, making it suitable for real-time applications. Peng et. al. present another attention-enhanced neural network model, which has demonstrated remarkable potential in simulating turbulence, particularly by addressing the multi-scale and chaotic nature of high-Reynolds-number flows [12]. By incorporating visual attention mechanisms, this model automatically adjusts the weights to capture non-equilibrium regions, significantly improving the accuracy of turbulence predictions. Compared to conventional models like the Fourier Neural Operator (FNO), the attention-enhanced model achieves approximately 40% error reduction with only a 1% increase in computational cost, making it both efficient and robust for high-dimensional fluid dynamics simulations. Kang et. al. proposes a CFDformer model that utilizes a combination of Vision Transformer (ViT) and U-Net architectures to approximate fluid flow solutions efficiently [13]. By leveraging ViT for capturing global attention and U-Net for local feature extraction, CFDformer can handle both local and global flow characteristics simultaneously. The model demonstrates superior performance in predicting steady-state velocity fields compared to baseline models, achieving faster computational times while maintaining high accuracy, even for unseen initial conditions. On the other hand, GeoANF (Geometric Attention Neural Field) combines geometric attention mechanisms with implicit neural field representations to predict flow fields over airfoils efficiently [14]. By utilizing point cloud inputs rather than grid-based data, GeoANF generalizes well across different airfoil shapes and operating conditions. The model significantly outperforms baseline methods in tasks such as volume flow field prediction and surface pressure measurement, demonstrating its ability to handle geometric variability while maintaining high accuracy. Deng et. al. introduce a modified Vision Transformer (ViT)-based encoder-decoder network for predicting compressible flow over supercritical airfoils [15]. To enhance accuracy near shock regions, the model incorporates geometric encoding methods and employs advanced loss functions, including gradient distribution and multilevel wavelet transformation. Additionally, transfer learning is utilized to improve the model's generalization, allowing rapid adaptation to new airfoil geometries with minimal training data.

Although Vision Transformers (ViTs) were originally developed for image classification and other computer vision tasks, they have found significant applications in aerodynamic modeling, as demonstrated by their use in predicting flow fields over airfoils and capturing complex aerodynamic characteristics. The adaptability of ViTs, due to their self-attention mechanisms, allows them to be applied in diverse fields beyond fluid dynamics. For instance, ViTs have shown competitive performance in industrial visual inspection tasks, including defect detection and object classification, where they often outperform traditional convolutional neural networks (CNNs) even when the training data is limited [16]. Even though Vision Transformers (ViTs) have demonstrated remarkable potential in aerodynamic modeling by leveraging self-attention mechanisms, the exploration of other deep learning paradigms has also yielded promising results. In particular, self-supervised and unsupervised learning approaches are gaining traction for their ability to enhance model robustness and generalization without requiring large labeled datasets. Self-supervised learning involves training models using automatically generated labels derived from the data itself, allowing the model to learn meaningful representations without human annotation. In contrast, unsupervised learning focuses on discovering patterns and structures within the data without any label information, making it particularly useful when labeled data is scarce or unavailable. For instance, employing auxiliary tasks, such as rotation prediction, not only improves robustness but also enhances the model's ability to detect out-of-distribution samples, outperforming fully supervised approaches in challenging scenarios [17]. Furthermore, Chen et. al. demonstrates that the use of contrastive learning frameworks, such as SimCLR, effectively improves the representation quality by maximizing the agreement between positive pairs while distinguishing them from negative pairs [18]. This method has proven particularly effective in scenarios where labeled data is limited, highlighting the value of self-supervised techniques in extracting meaningful patterns from raw data. Additionally, data augmentation plays a crucial role in unsupervised learning setups, particularly when handling limited or noisy datasets. As shown in recent studies, employing systematic data augmentation strategies can significantly enhance predictive performance, as seen in EEG signal modeling, where augmentation reduced overfitting and improved the robustness of the trained models [19]. By incorporating self-supervised and unsupervised learning techniques, deep learning models not only gain resilience against data variability but also achieve higher generalization, making them more adaptable to real-world applications. Consequently, aerospace studies have also adopted self-supervised and unsupervised techniques to enhance model robustness and efficiency. Drikakis et. al. demonstrate the use of self-supervised transformers for turbulent flow time series, highlighting how self-supervision improves model performance by allowing the network to learn from unlabeled time-series data [20]. This approach significantly enhances the ability to predict turbulent flow characteristics, particularly in high-speed regimes involving shock-boundary layer interactions. The model's capacity to reconstruct unsteady pressure signals demonstrates the advantage of leveraging self-supervised learning in fluid dynamics modeling. Kim et al present another notable study that applies unsupervised learning for the super-resolution reconstruction of turbulent flows, addressing the challenge of limited high-resolution data availability [20]. The proposed method utilizes a CycleGAN architecture to reconstruct high-resolution flow fields from low-resolution or unpaired turbulence data. This approach effectively bridges the gap between low-quality input and high-fidelity output, demonstrating that unsupervised learning can achieve comparable performance to supervised methods even when paired data is not available. De Fabritiis and Gryllias introduced an effective approach involving the use of self-supervised learning for anomaly detection in rotating machinery, particularly focusing on cycle-consistency representation learning to train an embedder network on univariate unlabeled data [21]. This method enforces a distance metric within the embedding space, enabling the model to distinguish between normal and anomalous vibration signals. The results indicate that this technique is highly effective for early fault detection in bearings, offering significant potential in predictive maintenance and condition monitoring applications. Tucci et. al. presents an unsupervised machine learning approach for high-fidelity compression of large experimental datasets related to HPT blade tip contouring. The method combines dimensionality reduction techniques such as Principal Component Analysis (PCA) and Autoencoders (AEs) to compress experimental measurements while retaining critical aerodynamic features [22]. This approach demonstrates that unsupervised learning can effectively correlate complex tip geometries with aero-thermodynamic performance, significantly enhancing the design optimization process for high-pressure turbines.

1.2 Purpose of Thesis

As revealed by the literature review, there is significant diversity in the application of deep learning methods within turbomachinery and aerodynamic modeling. Some studies focus on advanced attention-based architectures, which are particularly effective at capturing long-range dependencies and spatial interactions. In contrast, other studies employ more traditional neural network models, which tend to perform well in structured or simpler flow scenarios. Furthermore, there is considerable variation in the objectives of these models: while some aim to directly predict detailed flow fields, others focus on estimating derived performance metrics such as pressure loss or aerodynamic efficiency. Additionally, time-series-based models are frequently employed to capture dynamic behaviors, especially when real-time prediction or monitoring of transient flow phenomena is required. Recent studies have also shown an increasing trend in the adoption of self-supervised methods, particularly in cases where accurately capturing geometric differences is critical.

In this context, the present study aims to develop a model for predicting the radial profiles of fan geometries using the Vision Transformer (ViT) architecture. The choice of ViT is motivated by its capacity to efficiently process complex spatial patterns, which is crucial when analyzing fan blade profiles. The ViT's ability to model geometric variations and spatial coherence enhances the model's performance, especially when subtle differences can have significant aerodynamic impacts. Additionally, to further improve the model's sensitivity to geometric variations, self-supervised learning techniques are incorporated during the pretraining phase. This integration enables the model to accurately capture minor geometric differences that may influence aerodynamic performance.

One of the key aspects of the model's design is its adaptability to varying operational conditions. The architecture is specifically designed to ensure accurate predictions not only across different rotational speeds (RPMs) but also in challenging flow regimes such as stall and choke. This robustness is achieved through a thoughtful integration of domain knowledge into the training process, ensuring the model's ability to generalize effectively across diverse aerodynamic scenarios. The developed model is expected to provide reliable and precise predictions of fan geometry profiles, contributing to more efficient and accurate aerodynamic analysis in turbomachinery applications.

2. DATA PREPARATION

2.1 Database

The dataset consists of 1,600 separate CFD results that cover a variety of operating conditions and design configurations. These include different rotational speeds, blade geometries, number of rows (rotor and stator stages), and downstream static pressures, selected to align with specific speedlines. All simulations were performed using the NUMECA software, which was used throughout the entire process including mesh generation, solving, and post-processing.

2.1.1 CFD solutions

Mesh resolution was set to ensure y^+ values remained below 5, providing adequate near-wall accuracy for turbulence modeling. The Spalart–Allmaras turbulence model was chosen for all cases due to its balance between computational efficiency and suitability for turbomachinery applications.

As for boundary conditions, rotational speed (RPM), inlet total pressure and temperature, and outlet static pressure were defined for each case. These physical parameters, together with the corresponding blade geometry, were used as input to the deep learning model. This combined input allows the model to learn the relationship between both the operating conditions and the geometric configuration of the turbomachinery component.

This dataset includes 2-stage and 3-stage fan configurations, some of which feature Inlet Guide Vanes (IGV) and Outlet Guide Vanes (OGV). While a few geometries are represented by a single operating point at one specific speed, others have full speedline solutions computed at multiple RPMs. These speedline solutions are particularly important for capturing critical behaviors near stall and choke regions, which are essential for accurate performance prediction.

Approximately 70% of the simulations were carried out under sea-level conditions, while the remaining cases were generated using boundary values corresponding to specific altitude levels. The dataset covers a wide range of aerodynamic regimes, including both subsonic and transonic flow conditions, which enables the model to generalize across different physical behaviors.

It is important to note that this dataset was not intentionally generated for the purpose of this project. Instead, the CFD results were originally obtained throughout various design processes, reflecting realistic and application-oriented operating scenarios. The deep learning model was developed to make use of this existing pool of data, aiming to extract meaningful patterns and enable performance prediction without the need for additional costly simulations.

2.1.2 Streamtube calculation

In order to capture the flow characteristics in a more localized and physically meaningful manner, especially along the blade span, a custom post-processing algorithm was developed. For each CFD case, the flow planes located upstream and downstream of every blade row were automatically identified. These planes were then divided into 15 radial sections such that each section carried an equal portion of the total mass flow. During this process, an iterative approach was followed to determine the radial boundaries that satisfy the equal mass flow condition across the span. This mass-balanced segmentation enables a streamtube-based analysis that reflects the actual distribution of the flow. After the segmentation, key physical quantities (Total and Static Pressure and Temperatures, Absolute and Relative Mach Numbers, Discharge Coefficient) were extracted from each streamtube plane and stored for use in the model training process.

2.1.3 Blade geometry tensor construction

The blade geometries were exported in the standard `geomTurbo` format. The geometry dataset contains the Cartesian coordinates of the suction and pressure sides of the blades, as well as axial-radial coordinates of the hub and shroud surfaces. Prior

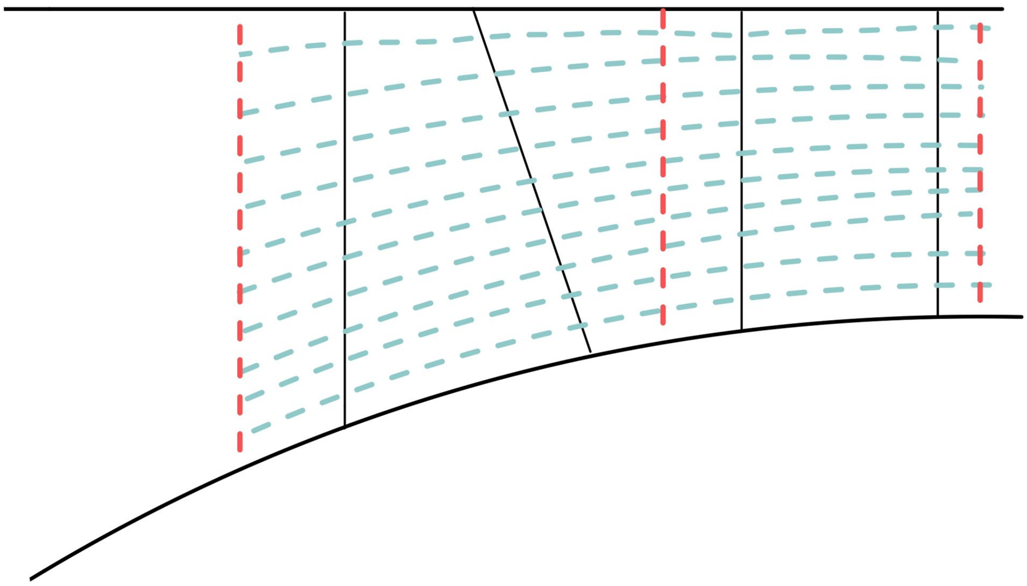


Figure 2.1: The representation of streamtubes(shown in blue) and rotor-stator planes from which the data are acquired(shown in red).

to tensor construction, all coordinate data were transformed into the polar coordinate system for better alignment with the rotational symmetry of the system.

The final input tensor representing blade geometry consists of 6 channels:

- R, θ, and Z coordinates for the suction side (3 channels),
- R, θ, and Z coordinates for the pressure side (3 channels).

In addition to the geometric tensor, two row-specific vectors are defined for each blade row:

- a vector indicating whether the row is a **rotor** (1) or a **stator** (-1),
- a vector specifying the **number of blades** in each row.

These vectors are later integrated into the model pipeline to provide context-specific information during the learning process.

The height of the input tensor corresponds to the number of spanwise sections (selected as 15), and the width is determined by the number of points exported per suction and pressure side (e.g., 225), multiplied by the number of rows. For instance, if each blade side contains 225 points and the geometry consists of 4 rows, the total width becomes $4 \times 225 = 900$.

In this study, the maximum number of blade rows is set to 7. Therefore, the maximum width of the input tensor becomes $7 \times 225 = 1575$. For geometries with fewer than 7 rows, zero-padding is applied to the end of the tensor to maintain uniform input dimensions. The padding strategy and the way this row-specific information is integrated into the model are detailed in the corresponding section on *Model Architecture*.

3. MODEL ARCHITECTURE

The proposed model consists of two main components: an encoder and a decoder. The encoder is responsible for extracting meaningful representations from the input geometrical data, while the decoder predicts the corresponding flow field variables based on these learned representations and boundary conditions. The encoder is based on the Vision Transformer (ViT) architecture, adapted to the specific requirements of turbomachinery geometries, whereas the decoder employs a physics-informed upsampling mechanism integrated with boundary condition conditioning.

3.1 Encoder

3.1.1 Attention mechanism

At the core of the encoder lies the self-attention mechanism, originally introduced in the Transformer architecture by Vaswani et. al. [23]. Self-attention allows the model to dynamically attend to different parts of the input sequence when computing representations, capturing long-range dependencies without relying on fixed convolutional receptive fields.

Given a sequence of input embeddings $X \in \mathbb{R}^{L\times d}$, the self-attention operation computes query (Q), key (K), and value (V) matrices as follows:

$$Q = XW^Q, \quad K = XW^K, \quad V = XW^V \tag{3.1}$$

where $W^Q, W^K, W^V \in \mathbb{R}^{d\times d}$ are learnable weight matrices. The attention output is then computed as:

$$\text{Attention}(Q,K,V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d}}\right)V \tag{3.2}$$

This formulation enables the model to assign dynamic importance scores between different patches, facilitating rich and flexible feature extraction from the geometrical data.

3.1.2 Vision transformer-based encoder

The encoder follows the general structure of the Vision Transformer [1] with several adaptations for the problem at hand. Unlike the original ViT, no class token is appended to the sequence. Instead, the model directly processes only the embedded geometrical patches.

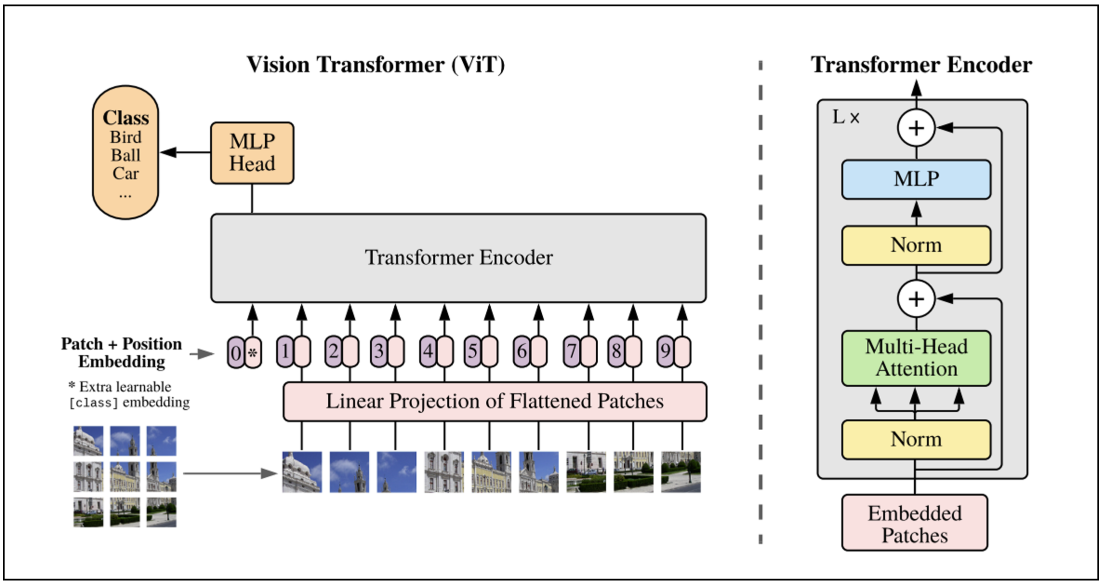


Figure 3.1: The original Vision-Transformer Architecture [1].

3.1.2.1 Patch embedding

The input geometrical tensor, representing the suction and pressure surfaces of multiple stages, is initially organized with dimensions $B \times 6 \times 15 \times 1575$. Here, B denotes the batch size, 6 represents the input channels corresponding to the radial, circumferential, and axial coordinates (r, θ, z) for both suction and pressure sides, while 15 and 1575 correspond to the radial and axial dimensions, respectively.

To effectively embed the geometrical information, the tensor is processed through a convolutional layer that acts as a patch embedding operation. This convolutional layer uses a kernel of size 3×225 and a stride of the same size, without padding, to ensure non-overlapping patch extraction. The kernel size is carefully chosen to capture spatial correlations within a patch while preserving the structural features along the radial (3) and axial (225) directions. In this way, the convolution operation itself divides the input geometrical tensor into non-overlapping patches.

The transformation from the input tensor to the embedded representation can be mathematically described as follows. The output dimensions after convolution can be calculated using the formula:

$$O = \left\lfloor \frac{I - K + 2P}{S} \right\rfloor + 1 \tag{3.3}$$

Here: - O is the output dimension, - I is the input dimension, - K is the kernel size, - P is the padding size, - S is the stride.

In our case, since the kernel size is 3×225, the stride is also 3×225, and no padding is applied ($P = 0$), the output dimensions along the radial and axial directions are calculated as follows:

$$O_r = \left\lfloor \frac{15 - 3 + 0}{3} \right\rfloor + 1 = 5 \tag{3.4}$$

$$O_a = \left\lfloor \frac{1575 - 225 + 0}{225} \right\rfloor + 1 = 7 \tag{3.5}$$

Therefore, the convolution layer transforms the original tensor shape from $B \times 6 \times 15 \times 1575$ to $B \times 192 \times 5 \times 7$, where 192 represents the number of output channels obtained through the embedding operation.

Following the convolution, the spatial dimensions are flattened to obtain the final shape of $B \times 35 \times 192$, where 35 is the total number of patches (5 x 7) arranged sequentially while preserving the embedded feature dimension. This embedding process ensures that each patch carries the local geometric information in a compact, 192-dimensional representation. The resulting sequence of embedded patches is then stacked to form the tensor Z of size $L \times d$, where $L = 35$ is the number of patches and $d = 192$ is the embedding dimension.

3.1.2.2 Positional embedding

To encode the spatial arrangement of patches, a learnable positional embedding matrix $P \in \mathbb{R}^{L \times d}$ is added:

$$Z_{\text{input}} = Z + P \tag{3.6}$$

This addition ensures that the model retains awareness of the relative locations of patches, a critical aspect for capturing geometric features influencing aerodynamic performance.

3.1.2.3 Transformer blocks

The embedded sequence Z_{input} is passed through six consecutive Transformer blocks, each composed of a multi-head self-attention module and a feed-forward network with layer normalization and residual connections.

3.1.2.4 Learnable inlet station

The model requires the output representation to include an additional station compared to the number of stages. To ensure consistency, a **learnable inlet station parameter**, denoted as (inlet_sta), is introduced. This parameter acts as an additional embedding and is trained to represent the inlet station independently from the encoded geometrical patches.

The final encoder output is obtained by concatenating the learnable inlet station with the encoded patches:

$$Z_{\text{out}} = [\text{inlet\_sta}, Z_{\text{encoder}}] \tag{3.7}$$

Where:

- $\left(\text{inlet\_sta} \in \mathbb{R}^{5\times d}\right)$ is the learnable embedding,
- $\left(Z_{\text{encoder}} \in \mathbb{R}^{35\times d}\right)$ is the encoded representation of the stages,
- The concatenation ensures that the output tensor (Z_{out}) has the correct shape $(40 \times d)$ for decoding.

This approach guarantees that the number of output stations matches the required configuration, maintaining geometric consistency in the model.

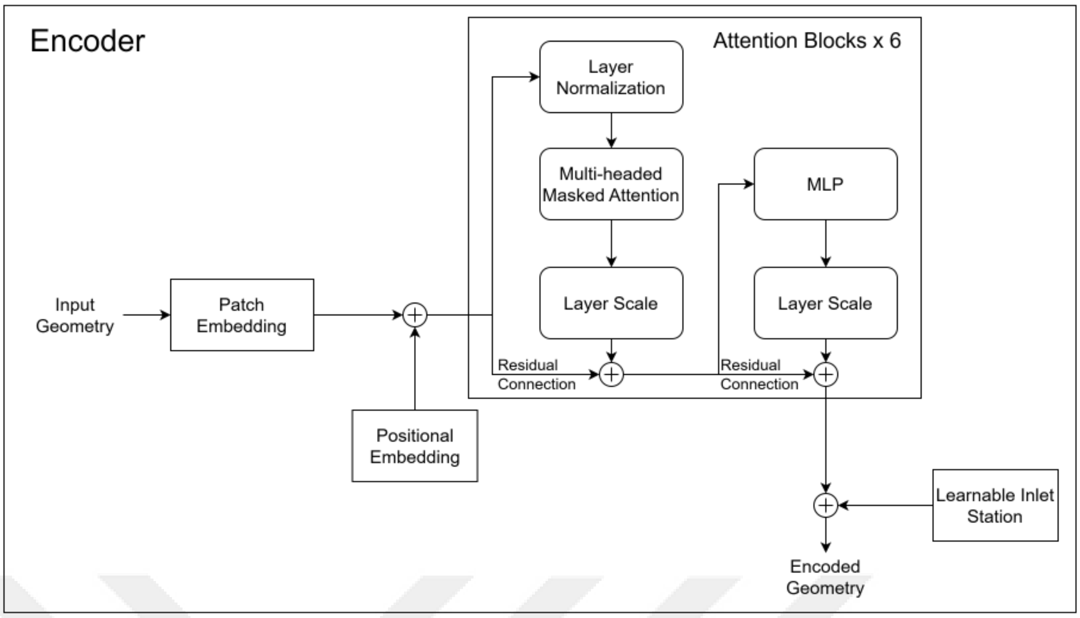


Figure 3.2: Vision Transformer employed Encoder Diagram.

3.1.3 Handling padding and attention masking

In practice, not all geometrical samples contain the same number of stages. To maintain consistent input shapes, samples with fewer than seven stages are padded by adding artificial rows. However, it is crucial to ensure that these padded regions do not influence the model's internal representations.

To achieve this, an attention mask $M \in \{0, -\infty\}^{L \times L}$ is applied to the attention score matrix before the softmax operation. Specifically, the raw attention logits are adjusted as:

$$\text{MaskedAttention}(Q,K,V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d}} + M\right)V \tag{3.8}$$

where the masking matrix M is defined as:

$$M_{ij} = \begin{cases} 0, & \text{if both patch } i \text{ and patch } j \text{ correspond to valid geometrical data,} \\ -\infty, & \text{if either patch } i \text{ or patch } j \text{ corresponds to padding.} \end{cases} \tag{3.9}$$

The role of the $-\infty$ values is to effectively eliminate any interaction involving padded regions during the attention computation. This behavior can be understood mathematically by examining the softmax function:

$$\text{softmax}(z_i) = \frac{e^{z_i}}{\sum_j e^{z_j}} \tag{3.10}$$

When a score z_i is set to $-\infty$, its exponentiated value $e^{-\infty}$ becomes zero. Consequently, after softmax normalization, any attention weight associated with padded regions becomes exactly zero:

$$e^{-\infty} = 0 \quad \Rightarrow \quad \text{softmax component} = 0 \tag{3.11}$$

Thus, padded patches neither contribute to nor receive information from other patches. This masking strategy ensures that the model focuses solely on the real geometrical features, maintaining the integrity and robustness of the learned representations.

Such masking approaches are widely used in Transformer-based architectures handling variable-length sequences, including natural language processing models like BERT [24].

3.2 Decoder

The decoder component of the proposed model is designed to generate physically consistent predictions from the encoded geometrical representations. Unlike the encoder, which focuses on capturing spatial features from the input geometry, the decoder's primary function is to map these features to flow field variables, incorporating both the geometric variations and the imposed boundary conditions. The challenge lies in accurately representing the complex aerodynamic behavior that arises from slight changes in geometry while ensuring the output adheres to physical laws. To address this, the decoder architecture is structured to maintain a fine balance between learned features and physics-based constraints, leveraging the encoded representations to produce flow predictions that are not only numerically accurate but also physically consistent.

3.2.1 Upsampler

The Attention Upsampler is responsible for increasing the spatial resolution of the encoded output, transforming the tensor from a shape of 40×192 to 120×192. This transformation is essential to align the encoder output with the expected structure of 8 stations and 15 radial values.

To preserve the distinction between rotor and stator stages, the encoded output is first augmented with a learnable rotor/stator embedding. This embedding acts as an additional feature channel, allowing the model to distinguish between different types of stages during the upsampling process.

After integrating the rotor/stator embedding, the upsampling is performed using an attention-based mechanism. The upsampler is designed to map the initial 40 spatial locations to 120 locations, ensuring that the increased resolution maintains spatial coherence between the input and the upsampled output.

By configuring the attention heads and their dimensionality appropriately, the model is able to generate three times the number of spatial locations while preserving the feature dimensionality. The resulting tensor after the upsampling process has the desired shape of 120×192, providing a higher resolution representation for the subsequent decoding stage.

This method ensures that the model effectively captures the geometric relationships inherent to the input while maintaining the correct number of spatial points required for downstream processing.

3.2.2 Boundary conditioning

The BC Conditioning module is essential for integrating boundary condition information into the latent representation obtained from the encoder. This integration is vital as the aerodynamic behavior of turbomachinery components is significantly influenced by inlet conditions, rotational speed, and outlet pressure. Therefore, capturing these influences within the latent space representation enhances the model's ability to generalize over varying operating conditions.

To achieve this integration, the BC Conditioning module employs a combination of **Feature-wise Linear Modulation (FiLM)** and **cross-attention mechanisms**. This approach effectively encodes both the inlet effects and the downstream influences from RPM and outlet static pressure.

3.2.2.1 FiLM modulation

Feature-wise Linear Modulation (FiLM), initially introduced by Perez et. al. , provides a flexible and efficient way to incorporate external conditioning signals into neural network representations [25]. In this context, the FiLM layer modulates the latent representation by generating scaling (γ) and shifting (β) parameters directly from the inlet boundary conditions. This technique ensures that the model can dynamically adjust the feature activations based on the input conditions.

In this model, the FiLM layer takes the inlet boundary condition vector, which contains RPM, inlet total pressure , and inlet total temperature , and projects it into two sets of parameters, γ and β, using learnable linear transformations:

$$\gamma = W_{\text{gamma}} \cdot \text{BC}_{\text{inlet}}, \quad \beta = W_{\text{beta}} \cdot \text{BC}_{\text{inlet}} \tag{3.12}$$

Here, W_{gamma} and W_{beta} are linear projection matrices specifically trained to extract the modulation factors from the boundary condition vector. These parameters are then applied to the encoded geometric features using the following modulation formula:

$$\text{modulated} = \gamma \odot \text{geom\_feats} + \beta \tag{3.13}$$

This formulation allows the model to directly influence the amplitude and bias of the latent features based on the inlet conditions. The key advantage of this method is its flexibility; the model can emphasize or attenuate specific spatial features depending on the input configuration. As shown by Perez et. al. , FiLM conditioning effectively adapts neural representations to external factors, making it highly suitable for scenarios where input conditions significantly affect the output [25].

3.2.2.2 Cross-attention for RPM and outlet conditioning

While the inlet boundary conditions are modulated through the FiLM layer, downstream conditions such as RPM and outlet static pressure require a different approach. These parameters influence the flow characteristics dynamically, often interacting with spatial patterns in the encoded features. To address this, the model utilizes a cross-attention mechanism.

The primary purpose of this cross-attention mechanism is to dynamically align the geometric representation with the operational point on the speedline. The position on the speedline is significantly influenced by the combination of RPM and outlet pressure, which together determine the operating regime of the turbomachinery. Therefore, accurately modeling the effects of these boundary conditions is crucial for capturing performance variations.

The cross-attention mechanism is configured to project the combined RPM and outlet pressure vector into the latent space dimensionality. This projection is performed using a learnable linear transformation, ensuring that the boundary condition vector is compatible with the geometric feature space. The resulting tokens are then expanded along the spatial dimension to match the encoder output size.

By leveraging cross-attention, the model effectively learns how the speedline position shifts in response to changes in RPM and back pressure. This formulation allows the network to dynamically adjust flow characteristics, particularly emphasizing regions where the aerodynamic performance is most sensitive to these boundary conditions. The boundary condition tokens act as both keys and values, while the modulated geometric features serve as queries, facilitating a flexible representation of speedline dynamics.

The final output from the BC Conditioning module has the same dimensionality as the upsampler output , but with enhanced boundary condition awareness. By combining FiLM-based scaling and shifting for inlet effects with cross-attention-based fusion for RPM and outlet effects, the model robustly encodes the boundary conditions for accurate aerodynamic predictions.

3.2.3 Projector

The final stage of the model architecture consists of a series of multi-layer perceptron (MLP) layers, collectively referred to as the projector. This module receives tensors of shape 120 × 192, where each of the 120 entries represents a spatial patch, and each patch is described by a 192-dimensional feature vector. The projector reduces the dimensionality of each 192-element vector down to 7, corresponding to the 7 physical quantities predicted by the model (e.g., pressure, temperature, Mach number, etc.).

As a result, the output of the projection step has a shape of 120 × 7. This tensor is then post-processed and reshaped into the final output format of 7 × 15 × 8, which represents 7 predicted quantities distributed across 15 radial spanwise positions and 8 streamwise stations. This restructuring step enables the model to produce physically interpretable field predictions aligned with the underlying geometry and mesh topology.

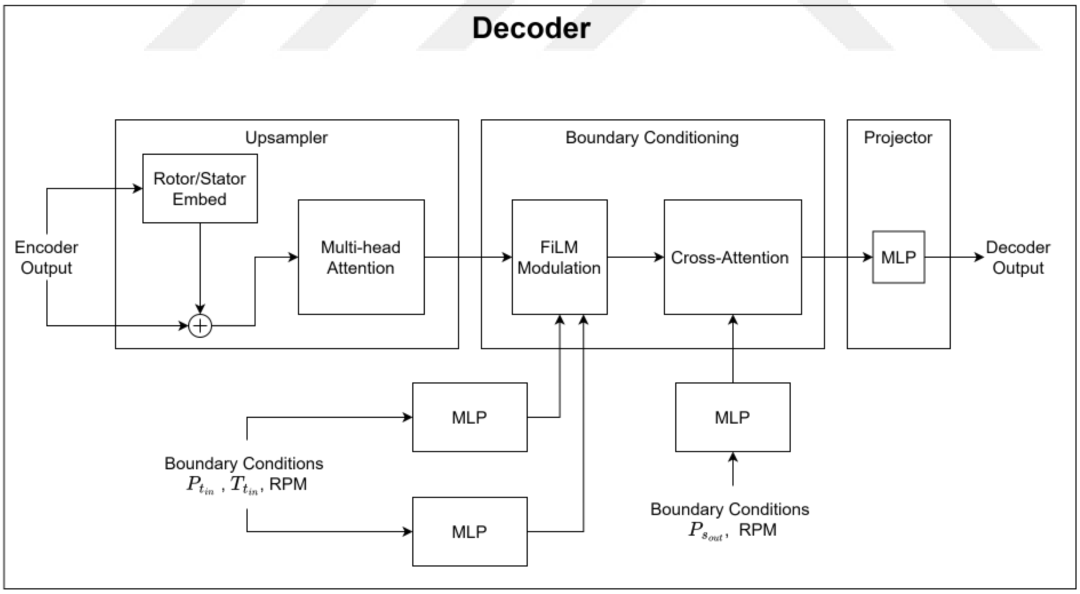


Figure 3.3: Decoder Diagram.

4. TRAINING

Developing a robust and accurate model for turbomachinery applications requires a well-structured training strategy that effectively captures aerodynamic behavior while maintaining sensitivity to geometric variations. Due to the inherent complexity and high dimensionality of the problem, a multi-stage training approach is adopted. The primary objective is to leverage both self-supervised and supervised learning techniques to ensure that the model accurately predicts aerodynamic parameters even when faced with subtle design changes.

To achieve this, the training process is divided into two main phases: an initial contrastive learning phase to pretrain the encoder and a subsequent fine-tuning phase to optimize the decoder. This structured approach allows the model to first learn geometric representations independently, before mapping them to physical outputs.

4.1 Contrastive Learning

4.1.1 NT-Xent loss

The Normalized Temperature-Scaled Cross-Entropy Loss (NT-Xent Loss) is a fundamental component of the contrastive learning framework, particularly introduced in the SimCLR method [18]. This loss function is designed to learn representations by promoting similarity between positive pairs and dissimilarity between negative pairs. In the context of turbomachinery modeling, it plays a crucial role in distinguishing between subtle geometric variations, as small differences in blade geometry can significantly impact aerodynamic performance.

NT-Xent Loss operates by taking paired embeddings and calculating their similarity based on cosine distance. The primary objective is to bring closer the embeddings of the same instance under different augmentations while pushing apart the embeddings of different instances. This approach is particularly relevant for the proposed model, as geometrically similar designs are expected to yield similar latent representations, while geometrically distinct designs should be well-separated in the embedding space.

In this framework, NT-Xent Loss is specifically applied to the encoder output, where the latent vectors produced by the encoder serve as the representations for contrastive learning. These latent vectors encapsulate the geometric features extracted from the input data, making them highly suitable for capturing nuanced differences between similar designs.

The mathematical formulation of NT-Xent Loss is as follows:

$$\mathscr{L}_{i,j} = -\log \frac{\exp(\mathrm{sim}(\mathbf{z}_i, \mathbf{z}_j)/\tau)}{\sum_{k=1}^{2N} 1_{[k \neq i]} \exp(\mathrm{sim}(\mathbf{z}_i, \mathbf{z}_k)/\tau)} \tag{4.1}$$

Here, z_i and z_j represent the latent vectors obtained from the encoder, corresponding to a positive pair. The similarity between these embeddings is computed using cosine similarity:

$$\mathrm{sim}(\mathbf{z}_i, \mathbf{z}_j) = \frac{\mathbf{z}_i \cdot \mathbf{z}_j}{\|\mathbf{z}_i\| \|\mathbf{z}_j\|} \tag{4.2}$$

The impact of the NT-Xent loss is clearly illustrated in Figure 4.1 and 4.2, providing an important insight into the effectiveness of contrastive learning. When the output of the model's encoder layer is evaluated under a flowpath modification scenario—details of which are presented in the *Results* chapter—a substantial difference is observed between models trained with and without contrastive pretraining.

In the absence of contrastive learning, where the model is trained solely using input–output mappings, the encoder outputs for two geometries that differ only in the flowpath region appear almost identical. This indicates that the model fails to distinguish subtle geometric variations at the feature representation level, which consequently leads to nearly identical final predictions from the decoder. However, as shown in the *Results* chapter, the ground-truth CFD results for these two designs differ meaningfully.

On the other hand, the model pretrained with the NT-Xent loss and subsequently fine-tuned on the decoder demonstrates a clear separation in the encoder's feature space for the two geometries. This confirms the utility of contrastive learning in enhancing

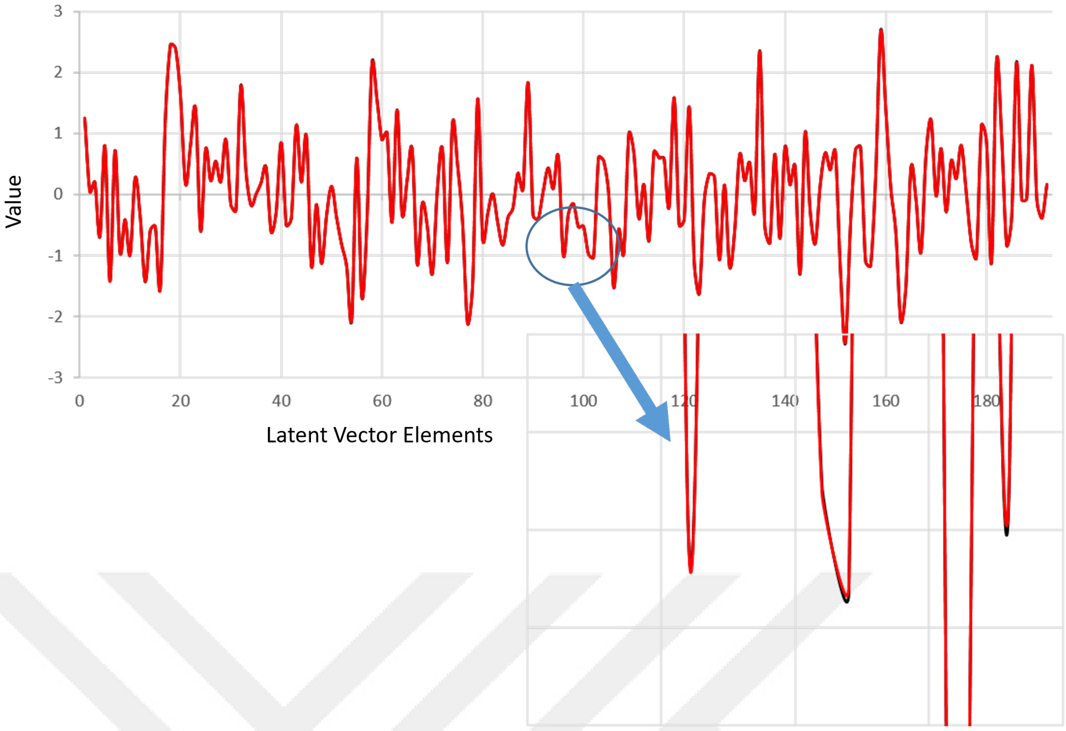


Figure 4.1: Encoder output comparision for two designs(red and black) without contrastive learning.

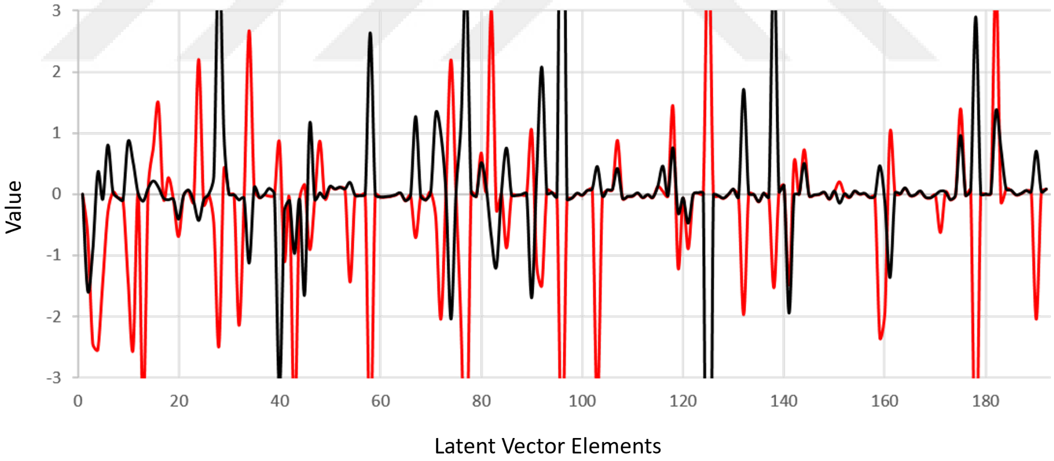


Figure 4.2: Encoder output comparision for two designs(red and black) with contrastive learning.

the model's sensitivity to fine-grained geometric differences and improving overall prediction fidelity. A more detailed assessment is provided in the *Results* chapter.

4.2 Physics-Informed Finetuning

Initially, the model's encoder was pretrained using the NT-Xent loss function within a contrastive learning framework. Following this stage, the encoder parameters were frozen, and the decoder module was fine-tuned using ground-truth CFD data.

During the fine-tuning phase, a percentage-based loss function was employed to minimize the discrepancy between the model predictions and the CFD reference results. In addition to this primary objective, supplementary loss terms were incorporated to further guide the learning process and enhance the physical consistency of the output. Additional loss functions were defined based on physical principles to improve the model's predictive consistency. Specifically, a continuity-based loss was introduced to ensure mass conservation across flow stations. Moreover, isentropic loss terms were constructed using absolute total pressure and total temperature values to enforce thermodynamic consistency.

For the continuity-based loss, the mass flow rate at each station was computed and penalized based on its absolute deviation from the inlet station's mass flow rate. This constraint acts as a physical regularization mechanism to promote global mass conservation throughout the domain.

$$\dot{m} = C_d \frac{A P_t}{\sqrt{T_t}} \sqrt{\frac{\gamma}{R}} M \left(1 + \frac{\gamma - 1}{2} M^2\right)^{-\frac{\gamma+1}{2(\gamma-1)}} \tag{4.3}$$

$$\mathscr{L}_{\text{cont}} = \frac{1}{N} \sum_{i=1}^{N} \left| \frac{\dot{m}_i - \dot{m}_{\text{inlet}}}{\dot{m}_{\text{inlet}}} \right| \tag{4.4}$$

Area is computed using the geometrical input and all the other variables involved in the mass flow calculation (P_t, T_t, M, C_d , $\gamma(T_s)$) were trained to receive feedback through this loss term. As a result, these quantities were adjusted to collectively produce more physically consistent predictions under the constraint of continuity.

To enforce thermodynamic consistency, additional isentropic loss terms were defined using only the model-predicted static pressure, static temperature, and Mach number. The total pressure and total temperature were recalculated from these predicted

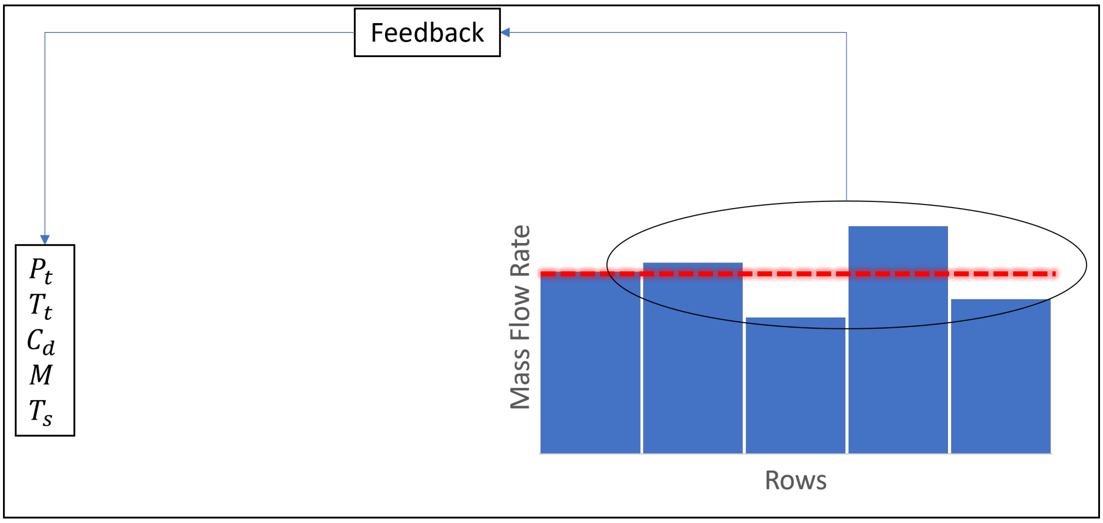


Figure 4.3: Feedback representation for the continuity loss.

quantities using standard isentropic relations. The resulting calculated values were then compared with the model's direct predictions for total quantities, and their relative errors were penalized.

$$P_{t,\text{calc}} = P_{s,\text{pred}} \cdot \left(1 + \frac{\gamma - 1}{2} M_{\text{pred}}^2\right)^{\frac{\gamma}{\gamma-1}} \tag{4.5}$$

$$T_{t,\text{calc}} = T_{s,\text{pred}} \cdot \left(1 + \frac{\gamma - 1}{2} M_{\text{pred}}^2\right) \tag{4.6}$$

$$\mathcal{L}_{P_t} = \frac{1}{N} \sum_{i=1}^{N} \left| \frac{P_{t,\text{calc},i} - P_{t,\text{pred},i}}{P_{t,\text{calc},i}} \right| \tag{4.7}$$

$$\mathcal{L}_{T_t} = \frac{1}{N} \sum_{i=1}^{N} \left| \frac{T_{t,\text{calc},i} - T_{t,\text{pred},i}}{T_{t,\text{calc},i}} \right| \tag{4.8}$$

$$\mathcal{L}_{\text{isent}} = \mathcal{L}_{P_t} + \mathcal{L}_{T_t} \tag{4.9}$$

These two loss functions play a critical role in making the model physics-informed. By embedding domain-specific physical constraints into the training process, the model is encouraged to learn mappings from inputs to outputs that not only minimize numerical error but also follow physically meaningful trends. As a result, and as will be demonstrated in the results section, the model achieves improved generalization

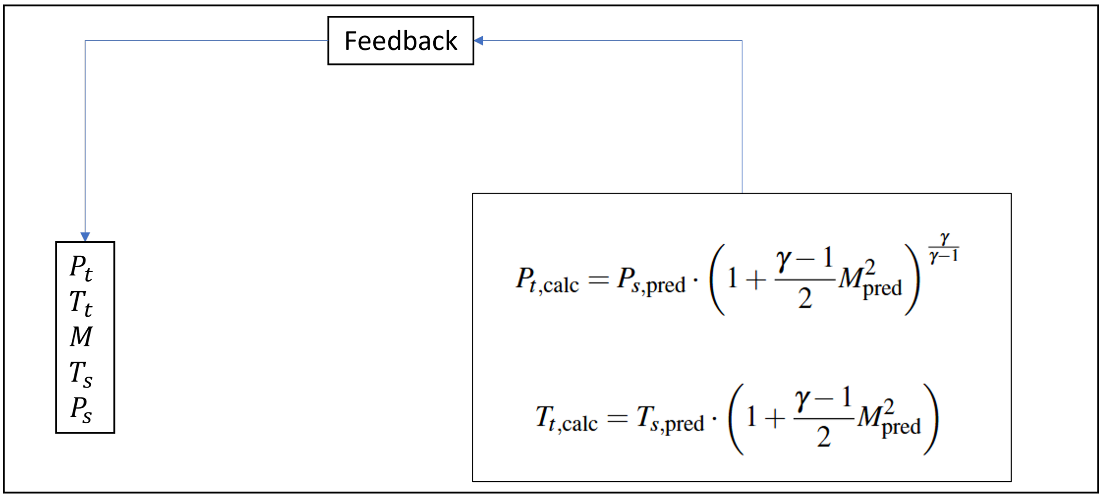


Figure 4.4: Feedback representation for the isentropic loss.

performance, particularly when exposed to unseen data or design variations outside the training distribution.

In the absence of such physical constraints, the model tends to overfit the training data, learning superficial patterns that may not generalize well. For instance, in the presence of geometric variations, the model may fail to produce consistent trends, misrepresenting whether a given parameter should increase or decrease in response to a design change. This undermines the model's interpretability and usefulness in real-world engineering applications.

From this perspective, incorporating additional physics-based equations into the training process can be viewed as a form of data augmentation. These constraints effectively enrich the information content of the dataset, guiding the model to prioritize solutions that are not only numerically accurate but also physically plausible—especially in data-scarce regimes where supervised signals alone may be insufficient.

During training, the model was optimized using a combination of a percentage-based loss function and two physics-based loss terms. These complementary components ensure that the model not only minimizes numerical discrepancies but also maintains physical consistency. The training process was terminated once the total loss fell below a predefined threshold corresponding to a 1% error margin.

$$\mathscr{L}_{\text{percent}} = \frac{1}{N}\sum_{i=1}^{N}\left|\frac{\hat{y}_i - y_i}{y_i}\right| \tag{4.10}$$

$$\mathscr{L}_{\text{total}} = \mathscr{L}_{\text{percent}} + \mathscr{L}_{\text{cont}} + \mathscr{L}_{\text{isent}} \tag{4.11}$$

5. RESULTS

In this chapter, the performance of the model will be comprehensively evaluated from multiple perspectives. First, the accuracy of the model will be analyzed in terms of both directly predicted parameters (pressure, Mach Numbers etc.) and those derived through computations (De Haller Number, Pr). Second, the consistency and reliability of the model's outputs will be assessed through visualizations on performance maps, providing insights into its behavior across various operating regimes. Finally, the model's response to design modifications will be examined, with a particular focus on its sensitivity to geometric variations and its ability to reflect physically meaningful differences across distinct geometries. These evaluations have been conducted for the set of geometries specified in the table below.

Table 5.1: Designs for the model testing.

Name	Baseline	Modification
DESIGN V1	-	-
DESIGN V2	*DESIGN V1*	R1/S1 outlet metal angle
DESIGN V3	-	-
DESIGN V4	*DESIGN V3*	R1 LE hub radius

Design V1 and *Design V3* serve as the baseline configurations, while the remaining designs are derived from these reference cases through specific geometric modifications. The transition from *Design V1* to *Design V2* involves a redistribution of the aerodynamic loading across the first and second stages, resulting in observable changes in the stage-wise flow behavior.

Between *Design V3* and *Design V4*, a minor geometric adjustment has been made specifically in the hub region of the flow path, introducing a localized change without significantly altering the overall blade design.

5.1 Model Capability on Radial Profiles

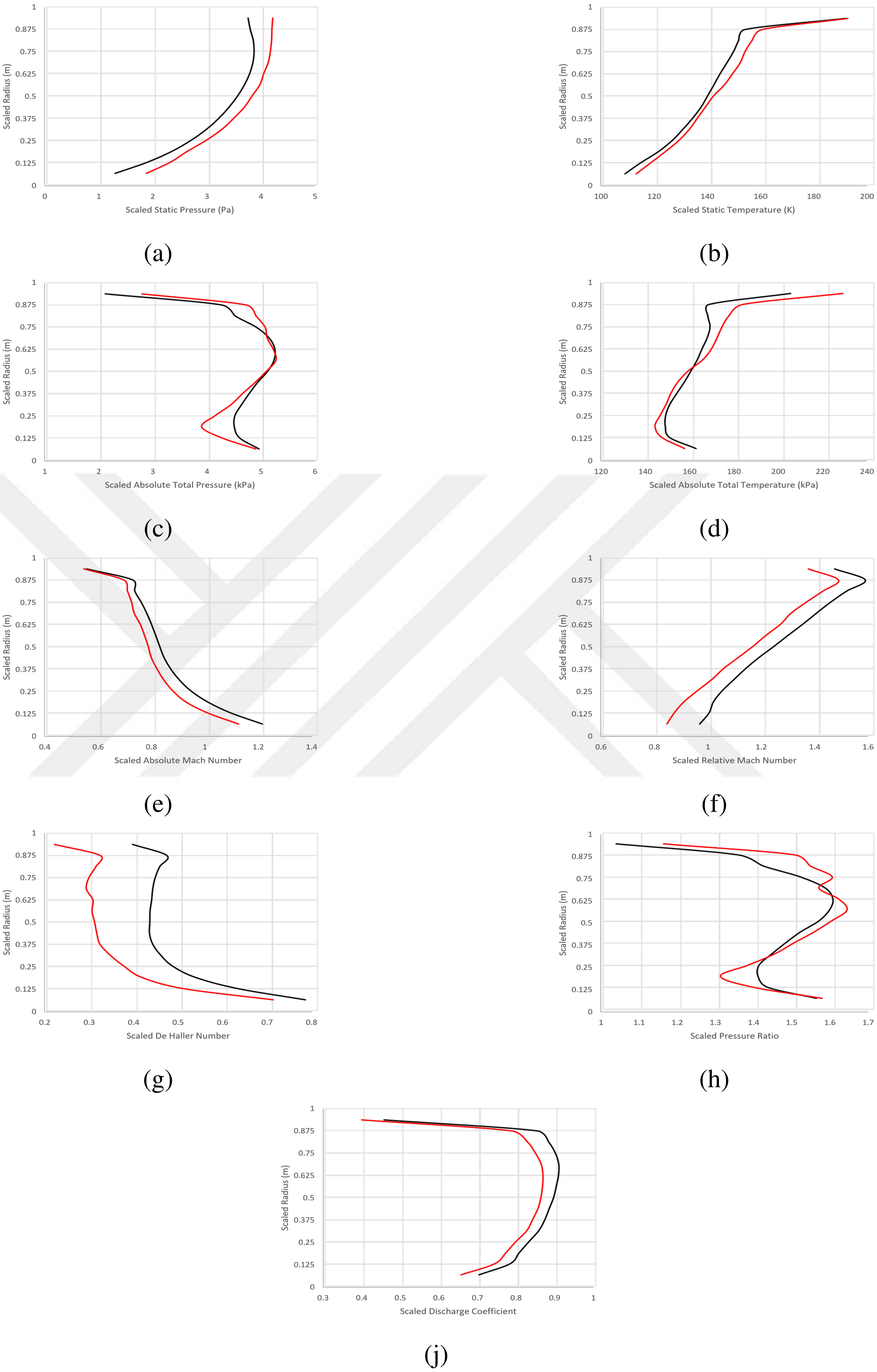
In this section, the model's predictive performance will be examined under high, part, and low speed operating conditions, with a focus on both spanwise and streamwise directions using the *DESIGN V3* geometry. This analysis aims to evaluate how well the model captures the underlying flow physics across different operational regimes, including changes in aerodynamic behavior along the blade span and through the axial flow path.

5.1.1 Design speeds

The "High Speeds" condition corresponds to 100% of the fan's design rotational speed, which not only represents the nominal operating point of the machine but also accounts for a significant portion of the training dataset. This provides a valuable benchmark for evaluating the model's performance under well-represented and physically meaningful conditions.

Figure 5.1 represents radial comparisions for the design matching point. From a spanwise perspective, the model demonstrates a strong ability to capture the general aerodynamic behavior of the turbomachinery, particularly in the mid-span regions. These regions are typically characterized by relatively uniform flow patterns, making them essential for assessing the core prediction capabilities of the model. Additionally, the model shows proficiency in handling the intricate flow dynamics near the endwalls, where secondary flows, boundary layer effects, and tip leakage phenomena dominate. Successfully predicting these complex features is key to determining whether the model can extend its applicability beyond the simpler mid-span flow and account for spanwise variations.

As shown in Figure 5.2 when examined in the streamwise direction (based on radial averaging) the model exhibits a high level of consistency with the CFD solutions. This alignment indicates that the model is not only capable of capturing local flow patterns but also successfully reproduces the overall streamwise evolution of the flow field.



(a)

(b)

(c)

(d)

(e)

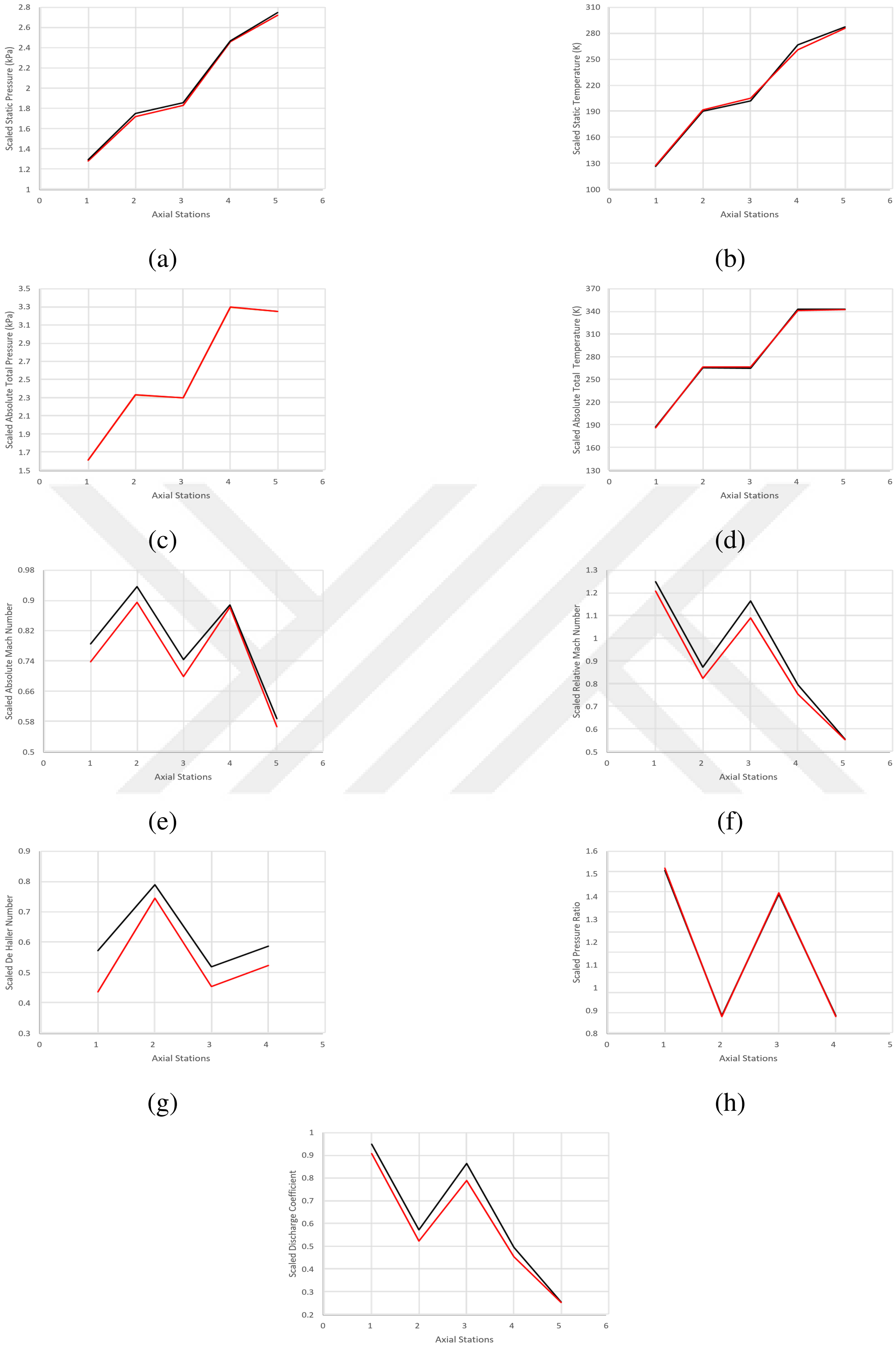
(f)

(g)

(h)

(j)

Figure 5.1: Spanwise distribution of the static pressure (a) and temperature (b), absolute total pressure (c) and temperature (d), absolute (e) and relative (f) Mach numbers, De Haller Number (g) and pressure ratio (h), discharge coefficient (j) for predicted (red) and CFD results (black).



(a) (b)

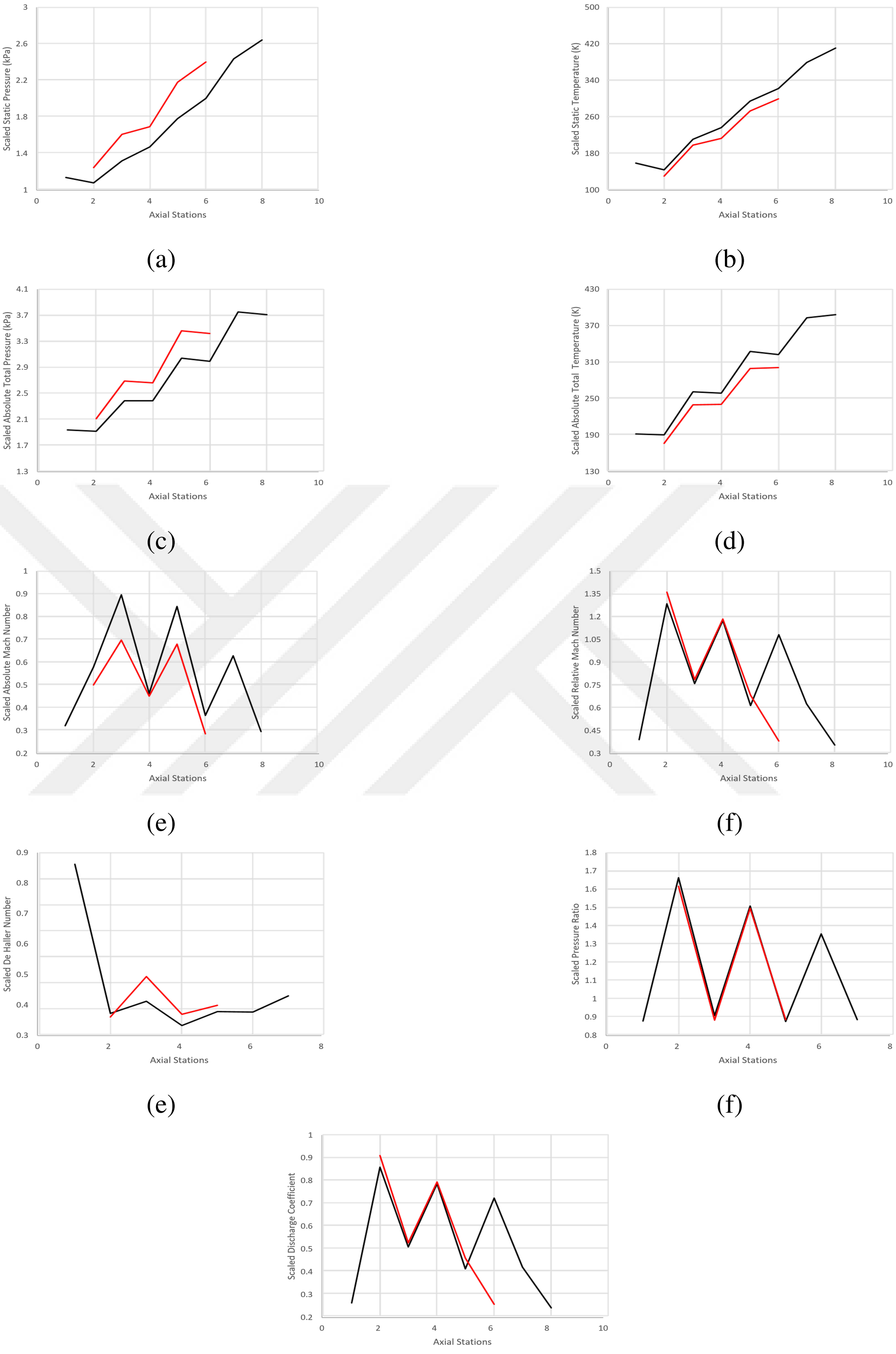
(c) (d)

(e) (f)

(g) (h)

(j)

Figure 5.2: Streamwise distribution of the static pressure (a) and temperature (b), absolute total pressure (c) and temperature (d), absolute (e) and relative (f) Mach numbers, De Haller Number (g) and pressure ratio (h), discharge coefficient (j) for predicted (red) and CFD results (black).



(a) (b)

(c) (d)

(e) (f)

(e) (f)

(f)

Figure 5.3: Streamwise distribution of the static pressure (a) and temperature (b), absolute total pressure (c) and temperature (d), absolute (e) and relative (f) Mach numbers, De Haller Number (g) and pressure ratio (h), discharge coefficient (j) for two-stage (red) and three-stage (black) configurations.

Such agreement is crucial for ensuring that the predicted quantities maintain physical coherence along the primary flow direction.

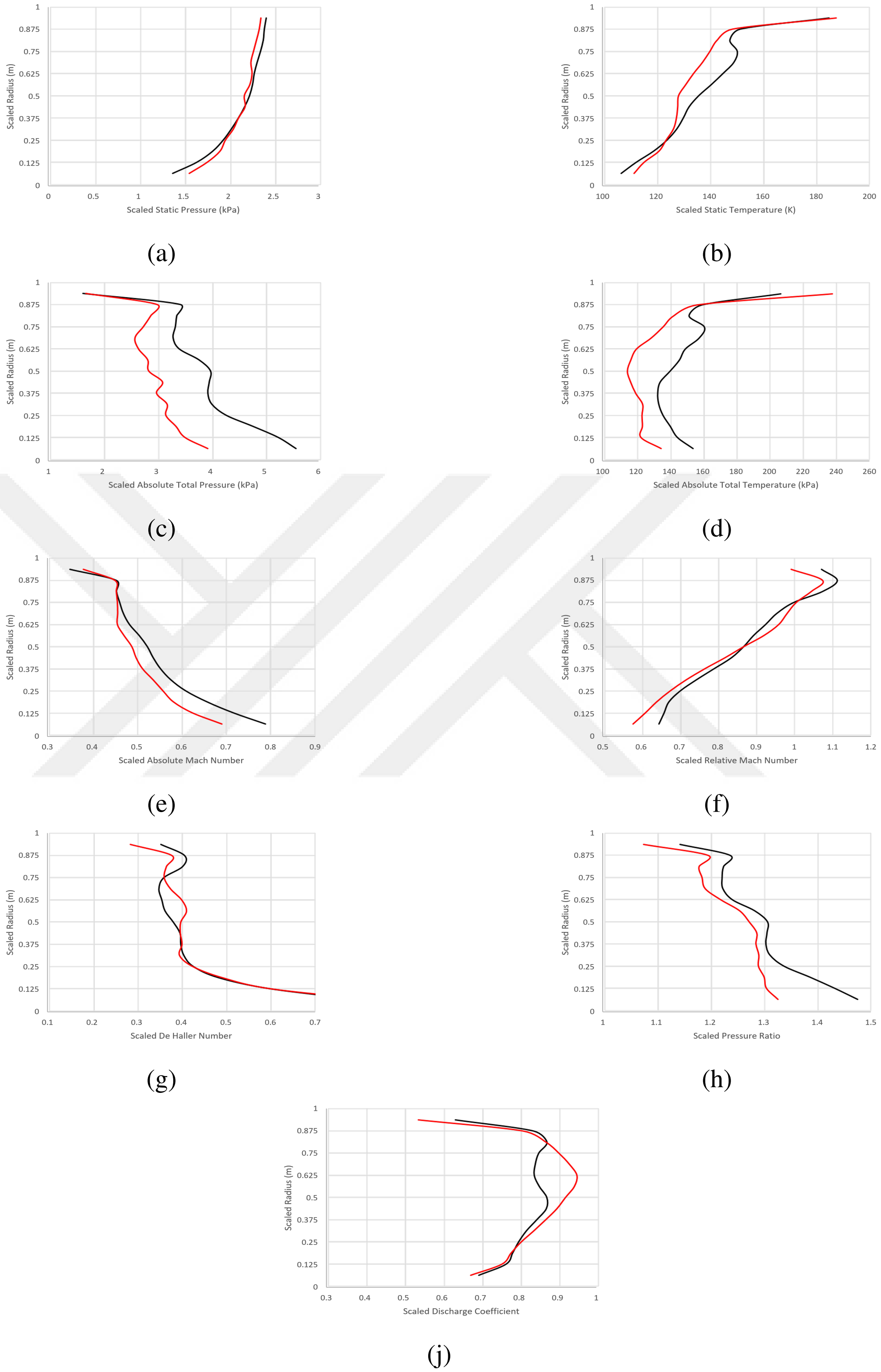
Although not all geometries in the dataset correspond to a uniform two-stage configuration without inlet guide vanes IGV or OGV, the model consistently demonstrates physically meaningful behavior across both types of configurations. In particular, it accurately captures the expected differences between rotor and stator components—such as an increase in absolute total pressure across the rotors and a corresponding pressure drop across the stators, given by the Figure 5.3. This consistency across heterogeneous geometries highlights the model's ability to generalize its understanding of fundamental turbomachinery physics, regardless of architectural variations.

The primary contributor to this robust behavior is the rotor–stator embedding layer within the decoder module, which enables the model to differentiate between the distinct physical roles of rotating and stationary components. By explicitly encoding rotor–stator identities, the model is guided to reflect the appropriate flow transformations occurring within each stage of the machine.

5.1.2 Off-Design speeds

Despite having relatively limited representation in the dataset, the model is able to produce consistent and physically plausible predictions at off-design speeds such as 70% and 40% of the nominal operating RPM, albeit with reduced accuracy. The results for the 70% speed case are presented in Figure 5.4 and 5.5, while the outcomes corresponding to the 40% speed case are shown in Figure 5.6 and 5.7.

Figure 5.8 illustrates the output of the model trained without incorporating physics-based information at 70 RPM. It can be clearly seen that the absence of physical guidance results in a noticeable decline in prediction accuracy.



(a) (b)

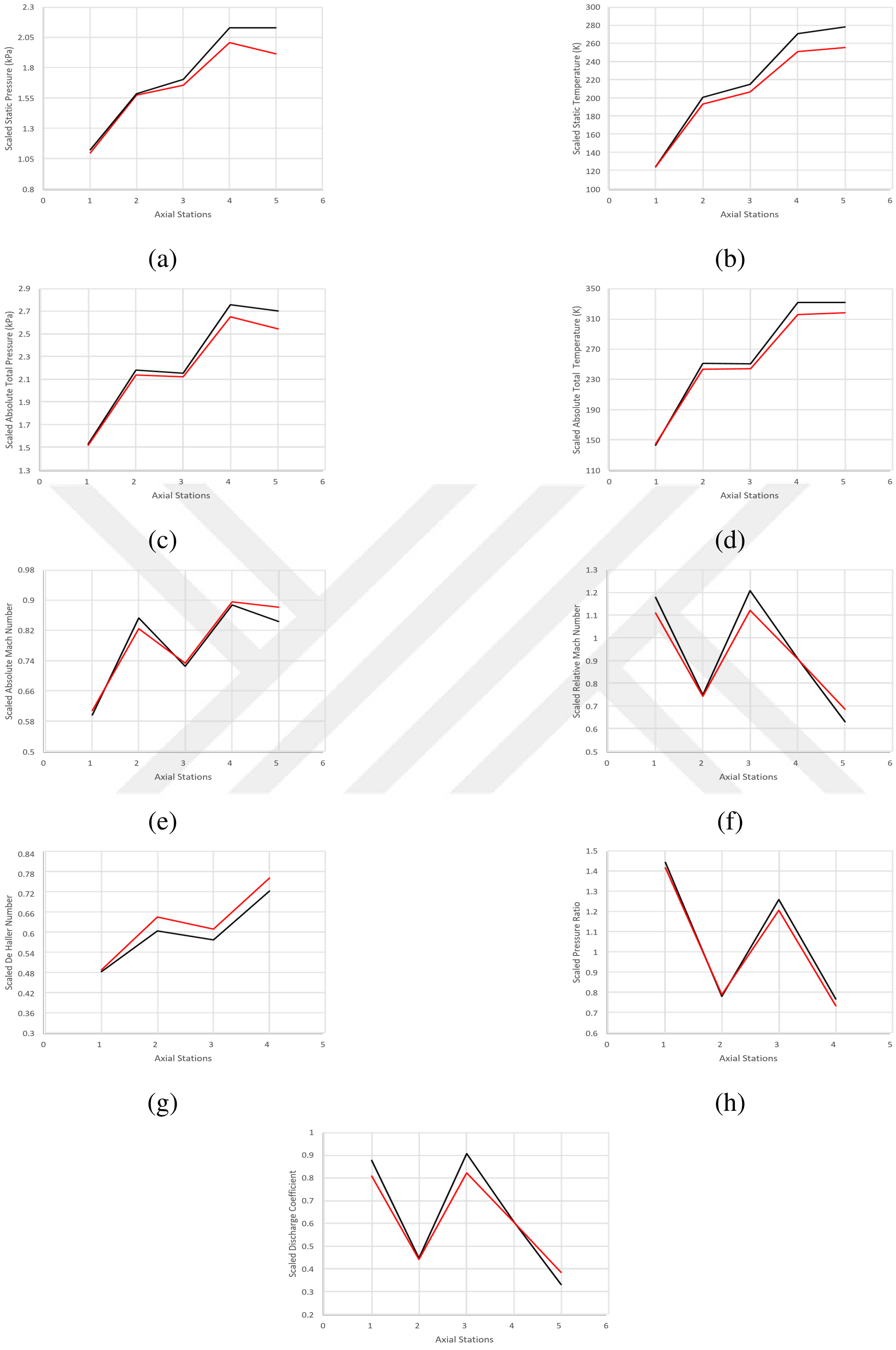
(c) (d)

(e) (f)

(g) (h)

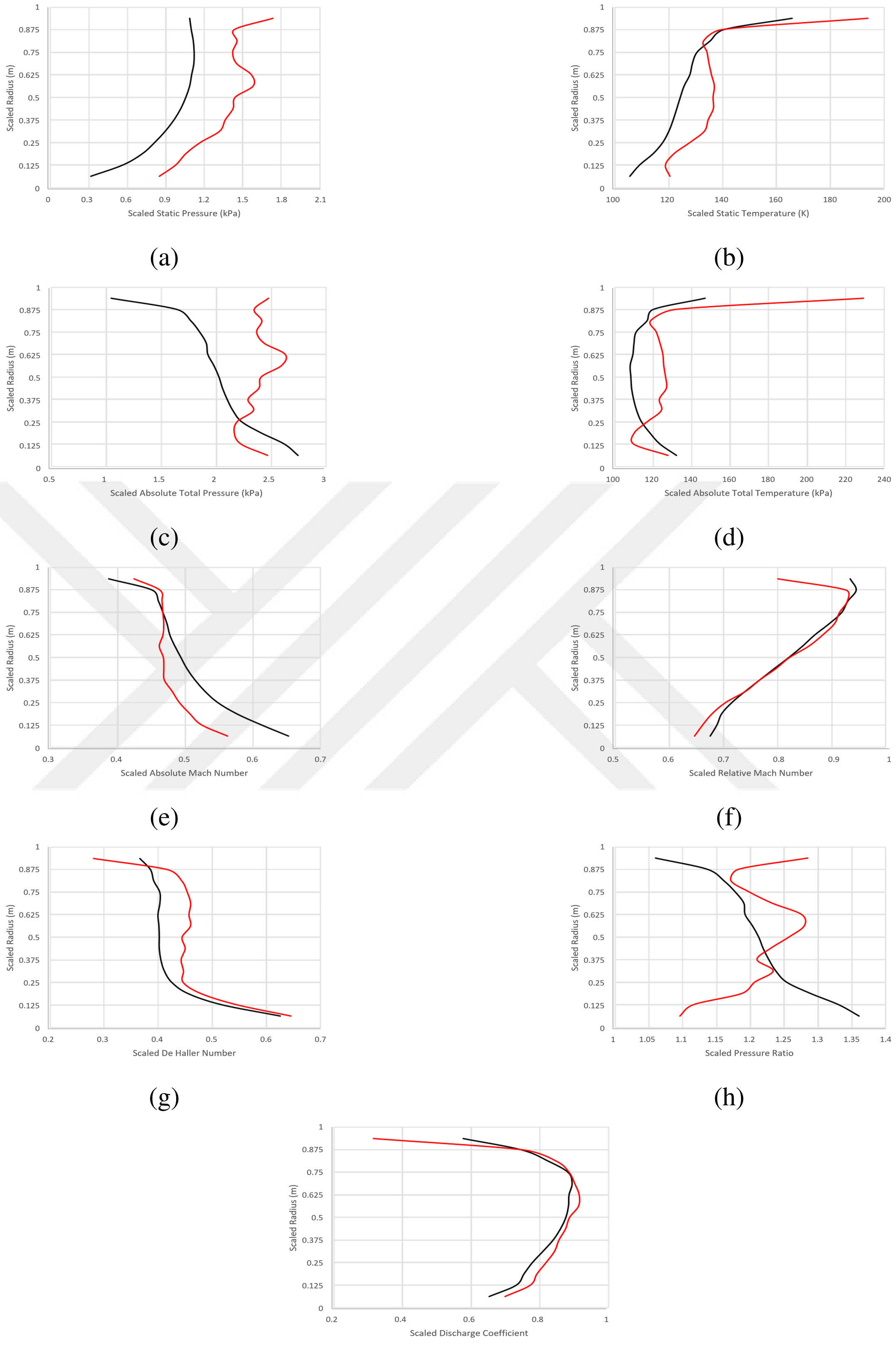
(j)

Figure 5.4: Spanwise distribution of the static pressure (a) and temperature (b), absolute total pressure (c) and temperature (d), absolute (e) and relative (f) Mach numbers, De Haller Number (g) and pressure ratio (h), discharge coefficient (j) for predicted (red) and CFD results (black) at 70% RPM.



(j)

Figure 5.5: Streamwise distribution of the static pressure (a) and temperature (b), absolute total pressure (c) and temperature (d), absolute (e) and relative (f) Mach numbers, De Haller Number (g) and pressure ratio (h), discharge coefficient (j) for predicted (red) and CFD results (black) at 70% RPM.



(a) (b)

(c) (d)

(e) (f)

(g) (h)

(j)

Figure 5.6: Spanwise distribution of the static pressure (a) and temperature (b), absolute total pressure (c) and temperature (d), absolute (e) and relative (f) Mach numbers, De Haller Number (g) and pressure ratio (h), discharge coefficient (j) for predicted (red) and CFD results (black) at 40% RPM.

In contrast, the physics-informed approach successfully provides the model with an additional layer of domain-specific insight, enhancing its ability to capture the underlying flow behavior more accurately and yielding results that are more consistent with expected aerodynamic patterns.

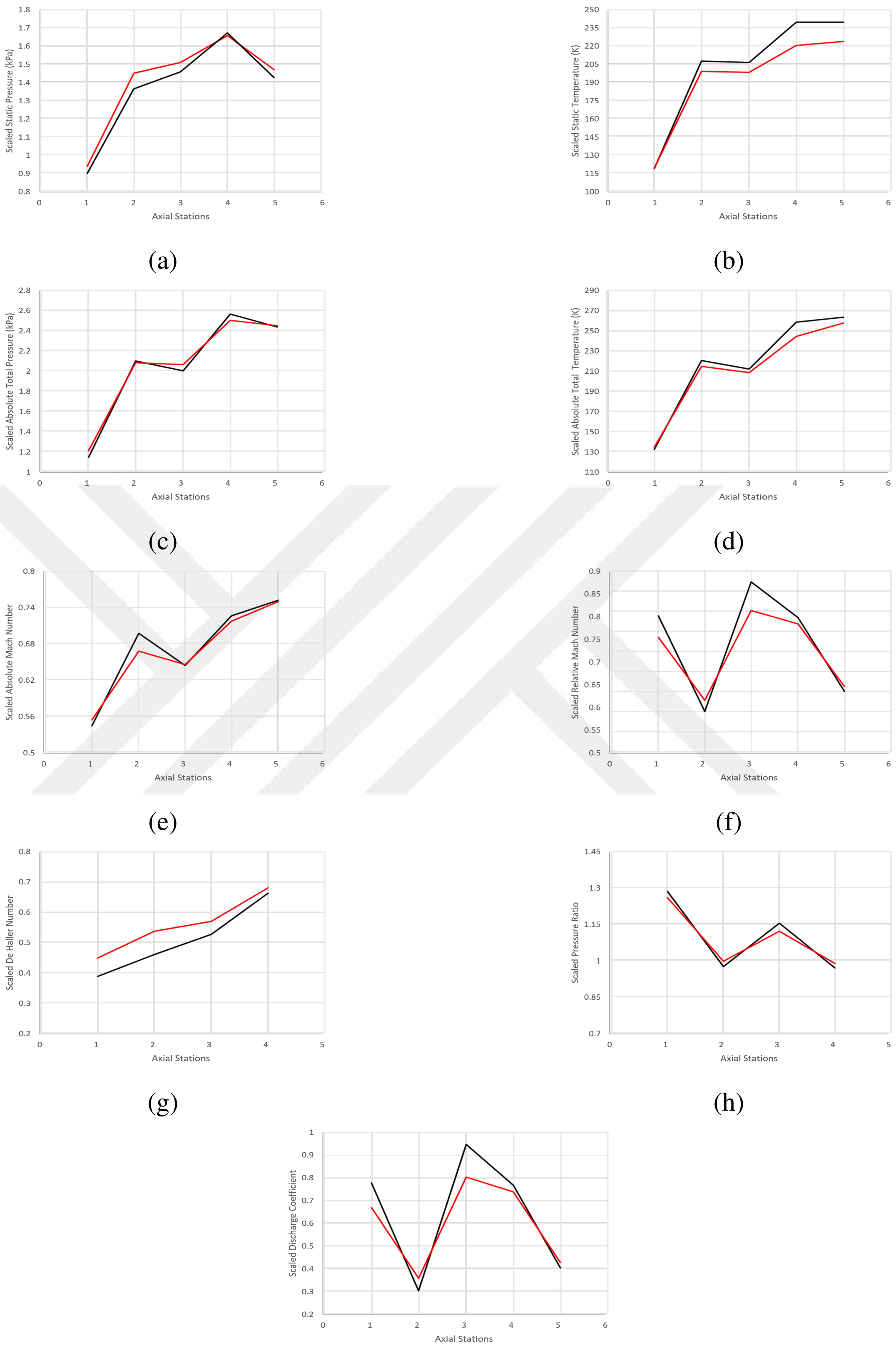
During model training, the continuity equation constraint was imposed based on the inlet station, ensuring consistency across the remaining flow stations. Accordingly, in the performance maps, the mass flow rate was also evaluated with reference to the inlet conditions to maintain coherence between the training setup and the performance evaluation framework.

$$\dot{m} = C_{d_{in}} \frac{A P_{t_{in}}}{\sqrt{T_{t_{in}}}} \sqrt{\frac{\gamma}{R}} M_{in} \left(1 + \frac{\gamma - 1}{2} M_{in}^2\right)^{-\frac{\gamma+1}{2(\gamma-1)}} \tag{5.1}$$

The mass flow rate, given in Equation 5.1, is computed based on five predicted quantities, including the specific heat ratio γ, which is derived from temperature-dependent relations. While the model yields low individual errors for each of these predicted variables, small discrepancies can accumulate and amplify when combined in a nonlinear formulation such as the mass flow rate equation.

The scarcity of training data in stall and choke regions reduces the model's ability to generate accurate predictions under these off-design operating conditions. This limitation is particularly evident in the performance maps, where the predicted flow behavior deviates more noticeably from the reference CFD results near the extremes of the speedline. In contrast, the model maintains higher fidelity in the radial direction, where the operating conditions are more uniformly represented and physically stable.

The FiLM-based boundary conditioning layer, together with the cross-attention mechanism within the model architecture, was specifically designed to ensure that predictions remain physically consistent with varying boundary conditions such as RPM and outlet static pressure. As demonstrated in the performance maps provided in Figure 5.9, these components are effective at capturing the physical influence of the specified inputs.



(a) (b)

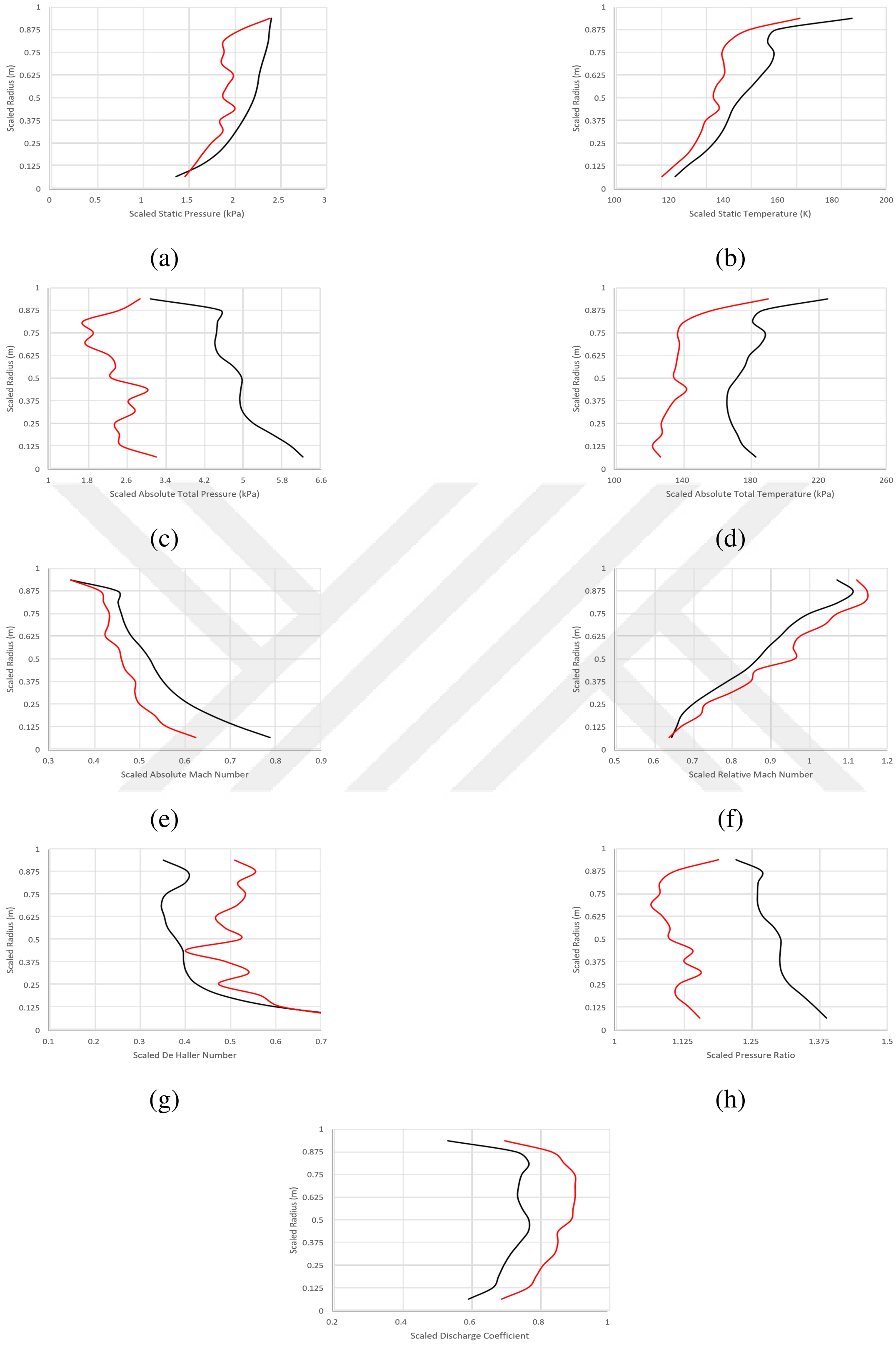
(c) (d)

(e) (f)

(g) (h)

(j)

Figure 5.7: Streamwise distribution of the static pressure (a) and temperature (b), absolute total pressure (c) and temperature (d), absolute (e) and relative (f) Mach numbers, De Haller Number (g) and pressure ratio (h), discharge coefficient (j) for predicted (red) and CFD results (black) at 40% RPM.



(a) (b)

(c) (d)

(e) (f)

(g) (h)

(j)

Figure 5.8: Spanwise distribution of the static pressure (a) and temperature (b), absolute total pressure (c) and temperature (d), absolute (e) and relative (f) Mach numbers, De Haller Number (g) and pressure ratio (h), discharge coefficient (j) for predicted (red) and CFD results (black) at 70% RPM using model that not uses physics-informed loss.

5.2 Model Capability on Speedlines

The model consistently identifies the location of the speedline on the performance map based on the given RPM, accurately reflecting its position in relation to the operating regime. Furthermore, it adjusts the placement of the operating point along the speedline in accordance with variations in back pressure. As the back pressure increases, the model captures the expected physical trends: an increase in pressure ratio accompanied by a reduction in mass flow rate.

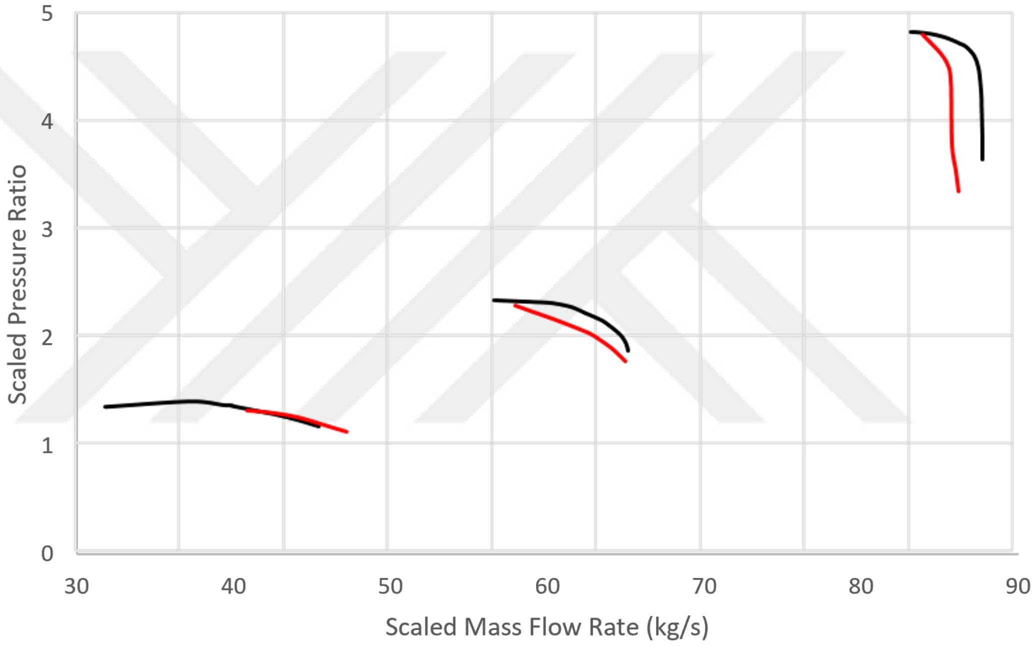


Figure 5.9: *DESIGN V3* performance map for different rpm values.

To further improve the accuracy of the model across the entire speedline, it may be beneficial to enrich the dataset with additional operating points or to incorporate supplementary physical constraints that can act as implicit regularization, especially in data-scarce regions.

5.3 Sensitivity of the Model to Geometric Variations

In this section, the focus shifts from evaluating the numerical accuracy of the model to examining its ability to predict the effects of design modifications. This aspect represents one of the most critical outcomes of the project. Achieving near-zero error when compared to CFD predictions is practically impossible; however, if the model can replicate the trend and relative changes in output parameters in response to design variations in a manner consistent with CFD results, it becomes a valuable tool for design iteration processes.

When comparing two different geometries, if the expected changes in key aerodynamic parameters observed in CFD outputs are also reflected in the model's predictions, the model may serve not only as a computationally efficient surrogate but also as a design-support system that significantly accelerates the overall development cycle.

Using the available geometries, the model's responses have been analyzed against two types of geometric variations: changes in flowpath radius and metal angle

5.3.1 Metal angle variation

In the transition from *Design V1* to *Design V2*, the first stage was off-loaded while the second stage was correspondingly loaded. Under these conditions, a increase in the De Haller distribution across the first stage is expected, along with an decrease across the second stage. As illustrated in Figure 5.10, the model successfully captures these expected changes, demonstrating its ability to reflect stage-specific flow behavior resulting from targeted design modifications.

(a) (b)

Figure 5.10: Model results showing the De Haller number for predicted(red) and CFD results(black) of *Design V1*(solid) and *Design V2*(dashed) on Rotor 1 (a) and Rotor 2 (b), metal angle variation case.

5.3.2 Flowpath variation

Between *Design V3* and *Design V4*, only a minor reduction in radius was introduced in the hub section of the leading edge of the first rotor's flowpath. This variation serves as a critical test case for evaluating the model's discriminative capability, as the geometric difference between the two variants is minimal from a mathematical standpoint.

Although the geometric change is minimal, a decrease in the Mach number distribution and a corresponding increase in the static pressure distribution are expected immediately downstream of the first rotor. As shown in Figure 5.11, the model successfully captures this behavior, clearly demonstrating its ability to distinguish even small geometric variations and translate them into physically consistent predictions.

(a) (b)

Figure 5.11: Model results showing the absolute Mach number (a) and static pressure (b) for predicted(red) and CFD results(black) of *Design V3*(solid) and *Design V4*(dashed).

In the section *Model Capability on Speedlines*, it was previously noted that the model exhibits higher error rates on performance maps compared to radial profiles, primarily due to the amplification of small prediction errors through derived quantities. However, the model is capable of capturing not only radial profile behavior but also the overall trends along the speedlines with reasonable physical consistency.

As illustrated in Figure 5.12, *Design V4* demonstrates the ability to reach more extreme mass flow values at both stall and choke limits, whereas *Design V3* exhibits a higher pressure ratio near the stall region. Despite the reduced accuracy typically observed in stall and choke predictions—owing to limited data in these regimes—the model still manages to reflect the correct physical behavior associated with these subtle geometric differences. This indicates that even in the presence of minimal design changes, the model preserves its ability to track the corresponding trends on the performance map.

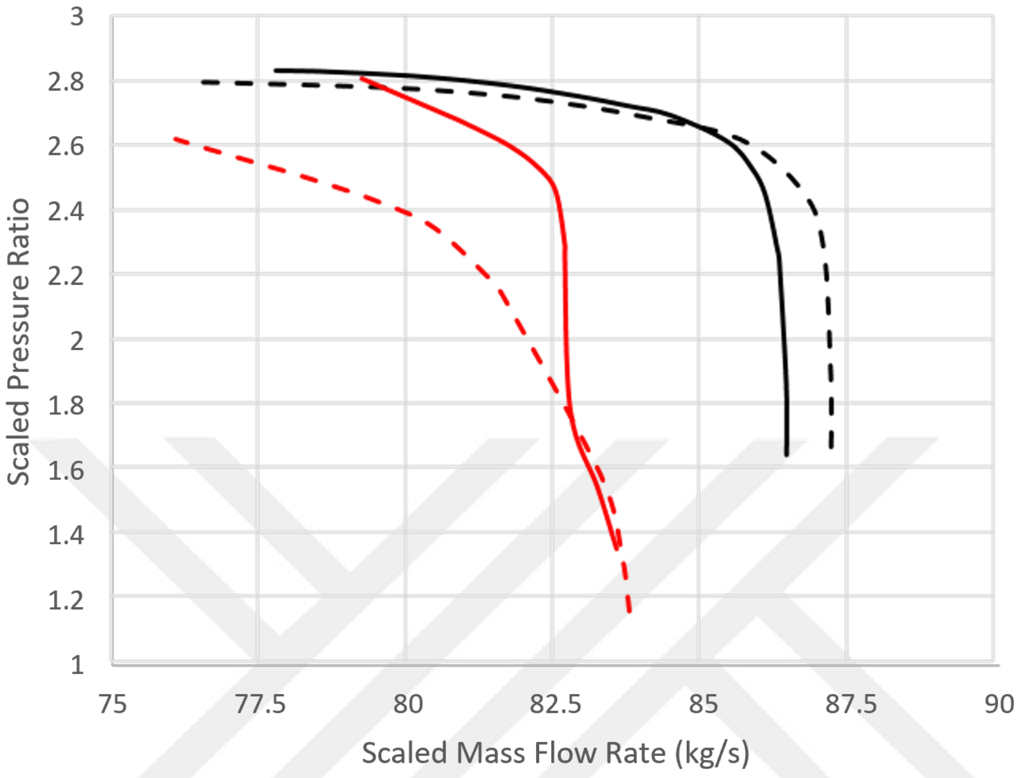


Figure 5.12: Model results showing the performance map for predicted(red) and CFD results(black) of *Design V3*(solid) and *Design V4*(dashed).

6. CONCLUSIONS

This thesis presented the development of a physics-informed deep learning framework for predicting aerodynamic quantities in multi-stage turbomachinery, with a particular emphasis on the sensitivity to geometric variations and operational conditions. By leveraging the Vision Transformer (ViT) architecture as the backbone of the encoder, the model successfully captured complex spatial dependencies and geometric structures from blade profiles. The proposed architecture was further enhanced with contrastive pretraining using the NT-Xent loss and a decoder that incorporated boundary condition conditioning via FiLM and cross-attention mechanisms.

The training procedure followed a two-stage strategy: first, the encoder was pretrained to distinguish subtle geometric variations in a self-supervised setting, and then the decoder was fine-tuned using CFD ground truth data. The use of physics-based loss functions—specifically, the continuity-based and isentropic loss terms—ensured that the predicted flow fields adhered to fundamental physical laws, promoting generalization and physical consistency even in underrepresented regions such as stall and choke conditions.

The model demonstrated robust predictive capabilities across various evaluation scenarios. On radial and streamwise profiles, the model reproduced complex aerodynamic trends with high fidelity, capturing both global flow evolution and local phenomena such as secondary flows near endwalls. On speedlines, the model consistently positioned the operating points according to back pressure and RPM inputs, revealing a solid grasp of performance map dynamics despite increased difficulty near flow regime boundaries. Most notably, the model was able to replicate the qualitative and trend-level changes in aerodynamic parameters resulting from minor design modifications, including changes in metal angle and hub radius. These results suggest that the model can be used not only for prediction but also as a surrogate model in design iterations, effectively accelerating the turbomachinery design process.

In conclusion, the integration of self-supervised contrastive learning and physics-based regularization into a transformer-based architecture has proven to be a powerful approach for generalizable, physically consistent aerodynamic modeling. While the model does not fully replace CFD in terms of fine-scale accuracy, it offers a compelling trade-off between computational efficiency and predictive performance. Future work may include expanding the dataset with additional design variations, refining the loss functions with domain-specific formulations, and extending the model to capture transient and 3D effects in more complex flow environments.

While the current model demonstrates strong generalization capabilities and physical consistency, several avenues remain for future enhancement. One immediate direction is to increase the diversity and quantity of training data, particularly by incorporating more complete speedlines and off-design operating conditions. A richer dataset would allow the model to better capture extreme aerodynamic regimes and reduce uncertainty in underrepresented regions, such as stall and choke.

Moreover, to further improve the fidelity of speedline predictions—especially in capturing mass flow and pressure ratio variations with higher accuracy—future studies could explore alternative network architectures and more sophisticated physics-informed loss functions. Such advancements could push the model beyond trend-level agreement toward more precise, CFD-comparable performance across the full design envelope.

REFERENCES

[1] **Dosovitskiy, A.**, **Beyer, L.**, **Kolesnikov, A.**, **Weissenborn, D.**, **Zhai, X.**, **Unterthiner, T.**, **Dehghani, M.**, **Minderer, M.**, **Heigold, G.**, **Gelly, S.**, **Uszkoreit, J. and Houlsby, N.** (2020). An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale, *CoRR*, *abs/2010.11929*, `https://arxiv.org/abs/2010.11929`, `2010.11929`.

[2] (2019). Surrogate Estimations of Complete Flow Fields of Fan Stage Designs via Deep Neural Networks, `https://doi.org/10.1115/GT2019-91258`, `https://asmedigitalcollection.asme.org/GT/proceedings-pdf/GT2019/58585/V02DT46A013/6441020/v02dt46a013-gt2019-91258.pdf`.

[3] **Feng, Y.**, **Song, X.**, **Yuan, W. and Lu, H.** (2023). Physics-informed deep learning cascade loss model, *Aerospace Science and Technology*, *134*, 108165, `https://www.sciencedirect.com/science/article/pii/S1270963823000627`.

[4] **Hipple, S.M.**, **Bonilla-Alvarado, H.**, **Pezzini, P.**, **Shadle, L. and Bryden, K.M.** (2020). Using Machine Learning Tools to Predict Compressor Stall, *Journal of Energy Resources Technology*, *142*(7), 070915, `https://doi.org/10.1115/1.4046458`, `https://asmedigitalcollection.asme.org/energyresources/article-pdf/142/7/070915/6539727/jert\_142\_7\_070915.pdf`.

[5] **Li, J.**, **Deng, Y.**, **Zhang, X.**, **Wang, W.**, **Fan, B. and Peng, F.** (2024). Instability prediction under distorted inflow based on deep learning neural networks in an axial flow compressor, *Physics of Fluids*, *36*(12), 127163, `https://doi.org/10.1063/5.0241142`, `https://pubs.aip.org/aip/pof/article-pdf/doi/10.1063/5.0241142/20308616/127163\_1\_5.0241142.pdf`.

[6] **Liu, T.**, **Gao, L. and Li, R.** (2024). Experimental data-driven flow field prediction for compressor cascade based on Deep Learning and 1 regularization, *Journal of Thermal Science*, *33*(5), 1867–1882.

[7] **Bruni, G.**, **Maleki, S. and Krishnababu, S.K.** (2025). Deep learning modeling of manufacturing and build variations on multistage axial compressors aerodynamics, *Data-Centric Engineering*, *6*, `http://dx.doi.org/10.1017/dce.2025.2`.

[8] **Aulich, M.**, **Goinis, G. and Voß, C.** (2024). Data-Driven AI Model for Turbomachinery Compressor Aerodynamics Enabling Rapid Approximation of 3D Flow Solutions, *Aerospace*, *11*(9), `https://www.mdpi.com/2226-4310/11/9/723`.

[9] **Ji, Z. and Du, G.** (2024). A tensor basis neural network-based turbulence model for transonic axial compressor flows, *Aerospace Science and Technology*, *149*, 109155, `https://www.sciencedirect.com/science/article/pii/S1270963824002888`.

[10] **Chen, L.W. and Thuerey, N.** (2023). Towards high-accuracy deep learning inference of compressible flows over aerofoils, *Computers Fluids*, *250*, 105707, `https://www.sciencedirect.com/science/article/pii/S0045793022003000`.

[11] **Xiang, H.**, **Ma, Y.**, **Dai, Z.**, **Wang, C. and Zhang, B.** (2024). *AeroDiT: Diffusion Transformers for Reynolds-Averaged Navier-Stokes Simulations of Airfoil Flows*, `https://arxiv.org/abs/2412.17394`, `2412.17394`.

[12] **Peng, W.**, **Yuan, Z. and Wang, J.** (2022). Attention-enhanced neural network models for turbulence simulation, *Physics of Fluids*, *34*(2), 025111, `https://doi.org/10.1063/5.0079302`, `https://pubs.aip.org/aip/pof/article-pdf/doi/10.1063/5.0079302/16643352/025111\_1\_online.pdf`.

[13] **Kang, H.**, **Kim, Y.**, **Le, T.T.H.**, **Choi, C.**, **Hong, Y.**, **Hong, S.**, **Chin, S.W. and Kim, H.** (2023). A new fluid flow approximation method using a vision transformer and a U-shaped convolutional neural network, *AIP Advances*, *13*(2), 025233, `https://doi.org/10.1063/5.0138515`, `https://pubs.aip.org/aip/adv/article-pdf/doi/10.1063/5.0138515/16759425/025233\_1\_online.pdf`.

[14] **Xiao, L.**, **Zhang, M. and Chang, X.** (2024). Learning Airfoil Flow Field Representation via Geometric Attention Neural Field, *Applied Sciences*, *14*(22), `https://www.mdpi.com/2076-3417/14/22/10685`.

[15] **Deng, Z.**, **Wang, J.**, **Liu, H.**, **Xie, H.**, **Li, B.**, **Zhang, M.**, **Jia, T.**, **Zhang, Y.**, **Wang, Z. and Dong, B.** (2023). Prediction of transonic flow over supercritical airfoils using geometric-encoding and deep-learning strategies, *Physics of Fluids*, *35*(7), `http://dx.doi.org/10.1063/5.0155383`.

[16] **Hütten, N.**, **Meyes, R. and Meisen, T.** (2022). Vision Transformer in Industrial Visual Inspection, *Applied Sciences*, *12*(23), `https://www.mdpi.com/2076-3417/12/23/11981`.

[17] **Hendrycks, D.**, **Mazeika, M.**, **Kadavath, S. and Song, D.** (2019). *Using Self-Supervised Learning Can Improve Model Robustness and Uncertainty*, `https://arxiv.org/abs/1906.12340`, `1906.12340`.

[18] **Chen, T.**, **Kornblith, S.**, **Norouzi, M. and Hinton, G.** (2020). *A Simple Framework for Contrastive Learning of Visual Representations*, `https://arxiv.org/abs/2002.05709`, `2002.05709`.

[19] **Rommel, C.**, **Paillard, J.**, **Moreau, T. and Gramfort, A.** (2022). Data augmentation for learning predictive models on EEG: a systematic comparison, *Journal of Neural Engineering*, *19*(6), 066020, `http://dx.doi.org/10.1088/1741-2552/aca220`.

[20] **Drikakis, D.**, **Kokkinakis, I.W.**, **Fung, D. and Spottswood, S.M.** (2024). Self-supervised transformers for turbulent flow time series, *Physics of Fluids*, *36*(6), 065113, `https://doi.org/10.1063/5.0214468`, `https://pubs.aip.org/aip/pof/article-pdf/doi/10.1063/5.0214468/19979767/065113\_1\_5.0214468.pdf`.

[21] **De Fabritiis, F. and Gryllias, K.** (2024). Self-supervised learning approach for anomaly detection in rotating machinery, *Annual Conference of the PHM Society*, *16*(1).

[22] **Tucci, F.**, **Delibra, G.**, **Tieghi, L.**, **Corsini, A. and Lavagnoli, S.** (2023). Unsupervised Learning for High-Fidelity Compression of Large Experimental Dataset: an Application to HPT Blade Tip Contouring.

[23] **Vaswani, A.**, **Shazeer, N.**, **Parmar, N.**, **Uszkoreit, J.**, **Jones, L.**, **Gomez, A.N.**, **Kaiser, L. and Polosukhin, I.** (2017). Attention Is All You Need, *CoRR*, *abs/1706.03762*, `http://arxiv.org/abs/1706.03762`, `1706.03762`.

[24] **Devlin, J.**, **Chang, M.**, **Lee, K. and Toutanova, K.** (2018). BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding, *CoRR*, *abs/1810.04805*, `http://arxiv.org/abs/1810.04805`, `1810.04805`.

[25] **Perez, E.**, **Strub, F.**, **de Vries, H.**, **Dumoulin, V. and Courville, A.** (2017). *FiLM: Visual Reasoning with a General Conditioning Layer*, `https://arxiv.org/abs/1709.07871`, `1709.07871`.

CURRICULUM VITAE

Mehmet Fatih Reyhan:

EDUCATION:

- **B.Sc.:** 2022, Istanbul Technical University, Faculty of Aeronautics and Astronautics, Aeronautical Engineering

PROFESSIONAL EXPERIENCE AND REWARDS:

- 2022-2025 TUSAŞ ENGINE INDUSTRIES Inc, Aerodynamics Engineer.

PUBLICATIONS, PRESENTATIONS AND PATENTS ON THE THESIS:

- **Reyhan, M. F.**, Mehel, Y. S., Zafer, B. (2025). A Deep Learning Based Modelling of Axial Turbomachinery. *4th International Graduate Research Symposium | IGRS’25 İstanbul Technical University*, May 12-14, 2025 Istanbul, Turkey. (Presentation Instance)