POLITECNICO MILANO 1863

SCUOLA DI INGEGNERIA INDUSTRIALE
E DELL'INFORMAZIONE

Executive Summary of the Thesis

A Study on Secondary Flows in Turbine Linear Cascades Using a Data-Driven Approach

Laurea Magistrale in Mechanical Engineering - Ingegneria Meccanica

Author: Matteo Bonanni

Advisor: Prof. Paolo Gaetani

Co-advisor: Sergio Lavagnoli

Academic year: 2025-2026

1. Introduction

In the context of modern turbomachinery design, the minimization of aerodynamic losses is of fundamental importance for increasing their efficiency and reducing fuel consumption. Among the loss mechanisms in axial turbines, complex secondary flows three-dimensional vortical structures, generated by the interaction between the endwall boundary layer and the blade, account for approximately one-third of their total loss.

Traditional design approaches mostly rely on empirical correlations (Ainley-Mathieson, Kacker-Okapuu, etc...[1], [2], [3], [4], [5], [6]), coupled with computationally expensive Computational Fluid Dynamics (CFD). Empirical correlations are fast but often inaccurate for novel geometries; on the other side, CFD is generally accurate but too slow for extensive design space exploration.

The objective of this work is to develop a **Data-Driven design strategy** that combines the accuracy of CFD with the speed of empirical models. This is achieved through the development of a custom automation tool capable of generating a massive, high-fidelity database to train Machine Learning (ML) surrogate models.

2. The Blade Processing Tool

The core contribution of this thesis is the development of the *Blade Processing Tool* (BPT), a modular software written in Python that automates the entire CFD simulation pipeline. The BPT was designed to be robust, modular, and capable of handling GPU-accelerated solvers to drastically reduce computational time.

2.1. Architecture

The BPT workflow is structured into nine sequential "Steps", orchestrated by a series of modules called on demand, and a main control script. This architecture ensures that hundreds of cascade configurations can be processed without User intervention. The workflow is depicted in Figure 1 and detailed below:

- **2D Geometry Generation (Steps 0-1):** The process begins with the bi-dimensional blade profile generation. Using the 2D Blade Generator [7], profiles are created based on aerodynamic loading distributions and style, namely, what are hereafter defined as the input design parameters. A filtering logic automatically discards geometries that do not meet specific criteria set by the author of the blade generator.

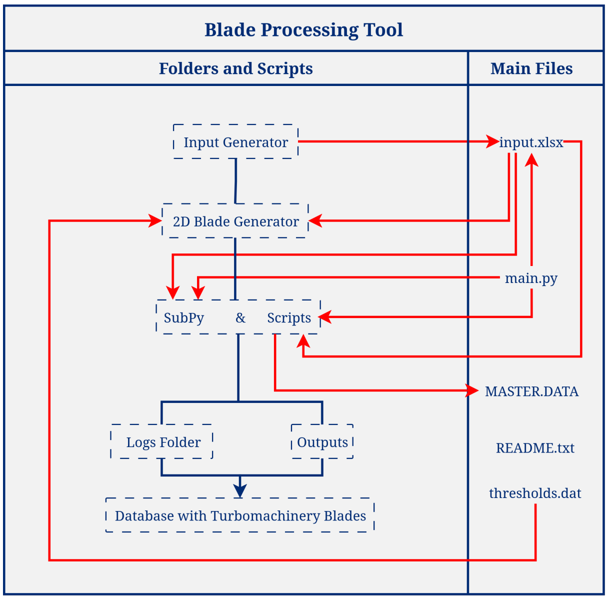

Figure 1: Logical workflow of the Blade Processing Tool (BPT), highlighting the interaction between files and folders in the tool's architecture.

- **3D Domain Construction (Steps 2-3):** The 2D profiles at hub, mid-span and shroud are stacked to create a 3D linear cascade. Crucially, the tool implements an optimization loop based on the Sharma & Butler model [8] to determine the optimal aspect ratio (h/C_{ax}). The work of Benner et al. [9] was also consulted in this context. An optimized aspect ratio ensures that the computational domain is high enough, along blade's radial direction, to prevent secondary flows from contaminating the mid-span region, a requirement for accurate loss decomposition [6].
- **Automated Meshing (Steps 4-5):** The tool interfaces with *ADS CodeWAND* to generate structured multi-block grids (O-H-H topology). A template-based approach allows the BPT to dynamically adapt the mesh to different blade geometrical requirements, while maintaining the closest cell size ensuring $y^+ < 1$, essential for accurate turbulence modeling. The template mesh has been chosen from the results of a mesh sensitivity analysis.
- **GPU-Accelerated Simulation (Steps 6-8):** Simulations are performed using *ADS CodeLEO*. A steady-state RANS approach with the $k-\omega$ Wilcox 98 turbulence model is employed. A key achievement for this work was the transition from CPU to GPU solver (NVIDIA RTX series), which yielded an acceleration factor of ≈ 14 (simulations are solved 14 times faster than on a CPU). The simulation time per case dropped from over 70 minutes (CPU) to approximately 5 minutes (GPU).
- **Post-Processing (Step 9):** A dedicated Python module supported by several sub-modules, extracts flow data at relevant planes (at inlet, outlet, through-passage, and mid-span). It computes mass/area/pitchwise-averaged quantities and separates profile losses from endwall losses as indicated in [6].

3. Database Construction

To train the surrogate models, used at the end of the work to predict secondary losses on a wider region of the space design, a comprehensive database of turbine cascades was generated. The sampling strategy adopted was a **Full-Factorial Design of Experiments (DoEs)**, chosen for project-related reasons.

3.1. Design Space Definition

The database explores a 6-dimensional design space defined by the following parameters:

- **Inlet Flow Angle (α_1):** It varies the orientation of the incoming flow vector, inducing a change in the blade's operative point, for a fixed geometry.
- **Outlet Flow Angle (α_2):** Determining the required flow turning and hence the blade's aerodynamic loading.
- **Isentropic Outlet Mach Number ($M_{2,is}$):** Strictly covering subsonic regimes.
- **Dimensionless Loading Parameters (l_p, M_p):** Controlling position and magnitude of the peak fluid velocity on suction side.
- **Inlet Boundary Layer:** Variation between "thin" and "thick" boundary layers to include their effect on secondary losses estimation.

The combination of these parameters resulted in **864 unique simulations**. The generation of this database required approximately one week of non-continuous simulations on a single workstation (running 7 simulation batches), a task that would have taken months with standard CPU solvers.

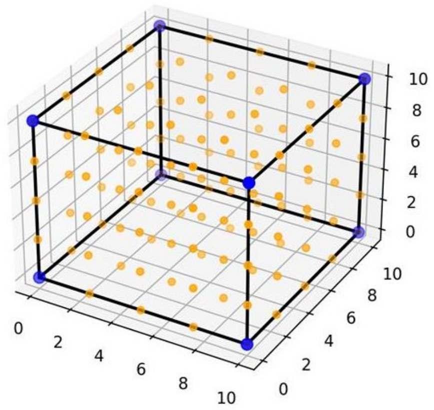

Figure 2: Graphical representation of a simple 3-dimensional full-factorial DoEs: the blue dots are the corner points, while the yellowish ones are the inner points.

4. Data-Driven Analysis

The generated database serves as ground truth for training Machine Learning models. The goal is to predict the **Endwall Loss Coefficient** ($\zeta_{endwall}$) as a function of the six input parameters. First, losses are extracted from CFD simulations with the methodology described in the previous section, and then stored in a text file (`losses.txt`). This file represents the "database" used for training the surrogate models.

Two surrogate models were implemented and compared:

1. **Gaussian Process Regression (GPR):** A probabilistic kernel-based method. The kernel was composed of a Radial Basis Function (encoding the physics of the problem) and a White Kernel (damping the noise in data, typical of CFD simulations results).
2. **Radial Basis Function Network (RBFN):** A neural network approach suitable for function approximation.

The models were trained on 70% of the dataset and tested on the remaining 30%.

5. Results and Validation

The framework was evaluated on three levels: validation of the CFD methodology, physical consistency of the database, and accuracy of the surrogate models.

5.1. CFD Validation

The BPT's numerical setup was validated against experimental data from the **SPLEEN** linear cascade [10], instrumented and tested for a long time at the von Karman Institute for Fluid Dynamics. As shown in Figure 3, the automated tool's CFD setup correctly predicts the blade loading distribution (M_{is}). Visible discrepancies near the trailing edge, suction side, are attributed to the fully turbulent nature of the RANS turbulence model, which does not capture laminar separation bubbles, but the global aerodynamic behavior is well-reproduced.

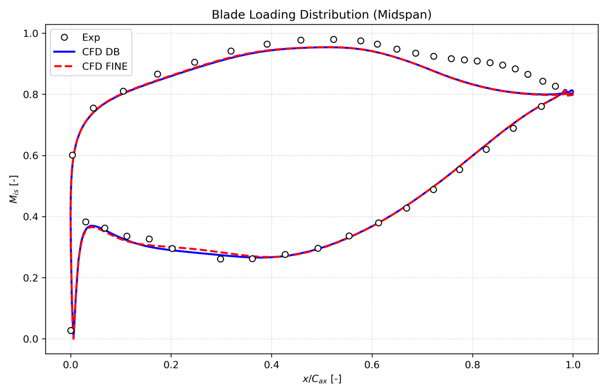

Figure 3: CFD Validation: Comparison of the Isentropic Mach Number distribution along blade's profile between two BPT simulations (one with the datbase mesh, one with a finer mesh), and SPLEEN experimental data, available in [10].

5.2. Database Physical Trends

Post-processing the 864 cases revealed coherent physical trends. Figure 4 displays the endwall losses map, as a function of deflection and outlet Mach number. The results confirm that **flow turning** (deflection) is the primary driver of secondary losses for the cases populating the database. Higher turning increases the cross-passage pressure gradient between the pressure side of a blade and the suction side of the immediately adjacent one, which strengthens the passage vortex, leading to higher mixing losses, and thus higher secondary losses. This aligns with theoretical expectations ([5], [6]), confirming the correct capture of the global physics of the database cascades. Nevertheless, Figure 4

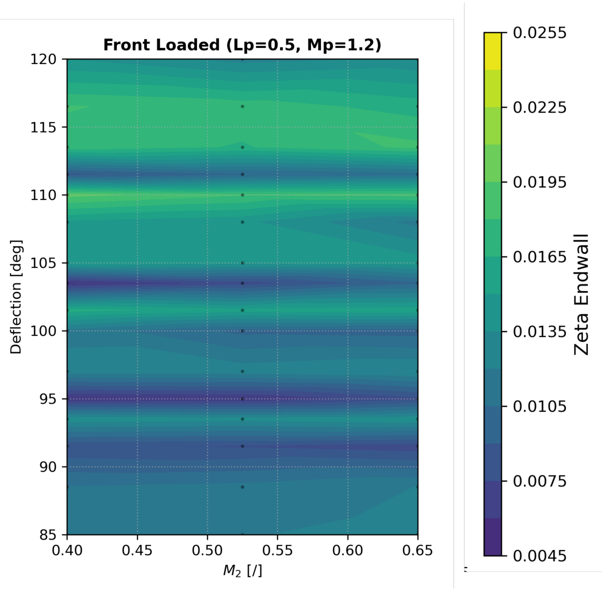

Figure 4: Database Analysis: Map of Endwall Kinetic Energy Loss Coefficient ($\zeta_{endwall}$) varying with deflection and outlet Mach number for a selected blade geometry: front-loaded configuration.

displays a non-continuous trend vertically. The potential causes include: blades parametrization, optimized for respecting a specific load condition, resulting in a profile shape potentially inducing high profile losses, reducing the related endwall share; nature of the template mesh topology (structured), creating discrete regions in the fluid domain with bad quality cells for specific blades geometrical configurations; poor sampling of the database due to limited time and computational resources.

5.3. Surrogate Models Performance

Both models achieved good predictive accuracy, with an $R^2 \approx 0.94$ and a Mean Squared Error (MSE) in the order of 10^{-6}. However, the GPR demonstrated superior performance for this specific study case.

As illustrated in Figure 5, both GPR (above) and RBFN (below) show the model's good capability to correctly predict test data. In the GPR plot can be noted that each prediction is coupled with a gray band along the prediction axis. It represents a **95% Confidence Interval**. This feature is important, allowing the designer to distinguish between a physical change in losses and a model prediction error due to lack of data in a specific region of the design space.

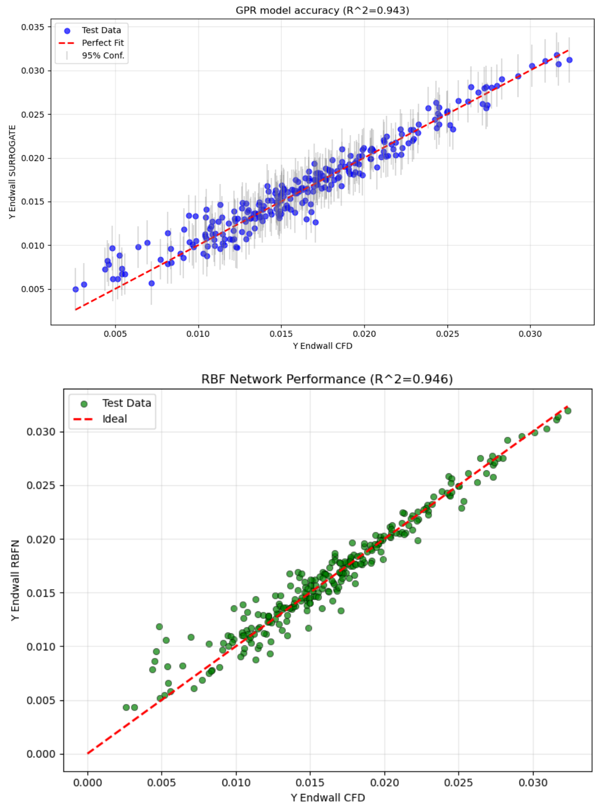

Figure 5: Models accuracy: Comparison of GPR and RBFN models accuracy. Both provide good results, showing predictions to be very close to test data.

6. Conclusions

This work aims to demonstrate the feasibility and effectiveness of a fully automated, data-driven framework for turbine cascade design.

- The Blade Processing Tool proved to be a robust solution, capable of autonomously generating and simulating a wide range of turbine blades, different in geometry and aerodynamic load.
- The integration of GPU acceleration was identified as a relevant technological improvement, reducing computational time by an order of magnitude.
- The Surrogate Models were capable of capturing the secondary flows physics, offering a rapid, yet viable alternative to compete with CFD-based inverse design.

Future developments will focus on extending the BPT to annular cascades to capture radial pressure gradient effects. Besides, the implementation of the logic required to process twisted geometries will augment the tool's complexity and functionalities. Future efforts could be devoted to the integration of a multiple mesh templates logic to enhance the quality of simulation results, but also the potential adaptation to compressor blade processing.

References

[1] Ainley, D. G., & Mathieson, G. C. R. (n.d.). A Method of Performance Estimation for Axial-Flow Turbines. *Areonautical Research Council.*

[2] Dunham, J., & Came, P. M. (1970). Improvements to the Ainley-Mathieson Method of Turbine Performance Prediction. *Journal of Engineering for Power.*

[3] Craig, H. R. M., & Cox, H. J. A. (1970). Performance Estimation of Axial Flow Turbines.

[4] Kacker, S. C., & Okapuu, U. (1982). A Mean Line Prediction Method for Axial Flow Turbine Efficiency. *Journal of Engineering for Power.*

[5] Coull, J. D. (2017). *Endwall Loss in Turbine Cascades.*

[6] Coull, J. D. (2025). *Turbomachinery Technical Conference and Exposition.*

[7] Porta, F. (2023). *Development of a Machine-Learning-Based Tool for Turbomachinery Blades* (Master's thesis). Politecnico di Milano.

[8] Sharma, O. P., & Butler, T. L. (1986). *International Gas Turbine Conference.*

[9] Benner, M. W., Sjolander, S. A., & Moustapha, S. H. (2006). An Empirical Prediction Method for Secondary Losses in Turbines—Part I: A New Loss Breakdown Scheme and Penetration Depth Correlation. *Journal of Turbomachinery*, 273-280.

[10] von Karman Institute for Fluid Dynamics. (2024). *SPLEEN Database.* [Online]. Available: https://zenodo.org/records/13712768

[11] Scarimbolo, C. (2025). *Development of a data-driven design tool for turbomachinery blades: Exploration of secondary flows in 3D turbine cascades* (Master's thesis). Politecnico di Torino.

[12] Coull, J. D., & Clark, C. J. (2021). The Effect of Inlet Conditions on Turbine Endwall Loss.

[13] Clark, C. J. (2019). A Step Towards an Intelligent Aerodynamic Design Process. *ASME Turbo Expo 2019.*

[14] Denton, J. D. (1993). Loss Mechanisms in Turbomachines. *Journal of Turbomachinery.*

[15] Roache, P. J., et al. (2008). Procedure for Estimation and Reporting of Uncertainty Due to Discretization in CFD Applications. *Journal of Fluids Engineering.*

[16] Yondo, R., Andres, E., & Valero, E. (2017). A review on design of experiments and surrogate models in aircraft real-time and many-query aerodynamic analyses. *Progress in Aerospace Sciences*, 1-39.

[17] Wang, H., et al. (2024). Recent Advances on Machine Learning for Computational Fluid Dynamics: A Survey.

[18] AeroDynamic Solutions, Inc. (2024). *ADS CFD Inc.* [Online]. Available: https://new.aerodynamic-solutions.com/