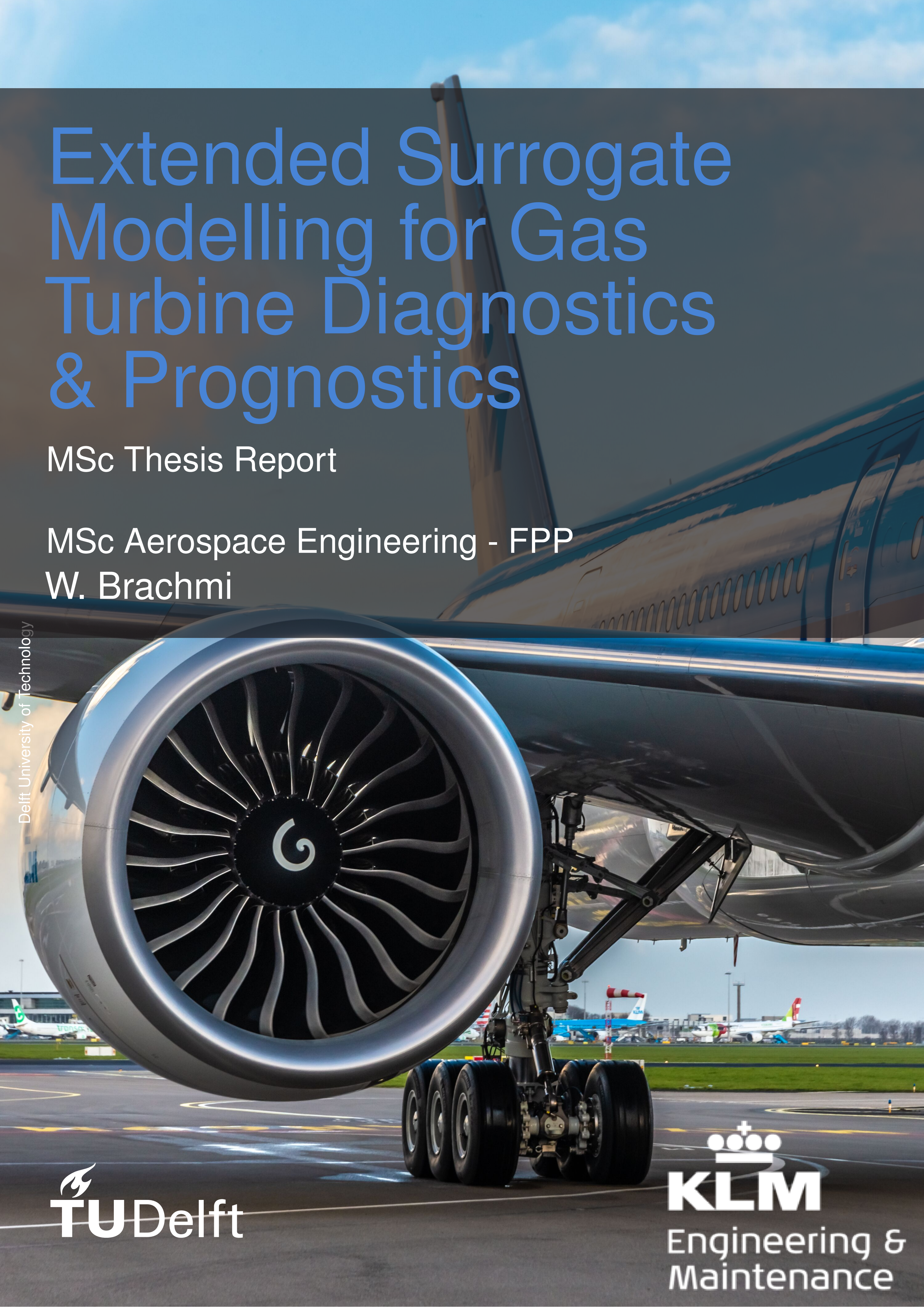


Extended Surrogate Modelling for Gas Turbine Diagnostics & Prognostics

MSc Thesis Report

MSc Aerospace Engineering - FPP
W. Brachmi

Delft University of Technology

TUDelft

KLM
Engineering &
Maintenance

Extended Surrogate Modelling for Gas Turbine Diagnostics & Prognostics

MSc Thesis Report

by

W. Brachmi

Student Number:	4554337
Project Duration:	May 2023 - Apr 2023
Institution:	Delft University of Technology
Place:	Faculty of Aerospace Engineering, Delft

Thesis Committee:	Prof.dr.ir. P. Colonna di Paliano	TU Delft, Chair
	Dr.ir. W.P.J Visser	TU Delft, Supervisor
	Ir. P.C. Roling	TU Delft
	Ir. T. Rootliep	KLM Engine Services, Supervisor

Cover Image: KLM Boeing B777 (Mark Wagtendonk)

TU Delft

Preface

This thesis marks the end of my graduate project at KLM Engine Services and my time as a Master student at the TU Delft. I would like to thank my supervisors at KLM, Ir. Tim Rootliep and Ir. Juan Reguiero, for there guidance and support throughout the project's duration. I am extremely grateful for having the opportunity to do this project at KLM Engine Services, from which I was introduced and learned so much about the MRO industry and given me the privilege of seeing aero-engines up close.

During my time at KLM ES, I had the privilege of meeting some outstanding colleagues at the Data Team. To Albert Timmer, Antonis Paradeisanos, Yair Brouwer and Leandro Rodrigues, thank you for having always made me feel at home and part of the team. I will most definitely miss our lunchtime walks through the hangars and the inexplicable amount of coffee we collectively had together. I also want to thank my thesis supervisor, Dr. Wilfried Visser, for his guidance and feedback whenever I needed it. My gratitude also goes out to Dr. Piero Colonna and Ir. Paul Roling for presiding in my graduation committee.

My deepest gratitude goes to my parents Bouchra and Mostapha, who always believed in me and provided me the opportunity to come here and achieve my dreams. I also would like to thank my brother, Aniss, for always being there when I needed a friend, and Teodora for your constant support.

W. Brachmi
Delft, April 2024

Executive Summary

As turbofan engine technology continues to evolve, aircraft engines still require periodic inspections and maintenance. Maintenance costs still represent a significant part of an aircraft's operational costs, therefore airlines are looking to implement the next step in maintenance strategies, called engine condition-based maintenance (or ECBM). ECBM involves assessing the current condition of engine parts and subsystems, while providing insight on its future condition and would contribute to reduced failure rates, increase operational safety and allow operators to plan maintenance tasks more effectively while reducing time engines spend on the ground. This involves continuous engine diagnostics, a steady stream of measurements and both sufficiently accurate numerical engine models for condition diagnostics and prognostics.

KLM Engine Services provides maintenance services for several engine types, such as the General Electric GEnx-1B. Research at KLM ES in ECBM has been ongoing for several years, focusing on building engine models for condition diagnostics based on non-linear gas path analysis (or GPA). An accurate model of the GEnx-1B was built using the Gas turbine Simulation Program (or GSP) based on test-cell data. The GEnx-1B provides large amounts of engine measurements in the form of Continuous Engine Operating Data (or CEOD) collected throughout flights, however fewer sensors available on the GEnx-1B compared to previous generation turbofans further complicates the application of physics-based methods. Development of a surrogate model was conducted at KLM ES based on the High Dimensional Model Representation (or HDMR) approach, capable of processing the large volume of data in CEOD at a lower computational cost, while being able to cope with measurement noise and bias. Given its low computational complexity, a surrogate model would also allow for real-time condition monitoring for all engines in the KLM fleet. The result of was a set of HDMRs capable of conducting component-wise diagnostics in terms of component efficiencies ($\Delta\eta$) and flow capacity (ΔW_c) and were trained using simulated data from the GEnx-1B's GSP model. Furthermore, HDMRs were also developed to provide engine system level diagnostics, in terms of component influence on the deviation of engine health indicators commonly used in MRO, such as the exhaust gas temperature (ΔEGT), fuel flow (ΔW_f) and thrust-specific fuel consumption ($\Delta TSFC$).

The goal of this thesis project is to continue development in surrogate modelling for engine condition-based maintenance at KLM ES and implement further improvements to extend the surrogate model's capabilities. Following a review of the research done so far at KLM ES, three aspects of the surrogate model are to be investigated: the implementation of secondary performance parameters, the extension of its operational envelope and an approach to extend its capabilities to prognostics.

The GEnx-1B is equipped with several subsystems, such as bleed flows, variable geometry and turbine clearance control systems that also have a significant effect on engine performance. KLM ES refers to these as secondary performance parameters and measurements of said parameters are also provided in on-wing data in the form of valve and actuator settings. A parameterization of these parameters was introduced in the form of quadratic relations to quantify their effects in terms of component map modifiers. The quadratic relations were fitted following an optimization using a genetic algorithm, with an objective function focused on reducing deviation between the GEnx-1B GSP model output and Snapshot data acquired from the GEnx-1B during the take-off phase. The optimization resulted in improved accuracy of the GSP model, such that the GSP model deviation was within $\pm 0.5\%$. Furthermore, the Standard Deviation in GSP output also showed a reduction but further highlighted the dependency in operating conditions for certain parameters. The quadratic relations were then integrated within the GEnx-1B's GSP model. This allowed for the effect of secondary performance parameters to be included within the the GSP simulated data such that the HDMRs can be trained to also account for them.

The accuracy of the HDMRs after training is measured in Root Mean Squared Error (or RMSE) between the HDMR output and the simulated target data. The implementation of secondary performance parameters and fan speed dependency led to mixed results, with RMSE reductions ranging from 5% to 15% for some HDMRs and RMSE increments ranging from 10% to 30% for most HDMRs. Overall, accounting for more engine system parameters did not result in a universal decrease in RMSE as was initially expected. Verification of the HDMRs was performed by feeding on-wing data to plot component condition over flight cycles and observe the trends. Snapshot data was again used for a particular GEnx-1B that has undergone several water-washes and two shop visits due to HPT-related failures. Despite the mixed results in RMSE, the HDMRs were still capable of capturing the expected changes in component condition. The effect of water-washes on component health deviations remain clearly visible in the trends for the HPC, while the progressive deterioration in HPT condition in terms of $\Delta\eta$, ΔW_c and system health indicators can be clearly identified more than two months prior to the two recorded HPT-related incident. Furthermore, the on-wing trending now sees a reduced number of outliers after the implementation of secondary performance parameters compared to baseline results.

The resulting HDMRs were then used as a starting point for the prognostics model. The same Snapshot data was fed to the HDMRs to generate a time-series dataset of average component-wise condition. The dataset was split into a training and test subset, where the test subset was used to assess the LSTM model's predictive capability prior to known engine component failures. Development began with a baseline model to be later tuned after a sweep of various model parameters, such as the number of cells, batch size and sequence length. Results showed that the LSTM model was able to capture long-term dependencies and learn the complex patterns and non-linear trends in the training dataset. Predictions remained relatively close to expected condition trends seen in the test subset, up to 10 flight cycles into the future before predictive accuracy started to decrease. This is particularly noticeable for components that exhibit large and sudden changes in component health. Furthermore, it is worth noting that the LSTM model's development was limited to a specific case scenario; data from one specific engine was used throughout since it experienced a series of documented component failures.

Overall, the framework proposed in this thesis has demonstrated the potential advantages of combining the surrogate and prediction model for engine diagnostics and prognostics. Accurate diagnostics can be performed using the HDMR-based surrogate model while the same information can be used to inform the LSTM-based prognostics model to make better predictions. While the current iteration of the LSTM-based prognostics model requires refinement and further research to validate these models using a wider selection of engine data encompassing various case scenarios, the conclusions drawn from this thesis still serve as a valuable starting point for future projects in the continuous development of engine condition-based maintenance at KLM Engine Services.

Contents

Nomenclature

Abbreviations

Abbreviation	Definition
ISA	International Standard Atmosphere
KLM ES	KLM Engine Services
KLM E&M	KLM Engineering & Maintenance
(E)CBM	(Engine) Condition-based Maintenance
GPA	Gas Path Analysis
CEOD	Continuous Engine Operating Data
EEC	Electronic Engine Control
MRO	Maintenance, Repair & Overhaul
LPC	Low Pressure Compressor
HPC	High Pressure Compressor
CC	Combustion Chamber
HPT	High Pressure Turbine
LPT	Low Pressure Turbine
GT	Gas Turbine
FOD	Foreign Object Damage
DOD	Domestic Object Damage
EGT(M)	Exhaust gas Temperature (Margin)
OEM	Original Equipment Manufacturer
FADEC	Full Authority Digital Engine Control
EMU	Engine Monitoring Unit
CFD	Computational Fluid Dynamics
SBO	Surrogate-Based Optimization
SBE	Surrogate-Based Emulation
ES	Expert System
HDMR	High Dimensional Model Representation
AI	Artifical Intellignece
EASA	European Aviation Safety Agency
ANN	Artificial Neural Network
KF	Kalman Filter
ML	Machine Learning
MLP	Multi-Layer Perceptrons
RMSE	Root Mean Square Error
RUL	Remaining Useful Life
GA	Genetic Algorithm
GSP	Gas (Turbine) Simulation Program
TSFC	Thrust-Specific Fuel Consumption
PHM	Predictive Health Management

Symbols

Symbol	Definition	Unit
A	Area	$[m^2]$
BPR	Bypass Ratio	[-]
C_p	Specific Heat Capacity @ Const. Pressure	[J/kgK]
C_v	Specific Heat Capacity @ Const. Volume	[J/kgK]
D	Characteristic Diameter	[m]
h	Enthalpy	[J]
LHV	Lower Heating Value	[J]
$\dot{m}$	Mass flow	[kg/s]
N	Rotational Speed	[1/s]
p	Pressure	[Pa] or [kg/ms^2]
PR	Pressure Ratio	[-]
$\dot{Q}$	Heat Rate	[W]
R	Gas Constant	[m^2/Ks^2]
T	Temperature	[° K]
$\dot{W}$	Power	[W]
v	Velocity	[m/s]
δ	Corrected Pressure	[-]
η	Efficiency	[-]
θ	Corrected Temperature	[-]
γ	Ratio of Specific Heats	[-]
ρ	Density	[kg/m^3]

List of Figures

List of Tables

1
Introduction

As the airline industry continues to grow year after year, aircraft operators continue to search for methods to reduce operational costs in order to maintain their positions in such a competitive market. One of the largest contributors of operational costs is maintenance, therefore airlines continue to invest in researching various improvements to their maintenance strategies. The conventional strategy is referred to as preventive maintenance, which involves periodic inspections and part replacement based on a fixed operational life. This approach, while the industry standard, may not be the most efficient. Therefore, airlines are looking to implement what is considered the next evolution, called condition-based maintenance.

Research at KLM Engine Services has been ongoing for several years on the development and implementation of condition-based maintenance, or CBM. For CBM to be successfully implemented, an accurate engine model is required in order to perform diagnostics of engine condition based on data collected. A popular approach for engine condition modelling is known as Gas Path Analysis or GPA, which is capable of quantifying the condition of key components of an engine. While KLM ES has access to a model based on non-linear GPA, its accuracy is lacking in certain conditions and will further degrade due to modern turbofans being equipped with fewer sensors resulting data progressively becoming unavailable. However, like the GEnx-1B, modern turbofans do record an exhaustive stream of continuous engine operating data, or CEOD. Unfortunately, GPA methods are not equipped to process such large amounts of data, therefore new modelling techniques are required to mitigate these limitations.

1.1. Research Question & Objective

Surrogate models have been the subject of research at KLM ES for the past few years. The result of this research was the development of a surrogate model capable of approximating engine deterioration based on CEOD. This method has proven to be quite accurate and its implementation within the current workflow at KLM ES is being investigated. However, this surrogate model is only capable of performing diagnostics based on data recorded from previous flights. Deterioration metrics can only be computed as data becomes available, hence only the current condition of an engine can be analyzed. Therefore, in order to continue towards predictive maintenance, the natural step forward is to revise the current model and investigate potential solutions and methods that would not only improve the current diagnostic accuracy but also to generate predictions of future engine condition based on the data at hand, resulting in a prognostics model.

Based on this, the main research question for this thesis can be formulated as follows:

How can the current surrogate model using CEOD designed for component-level diagnostics be improved to enhance condition-based maintenance?

Following the main research question, a series of sub-questions are provided in order to properly guide and structure this literature review. The sub-questions are listed below.

- What is the current state-of-the-art in the field of turbofan surrogate modelling for condition monitoring at KLM ES?
- What improvements could be implemented to the current surrogate model to increase its accuracy?
- To what extent do the improvements implemented to the surrogate model lead to more accurate results ?

- What is the current state-of-the-art in the field of condition prognostics using surrogate models?
- To what extent can the surrogate model be used to conduct engine condition prognostics ?

Following the research question and sub-questions formulated above, the research objective of this thesis is stated as follows:

The objective of this thesis project is to further the development of the surrogate modelling framework developed at KLM Engine Services for engine condition monitoring, by implementing modifications meant to extend the surrogate model's capabilities and investigate the resulting effects and potential improvements in modelling accuracy

1.2. Report Structure

This thesis report is composed of 11 chapters and is structured as follows. Chapter 2 provides some preliminary background information on gas turbine, a brief history and relevant modelling principles, along with a short description of KLM Engine Services. Chapter 3 outlines the principles behind gas turbine deterioration, which includes a description of the various deterioration mechanisms and how they are quantified, along with an introduction to Gas Path Analysis as the popular modelling approach. Chapter 4 provides an introduction to surrogate modelling and outlines several examples used in the field of gas turbine modelling. Chapter 5 outlines maintenance strategies and introduces the concept of Engine Condition-based Maintenance, followed by the body of research conducted at KLM ES on this topic so far. Then, Chapter 6 discusses several aspects with respect to modelling of the GEnx-1B while outlining the research to be conducted in this thesis with regards to extending the capabilities of the surrogate model, followed Chapter 7 with an overview of the various parameters and sources of data relevant to this research. Chapter 8 outlines the relevant theory and procedure for the parameterization of secondary performance parameters, followed by results on engine modelling accuracy. This parameterization is then carried over to be implemented within the surrogate model, along with other modifications to extend its capabilities. Results and verification of the extended surrogate model are discussed in Chapter 9. Furthermore, Chapter 10 introduces the LSTM model to be used in combination with the surrogate model for engine condition prognostics, along with the results of this experiment and further discussions. Lastly, this thesis report ends with Chapter 11 where conclusions and recommendations for future work are provided.

2

Background Knowledge

This chapter serves as a brief and concise introduction to gas turbines and the background knowledge required relating to the fields relevant for this literature study. A summary of the nomenclature used throughout this literature study is provided in Section 2.3, followed by Section 2.4 to provide the necessary definitions behind performance simulations.

2.1. Brief history of Gas Turbines

Gas turbines have been around since the early 1900s. The earliest iteration of a self-sustained gas turbine that managed to provide some power was in 1905, built by Auguste Rateau. Development on industrial gas turbines continued, especially in Switzerland by Prof. Aurel Stodola at the University of Zurich and the manufacturing company Asea Brown Boveri, known today as ABB. The first gas turbine designed for power generation was then first deployed in 1939 in Neufchateau, Switzerland [1].

Around the same period of development for industrial gas turbines, development of jet engines was already underway in Germany and England. Sir Frank Whittle, an office of the Royal Air Force, was the first to consider the concept of gas turbines as jet engines in 1929 and proceeded to claim a patent for it in 1930. However, the first jet engine to be built was by Hans von Ohain with the support of the aircraft manufacturer Ernest Heinkel, which later built the world's first turbojet powered aircraft in 1939, the Heinkel He-178. Shortly after, the technological advancements made so far led to the introduction of the Rolls-Royce Conway in the 1940; the world's first turbofan engine to enter service[1].

Over the next decades, with technological advancements in aerodynamics, material science and manufacturing methods progressed, turbofans became more and more efficient, making the most dominant type of jet engine for commercial aviation. Increasing fuel efficiency resulted in reduction in fuel consumption and fuel costs, making flying progressively cheaper and more accessible. Advancements in jet engine technology post-WWII was one of, if not the major driving factor that kickstarted the so-called "jet age".

2.2. About KLM ES

Following the merger with Air France in 2004, KLM has pursued the concept of one Air France-KLM Group, two airlines and three core activities: passengers, cargo and engineering & maintenance. The third activity, though not as well known as the airline, is handled by KLM Engineering & Maintenance, firm operating within the Air France-KLM Group alongside Air France Industries. KLM E&M is in fact a major player in the Maintenance, Repair & Overhaul (MRO) sector, with around 12,800 employees and offers technical support, such as (but not limited to) line maintenance, engine overhaul and air-frame maintenance to nearly 200 customers. Engineering & Maintenance is composed of five divisions: Aircraft Maintenance, Component Services, Engine Services, European Pneumatic Component Overhaul and Repair or EPCOR and KLM UK Engineering.

KLM Engine Services (ES) is a part of the Air France Industries - KLM Engineering & Maintenance Group. Engine Services performs nearly 150 overhauls for aircraft and terrestrial turbines annually. The facility can receive multiple engine types such as the popular CFM56-7B, CF6-80E1/C2, the recent GEnx-1B and gradually preparing to receive the LEAP-1A and 1B. Alongside the engine MRO facility, Engine Services also manages a test cell within the Schiphol complex. The test cell is used to conduct performance tests of engines when completing their shop visits. This is to ensure that the engine can provide sufficient power within its operating envelope according to certification rules and customer requirements before it returns to an on-wing nacelle.

2.3. Engine Nomenclature

Below in Figure 2.1 is a visual representation of the major components of a modern turbofan and their locations along the gas path. From left to right along the gas path, the major components typically seen on a twin-shaft turbofan are the Fan, the Low Pressure Compressor (LPC), High Pressure Compressor (HPC), Combustion Chamber (CC), High Pressure Turbine (HPT) and Low Pressure Turbine (LPT).

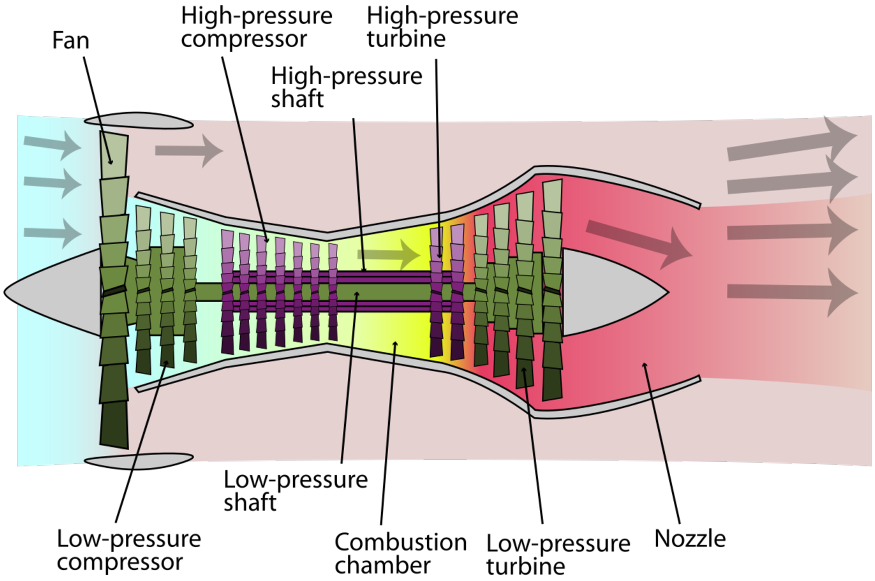


Figure 2.1: Modern turbofan engine architecture [2]

Though the exact numbering may vary in some publications, the numbering of each station along the gas path is also shown below in Figure 2.2 according to the widely accepted Aerospace Recommended Standard ARP 755A [3], which will be used from this point forward in this document.

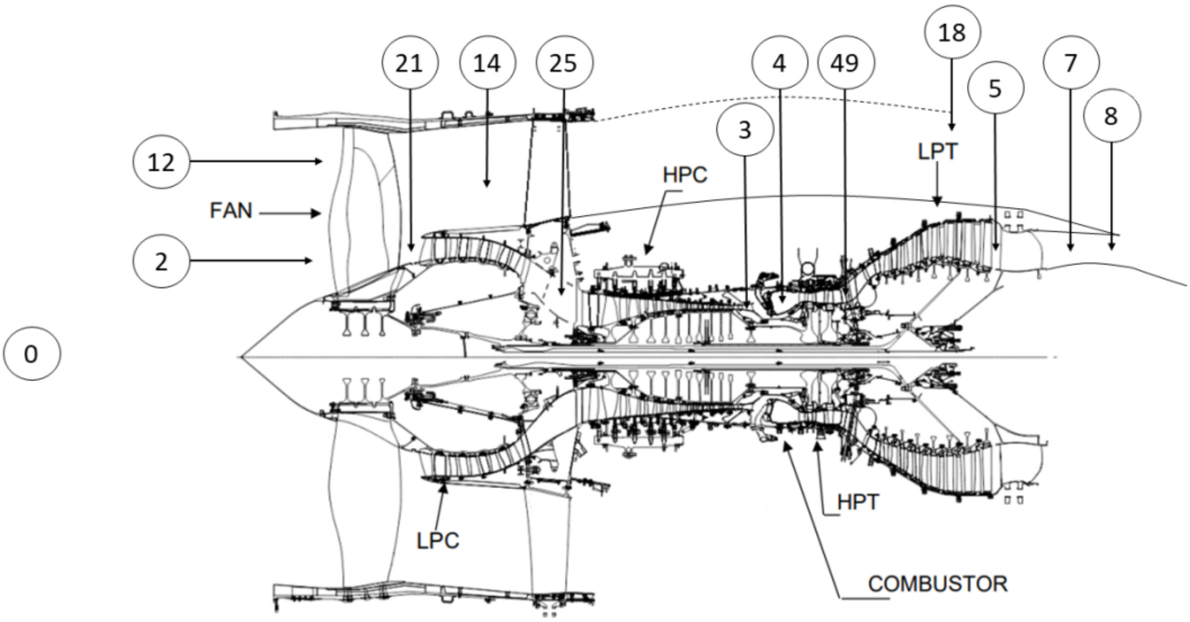


Figure 2.2: Standard gas path station numbering, adapted from GE Aviation [4][5]

Table 2.1: List of engine stations and locations

Station #	Location	Flow Path
0	Ambient Conditions (Free Stream)	Core/Bypass
12	Fan Intake @ Blade Tip	Bypass
2	Fan Intake @ Blade Root	Core
14	Flow Duct	Bypass
18	Nozzle Throat	Bypass
21	LPC Inlet	Core
25	HPC Inlet	Core
3	Combustion Chamber Inlet	Core
4	HPT Inlet	Core
49	LPT Inlet	Core
7	Nozzle Inlet	Core
8	Nozzle Throat	Core

2.4. Performance Simulation

Based on the theory previously outlined, the performance of a gas turbine engine can be modelled and simulated. In GT simulation, there are two simulation regimes that can be defined: design and off-design simulation.

2.4.1. Design Point Performance

In the design phase of a gas turbine, a design point is defined, also often referred to as the cycle reference point. This is the point at which the GT engine geometry is defined and sized. Parameters such as engine configuration, inlet geometry, mass flows, total pressures and temperatures and required component performance level are all defined and selected to meet the required specifications for the design point. In general, the design point for a GT engine is set for the operating condition at which an engine spends most of the time during operations. In the context of performance simulations, the design point as the reference operating point of the system, from which "off-design" simulations are conducted; operating conditions that deviate from the expected design point. The general cycle calculation scheme for design point simulation is outlined in Appendix B.

2.4.2. Off-Design Performance

The calculations presented in the previous section carry a significant assumption: the gas turbine is operating at what is referred to as "steady-state"; no major changes in operating conditions occur. In reality, gas turbines conditions do significantly vary during operation. Once the engine geometry is fixed by design point calculations, the engine performance at other flight conditions must also be evaluated, such as at Sea Level takeoff or during climb. When the operating conditions deviate from the design point, a gas turbine with a fixed design will shift to a new steady-state equilibrium point within its operating envelope. This behaviour can be modelled based on the governing equations from the laws of mass, momentum and energy conservation, shown below, which can then be iteratively solved to return a system of state variables that satisfies the laws [6].

2.4.3. Parameter Groups

Several state variables can be used to describe a system such as a gas turbine. However, in the context of gas turbine performance simulation, the most commonly used variables are listed below in Table 2.2. These are the most prominent system parameters used to describe the performance of the major gas turbine components.

Table 2.2: Performance parameters for gas turbine performance simulation, according to Walsh & Fletcher [3]

Parameter	Definition	Unit
Inlet Total Pressure	$p_{t,1}$	kg/ms^2 (or Pa)
Outlet Total Pressure	$p_{t,2}$	kg/ms^2 (or Pa)
Inlet Total Temperature	$T_{t,1}$	$^\circ K$
Outlet Total Temperature	$T_{t,2}$	$^\circ K$
Mass flow	$\dot{m}$	kg/s
Shaft rotational speed	N	1/s
Characteristic diameter	D	m
Gas Constant	R	m^2/Ks^2

The parameters shown above in Table 2.2 can be described in four distinct fundamental units: temperature, length, mass and time. However, it is often more convenient to work with dimensionless quantities. According to Walsh & Fletcher [3], reducing the relevant parameters into dimensionless parameter groups allows engineers to make better comparisons and quick judgements on the effects of various parameters on a design or performance of a system. In this case, they also help reduce the large number of relevant parameters involved in component performance into specific parameter groups. Specific to gas turbines, there are three groups that often appear in literature: the dimensionless, quasi-dimensionless and corrected parameter groups. According to the Buckingham-Pi theorem, if a system is well defined with n physical variables using k physical dimensions, then the system can be expressed with $n - k$ dimensionless parameters.

With 8 parameters listed in Table 2.2, using 4 physical dimensions, the entire system can be well represented using 4 dimensionless parameters, shown in Table 2.3. Pressure and temperature are now represented non-dimensional as the pressure and temperature ratios, based on the inlet and outlet values. The dimensionless mass flow and spool speed however require the use of other parameters to be made dimensionless. With this, all parameters listed in Table 2.2 are included, allowing for changes in geometry and fluid properties to also be accounted for.

Table 2.3: Non-dimensional parameter group for turbomachine maps [3]

Parameter	Definition	Unit
Pressure	$\frac{p_{t,2}}{p_{t,1}}$	-
Temperature	$\frac{T_{t,2}}{T_{t,1}}$	-
Mass flow	$\frac{\dot{m}\sqrt{RT_{t,1}}}{D^2 p_{t,1}}$	-
Spool speed	$\frac{ND}{\sqrt{RT_{t,1}}}$	-

Table 2.4: Quasi-dimensional parameter group for turbomachine maps [3]

Parameter	Definition	Unit
Pressure	$\frac{p_{t,2}}{p_{t,1}}$	-
Temperature	$\frac{T_{t,2}}{T_{t,1}}$	-
Mass flow	$\frac{\dot{m}\sqrt{T_{t,1}}}{p_{t,1}}$	$ms\sqrt{K}$
Spool speed	$\frac{N}{\sqrt{T_{t,1}}}$	$s^{-1}T^{-1/2}$

If changes in geometry or fluid properties are negligible, then the quasi-dimensionless group can be used. The quasi-dimensionless parameters, as seen in Table 2.4, are not dependent on said changes. This parameter group is similar to those from the fully dimensionless group, however in the case of mass flow and shaft speed, they are not dimensionless.

Furthermore, alongside the non-dimensional and quasi-dimensional groups, there is another group derived from the later that are quasi-dimensional as well. The corrected parameter group, shown below, is used to correct the pressure and temperature measured at altitude to a standard value. In most cases, they are corrected based on the sea level pressure and temperature defined by the International Standard Atmosphere (or ISA) model [7]. The correction ratios for pressure and temperature are shown below in Equation 2.1 and Equation 2.2 respectively. With the correction ratios, the corrected quasi-dimensional parameter group is defined as shown in Table 2.5.

$$\delta = \frac{P_{t,1}}{P_{std}} \tag{2.1}$$

$$\theta = \frac{T_{t,1}}{T_{std}} \tag{2.2}$$

Table 2.5: Corrected parameter group [6]

Parameter	Definition	Unit
Pressure	$\frac{p_{t,2}}{p_{t,1}}$	-
Temperature	$\frac{T_{t,2}}{T_{t,1}}$	-
Mass flow	$\frac{\dot{m}\sqrt{RT_{t,1}}}{Dp_{t,1}}$	-
Spool speed	$\frac{ND}{\sqrt{RT_{t,1}}}$	-

2.4.4. Component Maps

In gas turbine performance simulation, quantifying the performance at conditions away from the design operating point is generally done using what is called component operating maps. These define the whole operating envelope of a gas turbine as a function of four parameters: the pressure ratio on the vertical axis, the corrected mass flow on the horizontal axis, the efficiency of the component and corrected shaft speed. Below are examples of said component maps, with Figure 2.3 showing an axial compressor map and Figure 2.4 for an axial turbine.

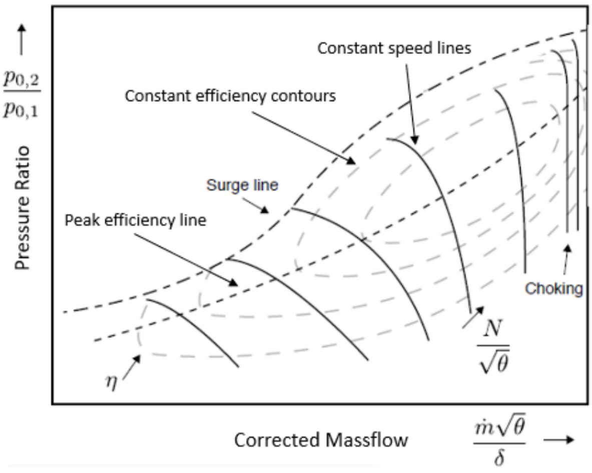


Figure 2.3: Typical operating map for a compressor [8]

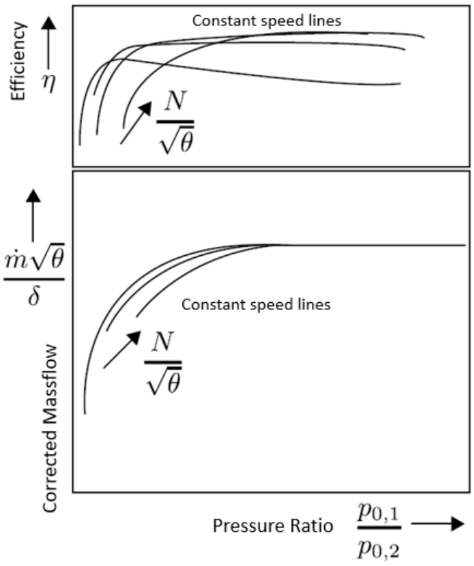


Figure 2.4: Typical operating map for a turbine [8]

3

Deterioration & GPA

Any machine ever designed and produced experiences deterioration of its components over its service life, including gas turbines. Over time, as component deterioration progresses, its performance will deteriorate as consequence, which is why consistent maintenance is required. In this chapter, the most prominent deterioration mechanisms of gas turbines are outlined in Section 3.1, followed by an explanation in Section 3.2 on their effects on engine performance and how they are quantified. Furthermore, this chapter introduces the framework of Gas Path Analysis in Section 3.3, a widely adopted approach in engine performance modelling.

3.1. Deterioration Mechanisms

Throughout their operational lifetime, gas turbines can experience various forms of component degradation, ranging from progressive material fatigue to unforeseen events such as foreign object damage resulting from a bird strike for example. In the context of gas turbine simulation however, only the most prominent deterioration mechanism resulting from operation will be considered. Furthermore, deterioration is generally classified in two categories:

- **Recoverable:** degradation that can be reversed by maintenance without requiring part repair or replacement.
- **Non-recoverable:** only reversible via part repair, replacement or complete engine overhaul.

The following subsections outline in more detail the most prominent deterioration mechanisms commonly seen throughout their operational lifetime.

3.1.1. Erosion

During operation, especially for gas turbine engines on aircraft, there is often various particles ingested. These solid foreign particles then impact components along the gas path, causing the removal of material via abrasion. This degradation mechanism is called erosion and since it involves material removal, it is classified as non-recoverable degradation [9]. Erosion generally involves ingestion of particles around 10 μm or larger in diameter [10] and can include sand, dirt, carbon particles such as ash, etc. Therefore, erosion is common for gas turbine engines operated by aircraft that commonly fly in sandy or dusty areas such as near deserts. To counteract erosion, a common method is to filter the ingested flow, often done with land-based turbines [9] but more difficult or simply infeasible for aero-engines.

3.1.2. Abrasion

Primary components along the gas path such as the compressors and turbines rotate and often experience abrasion with stationary parts of the engine, causing material to wear off. Abrasion generally leads to seal wear and tip clearance growth. Most aero-engines are designed with to have sufficient clearance between the blade tips and the surrounding shroud, however with thermal expansion and large rotational loads, blades can expand and abrade with the shroud. As the distance between the tip and shroud grows, a portion of the air bypasses the compressor leading to reduced efficiency [10]. There are tooth-shaped seals to reduce the clearance in areas where this occurs, however after a certain period, those seals wear off and the clearance starts to grow [10]. Given that abrasion involves material loss, it is considered a non-recoverable form of deterioration. Reversing performance deterioration due to it will require fully replacing damaged components.

3.1.3. Fouling

Depending on atmospheric and weather conditions, aero-engines often ingest various particulates. Over time, particles can accumulate and adhere to surface of components along the gas path. This is referred to as fouling [11] and it can alter geometry such as compressor blade profiles and increase surface roughness. In turn, the mass flow rate decreases leading to a reduction in pressure ratio and efficiency [10]. Fouling according to Ogaji [11] is responsible for more than 70% of the performance loss experienced during operation. Luckily, fouling is a reversible process, therefore the deterioration is for the most part recoverable via water washes [11].On-wing water washes are generally conducted when the mass flow rate has decreased by 2.5% from the nominal rate [12]. Over time however, contaminants can be transported downstream the engine and deposited into the later stages, where due to limited access, parts must be cleaned thus requiring disassembly.

3.1.4. Corrosion

All metals are susceptible to corrosion, which is a chemical reaction resulting in the material wearing off, leading to cracks and possible component failure . There are various types of corrosion depending on the conditions a given component experiences [11]. In sections along the gas path with increased temperatures, components can experience oxidation or sulfidation; when the metal reacts with oxygen and/or sulfides respectively in the flow and corrodes. Corrosion is generally classified as either cold or hot corrosion. An example of cold corrosion is when certain contaminants come into contact with water and corrode material, generally seen in compressor stages. Further downstream, hot corrosion is more common. High temperatures can help jump-start or accelerate reactions. Hot corrosion is often seen in stages after the combustor where high temperatures leading to the corrosion of turbine blades [9]. The result is increased surface roughness, increasing boundary layer loss, and changes in the cross-sectional area impacting mass flow [13]. Furthermore, as more material is lost, the structural integrity of turbine blades is also compromised if the corrosion levels are too high [12]. A common approach to mitigate corrosion is by applying protective coating on components designed to prevent excessive corrosion [9]. Since corrosion involves material loss, it is also classified as non-recoverable.

3.1.5. Object Damage

Most deterioration mechanisms act over time during operation, however there can also be more random and sudden sources such as object damage. Two types of object damage can be distinguished: foreign object damage (FOD), when outside objects such as runway debris or birds are ingested and damage engine internals, or domestic object damage (DOD), when internal components break loose and travel downstream of the gas path. Both of these cases can lead to serious and often irreversible engine damage [14].

3.2. Quantifying Deterioration

While the major deterioration mechanisms experienced by gas turbines are known, their effects can manifest in various ways, making it difficult to accurately model them. Since they manifest in different locations within the gas path, as well in different magnitudes and are highly dependent on both inside and outside flow conditions, their effects can vary during operation. However, a general statement that can be made is that all forms deterioration lead to an increase in fuel consumption, rotational speed and Exhaust Gas Temperature (or $T_{t,49}$) for a given thrust level. According to Greitzer [15], deterioration of engine components can be quantified by aerodynamic losses experienced within the internal flows. In the case of an aero-engine, deterioration can lead to an increase in losses via the following there mechanism :

- **Profile losses**: entropy increase in the boundary layer over surfaces
- **Secondary losses**: losses due to variation in velocity field within the flow, such as end-wall losses
- **Tip leakage losses**: clearance between blades and shroud induce vortices

A major obstacle however is that losses associated with this aerodynamic phenomena are complicated to model and require computationally expensive fluid dynamics simulations to properly quantify. That being said, in the field of engine performance simulations, component deterioration is generally quantified by relating the deterioration to the parameter groups that define the component maps, previously shown in Section 2.4.4. The most common deterioration parameters are the change in component efficiency $\delta\eta$ and corrected mass flow δW_c, however some authors make use of component efficiency with pressure ratio as well [16].

3.2.1. Effect on Operating Maps

Changes in efficiency and corrected mass flow then modify the shape of a given component map, resulting in a deviation of the expected performance at a given operating condition. These deviations are expressed in percentages and can be used to characterize deterioration of all five primary gas path components, being the fan, LPC. HPC, HPT and LPT. The various deterioration mechanisms outlined previously have the following effects shown visually in Figure 3.1 and Figure 3.2 and are summarized in Table 3.1.

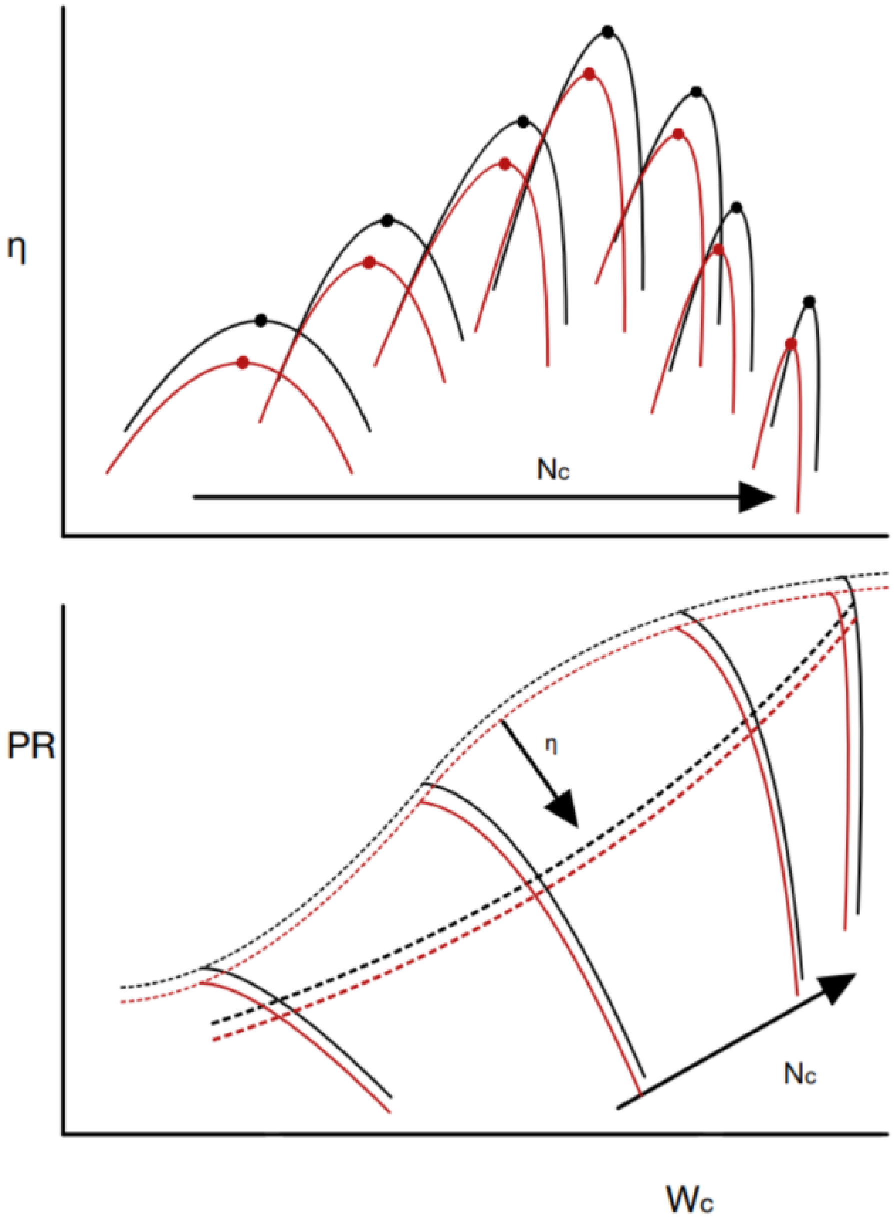


Figure 3.1: Compressor map showing deterioration effects [17]

In the case of a compressor, they all contribute to decreases in $\Delta\eta$ and ΔW_c. Changes in compressor blade geometry result in a reduction in achievable pressurization of the flow and reduced flow capacity in addition to reduced efficiency [17]. Given that the compressor section is the first to meet the oncoming air flow, it is more prone to fouling and FOD and since the compressor is driving the flow, any deterioration in the compressor section results in a decrease in flow capacity. Furthermore, as seen in Figure 3.1, the surge margin is now reduced, given that the surge line has now shifted lower down the PR axis, making it more likely to occur at lower pressure ratios.

As for a turbine, similar effects can be visualized in its operating map, shown in Figure 3.2. In the case of a turbine however, not all sources of deterioration lead to a decrease in $\Delta\eta$ and ΔW_c. Material loss in the turbine in fact increases the flow capacity ΔW_c. Since a turbine extracts energy from an expanding flow, it is driven by the flow and slows it down, as opposed to a compressor. In the case of DOD for example, if an object finds itself in the turbine section, it would be adding material blocking the flow thus reducing the flow capacity.

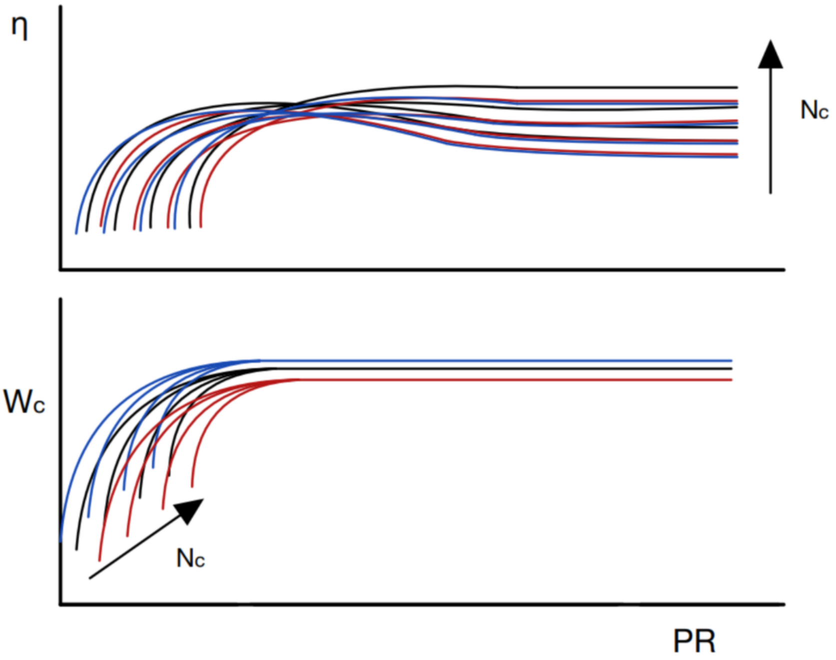


Figure 3.2: Turbine map showing deterioration effects in red [17]

Table 3.1: Effects of degradation mechanisms on components

Deterioration Mechanism	Effect on $\Delta\eta$ Compressor	Effect on $\Delta\eta$ Turbine	Effect on ΔW_c Compressor	Effect on ΔW_c Turbine
Erosion	$\downarrow$	$\downarrow$	$\downarrow$	$\uparrow$
Abrasion	$\downarrow$	$\downarrow$	$\downarrow$	$\uparrow$
Corrosion	$\downarrow$	$\downarrow$	$\downarrow$	$\uparrow$
Fouling	$\downarrow$	$\downarrow$	$\downarrow$	$\downarrow$
FOD	$\downarrow$	$\downarrow$	$\downarrow$	$\uparrow$ or $\downarrow$

3.2.2. Exhaust Gas Temperature

A common metric used in the field of gas turbine monitoring is the temperature of the exhaust gases downstream of the HPT, also denoted as $T_{t,49}$ in the established nomenclature. The Exhaust Gas Temperature (or EGT) is an important parameter since there is an allowable maximum temperature the engine core can reach established for all engines. If an engine were to be significantly deteriorated, a higher maximum core temperature would be detected for the same operating condition. The temperature at the HPT inlet, $T_{t,49}$, is generally not measured directly, given that the temperatures in that section of an engine are too high for any sensor to operate nominally. That being said, the EGT is measured further downstream and each engine has a maximum EGT, reported as the "redline EGT", over which an engine cannot exceed during normal operations. From this redline EGT, the engine health indicator generally in use is called the EGT margin or EGTM; a low EGT margin implies a deteriorated engine in need of a shop visit for maintenance.

$$EGTM = EGT_{redline} - EGT \tag{3.1}$$

EGTM is quite a convenient metric as a general engine health indicator, which is why it is commonly measured and recorded by data acquisition systems within the engine system. The EGTM recorded is corrected for hot-day conditions; the peak EGT at take-off is compared to the redline EGT, the result of which is used to compute the EGTM using physical relations that are not explicitly shared given that they are proprietary knowledge held by General Electric. That being said, an approximation of said relations was reverse-engineered by Röell at KLM ES [18]

3.3. GPA & GSP

3.3.1. Gas Path Analysis

Gas Path Analysis is a well-known method since the 1960s [19] to quantify the degradation of components in turbomachines. GPA is based on the assumption that component degradation leads to a change in engine performance parameters, which in turn can be linked to changes in the working state of a component. GPA has been extensively research since its introduction but its application to aero-engines came much after once the available computing power was no longer a limitation. Figure 3.3 is a visual representation of the principle behind GPA.

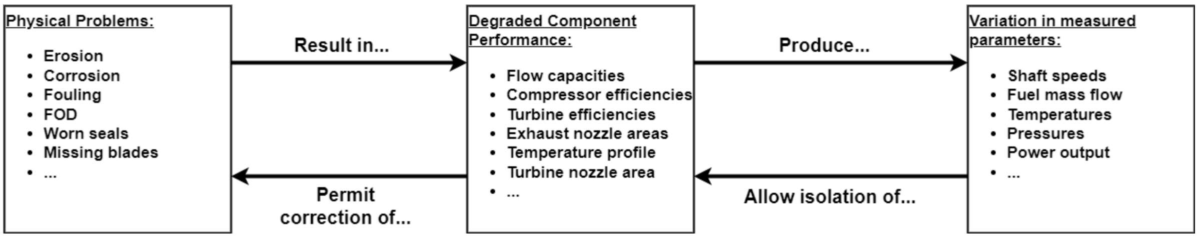


Figure 3.3: Visual representation of the principle behind GPA [20]

The main limitation however with GPA is that the accuracy of whatever engine model is produced is highly dependent on the knowledge and data available to the modellers. Extensive knowledge on the system is required to produce a sufficiently accurate model since it is based on physical relations. In addition, due to the presence of sensor noise and/or bias, measured data may often lead to infeasible or physically nonsensical results which further complicates the model's ability to converge on a suitable and physically consistent solution.

3.3.2. GSP

A popular implementation of gas path analysis is the Gas turbine Simulation Program or GSP, developed by Visser [6] from the Delft University of Technology in collaboration with the Netherlands Aerospace Centre (NLR). GSP was developed with the goal of creating a user-friendly, object oriented simulation program to model various gas turbine configurations, by simply dragging-&-dropping components and connecting them. Below in Figure 3.4 is an example of turbofan model built in GSP.

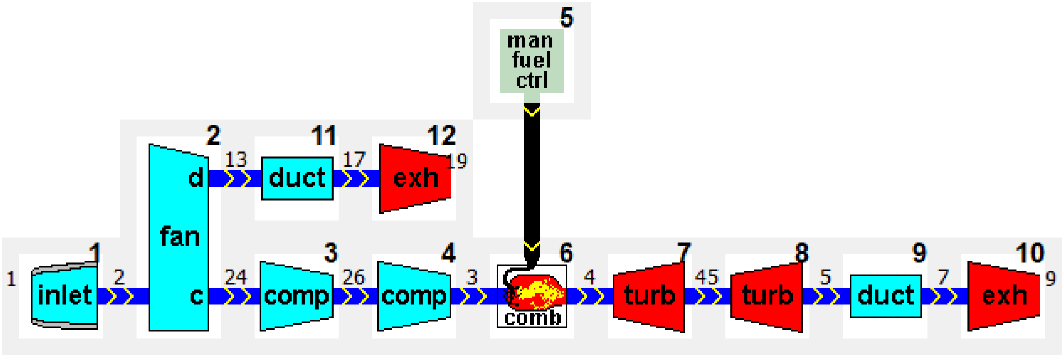


Figure 3.4: Example of turbofan model built in GSP

GSP conducts what is referred to as "0-D modelling", where properties of the modelled system as averaged at every component inlet and outlet along the gas path. This approach allows for simpler model development and low computation time. That being said, there are certain components that are modelled with a 1-D approach where directionality does in fact matter, such as recuperators or multi-reactor combustors. On-design calculations can be performed by directly specifying design point characteristics of the gas turbine model. once a design point is defined, off-design simulations can be conducted based on component maps. Via the component maps, engine degradation can be simulated by modifying the maps using scaling factors which scale the map around the established design point, as described in Section 2.4. Since the component maps of a given gas turbine are proprietary information held by the original equipment manufacturer (or OEM), they are not publicly available. Therefore the maps used in the program by an existing model must be adapted by tuning publicly available maps until the simulated performance matches of the actual engine based on collected data.

4

Surrogate Modelling

The various fields of engineering nowadays have become more and more dependant on computer-aided design and numerical simulations. As the level of complexity of a given system increase however, direct numerical simulations of a high-fidelity model of the system quickly becomes impractical due to the computational cost. A full three-dimensional CFD simulation for example of even a small aircraft in cruise conditions could potentially take weeks to process [21]. An alternative however would be to replace the high-fidelity model with a "surrogate" model. Surrogate models are simple analytical models that mimic the input/output behaviour of complex systems. The advantage is they can approximate the behaviour expected from complex and expensive simulations with reasonable precision, while being computationally cheaper to evaluate. Section 4.1 provides several examples of common applications of surrogate models, while Section 4.2 briefly outlines several types of relevant surrogate models, followed by Section 4.3 which explains the concept of hybrid models.

4.1. Common Applications

Surrogate models are most commonly used in two fields according to the recent literature: surrogate-based optimization and surrogate-based emulation

4.1.1. Surrogate-Based Optimization

In the field of optimization, it is common to see models of complex systems that result in computationally expensive evaluations. For example, the design optimization of an aircraft wing would require multiple iterations involving expensive CFD simulations that could take days to compute. Therefore, the main goal of Surrogate-Based Optimization (or SBO) is, in this example, to replace the current high-fidelity CFD model to reduce the computational cost by using an alternative, lower-fidelity yet reasonably accurate auxiliary model with significantly faster evaluations.[21]. After each iteration, the SBO process uses the results to update the surrogate model, thus continuously training it at each iteration. The SBO process can be represented visually as shown below in Figure 4.1.

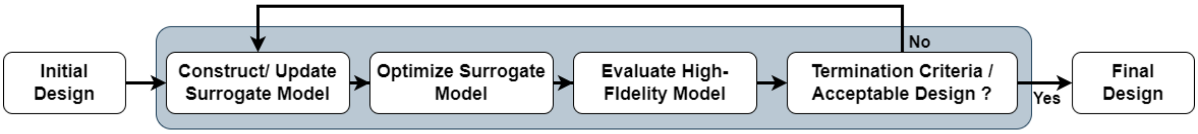


Figure 4.1: Visual representation of SBO [Adapted from [21]]

An example of SBO in literature was by Yan et al. [22], involving the structural optimization of a non-circular vent hole on an aero-engine turbine disk. As surrogate model, Yan constructed a surrogate-based optimization using an improved support vector regression as surrogate model. During this optimization, an initial sample database was continuously updated at each iteration, serving as training for the support vector regression model.According to Yan [22], not only the SBO scheme returned designs of the vent hole with reduced maximum stress but also achieved it with a reduction of time-consuming finite-element simulations by 92%, resulting in increase computational efficiency.

4.1.2. Surrogate-Based Emulation

While Surrogate-Based Optimization makes use of a surrogate model as a tool within the optimization loop, Surrogate-Based Emulation (or SBE) is the process of constructing the surrogate model itself. The goal of SBE is to construct a surrogate model that emulates a physical system or other model deemed too complex or computationally expensive for direct simulations or evaluations.

For example, Luo et al [23]. utilized a convolutional neural network as a surrogate model for the design of a turbo-machine blade. The goal was to accelerate design iterations that are generally held back by extensive CFD simulation with long computation time. The proposed CNN architecture performed reasonably well in predicting pressure loss and deviation angles, which are essential parameters in blade design [23]. Another was Liu et al [24], with the goal of predicting the performance of gas turbines for power generation. Liu et al. utilized multiple surrogate models, such as Artificial Neural Networks (or ANN) and High Dimensional Model Representation (or HDMR), and compared their performance. The resulting HDMRs and ANNs performed quite well in capturing gas turbine performance at part and full load performance conditions.

4.2. Categories of Surrogate Models

Surrogate models can come in different forms, ranging from simple approximation methods such as Least-Square Regression to more complex methods such as Expert Systems. That being said, two categories can be distinguished from literature: physics-based and data-driven models. That being said, there are also examples of models from either category that can be combined into hybrid models.

4.2.1. Physics-Based Methods

Physics-based surrogate models are constructed based on physical relations that define and/or govern the behaviour of the system in question. They are generally a lower fidelity model based on simplified physical relations, leading to simple and faster calculations.

Another advantage of physics-based surrogates is that they require less high-fidelity model data for training in order to return accurate results. This is due to the fact that knowledge of the system's behaviour is inherently embedded in the model. [21]. Despite those advantages, physics-based models still require extensive knowledge of the system at hand to be built, which may significant resources for development if the model at hand is meant to emulate a highly complex system. Another drawback of physics-based surrogates is that, since their response to inputs is based on physical relations, they tend to have a very low tolerance to noise and bias [25].

Furthermore, since the model is based on physical equations, the number of equations available must be equal or more than the amount of variables to be computed in order to fully define the system and obtain accurate results. This issue is particularly relevant in this case, given that most modern turbofans, such as the GEnx-1B, now come equipped with less sensors along the gas path compared to their older counterparts, further complicating the modelling process [26].

Gas Path Analysis

To reiterate what was outlined in Section 3.3.2, GSP makes use of GPA principles based on 0-dimensional, non-linear differential equations that simulate engine behaviour. Given that GSP employs a 0D approach, it operates as a lower fidelity physical model of a gas turbine, in contrast to a more complex, high fidelity model such as a full 3-dimensional CFD simulation. Therefore, GPA-based models, including GSP, can already be classified as surrogate models.

Expert System

Another form of surrogate modelling that has been around for quite some time are Expert Systems (ES). ESs have been around since the 1980s and continue to be used in the field of gas turbine diagnostics [20]. The idea behind an expert system is shown visually in Figure 4.2.

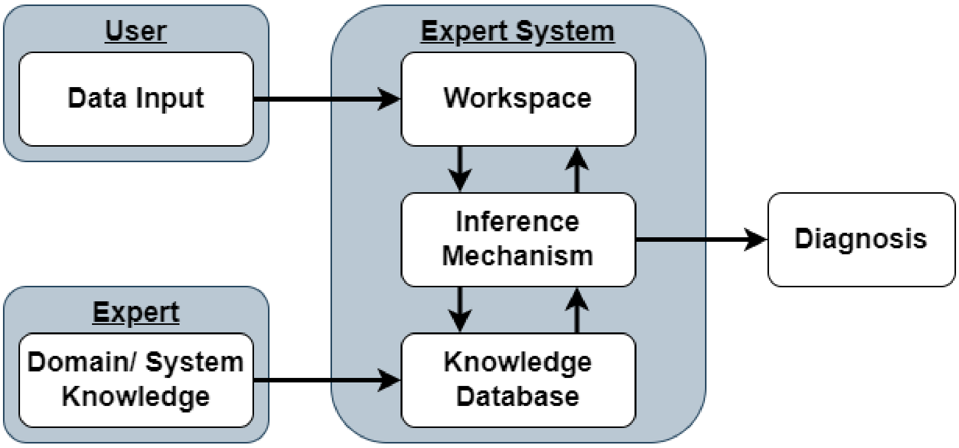


Figure 4.2: Visual representation of an expert system [Adapted from LaRocca [27]]

Expert systems contain a knowledge database built by the developer. An inference engine is then used which accesses the knowledge database in order to infer and formulate a response about the system's state based on the input. The philosophy is to mimic the steps in human reasoning and problem solving, similar to how an expert or engineer would assess the state of the system based on input data or initial information [28]. A current example would be at the expert system used at KLM ES, named PROGNOS: a rule-based expert system which returns failure warnings based on trends from engine measurements, as outlined in Section 5.3.

An expert system can return very accurate results and effective diagnostics, however that is strongly dependent on the accuracy of the content in the knowledge database. Therefore, significant effort and resources must be spent to build, maintain and update the knowledge base, especially for very complex systems such as a gas turbine. For example, tuning PROGNOS at KLM ES requires extensive work, in order to reduce the likelihood of false positives and undetected faults. The complexity of the knowledge database can range from simple "if-then" reasoning to more complicated conditions. However, more often than not, knowledge databases of ESs are mostly composed of rather Boolean logic. Measures of certain parameters such as degradation or anything within a range or spectrum of answers is very complicated to translate in "if/then" rules and often is done inaccurately.

Kalman Filter

The Kalman Filter (or KF) is a physics-based method that uses a set of system equations combined with data inputs to calculate the value of a measured parameter, while accounting for bias and uncertainty. In other words, it is a predictor-corrector technique that utilizes measurements of the current system state in order to approximate the system state at the next time step [25]. KFs are commonly applied in cases where multiple sources of data are present. A notable example is in the field of navigation. Multiple sensor measurements are combined and fused to determine parameters such as position and velocity while overcoming spatial and temporal inconsistencies. The data acquired is often noisy and prone to bias, which is remedied by combining physical relations with internal measurements in order to return physically consistent results [29].

In the field of gas turbines, KFs have also achieved successful results. An example would be that of Liping et al.[30], whom proposed an unscented Kalman-filter (UKF)-based simultaneous diagnostic scheme for gas-turbine gas path and sensor faults. The proposed scheme achieved an accuracy of 97.3% and 93% for sensor faults and gas path faults respectively, while achieving a high degree of accuracy when the magnitude of the detected faults.

The advantage of KFs is that they can still use measurements that likely have noise or bias. Additionally, not all variables must be measured; hidden variables can be estimated iteratively using both previous and current measurements [30]. Despite their relatively high degree of accuracy, the low computational cost and ability to deal with biased and noisy data, highly complex system require increasingly complex KFs in order to properly capture the system's behaviour. Furthermore, KFs are prone to smearing, where the effect of certain components is incorrectly attributed to others [25].

4.2.2. Data-Driven Methods

Functional surrogates, also known as approximation and/or data-driven surrogate models are based solely on data acquired from measurements of the system or from a high-fidelity model simulation [13]. The main advantage of data-driven models is, when compared to physics-based surrogate model, they require either very little to no knowledge of the system at hand. Additionally, since no system characteristics are embedded within the model, they are not restricted to that particular model, making widely applicable to any abstract system as long as sufficient data is available [21]. An example of this aspect would be that an accurate data-drive model of a certain engine can, in principle, be used for a different engine of the same type, provided that the data available is sufficient [31]. There is a wide variety of data-drive methods that can be used for surrogate modelling. The most common in literature have been compiled in the following sections.

While little to no system knowledge is needed, its main advantage also relates to a major inherent drawback. In order to properly capture the behaviour of a system without resorting to prior knowledge of it, sufficient data must be provided and often the requirements on the data-set may be significant. The accuracy of data-driven methods significantly drops when operating outside of the range of values or operating conditions they were trained within. Thus sufficient data is needed to capture a wide range of operating points [25]. Furthermore, data-driven methods while able to capture both linear and non-linear relations between variables, they are often difficult to interpret once trained, arriving at what is commonly referred to as the "black-box" problem [13] [21].

While the drawbacks mentioned are a cause for concern, the advantages of data-driven methods still make them an attractive choice. Compared to physics-based methods, they are more resistant to noise and bias within the data, as well as less computationally expensive by comparison. However, the benefit of low computational cost only appears after extensive and costly training during development [32]. It is worth noting however that, while it is not an insignificant drawback, it is not inherent only to data-driven methods; physics-based method will also require a significant investment of time for development.

Kriging

A common method to interpolating deterministic data is Kriging. The model resulting from this interpolation can then be used as a system model. The mathematical formulation of Kriging is shown below [21].

$$f(\bar{\boldsymbol{x}}) = \boldsymbol{g}(\bar{\boldsymbol{x}})^T\bar{\boldsymbol{\beta}} + Z(\bar{\boldsymbol{x}}) \tag{4.1}$$

Looking at Equation 4.1, $f(\boldsymbol{x})$ represent the model function that emulates the system in question, $\boldsymbol{g}(\boldsymbol{x})$ are known constant functions, $\bar{\boldsymbol{\beta}}$ are unknown model parameters and $Z(\bar{\boldsymbol{x}})$ is a realization of a normally distributed Gaussian random process [21]. The model parameters $\bar{\boldsymbol{\beta}}$ are fitted during training to match the training data such that the error between them is minimized. The regression term $\boldsymbol{g}(\bar{\boldsymbol{x}})^T\bar{\boldsymbol{\beta}}$ is then a global approximation of the function while $Z(\bar{\boldsymbol{x}})$ takes localized variations into account. Another advantage of the Krging approach is that once the model is trained, the random process $Z(\bar{\boldsymbol{x}})$ provides valuable information about the model error that later be used to improve the overall surrogate model [21].

Kriging is quite simple to implement and relatively inexpensive in terms of computational cost. Kriging has been utilized in several occasions in literature for surrogate modelling of gas turbine. Lai et al [33] constructed a Kriging surrogate model for health monitoring of the main fuel pump of a turbofan engine. The Kriging model was trained suing data generated from a physics-based model developed prior to this, resulting in the Kriging model achieving a maximum relative error of 5%. Aulich et al [34] also made use of Kriging within an optimization scheme for the design of a fan stage. However a major disadvantage outlined by Aulich was the the computational cost of training a Kriging model is proportional to N_p^3, where N_p is the number of data points. In other words, if the number of data points were to double, the cost of training would be eight times larger. This would then indicate that Kriging is not suitable for constructing surrogate models defined by very large datasets.

Fuzzy Logic

As opposed to commonly known Boolean Logic mostly seen in expert systems, an iteration of the expert system has led to what is referred to as "fuzzy" logic. Fuzzy Logic systems share many aspects with Expert Systems, as can be seen below in Figure 4.3.

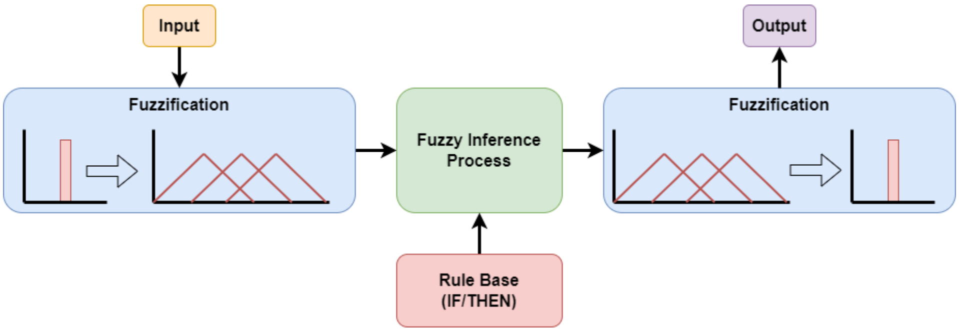


Figure 4.3: Visual representation of fuzzy logic system [Adapted from [25]]

Contrary to traditional expert systems that operate on knowledge databases built on Boolean logic, Fuzzy systems are capable of working with partial or multiple truths. Therefore they are able to infer conditions as expected from any expert system but also including a certain degree of uncertainty in its diagnosis. Fuzzy logic systems are one of the most commonly used forms of artificial intelligence. Generally, fuzzy systems consist of a set of fuzzy rules (generally IF-THEN statements), a "fuzzifier" mechanism which maps inputs to the fuzzy set, an inference engine similar to Expert Systems and a "defuzzifier" to convert the fuzzy output into output values [25].

Fuzzy logic systems have been implemented in field of gas turbine diagnostics in the past. A notable example is Ganguli [35] who constructed a fault detection and isolation model based on fuzzy logic. The goal of the model was to isolate system faults of a jet engine using measurements of N1 and N2 shaft speeds, EGT and fuel flow. The results showed that the model was capable of isolating faults from this limited set of measurements with an accuracy of over 95% accuracy [35], demonstrating that a fuzzy logic system was more than capable despite limited data. Similarly, Ogaji et al. [36] developed a diagnostic system for a modern military turbofan engine based on fuzzy logic. With this, Ogaji et al. were able to detect and identify single component faults with a reported accuracy of 92.5%. That being said, their inherent drawback is the strong dependency in the quality and robustness of the knowledge database.

In addition, fuzzy logic systems have shown to have low computational cost and can handle noisy and biased data. Despite these advantages, fuzzy logic systems suffer from disadvantages similar to those of Expert Systems. ESs depended on an extensive and quality knowledge base in order to establish the rule set [25]. Furthermore, they are limited to the operational range defined by the data and information present in the knowledge base, thus their capability of generalizing is limited to what the knowledge base contains. To mitigate this effect, sufficient data from multiple operating points and conditions is required [25].

Artificial Neural Network

Artificial neural networks (or ANNs), commonly referred to as simply Neural Networks (or NNs), are computing systems inspired by the human brain and nervous system. Neural networks emulate the electrical activity in the brain, where the processing elements known as neurodes or perceptrons are connected to one another in layers, similar to neurons in the brain [37]. Below is a visual representation of a neural network.

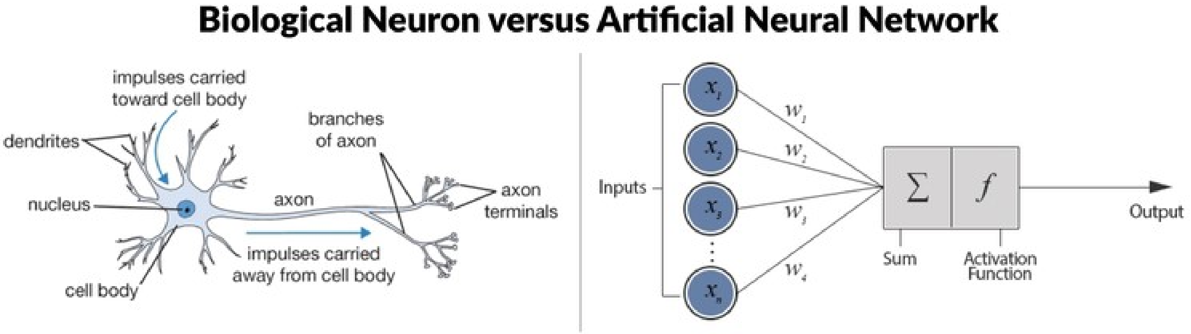


Figure 4.4: Visual representation of artificial neural network [38]

Neurodes are typically arranged in layers, where the output of one layer serves an input to the next layer or layers downstream, analogous to connections between synapses. A neurode can be connected to all neurodes or a subset of the neurodes of the next layer. Each connection between two neurodes has a weight, resulting in a weighted signal meant to simulate the level of electrical excitation of a nerve cell in the brain and subsequently the propagation of information through the network. The inputs to a neurode are then multiplied by the connection weight and either strengthens or weakens the activation of the following neurode, similar to the concept of neural pathways. By adjusting the connection weights between neurons, neural pathways change leading to different outputs, thus emulating the processing of learning [37]. Neural networks come in various forms depending on their purposes and abilities, such Multi-Layer Perceptrons (or MLPs), Convolutional Neural Networks (or CNNs), Probabilistic Neural Networks (or PNNs) or a Bayesian Belief Network [25].

ANNs have appeared several times in literature, often in gas turbine condition monitoring and design optimization For example, Safari et al. [39] presented an MLP network for predicting compressor and turbine outlet temperature of a gas turbine for power generation. The result was a prediction accuracy with an RMSE of 0.20% for the compressor outlet temperature and 0.29% for the turbine outlet. Similarly, Mortda et al. [40] proposed a two-layered feedforward network to predict the performance of a turbojet. The resulting model after training returned accurate predictions of metrics such as specific fuel consumption, specific thrust and a few others with good accuracy. The motivation behind this approach Twas to devise a model to predict turbojet performance over various operational conditions while reducing the need for tedious theoretical calculations or extensive experimental work [40].

Aside from performance prediction, several studies were conducted on gas turbine diagnostics powered by ANNs. Kong et al. [41] proposed an ANN-based system for fault identification in a gas turbine system. Matuck et al. [42] proposed a similar model trained on simulation data. Fast et al [42] developed an MLP-based diagnostics scheme for the purpose of optimizing the schedules of compressor washes. The popularity of neural networks in the field of gas turbine diagnostics increased as evidence of their better early detection ability, higher fault classification rate, and more efficient fault identification continued to be reported compared to other AI techniques. To summarize, ANNs have shown to be a suitable approach for performance prediction and fault diagnostics, assuming that sufficient operational data is available for training [25].

4.3. Hybrid Models

As discussed in the previous sections, physics-based and data-driven surrogate models both have their respective advantages and disadvantages. However hybrid methods; combinations of surrogate models, can provide better results by combining the advantages of some models to compensate for others' disadvantages. Hybrid methods allow the combination of two or more data-driven and/or physics-based methods, thus allowing the use of expert knowledge of the system in question while taking advantage of available data on the system.

4.3.1. Kalman Filter & ANN

An example of a hybrid method utilizing a physics-based a data driven approach is a neural network combined with a Kalman Filter. Peel [43] proposed a model consisting of MLPs and radial basis function (or RBF) neural networks for RUL prediction. The NN-based model would learn the nonlinear mapping between the parameter space and the sensor measurements. A Kalman filter was then implemented to fuse the multiple ensemble outputs over time and achieve an RUL prediction. In a similar fashion, Heimes [44] presented a recurrent neural network combined with an extended Kalman filter to filter the output. The RUL was simulated for a set of aircraft engines was calculated by averaging the outputs of the three best models.

4.3.2. Expert System, Fuzzy Logic & NN

Garga et a. [45] presented an automated inference system that made use of an explicit knowledge base and a feedforward neural network trained on sensor data. The goal was to assess the condition of an industrial gearbox and estimate its RUL, thought the exact method for RUL estimation is not fully described. IN this case, two neural networks were trained; one to learn the explicit rule set and the other for inferring system health based on operational data and measurements that were converted in fuzzy variables.

Another example of this sort of combination was presented by Chinman and Baruah [46], which involved a "neuro-fuzzy" inference approach. The inference model relied heavily on a solid knowledge base, which focused on monitoring and estimating RUL for cutting tools. Parameters such as the thrust force and torque applied on drill bits were predicted using focused time-lagged feedforward neural networks (or TLFN). The outputs of the focused TLFNs served as input to a fuzzy inference model which predicted RUL based on reliability estimations and expert knowledge used to define failure modes from degradation signals using fuzzy inference.

4.3.3. Neural Networks & Genetic Algorithms

Genetic algorithms or (GAs) are heuristic algorithms inspired from the process of natural selection. They are commonly used as optimization algorithms, with the goal of minimizing an objective function over a design space to reach a global minimum. GAs have been combined with ANNs in order to accelerate the learning process. Azzam et al. [47] an evolutionary ANN; a neural network trained using a genetic algorithm in order to avoid local minimas during training, with the goal of modelling NO_x emissions from gas turbines. Results showed that the GA-NN combination performed very well in capturing the non-linear relationship between the inputs and outputs. With respect to component diagnostics, notable examples are from Sampath et al. [48] and Kobayashi [49]. Both authors proposed hybrid models involving neural networks and genetic algorithms. In the case of Sampath, the ANN was used to reduce the number of fault classes to be explored. The GA-based diagnostic model would then explore the space of possible fault classes and locate the optimal fault class that best described the current state of the gas turbine. As for Kobayashi, the NN was constructed for diagnosing component faults while the GA was used to estimate the errors and bias within the measured data. While both authors have reported fairly well results, both have stressed that further investigation is needed and further refinement of both the GA and NN used is required in order to return more accurate results.

4.3.4. High Dimensional Model Representation

The High Dimensional Model Representation (or HDMR) is an approach for modelling high dimensional systems by utilizing expert knowledge and data. The model is essentially constructed using a linear combination of basis functions, defined beforehand based on expert knowledge. Once the basis functions defined, the weights of each term is adjusted through a data-driven optimization. For a system consisting of m dimensions and a vector of variables $\bar{x} = [x_1 x_2 ... x_m]$ as normalized inputs, the HMDR $f(\bar{x})$ can be expressed as shown in Equation 4.2, where a certain $g(...)$ represents a basis function of a certain dimension and g_0 the mean value of $f(\bar{x})$ within the domain in question.

$$f(\bar{x}) = g_0 + \sum_{i=1}^{m} g_i(x_i) + \sum_{1 \leq i_1 \leq i_2 \leq m} g_{i_1 i_2}(x_1, x_2) + \sum_{1 \leq i_1 \leq i_2 \leq i_3 \leq m} g_{i_1 i_2 i_3}(x_1, x_2, x_3) + ... \tag{4.2}$$

HDMRs have been used in the field of gas turbine modelling but are a fairly novel approach. A notable example found in literature is by Safari et al. [39], who also proposed a HDMR-based meta model for predicting the performance of a micro gas turbine based on measured data. The resulting HDMR was capable of predicting measured target value with relatively good accuracy, but required further tuning for certain parameters. While an HDMR itself constitutes a single model type, it can also be combined with other surrogate models to form hybrid models. An example of this is proposed by Jha et al [50] who devised a hybrid method to estimate the failure probability of a given structural subject such as a rotating disk. A first order HDMR was constructed to approximate a certain limit state function. An ANN was then trained using the sampling points from the HDMR in order to reduce the computational cost of evaluating the limit state function. In other words, the function approximated by the HDMR acts as an input of the ANN and the limit function response as the target for the ANN training [50]. Results showed good approximations of the failure probability obtained from direct Monte Carlo simulations at a significantly reduced computational cost.

Given that expert knowledge is inherently embedded into a HDMR, the resulting model has the advantage of transparency. In other words, unlike a deep neural network which is considered a black-box, relations within the model after training are recoverable and can be interpreted. In addition, modifications to the model can be applied with relative ease, such as accounting for additional parameters would only require appending additional basis functions given that the model is a linear combination. While HDMRs seem to have interesting qualities, they do come with drawbacks. The need for expert knowledge is further emphasized by the difficulty in selecting the proper basis functions. Choice of said functions is extremely important for the sake of the model's accuracy and will require extensive work.

5

Maintenance at KLM ES

The KLM ES facility conducts nearly 150 overhauls yearly on a variety of engine types, such as the popular CFM56-7B, CF6-80E1/C2 and the more recent turbofans such as the GEnx-1B, LEAP-1A and 1B. Section 5.1 provides an overview of the common maintenance strategies, while Section 5.2 introduces the Engine Condition-Based Maintenance paradigm. This concept is at the center of the research at KLM ES, known as Smart Workscoping and outlined in Section 5.3. Furthermore, a summary of the research done in engine modelling at KLM ES is provided in Section 5.4

5.1. Maintenance Strategies

In the airline industry, operations are only profitable as long as the aircraft fleet is flying; the longer they spend on the ground, the harder it is to stay competitive in the global industry. Therefore, to maintain aircraft availability and subsequently profitability, adequate maintenance is of the highest priority for an airline [51]. MRO activities account for approximately 10% to 15% of airlines' operational costs; an industry worth approximately $ 67.6 billion in 2016 and expected to grow to more than $ 100 billion by 2026 according to IATA [51]. As explained in Section 3.1, it is necessary for aero-engines to undergo regular maintenance checks and overhauls after a certain number of cycles in order to counteract the effects on performance due to non-recoverable deterioration mechanisms and ensure safe operation. The various strategies used in the MRO industry are visualized in Figure 5.1.

In the early days of the gas turbine and MRO industry, the most common strategy is referred to as corrective maintenance, which involved replacing or parts after failure occurred and was identified. For land-based gas turbines for power generation, this strategy may be still be valid, however for the aviation industry it poses significant safety risks. That being said, preventive maintenance is now the widely adopted strategy. Preventive maintenance introduces a well-defined maintenance intervals and limits on operational lifetime of components, thus reducing the chance of component or system failure during operation. The main drawback of this strategy, however, is that the actual condition of the component and/or module is not taken into account. Therefore it could be that maintenance is conducted on a healthy part but with an overdue shelf life or a part might fail before its schedule maintenance visit [14]. Therefore, the natural step forward from here would be to account for the current condition in the decision-making process when planning and scheduling maintenance tasks. This strategy falls within predictive maintenance and is referred to as "engine condition-based" maintenance (or ECBM).

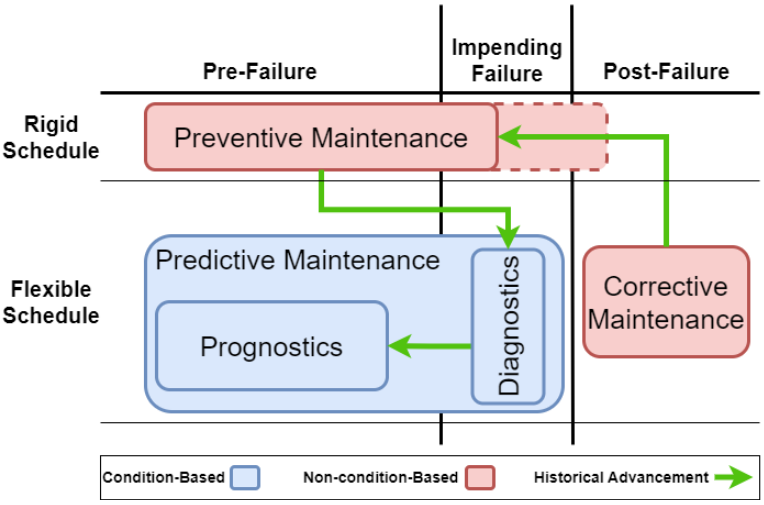


Figure 5.1: Evolution of maintenance strategies for gas turbines [Adapted from [14]]

5.2. Engine Condition-Based Maintenance

Aircraft maintenance, specifically aero-engine maintenance, is a substantial contributor to the operational costs for an airline. Over time, gas turbine maintenance costs can reach multiples of its initial purchase cost. A good example would be the V94.3A gas turbine designed by Siemens, which reached maintenance costs equal to 18 times its initial purchase over its operational life [52]. That being said, the MRO industry is now investing in predictive maintenance, often referred to as Engine Condition Based Maintenance (ECBM), with the goal of extending its operational lifetime and reduce overall costs [52]. This approach has already proven successful for extending the lifetime of terrestrial gas turbines for power generation [13].

The principle behind ECBM is that a given engine's health can be determined from real-time engine measurements. ECBM is categorized as a predictive maintenance strategy, involving two categories of activities: diagnostics and prognostics [14]. Diagnostics involve measuring the current health state of components and identifying faults based on a numerical model of the engine in order to convert sensor readings into quantifiable measurements of engine health. ECBM can also allow the implementation of a prognostics model to perform future projections of engine health based on its current state and previous deterioration trend. Prognostics if implemented correctly, allows for a more streamlined maintenance scheduling and significantly reduce both the economic losses due to system failure and the operational costs attributed to unnecessary repair and replacement of components [13]. The working behind ECBM is also illustrated below in Figure 5.2

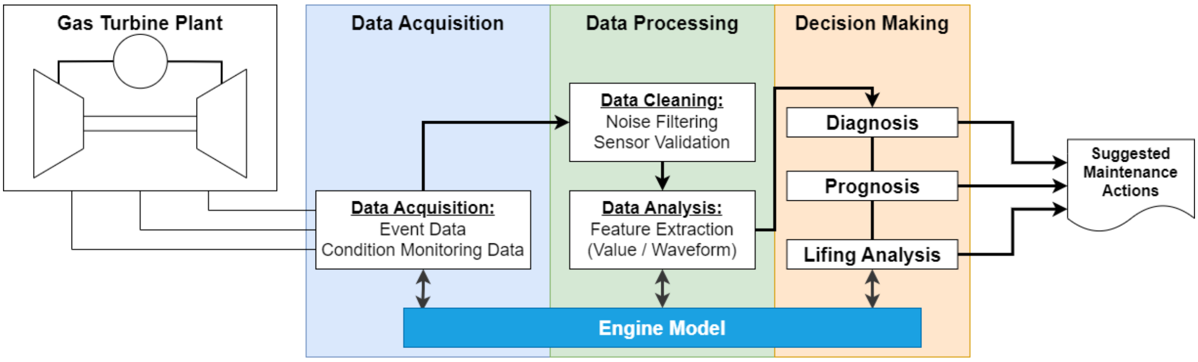


Figure 5.2: Visual representation of ECBM principle [Adapted from [53]]

For the MRO industry, the implementation of ECBM within the workflow has the potential to substantially reduce operational costs and has been the driving factor behind its development. Certain MRO providers operate on contracts where clients agree on an hourly rate throughout the engine's time in the facility (known as "Power-by-Hour" or PBH). This provides clients with the exact maintenance costs required in advance. That being said, this business model also acts an incentive to find solutions for reducing costs as much as possible and increasing profit margins [54]. Prognostics can therefore provide accurate estimations of future engine condition and performance to not only decrease operational disturbances and costs but also improve maintenance activities.

5.3. Smart Workscoping

The long-term vision at KLM ES is to establish a process known in the MRO field as "Smart Workscoping". While the process is relatively broad and encompasses several aspects of MRO activities, in this context the principle involves performing maintenance-related tasks only on components that actually require maintenance, as opposed more exhaustive workscopes which include preventive measures that may or may not be deemed entirely necessary depending on an engine's current state. This approach essentially minimizes or completely eliminates the need for unnecessary tasks within a workscope while allocating more time and resources to tasks and components that require more care. That being said, Smart Workscoping would then require sufficient knowledge of the system's current state in order to identify the necessary maintenance tasks, which can be obtained via ECBM.

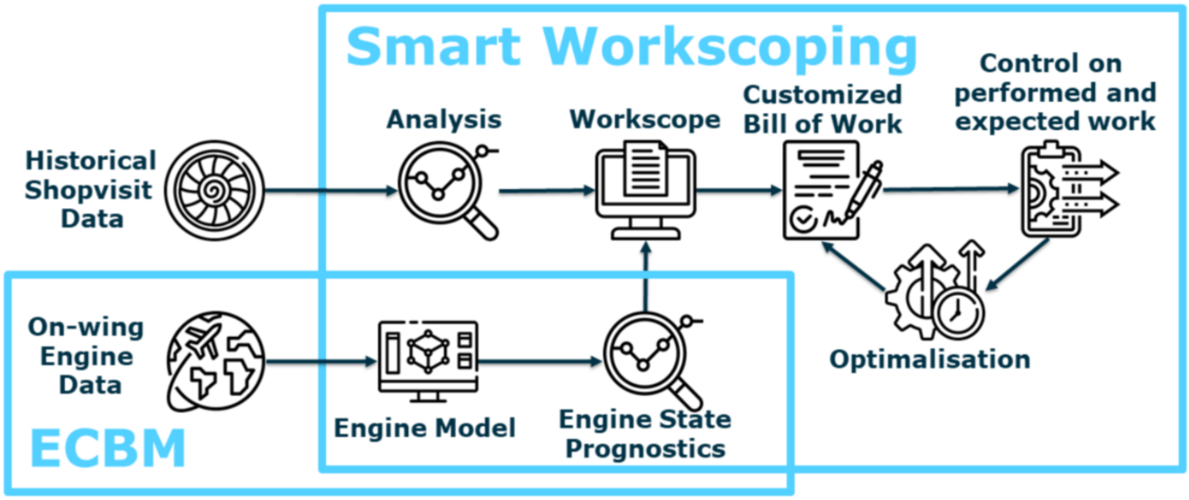


Figure 5.3: Visual representation of the concept of Smart Workscoping [5]

ECBM can provide measures of component degradation based on the data collected during flights. By making use of the ECBM system and historical data from previous shop visits, a workscope can be designed which provides the optimal maintenance with minimal work and cost. Furthermore, with access to such data, the ECBM system can also provide prognostics based on an engine's current state and account for future component failures within the workscope or define an entirely new workscope if a potential failure is imminent.

ECBM has in fact been implemented within KLM ES, albeit to a limited degree. KLM ES has at their disposal an on-wing prognostics system called PROGNOS, an expert system which uses engine data to perform trending of certain values measured. The rule-based system generates warnings if a trend indicates a potential failure or if deviations from historical data is also detected. While PROGNOS has proven its capability in detecting potential failures in the HPT for example, properly tuning the rules used by PROGNOS in order to balance false warnings and missed failures is a task revealing to be more complicated than expected according to the KLM E&M division. The use of PROGNOS so far has been limited to engines part of the KLM-Air France engine pool and to certain engines outside of it depending on the contract between ES and the client. Furthermore, it should also be noted that engine diagnostic models are not really used at KLM ES during servicing. However, if an engine does not pass its compliance tests post-visit during its test-cell run, then the engine performance parameters recorded at the test-cell serve as input parameters for the engine diagnostic models. This process allows KLM ES to determine which components require the most technical attention, however the facility still lacks the capacity to apply this strategy to every outgoing engine.

Below in Figure 5.4 is a conceptual road-map of the level of development in ECBM at KLM ES. The long-term goal of KLM ES's research is to further develop ECBM methods from "diagnostic" to "predictive" maintenance and eventually reach "prescriptive" maintenance. Prescriptive maintenance would involve integrating predictive methods and prescribe maintenance actions in advance based on the current state of a given system [55]. This would tie in directly to the overarching "Smart Workscoping" paradigm.

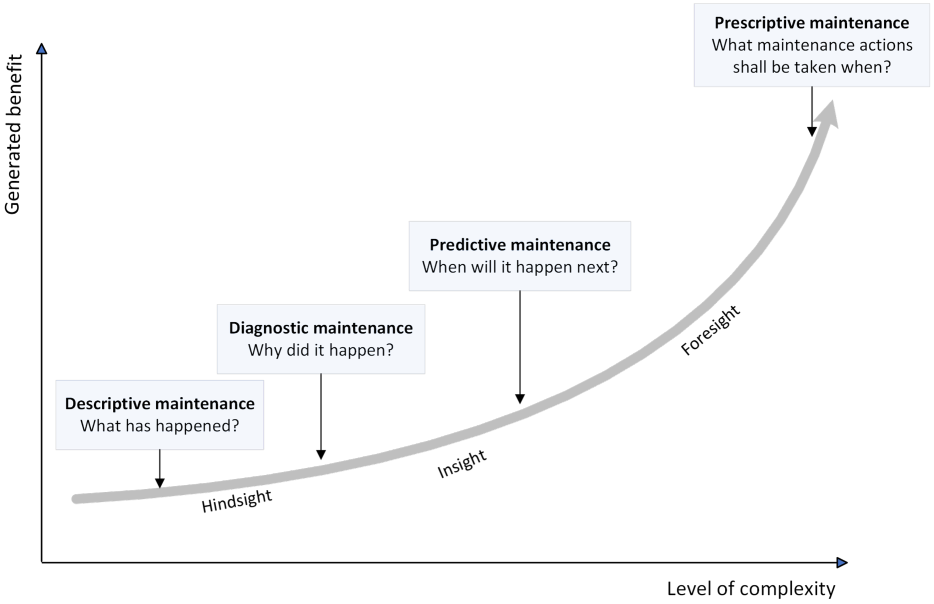


Figure 5.4: Conceptual progression of ECBM implementation at KLM ES

As discussed in Section 5.2, the MRO industry is investing heavily into the ECBM paradigm and predictive maintenance. This would allow MRO provides the ability to perform future projections of system health based on its current state and previous deterioration trend. According to Rootliep, GPA can serve as basis for development of predictive methods for component health using in-flight measurements of gas-path parameters [56].

5.4. Engine Modelling

Engineers at KLM Engine Services currently make use of GSP as a diagnostics tool based on principles from GPA in order to identify component deterioration in the case of unsatisfactory engine performance compliance following an engine's shop visit [56]. This approach allows engineers to locate faults on relevant components while reducing the need for a complete dismantling of the engine core, which can save significant time and resources. Development of GSP still continues to this day [6]. Research within KLM ES is still ongoing since the successful implementation of a GSP model of the CF6-50 for condition monitoring. Several academic projects have since then, dedicated to gas turbine modelling. An overview of said projects up until the time of writing is shown in Figure 5.5.

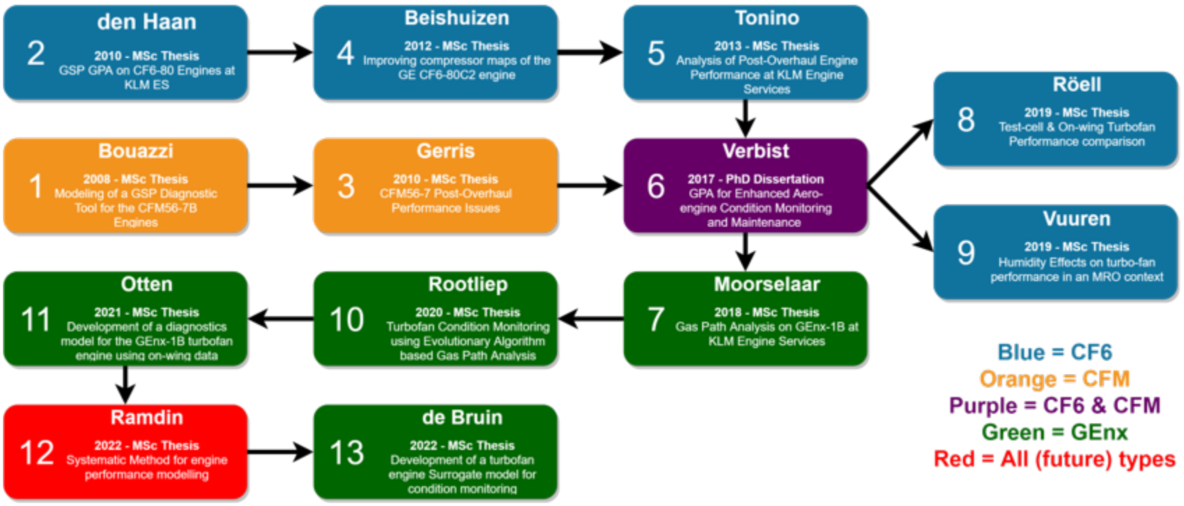


Figure 5.5: Overview of research conducted at KLM ES at the time or writing

As seen from Figure 5.5, most of the research done so far has focused on engine modelling via GSP based on the CF6 engine. Given that said model was found to be quite accurate, it served as a baseline for the later development of a model for the GEnx-1B turbofan designed by Moorselaar [26]. After that, work was continued by Rootliep [56], Otten [17] and Ramdin [57].

Otten's investigation was focused on mitigating the loss in accuracy of the GSP models due to what is referred to as secondary performance parameters. The result was an approach on how said secondary performance parameters could be properly quantified and accounted for in the GSP models. While Ramdin focused mainly on establishing a consistent framework for engine modelling applicable to any type, most of the research focused on methods utilizing the model and not development of the models themselves. Furthermore, research done so far has mainly remained in the field of diagnostics and monitoring purposes, such as Ramdin [57] and Rootliep [56]. However, the working principle behind GPA can also be utilized for prognostics according to Rootliep; in-flight recording of gas-path parameters can be used for real-time monitoring of engine component health, allowing for predictive methods to be implemented and prepare maintenance workscopes in advance [56].

Following Ramdin's research, de Bruin [5] focused on solving an issue that has continuously appeared since Moorselaar. Though KLM ES has access to GSP which utilizes nonlinear gas path analysis capable of performing engine diagnostics based on collected flight data, its accuracy for models of the GEnx-1B is significantly reduced. The reason for this is that the GEnx-1B contains fewer sensors along the gas path compared to its predecessors. The GEnx-1B does however, provide continuous engine operating data (or CEOD) which can help circumvent this issue. A breakdown of the available sources of data for modelling is provided in Chapter 7. However, due to the large amount of data and the inherent bias and noise found in sensor measurements, GPA methods, including GSP, are not able to properly make use of this data. Therefore, de Bruin proposed a surrogate model capable of utilizing CEOD based on HDMRs, as outlined in Section 4.3.4.

6

Extended Modelling of the GEnx-1B

Given the work conducted at KLM ES on engine modelling, the general objective of this thesis is to continue development and implement improvements to the established modelling framework surrounding the GEnx-1B. Section 6.1 opens with a more in depth introduction of the GEnx-1B as the engine of focus. Once the relevant subsystems and details are introduced, Section 6.3 discusses the results and insights provided during the development of the current surrogate model. Based on that, Section 6.5 will outline the methodology applied to implement the modifications as suggested from previous recommendations.

6.1. General Electric GEnx-1B

The General Electric GEnx-1B is a counter-rotating twin-shaft, high bypass turbofan engine first introduced in 2011. The GEnx-1B was designed for high-capacity long-haul aircraft, most notably the Boeing 787 Dreamliner, as a replacement for the widely used CF6 engine line. The GEnx-1B offers a reduction in fuel consumption of up to 15% compared to its predecessor, while offering reduced emissions during operation and requiring less intensive maintenance [58].

Figure 6.1: The General Electric GEnx-1B [58]

The GEnx-1B's performance is attributed to newer technologies also used in the GE90, such as curved composite fan blades leading to a reduction in engine weight, and blisk stages. "Blisk" stages are a combination of the rotors and blades into a single disk, contrary to the conventional blades which are manufactured as separate parts and attached to the rotating disk. Blisks offer increased efficiency by reducing the flow leakage typically occurring between the hub and rotor of compressor stages. Additionally, blisk stages offer a significant reduction in weight, possibly up to 30% compared to conventional compressor stages. With this increased compressor efficiency, the GEnx-1B can achieve higher pressure ratios with less compressor stages compared to previous generation turbofans [59]. Furthermore, the GEnx-1B is equipped with a Twin Annular Swirler combustor, allowing for sustained combustion at a lower temperature, resulting in significant decrease in NOx emissions [60].

The GEnx-1B is also equipped with several airflow control systems. These systems are present to regulate the airflow passing through the engine to ensure adequate operation across all flight conditions, increase operational margins or engine performance. The systems relevant for this thesis are presented in the following subsections.

6.1.1. Variable Stator Vanes

The GEnx-1B is equipped with Variable Stator Vanes (VSV) and Inlet Guide Vanes (IGV) as a method of airflow control to ensure stable operation during off-design conditions. The VSVs are placed at the first four stators in the HPC while the IGVs are located at the HPC inlet [4]. VSVs and IGVs vary their angles relative to the flow to induce higher tangential inflow flow at the rotors and reduce axial velocity. This is to avoid compressor stall, also known as surge. In low speed conditions such as taxi and take-off, the VSVs/IGVs are closed, resulting in a reduced mass flow and increased surge margin [3]. The VSV/IGV positions are generally controlled and scheduled based on the N2 speed, however they also operate on different schedules when accelerating or decelerating to ensure an adequate surge margin [3].

6.1.2. Bleed Valves

Variable Bleed Valves allow for air between the LPC and HPC to be discharged to the bypass flow. Discharging air from the compressor system results in a loss of energy extracted thus impacting engine performance, but in low speed conditions, it allows for an increased surge margin. The GEnx-1B is equipped with VBVs at the LPC outlet and are only used when necessary [61]. Transient Bleed Valves carry a similar purpose to the VBVs. Located aft of the HPC, core flow is discharged into the bypass flow when the TBVs open, which again leads to a decrease in engine performance due to lost energy. However, TBVs are generally used during start-up or when accelerating from idle in order to ensure sufficient surge margin [4].

6.1.3. Anti-Ice Bleeds

With regards to the GEnx-1B's anti-ice systems, two measurements are present within the on-wing data: the valve positions for the Booster Anti-Ice (or BAI) and Cowl Anti-Ice (or CAI) systems. The BAI and CAI system also consists of bleed valves, similar to the TBVs and VBVs, extract air flow from the core and route it to key systems to avoid the accumulation of ice on certain surfaces and reduce its impact on operational safety and performance. The BAI valve extracts air from the 7th stage of the HPC in order to heat the LPC, often referred to as booster, to avoid ice accumulation on blades and other surfaces. The same is the case for the CAI bleed valve, except it routes heated flow to the engine cowl instead. Both systems extract core flow, thus reducing the total energy extracted from the flow.

6.1.4. Core Compartment Cooling

The Core Compartment Cooling (CCC) system provides air from the fan outlet as cooling flow for the core compartment when high temperatures are experienced. Cooling flow is regulated using a dual position valve. This valve is generally fully open during low altitude conditions. As altitude increases, the valve closes to reduce air flow [4]

6.1.5. Turbine Active Clearance Control

The GEnx-1B is also equipped with an active clearance control system for both the High Pressure and Low Pressure turbines. A significant contributor to efficiency loss in turbo-machinery is tip leakage, which is further exacerbated with increased clearance between the turbine blade tips and the surrounding casing. The larger the clearance, the less energy is extracted from the flow by the turbines, which then has to be compensated by an increased fuel consumption. That being said, certain clearance must be allowed to avoid serious abrasion between the blades and the surrounding casing [61].

The highest temperatures occur in the combustor, followed right after by the turbine sections. Due to the extreme heat, the turbine blades and surrounding casing experience thermal expansion, which then alters the clearance. Therefore, the active clearance control systems in both the HPT and LPT feed cooling flow from the fan discharge and impinges on the turbine casing in order to cool down and reduce the clearance[4].

6.2. GSP Model of GEnx-1B

As shown in Section 5.4, work has been ongoing at KLM ES for years with regards to development of accurate engine models using GSP. Several models were developed for a variety of engines serviced by KLM ES and operated by KLM, such as the CFM56, CF6 and the GEnx-1B being the most recent.

The first GSP model of the GEnx-1B was initially developed by Moorselaar using test-cell and on-wing measurements [26], and was later used for further research in Gas Path Analysis by Rootliep [56] and Otten [17]. As mentioned in Section 3.3.2, GSP utilizes component maps for each turbomachine component to simulate both design and off-design engine performance. However, since the original component maps of the GEnx-1B are proprietary, initial tuning of the component maps was conducted by Moorselaar.

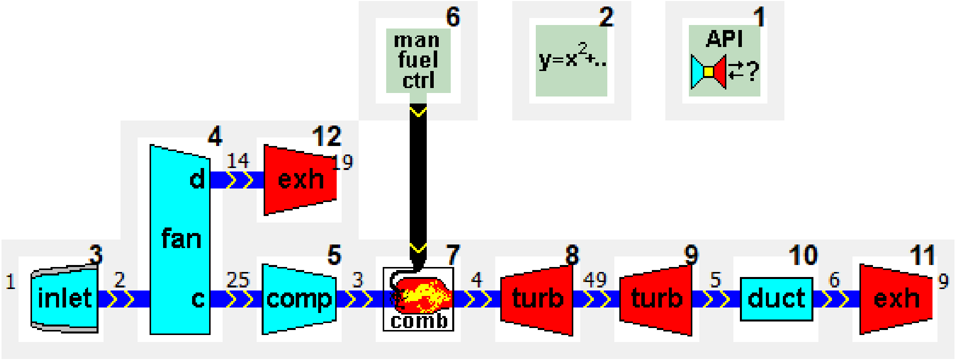


Figure 6.2: GSP model of GEnx-1B turbofan, developed by Ramdin [57]

Later, further development of the GSP model itself was later undertaken by Ramdin, with the goal of developing a systematic framework for engine performance modelling [57]. Ramdin encountered convergence issues when simulating certain operating conditions. The result of Ramdin's research was an improved GSP model of the GEnx-1B after implementing recommendations and adjustment proposed by the preceding work, such as combining the Fan and LPC components into one coupled component as shown in Figure 6.2. The resulting GSP model was later verified to be stable and sufficiently accurate in representing engine performance in take-off conditions.

6.2.1. Model Accuracy

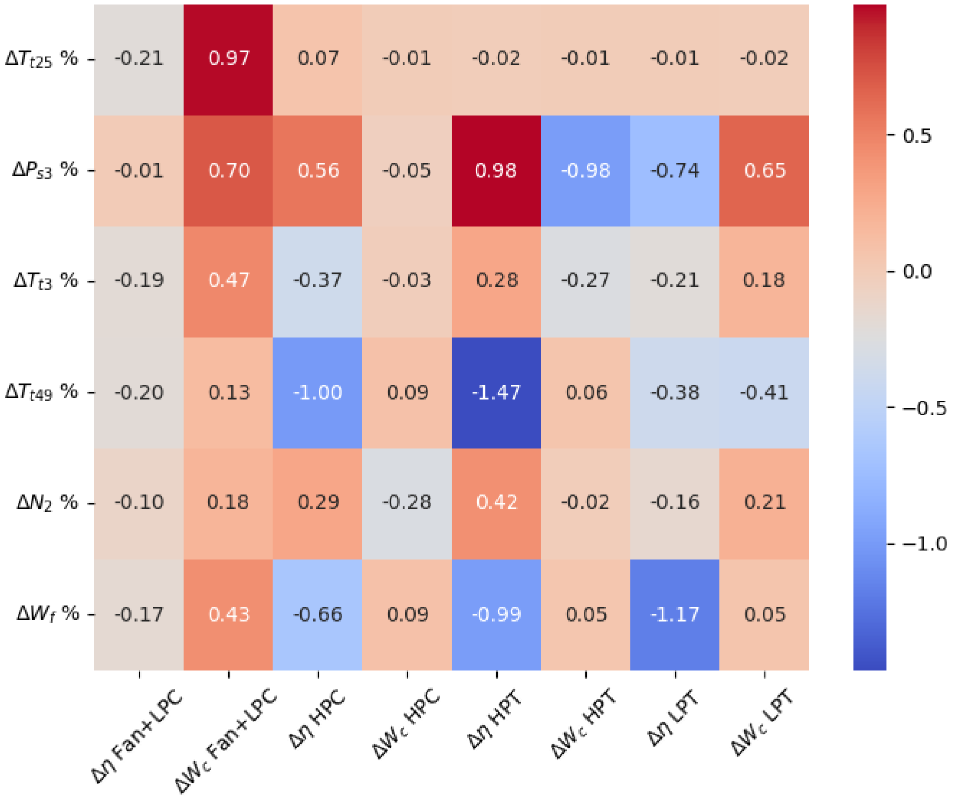
The relative deviation between the GSP model output and measurements taken from test-cell correlation reports were reported to fall below 1% for all relevant parameters for an operating range between 90 - 95 % N_1 speed when simulating near take-off conditions. This GSP model was then used by de Bruin [5] for the development of HDMRs at KLM ES. In the case of this thesis project, this will be the GSP model used going forwards. An analysis of the accuracy of this model at take-off was conducted as a baseline to assess later work. Based on the results of this analysis, shown below in Table 6.1, the current GSP model of the GEnx-1B is deemed sufficiently accurate for modelling engine performance during take-off. Note that the data shown below is based on simulations conducted on take-off data, while the deviation for cruise and climb were retrieved from results reported during the development of the currently available GSP model at KLM ES [57]

Table 6.1: Assessment of accuracy of baseline GSP model of GEnx-1B

Parameter	T_{t25}	P_{s3}	T_{t3}	T_{t49}	N_2	W_f
Avg Deviation @ Take-Off [%]	0.65	0.72	1.30	0.66	1.01	0.52
Avg Deviation @ Climb [%]	0.6	0.9	1.4	1.6	0.4	4.01
Avg Deviation @ Cruise [%]	3.3	0.7	3.8	7.2	3.1	7.99

6.2.2. Sensitivity Analysis of GSP Model

Given that the SPPs will directly affect the component performance on the measured parameters, it is important to first analyze said effects as obtained from the GSP model. To do that, a sensitivity analysis of the GSP model was conducted in order to gain insight on the influence of each component on the gas path parameters. This sensitivity analysis is presented in the form of a heatmap below in Figure 6.3, where the x-axis shows the component map modifiers and the y-axis the change in gas path parameters when the component map modifiers, $\Delta\eta$ and ΔW_c, are increased by 1%. This provides valuable insight about how the GSP model reacts to certain deterioration effects.



	$\Delta\eta$ Fan+LPC	ΔW_c Fan+LPC	$\Delta\eta$ HPC	ΔW_c HPC	$\Delta\eta$ HPT	ΔW_c HPT	$\Delta\eta$ LPT	ΔW_c LPT
ΔT_{t25} %	-0.21	0.97	0.07	-0.01	-0.02	-0.01	-0.01	-0.02
ΔP_{s3} %	-0.01	0.70	0.56	-0.05	0.98	-0.98	-0.74	0.65
ΔT_{t3} %	-0.19	0.47	-0.37	-0.03	0.28	-0.27	-0.21	0.18
ΔT_{t49} %	-0.20	0.13	-1.00	0.09	-1.47	0.06	-0.38	-0.41
ΔN_2 %	-0.10	0.18	0.29	-0.28	0.42	-0.02	-0.16	0.21
ΔW_f %	-0.17	0.43	-0.66	0.09	-0.99	0.05	-1.17	0.05

Figure 6.3: Sensitivity analysis of GSP model of GEnx-1B

For example, increasing the component efficiency of both the High and Low Pressure Turbines results in a decrease in fuel flow, which is in line with the expected physical behaviour. Additionally, a more efficient HPT leads to a larger reduction in EGT (or T_{t49}) as compared to the effect of the LPT on EGT. In contrast, an increase in flow capacity generally indicates a deterioration of performance for turbine sections, for example due to abrasion or erosion. An increase in flow capacity at the HPT results in a decrease in pressure at the HPC outlet, thus less energy to be extracted by the turbines following the combustor. The results from the sensitivity analysis also indicate that the HPC outlet pressure P_{s3}, the EGT T_{t49} and fuel flow W_f are the most sensitive to changes in component condition, specifically due to changes in condition of the HPC, HPT and LPT. This information provides valuable insight when interpreting results from the surrogate model in later chapters.

6.3. Current Surrogate Model

At the time of writing, the latest research was conducted by de Bruin, who devised a surrogate model for component-wise diagnostics of the GEnx-1B. Following a literature review, de Bruin concluded that HDMRs scored the best based on various characteristics. The motivation behind this project was for KLM ES to continue advancing towards the goal of condition-based maintenance using a computationally low-cost model and compatible with CEOD.

6.3.1. Development Results

The result of de Bruin's project was multiple HDMRs generated for component-wise condition diagnostics in terms of $\Delta\eta$ and ΔW_c and several health indicators, including Exhaust Gas Temperature (or EGT), Fuel Flow (labelled as W_f) and Thrust Specific Fuel Consumption (or TSFC). The motivation behind the development of HDMRs for health indicators was to provide a quantitative measure of engine condition in line with indicators conventionally used in the MRO industry. For example, a widely used indicator, as mentioned in Chapter 3, is the Exhaust Gas Temperature. If the average EGT is shown to increase over cycles, it is usually a clear sign of deteriorated engine condition. With that, the HDMR for ΔEGT was developed in order to quantify the level of deterioration of each component in terms of their respective effect on the change in engine EGT. The same approach is done for TSFC and fuel flow.

The resulting HDMRs developed by de Bruin [5] have been demonstrated to be an accurate modelling approach, as shown below in Figure 6.4 and in Table 6.2. The HDMRs generated were deemed to be sufficiently accurate; average error values were reported to be as low as 0.23%. That being said, several recommendations have been listed with regards to certain aspects of the HDMR development that show potential for further improvements.

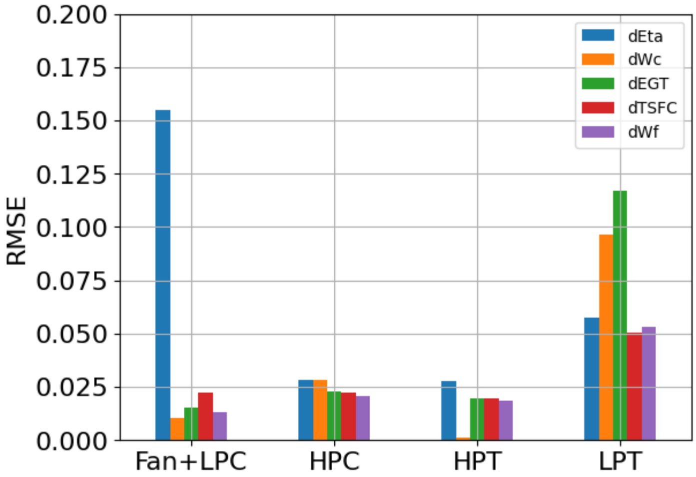


Figure 6.4: RMSEs for GEnx-1B of each HDMR type, redrawn from de Bruin [5]

Table 6.2: Table of RMSEs for GEnx-1B of each HDMR type

RMSE	$\Delta\eta$	ΔW_c	ΔEGT	$\Delta TSFC$	ΔW_f
Fan+LPC	0.15518	0.01054	0.01537	0.02252	0.01301
HPC	0.02852	0.02820	0.02305	0.02211	0.02053
HPT	0.02798	0.00119	0.01962	0.01944	0.01842
LPT	0.05749	0.09633	0.11726	0.05059	0.05343

6.3.2. Engine Condition Trending

Following the development of HDMRs, their capabilities are tested by observing their trending capabilities. On-wing data recorded over several flight cycles is fed to the HDMRs, of which the output is plotted over said time period. The results are time series plots used to visualize the progression of engine component deterioration over flight cycles. Below in Figure 6.5 and Figure 6.6 is an example of the outputs of the HDMRs, in this case the degradation of the Fan+LPC and HPC for a GEnx-1B labelled ESN 956XXX. The vertical axes represent the component deterioration in terms of $\Delta\eta$ and ΔW_c over time, shown in years on the horizontal axes. Maintenance events are also shown below, in the form of vertical lines representing water-washes and shaded areas representing shop visits taking place over a longer period. Within each plot is a red dashed line, representing a Gaussian rolling average of the resulting trending.

When inspecting Figure 6.6, the effects of shop visits and water-washes are sufficiently visible. In this case, the HDMRs do capture the restoring effect of water-washes for the HPC, even after shop visits. As for the Fan+LPC component, since water-washes are known to have little effect and shop visits often focus on the HPT and LPT sections, little to no change is expected. This is also adequately reflected by the HDMRs in Figure 6.5. Overall, the HDMRs showed good trending preceding certain reported failures when inspecting components such as the HPT, implying that the HDMRs could be used to identify said failure in advance.

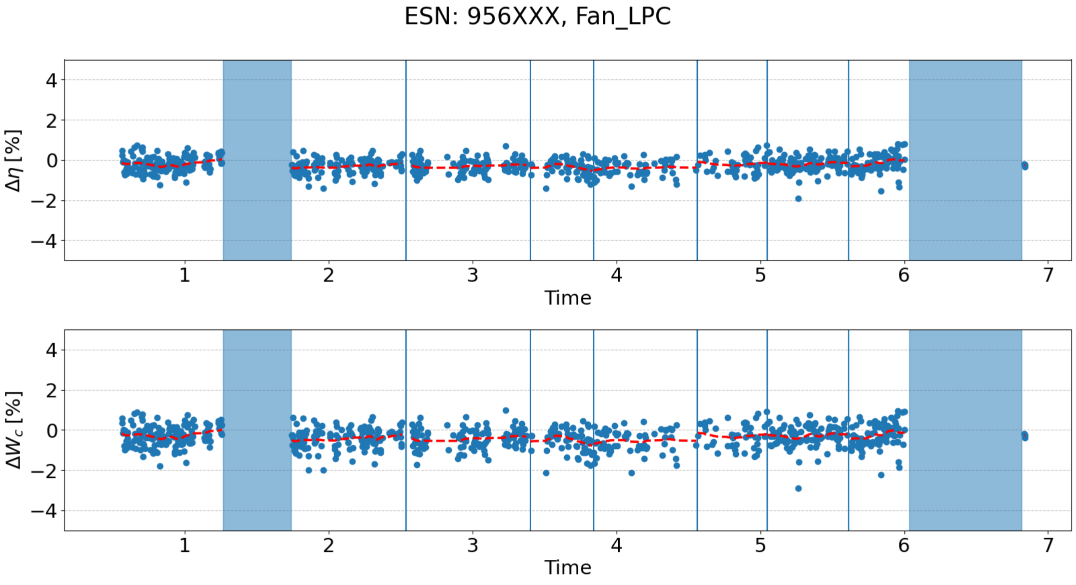


Figure 6.5: Example of HDMR results with on-wing data for GEnx-1B Fan+LPC, redrawn from de Bruin [5]

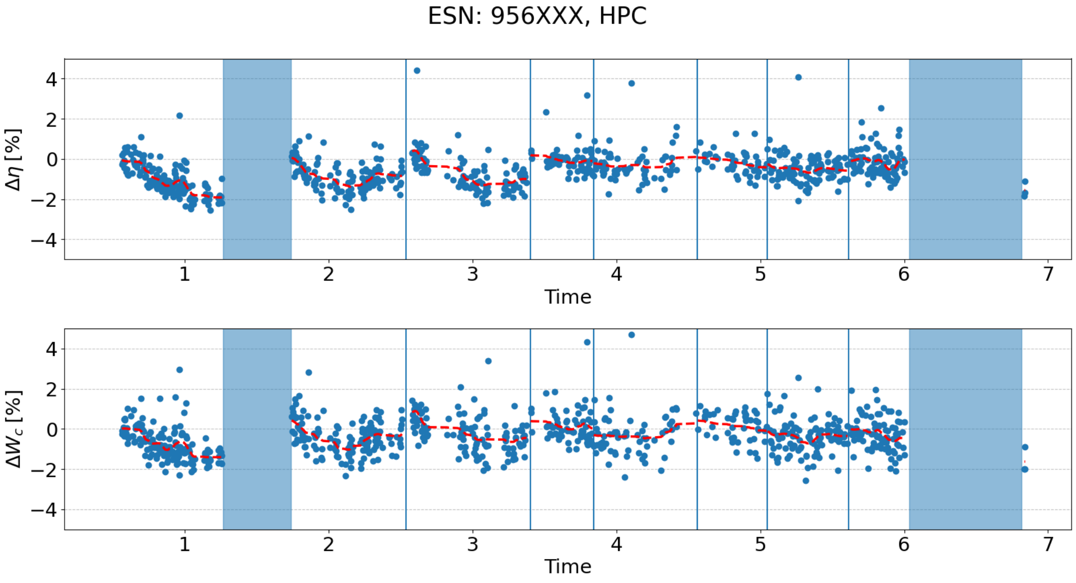


Figure 6.6: Example of HDMR results with on-wing data for GEnx-1B HPC, redrawn from de Bruin [5]

6.3.3. Secondary Performance Parameters

de Bruin recommended the implementation secondary performance parameters related to airflow control systems and active clearance control in the GEnx-1B. Secondary performance parameters, or SPPs, were not initially implemented as input parameters in the HDMRs. The justification for this omission was due to the initial modelling choice of maintaining a fixed corrected N1 speed. The effect of control systems such as the ACC and VSVs were not accounted for given that they vary little when the engine is operating at fixed spool speeds. That being said, implementing secondary parameters within the HDMRs could potentially yield higher accuracy since their effects would now be compensated for within the HDMRs. Given that the HDMRs are not treated as a black-box and that their inner-workings are known, implementing additional terms to the HDMRs should be sufficient in order to incorporate the effects of the SPPs on engine performance and component condition. Their effects on performance was outlined by Otten (2021) [17].

6.3.4. Constant N_{1c}

As mentioned above, the current HDMRs were developed with the assumption of a fixed corrected N1 speed. The justification for this decision was that it significantly reduced model complexity. However as a trade-off, it means that the resulting HDMRs are only accurate at that specific N_{1c}, thus greatly reducing the amount of points that can be used from CEOD. That being said, sufficient data is generally present even after the necessary filtering. If for any reason that is not the case, then it would be beneficial to implement a dependency on N_1 or N_{1c}., at the cost of increasing the complexity of the model.

6.3.5. High altitude corrections

In addition, given that the GSP model used for the GEnx-1B during the development of the HDMRs was constructed using correlation reports from test-cell runs, the GSP model is limited in accuracy the further operating conditions deviate from sea level. Therefore, the HDMRs can not be applied to data generated during cruise. This has been a recurring issue with engine models constructed in GSP.

At the time of writing, research is already ongoing internally in devising methods to implement high-altitude corrections to GSP models, focusing mostly on the relationship between certain parameters, nozzle coefficients and Reynolds number. Therefore, high altitude corrections will not be treated in this thesis.

6.3.6. Additional sensors

As mentioned in Section 5.4, one of the motivating factors for developing surrogate models for condition-based monitoring is to mitigate the limitation of GPA-based methods due to lack of information caused by a reduced number of sensors onboard newer generation turbofans. Notably, several static and total pressure sensors usually located in the bypass duct, the outlet of the LPC and HPT are now unavailable on the GEnx-1B, along with total temperature sensors at the engine exhaust [26]. This had a visible impact on the HDMRs' RMSE for the Fan+LPC component. Therefore, installing additional sensors could significantly benefit both the GSP model and HDMRs in terms of accuracy. For example, installing additional sensors at the LPC outlet could potentially decrease the reported RMSE for the Fan+LPC component and allow for both the Fan and LPC components to be uncoupled.

Unfortunately, additional sensors installed on all GEnx-1B operated by KLM would be a costly investment and difficult to justify. A solution would be to temporarily install the desired sensors during a test-cell run with a GEnx-1B after servicing. This would then provide updated test-cell correlation reports which could then further fine-tune the GSP model. However, without said measurements also included in the CEOD, this would mainly affect the accuracy of the GSP model.

6.4. Condition Prognostics

As previously discussed in Chapter 5, the high running costs of maintenance for an airline have been a driving factor in the development of diagnostics and prognostics method within the ECBM framework. Accurate prognostics could, for example, allow for more efficient maintenance schedules and reduce the need for component removals and/or disassembling, further reducing the risk of secondary damage in exposed systems. Additionally, MRO services would be able to anticipate and prepare for upcoming component replacements and repairs once certain components reach a certain deterioration level, such that engines can operate at a consistent efficiency and minimize fuel costs. According to Tchakoua [62], ECBM can lead to optimal maintenance strategies that result in lower running costs over a system's operational lifetime. This can be visualized below in Figure 6.7.

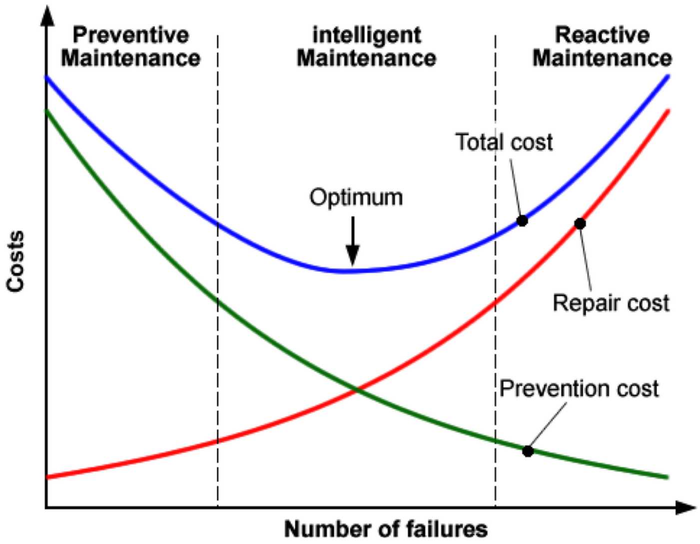


Figure 6.7: Comparison of different maintenance strategies [62]

The use of the HDMR method as a surrogate model for engine condition diagnostics provides several advantages, such as model transparency, the ability to implement expert knowledge and low computational cost. That being said, one of the downsides of this approach, as implemented by de Bruin, is that it can only act as a diagnostics method. Therefore, to perform condition prognostics, the approach must be reworked or modified to accommodate a secondary system or model to perform predictions on future component condition [5]. The HDMRs for condition diagnostics require engine data collected during flights as inputs to assess component condition, therefore predictions cannot be performed for flights of which data is unavailable of have yet to occur. Therefore, an additional model was be developed to work in tandem with the HDMR-based surrogate model.

Based on collected literature on ECBM and Predictive Health Management, several approaches were put forth as a prognostics solution for condition monitoring of engineering systems. According to Giorgi, several techniques have been used in the field of jet engine health prognostics, most notably Kalman Filters, Bayesian Belief Networks and Artificial Neural Networks [63]. The majority of publications focus on prognostics in terms of Remaining Useful Life. However, the exact definition of RUL can vary depending on the data available and the approach used. Given that a surrogate model is available, which can directly translate engine measurements into component condition in terms of map modifiers allows for a more physical interpretation of health condition and predictions. Given the nature of the outputs of the surrogate model, component condition data can be generated in order to train a separate model to handle component condition predictions. According to Giorgi [63] and Ntantis [12], artificial neural networks can handle non-linear effects, measurement noise and bias and avoid the smearing effect [12]. Furthermore, since HDMR outputs can be provided in the form of time-series, an ideal candidate for a prognostics model would be an neural network, specifically a Long Short Term Memory network. LSTM models have the same advantages as conventional ANNs, but have also been reported to show good performance when it comes to identifying long-term dependencies and complex patterns within data [64] [65]. Furthermore, LSTM networks have also demonstrated high performance for time-series predictions at a relatively low computational cost [66].

Based on the collected literature and the points made in the table above, it was decided that the prognostics model to be used for this thesis would be a LSTM-based machine learning model. Given that engine deterioration is known to exhibit a general downward trend over flight cycles, any prognostics model will have to be capable of capturing the temporal relationship between the relevant parameters [67].

6.5. Overview of Extended Modelling

To summarize, the objectives of this thesis are to implement the recommendations outlined in the preceding sections as improvements to the surrogate modelling framework used by KLM ES based on the HDMR approach. The main idea is to combine the results of the existing bodies of work completed by previous thesis research done at KLM ES. An visual overview of the framework employed in this thesis is provided in Figure 6.8 and is briefly explained below.

The GSP model for the GEnx-1B developed by Ramdin will act the baseline physical model, which is used to derive relationships for the effect on engine performance due to secondary performance parameters, using the GA-based optimization scheme proposed by Otten. The result of this exercise will be a modified version of the GEnx-1B model in GSP which can account for the effect on performance due to the secondary performance parameters based on additional inputs passed via the API. This phase and its results are further outlined in Chapter 8

The updated GSP model will then be used to generate simulated data of the GEnx-1B, for a variety of deteriorated conditions. The simulated data is then used as training data in order to re-train the HDMRs previously developed by de Bruin, but in this case accounting for the various secondary performance parameters. The SPPs will be implemented as additional input parameters to the basis functions that constitute the HDMRs. The results from these implementations is discussed in Chapter 9.

Furthermore, by using engine measurements as input data to the HDMRs, the component condition can be estimated per flight. An experiment will be conducted where the outputs of the HDMRs will be used as training data to develop a prognostics model. This model will consist of a neural network with LSTM cells, suited for time-series forecasting. In theory, this prognostics model will be capable of capturing the non-linear behaviour seen in turbofan component degradation and perform predictions on the future condition of said components. The results of this experiment are provided in Chapter 10.

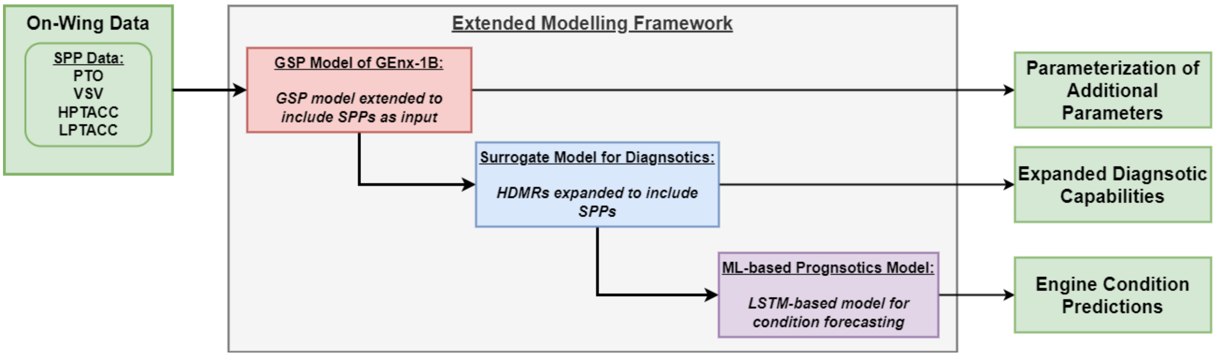


Figure 6.8: Flowchart of modelling scheme

7

Data Sources

Modern turbofan engines, such as the GEnx-1B, now come equipped with data acquisition systems on-board of aircraft. This new stream of data can be used for several applications, especially for engine modelling related tasks. This chapter outline the various sources of data available for this research. Furthermore, certain distinctions must be made, such as the difference between what is considered primary and secondary performance parameters. An overview of the parameters available is provided at the end of the chapter.

7.1. On-wing Data

During flight the FADEC performs measurements of engine operation throughout the flight. Depending on which generation of turbofan engine however, this data may be compiled in one of two ways, notably what is referred to as snaphsot data or CEOD.

7.1.1. Snapshot Data

During a flight, the FADEC stores what is referred to as a "snapshot": all primary, secondary and calculated parameters are recorded at key moments such as take-off, climb and cruise when specific conditions are reached. The advantage of snapshot data is that it is a common data acquisition approach in engines utilizing FADECs, thus three snapshots per engine per flight across both KLM and Air France's fleets results in a very large dataset at KLM ES's disposal [5]. A significant drawback of snapshot data however, is that it is highly sensitive to brief changes in the engine operating conditions. To mitigate this, snapshots are composed of data recorded over longer periods of time and post-processed. That being said, in the case of take-off and climb, it may be the case that a snapshot is taken when the engine is not operating in a steady-state condition where thermal equilibrium would be reached. That being said, it is reasonable to assume that snapshots taken during cruise are operating at, if not sufficiently close enough, to steady-state.

7.1.2. Continuous Engine Operating Data

The more modern turbofans such as the GEnx-1B not only can store snapshot data but can also gather Continuous Engine Operating Data, covering the entire flight with very small time steps thus including all operating conditions. In total, CEOD is composed of 640 different parameters measured by the FADEC. The downside however of CEOD is that it requires significant effort in post-processing due to the large amount of data taken per flight. Since no pre-processing is done, sensor noise will have to be filtered out. In addition, the refresh rate of the various sensors may not be the same, therefore certain slices of the data may be empty for a given time step. Furthermore, depending on the task, a sizeable chunk of the 640 parameters recorded might have to be discarded, such as activation bits, or FADEC and warning messages. That being said, CEOD acts as a very useful data source thanks to the amount of operating points recorded. Though extensive pre-processing will be needed, various methods were devised at KLM ES, notably by Otten [17] and Rootliep[56] in order to process the CEOD for condition monitoring and other related projects.

7.2. Performance Parameters

Engine performance parameters are parameters that provide a measure of a given engine's capabilities. In the context of turbofan engines, performance parameters are generally classified under the following categories: primary, secondary and calculated parameters [5].

7.2.1. Primary Performance Parameters

Most modern aircraft are equipped with engine control systems known as the Full Authority Digital Engine Control (or FADEC), which record data such as gas path parameters continuously throughout a flight via its Engine Monitoring Unit (or EMU). Direct performance measurements are classified as primary performance parameters, such as temperatures, pressures, mass flows, etc. For example, the General Electric GEnx-1B turbofan is equipped with sensors at various stations along the gas path, as shown below in Figure 7.1. Note that the sensors labelled in red are those longer provided compared to previous generation turbofans.

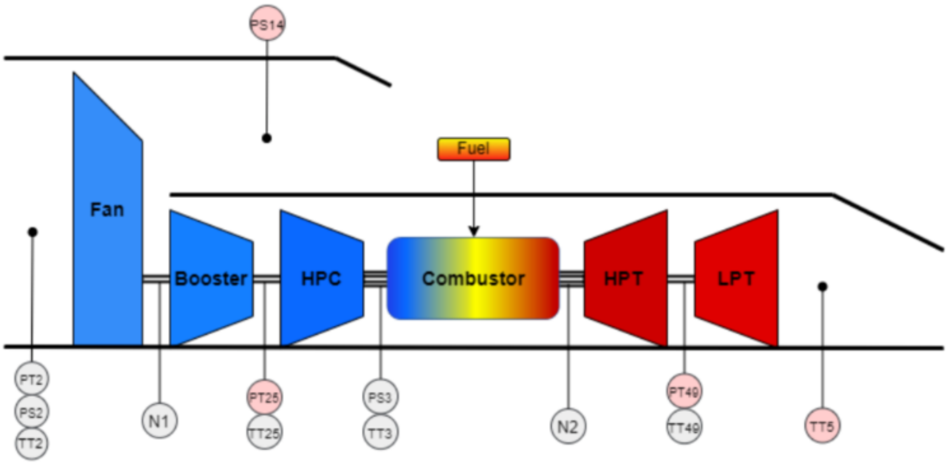


Figure 7.1: Gas path sensors and their locations on the GEnx-1B [57]

Temperature

Most temperature measurements in the GEnx-1B are done via either thermocouples or thermistors. Thermocouples are sensors consisting of two metals joined together that, when either cooled or heated, generate a voltage which can be translated to a temperature, whereas thermistors uses its electrical resistance which is dependent on temperature. Below is a table compiling the location, type and measurements accuracy of the temperature sensors onboard the GEnx-1B

Table 7.1: List of temperature sensors and relevant aspects

Gas Path Parameter	Sensor Type	Location	Accuracy
$T_{t,2}$	Thermistor	Nacelle inlet cowl	$\pm 1.1°C/\pm 0.4\%$
$T_{t,25}$	Thermistor	Fan hub frame	$\pm 1.1°C/\pm 0.4\%$
$T_{t,3}$	Thermocouple	Combustor diffuser nozzle	$\pm 1.1°C/\pm 0.4\%$
$T_{t,49}$ (EGT)	Thermocouple	Around HPT exhaust casing	$\pm 1.1°C/\pm 0.4\%$

Pressure

The GEnx is also equipped with sensors to measure the various pressures along the gas path. Most of the pressure sensors on-board consist of static ports in various locations with pneumatic line connecting them directly to the FADEC where the temperature-compensated sensor determines the pressure. The list of pressure sensors found onboard is shown in Table 7.2

Table 7.2: List of pressure sensors and relevant aspects

Gas Path Parameter	Sensor Type	Location	Accuracy
$p_{t,2}$	Total pressure sensor	Fan cowling	± 0.01 psi
$p_{s,3}$	Static pressure sensor	HPC outlet before CC	± 0.36 psi

Shaft Speed

Most modern turbofans have a twin-shaft (or two-spool) architecture; a low speed shaft powered by the LPT driving the LPC and fan, and a high speed shaft powered by the HPT which drives the HPC. The low and high speed shafts are often referred to as N1 and N2 respectively. The rotational speed of the shafts are measured by magnetic induction sensors with three coils present for redundancy.

Table 7.3: List of rotational shaft sensors and relevant aspects

Gas Path Parameter	Sensor Type	Location	Accuracy
N_1	Magnetic induction	Fan hub frame	± 0.12 RPM
N_2	Magnetic induction	Front of gearbox	± 0.12 RPM

Fuel Flow

Fuel flow is measured directly by a flow sensor relying on conservation of angular momentum. As the fuel is injected, it impinges on a turbine blade causing it to rotate. The rotational speed of this turbine is proportional to the fuel mass flow injected and is measured using magnetic induction coils as well. The accuracy of this flow meter is the largest compared to other sensors; with an accuracy of $\pm 3.5\%$. This increased level of inaccuracy must be accounted for in later modelling tasks.

7.2.2. Secondary Performance Parameters

Alongside the primary performance parameters discussed previously, the FADEC also records several other parameters involved in operating the engine, such as airflow control systems. These airflow control mechanisms use hydraulic fuel lines and valves to control actuators, thus the FADEC records their valve and actuator positions in percentages, from closed to fully opened. Within the CEOD recorded in flight, there are valve positions for the variable bleed valves (or VBVs), transient bleed valves (or TBVs), the HPT and LPT active clearance control valves (or LPTACC and HPTACC) and the variable stator vanes' (or VSVs) position setting. Additionally, the total power extracted or "power off-take" (PTO) from the N2 shaft is also recorded in the CEOD via a magnetic coil generator, recorded in terms of horespower [4].

7.2.3. Calculated Performance Parameters

Following primary and secondary parameters, other parameters are computed based on certain measurements recorded by the FADEC instead of being measured directly. An example of such parameter would be the HPT blade clearance, used later on for the HPT active clearance control system. Blade tip clearance is a significant loss factor and thus given an indication of the operating efficiency of the HPT. It should be noted however that the exact relations or correlations used to compute such parameters are not known and are proprietary information held by the OEM, in this case General Electric [5].

7.3. Test Cell

As mentioned before, KLM ES also has access to an engine test cell at Schiphol Oost. The test cell is capable of conducting compliance test runs for a variety of engines such as the CFM56-7B, CF6-80E1 and CF6-80C2. The GEnx-1B however cannot be tested at this facility due to its size and the test cell's flow constraints. Instead, GEnx-1B serviced by KLM ES are transported to the Zephyr test cell at Charles De Gaule airport in Paris, operating by Air France Industries. Engine performance is checked for compliance against multiple metric during a test cell run, such as vibrational loads, oil and fuel consumption, resulting thrust, shaft accelerations and so forth. Test cell runs are mainly used to verify whether an engine is compliant with the operational requirements after maintenance but also allows for data collection for other purposes.

Figure 7.2: Engine equipped with bell-mouth inlet in test cell [68]

During test-cell runs, engines are fitted with a bell-mouth inlet equipped with static pressure measurements located at station 12. Engines are attached to a load cell, made to measure the output thrust force. Along with the sensors found on board of the engine, the test cell collects additional parameters such as relative humidity, $p_{t,2}, p_{s,12}, T_{t,2}$ and thrust. The total pressure at ambient conditions is also measured from a sensor near the load cell, which is used to construct a correlation between that measurement and a pressure probe at station 2. Once the correlation is constructed, the station 2 probe is removed and the correlation is sued to measure $p_{t,2}$ based on the sensor from the load cell. During a test-cell run, the engine involved will follow a specific operating profile depending on which compliance tests it must pass. For example, in acceptance tests, KLM operates the engine at multiple regimes typically seen during normal operations, such as take-off, max continuous, idle in flight and on ground. Snapshot data is taken at each of these points as close to their nominal spool speeds.

The test cell has therefore the advantage of being able to provide measured data, collected within a controlled environment and operated at key operating points. The disadvantage however is that the number of operating points conducted during a test is fairly limited. Data gathered from correlation reports can be used to circumvent issue, although correlation runs are only conducted at an engine's first run or after the test cell has undergone modifications, thus the issue of limited data is still present.

7.4. GSP

In addition to the large dataset available from in-flight measurements acquired from the FADEC, the Gas turbine Simulation Program or GSP, introduced in Section 3.3 can also be exploited as a data source. The use of a simulation program has the advantage of generating physically consistent performance data of potentially any engine configuration at any operating condition. Measurements of any values at any engine station can therefore be acquired without the need of sensors.

The downside is that no model created in GSP is guaranteed to be fully accurate. As seen in Section 5.4, extensive research at KLM ES was conducted with the goal of developing accurate turbofan models using GSP. Models for the CF6 series were continuously researched and improved since den Haan in 2010 [69] and the same for models of the GEnx since Moorselaar [26]. That being said, the GSP model currently available still does not return sufficiently accurate results when simulating cruise conditions, thus limiting simulations to take-off conditions, and has yet to implement the secondary performance parameters relevant to the objectives of this thesis project.

That being said, simulated data generated by GSP was deemed the most appropriate source for training the HDMR-based surrogate model. The advantage of using GSP is that it allows for an extensive and structured sampling of the design space without any noise or bias. During development of the HDMR-based surrogate model at KLM ES, a Python-based application or "package" was created named "ParallelGSP", which employs virtual machines to run multiple instances of GSP for multiple simulations in parallel via the GSP API. More on ParallelGSP is discussed in Chapter 9

8

Secondary Performance Parameters

As outlined in Chapter 6, several recommendations were made with regards to improvements to the current surrogate modelling approach developed at KLM ES. One of the most prominent recommendations was the implementation of the secondary performance parameters. Section 8.2 opens with a more in-depth description of the relevant subsystems within the GEnx-1B and their effect on engine performance. Next, Section 8.3 outlines how said effects are quantified and implemented within the modelling framework. The results of this implementation are provided in Section 8.4.

8.1. Reference Engine & Flights

For this exercise, a set of on-wing measurements were retrieved from Snapshot data captured by a GEnx-1B engine, referred from now on as ESN956XXX. This engine was selected and deemed to be sufficiently healthy based on the recorded EGT Hot Day Margin (or EGTHDM). EGTHDM is considered to be a good indicator of the engine's general health state; the further it is from zero, the better. The selected flights for the optimization are the earliest segment of flights with the highest EGTHDM peak, as shown in Figure 8.1.

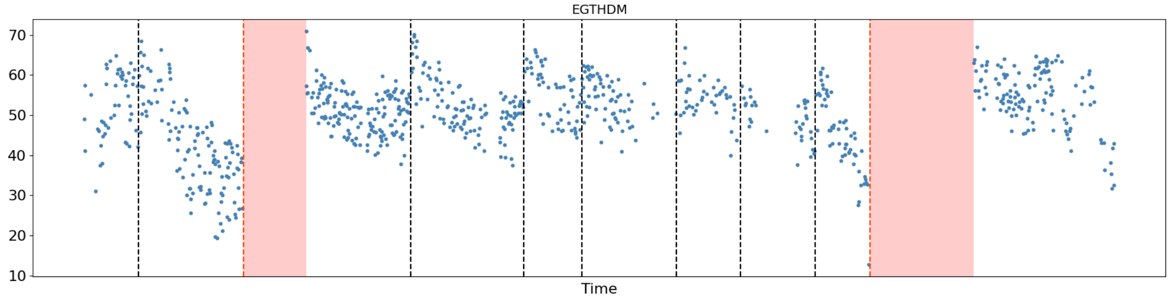


Figure 8.1: Time Series of EGT Hot Day Margin for ESN956XXX

Unfortunately, in the case of the GEnx-1B, KLM Engine Services only has access to CEOD after a certain number of cycles (number cannot be disclosed due to confidentiality). Therefore, on-wing measurements dating back to the first flights after entry into service are not available. This implies that the earliest datasets available are measurements from engines already with some level of component deterioration. While component deterioration is present at this stage of its lifetime, it is assumed to be a sufficient representation of a healthy engine.

That being said, for the sake of this thesis, Snapshot data of ESN956XXX will be used for two primary reasons. The first was stated above; CEOD is not available near the time of entry to service, however Snapshot data is available. Thus it can provide the necessary on-wing measurements early enough into the engine's operational life to be deemed healthy. The second reason is that, during the experimentation phase of this thesis, a substantial part of the CEOD database was not accessible. Certain datasets were still available for specific flights, however large gaps in the order of years were still present, thus a full uninterrupted dataset for ESN956XXX could not be retrieved. It was therefore decided that Snapshot data of ESN956XXX would be used moving forward.

8.2. Breakdown of SPPs

As mentioned briefly in Chapter 6, the GEnx-1B is equipped with several active systems to adapt engine performance depending on operating condition. Those discussed in this thesis are the bleed flow, variable geometry, active clearance control and power extraction systems.

8.2.1. Bleed Flows

While bleeding air from the core flow to other locations results in a decrease in total energy extracted by the engine core, it is often necessary in order to maintain a stable surge margin for safe operation of the compressor system. As listed in Chapter 6, two bleed control systems are found in the GEnx-1B: the variable and transient bleed valves (VBV & TBV). The VBVs are generally located between the LPC and HPC, in order to discharge core flow and prevent LPC from stalling. Meanwhile the TBVs are generally located at the outlet of the HPC with the same purpose; to reduce the load on the compressor and prevent stall usually during engine start-up and while idling [3].

Based on the distribution of the bleed valve positions shown in the snapshot data, it is clear that they are locked at fixed settings during the take-off phase and are rarely used. This is expected given that the bleeds are meant for stable operation of the engines during startup. A similar conclusion can be made for the Core Compartment Cooling valve.Similar to the TBVs, VBVs and CCC valves, the distribution of the anti-ice valves is also fairly limited. The CAI setting is recorded in values of either 0,1 or 2. No additional information on the meaning of these settings is available at KLM, therefore the exact relation between this system's settings and its effects on the engine performance remains ambiguous. As for the Booster Anti-Ice system, two measurements are present within the on-wing data. The first is a binary setting that indicates the operating mode of the system; it is either enabled or disabled. Additionally, the second is a pressure value measured in the flow duct leading to the booster. This pressure measurement is then used by the EEC system as an input to estimate the extracted bleed flow. Based on experience at KLM ES and after inspection of on-wing data, said pressure experiences spikes only during specific moments during flight and rarely at take-off. This would imply that the BAI valve is generally closed during the take-off phase [26].

Given that the bleed flow valve positions vary very little during the take-off phase and are primarily used during early phases such as start-up, taxiing and are rarely used during cruise, it is safe to consider that their effect on engine performance is minimal compared to others such as the Active Clearance Control systems and VSVs. Similar conclusions can also be made for the anti-ice bleed flow valves. Therefore, their measurements are filtered out from the on-wing data and will not be considered in further stages of this study.

8.2.2. Turbine Active Clearance Control

In order to rotate freely around the shaft axis, there is a certain amount of spacing or clearance between the blades and the surrounding shroud or casing. During operation, as the shaft is spinning, the turbine blades are under large induced loads, such as centrifugal, gyroscopic, aerodynamics and thermal loads [70]. In order to both maintain adequate efficiency and prevent damage due to abrasion, the turbine sections are equipped with thermal control systems to maintain the tip clearance within acceptable levels. Below in Figure 8.2 is a representation of the turbine rotor and casing radius over a complete flight for both the case of with and without the ACC system engaged.

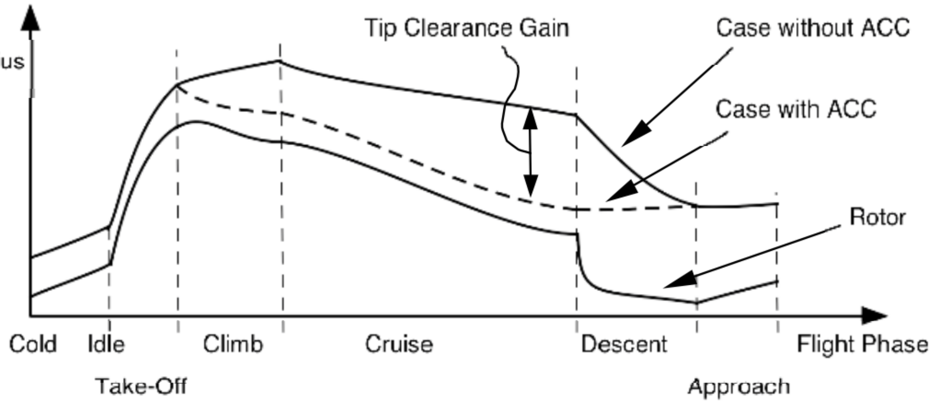


Figure 8.2: Graphical representation of effect of turbine active clearance control [61]

From Figure 8.2, it can be seen that the effect of the ACC is significant during cruise operation, where the largest clearances are experienced. The rotor radius also experiences a significant increase due to the centrifugal loads the rotor and blades experience during acceleration from idle to take-off and climb [61]. In certain flight phases such as take-off, the margin between the blade tips and casing can be quite tight. Therefore, it is clear to see why active clearance control is required in order to maintain the longevity of turbine components. However, clearance also plays a role in the efficiency of the system. The effect of tip clearance on engine efficiency is mostly related with entropy production (see Section B.1). With increasing tip clearance, more of the flow at the higher pressure side can spill over to the lower pressure or suction side of the blade, as shown in Figure 8.3, causing an increase in entropy through viscous friction as the flow mixes and creates a vortex flowing downstream [71]. Losses associated with tip clearance are reported to account for up to 30% of the total efficiency loss experienced by turbine stages [72].

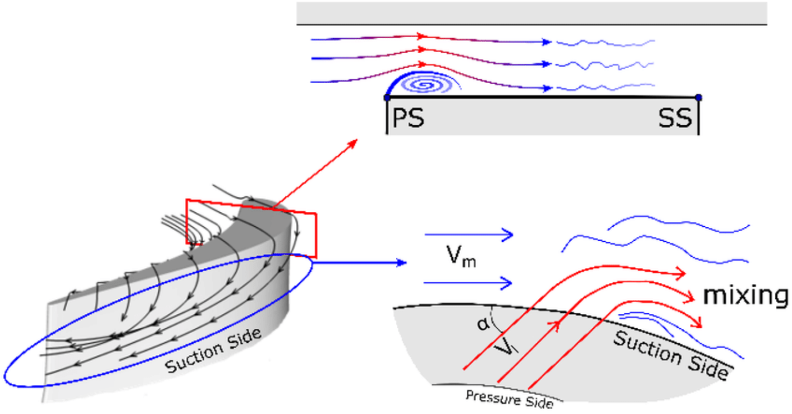


Figure 8.3: Visual example of blade tip leakage [73]

Based on the on-wing data collected from various flights during the take-off phase, a brief analysis was conducted on the variation of the measured actuator positions of the ACC system for both turbines. The distribution is provided below in Figure 8.4 for the HPT and Figure 8.5 for the LPT.

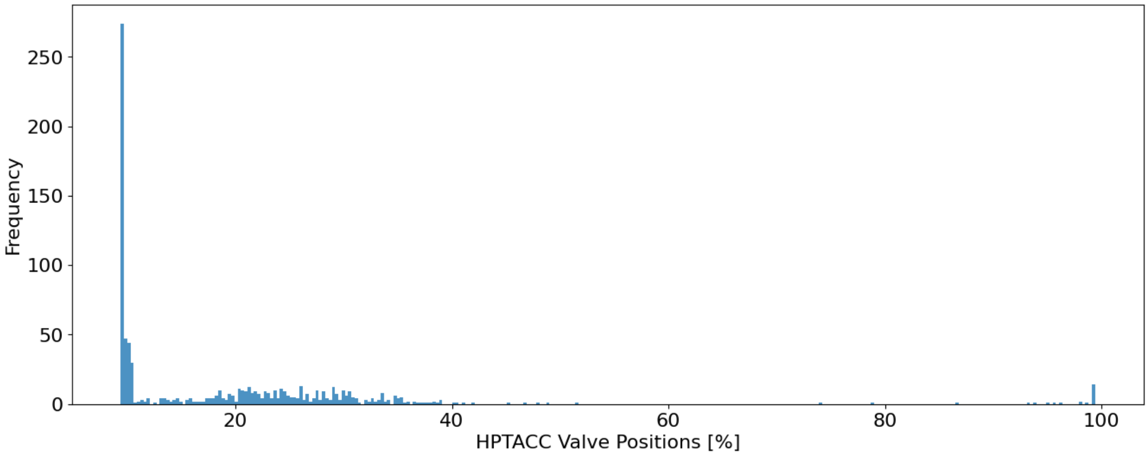


Figure 8.4: Distribution of HPTACC valve position from take-off data

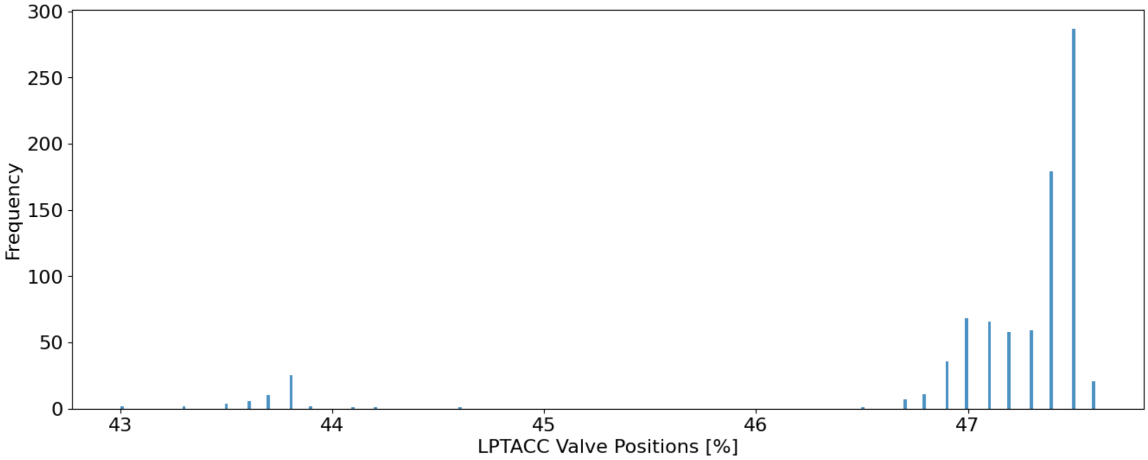


Figure 8.5: Distribution of LPTACC valve position from take-off data

As seen in Figure 8.4, the ACC valve position for the HPT varies significantly during the take-off phase. However, the most frequent setting is found at 9.5%, which implies that the HPTACC valve is more often operated at a low setting and rarely actuated. According to its system description, this particular setting still allows a continuous but limited cooling flow to pass through [4], meaning that the ACC system is generally acting passively at take-off. That being said, a substantial percentage of flights do employ higher valve settings, thus actively making use of the ACC system. As for the LPT ACC system, the distribution is rather small. The valve positions vary within a much smaller range compared to the HPT; between 43% and 48% during take-off. This implies that the LPT ACC generally operates at a steady regime during take-off. This makes sense given that the highest temperatures are experienced in the HPT, thus the LPT does not experience as large a variation in thermal expansion as compared to the HPT.

Furthermore, it should be noted that another parameter related to the HPT Active Clearance Control is also available within the on-wing data. The GEnx-1B also includes a "Calculated HPT Clearance" measured in inches. However, unlike the HPTACC valve position, this is not a direct measurement; based on other measurements, this clearance value in calculated using some relations programmed in the GEnx's EEC system. Unfortunately, General Electric considers this information to be proprietary, thus it is not known within KLM ES. Using this parameter as an input to the parameterization for the performance effect of the HPT's Active Clearance Control would introduce a significant amount of modelling inaccuracy, thus it was decided to continue working with the measured valve positions instead.

8.2.3. Variable Geometry

In the category of variable geometry, two systems are present on the GEnx-1B, notably the Variable Stator Vanes (VSVs) and Inlet Guide Vanes (IGVs). During off-design operation, both are present to prevent blade stall by changing their angle relative to the flow. When the VSVs are closing; the angle between the leading edge and the axial flow increases, inducing a higher flow velocity tangential to their upper surface and a reduced axial flow velocity.This effectively slows down the core flow. This effect is illustrated in the resulting velocity triangles shown in Figure 8.6

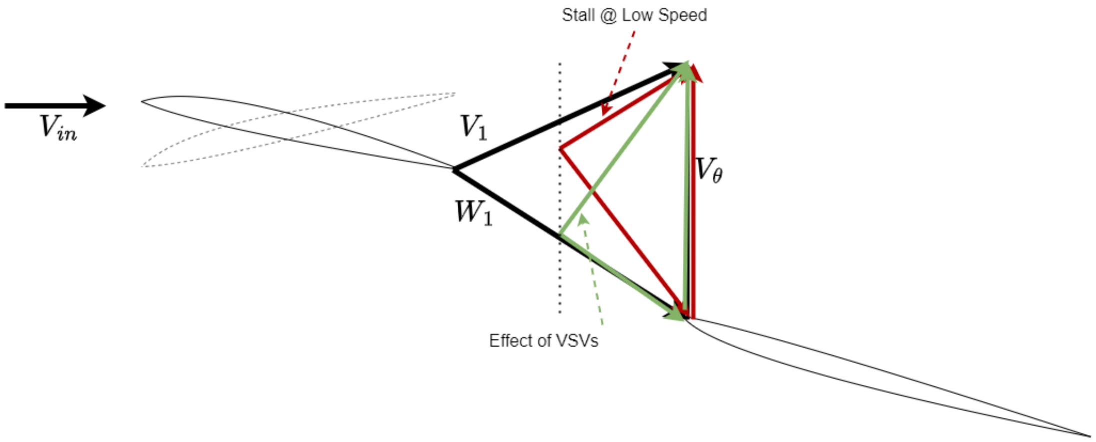


Figure 8.6: Velocity triangle showing effect of VSVs [74]

This mainly results in a decrease in corrected mass flow along the same corrected speed line on the respective component map [3]. The effect of the VSV/IGV position with respect to a compressor component map are shown below in Figure 8.7, where it can be seen that the operating line remains unaffected. The position of the VSVs and IGVs are schedule based on the engine corrected fan spool speed. Their schedules also varies with respect to altitude. In terms of impact on engine performance, the effect of the VSVs in this case are quantified in terms of $\Delta\eta$ and ΔW_c of the HPC.

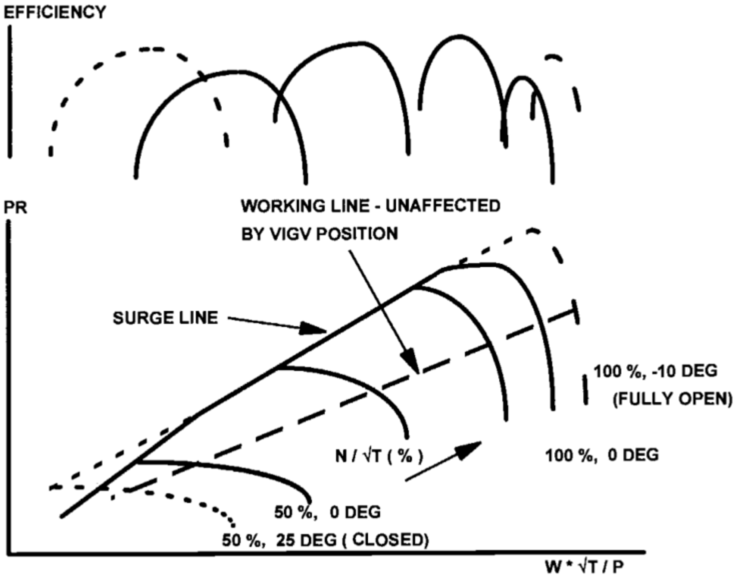


Figure 8.7: Example of compressor map showing effect of VSV/VIGV position, redrawn from [3]

For a given power setting, as the VSVs close, the corrected flow is reduced. This leads to a reduction in fuel consumption and EGT downstream, while increasing the resulting pressure P_{s3} at the outlet of the HPC. Additionally, in order to maintain a certain power setting, the N2 shaft speed will increase in order to continue driving the flow through the core. As for the efficiency, the VSVs would result in an increase in order to counteract the effect of a reduced corrected flow, since the operating point would move to a point of lower efficiency. A positive effect on the HPC efficiency would lead to an increase in resulting pressure at the outlet, P_{s3}, along with a decrease in resulting total temperature T_{t3} given that the compression process results in less entropy generated. Downstream, this effect would propagate to a reduced fuel consumption (W_f) and lower EGT (T_{t49}). This behaviour is also corroborated by the sensitivity analysis of the GSP model previously demonstrated in Section 6.2.2.

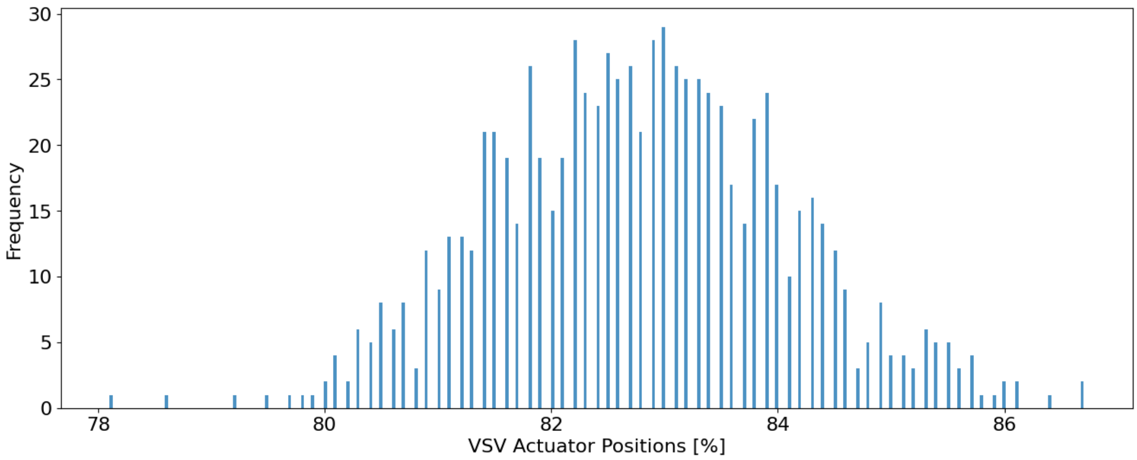


Figure 8.8: Distribution of VSV actuator position from take-off data

Based on Figure 8.8, the VSV actuator position generally lies within a relatively narrow range during take-offs yet varies significantly within it. This variation is attributed to the fact that the VSV positions are schedule with the core shaft speed N_2 and altitude. Since the only flight phase considered for this study is take-off, the effect of varying altitude should be entirely eliminated or at least sufficiently filtered out. Additionally, only flights leaving Amsterdam Schiphol airport (EHAM) are analyzed, in order to further eliminate the effect of high-altitude take-offs at airports, an example being Mexico International Airport (MMMX) with an elevation of about 2.2 km.

8.2.4. Power Off-take

Alongside parameters related to airflow control systems and active clearance control, another is the power extracted, also known as power off-take (or PTO). The onboard electrical system supplies the cabin air and other auxiliary systems aboard an aircraft. The electrical power is sourced from the rotating mechanical power of the high speed shaft, simply referred to as the N2 shaft. Extracting power from the N2 shaft results in more energy extracted from the flow, which must be then compensated by increased fuel consumption, resulting in higher turbine inlet temperature and a subsequent decrease in surge margin [75].

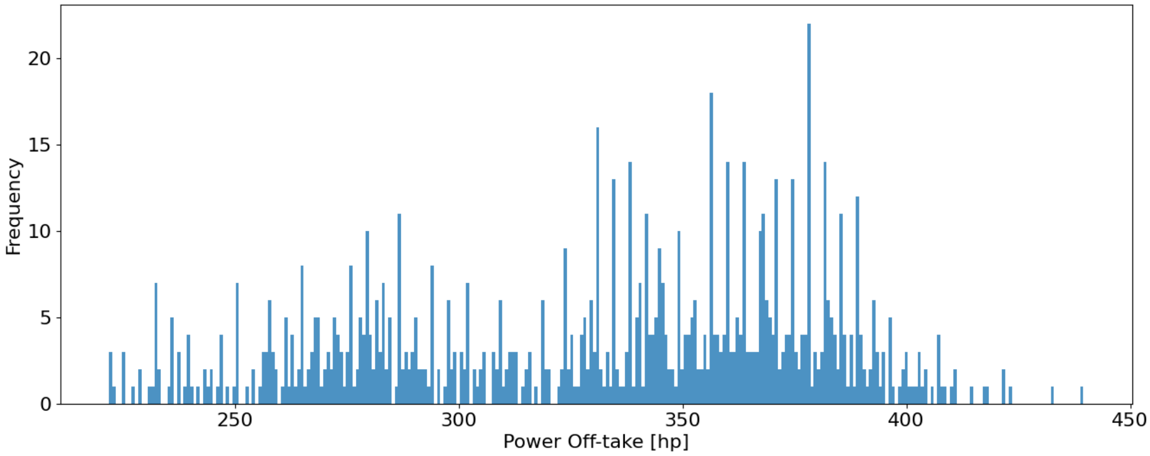


Figure 8.9: Distribution of Power Off-take

From Figure 8.9, the power extracted from the core shaft can vary significantly from one flight to another, depending on the power/thrust setting, ambient conditions and engine health condition. Since energy is extracted from the core flow, its impact on engine performance is quite apparent, thus is a relevant secondary performance parameter to implement.

8.3. Modelling SPPs

Following the conclusions made from the collected on-wing data, the SPPs to be analyzed in this thesis are the VSVs, both HPT and LPT active clearance control systems and the power off-take. For those measured in terms of actuator position, meaning the VSVs and turbine ACC systems, those are modelled using a quadratic parameterization discussed in the following subsections. The power off-take however, is the sole SPP that can be directly modelled in GSP. Using a GSP component called "Load Control", the user can specify either a power or torque load onto any shaft. In this case, a power load is imposed in terms of kW onto the high-speed (or N2) shaft. With this, the user can specify this load as an input to the GSP model via the API.

8.3.1. SPP Parameterization

Otten's [17] investigation was focused on mitigating the loss in accuracy of the GSP models when simulating operating conditions away from take-off conditions. One of the causes for these deviations was the effect of the SPPs, which prior to Otten's contribution were not initially accounted for in GSP. The result was an approach on how said secondary performance parameters could be properly quantified and accounted for in the GSP models. The approach involved an optimization using a genetic algorithm in order to fit coefficients for a quadratic relationship between a given SPP input and its effect on engine performance in terms of $\Delta\eta$ and ΔW_c. In both Equation 8.1 and Equation 8.2, the term "SPP_{OD}" refers to an input of a given SPP at a reference setting, while "SPP_{Ref}" refers to any other value away from the reference setting considered to be off-design respectively. Coefficients "a" and "b" define the magnitude of the relationship and are to be fitted.

$$\Delta\eta = a \cdot \left(\left| \frac{SPP_{OD} - SPP_{Ref}}{SPP_{Ref}} \right| \right) + b \cdot \left(\left| \frac{SPP_{OD} - SPP_{Ref}}{SPP_{Ref}} \right| \right)^2 \tag{8.1}$$

$$\Delta W_c = a \cdot \left(\left| \frac{SPP_{OD} - SPP_{Ref}}{SPP_{Ref}} \right| \right) + b \cdot \left(\left| \frac{SPP_{OD} - SPP_{Ref}}{SPP_{Ref}} \right| \right)^2 \tag{8.2}$$

This parameterization will be used in order to quantify the effects of the VSVs, HPT and LPT clearance control. Note that with this parameterization, the performance deviation will always go through the origin when and "SPP_{OD}" is equal to "SPP_{Ref}"; no deviation in SPP setting results in no performance deviation. Subsequently, the further the deviation, the steeper the performance deviation becomes.

8.3.2. Advantages & Limitations

As shown in Equation 8.1 and Equation 8.2, quadratic relations were chosen as parameteriziation for this problem. While there is no established method or decisive literature on what parameterization is appropriate, the motivation behind the choice of a quadratic relation is due to the following reasons:

- The first is that a simple linear relation is considered to be too simple of an approximation given that the intricate physical phenomena occurring with a gas turbine and between components are known to be non-linear. Therefore, it was decided that a straightforward linear approximation would not be able to sufficiently capture the dynamics involved between the gas turbine components.
- The second reason is the computational burden. While the differential evolution algorithm is said to be quite flexible, a higher degree relation would result in a larger set of coefficients to fit, thus requiring more iterations to reach a solution that both satisfies the set bounds and result in a sufficient reduction between simulation results and on-wing data
- Furthermore, during off-design operation it is assumed that if the deviation between the actual and nominal SPP setting is relatively small, the the impact on engine performance should also be small. Conversely, if the deviation continues to increase away from the nominal operational point, the resulting performance deviation should also reflect a non-linear change.

Therefore, a quadratic relation was chosen as a satisfactory trade-off between the arguments brought forward above. However, it is worth noting the limitations it brings forward.

- A quadratic relation is known to have two possible roots. Therefore, when plotted against a defined x-axis, the same curve would appear but mirrored along the y-axis. Such a parameterization would imply there is an inverted trend in performance as a function of SPP setting. For example, if the SPP setting for the HPT ACC is set to the middle of the domain of possible values and that an increasing setting is known to increase efficiency, it would imply that as the setting decreases past the reference setting the efficiency would start to increase again. This would then return a result not consistent with the known physics.
- While the effects of the SPPs are assumed to be non-linear, in this case quadratic in nature, it could lead to the scenario where if the deviation is sufficiently large, the slope of the resulting curve would significantly increase. This could then return very large performance deviation; larger than what is expected for a typical turbofan system.
- Given that the SPP settings are normalized based on a reference value, labelled SPP_{Ref}, the magnitude of the performance deviation will depend on the choice of said reference value. A sound choice would be to select the design point SPP setting, however this can vary between engine models. In this case, the design SPP settings for the GEnx-1B are not known.

8.3.3. Smearing

In addition to the limitations described above, another source of inaccuracy that could emerge from this approach is smearing. Smearing is an effect where a certain level of deterioration in component health is inappropriately attributed to spread between other components, usually between components operating on the same shaft. For example, if the deviation in performance of a certain turbofan is known to be due to a 2% decrease in efficiency of the HPC, GPA-based simulations could return results where the 2% deterioration in efficiency is split up as a 1% deterioration in efficiency in the HPT. This issue was discussed by Rootliep [56], where in Rootliep's case, turbine efficiency deterioration was simulated but the GA-based gas path analysis converged to a solution where deterioration was split equally between the compressor and turbine. According to Otten [17], this is likely caused by the lack of information in certain sections of the gas path. In this case, the lack of information is due to the reduced number of sensors. In this context, smearing could manifest during the optimization where, given the definition of the objective function, may attempt to offset the on-wing deviation by increasing the performance of certain components than others. In other words, the effect of deterioration in certain components may be attributed more importance than others.

8.3.4. Imposed Constraints

In order to mitigate the limitations of this parameterization, certain constraints have to be placed. Firstly, the bounds for each coefficient during the optimization must be constrained such that the resulting curve returns the expected slope in line with the expected physical behaviour. For example, as the SPP setting for the HPT and LPT ACC increases, an increase in efficiency is expected. A similar conclusion is made for the corrected flow. As for the VSVs, it is expected that as their setting increases, the VSVs get more and more closed; their angle relative to the axial flow increases. This would in turn decrease the resulting corrected flow with an increasing slope downward. The set of possible curves obtainable from this parameterization are shown below in Figure 8.10

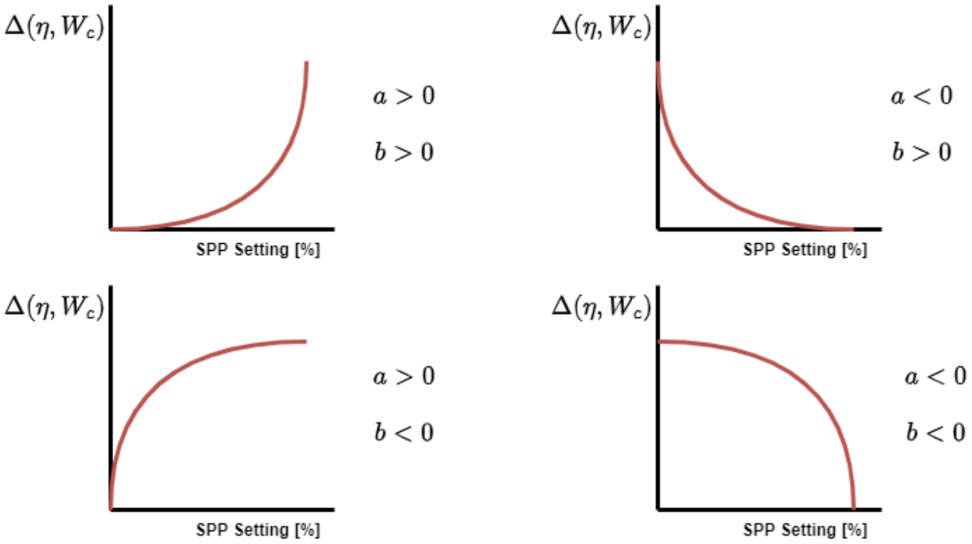


Figure 8.10: Possible curves for a parameterization of the form $ax + bx^2$

From the set of possible curves illustrated in Figure 8.10, the coefficients for the performance deviation of the SPPs to be studied in this thesis are constrained as follows:

Table 8.1: Constraints imposed on coefficients

	$\Delta\eta$	ΔW_c
VSV	$a > 0, b > 0$	$a < 0, b < 0$
HPTACC	$a > 0, b > 0$	$a < 0, b < 0$
LPTACC	$a > 0, b > 0$	$a < 0, b < 0$

The resulting performance deviation is expected to fall within a certain range. The bounds for said performance deviations are stated below in Table 8.2. Depending on the choice of bounds for the optimization, there could be cases where the algorithm returns a set of coefficients which result in exceedingly large performance deviation; higher than what is physically expected. The results obtained from the optimization are considered to be adequate if the resulting relations return performance deviation within the following ranges.

Table 8.2: Expected range of performance deviation

SPP	**Lower Bound [%]**		**Upper Bound [%]**	
	$\Delta\eta$	ΔW_c	$\Delta\eta$	ΔW_c
VSV	-2	-10	3	4
HPTACC	-2	-4	3	2
LPTACC	-2	-4	3	2

8.3.5. Reference Settings

The parameterization of the SPPs involves a term labelled SPP_{Ref}. This term is a reference value for a given SPP in order to compute the deviation from what is considered nominal operation, during off-design operation. Since the exact design settings for the GEnx-1B are not known, in order to capture the SPP effects, it is assumed that no other source of engine deterioration is present, except for those due to SPP variation from the reference setting. Therefore, the objective function must be evaluated using on-wing measurements from a "healthy" engine; assumed to have no component deterioration. Based on the EGTHDM of the reference engine selected sown in Section 8.1, it is deemed to be sufficiently healthy, thus it is assumed that any performance deviation is primarily due to the SPP settings.

As discussed in Section 8.3.2, the choice of SPP_{Ref} is important and can lead to different results. That being said, it is more important that the choice of SPP_{Ref} remains consistent throughout this analysis. Taking into account the limitations of the parameterization outlined in Section 8.3.2, after inspecting the distributions of the various SPPs from the on-wing data, the minimum values within the distributions were selected as the reference settings for the VSV actuator and turbine ACC valve positions.

This choice is based on the assumption that the SPPs are present to offset progressive deterioration of engine performance. For example, as an engine deteriorates, the EGT tends to progressively increase on average. Therefore, the EMU will demand more from the ACC system in order to offset the effect of tip clearance and increased thermal loads. Since the parameterization is quadratic, the optimization could risk reaching a solution that would return positive efficiency change for increasing clearance, which is contrary to its intended effect. Thus selecting the minimum values as reference settings will ensure the domain of input values remains on the positive side after the origin.

Table 8.3: Selected Reference SPP settings

SPP Position [%]	VSV_x	$HPTACC_x$	$LPTACC_x$
Reference Setting	78.1	9.4	43.0

8.3.6. GA-based Optimization

The coefficients "a" and "b" in Equation 8.1 and Equation 8.2 are then fitted via an optimization scheme using a genetic algorithm, known as differential evolution (or DE) as the optimizer. The differential evolution optimizer is provided in the SciPy Optimization package for Python. More details on the differential evolution algorithm are provided in Appendix C.

The driving hypothesis is that as the optimizer approaches a set of coefficients that returns an appropriate relationship between a given SPP setting and its effect on engine performance, the deviation between the GSP model output and on-wing measurements will progressively decrease. Therefore, the objective function is defined as follows in Equation 8.3:

$$f(x) = \sqrt{\frac{\sum_{m=1}^{meas} \left(\frac{y_{meas,m} - y_{out,m}}{y_{meas,m}}\right)^2}{N_{meas}}} \tag{8.3}$$

... where $y_{out,m}$ is the GSP model output at a given operating point and $y_{meas,m}$ is the measured on-wing parameters. The objective function is evaluated as various operating points from on-wing measurements by running the GSP model using the on-wing flight conditions as inputs. The resulting engine performance parameters are then compared with those from on-wing measurements, thus returning their deviation. The process is summarized in Figure 8.11. Before initiating the optimization, certain parameters must be set. The Differential Evolution algorithm requires the following parameters to be selected:

Table 8.4: List of DE settings

Parameter	Value	Description
N_i	30	Maximum number of iterations
N_P	8	Population size for each generation
X_{tol}	10^{-4}	Termination tolerance for objective function
CR	0.5	Probability of crossover
F	$[0.5, 1.0]$	Probability of mutation; two values for dithering

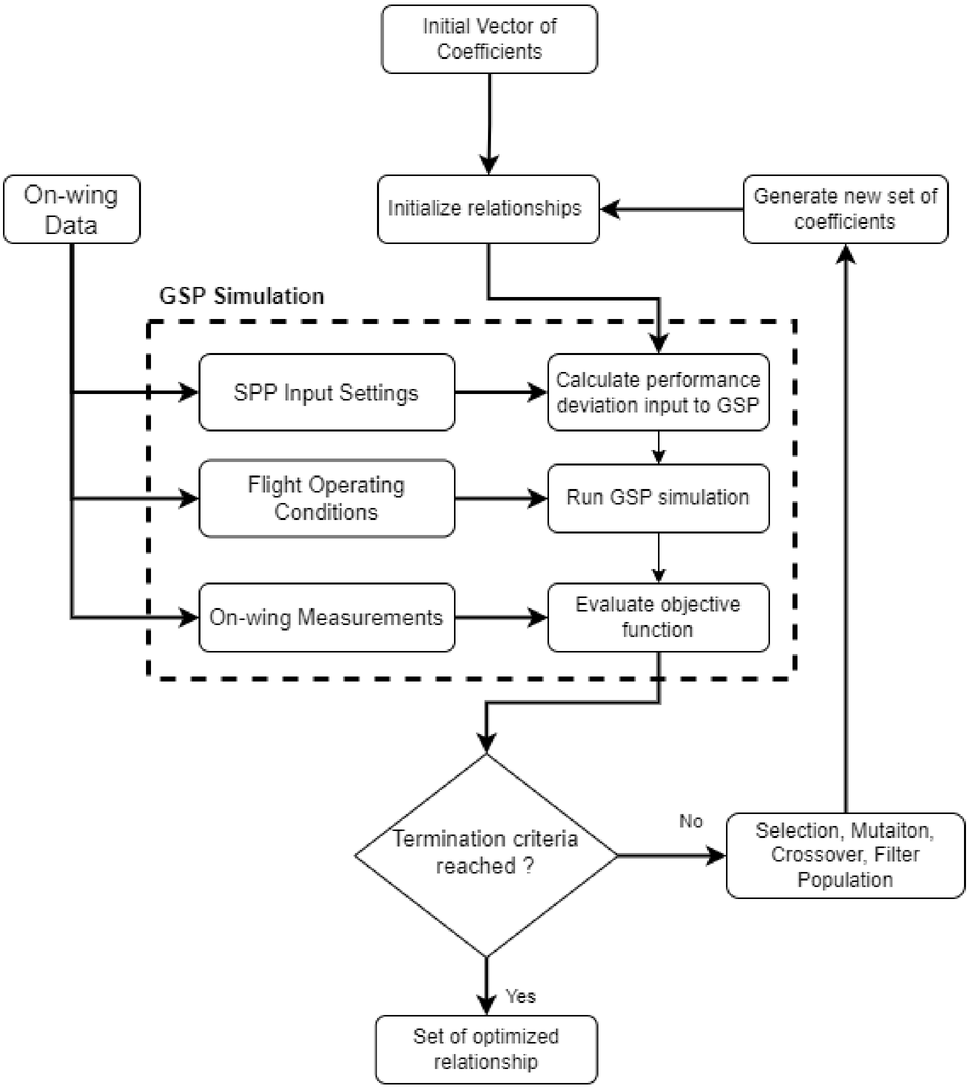


Figure 8.11: Flowchart of GA-based optimization of SPP relations

8.4. Results & Assessment

After having conducted the optimization using the genetic algorithm as the optimizer, the following results were obtained. The increase in accuracy due to the adapted relations is presented in terms of the percentage deviation between the GSP model output and the on-wing data. After that, the resulting curves from the relations are analyzed and discussed.

8.4.1. Model Deviation

Below in Figure 8.12 is bar chart of the resulting model deviation between the GSP model and the gas path measurements from on-wing data during take-off. The blue bars represent the deviation of the Baseline model; the standard GSP model of the GEnx-1B with only flight operating conditions as input. The orange bars represent the deviation from the GSP model but with deterioration effects active, where the magnitude of the deterioration is based on the various SPP settings recorded for that flight. Note that a positive deviation means the GSP model returns a higher value compared to on-wing measurements.

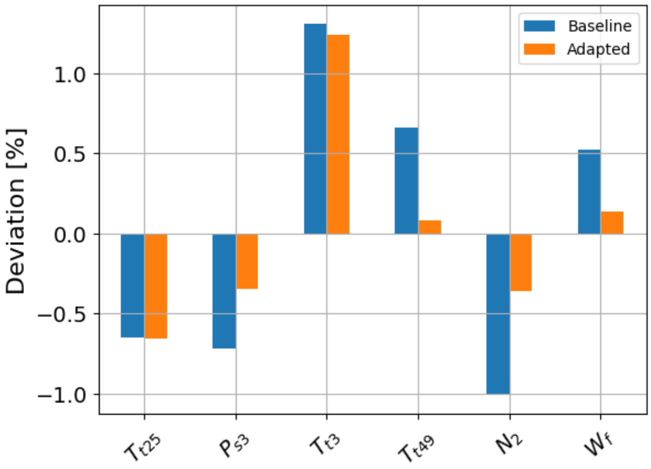


Figure 8.12: Deviation between GSP model output and on-wing data

From the results shown above, it can be seen that the adapted relations have contributed to a substantial decrease in the deviation between the GSP model output and the on-wing data. The most notable improvements are seen for T_{t49} and W_f, which have experienced the largest decrease in deviation. In contrast, the deviation in T_{t25} and T_{t3} were minimal. The magnitude of the decrease in deviation can be better seen in Figure 8.13, which shows the change relative to the baseline deviation.

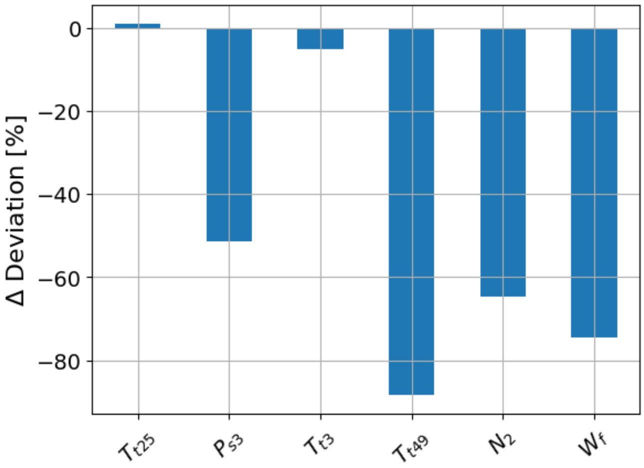


Figure 8.13: Deviation between adapted GSP model output and on-wing data in [%] relative to Baseline

From these results, it is clear that T_{t49} experienced the largest improvement in terms of on-wing deviation; almost a 90% decrease in deviation from on-wing data relative to the baseline, while N_2 and W_f have experienced a decrease of over 60% and 70% respectively. Similarly, a decrease in deviation almost 50% was reported for P_{s3}. Based on these results, the GA-based optimization seems to have returned relations that better approximate engine performance in terms of P_{s3}, T_{t49}, N_2 and W_f when accounting for the effects of SPPs. More is discussed in the following subsection regarding the resulting curves from the SPP relations.That being said, the reduction in deviation for the remaining parameters, T_{t25} and T_{t3}, are quite small by comparison. In the case of T_{t3}, only a reduction of approximately 5% was achieved. As for T_{t25}, the deviation seemingly increased, however by less than 1%.

Since it is measured at the Fan+LPC outlet, the minimal effect on T_{t25} is a consequence of the fact that none of the components in the engine core have an effect of the Fan+LPC located upstream. Furthermore, during the GSP simulations, no SPP is directly affecting the deterioration level of the Fan+LPC component, therefore its performance remains the same. This is also supported by the sensitivity analysis of the GSP model shown in Figure 6.3; since only $\Delta\eta$ and ΔW_c of the HPC, HPT and LPT are influenced by the SPPs being studied, the performance of the Fan+LPC component remains largely unchanged. Furthermore, it is worth noting that no gas path measurement is present at the LPT outlet. Since the Fan+LPC and LPT are connected via the N1 shaft, the contribution of that measurement deviation could have potentially help drive the optimization towards a reduced deviation in T_{t25}. That being said, the minor increase in deviation is attributed to the small modelling inaccuracies inherent within the adapted component map, since they can never be 100% accurate.

As for T_{t3}, the small decrease in deviation is likely due to the optimization's strategy in reducing the objective function. Since the objective function is defined as the sum of the individual deviations, the optimization converged towards to a solution that returns the lowest objective value. That being said, the lowest objective function value achieved in this case is one that had to balance the total deviation between certain parameters. In order to reach such large reductions with the established constraints, certain gas path parameters were focused on more than others.

8.4.2. Standard Deviation

Another method to assess the change in accuracy of the adapted model is to inspect how the standard deviation between the GSP model output and the on-wing measurements change. In this case, the standard deviation σ is a measure of how much the resulting data points are dispersed and spread out from the mean value.

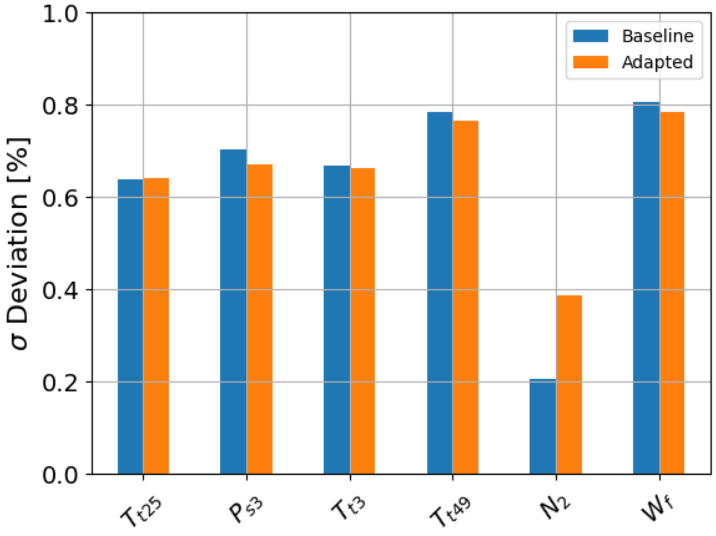


Figure 8.14: Standard deviation σ of model error against on-wing measurement

The standard deviation of error for each gas path sensor is shown in Figure 8.14. The Standard Deviations have experienced a relatively small but general decrease across most measurements, except for T_{t25} which has marginally changed, and the core shaft speed N_2 which experienced an increase. Similar to results shown in Section 8.4.1, T_{t25} has seen minimal change given the fact that the Fan+LPC component map was not directly affected by any of the SPPs. Therefore, any change seen in T_{t25} would be a result of small modelling inaccuracies in the GSP model's component maps or measurement noise. This would imply that this implementation of the SPP effects had a much smaller effect on the GSP model's accuracy at take-off. The relatively small decrease in standard deviation σ would mean that some level of inaccuracy, while relatively small, still persists when simulating take-off conditions. That being said, these results still show a general net positive, thus they are considered to be a satisfactory result.

8.4.3. Resulting relations

Using the optimized set of coefficients obtained from the optimization, the resulting curves from the SPP relations were generated and shown below for each SPP. Analyzing the resulting curves is done as a verification step; to collect insight on the results from the optimization before moving onto the next phase of the thesis. For clarification, the horizontal axes are shown in terms of the difference between the selected reference SPP setting and the setting recorded for a given flight. The vertical axes show the performance deviation for that SPP in terms of $\Delta\eta$ and ΔW_c.

Variable Stator Vanes

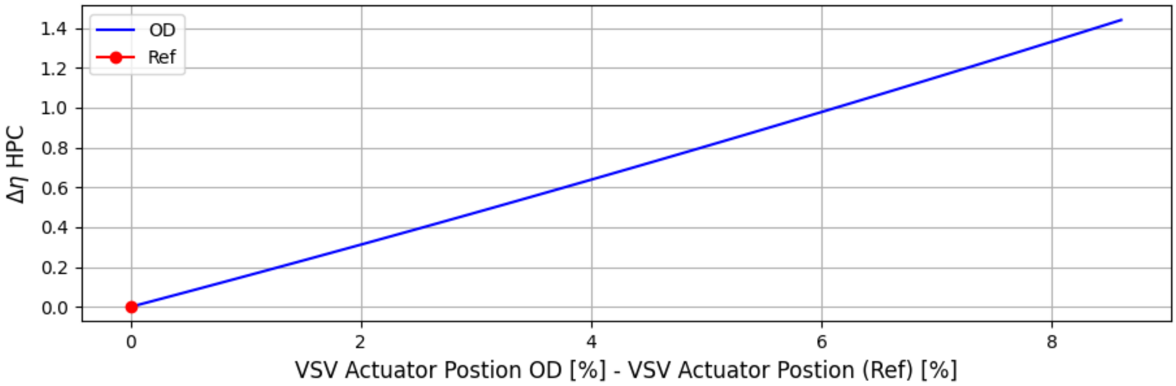


Figure 8.15: Resulting curve from SPP relation for effect of VSVs on HPC efficiency

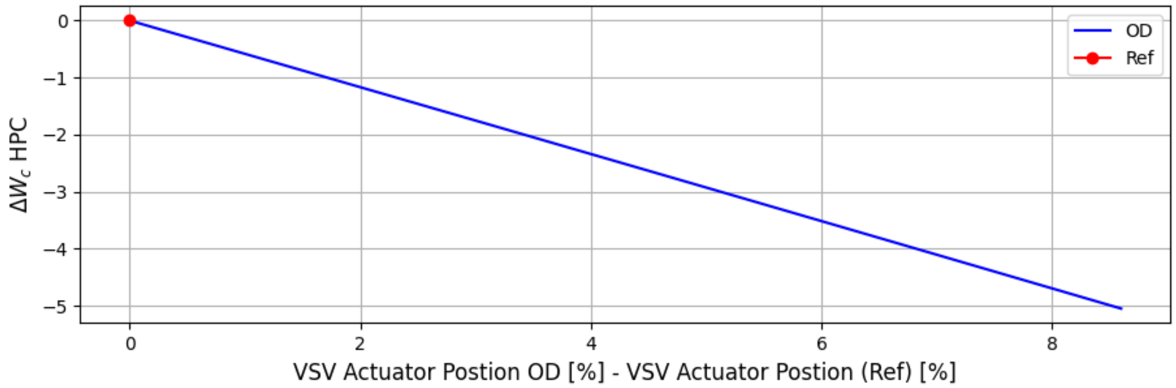


Figure 8.16: Resulting curve from SPP relation for effect of VSVs on HPC corrected flow

Looking at Figure 8.16, the GA-based optimization returned a relation where as the VSV actuator position increases resulting in the closing of the VSVs, the corrected mass flow through the HPC significantly decreases. This effect is consistent with the expected behaviour prescribed by the coefficient bounds set in Section 8.3.1. It is worth noting that the relation between the VSVs and their effect on the

HPC performance resulted in a seemingly linear relationship. Given that the parameterization allowed for a non-linear relation to emerge, the optimization arrived at a solution that exhibits a somewhat linear relation. However, after extrapolating to actuator positions past those shown in Figure 8.15 and Figure 8.16, the relation does reveal a quadratic curve. The combination of coefficients obtained from the optimization resulted in a quadratic curve of relatively shallow slope near the origin. Therefore, the relatively linear slope is due to the fact that the deviation in VSV settings at take-off from the VSV reference value covers a small interval, resulting in a locally linear relation.

Additionally, with decreasing corrected flow, the efficiency is also increased up to a maximum increase of about 1.4%, in order to compensate for the fact that the operating point on the HPC's component map has now shifted to a zone of lower efficiency but also further from the surge line. As a result, the HPC operates at roughly the same efficiency and corrected flow but with an increased N2 corrected shaft speed. During take-off, a certain thrust setting is demanded by the pilot, translating to a selected N1 speed. With a reduced corrected flow, the N2 shaft speed must also increase in order to provide the same power output for the same selected N1 speed. This is also visible in Figure 8.12, which indicates that the N2 speed was increased in order to close the gap between the on-wing data and the Baseline, which initially was underestimating the N2 speed. This suggests that the adapted GSP model reached the expected solution of increasing the N2 speed in order to match the required performance.

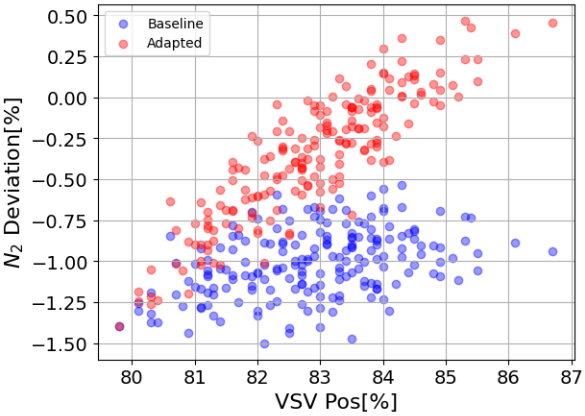


Figure 8.17: Deviation in N_2 as a function of VSV actuator position

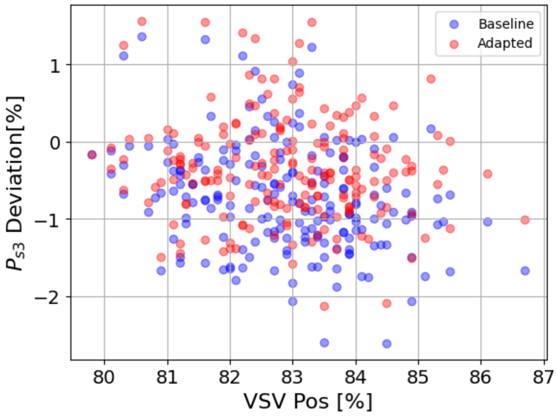


Figure 8.18: Deviation in P_{s3} as a function of VSV actuator position

Figure 8.17 and Figure 8.18 illustrate the model deviation in P_{s3} and N_2 against the actuator position of the VSV over the range of values analyzed at take-off. In the case of P_{s3}, as the VSV actuator position increases, the magnitude of the difference between the baseline and adapted model deviation increases. This is expected given that larger VSV actuator positions translate to an increase efficiency and decrease corrected flow modifiers. However, this also results in the case where the adapted GSP model has overestimated P_{s3} compared to on-wing measurements.

Furthermore, the data points in Figure 8.17 the Adapted case are now less spread out compared to the Baseline case. Furthermore, the mean N_2 model deviation was substantially decreased in the Adapted case, That being said, the standard deviation in the Adapted case has increased compared to the Baseline case. Therefore, since the majority of points in Figure 8.17 for the Adapted case are spread over a larger range along the y-axis, most are from the mean value, resulting in a larger standard deviation. This seems to be due to the N_2 model deviation in the Adapted case now exhibiting a more positive correlation with the actuator position; the N_2 model deviation increase from negative to positive values as the actuator position increases, similar to P_{s3} in Figure 8.18. The correlation then levels off near the higher actuator positions. This is likely due to the change in ambient conditions affecting the resulting corrected fan speed. VSV position is also set based on ambient conditions and how it affects engine performance during transient operations at take-off. As the VSV position increases, the resulting operating point on the component map shift further from the cycle reference point along its operating line where a different N_2 is likely needed to better match the on-wing data. Consequently, this splits the results into two regions where the model deviation is underestimated on one side of the interval and overestimated on the other, resulting in a larger standard deviation.

HPT Active Clearance Control

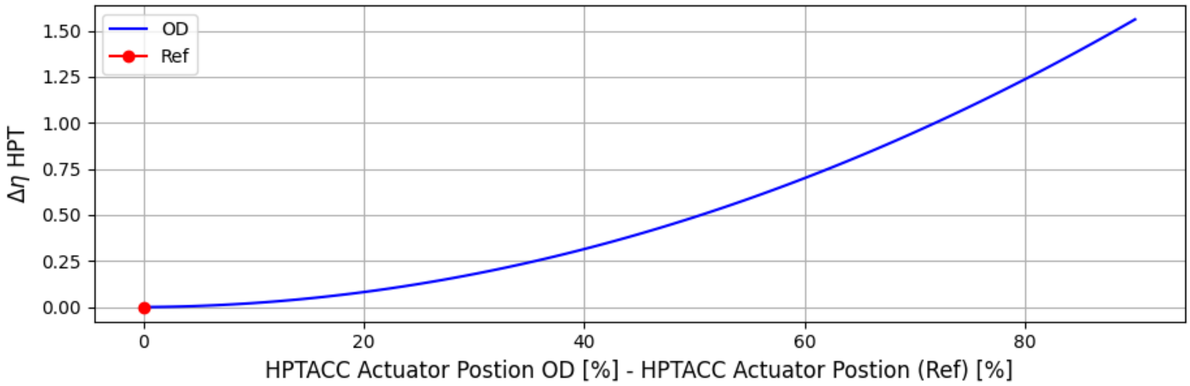


Figure 8.19: Resulting curve from SPP relation for effect of ACC on HPT efficiency

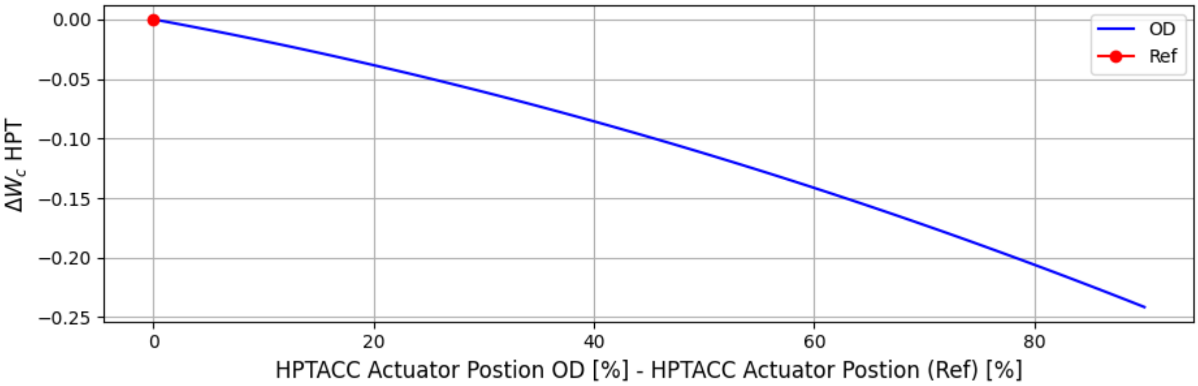


Figure 8.20: Resulting curve from SPP relation for effect of ACC on HPT corrected flow

Compared to the resulting curves for the VSV settings, the HPTACC setting vary over a larger interval, thus showing a clearer quadratic behaviour as expected from the imposed constraints stated in Table 8.1. As the HPTACC valve position increases, more cooling flow is directed to the engine casing surrounding the turbine section, thereby reducing the thermal expansion experienced by the casing and reducing the clearance between it and the turbine blades. This leads to an increase in efficiency and a subsequent decrease in corrected flow capacity, as stated in literature.

According to the sensitivity analysis of the GSP model, an increase in HPT efficiency leads to an increase in both N_2 speed and HPC output pressure P_{s3} while reducing the exhaust temperature T_{t49}. This is also substantiated by the decrease in on-wing deviation. Figure 8.13 shows that T_{t49} has benefited from the largest decrease in deviation, almost 90%. Figure 8.12 indicates that the Baseline model overestimated T_{t49}, therefore the Adapted model arrived at a solution that decreased the output EGT. Figure 8.21 through Figure 8.23 show the relationship between the HPTACC valve position and its effect on the model deviation of T_{t49}, N_2, W_f and P_{s3}. These parameters are particularly relevant to the HPT according to the model sensitivity response in Figure 6.3.

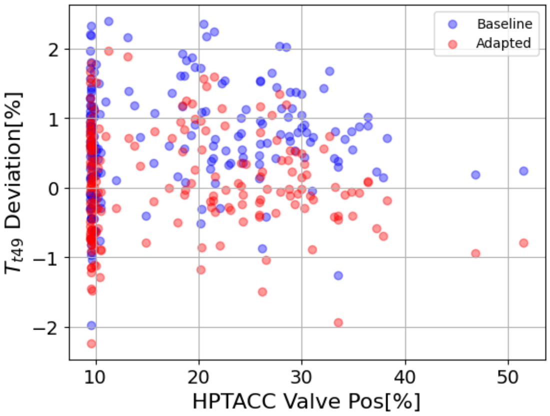


Figure 8.21: Deviation in T_{t49} as a function of HPTACC valve position

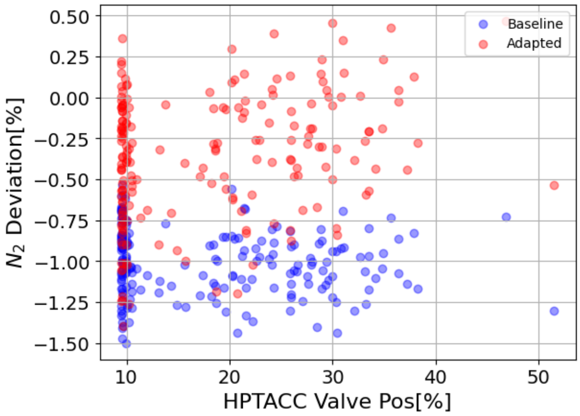


Figure 8.22: Deviation in N_2 as a function of HPTACC valve position

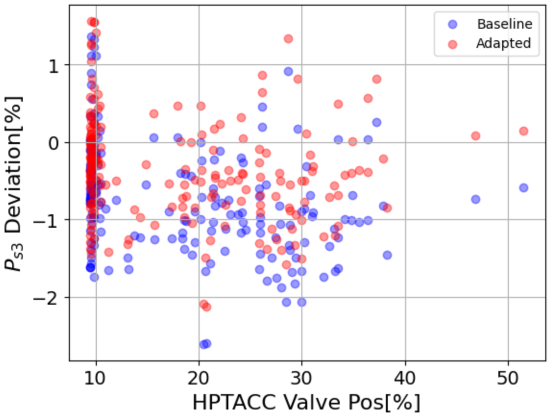


Figure 8.23: Deviation in P_{s3} as a function of HPTACC valve position

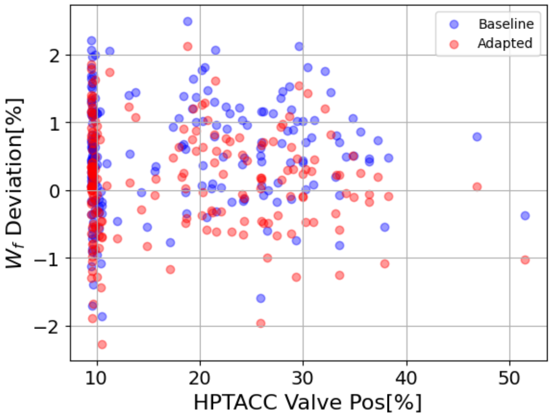


Figure 8.24: Deviation in W_f as a function of HPTACC valve position

The resulting model deviation for those parameters shows similar results to what is seen in the case of the VSVs. As the HPTACC valve position increases, the efficiency is expected to increase following a reduction in clearance, subsequently leading to a reduction in EGT (T_{t49}), as shown in the deviation in T_{t49} for the adapted model in Figure 8.21. A negative correlation is visible in the case of T_{t49}, where as the HPTACC valve continues to open, the resulting EGT on average start to be underestimated, shifting points below the 0% deviation line. A similar result in visible in Figure 8.22, where as the HPTACC valve position increases, the resulting N_2 increases towards 0% deviation. A more efficient HPT would result in more energy extracted from the core flow leading to a lower EGT. As for the N_2, a more efficient HPT can deliver faster rotational speeds since more work is extracted and converted to mechanical energy for the high-speed shaft to drive the HPC. A positive correlation is also visible with HPTACC position, indicating that as the ACC valve is more open, the increase in efficiency leads to lower EGT.However, similar to previous results, the increase in efficiency leads to an overestimation. Furthermore, as Figure 8.14 indicate, the resulting data points for the Adapted case are more dispersed compared to the Baseline model, still indicating a persisting mismatch between the adpated model and on-wing measurements.

Furthermore, a sizeable amount of flights have the HPTACC valve set to a position of approximately 10%, leading to the dense collection of points on the plots above. This follows a similar distribution of points seen when inspecting the calculated HPT clearance found in the on-wing data. Figure 8.25 shows said calculated clearance as a function of the HPTACC valve position. Note that the actual clearance is normalized for the sake of confidentiality.

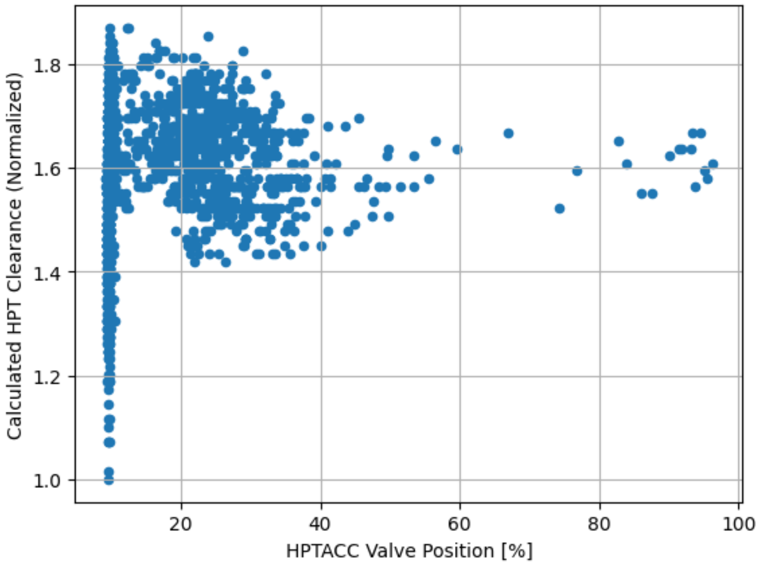


Figure 8.25: Calculated HPT Clearance as a function of HPTACC valve

Similar to Figure 8.21 through Figure 8.24, a large density of flights at take-off lie around the 10% valve position. More importantly, the calculated clearance varies significantly at that position over various flights, by over 90%. Figure 8.25 demonstrates that said calculated clearance varies significantly from flight to flight for the same HPTACC valve position. This would imply that the calculated clearance depends on several other factors. If the calculated HPT clearance as given by the EMU is assumed to be the most accurate metric, further reductions in model deviation from on-wing data would then require a more complex parameterization in order to more accurately capture the effects of the HPTACC during take-off.

LPT Active Clearance Control

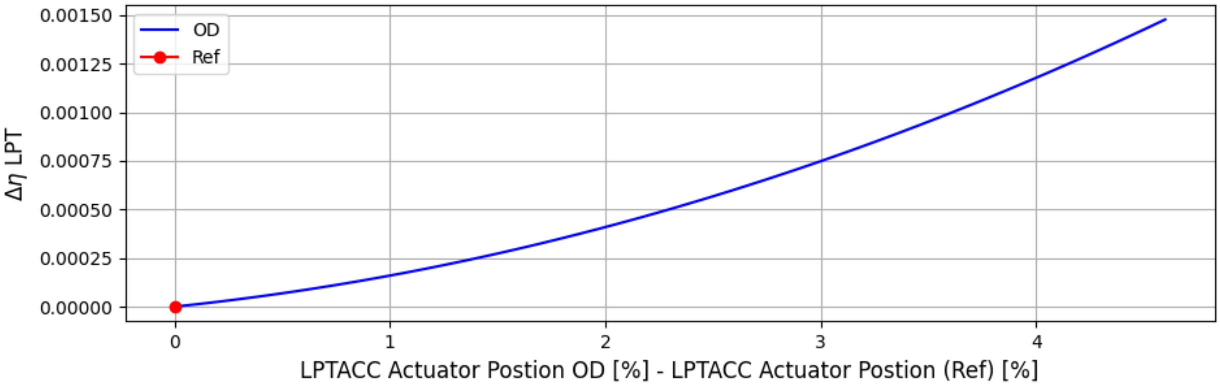


Figure 8.26: Resulting curve from SPP relation for effect of ACC on LPT efficiency

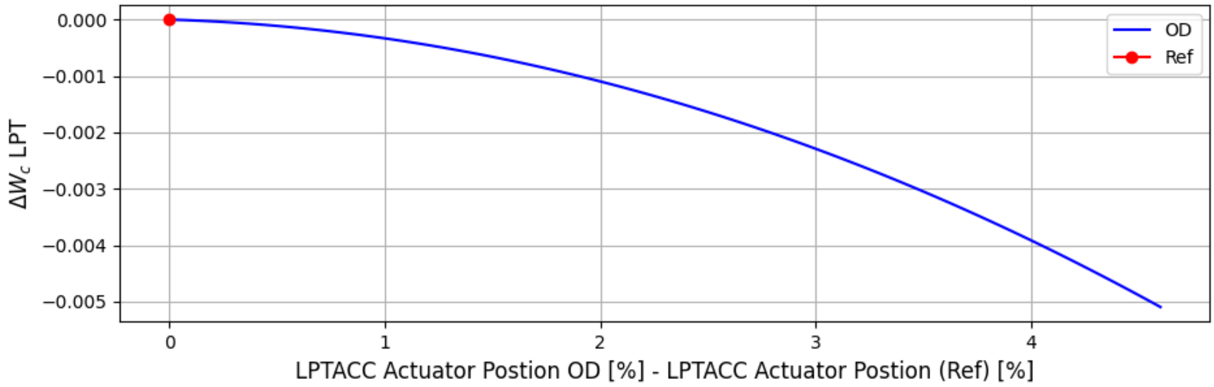


Figure 8.27: Resulting curve from SPP relation for effect of ACC on LPT corrected flow

As for the LPTACC, results similar to the HPTACC were obtained. In this case, the performance deviation in terms of $\Delta\eta$ and ΔW_c are substantially lower compared to those from the HPT. This is in part due to the smaller interval of valve positions reported in the on-wing data. Referring back to the distribution of valve settings seen at take-off shown in Figure 8.5, the LPTACC system is barely actuated throughout multiple flights, as compared to the HPTACC. This would imply that ACC is often not needed downstream in the LPT, due the lower temperatures found in the exhaust flow after crossing the HPT, implying the effects of thermal expansion are less serious. Due to this limited window, the optimization must have converged to a solution that resulted in smaller performance impacts with varying valve positions compared to the HPTACC as long at it results in a lower objective function value.

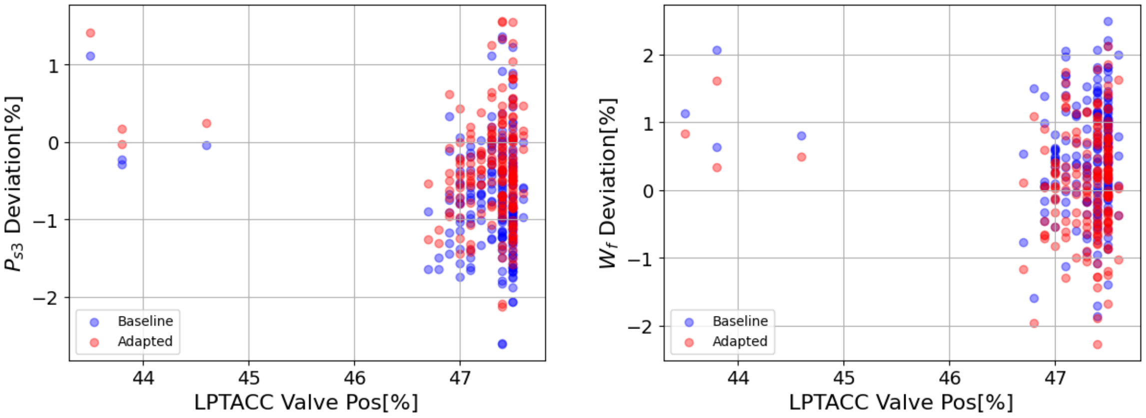


Figure 8.28: Deviation in P_{s3} as a function of LPTACC valve position

Figure 8.29: Deviation in W_f as a function of LPTACC valve position

That being said, based on the sensitivity analysis of the GSP model, a reduction in corrected flow and increase in efficiency of the LPT results in substantial decrease in fuel consumption and increase in HPC outlet pressure respectively. Figure 8.28 and Figure 8.29 show a general decrease in average model deviation for P_{s3} and W_f as the LPTACC valve is engaged, however relatively small compared to previous results. It is worth noting that, similar to the HPTACC, the majority of points are located in a dense region between positions 47% and 47.5%. If previous results from the HPTACC are of any indication, this is also likely due to the variation in actual LPT clearance as a function of other variables. Furthermore, since the temperature sensor for station 5 is not available for the GEnx-1B, this lack of information on the thermodyanmic state of the flow at the results in less information available for the optimization to better account for the impact of the LPT, thus explaining why the magnitude of the performance impact is quite small. Unfortunately, no information or calculated values for the LPT clearance are provided, thus no major comparisons can be made. However, it is safe to assume that the behaviour of the actual LPT clearance follows a similar pattern and correlation to that seen for the HPT, given that both their ACC systems function almost identically.

8.5. Discussion

In this chapter, the results from the GA-based optimization, designed to provide an adequate combination of coefficients for the proposed parameterization of the SPPs, were evaluated and discussed. The resulting deviation between the GSP model output and on-wing data showed a significant reduction for the measurements of P_{s3}, T_{t49}, N_2 and W_f, while T_{t3} showed a marginal improvement and practically little to no effect on T_{t25}. The behaviour of the associated quadratic curves were consistent with the known physics and the magnitude of the resulting component modifiers is within the range previously stated in Section 8.3.4. However, it is not as straightforward to draw conclusions on the resulting magnitude of the performance deviations, since there is no exact literature available in order to more directly translate the effects of these secondary systems to the engine's component maps.

Despite obtaining a general decrease in deviation, it can also be seen that the deviation of certain gas path parameters, such as T_{t25} and T_{t3}, have experienced either very little decrease or deviate more compared to the baseline. T_{t3} shows a marginal decrease, substantially lower than others found downstream of the engine. Meanwhile, T_{t25} has seen a slight increase in deviation. This behaviour is likely due to the effect known as smearing, as explained in Section 8.3.3. When running the GA-based optimization, smearing may manifest by converging to a solution where certain coefficients are given a larger magnitude to certain SPPs, hereby attributing a larger importance on the effect of some parameters and offsetting the effects of others. this is likely the cause behind the relatively small magnitude of the performance deviation associated with the LPT active clearance control, as seen in Figure 8.26 and Figure 8.27.

Furthermore, when inspecting the model deviation for individual gas path measurement as a function of the SPP settings, certain plots revealed correlations between the error and level of SPP actuation. In an ideal case, no correlation is preferred; the deviations would shift to near 0% no matter the actuator position, indicating an accurate one-to-one correction of the model output. However in some cases, some correlation is present, primarily attributed to the quadratic behaviour of the parameterization used.

Another potential source for the present correlations may be the fact that take-off was mainly analyzed. Results obtained by Otten [17] showed satisfactory estimations of the SPP effects, despite being limited to a less accurate GSP model of the GEnx-1B with no tuned maps. Otten focused on on-wing data measured at cruise, while the SPP reference settings selected were those measured at take-off, effectively selecting the take-off point as the representative design point. However, the inaccuracies present in those results were attributed to the mismatch in engine performance between measurements and GSP simulations caused by altitude-related effects. In the case of this thesis, it was initially intended to mitigate these effects by focusing on the take-off phase, where the current GSP model of the GEnx-1B developed by Ramdin was known to be the most accurate. Snapshot data was used for this analysis, which unfortunately does not include the ΔN_2 as measurement, therefore take-off measurements were assumed to be taken at an operating condition sufficiently close to steady-state. Possible transient effects may still be present within measurements, further contributing to the model deviation.

Fortunately, from a modelling perspective, the general decrease in on-wing deviation indicates that the resulting relations provide a sufficient numerical approximation of the SPP effects on engine performance in order to be implemented to GSP. Therefore, the following chapter will carry over these results to be later implemented into the HDMR development framework in order to expand it to include these additional parameters.

9

Modified Surrogate Model Results

Following the parameterization and optimization of the SPP relations, this chapter will focus on their implementation into the HDMR-based surrogate modelling framework developed at KLM ES. Section 9.1 will provide a quick overview of the mathematical structure behind HDMRs, followed by Section 9.3 which discusses the procedure employed to generate the appropriate training data for the generation of HDMRs. Finally Section 9.4 and Section 9.5 provides the results from the new implementations and condition trending, followed by concluding remarks in Section 9.6.

9.1. HDMR Generation

As stated before in Section 4.3.4, the High Dimensional Model Representation (or HDMR) approach allows for modelling of high dimensional systems, using expert knowledge and data. At its core, the model consists of a series of equations made of linear combinations of several basis functions. To reiterate, a HDMR $f(\bar{x})$ can be expressed as shown below in Equation 9.1, for a system consisting of m dimensions and a vector of variables $\bar{x} = [x_1 x_2 ... x_m]$.

$$f(\bar{x}) = g_0 + \sum_{i=1}^{m} g_i(x_i) + \sum_{1 \leq i_1 \leq i_2 \leq m} g_{i_1 i_2}(x_1, x_2) + \sum_{1 \leq i_1 \leq i_2 \leq i_3 \leq m} g_{i_1 i_2 i_3}(x_1, x_2, x_3) + ... \tag{9.1}$$

$$N_{terms} = \binom{m}{0} + \binom{m}{1} + ... + \binom{m}{m} = \Sigma_{i=0}^{m} \binom{m}{i} = 2^m \tag{9.2}$$

From Equation 9.1, it is clear that the amount of terms included in $f(\bar{x})$ increases rapidly with increasing number of dimensions m. To be exact, the number of terms can be expressed as such in Equation 9.2, indicating that the number of terms increases exponentially for every added dimension. For example, a system consisting of 4 dimensions would return an HDMR consisting of 16 terms, 5 dimensions would return 32 terms, 64 terms for 6 dimensions and so forth, hereby doubling for every new dimension included. Depending on the chosen order, a single term within a HDMR can then have a significantly large number of coefficients to be fitted, resulting in significant computational cost. Therefore, in order to limit the computational cost of generating and fitting several basis functions, the number of terms must be limited to some extent.

Additionally, an important distinction must be made regarding the terms "degree" and "order". In this context, "degree" refers to the exponent set in the basis function, where as "order" refers to the dimensionality of the basis function. Equation 9.3 shows an example of a second degree, first order basis function, whereas Equation 9.4 shows an example of a second order, first degree basis function. Moving forward, these definitions of degree and order will be respected for the remainder of this thesis. given that this convention is also adhered to in literature [76] [77] and subsequently de Bruin [5]. Based on existing literature involving HDMRs, several publications have reported that a sufficient level of accuracy is still achieved with second-order HDMRs while eliminating higher-order terms [78][79]. Similar results were achieved in publications where such a trade-off was also implemented for cases involving gas turbine modelling [24] [80]. Neglecting terms higher than second order thus results in a substantial decrease in terms. For example, a 5 dimensional second order system, originally with 32 terms, is now reduced to 16; effectively halving the amount in terms and substantially decreasing model complexity without having to incur a significant penalty to accuracy, according to de Bruin's results. Therefore, based on these findings, it was decided to limit the generation of HDMRs to only second-order terms in order to avoid rapidly increasing model complexity, given that additional parameters will be implemented in the form of additional basis functions.

9.1.1. Basis Functions

Several methods exist for constructing basis functions, such as Radial Basis Functions [77], Kriging [81], or various functions defined the user. User-defined functions is the most straightforward method, however it requires that the user already has extensive knowledge of the system at hand. That being said, user-defined functions is what allows the user to embed expert knowledge within the model. The basis functions $g(...)$ used in this case are standard polynomials of an arbitrary degree and and order.

$$a \cdot (x_1 - 1) \ , \ a \cdot \left(x_1^2 - 1\right) \tag{9.3}$$

$$a \cdot (x_1 \cdot x_2 - 1) \ , \ a \cdot \left(x_1^2 \cdot x_2 - 1\right) \ , \ a \cdot \left(x_1 \cdot x_2^2 - 1\right) \ , \ a \cdot \left(x_1^2 \cdot x_2^2 - 1\right) \tag{9.4}$$

The examples shown above also represent the form used for the HDMRs to be generated in this thesis. The coefficients "a" are to be fitted based on training data, whose generation is outlined below in Section 9.3. Note that the zeroth order term, g_0, is removed from the generation process. Given that the basis functions are of the form $x_n - 1$ and that the input parameters are baseline normalized, any sample point of a healthy engine located in the middle of the domain of ambient conditions would equal to 1 by definition, Hence, with no deviation from the baseline, the HDMR would return zero degradation given it matches the baseline measurements, taken from what is considered a healthy engine. At this point, the basis function is defined to equal zero, thus removing the need of a constant zeroth order term.

9.1.2. Data Normalization & Correction

During the HDMR generation process, the training data is normalized using two different methods depending on the situation. The range of the various parameters in the target data can vary significantly. Therefore, to allow for better model comparisons in terms of Root Mean Squared Error (or RMSE), the target data generated for all HDMRs is normalized using Min-Max normalization as shown in Equation 9.5.

$$x_i = \frac{x_i - \min(x_i)}{\max(x_i) - min(x_i)} \tag{9.5}$$

$$x_i = \frac{x_i}{x_{i,base}} \tag{9.6}$$

Alongside Min-Max normalization, in order to ensure better convergence and accuracy rates during HDMR generation, the GSP outputs for engine measurements are normalized using baseline normalization, as shown in Equation 9.6. This also allows for the HDMR inputs to be normalized against a reference healthy engine simulation. Baseline normalization helps to avoid large gradients during the coefficient fitting process in situations such as, for example. pressure values in the order of 10^6 or fuel flows in the order of 10^{-1}. The baseline values are sensor measurements taken from an engine considered to be healthy and under no degradation.

9.2. Recommended Improvements

As stated previously in Chapter 6, several recommendations were made following the development of the HDMR-based surrogate model for condition diagnostics.

The first recommendation made was to allow the corrected fan speed to vary. Initially, training data was restricted to operating points locked at a corrected fan speed N_{1c} of 91.6%. This value was selected by de Bruin to be both within the accurate operating range of the GSP model and to keep model complexity to a minimum. While the benefit in computational cost was apparent, the downside of this decision was the limitation on the acceptable input range of CEOD. Limiting the fan speed to a specific value reduces the amount of CEOD suitable for input; only operating points sufficiently close to the prescribed N_{1c} can be used. The further a given flight's corrected fan speed deviates from the N_{1c} the HDMR was trained on, the larger the inaccuracy. Therefore, in order to allow users at KLM ES sufficient flexibility in what on-wing datasets to use, the engine fan speed will be unconstrained and sampled between the range of N_1 speeds previously discussed in Section 6.2.

One of the other recommendations was to explore other parameters available in on-wing measurements that could potentially lead to more accurate results. Since the inner workings of a turbofan engine involve highly complex and non-linear phenomena, if more system-related parameters were to be incorporated and implemented then the resulting surrogate model would increase in complexity but potentially be more capable of reproducing system behaviour. Therefore, additional parameters from secondary systems on-board the GEnx-1B were parameterized and are implemented into an updated GSP model.

- **PTO:** Power extracted from the N_2 shaft can be directly passed through the GSP API using the "Load Control" block. Load Control allows the user to define either a Torque load (in Nm) or a Power load (in kW) on any defined shaft. In this case, a power load is passed to the high speed shaft.
- **Variable Geometry:** As outlined in earlier chapters and discussed in Chapter 8, systems involving variable geometry can have a significant impact on steady-state and transient engine performance. In this case, the actuator position defining the operating condition of the Variable Stator Vanes, located in the HPC, were analyzed and parameterized in order to be implemented as an additional parameter to the HDMRs.
- **Turbine Active Clearance Control:** Due to the extreme conditions within a turbofan's turbine sections, clearance between components and outer casing must be regulated to maintain longevity of the engine. The GEnx-1B's active clearance control system was defined and parameterized in terms of the cooling flow valve positions to be implemented to the HDMRs.

The results of such were discussed in Chapter 8. The SPPs derived from that exercise were embedded within the GSP model, allowing for the SPP settings to be passed as a direct input to the GSP API. This allows for smoother interfacing with the ParallelGSP and pyHDMR pacakges.

9.3. GSP Data Generation

As part of the HDMR generation process, training data is required in order to return an adequate fitting of the coefficients within the basis functions. Since the objective of the HDMRs is to return component-wise degradation as a function of engine parameters, the relationship between each component degradation and its effect on engine performance must be present within the training data for a variety of operating conditions. This is why a GSP model of the GEnx-1B is required; GSP being a GPA-based simulation program allows the user to define input degradation per component and obtain the engine performance from simulations.

9.3.1. Sample & Target Data

In the context of HDMR training, target data represent the intended output of the HDMRs. In this case HDMRs were trained for component condition diagnostics, therefore their outputs are measures of component degradation in terms of $\Delta\eta$ and ΔW_c. In addition to component-wise degradation, engine health can also be quantified based on the change of other parameters. It is common in the MRO industry to quantify engine health based on changes in fuel consumption and EGT margin. Therefore, an additional set of HDMRs is also created to quantify the change in EGT, fuel flow and Thrust Specific Fuel Consumption, labelled ΔEGT, ΔW_f and $\Delta TSFC$ respectively. These changes are based on the comparison between deteriorated simulations and a reference simulation representing an healthy engine.

Alongside target data, sample inputs for the HDMRs are also generated and referred to as sample data. The HDMRs generated in this study are considered to be an inverse model; sensor measurements are used as inputs to compute the respective component deterioration as opposed to GPA-based methods, such as GSP, which takes component deterioration as input to compute engine performance in terms of said measurements. The sampling strategy is discussed in Section 9.3.2. Sample measurements are baseline normalized and corrected. Since engine performance also depends on ambient conditions, they are also included as HDMR inputs, while sample measurements are corrected for ambient conditions. Ambient corrections are applied to pressure, temperature, shaft speed and fuel consumption measurements as shown below.

$$\theta = \frac{T_{0,1}}{T_{ref}} = \frac{T_{0,1}}{288.15} \qquad T_c = \frac{P}{\theta} \qquad W_{f,c} = \frac{W_f}{\delta\sqrt{\theta}}$$

$$\delta = \frac{p_{0,1}}{p_{ref}} = \frac{p_{0,1}}{1.01325} \qquad P_c = \frac{P}{\delta} \qquad TSFC_c = \frac{TSFC}{\delta\sqrt{\theta}}$$

$$N_c = \frac{N}{\sqrt{\theta}}$$

It is generally convention to use the total properties of pressure and temperature at the engine inlet. However, after some investigation and experimenting with the GEnx GSP model, de Bruin concluded that more accurate results can be obtained as long as all measurement parameters shall be corrected using the total properties at the fan face, except for the core shaft speed N_2, which is corrected using T_{t25}.

9.3.2. Sampling Strategy

The design space is defined by the range of input parameters passed through the GSP API and is shown below in Table 9.1 and Table 9.2. Ideally, the entire design space of values would be simulated in order to return an exhaustive and complete simulated dataset, by for example employing a "One Factor At a Time" strategy. However, such a strategy would return a rich training dataset but with a large computational cost, both in terms of simulation time and computational time during HDMR generation. Therefore, the ParallelGSP package employs a Latin Hypercube Sampling (or LHS) strategy as proposed by Williams et al. [82]. This strategy allows the sampling of a large design space in such a way that the most amount of information is sample with the least amount of sampling points. Additionally, de Bruin employed an additional strategy referred to as band sampling, in order for the LHS strategy to return combinations of component degradation values such that an approximate 1:2 ratio in efficiency to flow capacity is respected [12].

The range of degradation for each component was selected based on a combination of knowledge from literature and previous experience at KLM ES and maintenance workscope planning guides. Sampling of the degradation intervals is conducted over five separate simulations for a given operating point; one sampling where degradation of every component is varied and one sampling for every component where their respective degradations are set to zero. An example of this approach is a sample where every component has $\Delta\eta$ and ΔW_c varied over the prescribed range, except for HPC which is set to zero for both $\Delta\eta$ and ΔW_c. This strategy was implemented in order to generate sufficient simulated data where the deterioration effects of each component on others is isolated and identifiable during training [5]. As for the range of GSP input parameters, the range of N1 values was set based on previous verification conducted by de Bruin [5] regarding the GEnx-1B GSP model, which was concluded to be the most accurate at N1 speeds between 90% and 95%. The range set for ambient conditions is based on the common conditions observed at take-off from on-wing measurements.

Table 9.1: Sample bounds of GSP input parameters

Parameter	Bounds
N_1	$[90, 95]$
P_{sa}	$[0.6, 1.1]$
T_{sa}	$[250, 300]$
$Mach$	$[0.15, 0.6]$
PTO	$[0.0, 500]$
VSV_x	$[78.0, 84.0]$
$HPTACC_x$	$[9.0, 95.0]$
$LPTACC_x$	$[43.0, 47.0]$

Table 9.2: Sample bounds of GSP degradation parameters

Parameter	$\Delta\eta$	ΔW_c
Fan+LPC	$[0.0, -1.5]$	$[0.0, -3.0]$
HPC	$[0.0, -2.5]$	$[0.0, -5.0]$
HPT	$[0.0, -4.0]$	$[0.0, 8.0]$
LPT	$[0.0, -1.5]$	$[0.0, 3.0]$

Following data generation with GSP, the resulting dataset has to be post-processed. It often occurs that certain combination of input parameters from the initial Latin Hypercube sampling may lead GSP to not converge. Therefore, these and other potential outliers must be filtered out.

9.3.3. Simulations using ParallelGSP

In order to guarantee a high level of accuracy form the HDMR training, sufficient sample and target data must be generated from GSP. The target data for all HDMRs is the level of component degradation for each components in terms pf $\Delta\eta$ and ΔW_c. GSP can be operated from its user interface, but also via an Application Programming Interface (or API) developed by the NLR. The API allows users to load GSP models and execute simulations within a programming environment such as with Python or MATLAB. The main drawback of this API is that only one instance of GSP can be run at a given moment on a single device. While the current GEnx-1B model in GSP is quite stable and fast, this puts a serious limit on computational speed, given the large amount of operating conditions to be sample and simulated. Furthermore, the API is limited to Python environments running 32-bit versions.

Therefore, a Python package was developed at KLM ES by de Bruin [5] that allows multiple instances of GSP to be run within a 64-bit environment, using a series of virtual machines. ParallelGSP operates on a server-client principle, where a server allocates simulation inputs to various clients; in this case one local client and additional virtual machines running their separate instance of GSP. Figure 9.1 provides a visual overview of the functioning behind ParallelGSP.

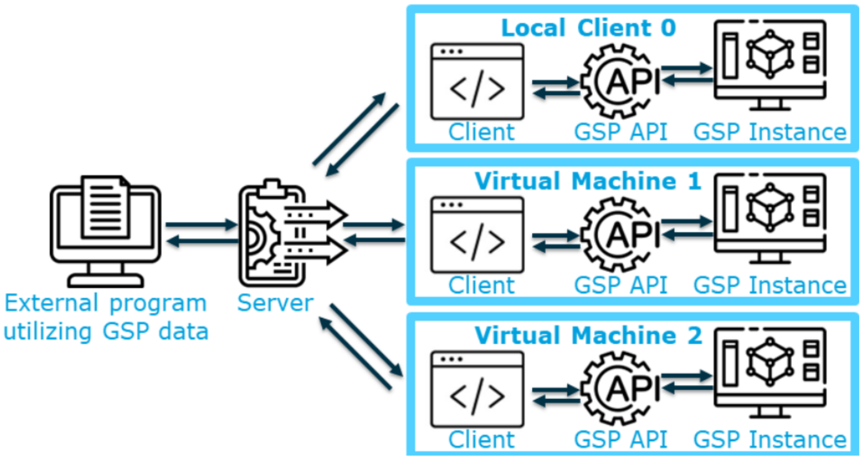


Figure 9.1: Visual representation of the ParallelGSP workflow developed by de Bruin [5]

As mentioned in previous chapters, the GSP model of the GEnx-1B developed by Ramdin will be utilized within this framework. The same engine gas path measurements are used as output, however several other parameters were added as model outputs to be later used for HDMR training. This includes the Secondary Performance Parameters, which are now embedded within the GSP model in order to account for their individual effects on engine performance. Note that the various sampled degradations are summed to the performance modifies due to the SPP settings, resulting in a total degradation term which is then passed to each respective component. Having the SPPs directly made available within the GSP model allows for the SPP settings to be passed as inputs through the GSP API and retrieved as outputs. This allows for simulations to input the various settings sampled between the range expected at take-off and retrieve those inputs as outputs structure the target data for HDMR training.

A summary of the GSP model inputs and outputs is provided below in Table 9.3, along with Figure 9.2 visualizing the flow of inputs and outputs during data generation with the SPPs implemented within the GSP model.

Table 9.3: List of input and output parameters in GSP model for HDMR generation

Inputs	Outputs
N_1	N_1
P_{sa}	P_{sa}
T_{sa}	T_{sa}
$Mach$	$Mach$
PTO	PTO
$\Delta\eta_{FanLPC}$	P_{t2}
$\Delta W_{c,FanLPC}$	T_{t2}
$\Delta\eta_{HPC}$	P_{t25}
$\Delta W_{c,HPC}$	T_{t25}
$\Delta\eta_{HPT}$	P_{s3}
$\Delta W_{c,HPT}$	T_{t3}
$\Delta\eta_{LPT}$	T_{t49}
$\Delta W_{c,LPT}$	N_2
	W_f
	$TSFC$
VSV_x	VSV_x
$HPTACC_x$	$HPTACC_x$
$LPTACC_x$	$LPTACC_x$

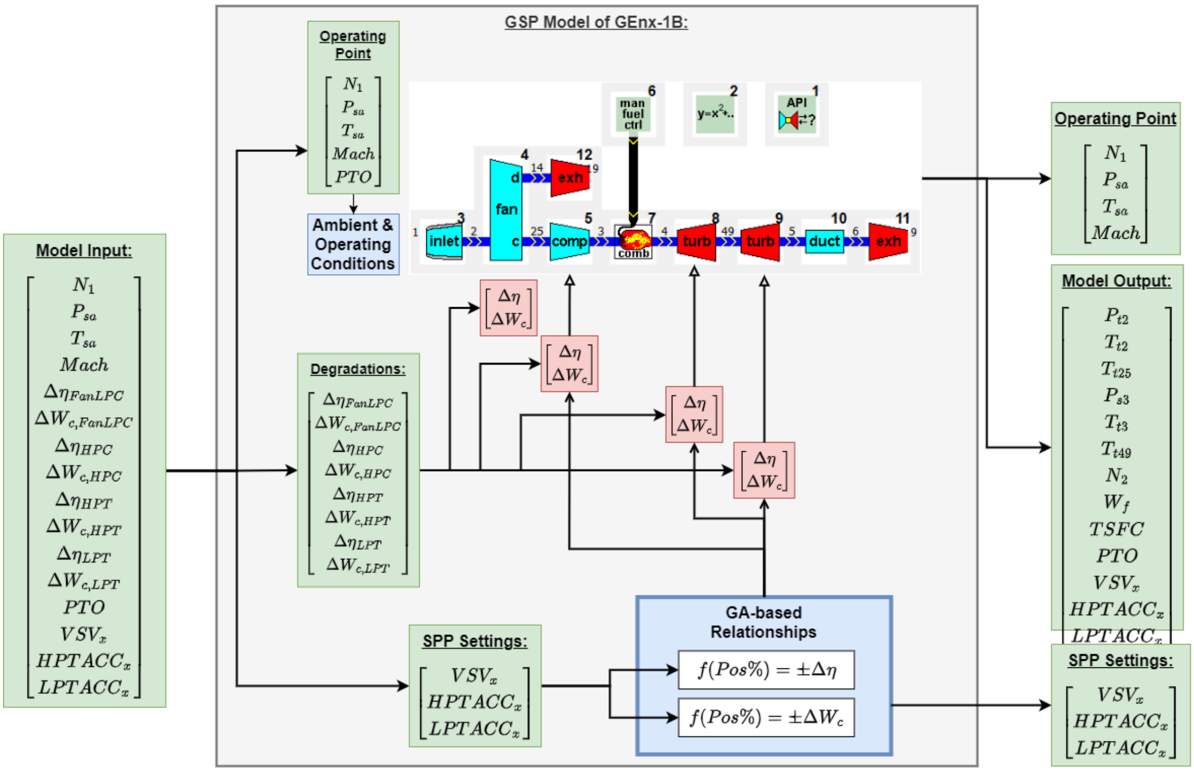


Figure 9.2: Flowchart of GSP model used during data generation with SPPs implemented

9.4. Results from Implementations

In this section, the results obtained from the recommended implementation are presented. During HDMR generation, the Root Mean Squared Error (or RMSE) is calculated in parallel as additional terms within a given HDMR's expansion are added. The RMSE is a widely used measure of accuracy, which indicates the degree of how far does the output of a given HDMR deviates from the target output. RMSEs are calculated for all five HDMRs; for component degradation in terms of $\Delta\eta$ and ΔW_c, along with the three engine health indicators, ΔEGT, ΔW_f and $\Delta TSFC$. As a reminder, each HDMR is generated on a component basis; each component has its component health in terms of $\Delta\eta$ and ΔW_c as conventionally used in GPA, alongside each component's contribution to the engine's ΔEGT, ΔW_f and $\Delta TSFC$.

The results obtained from the original HDMRs generated by de Bruin [5] are shown below as reference values. These RMSEs will be used as a baseline reference to compare the accuracy of the newly modified HDMRs.

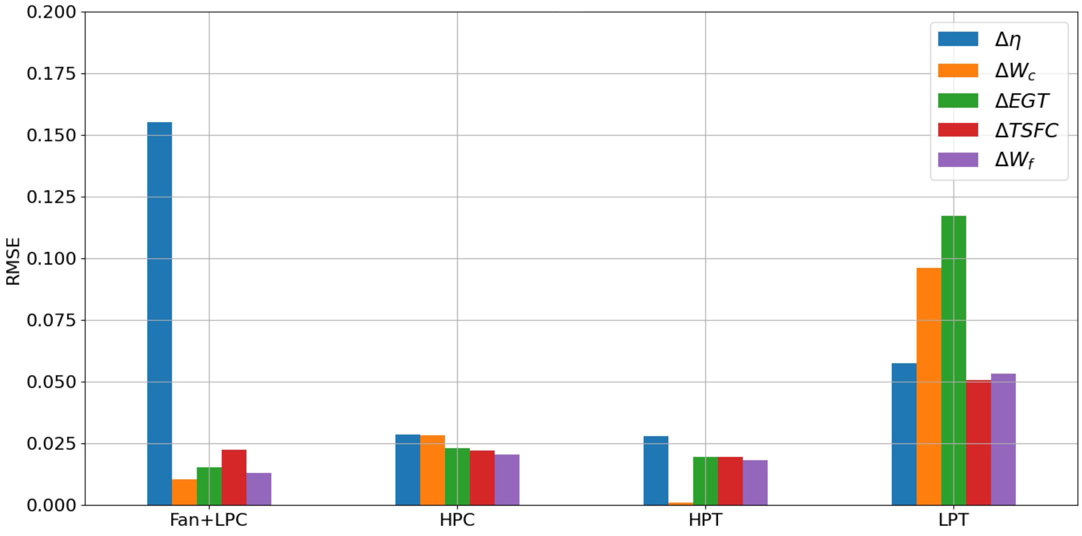


Figure 9.3: Original RMSEs for GEnx-1B of each HDMR type, redrawn from de Bruin [5]

9.4.1. Unconstrained Fan Speed

The first experiment to be conducted was to allow for an unconstrained fan speed. The HDMRs generated in de Bruin's study [5] were trained for a corrected N_1 shaft speed of 91.6%, since it was the GSP model's design point. Constraining it based a corrected fan speed allowed for a relatively small range of operating points to be used form CEOD, thus limiting its applicability to flights whose ambient conditions were within a sufficiently narrow range such that an N_{1c} close to 91.6% was maintained. That being said, to remedy this issue, the input N_1 was now set as a parameter to be sampled using the Latin Hypercube approach. The fan speed was sampled within the range of 90% and 95% in order to simulate a wider range of operating conditions generally seen around the take-off phase.

Below are the results of this experiment. Figure 9.4 shows the resulting RMSEs for the various HDMRs obtained when allowing for a wider range of N_1 conditions. For clarification, Figure 9.5 indicates the same results but relative to the original RMSEs obtained by de Bruin for the GEnx-1B, as shown earlier in Figure 9.3.

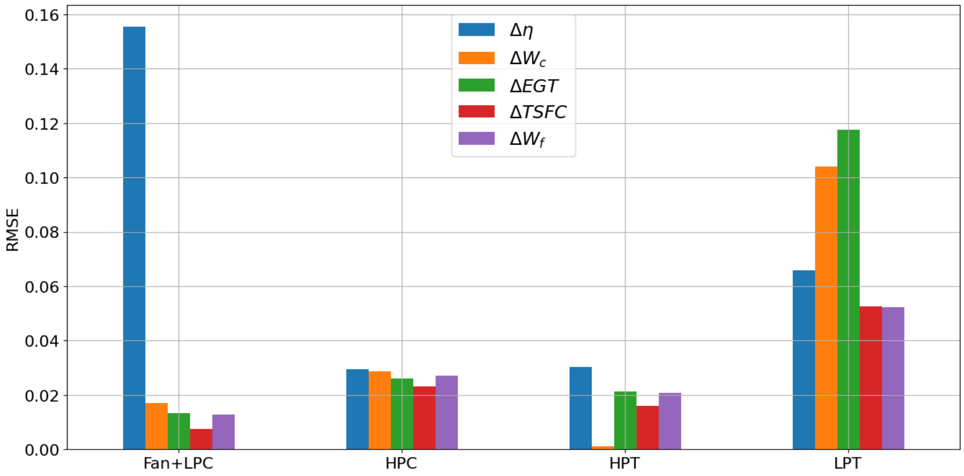


Figure 9.4: RMSEs for GEnx-1B HDMRs with varying N_1

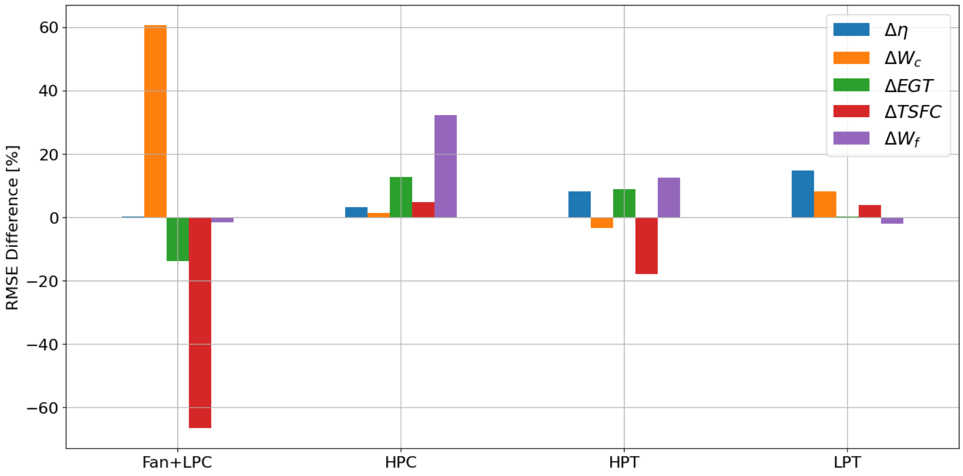


Figure 9.5: Percentage change in RMSEs for GEnx-1B HDMRs with varying N_1 compared to original results

Looking at Figure 9.4, we can see the RMSEs have experienced a general increase. Most notably, the RMSE for corrected flow at the Fan+LPC was found at $1.692 \cdot 10^{-2}$; an increase in RMSE of 60% compared to that obtained in the Baseline GEnx HDMRs. Moreover, a 32% increase was seen for the Conversely, the TSFC for the Fan+LPC was reduced by over 65%, down to $7.53 \cdot 10^{-3}$. Similarly, a decrease of approximately 20% was achieved for the contribution to $\Delta TSFC$ by the HPT. That being said, the majority of HDMRs have experienced a noticeable increase in RMSE compared to the Baseline results, as high as 15%.

In the case of the Fan+LPC component, the act of training HDMRs over a larger range of N_1 and including it as an input parameter to the HDMRs has resulted in a general increase in accuracy. All HDMRs have reported a reduction in RMSE compared to the Baseline set, as seen in Figure 9.5, except for $\Delta\eta$ and ΔW_c which report a marginal 0.3% and a substantial 60% increase respectively. Upon inspection of the generation logs provided by de Bruin, it seems that the majority of the RMSE reduction was reached before second order terms were being added. In this case, substantial reductions in RMSE were recorded when second order terms were being explored. This would then indicate that allowing a larger range of input N_1 speed and including it as an input parameter has increased model complexity, to the point where higher order terms might better reflect system behaviour.

Similar conclusions can be made to the remaining components. Certain HDMRs have benefited in terms of accuracy from a varying fan speed, while others have suffered a considerable penalty. Second order terms were also introduced for the remaining components which contributed to lower RMSEs according to their respective logs. Based on these findings, the implementation of N_1 as performed in this thesis has returned HDMRs with the ability to accept a larger range on on-wing data at the expense of some level of accuracy. While the relative penalties in RMSE suggest a significant loss in accuracy, the absolute RMSEs still remain in terms of magnitude within the range of those obtained by de Bruin. Furthermore, this suggests that, as more parameters are included within basis functions, second order terms start to show noticeable gains in accuracy, suggesting that perhaps third order terms may be worth the computational cost as model complexity increases.

9.4.2. Implementation of PTO

After experimenting with the fan speed as a varying parameter, the first secondary performance parameter to be implemented was the Power Off-Take. This parameter was implemented differently compared to the SPPs in the following subsection. PTO was directly modelled in GSP using a built-in component called the Load Control block. Load Control was used to pass an input power load on the N_2 shaft, simulating the mechanical power extracted to be converted to electrical power for other auxiliary systems. As the PTO was sampled, the HDMRs trained on the target parameters in order including PTO as a parameter in the basis functions. Below are the results from HDMRs generated with both varied N_1 and PTO implemented.

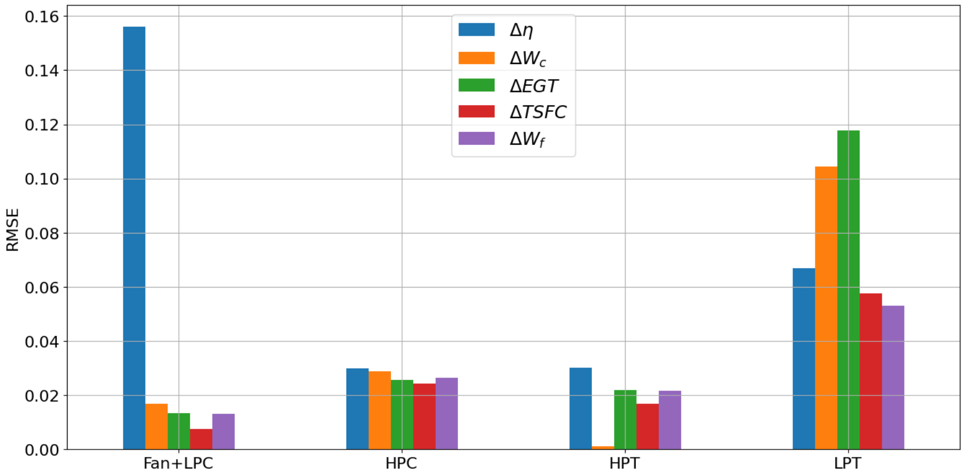


Figure 9.6: RMSEs for GEnx-1B HDMRs with varying N_1 & PTO

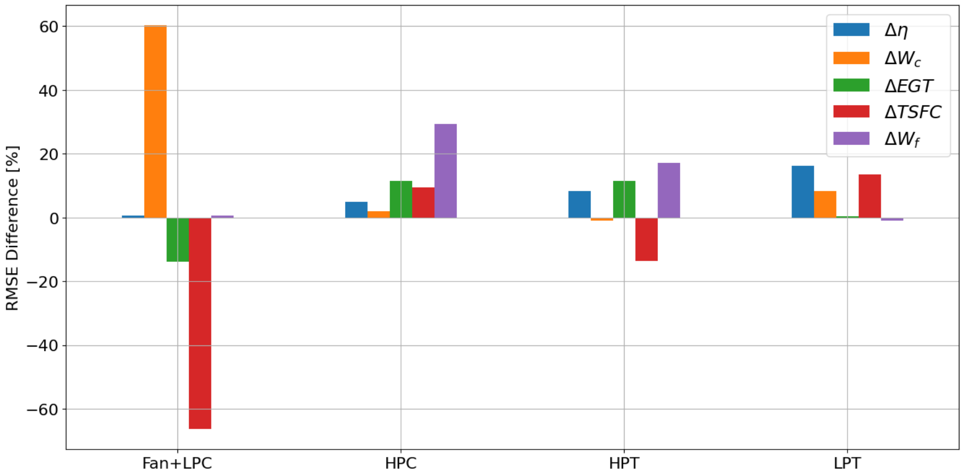


Figure 9.7: Percentage change in RMSEs for GEnx-1B HDMRs with varying N_1 & PTO compared to original results

The results obtained in Figure 9.6 and Figure 9.7 show a similar pattern to that seen when allow for

N_1 to vary. For better insight, Figure 9.8 provides the percentage difference in resulting RMSEs when compared to those obtained in the previous experiment in Figure 9.4.

From above, it is clear that PTO has lead to a further penalty in RMSE across most of the HDMRs, with the exception of Fan+LPC corrected flow, and EGT and TSFC at the HPC. What can be said is that, relative to previous results, the increase in RMSE ranges mostly below 5%, with the exception of TSFC at the LPT which records almost a 10% increase. What is unexpected however is the average increase in RMSE from the HPT after the implementation of PTO. It was initially expected that as the HPT deteriorates, less power can be extracted for a given operating point, thus accounting for the PTO would allow the HDMRs to better associate the level of power extraction from the HPT to its condition. That being said, accounting for this parameter has not lead to a general decrease in RMSE.

Similar to N_1, these results are also a consequence of increase model complexity. As more parameters are introduced, more constraints are imposed on the coefficient fitting process due to more parameters having to be satisfied. That being said, given that the implementation of PTO followed that of N_1, the decrease in accuracy is springily not as large. In reality, the impact of implementing PTO resulted in a decrease in RMSE for certain HDMRs while an increase of no more than 10% relative to results initially obtained for N_1. This is visually presented in Figure 9.8.

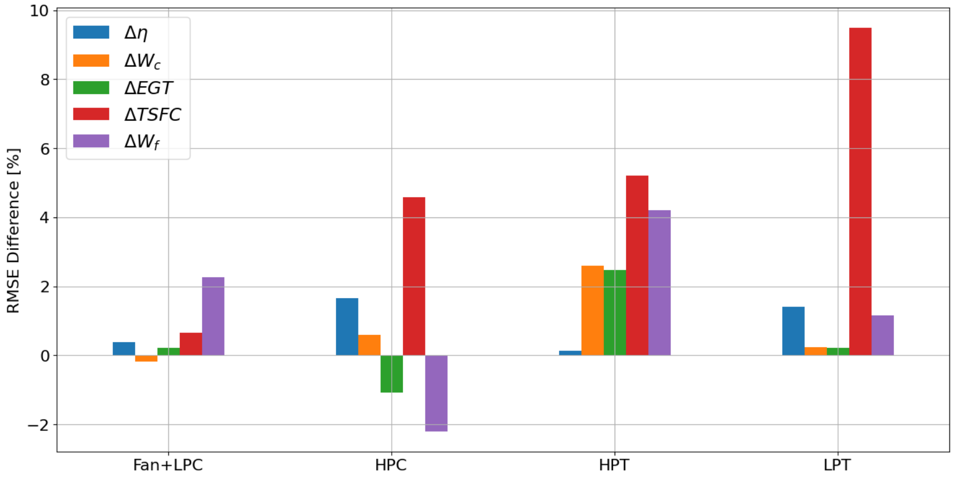


Figure 9.8: Percentage difference between RMSEs in Figure 9.6 and Table D.7

In this case, the addition of PTO has lead to a visible decrease in RMSE for ΔEGT and ΔW_f at the HPC, while others have experienced a minimal increase no larger than 5.5%, with the exception of $\Delta TSFC$ at about 9.5%. Therefore, while not having provided the expected gain in accuracy, PTO has not contributed to a substantial decrease in accuracy when combined with varying N_1. Furthermore, as stated in the previous subsection, the magnitude of absolute RMSEs still remain within range.

9.4.3. Implementation of SPPs

After implementing PTO within the HDMR generation framework the next steps was the implementation of the Secondary Performance Parameters. Out of the four SPPs analyzed in Chapter 8, only PTO was a parameter that could be defined within GSP, allowing for a more straightforward implementation within the generation process. In the case of VSVs and HPT and LPT ACC systems, a parameterization was required to be then implemented within GSP. Once the SPPs were parameterized and embedded within the GSP model, the acutator and valve settings could then be passed as inputs. With this, HDMRs were generated with all the recommendations put forth by de Bruin. The following results were obtained for HDMRs with varying N_1 speed and all four SPPs implemented as input parameters within the HDMR basis functions.

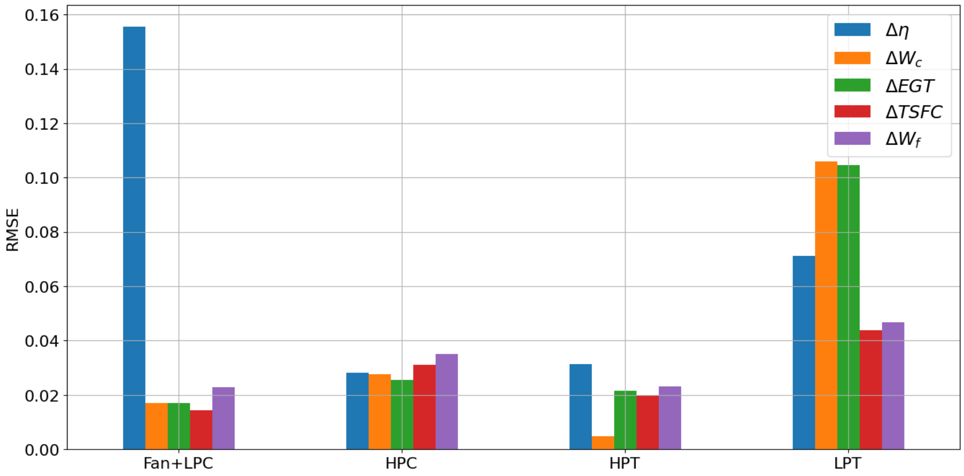


Figure 9.9: RMSEs for GEnx-1B HDMRs with varying N_1 and SPPs implemented

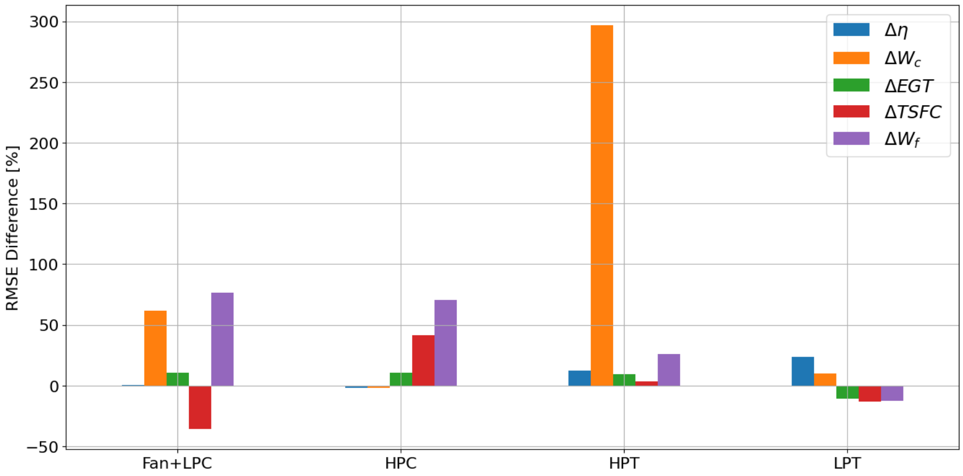


Figure 9.10: Percentage change in RMSEs for GEnx-1B HDMRs with varying N_1 and SPPs, compared to original results

From Figure 9.10, we can see that compared to the original RMSEs obtained for the GEnx-1B HDMRs, a general increase in RMSE was obtained. While many report a marginal increase despite the additional of several parameters and some even reported noticeable decrease, the consensus from these results indicate a general decrease in accuracy. Most notably, the RMSE for corrected flow at the HPT has seen an increase of almost 300%, from an RMSE of $1.19 \cdot 10^{-3}$ to $4.72 \cdot 10^{-3}$. Meanwhile, the HDMRs for the EGT, W_f and TSFC health indicators have reported a average decrease in RMSE of 12% for the LPT. The RMSEs for the fuel flow HDMRs have also increased by about 70% for the Fan+LPC, HPC and 25% at the HPT.

Based on the results obtained above, it is clear that the implementation of the recommendations to the HDMRs have lead to an overall decrease in accuracy, when put in direct numerical comparison with those obtained by de Bruin. The addition of PTO and widened range of values for N_1 already suggest that increasing model complexity can have a positive effect for some metrics but result in a significant penalty in accuracy for others.

An example of this is the fact that the SPPs implemented in this case directly affect the performance of the HPC, HPT and LPT, and is done in terms of the component map modifiers $\Delta\eta$ and ΔW_c. However, in the case of the HPC and LPT, the penalty in accuracy is relatively small; in the case of the HPC they in fact have decreased RMSEs, but the corrected flow for the HPT has most notably suffered in accuracy. Furthermore, the increase in RMSE for the Fan+LPC is an unexpected result, which then further supports that the decrease in accuracy from the updated HDMRs is a result of smearing.

Another source of uncertainty that may have been introduced to the HDMRs is the SPP relations themselves, previously outlined and derived in Chapter 8. While a general decrease in model deviation was reported with no significant penalty in accuracy, the increased spread in model outputs and correlation with the respective SPP settings suggest that the GA-base optimization struggled to converge to relations that eliminated certain dependencies. One notable example is the behaviour of the HPT clearance, which varies significantly at take-off due to other factors. Therefore, this sort of uncertainty has propagated down to the HDMR representing the HPT's corrected flow and lead to an almost 300% increase in RMSE.

Additionally, when speaking of increased model complexity, the increase in RMSEs also suggests underfitting of the target data. While increasing the size of the training dataset often leads to better results, after several iterations of the HDMR model generation, increasing the size of the dataset did not lead to a substantial decrease in RMSE and incurs a larger cost in terms of computation time given the growing size of the sampling for every parameter added. Furthermore, under-fitting can also be the result of model definition. The HDMRs were limited to second order terms, thereby limiting its scope to capturing second-order effects within the data. As suggested previously, given that large reductions in RMSEs were recorded for second-order terms, it suggests that if given enough time to explore third order terms, it is possible that model accuracy could improve.

9.5. On-Wing Condition Trending

Once HDMRs have been generated, the next step to assess their performance is to apply them to on-wing data to conduct component-wise diagnostics. Results are provided in the form of time series plots, indicating the various component condition metrics provided by each HDMR. Numerical verification and validation of the HDMRs is in reality a complicated task. For that, users would require exact data on component degradation, which unfortunately is not a direct engine measurement. While EGT, W_f and TSFC are in some way obtainable form measurements, the change in these indicators with respect to the contribution of the degradation of each individual component is not a straightforward metric to assess. This is one of the main advantages of the HDMRs.

While numerical verification and validation is not currently possible, another form of verification can be conducted when inspecting the trend in the time-series. It is worth noting that since every engine is different from one another, their conditions may not be directly comparable. Since the exact condition of the engine used during the development of the GSP model of the GEnx-1B is not known exactly, this difference must be corrected. Before trending, the first five flights are averaged and used as a constant initial offset to the HDMR output. This baseline offset acts a calibration factor to mitigate engine-to-engine differences [5]. Using on-wing data from the same reference engine mentioned earlier in Section 8.1, trends in component health can be observed and related to specific maintenance events. Since these maintenance events are known, their expected effect are also understood. Two categories of maintenance events can occur; water-washes and shop visits.

As explained in Section 3.1, two categories of component deterioration are present; recoverable and non-recoverable. One of the recoverable forms of deterioration is fouling, which can be recovered through water-washes. KLM conducts periodic water-washes of all engines, during shop visits and after a number of cycles,in order to reduce the amount of fouling accumulated during operation. The process involves spraying water, often mixed with detergents, into a running engine to wash off and remove any potential deposits of particulates on component surfaces. Water-washes will mostly affect the Fan, LPC and HPC, as fouling in the turbine sections is fairly limited, thus they should have a noticeable effect on the Fan+LPC and HPC trends.

Alongside water-washes, other maintenance events can appear during trending. Any clear drops in component health would be due to sudden component failures and be visible as outliers. Visible trends could be present in the case of gradual deterioration in the HPT. However, in the case of the reference engine selected, there are two major maintenance events present in the data. This particular GEnx-1B was sent for two shop visits five year apart due to significant HPT blade damage. Therefore, these should be visible as two distinct events, which will be referred to as Event 1 and Event 2 respectively.

The investigations following those shop visits revealed that Event 1 was attributed to contamination of the cooling channels in the first HPT blade stage, likely present since production. A lack of appropriate cooling resulted in a sudden blade burn-down. Given the rapid nature of this component failure, any trend related to it would appear in a narrow time span before the respective shop visit. As for Event 2, investigations concluded that failure of the clips ensuring proper attachment of the HPT's second stage shroud led to significant drooping off the shaft axis and subsequently causing an excessive amount of blade tip abrasion with the surrounding casing. This failure was considered to happen progressively, therefore a trend of progressive HPT component health deterioration should be visible. Following these events and their associated shop visits, data is also present, with only six flight cycles recorded post shop visits. Therefore, the effect of said shop visits is also visible in the trends. Note that past these repairs, further trending is not possible as no data is present following those events.

Each of the following subsections include the results obtained from the HDMRs for a specific component. In total, five figures are generated in the form of time-series plots over a certain period of the engine's operational lifetime. With this, the trend in deterioration in terms of the various health indicators for each HDMR is shown on their respective vertical axes. Note that the horizontal axes represent time but exact dates are withheld due to confidentiality. Furthermore, for the sake of this analysis, the trends generated from the Baseline HDMRs by de Bruin [5] are referenced in order to compare the effects of the updated HDMRs with the secondary performance parameters implemented. The thin blue vertical strips represent water-washes while the large blue shaded areas represent the time interval of a shop visit, where in this case the two shop visits are those prompted by Events 1 and 2.

9.5.1. Fan+LPC

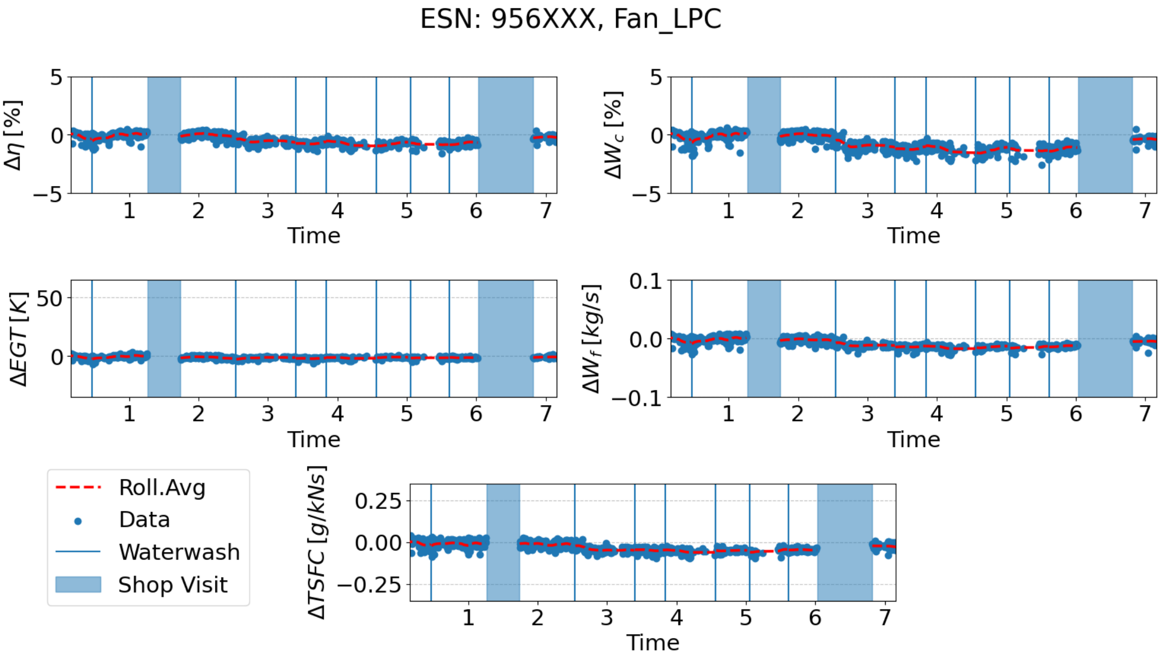


Figure 9.11: HDMR trending using on-wing data from ESN956XXX for the Fan+LPC

Figure 9.11 is composed of five figures, each showing the result of each HDMR for the Fan+LPC component. As can be seen, over this operational period, no major changes in component condition in terms of quantified degradation or the health indicators is present. No significant effect is observed following the waterwashes, which is consistent with the condition of the reference engine and consistent with de Bruin's initial results. Maintenance reports and borescope inspections of the Fan and LPC components show very little to no deposition of particulates were seen on blade surfaces, therefore it is expected that water-washes have a marginal effect on the general condition of the Fan+LPC component.

That being said, there is a small but noticeable change in efficiency and corrected after the first water-wash. A decreasing trend in both of those metrics is visible at the very beginning of the time series, followed by an increase after the first water-wash. The initially decreasing trend can be explained by two possible sources. The first is deterioration present in the engine prior to when the on-wing data was made available. Note that on-wing data is only made available at KLM after a certain number of cycles, therefore depending on operating conditions, fouling could already be present. That being said, the effect of the first waster wash results in an improvement in engine condition; efficiency increases along with corrected flow, which is the expected behaviour. The second source that could cause a progressively decreasing trend in component health is the accumulation of non-recoverable deterioration. Non-recoverable deterioration can be either abrasion with large particles or corrosion due to contact of other substances in the atmosphere depending on ambient conditions.

While it does not directly interact with the Fan+LPC component, the inaccuracy introduced by the additional parameters may have lead to smearing of the effect of Event 2 to other components such as the Fan+LPC. This effect can be seen following the second shop visit, which shows a slight increase in corrected flow and efficiency of the Fan+LPC. This effect is also present for ΔW_f and $\Delta TSFC$. That being said, this instance can be explained by the fact that during or following a shop visit, the engine does undergo a water-wash, thus the improvement in Fan+LPC condition following the shop visit for Event 2 can be explained. Overall, the resulting trends for the Fan+LPC demonstrates that the updated HDMRs, while reporting increases in RMSE, can still produce satisfactory condition trends with minimal smearing.

9.5.2. HPC

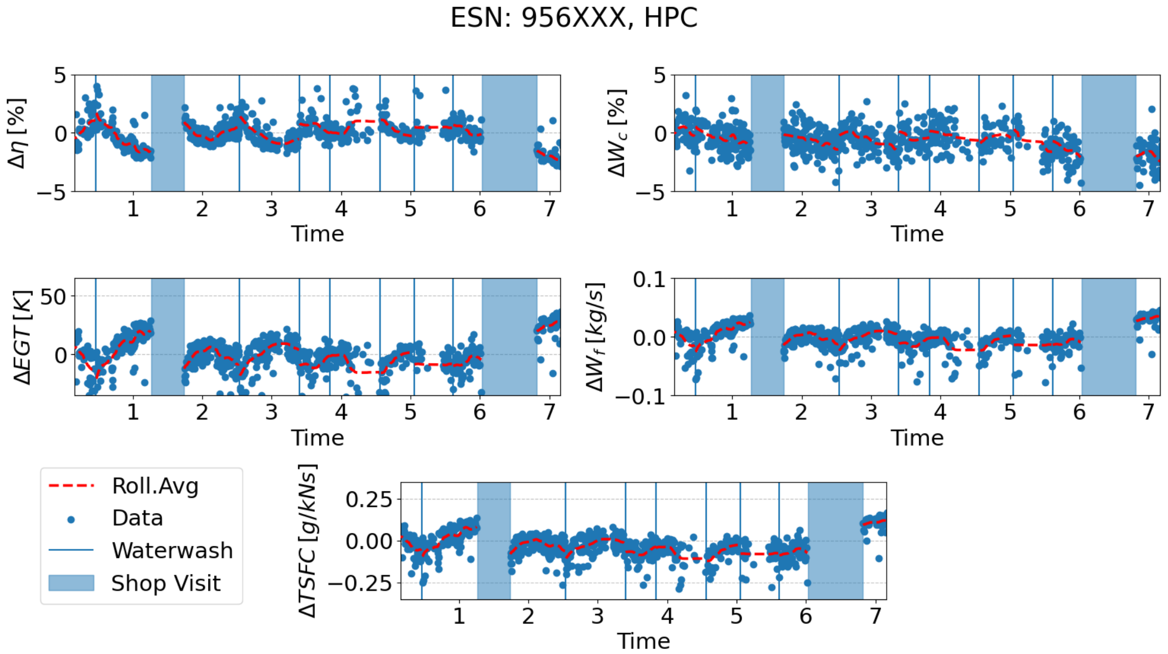


Figure 9.12: HDMR trending using on-wing data from ESN956XXX for the HPC

In the case of the HPC, the effect of the water-washes is expected to be more apparent. As can be seen in Figure 9.12, this is in fact the case. Fouling at the HPC would result in more turbulent flow forming on the blades, while effectively decreasing the total area available for air to flow through, thus resulting in decreased efficiency and corrected flow respectively. The updated HDMRs still reflect this behaviour, similar to the trending produced by de Bruin, yet on average their health restorative effect progressively becomes smaller after a few water-washes. This is likely due to the existence of non-recoverable deterioration accumulating over cycles. Additionally, there seems to be an increase in engine health across all HDMRs preceding certain water-washes. This effect was relatively minimal in de Bruin's results but now more prevalent and indicates a sort of seasonality. This is possibly due to the implementation of the VSV settings within the HDMRs. As the HPC performance slowly degrades after some cycles following a water-wash, this may be due to the variation in VSV settings at take-off over time in order to compensate for this performance decrease, leading up to the water-wash.

As for the effects of Events 1 and 2, a clear increase in component health is seen following the first shop visit. HPC health experiences an increase following Event 1 in both efficiency and change in EGT, fuel flow and TSFC. It should be noted however that the shop visit was focused on restoring the HPT, therefore no significant effect should be expected for the HPC. This then could be interpreted as smearing from the HPT to the HPC, however when inspecting the corrected flow of the HPC, no major difference is seen in trending following the first shop-vist. Therefore, the effect of Event 1 manifesting in HPC performance must be due to the water-wash conducted during the shop visit. Furthermore, since the HPT is connected to the HPC via the high-speed shaft, it is also understandable that deterioration of the HPT will lead to a performance decrease in the HPC.

Furthermore, following Event 2, a sudden decrease in HPC health is visible. This can be explained by the fact that this reference engine was transferred from a Boeing 787-9 to a 787-10 following its shop-visit since both aircraft operate the GEnx-1B. The 787-10 however is rated at a higher thrust level compared to the 787-9. Since the GSP model developed by Ramdin was based on data measured from a 787-9, the difference in engine rating from the measurements was interpreted as an engine operating at a lower performance level compared to its reference level. Given that the HDMRs were developed for component degradation, this change in engine performance must have been interpreted as further engine deterioration.

9.5.3. HPT

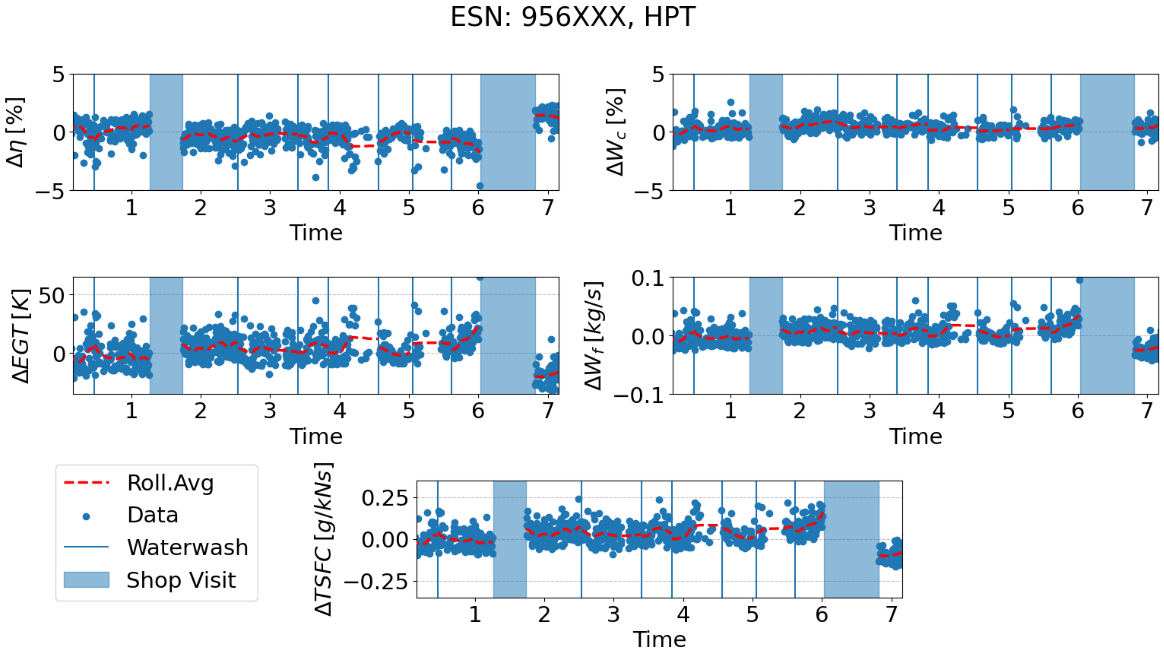


Figure 9.13: HDMR trending using on-wing data from ESN956XXX for the HPT

In the case of the HPT, the majority of trends remain relatively monotonous, however the interval between Event 1 and Event 2 show a small level of seasonality between water-washes for the HDMRs plotting $\Delta\eta$ and ΔW_c. This pattern emerges after the first shop visit where engine conditions remains relatively steady, until the first water wash where the engine progressively deteriorates afterwards. The subsequent water-washes seem to slow down this degradation, until we arrive at the second shop visit. This trend is attributed to Event 2, which was characterized by progressive deterioration, until the HPT blades suffered significant abrasion. The increase in tip leakage would result in decreased component efficiency due to tip vortices forming and entropy increase, while the corrected flow progressively increases due to more mass flow being allowed to flow around the area because of larger clearance between the blade tips and the surrounding casing.

This then implies that the HDMRs were able to identify the progressive deterioration leading up to Event 2. This progressive deterioration of the HPT was somewhat visible in de Bruin's results but is now more pronounced; a result of the implementation of the HPTACC which would be more active in order to counteract the increased EGT of a slowly deteriorating engine. As for the seasonality seen between certain water-washes, this is due to the implementation of the HPT Active Clearance Control system within the HDMRs. The water-washes are not expected to directly affect the HPT, however as the HPC fouls over time, more fuel is spent to maintain a sufficient pressure ratio, leading to an increase in EGT and thus more demand from the ACC system.

Furthermore, a clear trend of deterioration is visible in flights leading up to Event 2. In the case of $\Delta\eta$, a clear decrease in efficiency is visible in flights preceding the shop visit for Event 2. However, ΔW_c shows minimal changes by comparison. That being said, the trend leading up to Event 2 is clear for the HDMRs returning the health indicators. A steep rise in ΔEGT, ΔW_f and $\Delta TSFC$ is seen prior to Event 2, indicating that the HDMRs were able to identify the HPT condition leading up to the failure of Event 2.

9.5.4. LPT

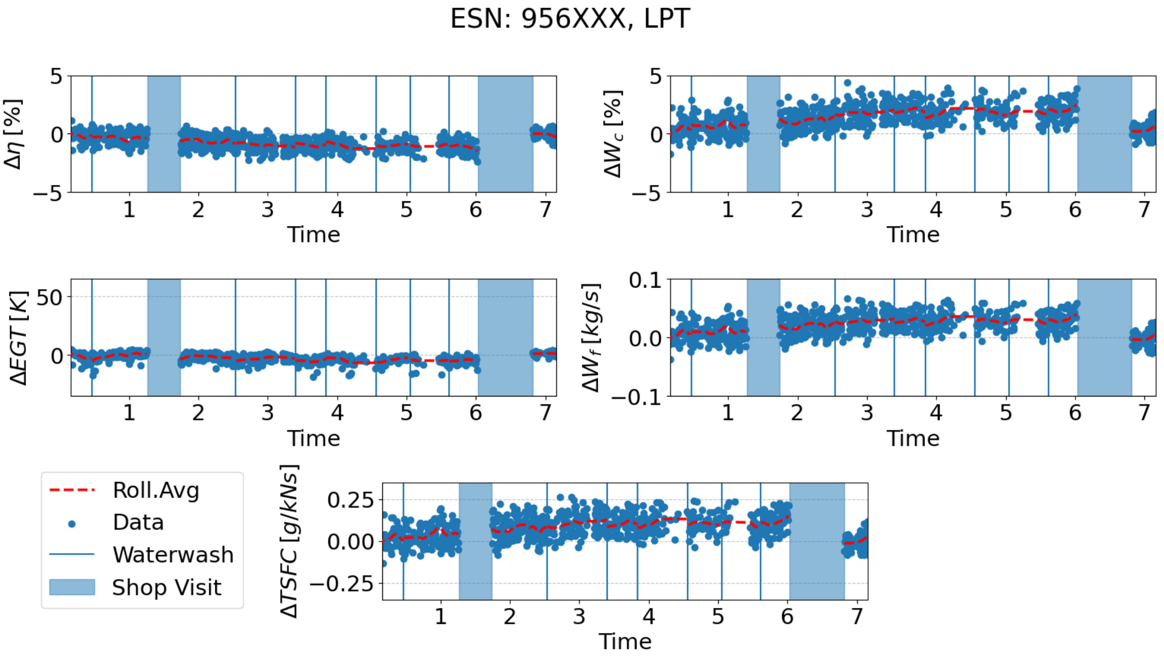


Figure 9.14: HDMR trending using on-wing data from ESN956XXX for the LPT

Similar to the Fan+LPC component, the LPT trend is relatively monotonous. No major changes in component condition is observable following water-washes. This is expected given that the restorative effects of water-washes is mostly focused on the Fan+LPC and HPC components upstream. As the engine is running, very little fouling is removed by the water flow by the time it reaches the LPT.

Contrary to previous results, no observable seasonality is seen for the LPT. The earlier seasonality observed for the HPC and HPT were interpreted as indications of the effect of the secondary performance parameters. More demand would be placed on the VSVs and ACC system as the engine starts to deteriorate again. However, this effect is quite muted for the case of the LPT. This can be explained the fact the relations for the LPT did not result in large changes in $\Delta\eta$ and ΔW_c over the interval of LPTACC settings seen at take-off.

That being said, the resulting HDMRs return a trend showing clear sings of progressive deterioration of the LPT. The restorative effect of the first shop visit in the case of the LPT is not as large compared to the HPT. While the LPT did suffer from object damage due to Event 1 at the HPT, that effect would not be visible or treated as an outlier. Furthermore, since Events 1 and 2 primarily involved servicing of the HPT, no major maintenance was conducted for the LPT, thus no major changes in component health are expected to be seen. Based on Figure 9.14, it seems that the LPT experiences a relatively monotonous rate of deterioration over this time period, with the exception of certain outliers. This constant rate of deterioration is expected over time, given that no major maintenance tasks are conducted for the LPT.

It is worth noting that the rate of deterioration increases mostly after the first shop-visit and levels off as it approaches Event 2. This can be observed in results for corrected flow, fuel flow and TSFC, which all experience an increase then remain at an certain level. This would imply some form of non-recoverable deterioration occurring over this period. Non-recoverable deterioration characterized by a steady increase in corrected flow would imply abrasion of the LPT blades. A reduction in LPT performance would then translate to less energy extracted to continue driving the LPC, further increasing the demand on the HPC and requiring more fuel to burn. Furthermore, following Event 2, this flat degradation level is then reset after the shop visit. This is unexpected given that it is assumed that no major maintenance tasks were conducted on the LPT. However it is worth mentioning that no information on reference maintenance events concerning the LPT are present at the time of writing, therefore no conclusive remarks can be made for now. That being said, the sudden restoration of LPT health after the second shop visit can be attributed to the change in thrust rating previously mentioned.

9.6. Discussion

The HDMR-based surrogate model developed at KLM ES was updated based on recommendations put forth after its development. Those recommendations were explored and a solution for their implementation was presented in previous chapters. The HDMRs were then re-generated with an updated set of input parameters to better make use of measurements provided provided in on-wing data. The resulting accuracy and performance of the HDMRs was analyzed both in terms of RMSE calculated during training and their ability to conduct trending of component condition using the on-wing data available.

The updated HDMRs returned mixed results. On one hand, the addition secondary performance parameters and the extension of the allowable range of operating points as inputs based on fan speed has lead to a noticeable increase in certain cases and a decrease in others. Initially, these recommendations were speculated to lead to a consistent increase in HDMR accuracy but results indicate otherwise. Given that a turbofan is a highly complex system and that performance of individual components is often difficult to isolate and analyze without affect others both upstream and downstream, it is not straightforward beforehand to identify where potential gains in accuracy will manifest. Following certain remarks made, it is then possible that further gains in accuracy can be made if the HDMR generation was allowed to explore third order terms. The justification for limiting the models to second order terms was also suggested from literature as a trade-off between accuracy and computational cost. Thus as a future recommendation, it would be wise to explore other means of expanding the HDMR's abilities by capturing more complex relationships between parameters via higher order terms.

That being said, the condition trending conducted in Section 9.5 revealed that, despite results obtained during generation, the HDMRs still show a satisfactory ability to identify trends in engine component conditions based on the expected behaviour based on information from known maintenance events. The updated HDMRs were able to detect the sudden trends in deterioration prior to two known events leading to shop visits. Additionally, the effect of water-washes were present and appropriately quantified for the relevant components. While generation logs imply a penalty in accuracy for the additional parameters, the trending ability of the HDMRs does not appear to be severely impacted. That being said, the presence of smearing was noted in certain cases and was not minimized by the increase complexity of the models.

Additionally, one of the major differences in this thesis compared to de Bruin's work was the use of Snapshot data for condition trending instead of CEOD. Condition trending conducted by de Bruin [5] made use of CEOD, which provides the rate of change of the core shaft ΔN_2 as a measurement. The advantage is that it allows for filtering of the CEOD to only include operating points that are considered to be at or near steady-state operating condition; with a ΔN_2 near zero or below a certain value. In this thesis however, snapshot data. The advantage is that snapshot data is available early on in an engine's operational lifetime. However, snapshot essentially includes three recordings of all measured parameters; one at take-off, climb and cruise. Since no ΔN_2 is available to confirm whether or not the engine was operating confidently at steady-state, it is assumed that the GEnx-1B's EMU acquires these snapshots at operating condition deemed close enough to steady-state. That being said, it is still possible that the core shaft is accelerating to some degree, resulting in transient effects seen in trending that explain the presence of outliers and more sudden changes in results from cycle to cycle.

Furthermore, with regards to the secondary performance parameters, their implementation within the HDMR framework proved to be successful. However, the resulting impact on HDMR performance was not as substantial as expected. Given that this analysis was limited to take-off for the sake of higher accuracy from the GSP model, which is at the core of the HDMR framework, the decision to focus on that practical phase of the flight has revealed certain limitations with regards to the available measurements. A notable example is the large spread of calculated HPT clearance according the GEnx-1B's EEC during take-off. Previous work on SPPs done by Otten did not suffer from this issue given that the focus of that study was for operating points at cruise.

10

Condition Prognostics Model

Following the results from the updated HDMRs, the following chapter focuses on the experiment involving the combination of the HDMR-based surrogate model with an LSTM-based prognostics model. Section 10.1 summarizes the LSTM model and how it interacts with the established surrogate model. Section 10.3 outlines the steps taken in the development of the LSTM model while Section 10.4 assess the resulting model's capability in component condition predictions. A discussion and remarks based on the obtained results is provided in Section 10.5.

10.1. Condition Prognostics

Referring back to Chapter 5, the ECBM paradigm also allows for the ability to conduct both diagnostics and prognostics on component and system condition. The HDMR-based surrogate model developed in this thesis has demonstrated it capability for engine component-wise diagnostics. Therefore, the next natural step is to explore its capability to conduct prognostics. The advantage of the HDMR-based surrogate model is that it allows for direct calculation of component condition in terms of map modifiers conventionally used in GPA-based approaches and applications, along with additional health indicators conventionally used in MRO such as ΔEGT, ΔW_f and $\Delta TSFC$. Therefore, as described in Section 6.4, the prognostics model proposed in thesis will also be developed using the same parameters, with the outputs of the HDMR-based surrogate model serving as training data. This is explained further in Section 10.1.2. Following the discussion in Section 6.4, a machine-learning model was selected given the context of the problem and the nature of the data at hand. Similar to previous chapter, a brief summary of the type of prognostics model used is given below.

10.1.1. Summary of LSTMs

Long Short-Term Memory (or LSTM) cells are a variant of Recurrent Neural Networks (or RNNs) introduced by Hochreiter and Schmidhuber [83], designed to more effectively capture long-term dependencies in a given sequential dataset. The LSTM was in fact presented as a solution to a common issue seen in RNN cells, namely the vanishing gradient problem [83]. Compared to general neural networks, LSTMs contain feedback connections in contrast traditional feed-forward networks. Below in Figure 10.1 is a visual representation of an LSTM cell.

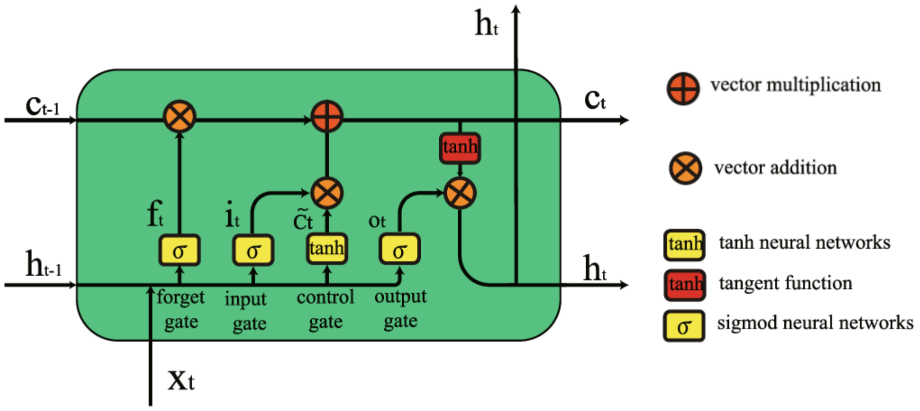


Figure 10.1: Diagram of LSTM cell [84]

LSTM cells are composed of a series of gates which manage how information is stored and flows in and out of the network. These are named the "forget" gate, "input" gate and "output" gate. [85]. The input gate handles what information is stored in the cell state while the purpose of the forget gate is to selectively neglect the information stored in the cell state and the output gate decides what value comes out [84]. The LSTM gates use a sigmoid as activation function, which returns either a 1, allowing everything to pass, or 0 which blocks all information. Mathematically, the process within an LSTM can be explained with the following equations [84]:

$$\begin{aligned} f_t &= \sigma\left(w_f \cdot [h_{t-1}, x_t] + b_f\right) & \tilde{C}_t &= \tanh w_C \cdot [h_{t-1}, x_t] + b_C \\ i_t &= \sigma\left(w_i \cdot [h_{t-1}, x_t] + b_i\right) & C_t &= f_t * C_{t-1} + i_t \cdot \tilde{C}_t \\ o_t &= \sigma\left(w_o \cdot [h_{t-1}, x_t] + b_o\right) & h_t &= \sigma_t * \tanh C_t \end{aligned}$$

On the left, f_t, i_t and σ_t represent the forget, input and output gates respectively. The sigmoid activation function is shown as σ, while w_x and b_x represents the weight of gate x and bias respectively for each LSTM cell. Furthermore, x_t represents the current input vector for time t while h_{t-1} is the output from the previous time step $t-1$. As for the right side, C_{t-1} and C_t denote the previous and new cell state (or memory) respectively for time steps $t-1$ and t. Furthermore, $\tilde{C}_t$ represents to the output of the hyperbolic tangent function, resulting in the new candidate state of the cell at time t. The symbol $*$ denotes element-wise multiplication of the relevant vectors [84].

10.1.2. Interaction with HDMRs

The goal of the prognostics model is to make predictions on component condition based on the diagnostics conducted by the surrogate model. As explained in previous chapters, the HDMR-based surrogate model outputs the condition of engine components based on on-wing data for a specific flight. More specifically, it estimates component condition based on engine performance on a flight-by-flight basis without considering component condition from previous flights. On the other hand, a prognostics model will have to be trained on some form of measurement of component condition in order to make predictions of future condition based on relationships it captures in the data. That being said, the combine modelling approach will involve making use of the outputs of the HDMR-based surrogate model in order to train the LSTM network, as shown in Figure 10.2. The idea for this approach was inspired by the combination of both an ANN and HDMR as surrogate models proposed by Li et al., where the ANN was used to generate samples from a combustion kinetic model to be later used by an RS-HDMR surrogate model for global sensitivity analysis [86]. The difference in the approach proposed here is that the samples are generated by the HDMR in order to train the LSTM-based neural network.

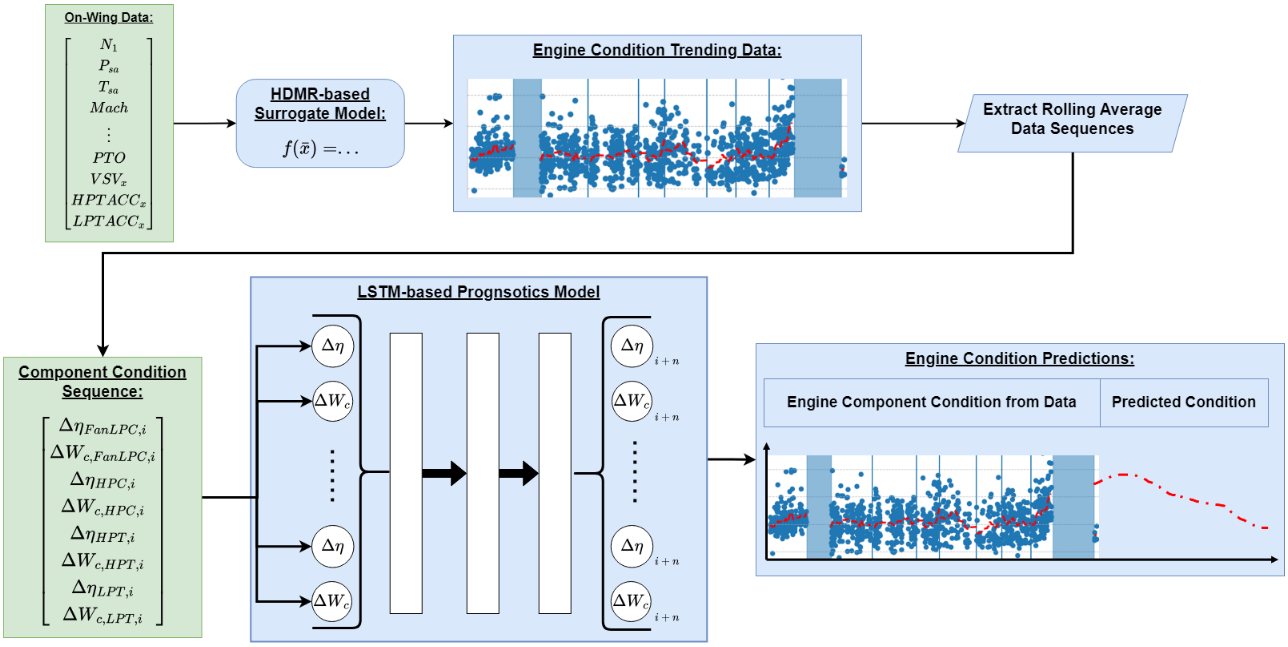


Figure 10.2: Flowchart of model interaction for LSTM development

10.2. Data Pre-Processing

When feeding flight data to the HDMRs, the trend in each individual component's condition in terms of $\Delta\eta$ and ΔW_c can be visualized, as seen in Section 9.5. Given that engine condition is quantified based on the working principle of GPA and measured in $\Delta\eta$ and ΔW_c and that the HDMRs return engine condition in those terms per component, the output of the HDMRs will serve as training data. Since component condition varies over time after every flight, this can be treated as a time-series forecasting problem; ideally suited for Long Short Term Memory cells.

As done in previous chapters, the same Snapshot dataset of engine ESN956XXX will be used to generate the training data. This dataset was also used to assess the HDMR surrogate model's condition trending capabilities. The following subsections outline further steps in data pre-processing required before moving onto model training and development.

10.2.1. Trajectories

The goal of the LSTM model is to capture the trend in component degradation over flight cycles. With the HDMR surrogate model, component condition data can now be generated in order to train any potential prognostics model. One critical aspect of machine learning models is the quality of the training data available. To reiterate, snapshot data from ESN956XXX is used to compute component condition, however there are discrete intervals in this engine's operational lifetime where maintenance events take place. The majority of research done in the field of machine learning for engine component condition prognostics, an example being Chao et al. [87], have made use of engine performance datasets generated by CMAPSS [88]. CMAPSS allows users to simulate engine data for a given failure case in order to train ML models to detect patterns in engine deterioration.

Therefore, given the dataset available for this thesis, the condition history generated for ESN956XXX will have to be segmented into intervals between shop visits. The reason for that is that shop visits lead to a large discontinuity in the time series, along with a sudden increase in component health. The goal of the LSTM model is to capture the trend in deterioration over flight cycles, therefore it must be trained to also capture the effect of significant maintenance events. Furthermore, the discontinuity in the sequential data caused by the presence of shop visits can further hamper the training process. Below is a visual representation of the EGTHDM of ESN956XXX along with each maintenance interval labelled a specify trajectory number.

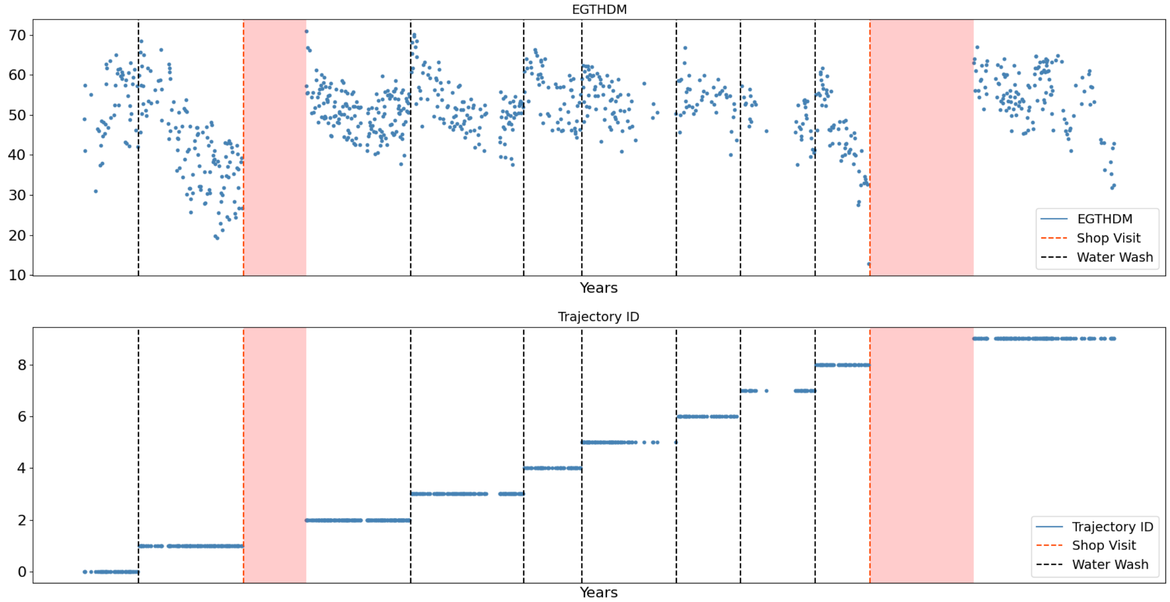


Figure 10.3: Plot of EGTHDM and trajectories for ESN956XXX

Trajectories in this case are defined as the intervals in the sequential data following a maintenance events. In this case, shop visits and water-washes are treated as maintenance events, therefore each trajectory number acts as an identifier within the sequential in order to separate and extract specific time intervals. Based on Figure 10.3, Trajectories 2 through 7 are used in order to provide sufficient data for training. Note that water-washes will be present within this interval but only occur over a very short window, therefore it is assumed that it will not significantly hamper the training process. In addition, water-washes only impact components with regards to recoverable deterioration, therefore this interval was selected in order for the prognostics model to also account for the effect of non-recoverable deterioration. Furthermore, water-washes still lead to sudden increases in component condition, however it is known that they barely affect the Fan+LPC and turbine sections and mostly affect the HPC. This will be discussed in further sections.

10.2.2. Rolling Average

Since the output of the HDMR surrogate model will be used to train the LSTM model, an important note must be made regarding the nature of the output. As seen in Section 9.5, the surrogate model outputs are generally quite scattered, similar to the measured EGTHDM over time. However, trends are still visible on average and that is distinguishable in the trending plots in Section 9.5 via the Gaussian rolling average shown as the red line.

The Gaussian rolling average provides a clearer trend for LSTMs to train on. Initial iterations of the LSTM model revealed that its prediction capability on unseen data returns a relatively flat line similar to a linear regression when using the scattered output. Essentially, the model could not handle the scatter within the data between flights. Therefore, the decision was made to extract the Gaussian rolling average instead as training data. This allowed for better convergence during training and the following prediction tests improved compared to earlier iterations.

10.2.3. Data Normalization

Once the data was compiled from the surrogate model outputs, the next step is to normalize the data. Data normalization reduces the dataset to the same scale, allowing for ML models to better identify relationships and patterns. Furthermore, it leads to better stability during training, mitigating issues that may arise from large or vanishing gradients. Normalization also ensures that the input features contribute proportionally to the training process, allowing for the model to identify patterns and relationships across features more effectively [89].

$$X_{norm} = \frac{X - X_{min}}{X_{max} - X_{min}} \tag{10.1}$$

In this case, Min-Max scaling is used as normalization technique, as shown above in Equation 10.1, where X represents any feature value, while X_{min} and X_{max} represent the minimum and maximum values of said feature respectively within the given dataset. .Min-max scaling reduces the features to a specific range, in this case between 0 an 1. This normalization technique is a widely used, especially for cases where the upper and lower bounds of the dataset are known and the distribution of the data is unknown but uniformly distributed within the range of values [89]. Min-max scaling in general also allows for faster convergence of the gradient descent compared to training where data is not scaled [90].

10.3. Model Development & Tuning

Similar to the HDMR surrogate model, the LSTM model was developed in Python 3, using the Keras; a widely used library for machine learning applications.

10.3.1. Model Parameters

Window Size

The window size, often referred to as the sequence length, refers to the length of the input vector the model processes at each time step. In the context of forecasting, it is how many steps in the past are used as input to compute the next step in a given sequence. It essentially dictates how much previous information the LSTM must take into account before making the following prediction.

The choice of window size has a direct impact on the performance of the model. If the window size is too narrow or too short, the input vector at a given time step may not provide sufficient information for the LSTM cells to capture long-term relationships and dependencies within the training data. However, too large of a window size risks introducing unnecessary complexity, hampering the training process and potentially lead to over-fitting and lower performance during inference.

Number of Layers & Cells
For any neural network, the model developer can specify any number of layers and cells within the layers, depending on hardware limitations. In the case of an LSTM network, the number of layers refers to how many LSTM layers are stacked, where each layer contains a number of LSTM cells.

Adding more layers to an LSTM network introduces more levels of abstraction and interpretation of the input data, further improving its ability to capture complex patterns within the sequential data. This can potentially lead to better results and predictions when inferring on unseen input data. That being said, a large number of layers can also lead to increased computational cost and training time, alongside increased risk of over-fitting. This can be mitigated if sufficient training data is available and the hardware employed is not as much of a limiting factor.

Learning Rate
In machine learning, the learning rate is a hyper-parameter defining the rate at which the model weights are adjusted with respect to the loss gradient during training. In terms of optimization, it essentially dictates the step size used when adjusting the weights at each iteration when moving towards a minimum of the objective function (often referred to as the loss function).

It is common for the learning rate to be tuned during development, as there is no indication of what the optimal learning rate is prior to training data. Increasing the learning rate results in faster convergence rate, however if set too large, the model weights can risk converging earlier towards a non-optimal solution or oscillate between minimums. Conversely, if the learning rate is set too low, the convergence rate can be significantly reduced, leading to slower training and also risk falling into a local minimum. Therefore, selecting an appropriate learning rate ensure efficient training and better chances of convergence towards an acceptable solution.

Batch Size
Alongside the learning rate, batch size is another hyper-parameter that also often requires tuning. Batch size defines the size of a subset of training data used during each iteration. In reality, it is this subset that is fed into the model when computing gradients during training.

Batch size can affect both training time and convergence. A large batch size generally provides more accurate estimations of the gradients but for a larger computational cost. Small batch sizes then allows for faster training time but can introduce noise in the gradient computations, thus compromising the training process.

10.3.2. Baseline Model

The first step in development was to establish a baseline model as a starting point. The results from the baseline model would later be used to inform further tuning of the parameters discussed in the previous section and provide some initial remarks on potential improvements.

The baseline model is constructed using 3 layers; the input layer, the LSTM layer and output layer. The input and output layer both consist of 8 standard fully connected neurons, one for each of the features being the $\Delta\eta$ and ΔW_c of all 4 components. The LSTM layer in between made up of 64 LSTM cells, which then return their outputs to the output layer. The window size was set to 10, thus at each time step the LSTM model would account for the 10 previous flight cycles. The batch size was set to 64 and training was left running for 350 epochs using ADAM as the optimizer. Below are the resulting training losses. The dataset was split $90\%/10\%$ for training and validation respectively, where the validation subset is used to evaluate the model between iterations in order to inform the training process.

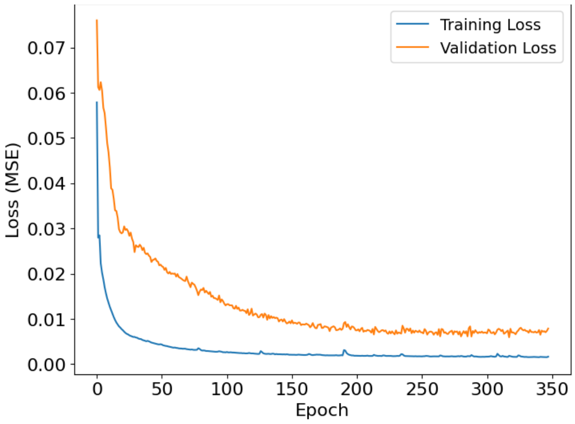


Figure 10.4: Training and validation loss based on MSE for Baseline model

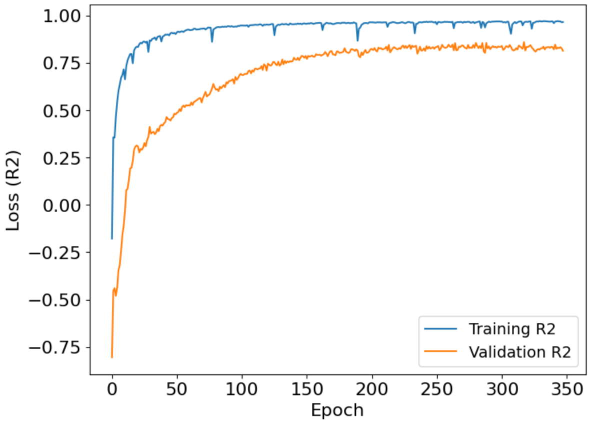


Figure 10.5: Training and validation loss based on R2 coefficient for Baseline model

Table 10.1: Final values from training and validation losses of Baseline model

Loss	Training	Validation
MSE	0.001662	0.007874
R2	0.964165	0.813042

Figure 10.4 show the training and validation loss based on Mean Squared Error. Both training and validation losses for MSE decrease significantly in early epochs, indicating an adequate training process and that the model is converging towards an adequate solution without falling into a local minimum. The training loss is significantly reduced early on, however the validation loss requires more epochs to decrease then reaches a point where it remains relatively constant. This would mean that the model is over-fitting; it has learned from the training data very well but will not perform as well on unseen data. As for Figure 10.5, the R^2 coefficient shows large improvements in early epochs as well for both training and validation. In this case, both the training and validation R^2 values converge to a value just below 1, suggesting a good fit to the data. However, the training R^2 is slightly higher than the validation R^2, which again points to over-fitting. The validation loss in terms of MSE remains quite low and relatively high for R^2. Model performance on the validation set is generally a better indicator of a model's ability to generalize and its potential performance on unseen data.

10.3.3. Model Tuning

Following the results discussed in Section 10.3.2 for the baseline model, the next step was to experiment with hyper-parameter tuning. The parameters that define the LSTM model are essentially selected based on the user's choice. In fact, there is no exact literature on how large the network is in terms of cells or layers should be, or what learning rate to use for example. Therefore, it is common in machine learning development to define said hyper-parameters and conduct experiments in order to asses the most suitable combinations. The parameters explored in this thesis are further discussed in the following subsections.

To summarize, below in Table 10.2 is an overview of the relevant model parameters and the range of values over which experiments are conducted. Note that the learning rate and amount of epochs for training was reduced to 200 epochs based on the results obtained in Section 10.3.2. This point was chosen for early stopping of training point, in order to save computational resources but can also help reduce over-fitting.

Table 10.2: List of parameters and range for model tuning

Parameter	**Range**
Batch Size	$[32, 64, 128, 256]$
LSTM Cells	$[8, 16, 32, 64, 128]$
Window Size	$[5, 10, 15, 20, 25]$

Table 10.3: Optimal results from parameter sweeps

Batch Size	**LSTM Cells**	**Window Size**	**MSE**	**RMSE**	R^2
32	128	5	$7.958 \cdot 10^{-4}$	0.022675	0.981847

Based on the results obtained from the parameter sweep, the combination of the model parameters shown below in Table 10.3 was identified to return the best results according to the selected performance metrics. The complete set of results form the parameter sweep is provided in Appendix E. From these findings, the model was retrained and returned the following results.

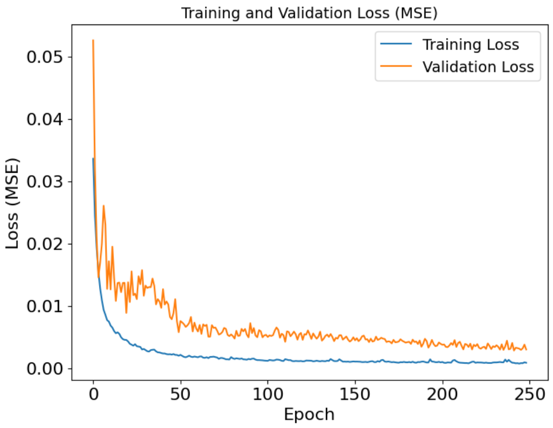


Figure 10.6: Training and validation loss based on MSE for Tuned model

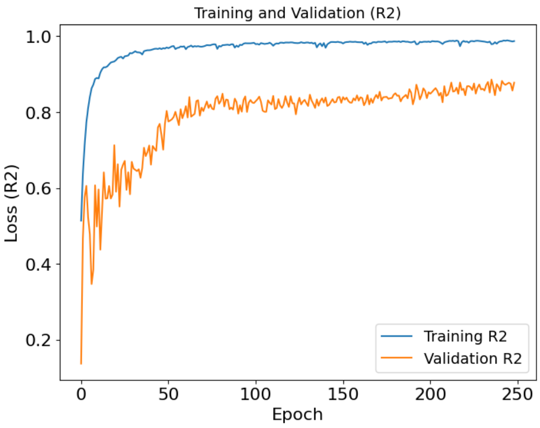


Figure 10.7: Training and validation loss based on R2 coefficient for Tuned model

10.4. Predictive Capability

Following the development of the LSTM model, the predictive capability was tested for both the Baseline and Tuned model. As explained previously in Section 10.2, the trajectories used for the training data were Trajectories 2 through 7 were used. The interval for Trajectory 8 was reserved as a test case to assess the LSTM model's prediction performance. The reason why this interval was selected is that it is the interval that preceded the second shop visit recorded for ESN956XXX. Recall from Section 9.5 that Event 2 was identified as a HPT related maintenance event due to progressive abrasion of the HPT blades. This component failure was said to be progressive, thus LSTM should have been able to capture the long-term deterioration pattern in the data that lead up to the HPT failure and the engine's subsequent shop visit.

The following sections discuss the results of the prediction tests for both the baseline and tuned model. Note that the horizontal axis represents time and each point represents one flight cycle. The time axis is once again anonymized due to confidentiality. Furthermore, the predictive capability was assessed for the condition of all four components in terms of $\Delta\eta$ and ΔW_c. As predictions are generated, each new prediction is added to the total vector in order for the rolling window predict the following points. The results a represented in red for the original condition trend and in blue for the predictions made by the LSTM model

10.4.1. Baseline Model

Starting with the baseline model, the window size was set to 10, thus the first 10 points within the test data are needed as inputs to predict the 11th point and so forth. The following figures were generated to showcase a series of predictions made for 50 flight cycles into the future, starting with the 11th point in the sequence. To demonstrate the results, this section will focus primarily on predictions made for the HPC and LPT. The full set of prediction results are provided in Appendix E, however results for other components will also be referenced for comparison.

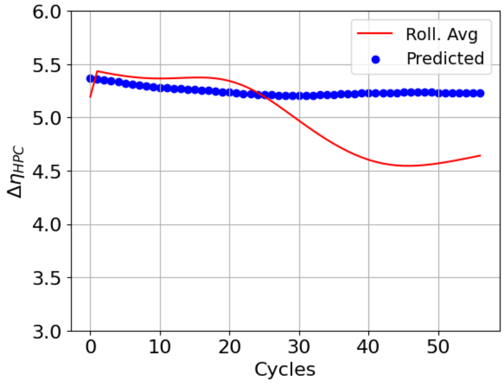


Figure 10.8: Resulting predictions from Baseline model for HPC $\Delta\eta$

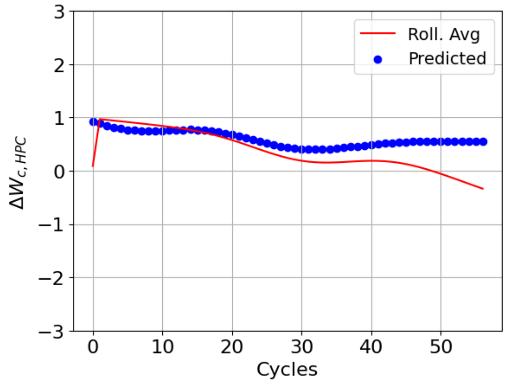


Figure 10.9: Resulting predictions from Baseline model for HPC ΔW_c

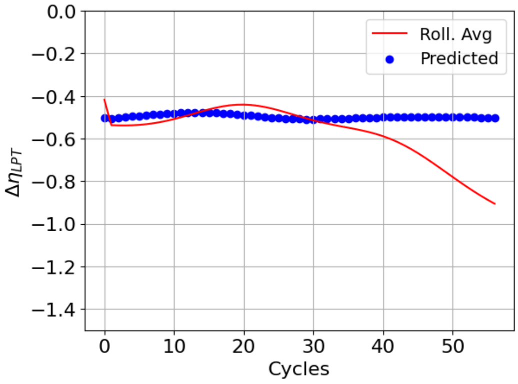


Figure 10.10: Resulting predictions from Baseline model for LPT $\Delta\eta$

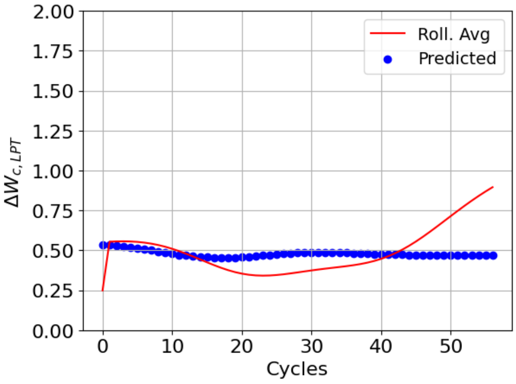


Figure 10.11: Resulting predictions from Baseline model for LPT ΔW_c

As you can see in Figure 10.8 and Figure 10.9, the baseline LSTM-based prognostics model tends to follow the underlying trend in component condition quite well in the initial time steps, indicating that the was able to capture the general patterns in the data. When inspecting Figure 10.9, the LSTM predictions are quite close to the expected rolling average compared to Figure 10.8, meaning that some patterns for certain components are easier to capture resulting in more stable predictions. However, within the first ten steps forward, there are instances where the model start to exhibit some lag between predictions and real data. This means that the model mostly reacts to changes rather than anticipate them, therefore requiring some time steps before reflecting oncoming changes. Similar behaviour is also seen in Figure 10.9, where the predicted trend shows an increase in HPC flow capacity around the 15th cycle only to later start decreasing after reacting to the change in slope of the rolling average line.

Another example of this issue is that the baseline model does not perform well to sudden changes. Sharp slopes in the rolling average are not properly reflected in the predictions. This would then suggest that the model either requires further tuning or more features must be introduced in order to better capture more complex changes. Similar behaviour was also seen for the HPT and Fan+LPC, where sudden peaks were not properly anticipated and thus predictions did not react to these changes. This is notable for $\Delta\eta$ of the HPT, where the trajectory start with a steep decrease which was not properly anticipated by the LSTM model, resulting in a relatively flat prediction line. Furthermore, following the first few time steps, the model predictions eventually flatten out, resulting in a smoothing effect on future predictions. This indicates that the model struggles in capturing the sudden changes and dealing with the underlying noise in the data, thus returning results lying closer to the mean value. From this, it can be concluded that the baseline, despite showing good results in the learning phase, does not sufficiently generalize to perform well on unseen data, resulting in inconsistent responses to changes and subsequent diffusion of the predicted trend.

10.4.2. Tuned Model

Following the results of the parameter sweep, the baseline model was then tuned for the parameters and range specific in Table 10.2. Based on the selected performance metrics, the best combination of model parameters was identified and the LSTM model was retrained as an attempt to obtained better results compared to the baseline model. Below are the results from predictions made for $\Delta\eta$ and ΔW_c for the Fan+LPC and LPT components. Once more, results for the remaining components are included in Appendix E.

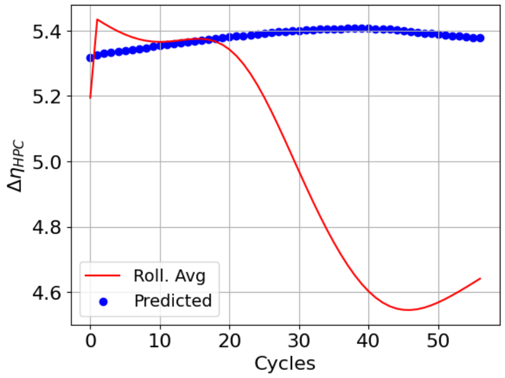


Figure 10.12: Resulting predictions from Tuned model for HPC $\Delta\eta$

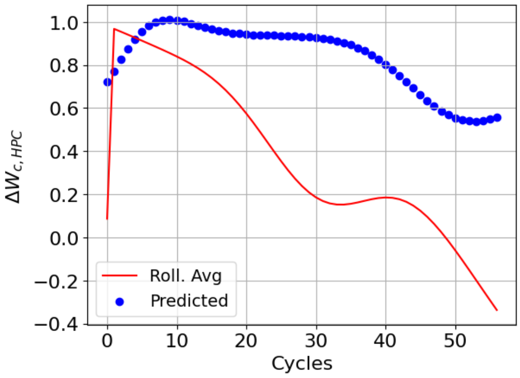


Figure 10.13: Resulting predictions from Baseline model for HPC ΔW_c

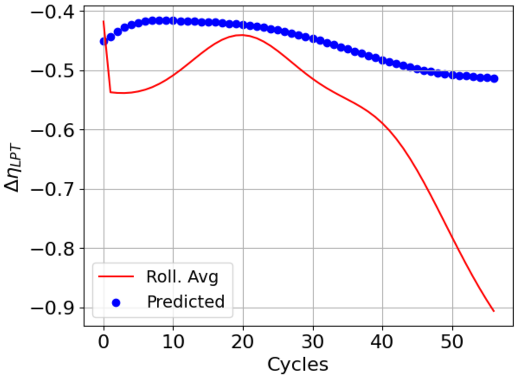


Figure 10.14: Resulting predictions from Tuned model for LPT $\Delta\eta$

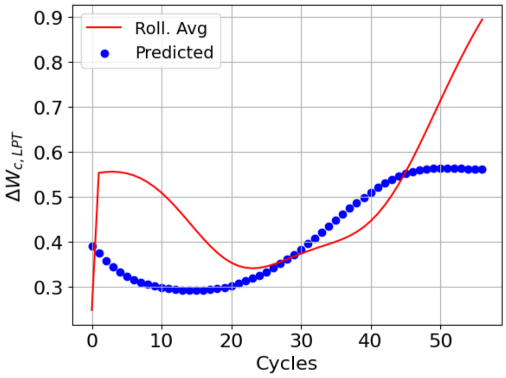


Figure 10.15: Resulting predictions from Tuned model for LPT ΔW_c

Based on the these results, the tuned model exhibits similar characteristics mentioned previously for the baseline model, with some notable improvements in some areas but similar shortcomings. In the case of Figure 10.12, the LSTM model was able to approximate the trend pretty closely for the first 20 cycles, and remained relatively close to the rolling average line in the case of Figure 10.15. However, the LSTM model still struggles to cope with sudden increases, as shown in both Figure 10.12 and Figure 10.13 around the 20th cycle, where a sudden downward slope in component condition was registered and predictions continued to follow a straight line. On the other hand, the increasing trend in the rolling average line followed by a relatively progressive decrease in Figure 10.13 indicates that the tuned model performed better at handling sudden changes in comparison to the baseline results. In further time steps its ability to follow the rapidly increasing trend quickly degrades, which is a common issue in ML models for forecasting. As a given model projects further in time, the mismatch between predictions and actual data starts to grow.

Referring back to remarks made about the baseline model, the predictive capability of the tuned LSTM model also varies between components, indicating that patterns within the data for some components are easier to learn than others. Notably, Figure 10.14 and Figure 10.15 show a better match with the overall trend compared to Figure 10.12 and Figure 10.13. Conversely, the tuned model does not perform as well for components such as the Fan+LPC and HPT. The rolling average for the condition of those two components involved sudden peaks and troughs throughout the interval observed. The tuned model, similarly to the baseline model, returned a generally liner trend further indicating that the LSTM model it is current form is not well equipped to handle the more complex deterioration trends seen in the training data. However, one notable improvement compared to the baseline model is that the LSTM model was able to return predictions further in time while still maintaining the general trend from the unseen data. This is particularly noticeable in Figure 10.14 and Figure 10.15, where despite a mismatch in magnitude, the pattern of the predictions for the first half of the total predictions still follows the general trend of the rolling average data.

10.5. Discussion

Based on the results above, it is clear that the neural network approach utilizing LSTMs has potential with regards to prognostics. Results from the training process at first indicates that the network is able to capture certain patterns within the training data and replicate certain trends in component degradation and showed signs of improvement after model tuning. That being said, this current iteration of the LSTM-based prognostics model does demonstrate certain drawbacks and shortcomings.

Training and validations losses during training for both the baseline and tuned model show good convergence during the training process, which suggests that the model is capable of learning the underlying complex patterns and relationships within the provided dataset. That being said, the tendency for the model to return a linear and flat trend points to over-fitting; the model has captured overly specific relationships within the data that can not be properly generalized to unseen data. This is a very common issue in machine learning, including applications involving time series forecasting [91]. Furthermore, in this approach, resulting predictions are added to the rolling window to make the following predictions. It is common with this approach that errors accumulate with every prediction, progressively resulting in less accurate predictions as the forecasting horizon extends.

Another aspect to consider is the training data used for this experiment. Recall that in Section 10.2, the snapshot data used to generate he training data from the surrogate model was segmented into trajectories. The goal was to isolate intervals between maintenance events in order to isolate data samples that would train the LSTM model to capture the component deterioration trend over time and flight cycles. While the LSTM model was able to capture some of the patterns during training, it struggled to reproduce the expected trends seen in other components on unseen data. Obtaining adequate performance during training but not when inferring on unseen data indicates that the training data did not include sufficient diversity in examples for the LSTM model to generalize. A solution would be to acquire a more diverse dataset in order to train the model on several examples of component deterioration trends.

Furthermore, it is worth noting that deterioration patterns for engine components over time and flight cycles are a function of multitude of variable and highly complex and non-linear phenomena. While LSTM models have demonstrated good performance in the field of forecasting, certain time-series are inherently complicated to predict. This is especially true in the case of engine condition, which based on engine performance can significantly vary between flight cycles, further complicating the task of making accurate predictions over long horizons. That being said, increasing the complexity of the LSTM model by introducing more features, layers and cells has the penitential of producing better outcomes and capturing the relevant relationships within the engine data.

11

Conclusion & Recommendations

To reiterate, the goal of this project was to contribute towards research at KLM Engine Services with regards to Engine Condition Based Maintenance for the GEnx-1B by expanding the capabilities of the surrogate model proposed by KLM ES, while exploring for a possible implementation of a prognostics model. This chapter outlines the conclusions in Section 11.1, while providing recommendations in Section 11.2 based on the outcome of this project.

11.1. Conclusions

The following sections outline the main conclusions drawn from the result of the implementation of the proposed extended modelling approach

- **Secondary Performance Parameters:** The effect of secondary performance parameters was investigated and implemented within the HDMR generation process. The parameterization used to quantify the effects of the secondary performance parameters proved to be effective. The resulting relations between SPP settings and their respective performance impact was derived from an optimization running on a genetic algorithm. The optimization returned a significant reduction in the deviation between the model outputs and the snapshot data. A reduction in model deviation was reported in engine performance simulated by the adapted GSP model of the GEnx-1B, while leading to a reduction in Standard Deviation in the GSP model output, concluding that the resulting relations were a satisfactory approximation of the effect of the secondary systems on the GEnx-1B's performance.

- **Extended Surrogate Model:** The surrogate model developed by KLM ES makes use the High Dimensional Model Representation framework, which allows for the introduction of expert knowledge within the surrogate model by implementing user defined basis functions. The list of basis function parameters used in this surrogate model was expanded to now account for the secondary performance parameters and wider range of operating points from on-wing data. This resulted in a more flexible surrogate model with respect to the acceptable range of operating cases and increasing accuracy for some component condition terms but at the expense of increased RMSE for others. Based on the results obtained, the various increments in RMSE are primarily due to the increased model complexity given a larger set of parameters are now included within the surrogate model.

- **Condition Trending:** The resulting HDMRs were verified by analyzing their diagnostic capability through verification of engine condition trending. Results show that the surrogate model was able to identify the expected trends in component-wise condition. The implementation of secondary performance parameters further highlighted certain effects like the seasonality in HPC condition between water washes compared to previous results and the progressive deterioration seen for the HPT and LPT trending further aided by the accounting of the behaviour of the active clearance control systems and power off-take.

- **LSTM-based Prognostics Model:** A prognostics model in the form of an LSTM-based machine learning model was proposed. Results from the training process indicates that the network is able to capture certain patterns within the training data and replicate certain trends in component degradation and showed signs of improvement after model tuning. The use of component deterioration data generated from the HDMR-based surrogate model shows promise as a source of training data, further indicating that the combined modelling approach proposed in this thesis is a good starting for future development of prognostics model for ECBM.

- **Predictive Capability:** Following development of the prognostics model, its predictive capability was put to the test on unseen data. Results obtained from the training process show that this type of model is capable of learning deterioration patterns but to a limited extent given the LSTM network's relatively low complexity. Further tuning lead to improved results, recreating the expected trends up to 10 flight cycles in advance before accuracy began to decrease. The predictive capability of the LSTM model suffers from the same common issues in machine learning, such as over-fitting and lack of extensive data which includes sufficient diverse examples to train on. Nonetheless, the LSTM framework has demonstrated its capability as a prognostics model.

- **Combined Modelling:** The goal of this thesis was to combine several aspects of engine modelling in the context of ECBM at KLM Engine Services. A method for accounting for the effects of measurements from secondary systems on-board the GEnx-1B was studied and implemented within a surrogate modelling framework. The result was a more complete surrogate model, capable of taking advantage of more of a wider range of on-wing, while providing clear trends in component condition for diagnostic purposes. A machine learning model was then introduced in order to take advantage of the surrogate model's diagnostics capabilities in order to inform its training and provide predictions on future engine component conditions. While certain aspects of this framework would require further research, it serves as a adequate reference for ongoing efforts in ECBM at KLM Engine Services.

11.2. Recommendations

Following the conclusions made from the results of this thesis, this section addresses the shortcomings reported by providing recommendations for future work.

- **Investigate cruise conditions:** During the exploration and implementation phase of the SPP parameterization, the choice was made to focus solely on the take-off phase. The justification was based on the result of the development of the GSP model of the GEnx-1B at KLM ES. The GSP model was constructed using data collected from test-cell runs at sea level, hereby providing the most accurate simulation results, within $\pm 1\%$ from the on-wing data, as long as the operating point simulated was close to sea level. The advantage of investigating cruise conditions is that operating conditions can be taken as steady-state, while the variation in ambient conditions is generally very small. Research has been ongoing at KLM ES on developing solutions to improve engine modelling accuracy at cruise, such as re-calibrated models and other approaches to account for effects such as Reynolds effects and nozzle flow effects. The combined modelling framework presented in this thesis is flexible enough to allow for the introduction of modified GSP models better equipped to simulate cruise conditions.

- **Explore other parameterizations:** While the quadratic parameterization of the SPP returned good results, the variation in SPP settings found during take-off was significant, further complicating the process of selecting adequate reference settings and requiring additional constraints in the GA-based optimization to avoid inconsistent solutions. It is therefore worth exploring other parameterization methods, such as cubic curves or splines, that could potentially be more robust in order to return better approximations of their effects on engine performance into the wider surrogate modelling framework.

- **Explore other HDMR basis functions:** The HDMR-based surrogate model consists of user defined basis functions, This allowed users to specific specific function inputs, in this case engine measurement parameters were used. This allowed for the introduction of expert knowledge within the surrogate model by selecting the most relevant engine measurement parameters. The choice of basis function also dictates the level of complexity that can be captured. As it stands, the current iteration of the HDMR-based surrogate model uses a relatively simple basis function with a given engine measurement as input parameter. It is therefore exploring the implementation of other basis functions in order to capture more complex relationships, but likely at the cost of higher computational time during generation.

- **Higher order terms:** During development of the surrogate model, the HDMR generation was limited to second-order terms based on literature and recommendations made in prior work, claiming sufficient accuracy is reached. However, given that the list of input parameters to the HDMRs has now been extended, the implementation of third or higher-order terms may provide some benefit. Second-order terms may not be sufficient to capture the complex relationships between the various combinations of parameters. If computational cost can be further reduced, higher order terms can potentially contribute to further reductions in RMSE.

- **Computational cost:** During development of the surrogate model, there are two integral processes that resulted in significant computational cost. The first was the data generation process for the HDMRs. Simulations were conducted using the ParallelGSP package which distributed the workload over several virtual machines. As the GSP model increased in complexity due to the implementation of SPPs, the simulation time significantly increased. Furthermore, the second process was the HDMR generation process, which depending on the computation hardware, bounds, selected tolerance and number of input parameters in the basis functions, generation times also significantly increased. It is therefore recommended for future work to explore methods or solutions that can drastically reduce the computational cost, allowing users to iterate faster during development. In addition, the HDMR generation was limited to second-order terms based on recommendations made in prior work, claiming sufficient accuracy is reached. If computational cost can be reduced, higher order terms can potentially contribute to further reductions in RMSE.

- **LSTM model complexity:** While the current iteration of the LSTM-based prognostics model proposed in this thesis showed promising results, further work will be required in order to produce more accurate results. Given the scope and resources available for this thesis project, tuning was conducted on a relatively brief set of model hyper-parameter since the model itself was not relatively complex. Furthermore, based on existing knowledge, the LSTM model was kept at a single layer, what many might refer to as a "vanilla" LSTM. Development of a more complex model, perhaps involving more layers, cells and features, might return better results as more complex patterns are learned.

- **Assess relevance of other features:** In this implementation, the input features of the LSTM model were chosen to the be the component-wise map modifiers as their health indicators. This decision was made to allow the direct use of the outputs of the HDMR-based surrogate model as input to the LSTM model. Therefore, it is recommended to explore the implementation of the LSTM model using the system-level health indicators, ΔEGT, $\Delta TSFC$ and ΔW_f, as input features in order to generate predictions of engine sysem-level condition.

- **Variety in failure cases:** The results obtained from this thesis were constrained to a single engine, labelled ESN956XXX. While Snapshot data obtained for this specific GEnx-1B engine covers a large period of time, it limits the number of failure cases present. This engine was selected due to the presence of two major HPT-related maintenance events, which serve as valuable information when assessing the condition trending capabilities of the surrogate model while providing training data for the LSTM model. However it is recommended to investigate and collect more data from various engines, covering a variety of failure cases, in order to improve the prognostics model's capability to generalize and further validate the accuracy of future models.

Bibliography

[1] van Buijtenen, J.P, Visser, W., Gangoli Rao, A. *Aero Engine Technology - AE4238*. 5th ed. Delft, The Netherlands: Faculty of Aerospace Engineering, TU Delft, 2021.

[2] *The Messy Battle to Build the F-35's Next Engine — popularmechanics.com*. `https://www.popularmechanics.com/military/aviation/a43128883/f-35-needs-new-engine/`. [Accessed 21-07-2023]. 2023.

[3] Walsh, P.P, Fletcher, P. *Gas Turbine Performance*. 2nd ed. Oxford, England: Blackwell Science Ltd., 2004.

[4] GE Aviation. *Genx-1B - Turbofan Engine Installation Manual*. 2021.

[5] de Bruin, B. *Surrogate Model Based Diagnostics for the GEnx-1B Turbofan Engine*. 2025. URL: `http://resolver.tudelft.nl/uuid:f2424334-caeb-410e-a3d6-4ed642288fa7` (visited on 05/13/2023).

[6] Visser, W.P.J. "Generic Analysis Methods for Gas Turbine Engine Performance". 2015. DOI: `10.4233/UUID:F95DA308-E7EF-47DE-ABF2-AEDBFA30CF63`. URL: `http://resolver.tudelft.nl/uuid:f95da308-e7ef-47de-abf2-aedbfa30cf63`.

[7] Cavcar, M. *The International Standard Atmosphere (ISA)*. URL: `http://fisicaatmo.at.fcen.uba.ar/practicas/ISAweb.pdf` (visited on 07/15/2023).

[8] Joachim Kurzke and Ian Halliwell. *Propulsion and Power*. Springer International Publishing, 2018. ISBN: 9783319759791. DOI: `10.1007/978-3-319-75979-1`. URL: `http://dx.doi.org/10.1007/978-3-319-75979-1`.

[9] Rainer Kurz et al. *GAS TURBINE DEGRADATION*. 2014.

[10] C.B. Meher-Momji, M. A Chaker, and H.M. Motiwala. "Gas Turbine Performance Deterioration". In: ().

[11] S O T Ogaji et al. *Parameter selection for diagnosing a gas-turbine's performance-deterioration*. 2002. URL: `www.elsevier.com/locate/apenergy`.

[12] E L Ntantis, P N Botsaris, and wwwjestrorg Jestr. "Diagnostic Methods for an Aircraft Engine Performance Engineering Science and Technology Review". In: *Journal of Engineering Science and Technology Review* 8 (4 2015), pp. 64–72. ISSN: 1791-2377. URL: `www.jestr.org`.

[13] Mohammadreza Tahan et al. "Performance-based health monitoring, diagnostics and prognostics for condition-based maintenance of gas turbines: A review". In: *Applied Energy* 198 (2017), pp. 122–144. ISSN: 03062619. DOI: `10.1016/j.apenergy.2017.04.048`.

[14] Houman Hanachi et al. "Performance-Based Gas Turbine Health Monitoring, Diagnostics, and Prognostics: A Survey". In: *IEEE Transactions on Reliability* 67 (3 Sept. 2018), pp. 1340–1363. ISSN: 00189529. DOI: `10.1109/TR.2018.2822702`.

[15] E. M. Greitzer, C. S. Tan, and M. B. Graf. *Internal Flow*. Cambridge University Press, Apr. 2004. DOI: `10.1017/cbo9780511616709`. URL: `https://doi.org/10.1017/cbo9780511616709`.

[16] Tomas U. J. Gro¨nstedt. "Identifiability in Multi-Point Gas Turbine Parameter Estimation Problems". In: *Volume 2: Turbo Expo 2002, Parts A and B*. ASMEDC, Jan. 2002. DOI: `10.1115/gt2002-30020`. URL: `https://doi.org/10.1115/gt2002-30020`.

[17] Ramdin, S.S. *Development of a diagnostics model for the GEnx-1B turbofan engine using on-wing performance data*. 2021. URL: `http://resolver.tudelft.nl/uuid:3269e573-31b7-4bf4-91f4-9cfd08166c63` (visited on 06/15/2023).

[18] Roell, B. *Test-cell and On-wing Turbofan Performance Comparison at KLM Engine Services*. 2019. URL: `http://resolver.tudelft.nl/uuid:8b882e0a-aa26-4785-90ce-27b5fed7b542` (visited on 06/16/2023).

[19] Barrett, R.F, Smetana, F.O. “State-of-the-art Review of Turbine Engine Chkecout Techniques”. In: *AIAA* (1969).

[20] Y. G. Li. “Performance-analysis-based gas turbine diagnostics: A review”. In: *Proceedings of the Institution of Mechanical Engineers, Part A: Journal of Power and Energy* 216 (5 2002), pp. 363–377. ISSN: 09576509. DOI: `10.1243/095765002320877856`.

[21] Slawomir Koziel and Xin-She Yang, eds. *Computational Optimization, Methods and Algorithms*. Springer Berlin Heidelberg, 2011. DOI: `10.1007/978-3-642-20859-1`. URL: `https://doi.org/10.1007/978-3-642-20859-1`.

[22] Cheng Yan et al. “Surrogate-based optimization with improved support vector regression for non-circular vent hole on aero-engine turbine disk”. In: *Aerospace Science and Technology* 96 (Jan. 2020). ISSN: 12709638. DOI: `10.1016/j.ast.2019.105332`.

[23] Shirui Luo et al. “Turbomachinery Blade Surrogate Modeling using Deep Learning”. In: June 2021.

[24] Zuming Liu and Iftekhar A. Karimi. “Gas turbine performance prediction via machine learning”. In: *Energy* 192 (Feb. 2020). ISSN: 03605442. DOI: `10.1016/j.energy.2019.116627`.

[25] Amare D. Fentaye et al. “A review on gas turbine gas-path diagnostics: State-of-the-art methods, challenges and opportunities”. In: *Aerospace* 6 (7 July 2019). ISSN: 22264310. DOI: `10.3390/aerospace6070083`.

[26] Moorselaar, M.P.R. *Gas Path Analysis on the GEnx-1B at KLM Engine Services*. 2018. URL: `http://resolver.tudelft.nl/uuid:51b66a01-a485-4aac-a6e6-544f619b9aa1` (visited on 05/10/2023).

[27] G. LaRocca - TU Delft. *Lecture 1 - The Artificial Intelligence roots (part 1)*. 2021. URL: `https://brightspace.tudelft.nl/d2l/le/content/397952/viewContent/2274228/View` (visited on 07/19/2023).

[28] Changduk Kong. “Review on Advanced Health Monitoring Methods for Aero Gas Turbines using Model Based Methods and Artificial Intelligent Methods”. In: *International Journal of Aeronautical and Space Sciences* 15.2 (June 2014), pp. 123–137. DOI: `10.5139/ijass.2014.15.2.123`. URL: `https://doi.org/10.5139/ijass.2014.15.2.123`.

[29] Xixiang Liu et al. “A Simplified Kalman Filter for Integrated Navigation System with Low-Dynamic Movement”. In: *Mathematical Problems in Engineering* 2016 (Jan. 2016), pp. 1–9. DOI: `10.1155/2016/3528146`.

[30] Liping Yan et al. “Unscented Kalman-filter-based simultaneous diagnostic scheme for gas-turbine gas path and sensor faults”. In: *Measurement Science and Technology* 32.9 (June 2021), p. 095905. DOI: `10.1088/1361-6501/abfd67`. URL: `https://doi.org/10.1088/1361-6501/abfd67`.

[31] A.G Stamatis. “Engine Condition Monitoring and Diagnostics”. In: *Progress in Gas Turbine Performance*. InTech, June 2013. DOI: `10.5772/54409`. URL: `https://doi.org/10.5772/54409`.

[32] Stephen O.T Ogaji and Riti Singh. “Advanced engine diagnostics using artificial neural networks”. In: *Applied Soft Computing* 3.3 (Nov. 2003), pp. 259–271. DOI: `10.1016/s1568-4946(03)00038-3`. URL: `https://doi.org/10.1016/s1568-4946(03)00038-3`.

[33] Lai Chenyang, Yingqing Guo, and Haotian Mao. “Main fuel pump health monitoring method based on Kriging surrogate model”. In: July 2019, pp. 5112–5117. DOI: `10.23919/ChiCC.2019.8865827`.

[34] Marcel Aulich and Ulrich Siller. “High-Dimensional Constrained Multiobjective Optimization of a Fan Stage”. In: *Volume 7: Turbomachinery, Parts A, B, and C*. ASMEDC, Jan. 2011. DOI: `10.1115/gt2011-45618`. URL: `https://doi.org/10.1115/gt2011-45618`.

[35] R. Ganguli. “Data Rectification and Detection of Trend Shifts in Jet Engine Path Measurements Using Median Filters and Fuzzy Logic”. In: *Journal of Engineering for Gas Turbines and Power* 124.4 (Sept. 2002), pp. 809–816. DOI: `10.1115/1.1470482`. URL: `https://doi.org/10.1115/1.1470482`.

[36] S.O.T. Ogaji et al. “Gas-turbine fault diagnostics: a fuzzy-logic approach”. In: *Applied Energy* 82.1 (2005), pp. 81–89. ISSN: 0306-2619. DOI: `https://doi.org/10.1016/j.apenergy.2004.07.004`. URL: `https://www.sciencedirect.com/science/article/pii/S0306261904001035`.

[37] Igor Loboda. *Neural Networks for Gas Turbine Diagnosis*. Ed. by Joao Luis G. Rosa. Rijeka: IntechOpen, 2016. Chap. 8. DOI: `10.5772/63107`. URL: `https://doi.org/10.5772/63107`.

[38] *TensorFlow Keras Model | TensorFlow Keras Model and Method — educba.com*. `https://www.educba.com/tensorflow-keras-model/`. [Accessed 31-08-2023].

[39] Amir Safari et al. “Application of an Experimentally Validated High-dimensional Surrogate Model to Performance Prediction of a CHP Unit”. In: *Energy Procedia* 142 (2017). Proceedings of the 9th International Conference on Applied Energy, pp. 1106–1111. ISSN: 1876-6102. DOI: `https://doi.org/10.1016/j.egypro.2017.12.635`. URL: `https://www.sciencedirect.com/science/article/pii/S1876610217363981`.

[40] Mortda Mohammed, Maher K. Taher, and Saleh khudhair. “Prediction of turbojet performance by using artificial neural network”. In: *Materials Today: Proceedings* 60 (Jan. 2022), pp. 1513–1522. ISSN: 22147853. DOI: `10.1016/j.matpr.2021.12.027`.

[41] Chang-duk Kong et al. “Intelligent performance diagnostics of a gas turbine engine using user-friendly interface neural networks”. In: *Aircraft Engineering and Aerospace Technology* (2004). ISSN: 0002-2667. DOI: `10.1108/00022660410545500`. URL: `https://doi.org/10.1108/00022660410545500`.

[42] Gustavo Matuck et al. “Multiple Faults Detection of Gas Turbine by MLP Neural Network”. In: *Proceedings of the ASME Turbo Expo* 1 (Jan. 2009). DOI: `10.1115/GT2009-59964`.

[43] Leto Peel. “Data driven prognostics using a Kalman filter ensemble of neural network models”. In: (2008), pp. 1–6. DOI: `10.1109/PHM.2008.4711423`.

[44] Felix Heimes. “Recurrent neural networks for remaining useful life estimation”. In: Nov. 2008, pp. 1–6. DOI: `10.1109/PHM.2008.4711422`.

[45] A.K. Garga et al. “Hybrid reasoning for prognostic learning in CBM systems”. In: 6 (2001), 2957–2969 vol.6. DOI: `10.1109/AERO.2001.931316`.

[46] Ratna Babu Chinnam and Pundarikaksha Baruah. “A neuro-fuzzy approach for estimating mean residual life in condition-based maintenance systems”. In: *International Journal of Materials and Product Technology* 20.1/2/3 (2004), p. 166. DOI: `10.1504/ijmpt.2004.003920`. URL: `https://doi.org/10.1504/ijmpt.2004.003920`.

[47] Mazen Azzam, Mariette Awad, and Joseph Zeaiter. “Application of evolutionary neural networks and support vector machines to model NOx emissions from gas turbines”. In: *Journal of Environmental Chemical Engineering* 6 (1 Feb. 2018), pp. 1044–1052. ISSN: 22133437. DOI: `10.1016/j.jece.2018.01.020`.

[48] Suresh Sampath and Riti Singh. “An Integrated Fault Diagnostics Model Using Genetic Algorithm and Neural Networks”. In: *Journal of Engineering for Gas Turbines and Power* 128.1 (Mar. 2004), pp. 49–56. DOI: `10.1115/1.1995771`. URL: `https://doi.org/10.1115/1.1995771`.

[49] Takahisa Kobayashi and Donald Simon. “A Hybrid Neural Network-Genetic Algorithm Technique for Aircraft Engine Performance Diagnostics”. In: *Journal of Propulsion and Power* 21 (Aug. 2001). DOI: `10.2514/1.9881`.

[50] Bhaw Nath Jha and Hong shuang Li. “Structural reliability analysis using a hybrid HDMR-ANN method”. In: *Journal of Central South University* 24 (11 Nov. 2017), pp. 2532–2541. ISSN: 22275223. DOI: `10.1007/s11771-017-3666-7`.

[51] J.P. Sprong, X. Jiang, and H. Polinder. “Deployment of Prognostics to Optimize Aircraft Maintenance – A Literature Review”. In: *JOURNAL OF INTERNATIONAL BUSINESS RESEARCH AND MARKETING* 5 (4 2020), pp. 26–37. ISSN: 18498558. DOI: `10.18775/jibrm.1849-8558.2015.54.3004`.

[52] Anping Wan et al. “Prognostics of gas turbine: A condition-based maintenance approach based on multi-environmental time similarity”. In: *Mechanical Systems and Signal Processing* 109 (Sept. 2018), pp. 150–165. DOI: `10.1016/j.ymssp.2018.02.027`. URL: `https://doi.org/10.1016/j.ymssp.2018.02.027`.

[53] C Romesis. *EXPERIENCE WITH CONDITION-BASED MAINTENANCE RELATED METHODS AND TOOLS FOR GAS TURBINES*. 2014.

[54] Allan Volponi. "Gas Turbine Engine Health Management: Past, Present and Future Trends". In: vol. 136. June 2013. DOI: 10.1115/GT2013-96026.

[55] Tim Rootliep. "ECBM Vision - Current State-of-the-art ICCBM". Internal, KLM ES. 2022.

[56] Rootliep, T. *Turbofan Condition Monitoring using Evolutionary Algorithm based Gas Path Analysis*. 2020. URL: `http://resolver.tudelft.nl/uuid:7ff9b284-0b7b-4f70-8ed5-4d81f1eb113f` (visited on 05/02/2023).

[57] Ramdin, S.S. *A Systematic Approach For Modelling Modern Turbofan Engines*. 2022. URL: `http://resolver.tudelft.nl/uuid:aa5efca4-8f4a-44d6-8d15-0eeedd45b915` (visited on 06/08/2023).

[58] General Electric. *GEnx - High Bypass Turbofan Engines*. URL: `https://www.geaerospace.com/sites/default/files/datasheet-genx.pdf`.

[59] Gina Daugherty. *The magnitude of manufacturing in the GEnx - The GE Aerospace Blog | Aviation & Flight News — blog.geaerospace.com*. `https://blog.geaerospace.com/manufacturing/the-magnitude-of-manufacturing/`. 2016.

[60] Michael Foust et al. *Development of the GE Aviation Low Emissions TAPS Combustor for Next Generation Aircraft Engines*. Jan. 2012. DOI: 10.2514/6.2012-936. URL: `http://dx.doi.org/10.2514/6.2012-936`.

[61] Andreas Linke-Diesinger. *Systems of Commercial Turbofan Engines*. Springer Berlin Heidelberg, 2008, pp. 1–21.

[62] Pierre Tchakoua et al. "Wind Turbine Condition Monitoring: State-of-the-Art Review, New Trends, and Future Challenges". In: *Energies* 7.4 (Apr. 2014), pp. 2595–2630. DOI: 10.3390/en7042595. URL: `https://doi.org/10.3390/en7042595`.

[63] Maria Grazia De Giorgi, Nicola Menga, and Antonio Ficarella. "Exploring Prognostic and Diagnostic Techniques for Jet Engine Health Monitoring: A Review of Degradation Mechanisms and Advanced Prediction Strategies". In: *Energies* 16.6 (Mar. 2023), p. 2711. ISSN: 1996-1073. DOI: 10.3390/en16062711. URL: `http://dx.doi.org/10.3390/en16062711`.

[64] Alex Graves. *Supervised Sequence Labelling with Recurrent Neural Networks*. Springer Berlin Heidelberg, 2012. ISBN: 9783642247972. DOI: 10.1007/978-3-642-24797-2. URL: `http://dx.doi.org/10.1007/978-3-642-24797-2`.

[65] N. Borst and W.J.C. Verhagen. "Introducing CNN-LSTM network adaptations to improve remaining useful life prediction of complex systems". In: *The Aeronautical Journal* 127.1318 (2023), pp. 2143–2153. DOI: 10.1017/aer.2023.84.

[66] Artur Pessoa, Ruslan Sadykov, and Eduardo Uchoa. "Enhanced Branch-Cut-and-Price algorithm for heterogeneous fleet vehicle routing problems". In: *European Journal of Operational Research* 270.2 (Oct. 2018), pp. 530–543. ISSN: 0377-2217. DOI: 10.1016/j.ejor.2018.04.009. URL: `http://dx.doi.org/10.1016/j.ejor.2018.04.009`.

[67] Greg Van Houdt, Carlos Mosquera, and Gonzalo Nápoles. "A Review on the Long Short-Term Memory Model". In: *Artificial Intelligence Review* 53 (Dec. 2020). DOI: 10.1007/s10462-020-09838-1.

[68] Asian Aviation Staff. *Rolls-Royce, HAECO sign deal for engine services - Asian Aviation — asianaviation.com*. `https://asianaviation.com/rolls-royce-haeco-sign-deal-for-engine-services/`. [Accessed 30-08-2023].

[69] den Haan, D.M. *GSP GPA on CF6-80 Engines at KLM Engine Services*. 2010.

[70] Scott B. Lattime and Bruce M. Steinetz. "High-Pressure-Turbine Clearance Control Systems: Current Practices and Future Directions". In: *Journal of Propulsion and Power* 20.2 (Mar. 2004), pp. 302–311. ISSN: 1533-3876. DOI: 10.2514/1.9255. URL: `http://dx.doi.org/10.2514/1.9255`.

[71] J. D. Denton. "The 1993 IGTI Scholar Lecture: Loss Mechanisms in Turbomachines". In: *Journal of Turbomachinery* 115.4 (Oct. 1993), pp. 621–656. ISSN: 1528-8900. DOI: 10.1115/1.2929299. URL: `http://dx.doi.org/10.1115/1.2929299`.

[72] Moustapha S. H. Schaub U. W. Vlasic E. "Effect of Tip Clearance on the Performance of a Highly Loaded Turbine Stage". In: *NATO Advisory Group for Aerospace Research and Development*

(1994). ISSN: 0549-7191. URL: https://www.tib.eu/en/search/id/BLCP%3ACN002811923/Effect-of-Tip-Clearance-on-the-Performance-of-a/.

[73] Cis De Maesschalck et al. “Experimental and Numerical Investigation of Optimized Blade Tip Shapes - Part II: Tip Flow Analysis and Loss Mechanisms”. In: June 2018. DOI: 10.1115/GT2018-76567.

[74] A.L. Tamiru, F.M. Hashim, and C. Rangkuti. “Generating Gas Turbine Component Maps Relying on Partially Known Overall System Characteristics”. In: *Journal of Applied Sciences* 11.11 (May 2011), pp. 1885–1894. ISSN: 1812-5654. DOI: 10.3923/jas.2011.1885.1894. URL: http://dx.doi.org/10.3923/jas.2011.1885.1894.

[75] Saravanamutto, H,I,H, Rogers, G, Cohen, H. *Gas Turbine Theory*. 7th ed. Pearson, 2017.

[76] C. Romesis and K. Mathioudakis. “Setting Up of a Probabilistic Neural Network for Sensor Fault Detection Including Operation With Component Faults”. In: *Journal of Engineering for Gas Turbines and Power* 125.3 (July 2003), pp. 634–641. ISSN: 1528-8919. DOI: 10.1115/1.1582493. URL: http://dx.doi.org/10.1115/1.1582493.

[77] Wang Q. “AN ITERATIVE HDMR APPROACH FOR ENGINEERING RELIABILITY ANALYSIS”. In: *Interdisciplinary Civl and Construction Engineering Project* (2021). ISSN: 2644-108X. DOI: www.doi.org/10.14455/ISEC.2021.8(1).RAD-04.

[78] Herschel Rabitz and Ömer F. Aliş. In: *Journal of Mathematical Chemistry* 25.2/3 (1999), pp. 197–233. ISSN: 0259-9791. DOI: 10.1023/a:1019188517934. URL: http://dx.doi.org/10.1023/A:1019188517934.

[79] Genyuan Li et al. In: *Journal of Mathematical Chemistry* 30.1 (2001), pp. 1–30. ISSN: 0259-9791. DOI: 10.1023/a:1013172329778. URL: http://dx.doi.org/10.1023/A:1013172329778.

[80] Zhixue Tan, Shisheng Zhong, and Lin Lin. “A model learning strategy adapted to health assessment of multi-component systems”. In: July 2017, pp. 1–7. DOI: 10.1109/PHM.2017.8079223.

[81] Donghyun Kim and Ikjin Lee. “Efficient high-dimensional metamodeling strategy using selectively high-ordered kriging HDMR (SH-K-HDMR)”. In: *Journal of Mechanical Science and Technology* 35.11 (Oct. 2021), pp. 5099–5105. ISSN: 1976-3824. DOI: 10.1007/s12206-021-1026-x. URL: http://dx.doi.org/10.1007/s12206-021-1026-x.

[82] B.A. Williams and S. Cremaschi. “Surrogate Model Selection for Design Space Approximation And Surrogatebased Optimization”. In: *Proceedings of the 9th International Conference on Foundations of Computer-Aided Process Design*. Elsevier, 2019, pp. 353–358. DOI: 10.1016/b978-0-12-818597-1.50056-4. URL: http://dx.doi.org/10.1016/B978-0-12-818597-1.50056-4.

[83] Sepp Hochreiter and Jürgen Schmidhuber. “Long Short-Term Memory”. In: *Neural Computation* 9.8 (Nov. 1997), pp. 1735–1780. ISSN: 1530-888X. DOI: 10.1162/neco.1997.9.8.1735. URL: http://dx.doi.org/10.1162/neco.1997.9.8.1735.

[84] Junling Luo et al. “Time series prediction of COVID-19 transmission in America using LSTM and XGBoost algorithms”. In: *Results in Physics* 27 (June 2021), p. 104462. DOI: 10.1016/j.rinp.2021.104462.

[85] Alex Sherstinsky. “Fundamentals of Recurrent Neural Network (RNN) and Long Short-Term Memory (LSTM) network”. In: *Physica D: Nonlinear Phenomena* 404 (Mar. 2020), p. 132306. ISSN: 0167-2789. DOI: 10.1016/j.physd.2019.132306. URL: http://dx.doi.org/10.1016/j.physd.2019.132306.

[86] Shuang Li, Bin Yang, and Fei Qi. “Accelerate global sensitivity analysis using artificial neural network algorithm: Case studies for combustion kinetic model”. In: *Combustion and Flame* 168 (June 2016), pp. 53–64. DOI: 10.1016/j.combustflame.2016.03.028.

[87] Manuel Arias Chao et al. “Aircraft engine run-to-failure dataset under real flight conditions for prognostics and diagnostics”. In: *Data* 6 (1 Jan. 2021), pp. 1–14. ISSN: 23065729. DOI: 10.3390/data6010005.

[88] NASA’s Open Data Portal. *C-MAPSS Aircraft Engine Simulator Data*. URL: https://data.nasa.gov/dataset/C-MAPSS-Aircraft-Engine-Simulator-Data/xaut-bemq/about_data.

[89] Sejal Jalswal. *What is Normalization in Machine Learning? A Comprehensive Guide to Data Rescaling — datacamp.com*. `https://www.datacamp.com/tutorial/normalization-in-machine-learning`. [Accessed 16-12-2024].

[90] Sergey Ioffe and Christian Szegedy. *Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift*. 2015. arXiv: `1502.03167 [cs.LG]`.

[91] Xue Ying. "An Overview of Overfitting and its Solutions". In: *Journal of Physics: Conference Series* 1168 (Feb. 2019), p. 022022. DOI: `10.1088/1742-6596/1168/2/022022`.

[92] Benson, T. - Glenn Research Center, NASA. *Ideal Brayton Cycle*. 2021. URL: `https://www.grc.nasa.gov/www/k-12/BGP/brayton.html` (visited on 06/15/2023).

[93] Hall, N. - Glenn Research Center, NASA. *PV and TS Diagrams*. 2021. URL: `https://www.grc.nasa.gov/www/k-12/airplane/pvtsplot.html` (visited on 06/10/2023).

[94] Afshin Ghajar, Ronald Delahoussaye, and Vandan Nayak. "Development and Implementation of Interactive/Visual Software for Simple Aircraft Gas Turbine Design". In: (Aug. 2023).

[95] Rainer Storn and Kenneth Price. "Differential Evolution - A Simple and Efficient Heuristic for Global Optimization over Continuous Spaces". In: *Journal of Global Optimization* 11 (Jan. 1997), pp. 341–359. DOI: `10.1023/A:1008202821328`.

[96] Swagatam Das and Ponnuthurai Suganthan. "Differential Evolution: A Survey of the State-of-the-Art." In: *IEEE Trans. Evolutionary Computation* 15 (Jan. 2011), pp. 4–31.

A

Thesis Assignment

Enhanced Turbofan Surrogate Modelling for Data Analytics use at KLM ES

MSc Assignment for Walid Brachmi, Propulsion & Power (FPP), Faculty of Aerospace Engineering

Introduction

KLM Engine Services (ES) is part of Air France Industries KLM Engineering & Maintenance Group, overhauling approximately 140 aircraft engines annually. The overhaul shop visit ends with a standardized performance test, to assess compliance to certification rules and customer contracts before release for operation on-wing. At two different locations, the following turbofan engine types are tested:

- CFM56-7B KLM E&M Testcell / Schiphol-Oost
- CF6-80E1 KLM E&M Testcell / Schiphol-Oost
- CF6-80C2 KLM E&M Testcell / Schiphol-Oost
- GEnx-1B Zephyr Testcell / Charles de Gaulle Airport, Paris
- CFMI LEAP-1A & -1B (in gradual introduction)

Over the years, KLM ES Engineering has used GSP (Gas turbine Simulation Program) as a supporting tool to analyze and evaluate engine performance data. Gas Path Analysis (GPA) techniques are used to translate engine performance data into component condition information. For optimal accuracy, parameter inputs from all gas path sensors at the various engine stations are required. Unfortunately, new engine types such as the GEnx and LEAP no longer allow installing additional sensors at the various stations, reducing performance analysis accuracy. However, this can be compensated for by using an accurate and detailed surrogate model, which provides relationships between measured and unmeasured parameters using Continuous Engine Operating Data (CEOD), nowadays collected in-flight. Surrogate models can process large amounts of data in real-time due to their low computational complexity. Therefore KLM ES is continuing development of a surrogate model.

Key objectives

- Conduct a literature study to determine the most effective improvements of the current surrogate model for increased accuracy. Potential topics include:
    - Reynolds corrections for high Mach numbers/high altitude
    - Modelling of Nozzles (velocity coefficients)
    - Correcting for Bleed flow and active flow control systems (VSVs, ACC)
    - Inlet corrections
    - Power off-take
- Develop an enhanced surrogate model for the GEnx-1B and validate it using available testing and/or on-wing data to assess the increased potential of the implemented improvements

Assignment

Your work will include the following elements:

1) A literature study on current trends in using surrogate models for gas turbine performance modelling.
2) Introduction to current KLM performance and condition monitoring practice and relation to the maintenance concept.
3) Introduction to GSP (test analysis and gas path analysis models) and KLM Big Data projects.
4) Assessment of how to best capture engine performance (failure/deterioration modes) with a small set of calculated parameter curves using the available sensor data.
5) Verify the model results using existing GSP models and simulated data.
6) Validation of model using on-wing data and engine maintenance events.

Report

Results of the work must be reported in English, with a copy of this assignment and an executive summary.

Coaching

The work will be performed in close collaboration with KLM Engine Services

Professor,	Delft University supervisor,	Supervisors at KLM
Prof. dr. ir. P. Colonna	Dr. ir. W.P.J. Visser	Juan Regueiro, Tim Rootliep
Piero Colonna	25-5-2023	Tim Rootliep 30-5-2023

B

Gas Turbine Thermodynamics & Cycle Calculations

B.1. Thermodynamics of Gas Turbines

Before moving onto the thermodynamics behind gas turbines, here is a list of relevant terms that are used in this section along with their definitions.

- **Isentropic:** state of constant entropy; entropy does not increase
- **Isobaric:** state of constant pressure
- **Compression:** process of increasing a fluid's pressure
- **Expansion:** process of decreasing a fluid's pressure
- **Adiabatic:** no transfer of heat occurs

Gas turbines operate based on the Brayton cycle, which consists of four basic processes, shown below in Figure B.1 for the case of a simple cycle:

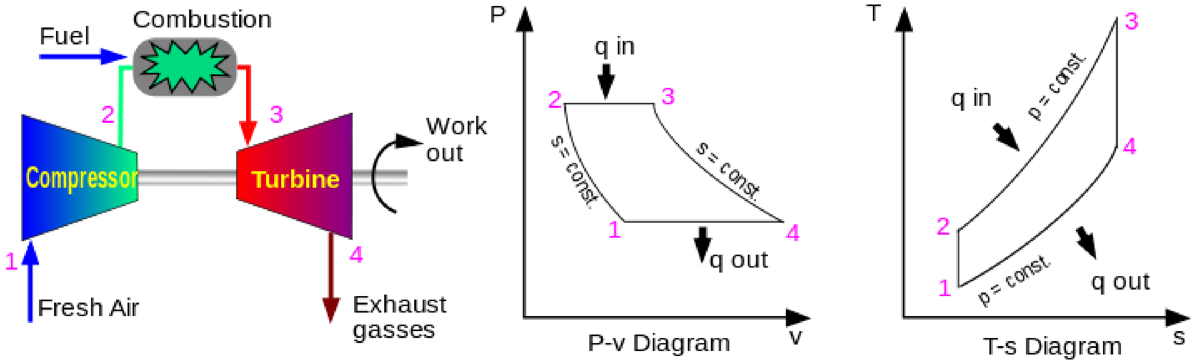


Figure B.1: T-S & P-V diagram of Brayton Cycle [92]

Air is drawn in at the intake (1) and is compressed and exits the compressor (2) at a higher pressure than it entered. Then it moves to the combustion chamber where heat is added, resulting in a hot flow of gas. The hot gas flows into the turbine (3) where it expands, causing the turbine to spin. After that, the gas is exhausted out into the atmosphere (4). Figure B.1 shows a Pressure-Volume (or PV) and Temperature-Entropy (or TS) diagram that describes the same process. In this case, the numbers on these diagrams correspond to the gas path stations shown in the figure on the left of a basic gas turbine machine. With respect to a modern turbofan, as shown in Figure 2.2, the processes shown above occur in the following stations according to the established station convention.

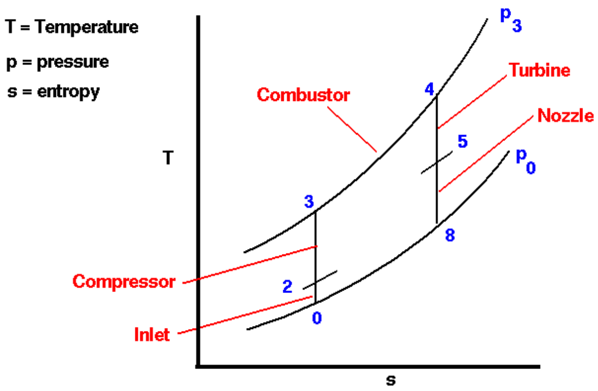


Figure B.2: T-S diagram for turbofans [93]

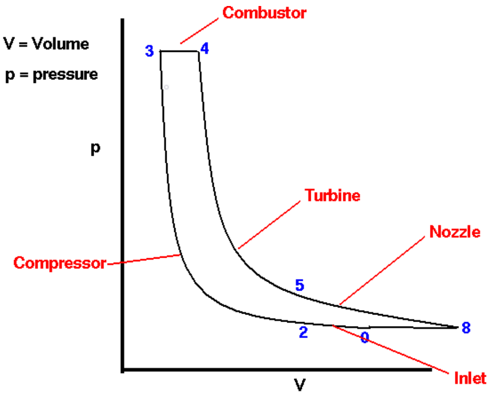


Figure B.3: P-V diagram for turbofans [93]

- **0 -> 2 - Inlet:** isentropic compression
- **2 -> 3 - Fan Core & LPC:** isentropic compression
- **3 -> 4 - Combustor** isobaric heat addition
- **4 -> 5 - Turbines** isentropic expansion
- **5 -> 8 - Exhaust Nozzle:** isentropic expansion
- **8 -> 0 - Atmospheric conditions:** isobaric return to inlet conditions

To clarify, the additional point labeled "gg" between point 4 and 5 represents the point at which the energy needed to drive the core flow has been produced. In the case of a twin-shaft turbofan, the HPT drives the HPC via the high-speed shaft. All the energy extract from the flow by the HPT is then used to drive the HPC and subsequently the core flow. The energy extracted from the LPT also drives the LPC and the core flow but part of the energy is also used to drive the fan and subsequently the bypass flow as well. Therefore, Point "gg" along the gas path lies somewhere in the LPT where its exact position varies depending on engine design parameters, such that the distance between stations 2 and 3 is equal to between 4 and "gg". The remaining power generated, "gg" to 5, is referred to as the specific power, which is the power available to be extracted from the gas generator or the maximum kinetic power to be converted to thrust for an aero-engine [6].

Note that the cycle shown above is in fact a closed cycle, whereas turbofans operate in an open cycle; new air is constantly fed in and exits the nozzle. However, for the sake of thermodynamic analysis and calculations, it is common to refer to the system as a closed cycle. Most importantly, the cycle shown above is the **ideal** Brayton cycle. Ideal cycles take place with no loss of energy or pressure in either process. This can only be possible with a system under these assumptions:

- **Ideal gas:** Working fluid is of constant C_p, C_v and chemical composition
- **Closed cycle:** The ideal cycle is modelled as a closed cycle, where the atmosphere is modelled as a heat exchanger that cools the working fluid along the line 8-0
- **Constant mass flow:** Gas turbine is assumed to operate at steady-state, where the mass flow through it remains constant
- **Isentropic compression & expansion:** No entropy increase during compression and expansion
- **No losses:** The total energy in the system is assumed constant and none is lost to the outside
- **No pressure losses:** Pressure is constant during the heat addition in the chamber, thus those points lie on the same constant pressure line
- **No mechanical losses;** No energy is lost in the transmission of power from the turbine to the compressor via the rotating shaft

In practice, real thermodynamic cycles can never achieve isentropic processes, therefore entropy generation at each process is almost inevitable. Steps considered to be isentropic in fact still generate entropy such as compression, as seen below in Figure B.4.

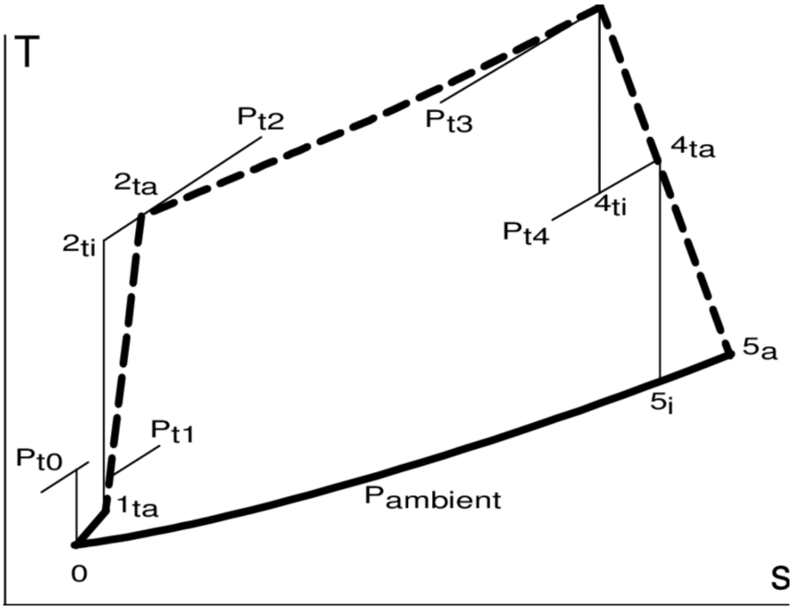


Figure B.4: T-S diagram of Brayton Cycle [94]

For real cycles, the assumptions previously outlined in the ideal case are now longer valid. Energy losses are now present, generally in the form of heat radiating outside the system or fuel particles left in the exhaust flow due to incomplete combustion. Energy losses are characterized by entropy generation, which can be seen for example in the compression stage, initially assumed isentropic now generates entropy; the line is no longer vertical thus $\Delta s \neq 0$. Additionally, the heat addition process in the combustion chamber, previously assumed isobaric, now experiences pressure losses due to friction of the working fluid with the walls; points 2 and 3 no longer lie on the constant pressure line. Mechanical losses within the shafts are now present, resulting in more energy needed from the LPT and HPT to keep driving the flows and point "gg" to shift downwards.

Losses occurring within a real Brayton cycle can be represented by the efficiency of each major component within the cycle. Component efficiencies provide a way to quantify the difference between the work extracted from the real cycle and that expected in an ideal case.

B.2. Cycle Calculations

In this section, the general cycle calculation scheme is outlined for the case of a twin-shaft turbofan running at steady-state. This calculation is considered to be 0-D; values at each engine station do not depend on any spatial dimension. Furthermore, the gas is assumed to be ideal with constant C_p & C_v, however different values of C_p will be used for fresh air entering and flue gas exiting after combustion, denoted as $C_{p,a}$ and $C_{p,g}$ respectively.

B.2.1. Relevant Terms

Before outlining the calculations scheme, it is worth introducing the relevant terms, most of which are important engine properties that are assumed to be known, such as the isentropic efficiencies and pressure ratios of the relevant gas path components, along with ambient conditions and gas properties.

The first set of relevant terms are the component efficiencies. these terms are used to quantify the losses in the context of a non-ideal cycle. Losses in real cycles are quantified in entropy increase, thus every major component is attributed what is referred to as "isentropic" efficiency; a measure of how far it deviates from ideal conditions with no entropy generated. Isentropic efficiency is defined as the ratio of the work generated by a given component in ideal isentropic conditions, to the real work delivered.

Starting with the first component along the gas path, the fan, the isentropic efficiency of the fan is defined as follows in Equation B.1:

$$\eta_{is,fan} - \frac{\dot{W}_{is}}{\dot{W}} = \frac{(\dot{m}c_{p,a}\Delta T)_{is}}{\dot{m}c_{p,a}\Delta T} = \frac{\Delta T_{is}}{\Delta T} = \frac{T_{t,21,is)-T_{t,2}}}{T_{t,21} - T_{t,2}} \; [-] \tag{B.1}$$

...where η denotes the efficiency, $\dot{W}$ the work generated, $\dot{m}$ the mass flow and T the temperature. Similarily to the fan, the isentropic efficiency for the following components such as the LPC and HPC are defined below on the left in Equation B.2 and Equation B.3 respectively.

$$\eta_{is,LPC} = \frac{T_{t,25,is} - T_{t,21}}{T_{t,25} - T_{t,21}} \; [-] \tag{B.2}$$

$$\eta_{is,HPT} = \frac{T_{t,49} - T_{t,4}}{T_{t,49,is} - T_{t,4}} \; [-] \tag{B.4}$$

$$\eta_{is,HPC} = \frac{T_{t,3,is} - T_{t,25}}{T_{t,3} - T_{t,25}} \; [-] \tag{B.3}$$

$$\eta_{is,LPT} = \frac{T_{t,5} - T_{t,49}}{T_{t,5,is} - T_{t,49}} \; [-] \tag{B.5}$$

On the right, in Equation B.3 and Equation B.5, are the isentropic efficiencies of the HPT and LPT. For the turbines, they are defined similarily to the compressor stages, however they are inverted. Since the turbines extract energy from the flow, the usable work extracted is lower than the total work available in the flow. Therefore, the temperature drop in the ideal case is larger than that in the real case.

In the Joule-Brayton cycle, there is a heat addition stage, which is handled by the Combustion Chamber (CC). The job of the combustor is to transfer heat energy in the form of fuel to the flow. The efficiency of the combustor is defined as the ratio of the heat energy absorbed by the flow to the heat energy contained in the fuel. The combustor efficiency is shown below in Equation B.6.

$$\eta_{CC} = \frac{\dot{Q}}{\dot{Q}_f} = \frac{\dot{m}c_{p,g}(T_{t,4} - T_{t,3})}{\dot{m}_f \cdot LHV_f} \tag{B.6}$$

As the flow exits the turbofan, it reaches the nozzles. For a turbofan, the flow is separated in two directions, as shown in Figure 2.1. The flow passing through the LPC, HPC, HPT and LPT is the core flow, however the larger fraction of the air at the inlet is propelled by the fan through the bypass duct. Since the flow is expanding at that point, the nozzle must also be designed to allow the gas to expand efficiently for maximum thrust output. Therefore, nozzles do also contribute losses in certain conditions. Given expansion is involved, nozzle efficiencies are defined similarly to the turbine stages.

$$\eta_{is,noz,core} = \frac{T_{t,8} - T_{t,7}}{T_{t,8,is} - T_{t,7}} \; [-] \tag{B.7}$$

$$\eta_{is,noz,bypass} = \frac{T_{t,18} - T_{t,17}}{T_{t,18,is} - T_{t,17}} \; [-] \tag{B.8}$$

Furthermore, mechanical losses must involved be accounted for. The most prominent mechanical losses involve the shafts. In twin-shaft (or two-spool) turbofans, the HPT and LPT are connected and drive the HPC and LPC respectively via the high-speed (N2) and low-speed (N1) shafts. The associated mechanical efficiency is defined as the ratio of work delivered to the work generated by the respective turbines. Note that a negative sign is present given that the turbine work terms are defined as negative since they extract energy out of the flow as opposed to the compressor stages which transfer energy into the flow.

$$\eta_{mech,N1} = -\frac{W_{fan} + W_{FPC}}{W_{LPT}} \tag{B.9}$$

$$\eta_{mech,N2} = -\frac{W_{HPC}}{W_{HPT}} \tag{B.10}$$

B.2.2. Cycle Calculations Framework

For the sake of thermodynamic calculations, it is generally more convenient to work with total quantities, such as total pressure and temperature. Total quantities are derived from the energy balance of a gas turbine. If potential energy changes are neglected, leaving only kinetic energy, then the energy balance reduces to the total enthalpy in the gas turbine, as shown below:

$$0 = h_2 - h_1 + \frac{1}{2}(0 - v_1^2) \;\rightarrow\; h_0 = h + \frac{1}{2}v^2$$

Total enthalpy h_0 is defined as the enthalpy of a fluid if it was adiabatically brought to rest with no change of work [75]. The advantage of using this property is that kinetic energy is now implicitly included and does not need to be explicitly accounted for in further calculations, thus reducing their complexity when modelling. Based on total enthalpy. the same advantage can be extended to other properties such as temperature and pressure, which are related as follows:

$$h_0 = c_p \cdot T + \frac{1}{2}v^2 \qquad T_0 = T + \frac{v^2}{2 \cdot c_p} \qquad p_0 = p \cdot \frac{T_0}{T}^{\frac{\gamma}{\gamma-1}}$$

Intake

For the sake of cycle calculations it is often more convenient to use total gas path quantities rather than static quantities, specifically for pressure and temperature. The first step is then to convert the inlet conditions generally given as static qunatities to total quantities as shown below:

$$p_{t,0} = p_{s,0} \cdot \left(1 + \frac{\gamma_a - 1}{2}M^2\right)^{\frac{\gamma_a}{\gamma_a - 1}} \quad \text{(B.11)} \qquad T_{t,0} = T_{s,0} \cdot \left(1 + \frac{\gamma_a - 1}{2}M^2\right) \quad \text{(B.12)}$$

Once the inlet conditions (Station 0) have been converted to total quantities, the conditions after the inlet or at the fan face (Station 2) are computed as shown below. The engine intake is generally assumed ideal (ie: isentropic) for the sake of these calculations, thus the total temperature does not change.

$$p_{t,2} = PR_{in} \cdot p_{t,0} \qquad T_{t,2} = T_{t,0}$$

Fan

Once through the fan, the flow is compressed by a relatively small amount. Across the fan, it is assumed for 0-D calculations that the flow is homogeneous, thus the conditions after the fan for the fraction of the flow entering the core section is the same as the fraction of the flow heading towards the bypass ducts. For now, we start with the part heading to the core flow (Station 21). More importantly, the temperature now changes due to the compression done by the fan given that this is a real cycle thus no process is considered isentropic.

$$p_{t,21} = PR_{21} \cdot p_{t,2} \qquad T_{t,21} = T_{t,2} \cdot \left(1 + \frac{1}{\eta_{is,fan}} \cdot \left(\left(\frac{p_{t,21}}{p_{t,2}}^{\frac{\gamma_a - 1}{\gamma_a}}\right) - 1\right)\right)$$

Since the fan can be considered a compressor, it has transfered work to the flow. The work done on the flow by the fan can be computed using the difference in total temperatures across the fan. $\dot{W}_{fan}$ will be relevant in later stages when power balances are involved.

$$\dot{W}_{fan} = \dot{m} \cdot c_p \Delta T = \dot{m}_0 \cdot c_{p,a} \cdot (T_{t,21} - T_{t,2})$$

LPC

Following the core gas path after the fan, the flow arrives at the LPC inlet (Station 21). After crossing the LPC, the pressure increases and so does the temperature given that the compression generated entropy. The pressure and temperature at the LPC exit (Station 25) is shown below:

$$p_{t,25} = PR_{LPC} \cdot p_{t,21} \qquad T_{t,25} = T_{t,21} \cdot \left(1 + \frac{1}{\eta_{is,fan}} \cdot \left(\left(\frac{p_{t,25}}{p_{t,21}}^{\frac{\gamma_a - 1}{\gamma_a}}\right) - 1\right)\right)$$

Similar to the fan, the LPC also delivered work to the flow, which is computed using the mass flow, specific heat and temperature difference across the LPC. At this stage, it is also worth noting that the core mass flow,$\dot{m}_{21}$, is computed as shown below based on the turbofan's bypass ratio. The other mass flow fraction was diverted to the bypass duct, shown later.

$$\dot{m}_{21} = \frac{\dot{m}_0}{BPR + 1} \qquad \dot{W}_{LPC} = \dot{m}_{21} \cdot c_{p,a} \cdot (T_{t,25} - T_{t,21})$$

HPC

After crossing the LPC, the flow arrives at the HPC inlet (Station 3) and crosses to the HPC exit, also known as the combustion chamber inlet (Station 3). The process is the same as through the LPC, except with a higher pressure increase given it is the HPC.

$$p_{t,3} = PR_{HPC} \cdot p_{t,25} \qquad T_{t,3} = T_{t,25} \cdot \left(1 + \frac{1}{\eta_{is,fan}} \cdot \left(\left(\frac{p_{t,3}}{p_{t,25}}^{\frac{\gamma_a - 1}{\gamma_a}}\right) - 1\right)\right)$$

$$\dot{W}_{HPC} = \dot{m}_{25} \cdot c_{p,a} \cdot (T_{t,3} - T_{t,25})$$

Combustor

The pressure rise in the combustor is relatively straightforward to compute. The temperature rise however is computed based on the energy balance across the combustor.

$$p_{t,4} = PR_{CC} \cdot p_{t,3} \qquad T_{t,4} = T_{t,3} + \frac{\dot{m}_f LHV_f}{\dot{m}_3 c - p, g}$$

HPT

After the combustion chamber, the flow arrives at the turbine inlet (Station 4). The temperature drop experienced across the turbines is computed based on the enrgy balance. In this case, work is now extracted from the flow and drives the shafts. Therefore, the sum of the work extracted from the HPT and the work done by the HPC should be zero, accounting for mechanical losses along the N2 shaft.

$$\dot{W}_{HPT} + \frac{\dot{W}_{HPC}}{\eta_{mech,N2}} = 0$$

With this power balance, the gas path conditions after the HPT (Station 49) is computed similarly to the previous stations. Note that after the combustion chamber, $c_{p,g}$ is used given that flue gas from the combustion process are now present in the flow.

$$p_{t,49} = p_{t,4} \cdot \left(\frac{1}{\eta_{is,HPT}} \cdot \left(\frac{T_{t,49}}{T_{t,4}} - 1\right) + 1\right)^{\frac{\gamma_g}{\gamma_g - 1}} \qquad T_{t,49} = T_{t,4} \cdot \frac{\dot{W}_{HPT}}{c_{p,g} \cdot (\dot{m}_3 + \dot{m}_f)}$$

LPT

After the HPT, the flow is brought to a lower pressure and continues to expand in the LPT. The cycle calculations for the LPT are almost the same as the HPT, except for the power balance now involving the LPC and the fan, given both are dirven by the N1 shaft.

$$\dot{W}_{LPT} + \frac{\dot{W}_{fan} + \dot{W}_{LPC}}{\eta_{mech,N1}} = 0$$

$$p_{t,5} = p_{t,49} \cdot \left(\frac{1}{\eta_{is,HPT}} \cdot \left(\frac{T_{t,5}}{T_{t,49}} - 1\right) + 1\right)^{\frac{\gamma_g}{\gamma_g - 1}} \qquad T_{t,49} = T_{t,4} \cdot \frac{\dot{W}_{HPT}}{c_{p,g} \cdot (\dot{m}_3 + \dot{m}_f)}$$

Exhaust

At this stage, the flow ix exiting the core and arriving at the nozzles, after which the core flow and bypass flow meet again and flow out to the atmosphere. Based on the flow conditions at the nozzle, the core and bypass nozzle contrinbutions to the resulting thrust can be computed. Depending on the nozzle design, the flow can be in either two flow conditions. The critical pressure ratio PR_{crit} is a condition that helps distinguish whether the flow at a nozzle is either choked or not.

$$\frac{p_{t,7}}{p_{t,0}} \geq PR_{crit} = \left(1 - \frac{1}{\eta_{is,nozz,core}} \cdot \frac{\gamma_g - 1}{\gamma_g + 1}\right)^{\frac{\gamma_g}{1 - \gamma_g}}$$

$$\frac{p_{t,18}}{p_{t,0}} \geq PR_{crit} = \left(1 - \frac{1}{\eta_{is,nozz,bypass}} \cdot \frac{\gamma_a - 1}{\gamma_a + 1}\right)^{\frac{\gamma_a}{1-\gamma_a}}$$

$$F_{core} = (\dot{m}_3 + \dot{m}_f) \cdot (v_8 - v_0) + A_8 \cdot (p_{s,8} - p_{s,0})$$

$$F_{bypass} = \frac{BPR \cdot \dot{m}_0}{BPR + 1} \cdot (v_18 - v_0) + A_18 \cdot (p_{s,18} - p_{s,0})$$

For unchoked nozzle If the ratio of the total pressure at the nozzle to the ambient conditions is lower than the critical pressure ratio, then the flow is considered unchoked. The flow can still expand fully to ambient conditions, where the flow's static pressure matches the ambient. In this case, no pressure term is present in the thrust contribution given that no pressure difference is present between the exhaust flow and the ambient pressure.

$$T_{s,8} = T_{t,7} \cdot \left(1 - \eta_{is,noz,core} \cdot \left(\frac{p_{s,8}}{p_{t,7}}\right)^{\frac{\gamma_g - 1}{\gamma_g}}\right) \qquad T_{s,18} = T_{t,17} \cdot \left(1 - \eta_{is,noz,bypass} \cdot \left(\frac{p_{s,18}}{p_{t,17}}\right)^{\frac{\gamma_a - 1}{\gamma_a}}\right)$$

Thus the thrust is a result of the high exhaust velocity and can be calculated using the difference between the total and static temperatures as shown below for each nozzle.

$$v_8 = \sqrt{2c_{p,g}(T_{t,7} - T_{s,8})} \qquad v_{18} = \sqrt{2c_{p,a}(T_{t,17} - T_{s,18})}$$

For choked nozzle If the ratio of the total pressure at the nozzle to the ambient conditions is greater than the critical pressure ratio, then the flow is considered choked. In the choked condition, the flow reaches sonic conditions (Mach 1), forming a shock front in the nozzle. The flow is unable to fully expand to ambient conditions, therefore a difference in pressure between the flow and ambient conditions is present, resulting in a pressure in the thrust contribution. The flow velocity exiting the nozzle is now the speed of sound and the pressure ratio over the nozzle is equal to the critical pressure ratio. With this information, the temperature at the nozzle can be calculated as shown below.

$$\begin{aligned} T_{s,8} &= T_{t,7}\frac{2}{\gamma_g + 1} \\ v_8 &= \sqrt{\gamma_g R T_{s,8}} \\ A_8 &= \frac{\dot{m}_8 R T_{s,8}}{v_8} \cdot \frac{PR_{crit}}{p_{s,8}} \end{aligned} \qquad \begin{aligned} T_{s,18} &= T_{t,17}\frac{2}{\gamma_a + 1} \\ v_{18} &= \sqrt{\gamma_a R T_{s,18}} \\ A_{18} &= \frac{\dot{m}_8 R T_{s,18}}{v_18} \cdot \frac{PR_{crit}}{p_{s,18}} \end{aligned}$$

C

Optimization using Differential Evolution

During the parameterization of the Secondary Performance Parameters, an optimization was conducted in order to obtain appropriate coefficients for the quadratic relations meant to approximate the effects of SPPs. The SPPs analyzed in this thesis were the VSV actuator position and the valve position of the active clearance control system for both the Low and High Pressure Turbine. Section C.1 provides a brief explanation of the differential evolution algorithm used for this thesis, while Section C.2 provides the additional results from said optimization.

C.1. Differential Evolution

Originally introduced by Storn and Price [95], Differential evolution (or DE) is classified as an evolutionary algorithm, generally used in optimization problems. Similar to genetic algorithms, which work form principles inspired by nature, the DE algorithm starts with a set of candidate solutions and improving them at each iteration through an evolutionary process [96]. The advantage of DE is that it is a gradient-free optimizer; it does not use any gradient descent method, thus the optimization problem does not need to be differentiable which is why it is often used for global optimization problems. The DE algorithm involves four distinct steps, analogous of natural evolution and similar to genetic algorithms. These steps are outlined below in order of operation.

Initialization

The algorithm begins with an population of target vectors $x_{i,G}$, where $i = 1, 2, 3, ..., NP$ and NP represents the population size set by the user. In the context of this thesis, the target vectors contain the coefficients a and b, for the quadratic relations for each of the SPPs. This results in target vectors with 6 elements each. An initial evaluation of the objective function is made, which then leads to the next step.

Mutation

Following initialization, the target vectors in the population undergo what is called mutation. For each target vector, a mutant of each vector $v_{i,G+1}$ is created using the equation shown below, where F is referred to as the amplification factor set by the user which influences the degree of mutation and ranges between 0 and 1.

$$v_{i,G+1} = x_{r_1,G} + F \cdot (x_{r_2,G} - x_{r_3,G}) \ , \ i = 1, 2, .., NP$$

$$r_1, r_2, r_3 \in [1, 2, ..., NP] \ , \ r_1 \neq r_2 \neq r_3$$

The terms r_1, r_2, r_3 are randomly selected integers such that they are not equal to each other, which represent the index of the individual target vectors to be combined to create the resulting mutant vector. Once the mutant vectors are generated, the algorithm conducts the crossover.

Crossover

Crossover involves combining the target and mutant vectors to create new trial vectors. This is done to ensure a certain level of diversity in the variables held in each individual vector in the following generations. Crossover is performed on the mutant vector $v_{i,G+1}$ and target vector $x_{j,i,G}$ to then form the trial vector $u_{j,i,G+1}$.

$$u_{j,i,G+1} = (u_{1,i,G+1}, u_{2,i,G+1}, ..., u_{D,i,G+1})$$

$$u_{j,i,G+1} = \begin{cases} v_{j,i,G+1}, & \text{if } rand\,[0,1] \leq CR \vee j = k \\ x_{j,i,G}, & \text{otherwise} \end{cases}$$

The decision to swap elements from the mutated to a target vector depends on the choice of the crossover constant, CR ranging between 0 and 1, which is also set by the user.

Selection

After conducting the crossover, an evaluation of the objective function is conducted using these trial vectors in order to determine if they should be included in the next generation. The algorithm follows the greedy approach; if the objective function value obtained from the trial vectors is smaller than that of the current vector, then the current vector is replaced by the trial vector and continues to the next generation.

$$X_{i,G+1} = \begin{cases} U_{i,G+1}, & \text{if } f\left(U_{i,G+1}\right) \leq f(X_{i,G}) \\ X_{i,G}, & \text{otherwise} \end{cases}$$

Throughout the optimization, the steps involving mutation, crossover and selection are repeated based on the set number of iterations until a termination criteria is reached. In this case the termination criteria was the number of iterations.

C.2. Optimization History

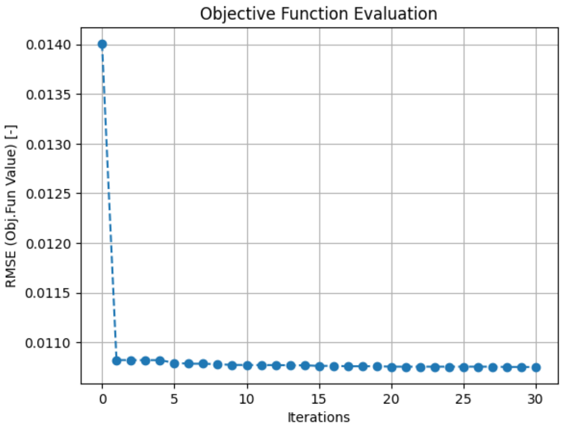


Figure C.1: History of objective function value per iteration

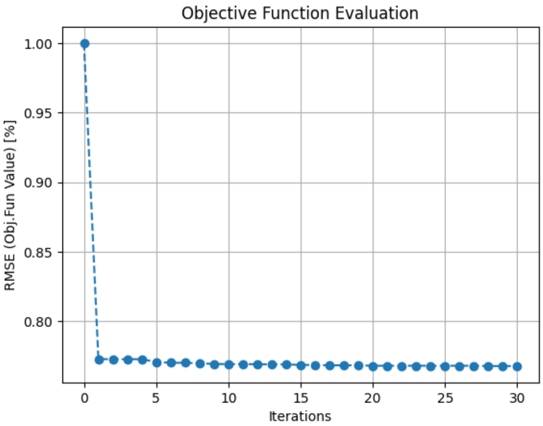


Figure C.2: Obj.Fun history in percentage relative to $0th$ iteration

Table C.1: Coefficients for SPP relations obtained from optimization

Coefficient	$\Delta\eta$ **HPC**	ΔW_c **HPC**	$\Delta\eta$ **HPT**	ΔW_c **HPT**	$\Delta\eta$ **LPT**	ΔW_c **LPT**
a	11.943	-45.503	$2.610 \cdot 10^{-3}$	$-1.602 \cdot 10^{-2}$	$4.939 \cdot 10^{-3}$	$-5.095 \cdot 10^{-3}$
b	10.313	-3.068	$1.676 \cdot 10^{-2}$	$-9.605 \cdot 10^{-4}$	$8.287 \cdot 10^{-2}$	-0.397

D

Additional Results from Enhanced HDMRs

Resulting RMSEs

Table D.1: Table of original RMSEs for GEnx-1B of each HDMR type, as reported by de Bruin [5]

RMSE	$\Delta\eta$	ΔW_c	ΔEGT	$\Delta TSFC$	ΔW_f
Fan+LPC	0.15518	0.01054	0.01537	0.02252	0.01301
HPC	0.02852	0.02820	0.02305	0.02211	0.02053
HPT	0.02798	0.00119	0.01962	0.01944	0.01842
LPT	0.05749	0.09633	0.11726	0.05059	0.05343

Table D.2: Table of RMSEs for GEnx-1B HDMRs with Fan Speed dependency implemented

RMSE	$\Delta\eta$	ΔW_c	ΔEGT	$\Delta TSFC$	ΔW_f
Fan+LPC	0.15560	0.01692	0.01323	0.00753	0.01280
HPC	0.02942	0.02859	0.02596	0.02315	0.02715
HPT	0.03024	0.00115	0.02134	0.01596	0.02071
LPT	0.06591	0.10412	0.11744	0.05250	0.05234

Table D.3: Table of Percentage change in RMSEs for GEnx-1B HDMRs with Fan Speed dependency implemented

RMSE	$\Delta\eta$	ΔW_c	ΔEGT	$\Delta TSFC$	ΔW_f
Fan+LPC	+0.6637	+60.246	-13.728	-66.341	+0.6149
HPC	+4.8737	+1.9858	+11.409	+9.4979	+29.322
HPT	+8.2201	-0.8403	+11.467	-13.631	+17.155
LPT	+16.263	+8.3566	+0.3667	+13.619	-0.8983

Table D.4: Table of RMSEs for GEnx-1B HDMRs with varying N_1 & PTO

RMSE	$\Delta\eta$	ΔW_c	ΔEGT	$\Delta TSFC$	ΔW_f
Fan+LPC	0.15621	0.01689	0.01326	0.00758	0.01309
HPC	0.02991	0.02876	0.02568	0.02421	0.02655
HPT	0.03028	0.00118	0.02187	0.01679	0.02158
LPT	0.06684	0.10438	0.11769	0.05748	0.05295

Table D.5: Table of Percentage change in RMSEs for GEnx-1B HDMRs with varying N_1 & PTO compared to original results

RMSE	$\Delta\eta$	ΔW_c	ΔEGT	$\Delta TSFC$	ΔW_f
Fan+LPC	+0.6637	+60.246	-13.728	-66.341	+0.6149
HPC	+4.8737	+1.9858	+11.409	+9.4979	+29.322
HPT	+8.2201	-0.8403	+11.467	-13.631	+17.155
LPT	+16.263	+8.3566	+0.3667	+13.619	-0.8983

Table D.6: Table of RMSEs for GEnx-1B HDMRs with Fan Speed dependency, PTO and SPPs implemented

RMSE	$\Delta\eta$	ΔW_c	ΔEGT	$\Delta TSFC$	ΔW_f
Fan+LPC	0.15563	0.01703	0.01701	0.01445	0.02300
HPC	0.02807	0.02769	0.02547	0.03124	0.03501
HPT	0.03140	0.00472	0.02148	0.02007	0.02319
LPT	0.07120	0.10592	0.10450	0.04397	0.04685

Table D.7: Table of Percentage change in RMSEs for GEnx-1B HDMRs with Fan Speed dependency, PTO and SPPs implemented, compared to original results

RMSE	$\Delta\eta$	ΔW_c	ΔEGT	$\Delta TSFC$	ΔW_f
Fan+LPC	+0.2899	+61.574	+10.670	-35.834	+76.787
HPC	-1.5778	-1.8085	+10.498	+41.293	+70.530
HPT	+12.233	+296.63	+9.4801	+3.2407	+25.895
LPT	+23.847	+9.9553	-10.881	-13.085	-12.315

Trending with Updated HDMRs

Fan+LPC

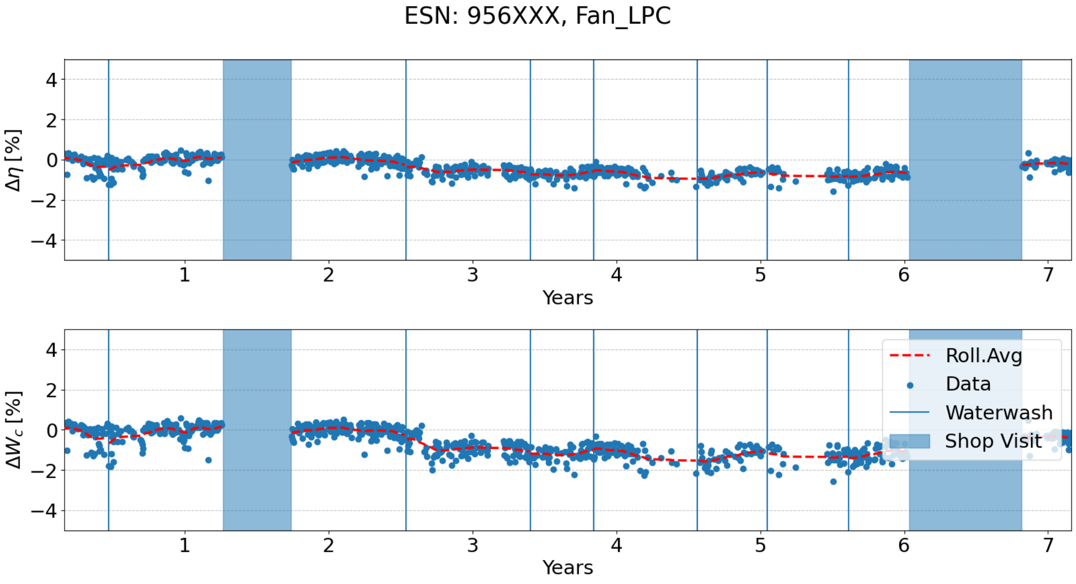


Figure D.1: HDMR trending using on-wing data from ESN956XXX for the Fan+LPC Degradations

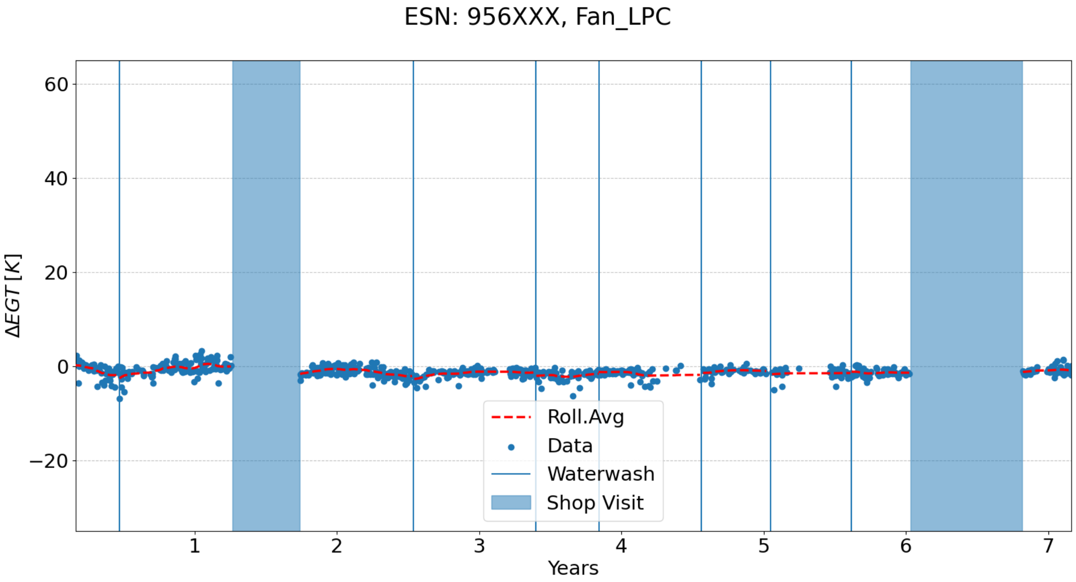


Figure D.2: HDMR trending using on-wing data from ESN956XXX for the Fan+LPC ΔEGT

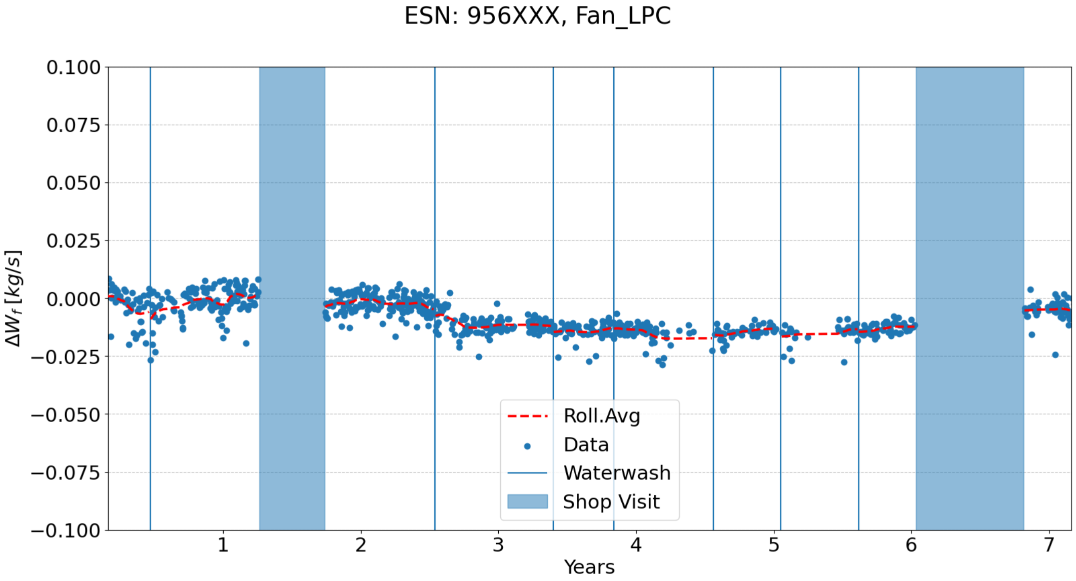


Figure D.3: HDMR trending using on-wing data from ESN956XXX for the Fan+LPC ΔW_f

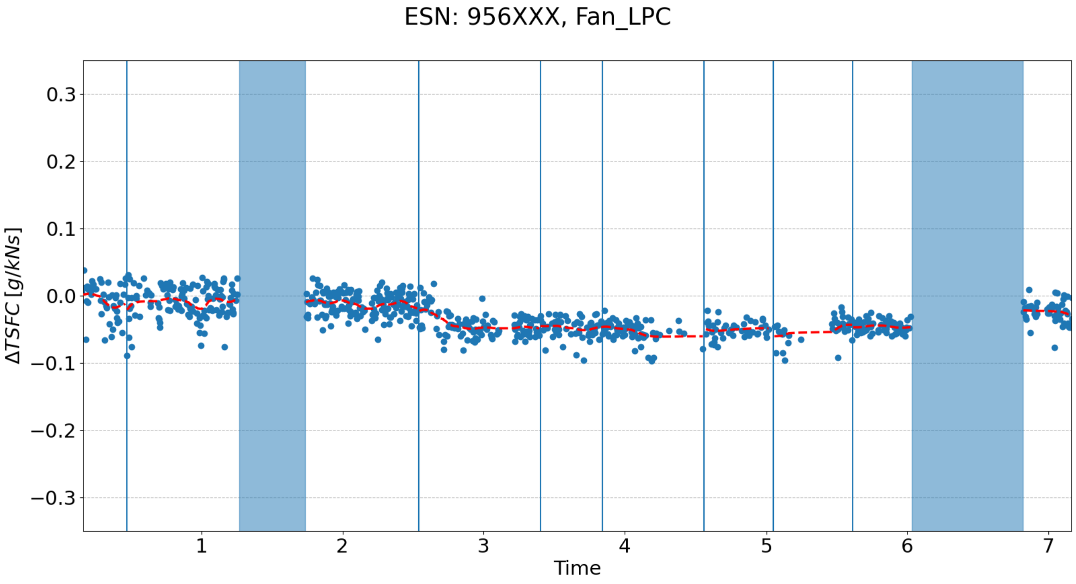


Figure D.4: HDMR trending using on-wing data from ESN956XXX for the Fan+LPC ΔW_f

HPC

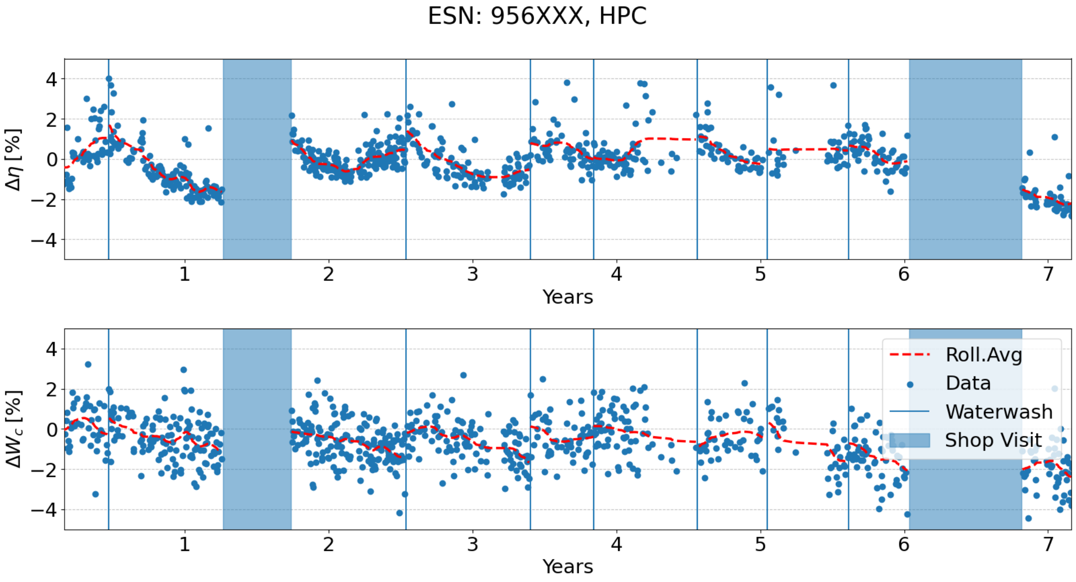


Figure D.5: HDMR trending using on-wing data from ESN956XXX for the HPC Degradations

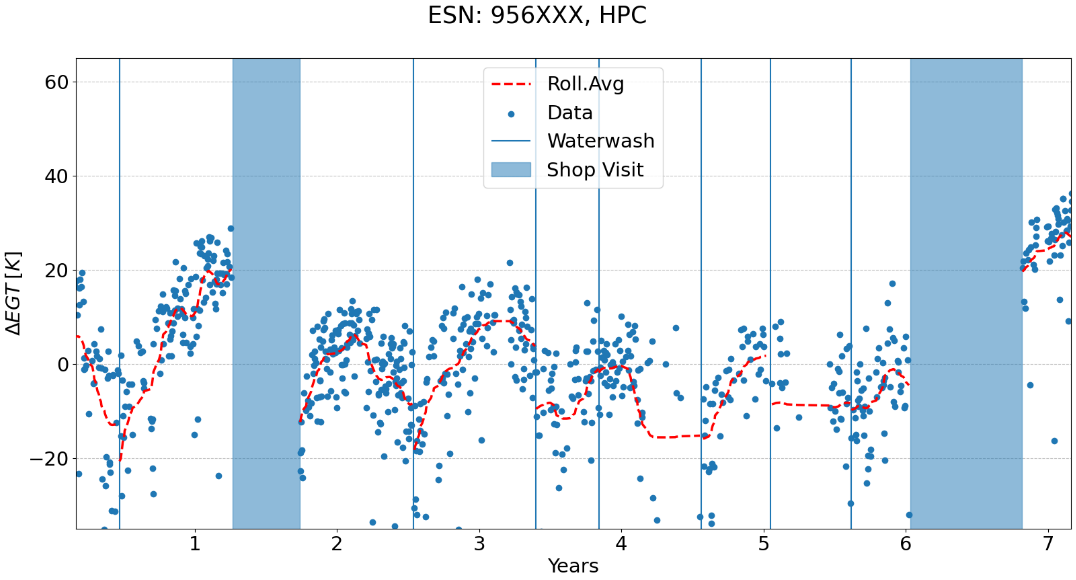


Figure D.6: HDMR trending using on-wing data from ESN956XXX for the HPC ΔEGT

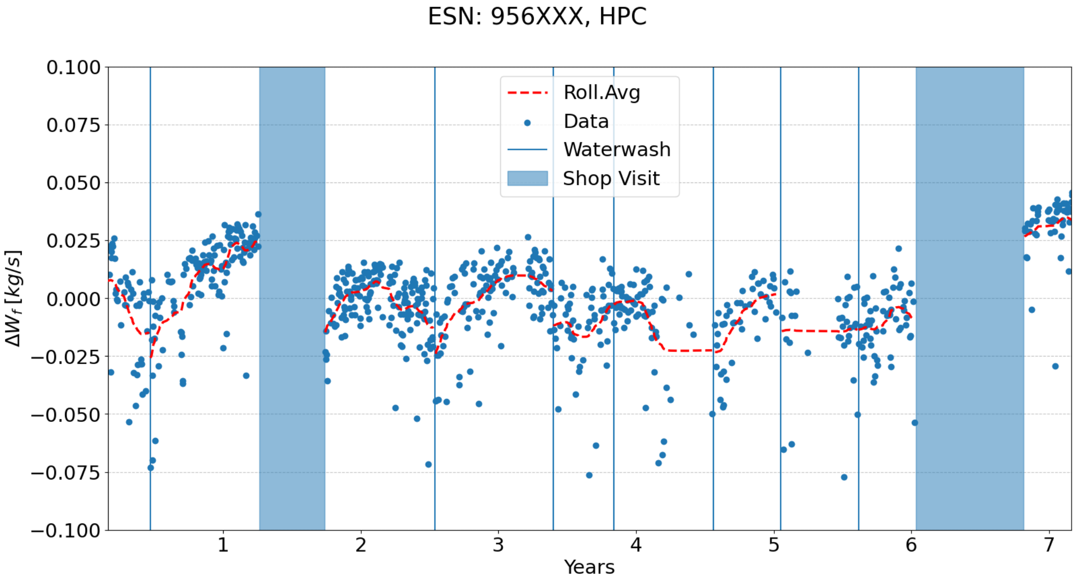


Figure D.7: HDMR trending using on-wing data from ESN956XXX for the HPC ΔW_f

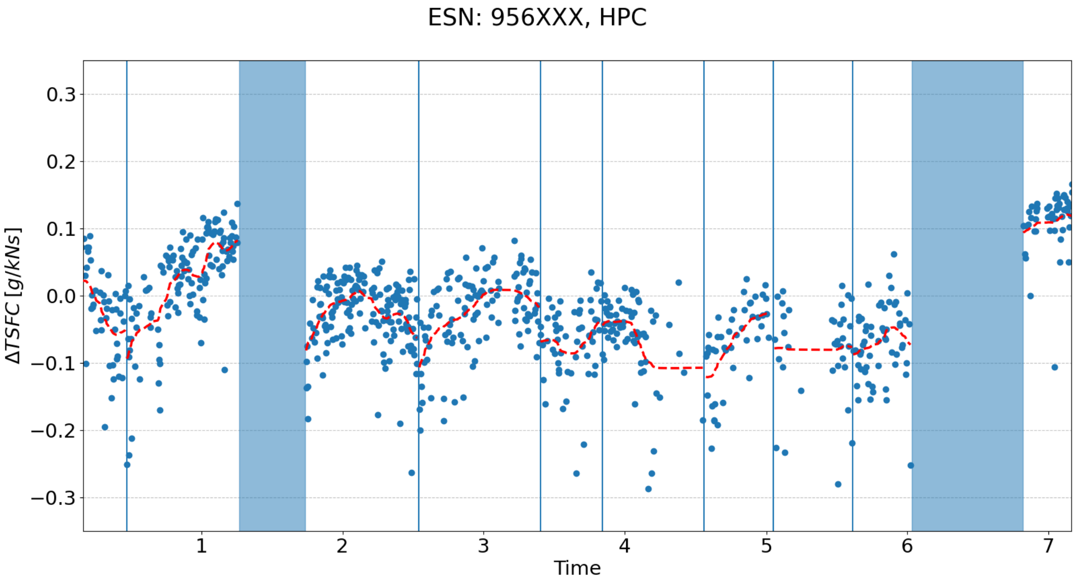


Figure D.8: HDMR trending using on-wing data from ESN956XXX for the HPC ΔW_f

HPT

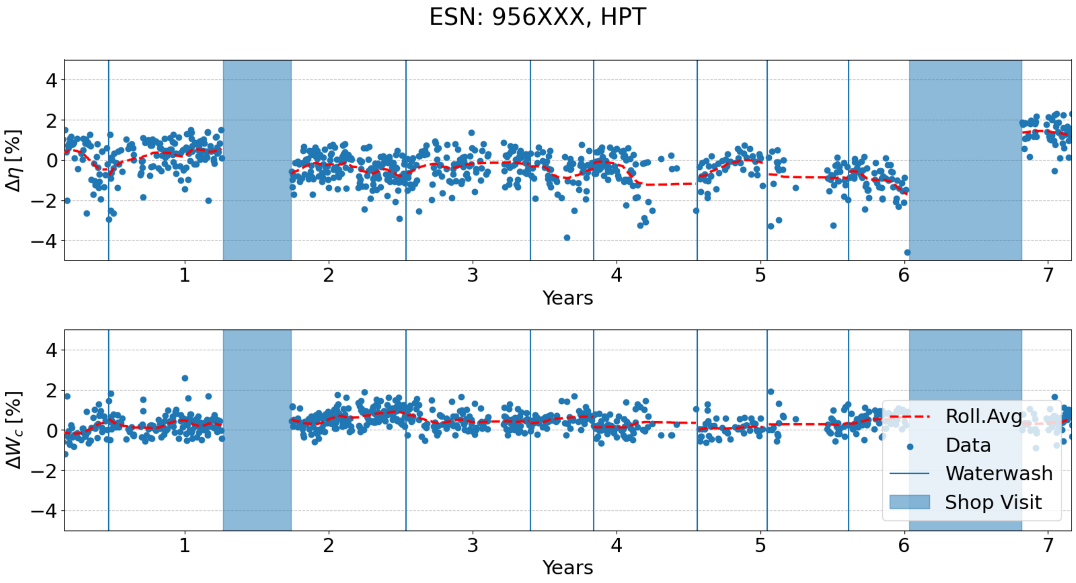


Figure D.9: HDMR trending using on-wing data from ESN956XXX for the HPT Degradations

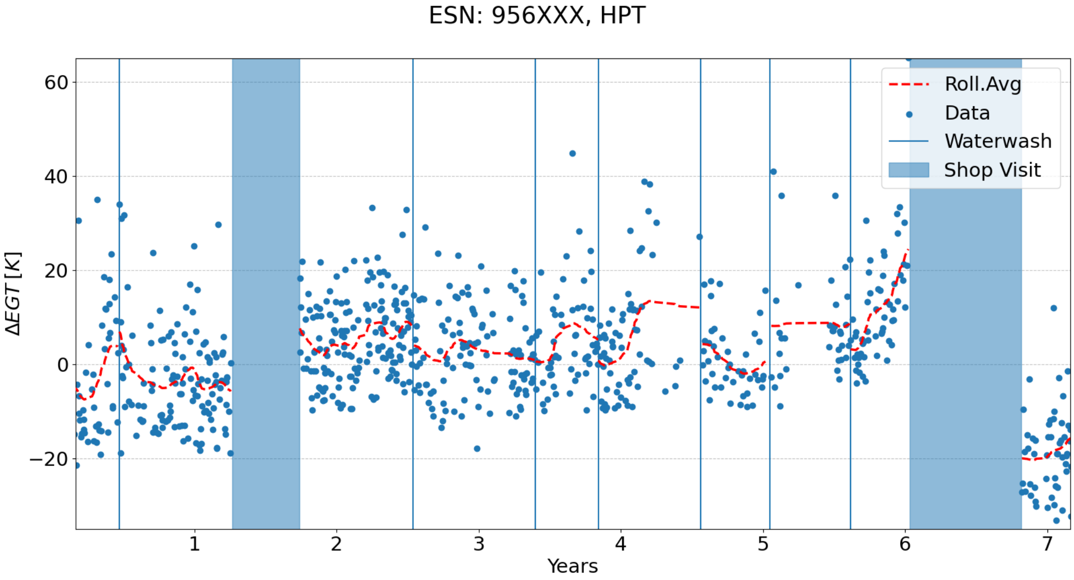


Figure D.10: HDMR trending using on-wing data from ESN956XXX for the HPT ΔEGT

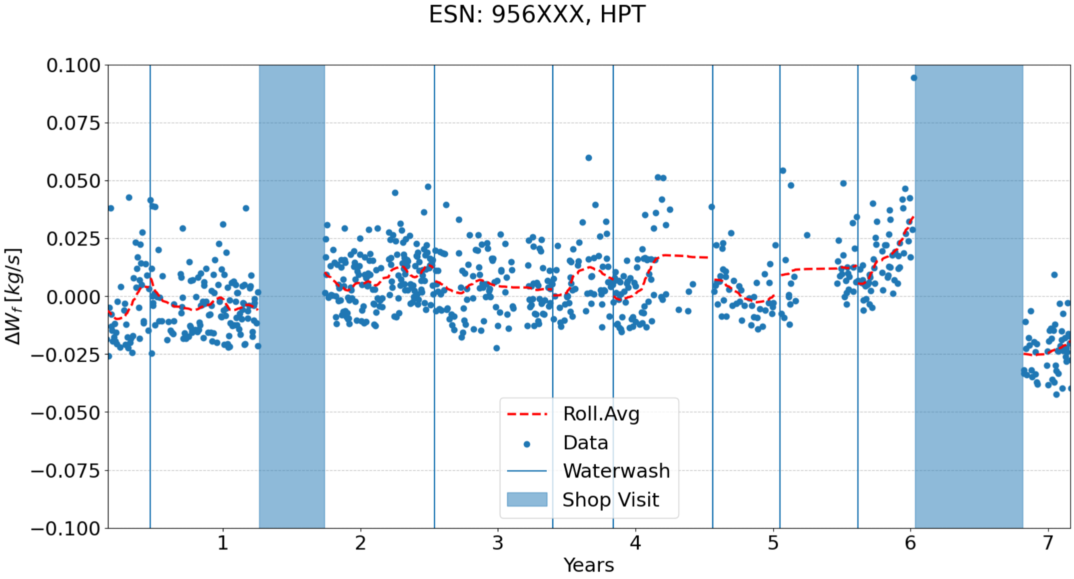


Figure D.11: HDMR trending using on-wing data from ESN956XXX for the HPT ΔW_f

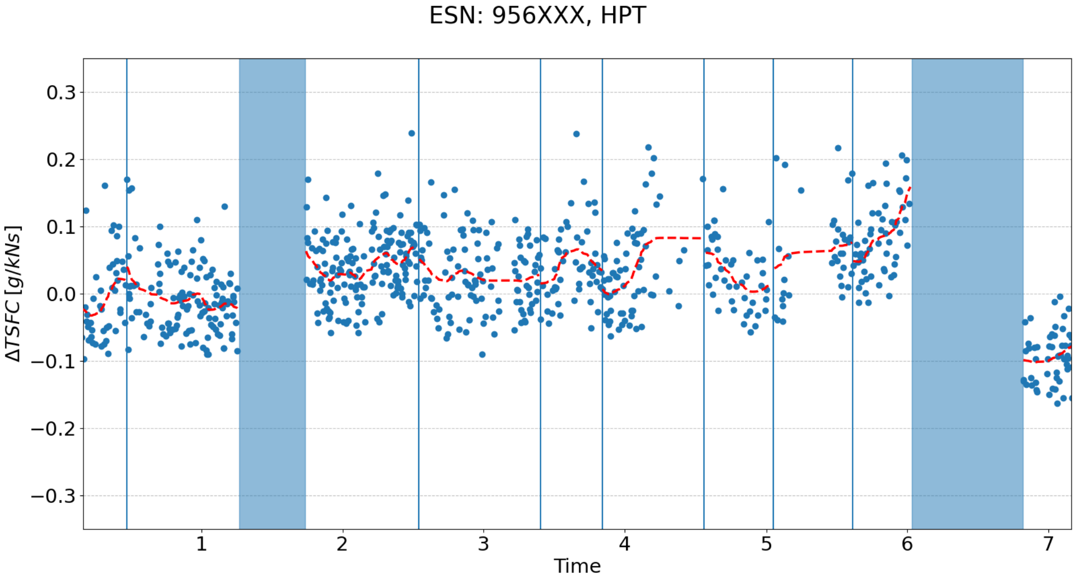


Figure D.12: HDMR trending using on-wing data from ESN956XXX for the HPT ΔW_f

LPT

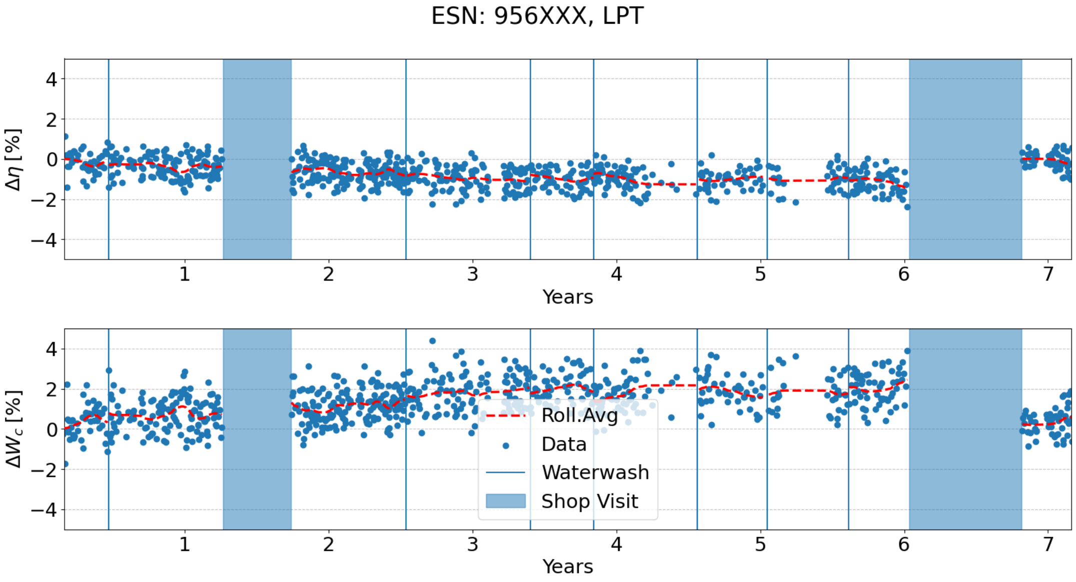


Figure D.13: HDMR trending using on-wing data from ESN956XXX for the LPT Degradations

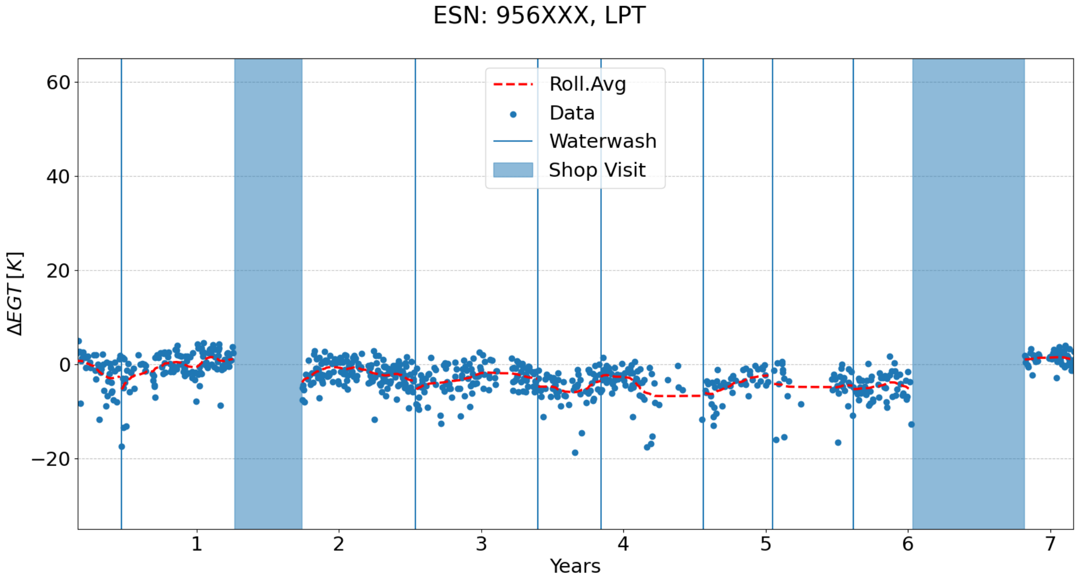


Figure D.14: HDMR trending using on-wing data from ESN956XXX for the LPT ΔEGT

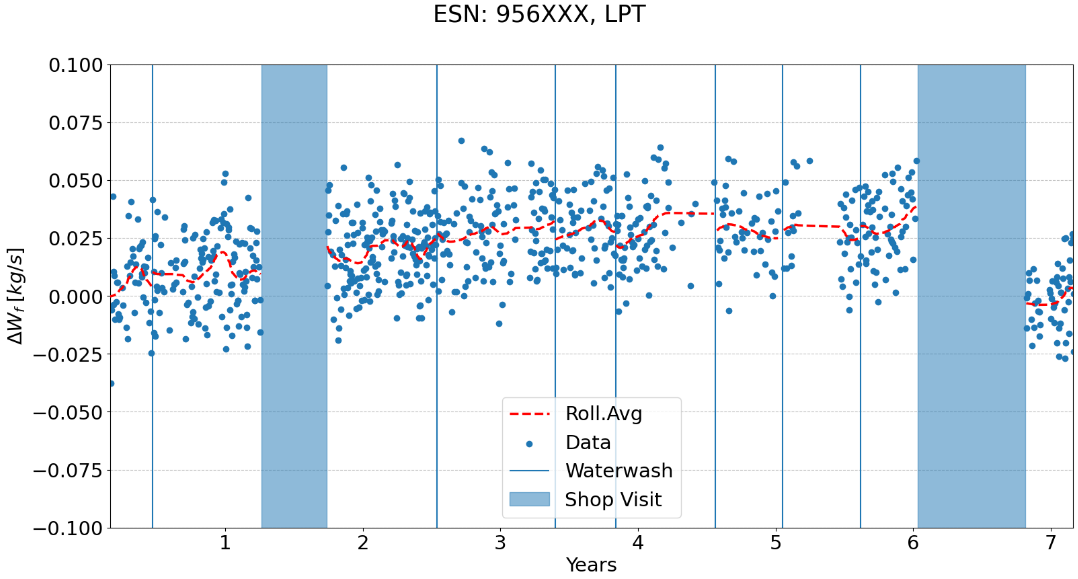


Figure D.15: HDMR trending using on-wing data from ESN956XXX for the LPT ΔW_f

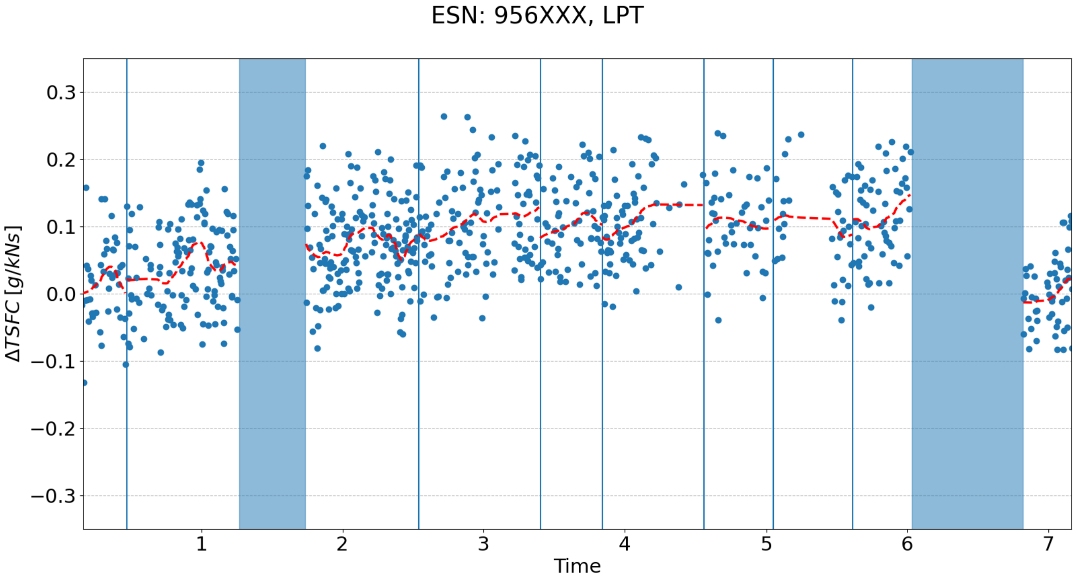


Figure D.16: HDMR trending using on-wing data from ESN956XXX for the LPT ΔW_f

E

Supporting Results for LSTM-based Prognostics Model

E.1. Model Tuning Results

E.1.1. Root Mean Squared Error

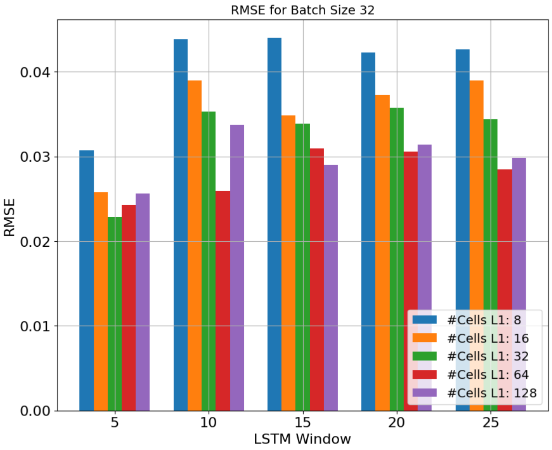


Figure E.1: RMSE results for Batch Size 32

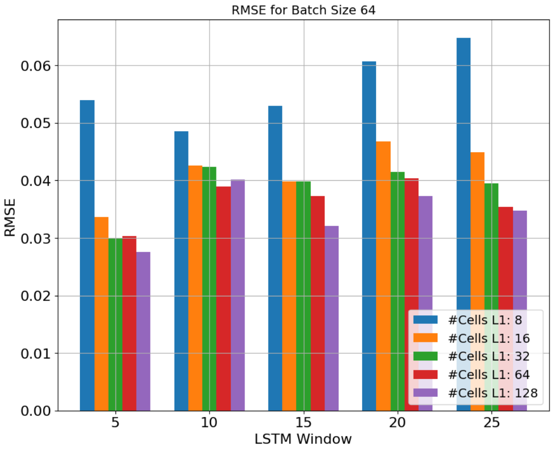


Figure E.2: RMSE results for Batch Size 64

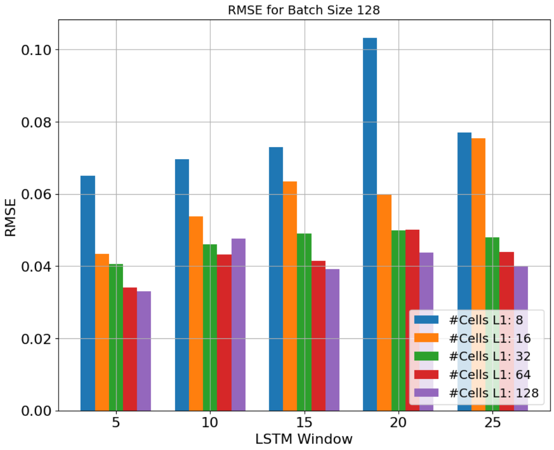


Figure E.3: RMSE results for Batch Size 128

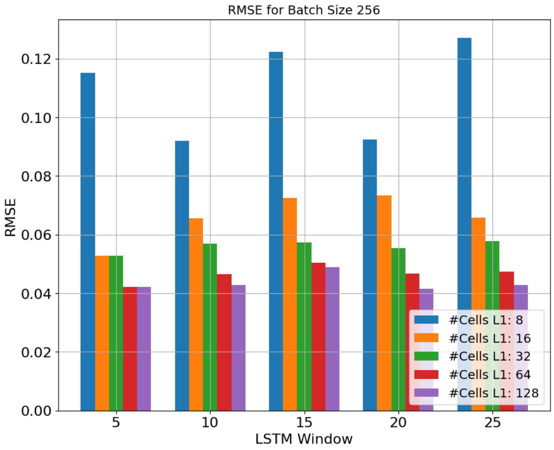


Figure E.4: RMSE results for Batch Size 256

E.1.2. Mean Squared Error

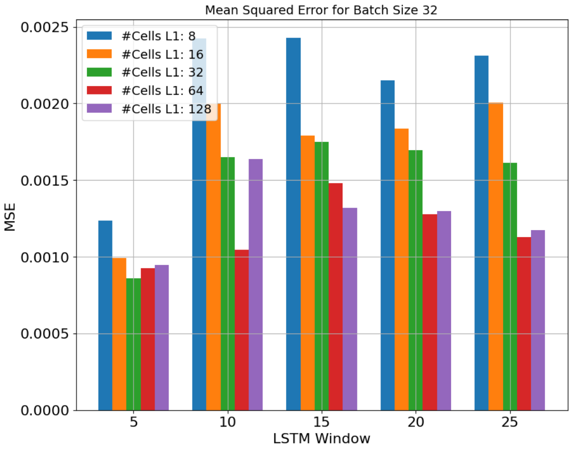


Figure E.5: MSE results for Batch Size 32

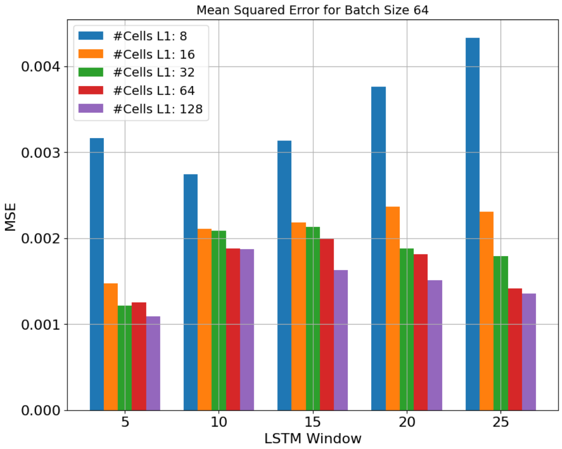


Figure E.6: MSE results for Batch Size 64

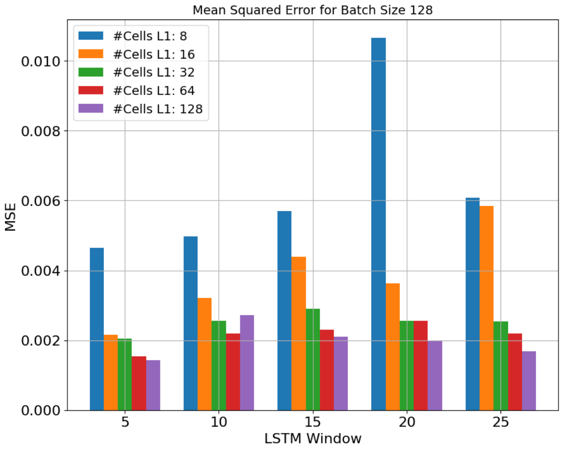


Figure E.7: MSE results for Batch Size 128

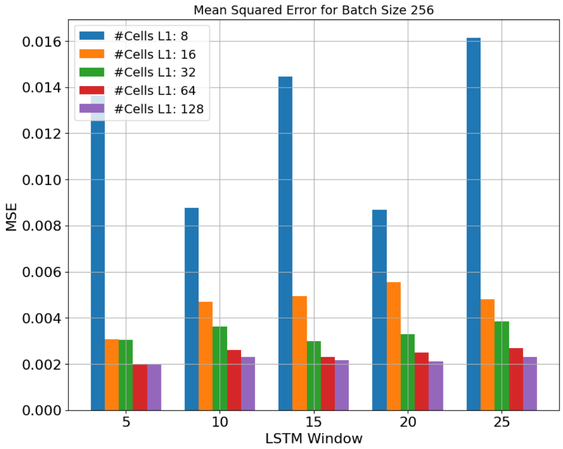


Figure E.8: MSE results for Batch Size 256

E.1.3. R2 Metric

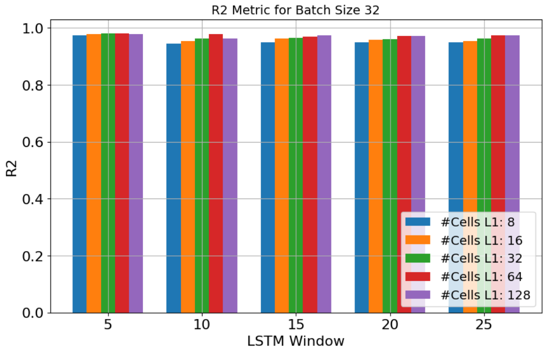


Figure E.9: R2 results for Batch Size 32

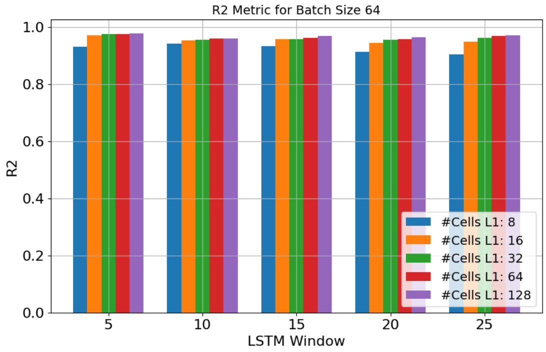


Figure E.10: R2 results for Batch Size 64

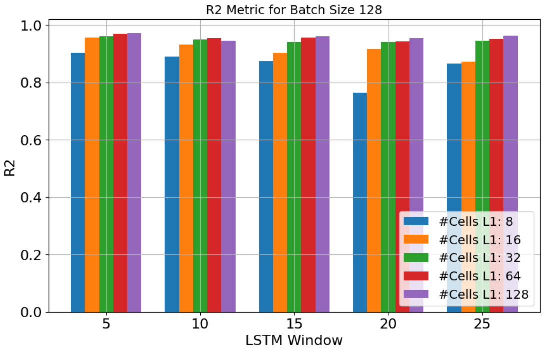


Figure E.11: R2 results for Batch Size 128

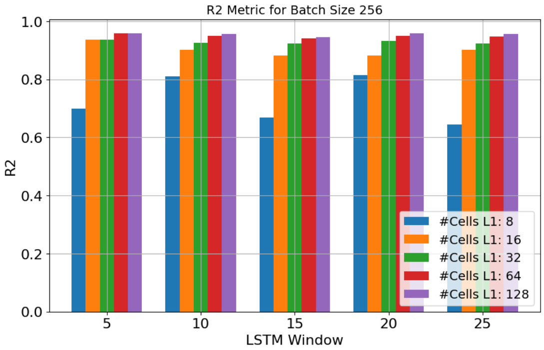


Figure E.12: R2 results for Batch Size 256

E.2. Prediction Tests

E.2.1. Baseline Model

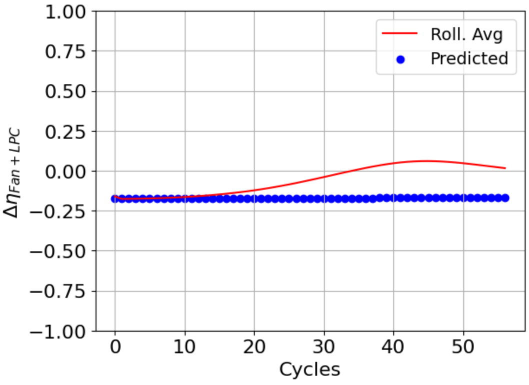


Figure E.13: Resulting predictions from Baseline model for Fan+LPC $\Delta\eta$

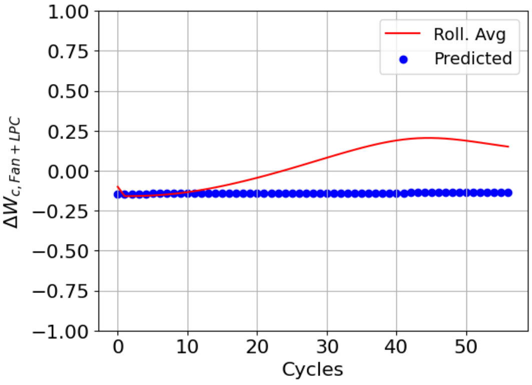


Figure E.14: Resulting predictions from Baseline model for Fan+LPC ΔW_c

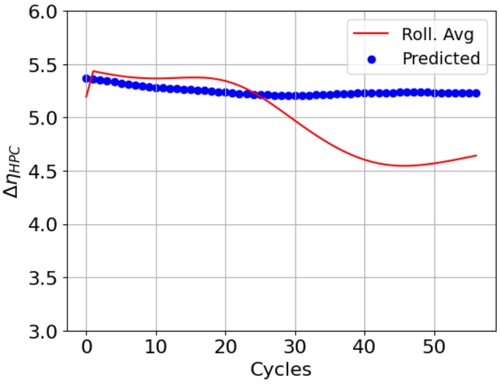


Figure E.15: Resulting predictions from Baseline model for HPC $\Delta\eta$

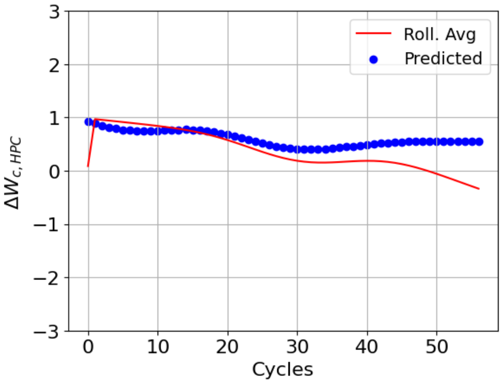


Figure E.16: Resulting predictions from Baseline model for HPC ΔW_c

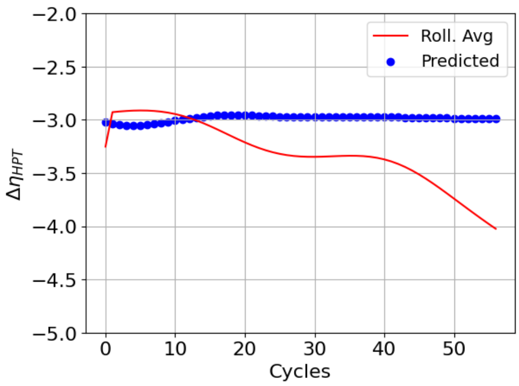


Figure E.17: Resulting predictions from Baseline model for LPT $\Delta\eta$

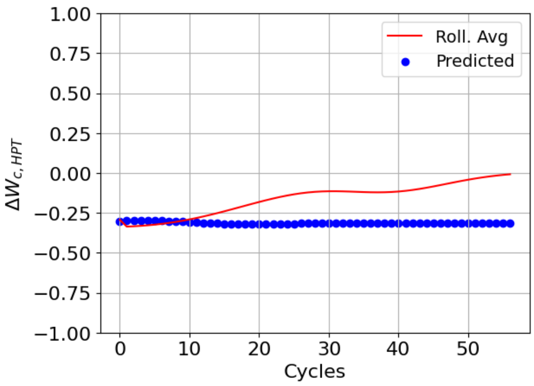


Figure E.18: Resulting predictions from Baseline model for LPT ΔW_c

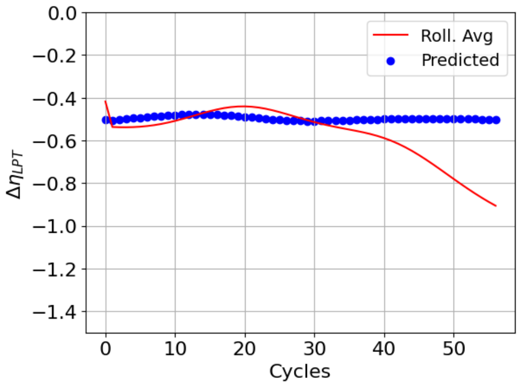


Figure E.19: Resulting predictions from Baseline model for LPT $\Delta\eta$

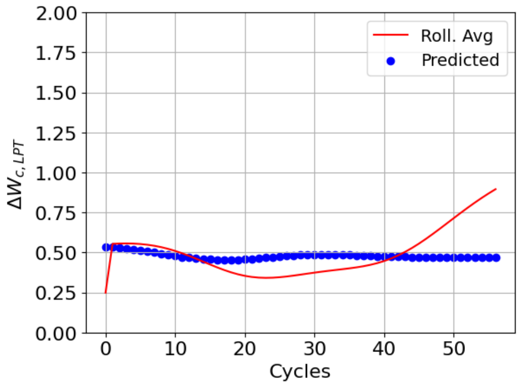


Figure E.20: Resulting predictions from Baseline model for LPT ΔW_c

E.2.2. Tuned Model

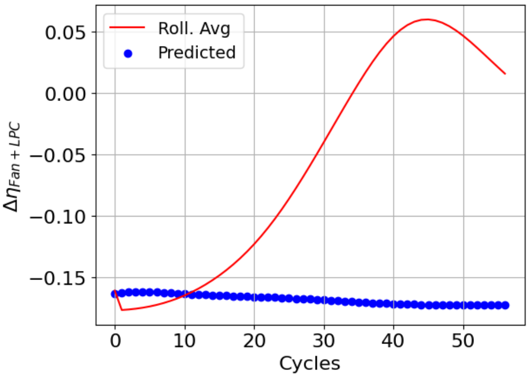


Figure E.21: Resulting predictions from Tuned model for Fan+LPC $\Delta\eta$

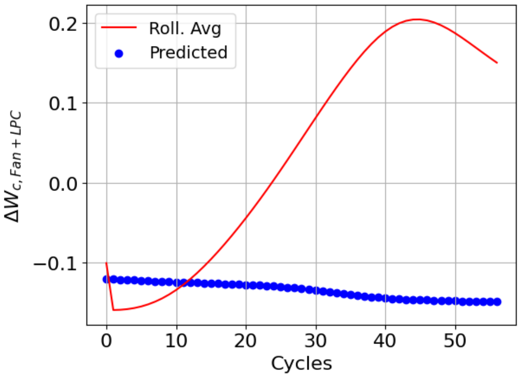


Figure E.22: Resulting predictions from Baseline model for Fan+LPC ΔW_c

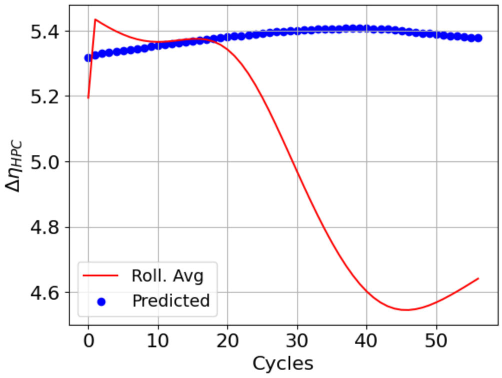


Figure E.23: Resulting predictions from Tuned model for HPC $\Delta\eta$

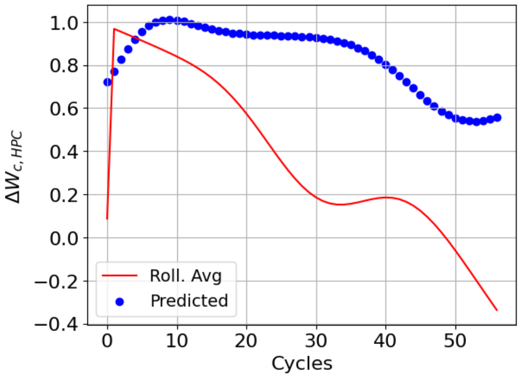


Figure E.24: Resulting predictions from Baseline model for HPC ΔW_c

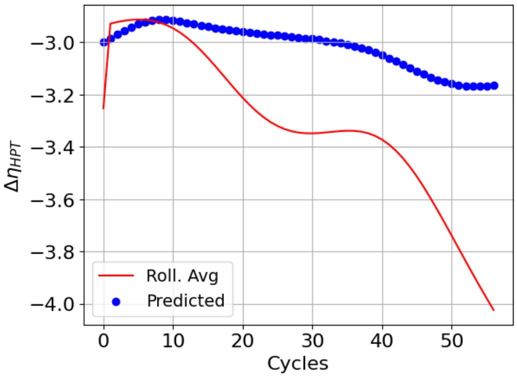


Figure E.25: Resulting predictions from Tuned model for HPT $\Delta\eta$

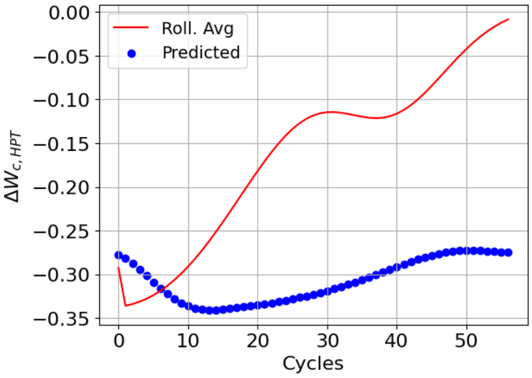


Figure E.26: Resulting predictions from Tuned model for HPT ΔW_c

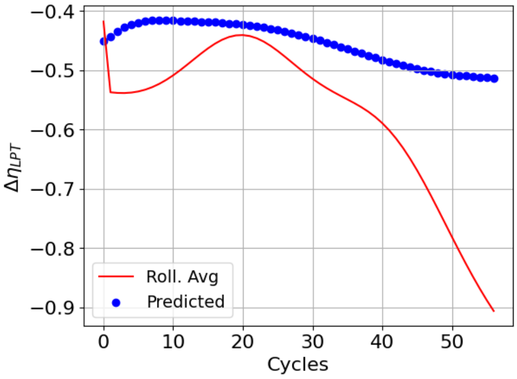


Figure E.27: Resulting predictions from Tuned model for LPT $\Delta\eta$

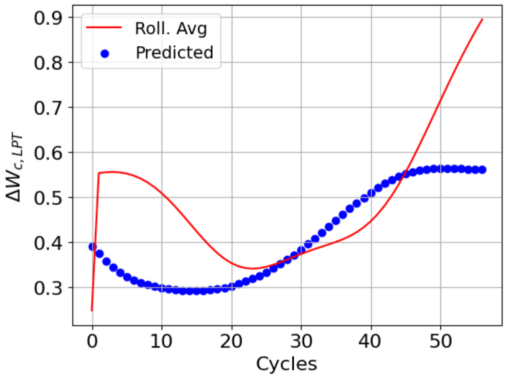


Figure E.28: Resulting predictions from Tuned model for LPT ΔW_c