UNIVERSITAT POLITÈCNICA
DE CATALUNYA
BARCELONATECH

DA&AI supporting tools for gas turbine's efficiency improvement: maintenance, operation modes and performance enhancement

Martí de Castro Cros

Universitat Politècnica de Catalunya (UPC)
Intelligence Data Science and Artificial Intelligence Research Centre (IDEAI-UPC)

DA&AI Supporting Tools for Gas Turbine's Efficiency Improvement: Maintenance, Operation Modes and Performance Enhancement

by
Martí de Castro Cros
Programa de Doctorat en Intel·ligència Artificial

Supervised By
Prof. Cecilio Angulo Bahón
Dr. Manel Velasco Garcia

Barcelona, 2023

We cannot solve our problems
with the same thinking we used
when we created them
Albert Einstein

Abstract

Digitalization has revolutionized many industries, including the power generation sector. The availability of a vast amount of data from various systems has transformed decision-making processes in Industry. Advances in artificial intelligence and machine learning have enabled the development of sophisticated algorithms that can process large datasets and uncover patterns and insights that were previously difficult to detect.

This thesis is part of a collaborative project between the Universitat Politècnica de Catalunya (UPC) and Siemens Energy (SE), aimed at creating digital tools for monitoring and improving the efficiency of industrial gas turbines used in power generation. The focus of this study is on developing maintenance-related support tools, as it is a key factor in the equipment performance as well as the cost of it is a significant expense for gas turbine operators.

Maintenance is a critical process for ensuring the reliability and availability of industrial systems. Key Performance Indicators (KPIs) and soft sensors have become increasingly popular for monitoring industrial processes and predicting variables that are difficult to measure. Therefore, the main goal driving this thesis is to develop an AI-based indicator that can help assess equipment performance and recommend maintenance actions.

To achieve this goal, an autoencoder-based architecture is used, incorporating several different structures and two types of autoencoder. These models are tested to determine which performs the best and is most suited to the equipment requirements. A detailed study is

presented, evaluating the performance of the models using two different metrics: absolute error and Fréchet distance, combined with two time-averaging calculations: moving average and incremental window average. The results of this analysis reveal a clear drift in the model output. Moreover, further results are obtained by modifying the autoencoder structure which lead to detect significant changes in equipment performance associated with major maintenance events.

Time series decomposition, wavelet transform, and clustering methods are used to further analyze the findings and obtain additional insights into gas turbine performance. The outcome derived from this study strengthen the drift detection in gas turbine performance and the identification of significant change in its behavior due to major maintenance events. This research can serve as a foundation for future studies and investigations in this area, as it has laid the groundwork for the potential development of more sophisticated and accurate models that can effectively monitor and diagnose potential issues in Siemens Energy gas turbines.

This doctoral thesis contributes to reducing the gap between academia and industry by applying novel technologies and algorithms to real plant problems. It provides valuable insights and understanding of gas turbine systems and presents a framework for developing more accurate and targeted models. By leveraging the power of machine learning and advanced analytics, researchers and industry professionals can work together to improve the efficiency, reliability, and safety of gas turbines in a wide range of industrial applications.

Resum

La digitalització ha revolucionat el sector industrial. L'augment en la captura de dades i el desenvolupament en els camps de la intel·ligència artificial i l'aprenetatge automàtic han significat un canvi transformador en la forma com es prenen decisions a la indústria. Aquests avenços han permès el desenvolupament d'algorismes sofisticats que poden processar grans quantitats de dades i descobrir patrons i estructures que abans eren difícils de detectar

Aquesta tesi forma part d'un projecte de col·laboració entre la Universitat Politècnica de Catalunya (UPC) i Siemens Energy (SE), amb la idea de crear eines digitals pel seguiment i la millora de l'eficiència de les turbines de gas en aplicacions industrials. L'objectiu d'aquest estudi es centra en el desenvolupament d'eines de suport relacionades amb el manteniment, ja que és un dels factors clau en el rendiment de l'equipament i el seu cost és una despesa important per als operadors d'aquest sector.

El manteniment és un procés crític per garantir la fiabilitat i disponibilitat dels sistemes industrials. Els indicadors (KPI) i els sensors digitals s'han tornat cada cop més populars per monitoritzar processos industrials i predir variables difícils de mesurar. Per tant, l'objectiu principal d'aquesta tesi és desenvolupar un indicador basat en IA que pugui ajudar a avaluar el rendiment dels equips i recomanar accions de manteniment.

Per aconseguir aquest objectiu, s'utilitza una arquitectura basada en autoencoder, que incorpora estructures i tipus diversos. Aquests models s'han posat a prova per determinar quin ofereix un millor rendiment i s'adapta bé més als requisits de les màquines. Es

presenta un estudi detallat, avaluant el rendiment dels models mitjançant dues mètriques diferents: l'error absolut i la distància de Fréchet, combinats amb dos càlculs de mitjana temporal: mitjana mòbil i mitjana incremental. Els resultats d'aquest anàlisis revelen una clara desviació en els resultats del model. A més, s'obtenen resultats addicionals modificant l'estructura de l'espai latent que permeten detectar canvis significatius en el rendiment de la màquina associats a esdeveniments significatius de manteniment.

S'utilitzen mètodes de descomposició de sèries temporals, transformació de wavelet i d'agrupació de dades per analitzar més els resultats i obtenir informació addicional sobre el rendiment de les turbines de gas. El resultat derivat d'aquest estudi consisteix en la detecció d'una desviació en el rendiment de la màquina i la identificació de canvis significatius en el seu comportament a causa d'activitats significatives de manteniment. Aquesta investigació pot servir de base per a futurs estudis i investigacions en aquesta àrea, ja que ha assentat les bases per al desenvolupament potencial de models més sofisticats i precisos que puguin controlar i diagnosticar eficaçment problemes potencials a les turbines de gas de Siemens Energy.

Aquesta tesi doctoral contribueix a reduir la diferència entre la investigació i la indústria mitjançant l'aplicació de noves tecnologies i algorismes a problemes industrials reals. També pretén aportar informació valuosa sobre el funcionament dels sistemes de turbines de gas des d'un punt de vista de dades i presenta unes eines per desenvolupar models més precisos i orientats a tasques específiques i necessàries de les turbines de gas. Aprofitant el desenvolupament de l'aprenentatge automàtic i l'anàlisi avançada, es pretén millorar l'eficiència, la fiabilitat i la seguretat d'aquestes màquines per ampli rang d'usos en aplicacions industrials.

Keywords: Artificial Intelligence, Autoencoder, Compressor, Industrial Gas Turbine, KPI, Machine Learning, Maintenance, Soft Sensor

Table of Contents

List of Figures

List of Tables

List of Abbreviations

AD	**A**bsolute **D**ifference
ADIWA	**A**bsolute **D**ifference on **I**ncremental **W**indow **A**verage
ADMA	**A**bsolute **D**ifference on **M**oving **A**verage
AE	**A**uto**E**ncoder
AI	**A**rtificial **I**ntelligence
ANN	**A**rtificial **N**eural **N**etwork
CBM	**C**ondition-**B**ased **M**aintenance
CC	**C**ombined **C**ycle
CCUS	**C**arbon, **C**apture, **U**tilisation and **S**torage
CE	binary-**C**ross**E**ntropy
CHP	**C**ombined **H**eat and **P**ower
CM	**C**ondition **M**onitoring
CMAPSS	**C**ommercial **M**odular **A**ero-**P**ropulsion **S**ystem **S**imulation
CNN	**C**onvolutional **N**eural **N**etwork
CV	**C**ontrol **V**olume
DA	**D**ata **A**nalytics
DBSCAN	**D**ensity-based **S**patial **C**lustering of **A**pplications with **N**oise
DWT	**D**iscrete **W**avelet **T**ransform
EKF	**E**xtended **K**alman **F**ilter
ELU	**E**xponential **L**inear **U**nit
FC	**F**ully **C**onnected autoencoder
FD	**F**réchet **D**istance

FDIWA	**F**réchet **D**istance on **I**ncremental **W**indow **A**verage
FDMA	**F**réchet **D**istance on **M**oving **A**verage
GA	**G**enetic **A**lgorithm
GD	**G**radient **D**escent
GT	**G**as **T**urbine
GPA	**G**as **P**ath **A**nalysis
ICM	**I**nfluence **C**oefficient **M**atrix
IGT	**I**ndustrial **G**as **T**urbine
IWA	**I**ncremental **W**indow **A**verage
KF	**K**alman **F**ilter
KPI	**K**ey **P**erformance **I**ndicator
LeakyReLU	**L**eaky **Re**ctified **L**inear **U**nit
LKF	**L**inear **K**alman **F**ilter
MA	**M**oving **A**verage
MAD	**M**edian **A**bsolute **D**eviation
ML	**M**achine **L**earning
PHM	**P**rognostics and **H**ealth **M**anagement
ReLU	**Re**ctified **L**inear **U**nit
RL	**R**einforcement **L**earning
RNN	**R**ecurent **N**eural **N**etwork
RUL	**R**emaining **U**seful **L**ife
SE	**S**iemens **E**nergy
SGD	**S**tochastic **G**radient **D**escent
SL	**S**upervised **L**earning
SSL	**S**emi-**S**upervised **L**earning
TBM	**T**ime **B**ased **M**aintenance
TIT	**T**urbine **I**nlet **T**emperature
t-SNE	**t**-distributed **S**tochastic **N**eighbor **E**mbedding
UBM	**U**sage **B**ased **M**aintenance

UKF	**U**nscented **K**alman **F**ilter
UL	**U**nsupervised **L**earning
UMAP	**U**niform **M**anifold **A**pproximation and **P**rojection
UPC	**U**niversitat **P**olitècnica de **C**atalunya
SHM	**S**tructural **H**ealth **M**onitoring
STD	**ST**andard **D**eviation
VGV	**V**ariable **G**uide **V**ane
WT	**W**avelet **T**ransform

List of Symbols

A	Surface	m^2
C_p	Specific heat at constant Pressure	J / (kg K)
C_v	Specific heat at constant Volume	J / (kg K)
$\mathcal{DM}$	Differential Model	
E	Energy	J
$\mathcal{EM}$	Expansion Model	
F	Decoder Network	
G	Encoder Network	
h	Enthalpy	J / kg
H_a	Air Humidity	%
KL	Kullback-Leible divergence	
$\mathcal{M}$	Fresh Training Data	
P	Pressure	kPa
Q	Heat	J
$\mathcal{RD}$	Fresh Reconstruction Discrepancy	
rP	Pressure ratio	
rT	Temperature ratio	
S	Entropy	J / K
T	Temperature	^{o}C
$\mathcal{T}$	Test data	
$Tanh$	Hyperbolic Tangent	
T_s	Sampling time	min

V	Volume	m^3
W	Compressor / Turbine Work	J
ΔK	Temporal gradient of Kinetic Energy	J
ΔP	Temporal gradient of Potential Energy	J
ΔU	Temporal gradient of Internal Energy	J
ϵ	Loss function	
η_C	Compressor Isentropic Efficiency	
η_T	Turbine Isentropic Efficiency	
$\psi_{a,b}$	Wavelet Function	

Chapter 1

Introduction

This section presents a comprehensive overview of machine learning, Gas Turbine (GT) systems, and key concepts like thermodynamics and maintenance policies associated to this thesis. It underscores the significance of maintaining optimal performance of GT systems for achieving high reliability, system availability, environmental sustainability, and human safety. Additionally, it highlights the usage of major maintenance operations for restoring Industrial Gas Turbines (IGTs) to a new operational point. Furthermore, the section discuss the limitations of traditional maintenance methods and the potential benefits of implementing artificial intelligence techniques to improve maintenance decision-making. Finally, it emphasizes the need of conducting research to explore the application of artificial intelligence in GT maintenance for enhancing system reliability and reducing maintenance expenses.

1.1 Artificial Intelligence

The precise definition and meaning of Artificial Intelligence (AI) is a subject of much discussion and has caused a lot of confusion [1]. In

the first half of the 20th century, the concept of AI began with some fantasies and fiction writers however, it was not until 1950 that Alan Turing suggested a logical framework called *Computing Machinery and Intelligence* [2] in which he discussed how to build intelligent machines and how to test it. The main idea behind the test was that if a human interrogator, after posing some written questions, cannot tell the difference amongst human and computer responses, the machine was intelligent. The capabilities of the computer to pass the test were natural language processing, knowledge representation, automated reasoning, machine learning, computer vision and robotics [3]. In 1956, what is considered the first AI program was presented at the *Dartmouth Summer Research Project on Artificial Intelligence (DSRPAI)* hosted by John McCarthy and Marvin Minsky [4]. While the hype was too high, after some years of research, the most ambitious goals were not met and expectations of AI did not match the reality since the computers could not store enough information or process it fast enough. It was during the 1900s that many landmarks goals of AI were achieved and when machine learning algorithms were successfully applied to many problems in academia and industry with the use of powerful hardware, and immense collections of data [5]. From then on, this has been a huge field of research and new applications have flourished.

1.2 Machine Learning

Machine Learning (ML) is a subfield of artificial intelligence which is a rapidly growing area that is being applied across various industries to build automated analytical models. By utilizing algorithms to analyze data, ML is able to automatically recognize patterns and

learn valuable insights that are used to improve decision-making [6]. These sophisticated data analytic algorithms are capable of processing large amounts of data, and provide significant improvement of machine performance. ML methodologies are frequently employed to extract knowledge from multidimensional time series data, identifying and revealing hidden structures [7].

The field of ML has grown significantly in recent years, thanks to the development of powerful computing resources, sophisticated algorithms, and the availability of large datasets [8]. These techniques are being used in various fields, including image and speech recognition, natural language processing, recommender systems, and autonomous vehicles.

In today's classification, four different types of ML can be described: supervised learning, unsupervised learning, semi-supervised learning, and reinforcement learning [9]. These are further explained next.

1.2.1 Supervised Learning

Supervised Learning (SL) is a subfield of ML in which the data used to train the algorithm includes their desired solutions, called labels. The goal is to create a model that can predict the correct output for new input data. This is performed by tuning the parameters of the model to minimize the difference between its predictions and the true output values in the training data [10].

SL can be used for a variety of tasks, but the most common usage is for classification purposes. Some popular algorithms for SL include decision trees, random forests, support vector machines, and neural networks.

One example of the use of these methods is in predicting the energy output of a wind turbine based on various environmental factors such as wind speed and temperature. Researchers have used a variety of supervised techniques, such as Artificial Neural Networks (ANNs) and support vector regression, to create models that can accurately predict wind turbine output with high accuracy [11].

Another application of SL in a completely different field is for medical diagnosis. Researchers have used supervised learning algorithms to analyze medical imaging data, such as X-rays and MRI scans, to detect and diagnose diseases such as cancer, amongst others [12].

Finally, this last example proposes a decision tree model to analyse the most important factor that will result in the improved education level of India [13].

1.2.2 Unsupervised Learning

Unsupervised Learning (UL) is a subfield of ML in which the learning algorithm operates with unlabelled data to discover data structures and to identify hidden patterns. This approach is useful either when there is not any prior knowledge about the composition of the data or when different relationship must be drawn from a group of points, also known as clusters [14].

The way how these algorithms learn is by defining a cost function that constraint either the number of clusters or some prior metrics threshold, usually some distance metrics. Then, an interative process is started to optimize the cost function until its convergence [15]. Unlike SL methods, which are prone to minimizing error or misclassifications of the pair input / output data, UL methods are devoted

to maximizing the similarities between data items and cluster prototypes.

Common applications of UL include clustering, dimensionality reduction, visualization, and anomaly detection.

1.2.3 Semi-Supervised Learning

Semi-Supervised Learning (SSL) is a subfield of ML that combines labelled and unlabelled data to train a model. The availability of large amounts of unlabelled data makes SSL particularly useful in situations where labelled data is scarce or expensive to obtain [16].

In SSL, a small amount of labelled data is used to guide the learning process, while the majority of the training data is unlabelled. Then, the model attempts to generalize from the labelled data to make predictions on the unlabelled data [17].

There are three main approaches to SSL, including generative models, graph based methods, and self-training. Generative models assume that the labelled and unlabelled data share a common underlying distribution and use this assumption to infer the labels of the unlabelled data. Graph based methods use the similarity between data points to propagate labels from the labelled data to the unlabelled data. Self-training involves using the model's own predictions on the unlabelled data to generate additional labelled data [18]. This kind of learning has been successfully applied in various fields, such as natural language processing, computer vision, and bioinformatics.

Overall, SSL provides a promising approach to training machine learning models with limited labelled data, and it is an active area of research in the field of ML [19].

1.2.4 Reinforcement Learning

Reinforcement Learning (RL) is a subfield of ML that involves an agent interacting with an environment, taking actions and receiving feedback in the form of rewards or penalties, with the goal of maximizing its long-term cumulative reward. The agent learns to make decisions through trial and error, exploring different actions and learning from the feedback it receives from the environment. The agent's goal is to learn a policy, a function that maps states to actions, maximizing its expected cumulative reward over time. In short, the basic idea of the algorithm is to learn through trial and error to make decisions regarding the current state and maximize rewards or minimize penalties [20].

One of the key challenges in RL is balancing both exploration, meaning to favour new actions to learn more about the unknown part of th environment, and exploitation, using the agent's current knowledge to take actions that maximize reward [21]. Various algorithms have been developed to address this challenge, such as Q-learning [22], SARSA [23], actor-critic [24], and deep reinforcement learning [25].

1.3 Gas Turbine Industry

The current energy crisis has exponentially increased concerns about the vulnerability, fragility, and unsustainability of our existing energy system. As the demand for energy continues to increase, the limitations of the current system have become more apparent. Despite the development of renewable energies, the share of fossil fuels in the global energy mix has been stubbornly high, at around 80% for decades [26]. The importance of reducing the world fossil fuel use

and the emission produced by them are highlighted in by the research ecosystem due to the uncertainty about their future availability and to avoid the severe impact on the environment.

Concerns over the negative environmental impact of fossil fuels, particularly coal and oil, have led to increased calls to transition to more sustainable energy sources [27]. Natural gas has been recommended as it emits fewer pollutants than other fossil fuels and can provide more flexibility in power generation [28]. Combined cycle (CC) and combined heat and power (CHP) plants are increasingly being used to harness the potential of natural gas in a more efficient and cost-effective manner, resulting in lower fuel costs and emissions per MWh [29].

Gas turbines (GTs) are the primary components of CC and CHP power plants. In recent years, GT units have undergone significant efficiency enhancements, with improvements in manufacturing processes and the use of new materials contributing to efficiency gains up to 45% in the simple cycle [30]. However, some of these improvements, such as increasing the Turbine Inlet Temperature (TIT) to improve turbine blade performance, can lead to higher degradation and a reduction in reliability [31]. As a result, researchers are exploring new ways to optimize the performance of GT and enhance their reliability, including through the use of advanced materials and coatings [32].

GTs are widely used in various industries such as power generation, aviation, oil and gas, marine, and industrial processing. The global market is expected to grow in the coming years, driven by its storability, its ability to be delivered through pipelines or liquefied and sent by ship. Moreover, the ability to provide a quick and reliable

source of energy, thanks to their efficiency and low emissions, regarding other fossil fuels, allows natural gas to respond to both seasonal and short-term demand fluctuations. It is one of the most optimal choices available to provide back-up to the growing use of variable renewables such as wind and power. While GTs are not considered as clean and sustainable as renewable energy sources such as wind and solar power, they are still necessary to support the growing energy demands of modern society. Still, as part of global efforts to reach net zero emissions, natural gas use is expected to come under pressure in some countries [33]. As a result, research and development efforts are underway to improve the efficiency of GTs, reduce emissions by deploying Carbon Capture, Utilisation and Storage (CCUS) equipment, and explore the potential for using in combination with cleaner fuels such as hydrogen.

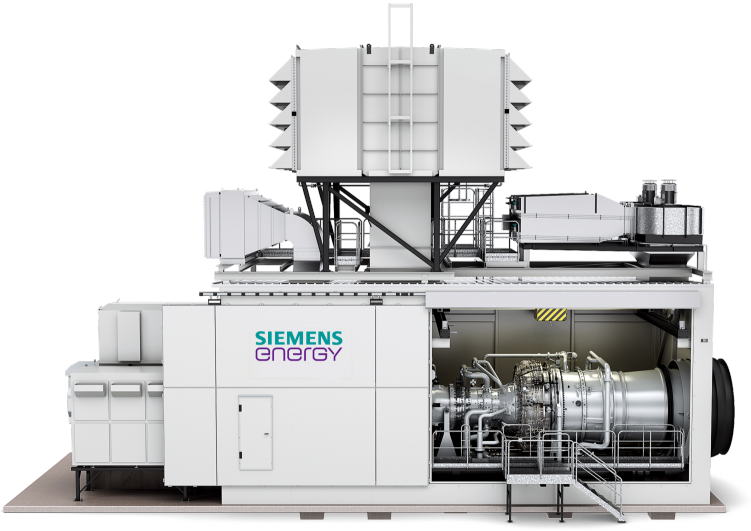


Figure 1.1: SGT-800 Industrial Gas Turbine (courtesy of Siemens Energy AG)

1.4 Industrial Gas Turbine System

IGTs are a versatile technology used for power generation in a variety of industries including oil and gas, energy generation, and aviation. They are internal combustion engines that convert the chemical energy of fuel into mechanical energy which is then converted into electrical power. The three main components of an IGTs are compressor, combustor, and power turbine.

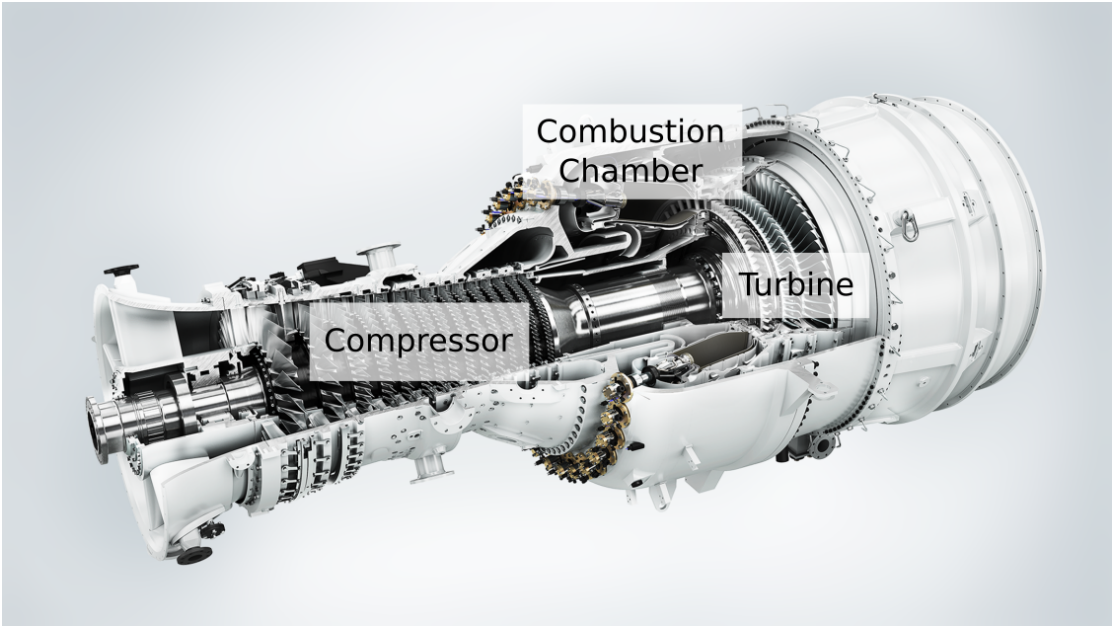


Figure 1.2: The core engine of the Industrial Gas Turbine model used to perform the current analysis (courtesy of Siemens Energy AG)

The *compressor* is responsible for compressing and providing sufficient quantity of air to the system to ensure its proper operation. Its main function is to increase the pressure and temperature of the incoming air to facilitate efficient combustion in the combustor. It is composed of a series of blades that rotate at high speed, drawing in and compressing air. The compressor plays a crucial role in determining the overall efficiency and performance of the GT, as it directly affects the power output and fuel consumption. Any degradation or inefficiency in the compressor can lead to reduced performance,

higher maintenance costs, and shorter service life. Therefore, the proper maintenance and operation of the compressor is essential for maximizing the efficiency and longevity of GTs [34].

The *combustion chamber*, which contains several combustors, is where the combustion process takes place. The main function of the combustion chamber is to mix fuel and air, and then ignite them by using the burners to produce high-temperature, high-pressure gas. The heat produced from the exothermic reaction is then used to drive the turbine and generate electricity. The combustion chamber must be carefully designed to ensure it provides a complete combustion process and avoid malfunctioning events. Moreover, a precise control is required to ensure that temperature and pressure values within the chamber are within safe operating limits, and to minimize the emissions. Thereby, efficient and reliable combustion chambers are essential for the overall performance and reliability of GT systems.

The *turbine* is the last component of the chain as it is the responsible to convert the high-temperature and high-pressure gas generated in the combustion chamber into mechanical energy to drive the generator and produce electricity. The gas expands as it passes through the turbine blades, causing them to rotate at high speeds. The energy from the rotating blades is then transmitted to the generator through a shaft, which converts the mechanical energy into electrical energy. The design of the turbine is crucial to maximize the energy conversion efficiency of the GT, as well as to ensure its reliability and durability under harsh operating conditions.

GTs are suitable for a wide range of applications, from small mobile power plants to large, stationary systems with high power production. They are highly efficient and have a short startup time,

making them ideal for supplying fast, flexible energy to meet peak demands [35]. However, there are concerns about the environmental impact of GTs, including toxic and noxious emissions. To address these concerns, there is a need for ongoing monitoring, condition control, system diagnosis and improvements in emission reduction technologies.

GTs are also known for their ability to operate under extreme and harsh conditions. These situations also lead to various types of deterioration, which mainly cause malfunctioning events and decreased performance. These include fouling, corrosion, erosion, abrasion, and unexpected particles. Further description of each fault can be found in Table 1.1.

Table 1.1: Main deterioration causes of a gas turbine

Name	**Cause**	**Prevention**	**Recoverable**
Fouling	Adherence of particle to airfoils and annulus surfaces. Increases surface roughness and changes airfoils shape.	Filtration system	By washing components
Corrosion	Loss or deterioration of materials caused by inlet air contaminants and by derivative of fuel combustion.	Complete combustion process	No
Hot Corrosion	Loss or deterioration of materials from flow path caused by chemical reactions at high temperature.	Filtration system	No
Erosion	Material removal from the flow path by hard or incompressible particles impinging on flow surface.	Filtration system	No
Abrasion	Material removal caused by a rotating surface rubbing a static surface to establish the proper clearances.	–	No
Foreign objects	Unexpected particles striking components along the flow path: particles in air or broken pieces.	Filtration system	No

In addition, changes in the standard operational regime, such as the number of starts and stops or modifications in the output power set

point, can also cause fatigue and impact the system's performance. Therefore, reducing the operational and maintenance expenses, since a huge amount of money is invested in operational and maintenance actions [36], while increasing the reliability, availability, and safety of equipment are key factors of profitability and competitiveness. To address these issues, health-monitoring practices have become mandatory in order to avoid economic, environmental and security defects. Thanks to the development of new and more advanced technologies, more sophisticated intelligent tools for condition and fault assessment have been emerged over the past decade. In this context, data collectors have played an important role to improve health management strategies [37].

The big amount of data captured by industrial systems contains information about components, events, and alarms related to industrial processes. All these data can provide significant knowledge and information about system processes and their dynamics. Thereby, the condition assessment of the system can be improved leading to maintenance cost and machine fault reduction. This can also be translated as an increase in production and improvement of operator safety [38].

1.5 Condition Assessment

Condition assessment consists of a systematic inspection, review, and report of the state of the equipment. Inspection procedures have evolved as recent and more effective techniques have been developed, with the increase of data availability. Hence, they can be distilled into three main components: Condition Monitoring (CM), diagnostics, and prognosis [39].

Firstly, CM involves periodically checking the status of the equipment while keeping a record of observations and analyzing data from various sensors located on the system. Monitoring can be developed near the machine or remotely, depending on the setup. Data records are collected and sent to engineering office where they are analyzed and processed to identify any potential issues or unforeseen event. CM helps to detect and diagnose faults and deterioration in the equipment allowing for preventive maintenance and avoiding costly repairs or down times. With the advance in technologies, this process has become more sophisticated and effective in identifying and preventing equipment issues [40].

Secondly, diagnostic systems process the information gathered from the equipment status to determine and identify risks that impact its operational integrity. Experienced engineers typically provides a report that outlines a plan or recommended actions to follow in case it is not working as expected. The main purpose of this data analytics process is to make the machine more reliable, available all of the time, and safer, thereby improving its performance [41]. This complements the CM in such a way that one aims to capture the performance of the system, while the other aims to improve it. Thereby, both of the systems need to work in harmony to achieve their goals.

Finally, prognosis can help on predicting when maintenance should be performed, which components need to be replaced, and when the equipment might fail. The objective of future condition forecasting is usually defined in terms of the prevention of hard failures of the components or reducing the performance degradation related to the equipment's operation. Failure prognosis puts the focus on forecasting the damage state or failure rate of a component or system

of components in an engine, whereas degradation prognosis is associated with the slower decrease of performance throughout its life. This process is a step ahead of both CM and diagnostics, but at the same time, there is some dependency on how the diagnosis is performed [42]. The relationship among all these processes is represented in Figure 1.3.

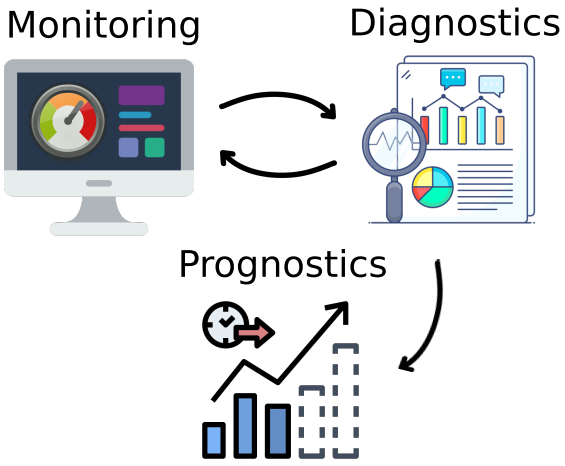


Figure 1.3: Relationship representation between condition monitoring, diagnostic and prognosis systems

1.6 Gas Turbine Thermodynamics

This section presents the basic principle that drive the functioning of GT. The main focus of this thesis is the enhancement of the performance of a single GT, and more precisely, by analysing the functioning of the compressor. Therefore, an overview of the air-standard Brayton cycle for the whole system of gas turbines and also specifically to each component is provided [43].

Before going in depth with the mentioned processes, some key concepts must be introduced. The concepts explained below can be further expanded in [44].

1.6.1 First Thermodynamic Law

As any other thermal machine, GT is governed by the most essential thermodynamics. Since the thermodynamics of these kind of system is beyond this thesis, only a small intuition will be given in order to understand some decisions taken during this study.

The first law of thermodynamics states that energy is conserved. Thereby, it states that the change in the amount of energy contained within a system during some time interval is equal to the difference between the net amount of energy transferred *in* across the system boundary during the time interval and the amount of energy transferred *out* across the system boundary by *work* during the time interval. This can be expressed as:

$$E_2 - E_1 = Q - W \tag{1.1}$$

Where E_i is the system's energy in/out, Q is the heat coming in the system, and W is the work performed by the system.

An alternative form of this expression is:

$$\Delta K + \Delta P + \Delta U = Q - W = \Delta E \tag{1.2}$$

where ΔK is the temporal gradient of kinetic energy, ΔP is the temporal gradient of potential energy and ΔU is the temporal gradient of internal energy. Also note that the algebraic sign corresponds to the direction of the transfer. If the energy is transferred from the system to the surroundings it is negative sign, otherwise it is positive.

1.6.2 Second Thermodynamic Law

The second law of thermodynamics introduce the meaning of entropy (S). In short, the entropy is an extensive property like mass and energy that measures the molecular disorder or randomness of a system. The second law states that it cannot be destroyed but instead it can be created. The entropy, just as mass and energy, accounted for by an entropy balance. In open systems, it is expressed as:

$$\Delta S = S_{transfer} + S_{gen} \tag{1.3}$$

where ΔS is the change of entropy within the system, $S_{transfer}$ is the transferred entropy across the system boundary during time t, and S_{gen} is the amount of entropy generated within the system during time t.

This property allows to determine if a process is reversible ($\Delta S = 0$), meaning that the system and the surroundings can be returned to their initial states, or irreversible ($\Delta S > 0$), meaning that they cannot be recovered. This application is highly important in open systems to compare the efficiency of thermodynamic processes.

1.6.3 Control Volume

A Control Volume (CV) is strictly defined as a delimited region of space which mass may flow. This concept is the key to understand open systems.

Let $A \in \mathbb{R}^3$ be a volume:

$$A(\mathbb{R}^3) = \int_{\mathbb{R}^3} dV = \int_{\mathbb{R}^3} dx\, dy\, dz = \int_{\mathbb{R}^3} X_{\mathbb{R}}(x, y, z)\, dx\, dy\, dz \tag{1.4}$$

where:

$$X_{\mathbb{R}}(x, y, z) \begin{cases} 1, & (x, y, z) \in V \\ 0, & \text{otherwise} \end{cases} \tag{1.5}$$

A graphical representation of A during the instant t is shown in Figure 1.4.

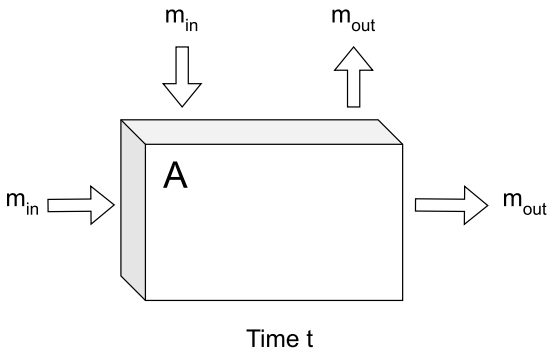


Figure 1.4: Graphical representation of control volume definition

The energy rate balance for CV must follow the conservation of energy principle explained in the first thermodynamic law. The difference between the closed and open system is that now the energy can be transferred into or out of a CV as a result of mass crossing boundaries. Therefore, the conservation of energy principle must be modified by taking into account these energy transfers. Thus, an extra component must be added in the previous expression that is the net rate of energy transfer into the system accompanying mass flow. Thereby, the modified expression is defined as:

$$\begin{aligned} \frac{dE_{cv}}{dt} &= \dot{Q} - \dot{W} \quad + \dot{m}_{in}\left(u_{in} + \frac{v_{in}^2}{2} + g\, z_{in}\right) \\ &\qquad - \dot{m}_{out}\left(u_{out} + \frac{v_{out}^2}{2} + g\, z_{out}\right) \end{aligned} \tag{1.6}$$

where E_{cv} denotes the energy of the CV at time t. The terms $\dot{Q}$ and $\dot{W}$ correspond to the net rate of energy transfer by heat and work across the boundary of the system at time t, respectively. The last two elements correspond to the rates of transfer of internal, kinetic and potential energy of entering and exiting streams. These are the added components regarding expression (1.1).

Since the work is always performed on or by CV, it is convenient to separate it in two contributions: the first contribution regarding the work associated to the fluid pressure as mass is introduced at inlets and removed at exits. The other include all other works effect associated to rotating shafts, displacement of boundaries, electrical effects, and so on. The first work can also be expressed as the required energy to push the fluid into or out the control volume. Thereby, the normal force can be expressed as the pressure P of the fluid on the enter / exit surface A, and the fluid displacement d. Thereby, the work performed on CV can be expressed as:

$$\dot{W} = \dot{W}_{vc} + (P_{out}A_{out})d_{out} - (P_{in}A_{in})d_{in} \tag{1.7}$$

where P, A, and d are the pressure, the area, and displacement of fluid in control volume, and the index *out* and *in* means at exit and inlet, respectively.

Expressing the A and d as a volume and in a unit basis, it remains:

$$\dot{W} = \dot{W}_{vc} + (P_{out}v_{out}) - (P_{in}v_{in}) \tag{1.8}$$

where v refers to the volume per unit of mass.

Evaluating the rate of work in the energy in equation (1.1) and applying the definition of specific enthalpy ($h_i = u_i + P_i\, v_i$), it concludes

to:

$$\begin{aligned} \frac{dE_{cv}}{dt} &= \dot{Q} - \dot{W}_{vc} \quad + \dot{m}_{in}\left(h_{in} + \frac{v_{in}^2}{2} + g\, z_{in}\right) \\ &\qquad\qquad - \dot{m}_{out}\left(h_{out} + \frac{v_{out}^2}{2} + g\, z_{out}\right) \end{aligned} \tag{1.9}$$

In practice, there may be several locations on the boundary through which the mass enters or exits and considering the same height, same specific velocity and steady state conditions:

$$\frac{dE_{cv}}{dt} = 0 = \dot{Q} - \dot{W}_{vc} + \sum \dot{m}_{in}\, h_{in} - \sum \dot{m}_{out}\, h_{out} \tag{1.10}$$

1.6.4 The Brayton Cycle

The Brayton cycle was proposed by George Brayton. It is a thermodynamic cycle that describes the operation of gas turbine in its four processes: compression, heating, expansion and recover.

In ideal conditions, both compression and expansion are treated as isentropic processes, while heating and recover are considered isobaric, that means the pressure is considered constant along the process. A schematic diagram of the most simple gas turbine power plan system is shown in Figure 1.5.

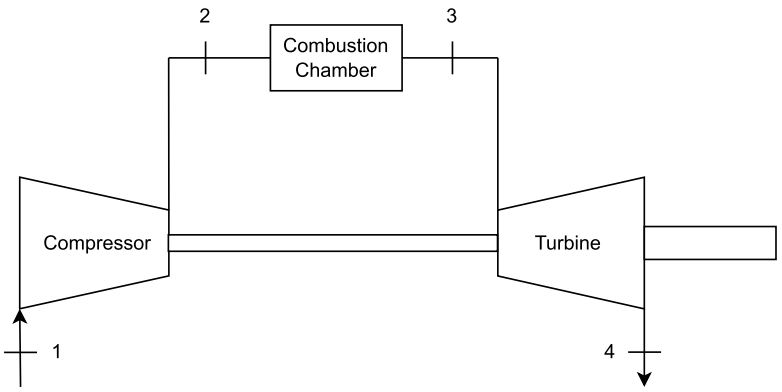


Figure 1.5: Schematic of basic gas turbine components

Figure 1.6 shows the two most common representations of Bryton cycle considering either Pressure P and Volume V or Temperature T and Entropy S on the axis.

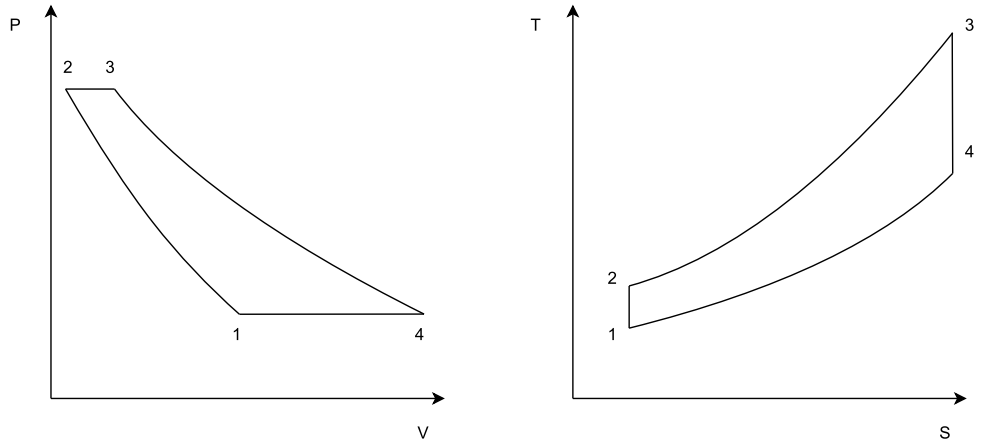


Figure 1.6: Ideal Brayton cycle

The two most important thermodynamic processes in GT functioning are:

1. **Compression**. It takes place in compressors, pumps and diffusers. The fluid pressure is increased by reducing its volume.
2. **Expansion**. It takes place in turbines, nozzles, etc. The fluid pressure is reduced to extract the net work.

In real life, the ideal process for compression and expansion are isentropic processes. This means that there is no heat transfer to or from the working fluid and the work transfers of the system are frictionless ($Q = 0$ and $\Delta S = 0$), indeed adiabatic and reversible processes. The differences between real and ideal processes are shown in Figure 1.7, where s suffix correspond ideal values, also named isentropic values.

The way to compute the efficiency in this kind of processes is by comparing the reversible and irreversible adiabatic process. In the compression, the irreversible process ($1 \rightarrow 2$) needs more work to

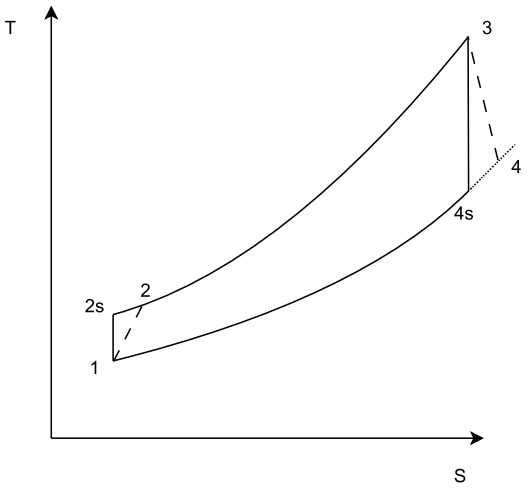


Figure 1.7: Graphical representation of the differences between real and ideal adiabatic processes

drive the compressor than the reversible process ($1 \rightarrow 2s$). This is expressed as:

$$\eta_C = \frac{W_{as}}{W_a} = \frac{h_{2s} - h_1}{h_2 - h_1} \tag{1.11}$$

In the expansion, the adiabatic irreversible process ($3 \rightarrow 4$) obtains less net power than the reversible adiabatic process ($3 \rightarrow 4s$). This is expressed as:

$$\eta_T = \frac{W_a}{W_{as}} = \frac{h_3 - h_4}{h_3 - h_{4s}} \tag{1.12}$$

The above expressions can be rewritten for thermally and calorifically perfect gas, i.e. the ideal gas law, in terms of total pressure and temperature. Thereby, the compressor efficiency can be expressed as:

$$\eta_C = \frac{\left(\frac{P_2}{P_1}\right)^{\frac{\gamma-1}{\gamma}} - 1}{\frac{T_2}{T_1 - 1}} \tag{1.13}$$

The process between 1 and 2s can be defined by the following equation state:

$$\frac{P}{\rho^n} = constant \tag{1.14}$$

where n defines the process from point 1 to point 2s in Figure 1.7. Therefore, rewritting 1.13

$$\eta_C = \frac{\left(\frac{P_2}{P_1}\right)^{\frac{\gamma-1}{\gamma}} - 1}{\left(\frac{P_2}{P_1}\right)^{\frac{n-1}{n}} - 1} \tag{1.15}$$

The turbine efficiency can be expressed as:

$$\eta_T = \frac{1 - \frac{T_4}{T_3}}{1 - \left(\frac{P_4}{P_3}\right)^{\frac{\gamma-1}{\gamma}}} \tag{1.16}$$

Moving from the components efficiency, the overall gas turbine efficiency is given by the ratio of the net energy output to the energy input to the turbine.

$$\eta_{cyc} = \frac{W_{cyc}}{Q_{cc}} \tag{1.17}$$

where W_{cyc} is the work completed by the gas turbine, indeed the net work produced by the turbine and the work consumed by the compressor, and Q_{cc} is the heat added by the combustor. Then, according to the theory explained before and the standard Brayton cycle shown in Figure 1.6, the expression (1.17) can be transformed to:

$$\eta_{cyc} = \frac{W_T - W_C}{Q_{cc}} = \frac{(\dot{m}_a + \dot{m}_f)(h_3 - h_4) - \dot{m}_a (h_2 - h_1)}{(\dot{m}_a + \dot{m}_f) h_3 - \dot{m}_a h_2} \tag{1.18}$$

where $\dot{m}_a$ is the air mass flow, $\dot{m}_f$ is the fuel mass flow, and h_i is the corresponding entropy. Assuming that $\dot{m}_a \gg \dot{m}_f$, that there is not mass losses, and defining the CV around the turbine and compressor in one block, can be obtained:

$$\eta_{GT} = \frac{W_{net}}{Q_{in}} = \frac{Q_{in} - Q_{out}}{Q_{in}} = 1 - \frac{h_4 - h_1}{h_3 - h_2} \tag{1.19}$$

Considering the air an ideal gas and rewriting equation (1.19) using isentropic performance of the compressor and turbine (η_T and η_C, respectively), it is easy to get:

$$\eta_{GT} = \frac{\eta_T \left(1 - r_p^{\frac{1-\gamma}{\gamma}}\right) - \frac{\frac{T_1}{T_3}}{\eta_C}\left(r_p^{\frac{1-\gamma}{\gamma}} - 1\right)}{1 - \frac{T_1}{T_3} - \frac{\frac{T_1}{T_3}}{\eta_C}\left(r_p^{\frac{1-\gamma}{\gamma}} - 1\right)} \tag{1.20}$$

where r_p is the pressure ratio of the compressor, indeed the ratio between output and input pressure, γ is the specific heat rate C_p/C_v, T_1 is the compressor inlet temperature, and T_3 is the turbine inlet temperature.

From the expression above it can be drawn that increasing the pressure ratio and the turbine firing temperature increases the Brayton cycle's adiabatic thermal efficiency. Moreover, there is an other important variable that can affect the thermal efficiency of the cycle that is the inlet temperature in the compressor.

In this section, a brief intuition has been given about the analysis of the most simple gas turbine operation. Further variations of gas turbines configurations exists but these are beyond this thesis. Short explanations about the gas turbines configurations can be found in Appendix A.

1.7 Maintenance Operation

In order to achieve profitability and competitiveness in industry, it is crucial to maintain high reliability, constant system availability, and prioritize human safety while minimizing environmental impact [45]. A real concerns in design and process engineering are the operational changes related to challenging weather conditions, user decisions beyond the nominal operational regime, and, especially, non-reported hash conditions [46]. The current trend of effective management strategies from unexpected events has been sharpened to avoid major economical, environmental, and security defects [47]. Thereby, they are beyond the restoration of the system to a profitable operational point. A proper health management system as well as a good condition assessment method are mandatory.

Maintenance is a critical activity that helps to maintain optimal machine performance and prevent structural degradation, which can result in system failures. While major maintenance operations can restore IGTs to a functional state, they do not return the equipment to its original condition. Instead, they bring the IGTs to a new operational point that ensures smooth functioning and minimizes further deterioration [48].

Maintenance strategies can be classified into three main categories, in increasing order of complexity: corrective, preventive, and predictive maintenance [49].

1.7.1 Corrective Maintenance

Corrective maintenance is the simplest approach to deal with system conditions. It is performed only when a component of the system

breaks down [50]. It is also known as unplanned, run-to-failure, or reactive maintenance [51]. As well as being the simplest approach, it is also the less effective one, as the cost of interventions and the associated downtime are usually more substantial than those with planned corrective actions taken in advance.

1.7.2 Preventive Maintenance

Preventive maintenance is regularly performed according to a plan scheduled in advance, regardless of the health status of the equipment. The plan is based on suggestions of experienced equipment manufacturers, historic breakdowns or failure data, operating experience, and the judgement of maintenance staff and technicians. Therefore, the goal of this strategy, also called planned maintenance, is to increase the availability of the system by slowing down the deterioration processes leading to faults [52]. Maintenance actions are planned on a time or usage trigger [53]: maintenance based on a time trigger, also known as Time Based Maintenance (TBM), includes actions that are carried out periodically, whereas maintenance based on a usage based trigger, named Usage Based Maintenance (UBM), includes actions that are planned according to a process iteration such as a certain amount of production cycles. Although preventive maintenance can reduce the probability of system failures and the frequency of unplanned emergency repairs, it cannot completely eliminate the occurrences of random failures [54].

1.7.3 Predictive Maintenance

This maintenance strategy aims at detecting possible defects, based on estimate health of the equipment, and fixing them before they result in failure. It attempts to avoid unnecessary maintenance tasks by taking actions reactively. Indeed, it is a proactive process which requires the development of a system model that can trigger an alarm for corresponding maintenance [55]. Several disciplines have been developed on this strategy along time, such as Condition Based Maintenance (CBM) [56], Structural Health Monitoring (SHM) [57], and Prognostics and Health Management (PHM) [58]. The most common approaches that are used nowadays in industry are based on ML and, more precisely, ANN techniques [59].

1.7.4 Maintenance Policy

The selection of the most suitable maintenance strategy for IGTs should be based on the specific needs of the equipment and the organization, taking into account the cost-benefit analysis, which involves comparing the costs of implementing the maintenance strategy against the potential savings that can be achieved by avoiding equipment failures and downtime [60]. For instance, a high-risk system that is critical to the operation of the organization may require a more proactive approach, such as preventive or even, predictive maintenance, to minimize downtime and prevent catastrophic failure. On the other hand, a low-risk system that has a minimal impact on the operation of the organization may be better suited for a more reactive approach, such as corrective maintenance, to reduce maintenance costs

Several criteria can be used to determine the best maintenance policy for an industrial system. In a failure based approach, it can be decided by two factors: frequency and development time, as indicated in Table 1.2.

Table 1.2: Maintenance policy can be decided by two factors: failure frequency and failure development time

Frequency vs Dev. Time	Monitored	No Development Time
Regular Frequency	Preventive / Predictive	Preventive
Random Failures	Predictive	Corrective

For failures occurring in a reasonable regular frequency, preventive maintenance in any method could be a proper policy. If failures can be monitored in any way, either they have development time or not, a predictive maintenance approach could be a proper choice. Therefore, failures occurring randomly and having a development time can be properly solved using a predictive maintenance. However, if there is neither a development time nor a CM system, a corrective maintenance approach must be employed [61]. With the increasing availability of data and the development of intelligent tools for condition and fault assessment, it has become possible to apply advanced data analytics techniques to industrial gas turbine maintenance operations. The large volume of data obtained from IGTs provides insights into the condition of components, malfunctions, and warnings, which can be used to determine the state of the system and identify potential issues before they lead to failures. ML algorithms and other advanced analytical techniques can be applied to this data to automatically learn insights and recognize hidden patterns, which can help improve maintenance decision-making and optimize maintenance schedules. This can lead to more efficient and cost-effective

maintenance operations, ultimately improving the reliability, availability, and safety of industrial gas turbines [62].

The current maintenance policy of SE gas turbines is preventive maintenance. It consists on two main categories: operation and periodic maintenance. The *operation maintenance* is a regular attention given to the gas turbine where trends and abnormalities can be determined at an early stage. It is usually done by the site personnel and basically, it consists of monitoring the drive train operation, which includes lubrication, calibration, and function checks as well as the required filter replacement and cleaning. The *periodic maintenance* is defined for a gas turbine regarding a Key Performance Indicator (KPI). It determine approximately when the maintenance must be performed according to the system working time combined with several fatigue factors, for instance, the number of starts and stops, the output power set point, and so on. This maintenance includes necessary services activities form many important functions in the gas turbine system.

1.8 Objective and Research Questions

This study is part of a collaboration between SE and UPC in the shared business-academic approach for project identification, implementation and management related to the digitalization on power generation. SE gas turbines are equipped with a plethora of sensors that collect vast amounts of data, which is sent back to the office. The engineering perspective of this information can be limited but, by leveraging the the potential of advanced machine learning techniques, intelligent supporting tools for gas turbine assessment can be developed to help engineers on this task.

The gas turbine performance improvement can be driven from many different perspective. The framework of this study will put the focus on SE maintenance policy. With the current data acquisition system in SE gas turbines, either unexpected events or unappropriated behaviors of the system can be captured. Then, Data Analytics (DA) and ML algorithms can be employed in order to design and implement engineer's supporting tools for IGT operation as well as for a better understanding of the equipment condition.

Thereby, the present work addresses the following objective:

Main objective: Improve the assessment system of the IGT to ensure more reliability and working time availability of the system by building a proper diagnosis intelligent tool to implement a predictive maintenance policy.

This goal has been breakdown in the following sub-objective:

1. Review the published scientific literature on the application of AI algorithms to enhance IGT's performance
2. Analyze the SE fleet to select similar and significant GTs to work with
3. Study the thermodynamic performance of selected GTs for gaining further insights about its operation
4. Define data-driven indicator to assess the GT performance
5. Validate the indicator using both curated and SE gas turbines fleet data
6. Develop mathematical models to capture the GT behavior using novel AI techniques

7. Develop an analytical framework to assess models performance
8. Verify models using long-term data

To achieve these goals, the procedure has been breakdown in four specific tasked to be performed:

Firstly, sets of sensor data will be analyzed from the same machine in the same operation mode, indeed working on full load, to evaluate how it works in steady state conditions.

Secondly, the degradation of machine elements and weather conditions can cause the GT to operate outside its ideal operating point over time. ML models will be developed to identify trends that appropriately represents the behavior of the machine over the time.

Then, an algorithm will be designed to fine-tune and reproduce the results for a set of IGTs. By analyzing the data obtained from each machine sensors, ML algorithms will be applied to observe how the machine's behavior changes under similar conditions in the historical data.

Finally, extra data regarding warnings and faults events will be used to drawn insights about observed machine operation changes. This will give a further understanding about the main reasons why the machine operation changes over time.

The expected results are:

- Identify significant trends in the degradation of gas turbine performance over time, which will serve as a starting point for the development of an appropriate mathematical model to accurately assess its behavior

- An AI-based indicator to evaluate the performance of gas turbine operation and provide decision-making support for maintenance operations
- An advanced approach to evaluate and analyze the structure of the model, as well as its performance giving further insights about GT operation

1.9 Scope of the Study

The scope of this study is focused on developing and implementing various DA and AI based algorithms to assess IGT performance when operating at full load. The research seeks to evaluate the use of various ML tools, including supervised and unsupervised algorithms, as well as statistical process control techniques, to detect and diagnose issues related to GT operation. By utilizing these tools and techniques, the research aims to provide a more comprehensive understanding of the complex behavior equipment when operating at full load, offering insights and recommendations for improved GT performance and maintenance strategies. The main SE maintenance operation is done in the compressor, therefore the data used to perform all the analysis will be focused on this component. Overall, the study provides a valuable contribution to have further understanding of SE gas turbine maintenance operation, more precisely maintenance operation in the compressor, and new KPI to assess it. Any other data or activity beyond compressor maintenance performance is out of scope of this thesis.

1.10 Limitations and Assumptions

One of the limitations of using real plant data to analyze gas turbine performance is the potential privacy concerns surrounding the data used. Many IGTs are owned and operated by different customers, and the data collected from these systems may contain sensitive information that needs to be protected. While anonymizing the data can help to mitigate privacy concerns, there is still a risk of data breaches or unauthorized access to the data. Therefore, it is important to understand that part of the data will not be delivered from SE and thus, some gaps on the research will be considered regarding the components types.

On the other hand, it must be noted that the company's objective of obtaining simple models to maintain human interpretability of the results may also pose a limitation in finding the optimal model that can fit the data using fewer parameters while maintaining high accuracy and providing meaningful insights into the reasoning behind the model. Balancing the trade-off between model complexity and interpretability has been a crucial aspect of this research, as it seeks to develop models that are not only effective but also understandable for practical applications and customer implementation

Two main assumption lead the focus of this study. Firstly, the SE gas turbine operation is limited by the TIT. It is limited because if it exceeds a certain point, it can cause serious damage to the system. In this study, it is assumed that when the machine is operating at it full load, indeed at its maximum capacity, it is using most of its available resource. Thereby, it is when is easy to capture and visualize the degradation of the system. This hypothesis will lead the way how the data is treated in this research.

Secondly, the operation of SE IGT during the first year is considered that is working at its optimal regime. This will help on properly create a fresh model as a reference to study the deviation of the IGT performance in long-term.

1.11 Significance of the Study

The proposed study aims to significant impact on the field of gas turbine performance analysis and enhancement. First, by bridging the gap between academia and industry, the study promotes a better understanding and collaboration between these two important sectors. Second, by using state-of-the-art ML algorithms, the study offers a more accurate and efficient way of analyzing and enhancing IGT performance. Finally, by improving the tools used to assess GT behavior, the study provides a more comprehensive understanding of these complex systems, leading to better decisions and actions that can optimize performance and reduce downtime. Overall, the study offers a valuable contribution to the field of IGT technology, opening up new possibilities for enhanced maintenance operation performance and more sustainable energy production.

1.12 Implication for Industry and Academia

The benefits of conducting research in the area of gas turbine performance analysis and enhancement are significant for both industry and academia. For industry, such research can lead to more efficient and cost-effective GT operation, improving competitiveness and profitability. By using advanced ML algorithms and tools to assess IGT behavior, industry can identify potential issues and take

corrective action more quickly, reducing downtime and maintenance costs. In addition, a better understanding of GT performance can lead to the development of more efficient and sustainable energy systems, reducing environmental impact and improving public perception of the industry.

For academia, research in this area provides an opportunity to advance scientific knowledge and develop new technologies, contributing to a better understanding of complex systems and enhancing the field of energy engineering. By collaborating with industry, academia can also contribute to the development of real-world solutions to practical problems, leading to more impactful research outcomes and new avenues for research and innovation. In conclusion, research in the area of IGT performance analysis and enhancement has significant benefits for both industry and academia, offering opportunities for technological advancement, cost reduction, and sustainability improvements.

1.13 Structure of this Document

The structure of the dissertations is as follows.

In Chapter 1, a comprehensive overview of AI and its various types is presented. Furthermore, the chapter provides an extensive introduction to IGTs, encompassing the power generation industry, the equipment with its constituent components, the fundamental thermodynamic cycles, and the diverse types of maintenance. Finally, an explanation is offered regarding the limitations of conventional maintenance techniques and the potential benefits of incorporating ML methodologies to enhance maintenance decision-making.

In Chapter 2, the primary components necessary for the successful implementation and evaluation of digitalization in IGTs, basically focusing on maintenance policies, are mentioned. These components encompass data acquisition systems, which involve the mechanism of capturing and returning data to the engineering office for performing specific analyses, data processing, involving the application of algorithms to process the data for machine health status analysis and anomaly detection, and maintenance decision support, which entails the evaluation of machine operations based on diagnosis and prognosis techniques.

Chapter 3 delineates the methodology employed in this study, along with the algorithms utilized for modeling and analyzing the IGT system. A comprehensive explanation of the data treatment methodology, including the filtering methods and the partitioning of the dataset, is provided. Additionally, the main ANN based architecture to model the equipment, indeed an Autoencoder (AE) based architecture, including their architecture, types, and structure, are elaborately explicated. A detailed examination of the key parameters of the proposed model is also presented. Finally, the DA and ML methodologies utilized in this study to analyze the performance of the proposed model are discussed.

Chapter 4 showcases the outcomes obtained by implementing the aforementioned methodologies to IGT data. Firstly, the primary results pertaining to the thermodynamics are presented. Subsequently, the performance of the AE models based on their various types and structures is explicated in detail. Finally, the results of applying significant ML methods with the objective of analyzing the model performance are also presented.

In Chapter 5, an in-depth discussion regarding the data treatment methods utilized in the study are presented. The thermodynamic evaluation and the performance of the AE concerning IGT data are also analyzed and discussed in detail. Furthermore, the ML model architectures proposed in the study are thoroughly denoted.

Chapter 6 offers the conclusion of the study, elucidating the key processes and findings derived from this research. Additionally, based on the outcomes obtained, some recommendations and potential areas for future research are presented. Furthermore, the publications resulting from this study are also presented.

Appendix A and Appendix B serve to clarify or expand over important concepts presented in this study. Appendix A provides an elaboration on the various enhancements of the basic Brayton cycle that can be employed to enhance the accuracy of gas turbines. On the other hand, Appendix B offers a more in-depth explanation of significant ANN processes, specifically the learning process involved in ANN.

1.14 Main Outputs of the Ph.D. Thesis

The dissemination of the research findings has been conducted through multiple channels to ensure that relevant stakeholders both within the company and the academic sector have been reached.

Primarily, the research has been presented in several internal sessions at Siemens Energy company, which were attended by managers and experts in the field. The purpose of these sessions was to increase the visibility of the research within the company and propose potential applications for custom solutions to customers.

In addition, the research results have been presented at various industry conferences, such as the Advanced Factories Expo & Congress [63], and academic workshops, such as the VI Conference of the Predoctoral Researchers of Universitat de Girona (UdG) [64]. These conferences provided an opportunity to disseminate the findings to a broader audience and engage in discussions with other experts in the field.

Furthermore, the research has been published in several peer-reviewed journals that are highly regarded in the field. These publications have helped to disseminate the research findings to a wider audience and contribute to the academic discourse on the topic.

1. Martí de Castro-Cros, Stefano Rosso, Edgar Bahilo, Manel Velasco, and Cecilio Angulo. Condition assessment of industrial gas turbine compressor using a drift soft sensor based in autoencoder. Sensors, 21(8):2708, 2021.

   - **Journal**: MDPI Sensors (ISSN: 1424-8220)
   - **Impact Factor**: 3847
   - **5-Year Impact Factor**: 4050
   - **Quartile**: Q2
   - **Visualizations**: 1398
   - **Cites**: 8

2. Martí de Castro-Cros, Manel Velasco, and Cecilio Angulo. Machine-learning-based condition assessment of gas turbines—a review. Energies, 14(24):8468, 2021.

   - **Journal**: MDPI Sensors (ISSN: 1996-1073)

- **Impact Factor**: 3252
- **5-Year Impact Factor**: 3333
- **Quartile**: Q3
- **Visualizations**: 2394
- **Cites**: 8

3. Martí de Castro-Cros, Manel Velasco, and Cecilio Angulo. Analysis of gas turbine compressor performance after a major maintenance operation using an autoencoder architecture. Sensors, 23(3):1236, 2023

   - **Journal**: MDPI Sensors (ISSN: 1424-8220)
   - **Impact Factor**: 3847
   - **5-Year Impact Factor**: 4050
   - **Quartile**: Q2
   - **Visualizations**: 451
   - **Cites**: 0

Overall, the dissemination efforts have ensured that the research findings have been widely disseminated and are accessible to a variety of stakeholders in both the industry and academic sectors.

Chapter 2

State of the Art

The literature about maintenance policy is huge and diverse basically for its wide variety of systems, components and parts. Significant papers appear every year in academic journals, conference proceedings and technical reports that includes theories and practical applications. The strategy to follow for a proper maintenance policy may vary according to the equipment type and the customers needs. Therefore, the maintenance strategy must be carefully selected taking into account all possible requirements. When a proper maintenance program is established and effectively implemented, it can significantly reduce maintenance cost by reducing the number of unnecessary scheduled maintenance operations leading to a reduction in the cost [65].

The main components to define, implement and evaluate maintenance policies are: (i) data acquisition, indeed digitalization, (ii) data processing, and (iii) maintenance decision support, which includes diagnosis and prognosis.

2.1 Data Acquisition

Data acquisition is the process of collecting and storing useful data from the equipment. This process is an essential step to assess the maintenance policy and to implement new maintenance programs for machinery failure diagnostics and prognostics. Data collected can be categorized into two main types: event data and condition monitoring data.

Event data include useful information of the maintenance actions performed to the targeted equipment such as equipment installation time, performed last maintenance time, components change, and so on, where the main data source are from operators and maintenance personnel.

Condition maintenance (CM) data is about measurements that are related to determining the condition of the IGT or any of its components, i.e., its health status. Main data source are sensors placed strategically in the IGT system. This information is relevant because it can capture significant changes in the behavior of the system to detect and identify potential failures. Specifically in IGT, CM systems can be implemented using various methods and technologies with varying levels of fidelity and unique advantages [66]. This step is quite important because is when the component information is captured and this is the data that is going to be used for the next algorithms.

Another important issue that these techniques must face are the detection of anomalies in the captured data in order to avoid false alarms and to improve the reliability. Both of the processes, health status monitoring and anomaly detection, can use similar data to

achieve their goal. However, for each case data is processed in different ways. This data processing phase is explained in the next section.

2.2 Data Processing

Before going in depth in data treatment for health status and anomaly detection processes, it is important to point out that the first step for data processing is data cleaning. It is mandatory since data always contains errors. Data cleaning ensures, or at least increases the chance, that good data is used for further analysis and modeling. Data errors are caused by many factors including the human factor as long as some event data is generate manually. For CM, data errors may be caused by sensor faults. In this case, sensor fault isolation must be performed. In general, however, there is no simple way to clean data. Usually it requires manual examination of data and sometimes graphical tools can be very helpful to finding and removing data errors. Data cleaning is, indeed, a big area.

2.2.1 Health Status

The process of going from observed data to a mathematical model has become fundamental in industry in order to properly identify the optimal operation. This issue has been faced from several points of view. Some authors have shown promising results by using the information of the entire system, while some others have broken down the machine into its components and assessed the condition locally.

When considering the entire system, as explained in Section 1.4, one of the main feature to identify and evaluate the performance of a

GT is the exhaust temperature [67]. Several methods together with various metrics have been used to proper determine the condition of the system as well as to design its own self-learning features [68]. Moreover, with the enhancement of novel advanced technologies, it have been easy to generalize the solutions for varying working conditions [69]. In all the cases, promising results are showed in operationally changing context and showed that it is possible to train accurate models.

From the component point of view, the main considerations are focused on the combustion chamber and exhaust gas temperature, as well. The most common issue to tackle is that the way how the combustor operates may vary and can considerably affect the GT performance. Therefore, having proper methods to identify the combustion stability as well as to optimze the operating conditions can be crucial [70]. Moreover, capturing the transition from the stable to unstable regime can also be differential in understanding the operational mode of GT [71]. Regarding other main components, there is not much literature that aims to extrapolate the enhancement of the local performance to the whole system.

Finally, more complex frameworks exist aiming to tackle condition monitoring and to handle some extra issues such as anomaly detection, diagnosis and prognosis all in one and for several components [72]. This kind of frameworks needs a practical investigation through strategic, tactical, and operational levels [73], but although some integrative proposals demonstrate huge potential in this kind of frameworks, a lot of research needs to be done in order to have a proper implementation in industry [56].

2.2.2 Anomaly Detection

Access to reliable data captured from GT sensors is essential to obtain good monitoring practices. Anomaly detection has been widely applied to monitor asset operation status, as well as to provide its health status. However, due to the evolution to extreme complex industrial systems, many challenges have arisen for the classical anomaly detection approaches. Therefore, having ML models that are able to automate the construction of anomaly detection approaches from available data ease the concern about having a proper model.

Several approaches are proposed for this kind of solutions regarding ML types. The most evident and more understandable methods are the supervised and semi-supervised learning, where faulty data is added in the training of the models to properly detect the anomalies [74].

Unsupervised learning has also shown satisfactory applications in the field of anomaly detection where some clustering based methods are proposed to detect and filter outliers [75]. Moreover, more advanced methods are presented regarding the feature extraction capabilities of certain ANN models [76], indeed Convolutional Neural Network (CNN), or the ability of finding hidden patterns in input data [77], indeed autoencoder (AE) models.

Finally, some other systems are proposed to tackle with anomaly detection solutions that also proposes future possibilities for including diagnostics and health management methods to the system [78].

2.3 Maintenance Decision Support Process

In order to make informed decisions regarding maintenance tasks, it is crucial to provide maintenance personnel with effective and comprehensive tools to evaluate the condition of the system [79]. These are divided in two categories: diagnostics and prognostics [80]. *Diagnostics* consists of three main steps: fault detection, isolation and identification. The objective of *fault detection* is to identify any deviations from the expected behavior of a system using the CM approach. On the other hand, *fault isolation* involves the task of pinpointing the specific component or subsystem responsible for the detected anomaly. Lastly, *fault identification* entails determining the root cause of the fault. *Prognostics* is a more advanced diagnosis that aims to detect future faults before it occurs. Fault prediction is a task to determine whether a fault is impending and estimate when a fault will occur. Diagnostics is posterior event analysis and prognostics is prior event analysis. Prognosis is a much more efficient task than diagnosis to achieve zero-downtime performance but it can be treated as an improvement in the maintenance decision support process. However, diagnosis is required to determine the nature of an unexpected failure. Next, some review according to this two types of conditioning are provided.

2.3.1 Diagnosis

The primary objective of diagnosis is to determine the underlying cause of a failure, which needs the use of an effective CM system. The two main characteristics that differentiate diagnostic methods are the parameters used and the methodology employed for analysis.

Typically, three categories of performance monitoring approaches are used: model based, data-driven, and experience based methods.

Model based methods involve analytical modelling of system operation and are quite promising for real-time condition monitoring. Model based approaches have proven their ability to detect both abrupt and more notably, gradual degradation in engine performance in real-time online implementations. However, as the modeling uncertainties and system complexity increase, the monitoring accuracy decreases. Kamboukos and Mathioudakis [81] compare linear and non-linear GTs diagnostic approaches and present an overall assessment of their merits and weakness with the conclusion that the use of linear methods might lead to substantial inaccuracies in estimation of degradation. The inadequacy of the linearity assumption has led to the development of non-linear alternatives. Although all data-driven approaches are non-linear, model based methods involve both linear and non-linear methods.

The most common method is Gas Path Analysis (GPA), which by monitoring the deviation of engine health variables or independent parameters aims to detect the deterioration of the equipment. Health components are not directly measurable however there is a thermodynamics correlation with these parameters such as the pressure and temperature. Therefore, the gas-path faults have an observable effect on the measurements. The first GPA method was introduced by Urban [82], which is now referred to as linear GPA, using a Influence Coefficient Matrix (ICM). Nevertheless, the engine operates under a non-linear thermodynamics behavior, thus the linear performance was not able to deliver an accurate estimation of the deviation. A promising tool was developed using this method by David L. Doel named TEMPER [83]. Escher and Singh [84] present an approach

using a Newton-Raphson iterative method to solve the non-linearity relationship between the health parameters and the sensor measurements. This solution obtain better results and multiple fault diagnoses were able to be performed. However, the limitation is that degradation is only detected in small scale. Moreover, the accuracy is highly affected by an inaccurate ICM, a high measurement uncertainty and correlated measurements. Another approach is studied by Doel [85] and implemented in the General Electrics GPA tool. This method is based on the weighted least square technique, where a parametric model is fitted by minimizing the sum of weighted squared deviation between the actual and predicted measurements. In addition, the measurement uncertainties associated with the gas-path measurement are also considered using a weighting matrix with their respective sample variances. In a similar way, this technique was before used by Rolls-Royce in the COMPASS diagnostic system presented in the Barweli study [86].

Another method is the state variable estimation of dynamic systems using the Kalman Filter (KF) technique. It is an optimal recursive data processing algorithm used to estimate the health of the engine components in the presence of measurements noise and bias. KF evaluates all available measurements data and prior knowledge about the system to produce an estimate of desired variables with the statistically minimized error. The most common variant for engine diagnostics is the Linear Kalman Filter (LKF) [87]. This kind of technique can be either used to diagnose the whole system as well as detect specific component, sensor or actuator fault [88]. To treat with the non-linearities using the KF technique, two variants were also developed: the Extended Kalman Filter (EKF) and the Unscented Kalman Filter (UKF). EKF uses Taylor series to linearly

approximate a non-linear function around the mean and covariance. UKF is a step forward of the EKF, which uses deterministic sampling to form a new mean and covariance estimation. Both algorithms have been used in GT's diagnosis as presented in [89] and [90], respectively.

Finally, a last method present in model based performance is Genetic Algorithm (GA). It is an optimization approach based on an heuristic search that follows the procedure of Darwin's natural selection theory. The goal in applications for IGT fault diagnosis is to obtain a set of component parameters that after mutation produce a non-linear model by creating a set of predicted dependent parameters that fits the measurements. This method was first introduced by Stamantis et al [91]. By detecting the shift between the real measurements and the model, the fault can be identified. It shows promising results in comparison with linear and non-liner GPA models [92], also working in presence of uncertainties [93], and in combination with other methods such as neural networks [94].

The main advantage of model based approaches is that they incorporate physical understanding of the monitored system. In many situations, the changes in features vector are closely related to model parameters, thus a functional mapping between the drifting parameters and the selected prognostic features can be established. Moreover, if the understanding of the system degradation improves, the model can be adapted to increase its accuracy and to address subtle performance problems. Consequently, they can significantly outperform data-driven approaches. However, this closed relation with a mathematical model may also be a strong weakness: it can be difficult, even impossible to catch the system's behavior.

The second group refers to *data-driven* models such as ANN systems that 'learn' from examination of real data containing nominal and known faulty conditions. Research results have shown that these methods offer a flexible tool for addressing the complex and non-linear characteristics of dynamical systems. Therefore, they have been extensively exploited and used in diagnosis and prognosis of GTs systems [95]. Moreover, ANN works well with measurements uncertainty and rapid computational speed, hence a wide range of techniques and various ANN architecture types have been developed, mostly for fault detection. This kind of methods, indeed ANN, are further explained in the next chapter.

To perform the analysis of a IGT, prior data is needed to train the model. A common procedure for this analysis is to develop a model that capture the healthy behavior of the equipment in order to accurately detect a shift in the output values according to the real IGT performance. This model can be built by using several types of ANN architecture which can lead to different results [96].

The last group, indeed *experience based* models, includes systems that embody 'rules of thumb' that have been developed and refined by human maintenance experts. Examples of these systems are rule based expert systems and fuzzy logic approaches [97]. Although these methods are capable of offering explanations and methods that reach a particular solution, it is highly complex to find a proper set of rules, functions, and tuning that can obtain a satisfactory solution as the system complexity increases.

2.3.2 Prognosis

Until approximately a decade ago, the focus of engine performance-based health monitoring technologies has been on diagnosis of critical faults. Moving towards an improved insight into engine health, the IGT research community has recently decided to incorporate time evolution into monitoring systems by adding prognostic capabilities. Both model based and data-driven systems are used. Model based approach relies on an analytical analysis of the system whereas data-driven models rely only on previously observed data to predict the projection of a system state or to match similar patterns in the history to infer the Remaining Useful Life (RUL). Because sensors are sources of considerable noise, pre-processing smoothing algorithms can be applied to extract information from historically acquired data.

Physical model based approaches usually employ mathematical models that are directly tied to physical processes which have direct or indirect effects on health of related components. Physical models are usually developed by domain experts, and the parameters in the model are validated by large sets of data. Physical model based approaches used for prognostics require specific mechanistic knowledge and theories relevant to the monitored systems [98].

Physical models are useful in accounting for different operating conditions. With an intelligent monitoring system, most often, they work well under any load profile, including steady-state and transient performance and unanticipated conditions, loads, and operational regimes. Since they incorporate physical understanding of the system for monitoring, in many situations, the changes in feature vectors are closely related to model parameters. Therefore, a functional mapping between drifting parameters and selected prognosis features

can be established [99]. Moreover, if the understanding of system degradation improves, these models can be adapted to increase their accuracy and address subtle performance problems. The limitations are their higher costs and component speciality, which means that they cannot be generalized to other types of components. Furthermore, it is very difficult to build a good physical system. Generally, a combination of various methods must be used to improve the system performance and accuracy [100].

Data-driven prognostic methodology is based upon statistical and self-learning techniques, most of which originated from the theory of pattern recognition. Data-driven methods can be classified into two categories: statistical approaches and ML approaches. Statistical approaches include multivariate statistical methods, state space models and regressive models. Now most of the existing data-driven ML approaches for prognosis have employed ANN and its variants. Data-driven models are usually developed from collected input/output data. These models can process a wide variety of data types and exploit the nuances in the data that cannot be discovered by rule based systems [98].

From a statistical point of view, using the probability density functions of prior observations, the future is forecasted and by detecting the shift between distributions, the fault is predicted [101]. Of course, a proper historical data is needed to model the operating stat and obtain data trends [102].

From a ML point of view, much more applications have been recently developed using a similar approach that statistics. The aim is to proper model the condition of the equipment and by runnning it in future steps, compare the values with real output data [103]. Since

the treatment of data is mainly completed by using time series, RNN based architectures are one of the most tested methods for this kind of job [104].

Finally, some hybrid methods are also developed from the combination of physical and data-driven models [105] as well as, statistics and ML approaches [106].

The field of maintenance policy research is expanding rapidly as technology continues to improve its performance. This applies to each component of the GT as well as the entire system. Promising results have been shown, but there remains a lack of optimal solutions for maintenance strategies. Each solution is custom and specific to a simple task, but not necessarily applicable to the entire system. With the use of ML, some improvements in certain areas are presented leading to a promising research path. However, there remains a pressing need for additional investigation to establish a comprehensive approach to maintenance policy that can be implemented throughout the entire GT system.

Chapter 3

Methodology

In previous chapters most of the relevant theory that is applied to this study has been described. This chapter deals about the data pipeline the global model follows and about the methods that have been carried out in this research. Firstly, insights about the data resources will be presented, as well as the data treatment procedures, and the whole pipeline created around it. Then, the features and the models computed for monitoring the current used IGT system will be exposed. Finally, the last part of this chapter will refer to the methods developed to properly analyse the system performance. Figure 3.1 illustrates a diagram summarizing the employed methodology.

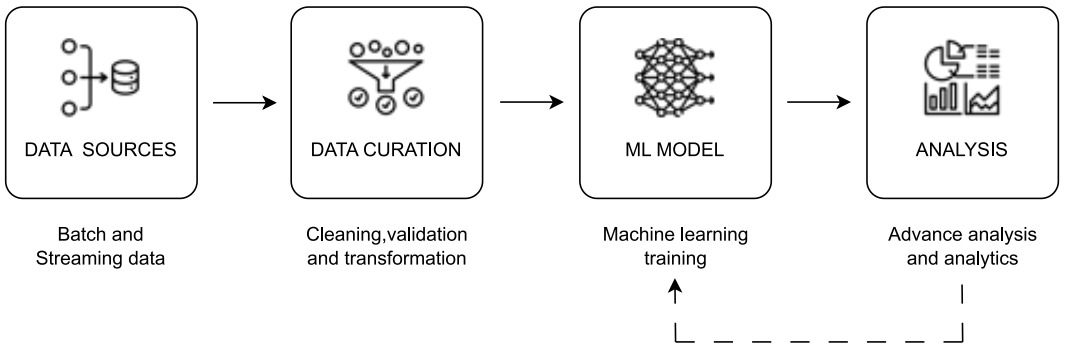


Figure 3.1: Schematic diagram of the methodology

3.1 Data Sources

A significant amount of data is generated continuously in various ways for systems composing industrial gas turbines (IGTs). The maintenance reports database provides valuable information about unexpected events, component repairs, and operation history. Multiple sensors located throughout the turbine gather data on thermodynamic parameters (e.g., compressor inlet temperature, outlet turbine temperature, and emissions) and production parameters (e.g., output power and fuel usage). Despite potential noise and measurement errors, this data reflects the turbine's current state and situation. Therefore, collecting and maintaining this information in a database format for knowledge discovery has become a subject of great interest. Data has always been essential for organizations, but the evolution of data generation and acquisition is transforming industries' working methods. Some experts even suggest that data-driven industries will become commonplace in the future [107].

IGT's data sets consist of a list of multiple multivariate time series. These time series are obtained from sensors placed strategically on the IGT to properly capture the most important features, indeed temperature and pressure, of the system. In this research study, two different data sets are used. On the one hand, a public data set for asset degradation modeling from NASA is used to validate the proposed ML model. On the other hand, a huge database is provided by Siemens Energy that contains hundreds of time series from real IGTs operation.

3.1.1 NASA Turbofan Jet Engine Data Set

The NASA Ames Intelligent System Division with external partners released a Turbofan Engine Degradation Simulation Data set for research purposes [108]. The data set was generated with the Commercial Modular Aero-Propulsion System Simulation (C-MAPSS) dynamic model where real flight conditions recorded onboard a commercial jet were given as input.

The data is composed by normal and faulty operation modes. The training data contains normal operation information of the engine, indeed normal operation mode, until it develops a fault at some point and the system fails, indeed faulty operation mode. The test data contains the same time series but these ends some time prior to system failure. The data is provided with 26 columns of different feature where each row corresponds to a single operational cycle.

The structure of this data is prepared to predict the RUL of each engine in the test dataset. However, in this study, this data will only be used to validate that the proposed ML based models can properly run in already curated data. The results of this validation is presented in the Results chapter.

3.1.2 SE Gas Turbines Data Set

As it was mentioned in Section 1, this thesis is related with a research project developed jointly with Siemens Energy (SE). The company provides us with anonymized data obtained from hundreds of sensors placed in different IGTs all over the world. Access to original data is limited to a table view and two descriptive tables.

Snowflake (cloud database) is the main resource which contains all the useful information for all turbines' operations. SQL statements are used to access to the database as well as to query information. The most used features for the table view are: *sensor timestamp*, when the sensor value was captured; *sensor key*, from what sensor the data is retrieved; *sensor value*, the captured value strictly; *unit key*, the GT where retrieved data come from; and, *good industrial quality*, a reliability index for the captured value.

Moreover, the other two tables contain the description for each sensor and each IGT. The sensor descriptive table match the sensor key value with the a proper description of it and, IGT descriptive table match the unit key value with its nameable type. SE build up different types of equipment regarding various parameters such as the power generation capacity, building components, dimensions, and so on. This is confidential information thus the exact name of the model is not provided. Instead, a nameable type is used to identify the common systems' type.

To better understand the nature of data two different analysis were performed. The first one was a metadata analysis and the second was a sensor analysis. A *metadata analysis* is an structural analysis about how data is organized. It consists of analyzing the number of turbines' types, the number of sensors in each IGT and find patterns between them, before going in-depth in sensor value data. *Sensor data analysis* is a more accurate analysis considering value ranges captured by the sensor. This information could lead to some similar working conditions between different machines and even, between different types of families of IGTs.

3.1.2.1 Metadata Analysis

The aim of the metadata analysis is to determine the most interesting IGT type according to the amount of data stored. Moreover, a set of systems to start working with will be decided by defining a similarity index, that give a way of classifying the turbine inside each type. As mentioned above, the different types of IGTs are confidential information, therefore identification labels are defined, called package types.

The first approach is computed by counting the number of instances in each package type and also, the number of sensors in each instance. Instances from the same package type could have different number of sensors, thus the average of whole equipment that belongs to the same package type is considered. This indicator determines which is the package type with larger data, and the one to start working with.

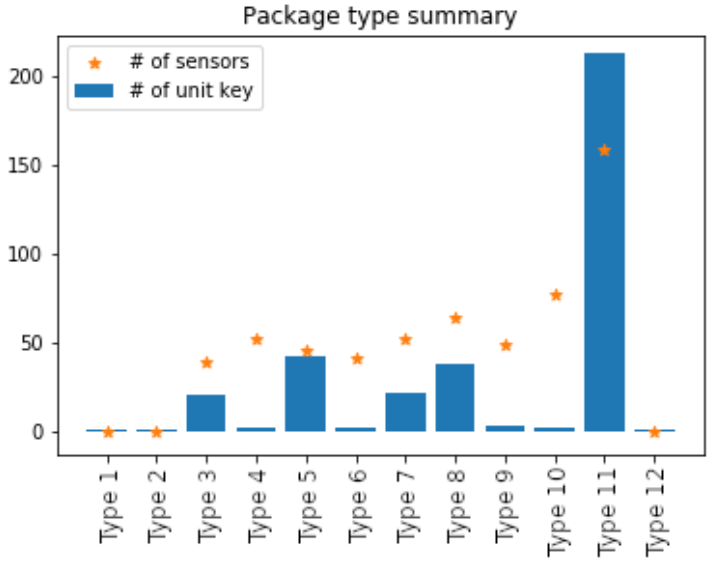


Figure 3.2: Summary of the amount of gas turbines and sensors per package type

Figure 3.2 shows an overview of the distribution of gas turbine units according to their respective package types. The blue bars represent the number of units available for each package type, while the orange stars denote the corresponding count of available sensors. Notably, *Package Type 11* emerges as the most prevalent type, boasting the highest number of both gas turbine units and available sensors.

The first criterion to select the most interesting package type has been conducted only using metadata. Next, a more restrictive criterion is introduced considering sensor values to select specific units.

3.1.2.2 Sensor Data Analysis

In this section a criterion is defined regarding the values captured by the sensors. This will lead to a set of similar machines regarding its performance and its working conditions.

An important concern in IGTs systems are ambient/weather conditions because they can considerably modify its performance. Therefore, a clustering criterion is defined related with similar working ambient conditions for equipment. The working mode for the IGT changes considering the inlet temperature. For inlet temperature values lower than $5°C$, an *anti-freezing* mode is selected. In this situation, ice formation is very plausible and the machine could get damaged. Conversely, when the inlet temperature is above $30°C$ the operation mode is changed to *grid code regulation* mode. In these both cases the machine operation differs from the *normal conditions* mode in order to avoid harmful events. Hence, machines that are working beyond the normal conditions boundaries are rejected (exclusion criteria). Figure 3.3 presents the machine selection among GTs in Package Type 11.

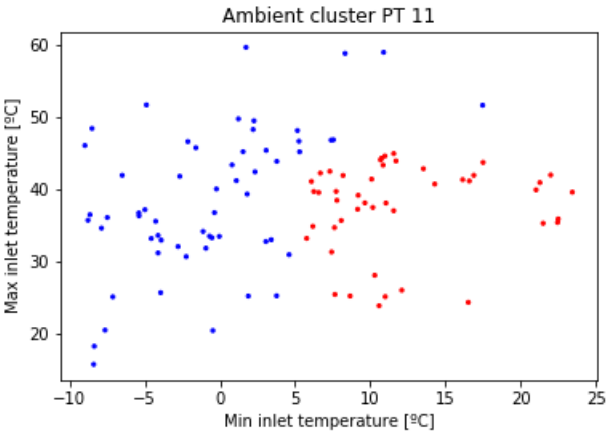


Figure 3.3: Gas turbines' distribution according to the minima and maxima values of the inlet temperature. The blue dot represents the excluded machines, whereas the red dots represent the included gas turbines

For current selection, boundaries are specified between 5 and 30°C for its minimum value and lower than 45°C for its maximum value. Even though the maximum value could enable the grid code regulation mode, using a more restrictive condition, the number of machines is heavily reduced. Moreover, after a in-depth study, the number of points that exceed this limit is not a representative set in the whole time series. Thus, all of the points above this limit are filtered in the pre-processing step. At the end, nineteen (19) IGTs are selected according to this sensor value criterion.

3.2 Features Set

The *features set* refers to the measurements selected to train and test the model. As mentioned in the scope section, the main focus of this study is the compressor component, therefore only compressors' features are used. The main characteristics of the compressor are inlet and outlet pressure and temperature. As previously highlighted, also ambient conditions can considerably affect IGTs performance, and indeed compressor performance. Hence, ambient temperature

and humidity are also selected, leading to six (6) raw features to start working with:

- *Inlet Pressure*, $P_{c,i}$. Input air pressure into the compressor.
- *Outlet Pressure*, $P_{c,o}$. Output air pressure from the compressor.
- *Inlet Temperature*, $T_{c,i}$. Input air temperature into the compressor.
- *Outlet Temperature*, $T_{c,o}$. Output air temperature from the compressor.
- *Ambient Temperature*, T_a. Air temperature in the external.
- *Ambient Humidity*, H_a. Air humidity in the external.

From a machine learning point of view, the number of variables in the feature's subset is very small, so it will be increased to capture not evident information. Exploring in the field of thermodynamics and dynamic systems, a few more features are proposed based on proportionality and derivative states. The ratio features are defined as the proportion between the outlet and inlet; inverse features are included based on the ideal gas law relationship; differential features are obtained from the subtraction of two consecutive points in the corresponding time-series.

Thereby, the subset will also include inlet pressure and temperature, differential inlet and outlet pressure and temperature, ambient humidity, pressure and temperature ratio, and inverse inlet and outlet temperature, that is eleven (11) variables in total,

- *Inlet pressure/temperature*, $P_{c,i}$, $T_{c,i}$
- *Ambient humidity*, H_a

- *Differential inlet/outlet pressure/temperature*, $\Delta P_{c,i}$, $\Delta P_{c,o}$, $\Delta T_{c,i}$, $\Delta T_{c,o}$. Values obtained from the subtraction of two consecutive points in the inlet pressure measurement as well as in inlet temperature and outlet pressure and temperature.
- *Pressure/temperature ratio*, rP_c, rT_c. Defined as the proportion between outlet and inlet pressure as well as outlet and inlet temperature.
- *Inverse inlet/outlet temperature*, $T_{c,i}^{-1}$, $T_{c,o}^{-1}$. They are its inverse value.

This list of features is considered individually into the subset, indeed assuming that there is no correlation among them. Therefore, a correlation test is computed to check whether the information given from each feature is actually independent.

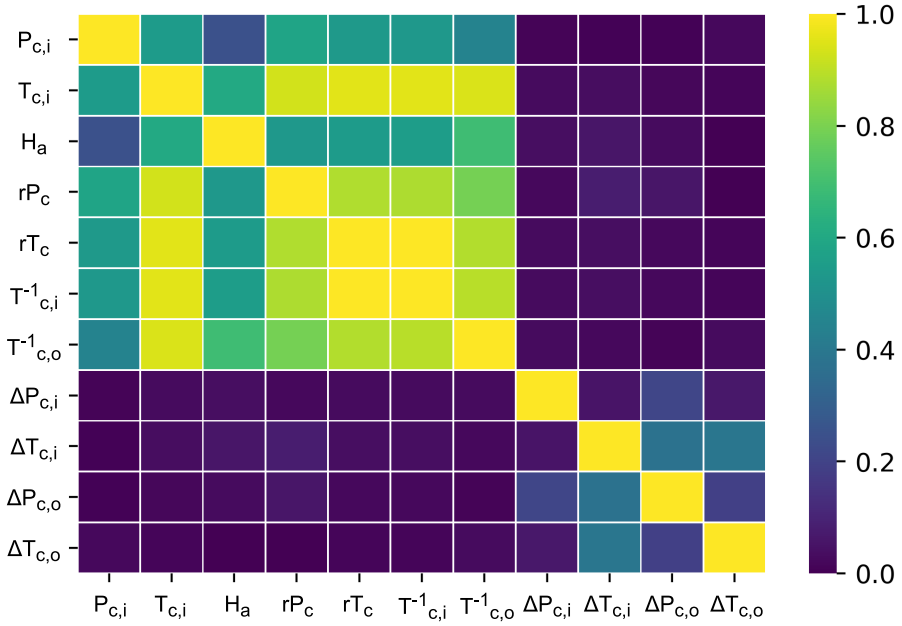


Figure 3.4: Graphical matrix representation for the values obtained from the correlation test between the 11 features considered for the data set. Each label corresponds to a single feature in an abbreviate form

Figure 3.4 shows graphically the results of this correlation test. Two features would be considered dependants whether they exceed a threshold in the correlation test, thus one of them is obviated. The threshold of the correlation value is usually fixed to 0.95. Nine features remain after passing the correlation test, where the temperature ratio rT_c and the inverse inlet temperature $T_{c,i}^{-1}$ are discarded because their correlation factor is up to 0.956 and 0.957, respectively, with the inlet temperature time series. Therefore, the model input is reduced to nine (9) features, $\mathbf{x} = (x^{(1)}, \ldots, x^{(9)})$.

The selected features set is not a static variables selection/definition, instead it is the beginning set point to start working with. Along the research project, this variables selection can be modified according to the accuracy and analysis of the models.

3.3 Data Curation

Data quality is an important aspect which should be considered before any actual analysis. After features selection processing, the values of the features are preprocessed before the model training procedure. Ranges for the values in the features are large and diverse from each other. For instance, outlet pressure values can be 2000 times higher than the differential inlet pressure values. This relative variation could lead to model error because the highest values may get more influence than the smallest ones. Thus, standard scaling is used to normalize the range of independent variables or features of data with mean zero and standard deviation one.

Moreover, a filtering method is also applied for data cleaning to erase the most evident outlier elements. It is carefully defined to avoid

finding out anomalies inside data. The filtering method consists of computing the median for each feature and adding or subtracting k-times the Median Absolute Deviation (MAD) to define both upper and lower boundaries. The median is the value separating the half of a probability distribution, while the MAD is a measure of the variability of a univariate sample of quantitative data. They can be related with the mean and the Standard Deviation (STD) measures; however, MAD is considered a more robust estimator in presence of outliers than STD as well as the median in front of the mean. The k-times constant is heuristically determined by comparing IGT's sensor graphs. It has been set to seven (7). Therefore, the upper boundary is set to 7 times greater than the median, whereas the lower boundary is set to 7 times lower than the median.

Next, data are also filtered by the *full load* working regime. Full load is an operation mode such that the machine is working at its limit condition, where degradation becomes apparent. Therefore, data are filtered at full load mode under the assumption that machine degradation can be clearly captured.

Finally, as explained in Section 3.1.2.2, the *grid code regulation* must be considered. Therefore, this event is also filtered by applying a threshold in the Variable Guide Vane (VGV) aperture. The threshold has been defined using the median of the VGV aperture in full load conditions for each equipment.

Following the process of data curation, the number of IGTs available for analysis was reduced from nineteen (19) to twelve (12) due to the limited amount of available data. Consequently, the ensuing analysis and subsequent results will be based on these 12 equipment.

3.3.1 Data Set

Data are collected from sensors at different locations in IGT's compressor for its real operation. Data are curated to ensure adequate quality for developing the ML models, and afterwards, it is divided in three different subsets to train, validate and test these ML models.

3.3.1.1 Training Set

The training set is composed by the first year of IGT data,

$$\mathcal{M} = \{\mathbf{x}(t_k^{(m)})\}_{k=1}^{N_m} = \{\mathbf{x}_k\}_{k=n_m}^{n_m+N_m} = \{\mathbf{m}_k\}_{k=1}^{N_m} \subset \mathcal{D} \subset \mathcal{X}, \tag{3.1}$$

with inputs $\mathbf{x} \in \mathcal{X}$ as training features, and $t_k^{(m)}$ notes (m) to refer to the time where the model is trained, that is the first year of the fresh machine. The aim when selecting this time period for training the model is to capture the behavior of the fresh machine at the beginning of its life. This model will serve as a behavioral baseline to be compared with the performance in the long-term. In case discrepancies are found, it allows one to assess that a drift exists between the model behavior and the real IGT plant. Sampling time T_s to generate the training dataset is 1 min. Therefore, note that $\mathbf{x}_k = \mathbf{x}(t_k)$ and $\mathbf{x}_{k+1} = \mathbf{x}(t_{k+1}) = \mathbf{x}(t_k + T_s)$.

A histogram of both the training data set and the whole data set is computed to assess fairness of data, i.e If the training data is sufficiently representative of the entire dataset. The bounds of the histogram (its maxima and minima values), and its shape determine whether the training set is significant enough. In Figure 3.5, an example is shown. On the left column, the whole system histograms are represented, whereas on the right column, only histograms from the

training data are displayed. Both are thoroughly analyzed and compared before training any model. This analysis determines whether data used in the training and the number of instances are significant enough. If so, the model would be treated as a representative plant for short-term performance.

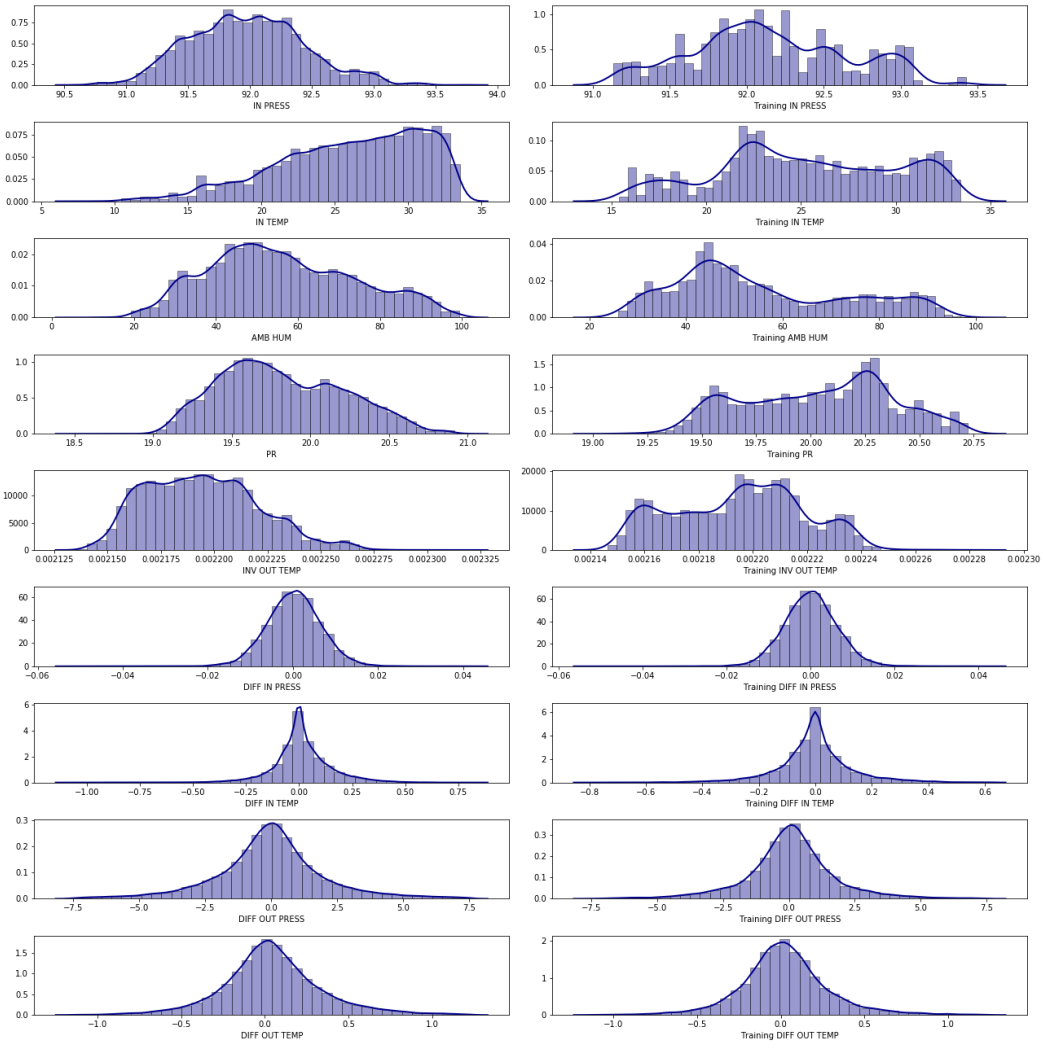


Figure 3.5: Representation of instances in the training data relative to the entire operational dataset

Regarding the distribution of training data, it was determined that two additional machines could not be considered due to inadequate representation. As a result, these systems were deprecated since the available data points were deemed insufficient to effectively train the model.

3.3.1.2 Validation Set

In this study, the validation set is the same as the training set, as expressed in Equation (3.1). The model is constructed under the assumption that the equipment is working under fresh conditions for the first working year. It means that the working mode during this period of time is assumed as the correct one. Therefore, to validate the model, the density plot of each feature is used to ensure an entire instance representation in the training set.

3.3.1.3 Testing Set

Test data, defined as

$$\mathcal{T} = \{\mathbf{x}(t_k^{(t)})\}_{k=1}^{N_t} = \{\mathbf{x}_k\}_{k=n_t}^{n_t+N_t} = \{\mathbf{t}_k\}_{k=1}^{N_t} \subset \mathcal{D} \subset \mathcal{X}, \tag{3.2}$$

with $n_t = 1$ and $n_t + N_t = N$, represent the whole data set, i.e., the entire available data for the assessed IGT is used to test the model of the compressor to its long-term performance. $t_k^{(t)}$ notes (t) to refer to the time where the model is tested, that is the entire available data in our case.

3.3.1.4 Post-Maintenance Set

Maintenance events can be classified as inspections $\mathcal{I}$ and replacements $\mathcal{R}$. $\mathcal{I}$ refers to a periodic revision performed in the compressor to ensure the minimum quality component conditions, whereas $\mathcal{R}$ corresponds to a more complex operation, during which some components are changed and, therefore, the performance of the IGT is altered. Hence, post-maintenace sets are part of the testing set, but

we are taken into account the moment a maintenance, inspection or replacement, has been produced.

3.4 Model Description

It is crucial to have a clear understanding of the basic concepts and techniques used to construct and define the structure of the models in the field of data analysis and modeling. The purpose of this section is to provide a comprehensive overview of these fundamental concepts and methods. Therefore, the key concepts behind the models as well as the way how these are constructed to both analyze complex data sets and obtain the study results are introduced.

3.4.1 Artificial Neural Network

Artificial neural network (ANN) is a black-box model of non-linear, multivariable static and dynamic system which can be treated using input-output information measured from the system. The basic element is a *neuron*, each one performing a weighted sum of its inputs and, the output is then passed through, usually, a non-linear function, the so called activation function. This can be expressed as,

$$f_{W,b}(\mathbf{x}) = \sigma(W^T \cdot \mathbf{x} + b) \tag{3.3}$$

where $f \in \mathbb{R}^n \times \mathbb{R}$ maps the output of the neuron, indeed the activation function $\sigma(\cdot)$, and $\mathbf{x} \in \mathbb{R}^n$ is the input to the neuron. Thereby, the output of a single neuron corresponds to the input-output mapping defined by the activation function. This can be described in many different ways, which the most common ones will be shown in Section 3.4.2.3.

Neurons are grouped by layers according to the type of architecture and the learning algorithm. Then, each neuron is connected by links known as synapses to all the other neurons in the forward layer characterized by weights which store the knowledge of the network acquired through the learning process. A normal structure of ANN is formed by an input layer, one or more hidden layers and an output layer. The configuration of the number of neurons and layers is determined arbitrary based on a specific application and previous experience. An illustration of a single neuron and more complex neural network is shown in Figure 3.6. Each square refers to a neuron, also known as node, of the ANN model. The leftmost layer, which is colored green, is referred to as the input layer. On the other hand, the rightmost layer, which is colored orange, is known as the output layer. The layers between the input and output layers are called hidden layers, as their values are not observable during training. In this particular example, the neural network comprises 7 input nodes, 4 and 3 hidden nodes in the first and second hidden layers respectively, and a single output unit.

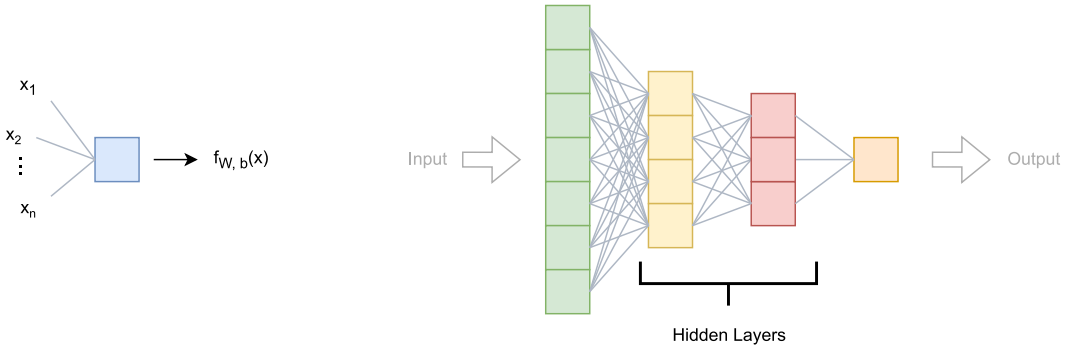


Figure 3.6: (Left) Representation of a single neuron. (Right) more complex Artificial Neural Network

To train this kind of models, a cost function J must be defined, which is then expressed as a function of weights W and bias b. Reformulating to $J(W, b)$, the network coefficients are updated following the iterations of gradient descent. These are expressed as partial

derivatives and, to efficiently compute them, the backpropagation algorithm is computed.

With this intuition, it can be noted that the structure, which includes the number of layers, neurons, and the chosen activation function, and the definition of the cost function determines the complexity of the ANN, which can vary significantly.

3.4.2 Autoencoder Architecture

An autoencoder (AE) architecture based on ANNs is proposed to identify significant hidden patterns to determine operational changes in IGT performance based on compressor data. Autoencoder is a type of unsupervised learning architecture that is widely used across various fields. Its main objective is to reconstruct the original input data by effectively encoding it as input vectors. This is achieved using a two-faced ANN structure consisting of an encoder and a decoder, as illustrated in Figure 3.7. The encoder network $G(\cdot)$, or simply the encoder, is defined as an encoding function,

$$\mathbf{z} = G(\mathbf{x}) \tag{3.4}$$

where $\mathbf{x}$ is the model input and $\mathbf{z}$ is a set of latent variables. The decoder network $F(\cdot)$, i.e., the decoder, is defined to reconstruct the encoded signal,

$$\mathbf{x}' = F(\mathbf{z}) = F(G(\mathbf{x})) \tag{3.5}$$

where $\mathbf{x}'$ refers to the reconstructed input signals.

Therefore, the model to be trained in our research on fresh training data can defined as

$$AE_{\mathcal{M}} = (G_{\mathcal{M}}, H_{\mathcal{M}}), \tag{3.6}$$

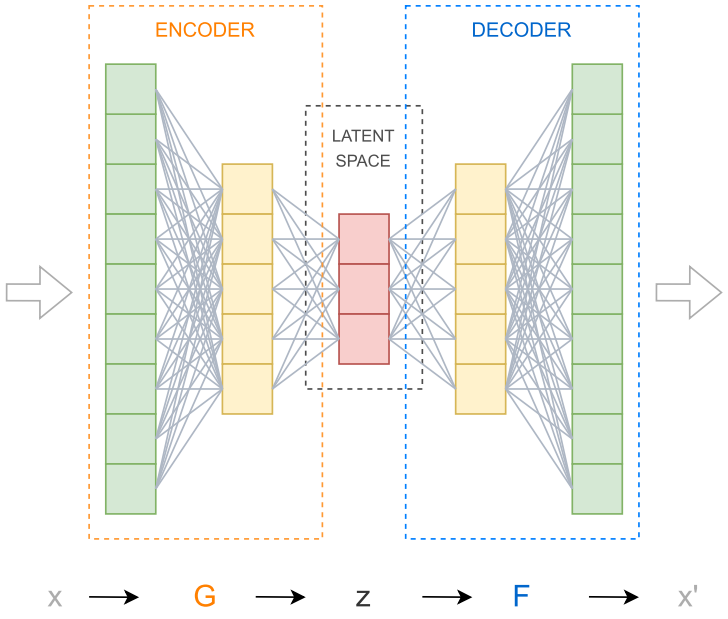


Figure 3.7: General example of an autoencoder with a three-dimensional latent space. Network G is an encoding function $z = G(x)$, where x is the model input, and z are named latent variables. The decoder F is defined to reconstruct the encoded signal, $x' = F(z)$. The set of weights for both networks is simultaneously learned by minimising a loss function $\epsilon = d(x, x')$ according to some distance metric

where $\mathcal{M} \subset \mathcal{X}$ is the training data defined in Equation (3.1).

To build an AE, several hyperparameters must be set, including the number of hidden layers, the number of nodes per layer, the activation function, and the loss function used to measure the match between input and recovered data, indeed the cost function. In addition, an extra parameter that needs to be set is the code size in the latent space.

The set of weights for both ANNs are simultaneously learned by minimising a loss function $\epsilon = d(\mathbf{x}, \mathbf{x}')$ according to some distance metric [109]. The output layer of the encoder (input layer of the decoder) represents the compressed/transformation of the input data, which is of interest because this latent space captures all the relevant information of the input data in a reduced dimensionality. This layer is also referred to as the *code*.

In our research path, the validation set is the same that the training test, hence the loss metric can be defined as,

$$\epsilon_{\mathcal{M}}(\mathbf{x}) = d(\mathbf{x}_{\mathcal{M}}, AE_{\mathcal{M}}(\mathbf{x})) \tag{3.7}$$

As far as the reconstruction distance refers to the AE model that has been trained with modeling data $\mathcal{M} = \{\mathbf{m_k}\}_{k=1}^{N_m}$, it is expected that $\epsilon_{\mathcal{M}}(\mathbf{m_k}) \approx 0$, for $\mathbf{m_k} \in \mathcal{M}$, as data are validated in similar conditions that the model was trained, indeed in short-term performance.

In the other side, it is expected that our AE trained on new first-year data should be able to identify changes in the operational point in data after major maintenance (post-maintenance data) by showing a certain degree of discrepancy between the original data and the reconstructed data. By contrast, no abrupt changes in the operational point are expected after inspections; that is, the discrepancy should remain almost stable after inspections.

In this research study, two different type of AE are used, several structures are also tested, and a further analysis of activation functions are done in order to determine the one that performs the best. Next, a more detailed information about each aspect is presented.

3.4.2.1 Autoencoder Types

AE can be defined in various ways depending on the problem at hand. In this thesis, two approximations based on the structure and cost function are used: a Fully Connected (FC) and a sparse autoencoder. FC models are the simplest ones since all neurons in the ANN are connected amongst layers without any constraints. The

cost function is defined as:

$$J(x) = \min \frac{1}{N} \sum_{k=1}^{N} ||\mathbf{x}_k - \mathbf{x}'_k||^2 \tag{3.8}$$

where $\mathbf{x}_k$ are the input features, and $\mathbf{x}'_k$ are the reconstructed input signals.

The sparse version impose a sparsity constraint by reducing node activation at all network levels, which is controlled by a sparsity parameter, ρ. This parameter is included in the cost function as:

$$J_{sparse}(x) = \min J(x) + \beta \sum_{j=1}^{l} KL(\rho||\rho'_j) \tag{3.9}$$

where $KL(\rho||\rho'_j) = \rho \log \frac{\rho}{\rho'_j} + (1-\rho) \log \frac{1-\rho}{1-\rho'_j}$ is the Kullback-Leibler (KL) divergence between a Bernoulli random variable with mean ρ and ρ'_j is the average activation of hidden units j, and β is the weight of the sparse penalty term. This penalty function has the property that $KL(\rho||\rho') = 0$ if $\rho = \rho'$. Therefore, by defining an small value for ρ, the optimization constrict that the average activation of hidden units j must be close to it. Thereby, high value in activation units can be avoid.

It becomes particularly intriguing to apply a sparse constraint to avoid a mere replication of the existing inputs when expanding the number of features beyond the input layer. Such a constraint promotes the creation of new and informative features while suppressing the irrelevant ones, leading to a more concise and effective representation of the data.

3.4.2.2 Autoencoder Structures

The AE structure is defined as the number of layers in the model and the number of neurons per layer. Two main different structures are proposed based on the features used. The number of layers is fixed to one hidden layer in each network, indeed the encoder and decoder. Therefore, the structure variation is based on the number of neurons per layer. This design choice strikes a balance between added computational complexity and interpretability of the results.

On the one hand, the first model structure is developed by considering all features explained in Section 3.2. Thereby, the AE structure is set to $9-6-4-6-9$, as shown in Figure 3.8. This structure presents a compressed model having a middle step of compression, as well. From now on, it will be called *Differential Model ($\mathcal{DM}$)*.

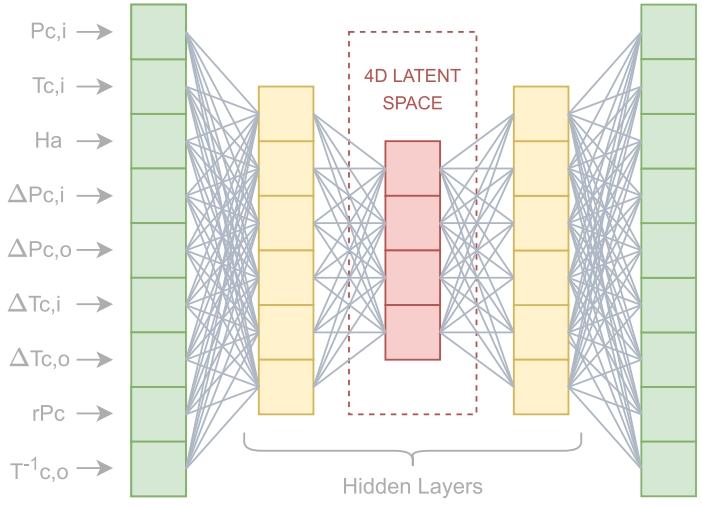
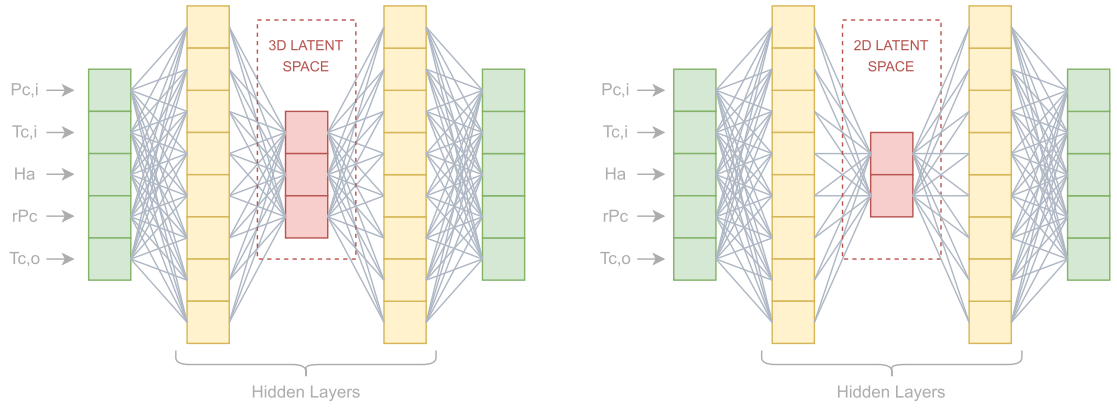


Figure 3.8: Representation of the Differential Model $\mathcal{DM}$ employed to capture gas turbine's compressor performance

On the other hand, the features used are the raw values of the most important compressor features, which include inlet pressure and temperature, pressure ratio, outlet temperature, and ambient humidity. Moreover, the structure presents an expansion after the input layer, indeed the hidden layer is set to eight (8). Finally, the code size

is evaluated in two different values to perform a model analysis. Therefore, the two proposed models for this structure are set to $5-8-3-8-5$ and $5-8-2-8-5$, respectively, as depicted in Figure 3.9a and Figure 3.9b. From now one, these will be called *Expansion Model ($\mathcal{EM}$).*



a) 5-8-3-8-5 Autoencoder structure representation

b) 5-8-2-8-5 Autoencoder structure representation

Figure 3.9: Representation of the Expansion Model $\mathcal{EM}$ employed to capture gas turbine's compressor performance

The second approach, as well as some dimensionality reduction methods explained later, were intentionally developed because our industrial data owner partner was interested in further research about explainability using the information from the latent space. The criterion that determined the latent space dimension was defined by the lowest possible values with maximum model performance based on experimental results.

The models presented in this study have been extensively utilized; however, additional significant analyses were conducted to accurately determine the proposed structure. The complementary models employed to identify the most effective models are stated in each specific section

3.4.2.3 Activation Function Analysis

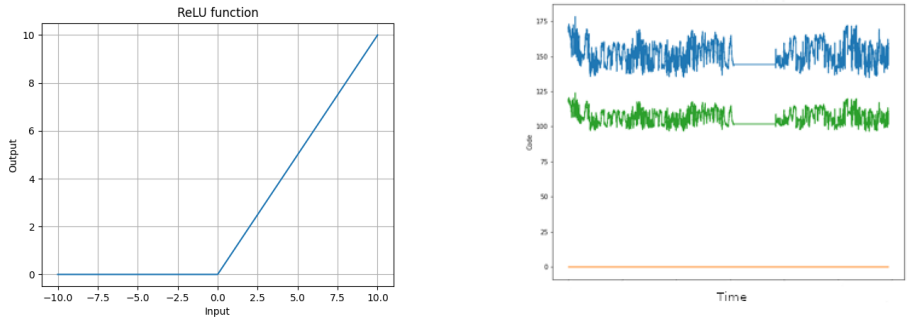
Activation functions are a set of mathematical operations used to compute the output of each layer in an ANN based model, in particular when an autoencoder architecture is composed. The choice of the activation function have a significant impact on the outcome of the code section, that is the latent space. This area is of particular interest and will be subjected to further analysis and investigation in the section devoted to presenting the results. Therefore, selecting the appropriate activation function can lead to significant improvements in the performance of the ANN. Hence, a series of tests are conducted to evaluate the effectiveness of different activation functions as well as the representation of the code in each case. To assess the activation functions the Expansion Model for a Fully Connected autoencoder with three nodes in the latent space, $\mathcal{EM}^{FC3D}$ is used. The presented results uses lineal time, meaning that the each point is treated as a time unit. In Section 4.7, time samples are placed in the timeline, thus the visualization is slightly different. Moreover, units used to perform the following analysis and the one presented in the results section are different, this is why it can present differences in the y-axis values.

The first activation function analyzed is the Rectified Linear Unit (ReLU). It is defined as

$$\mathrm{ReLU}(x) = \max(0, x) \tag{3.10}$$

where x is the real output of each node in the layer. Figure 3.10 shows its representation on the left. The code behaviour $(z_1, z_2, z_3)^\intercal$ is depicted on the right in the same figure, when using ReLU as

activation function in the encoder and decoder elements of the autoencoder.



(a) ReLU function representation

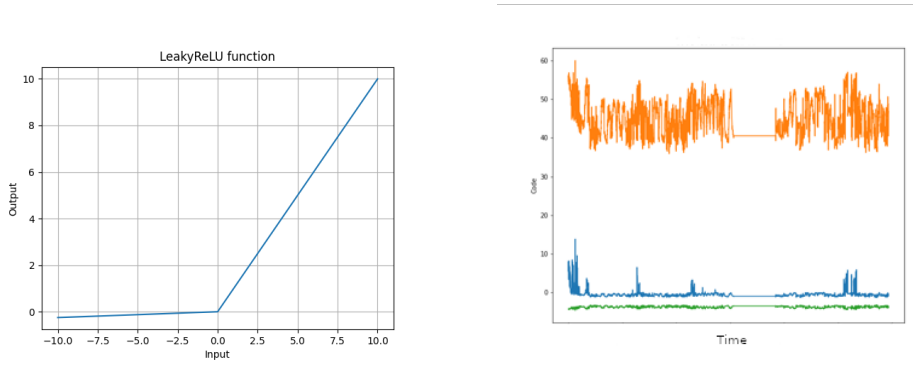
(b) Code time-series representation

Figure 3.10: Auto-encoder code performance using ReLU as activation function

The second activation function is the Leaky Rectified Linear Unit (LeakyReLU). It is defined as

$$\text{LeakyReLU}(x) = \begin{cases} x & \text{if } x \geq 0 \\ \alpha \cdot x & \text{otherwise} \end{cases} \tag{3.11}$$

where x is the output of each node in the layer. Figure 3.11 shows its representation and the code performance using it as activation function.



(a) Leaky ReLU function representation

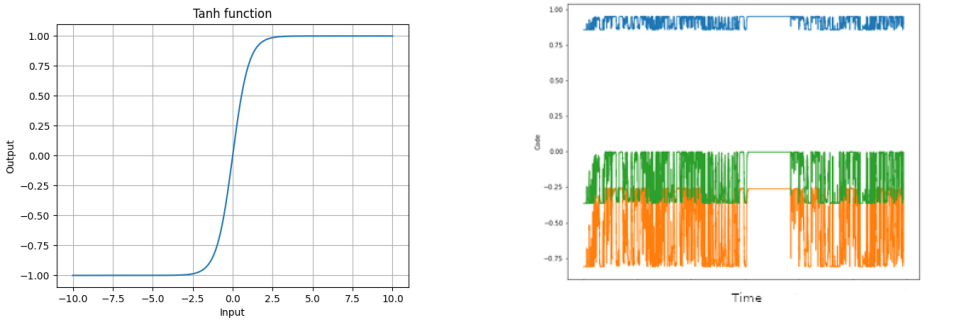
(b) Code time-series representation

Figure 3.11: Auto-encoder code performance using Leaky ReLU as activation function

The third activation function is the Hyperbolic Tangent (tanh). It is defined as

$$\tanh(x) = \frac{\exp(x) - \exp(-x))}{\exp(x) + \exp(-x)} \tag{3.12}$$

where x is the output of each node in the layer. Figure 3.12 shows its representation and the code performance using it as activation function.



(a) tanh function representation

(b) Code time-series representation

Figure 3.12: Auto-encoder code performance using Tanh as activation function

The fourth function under inspection is the Exponential Linear Unit (ELU). It is defined as

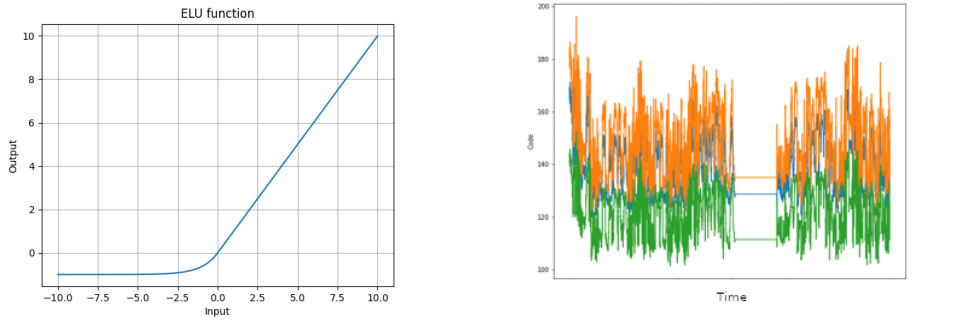
$$\mathrm{ELU}(x) = \begin{cases} x & \text{if } x \geq 0 \\ \alpha\left(\exp(x) - 1\right) & \text{if } x \leq 0 \end{cases} \tag{3.13}$$

where x is the output of each node in the layer. Figure 3.13 shows its representation and the code performance using it as activation function.

Finally, the sigmoid activation function is analyzed. It is defined as

$$\mathrm{sigmoid}(x) = \frac{1}{1 + \exp(-x)} \tag{3.14}$$

where x is the output of each node in the layer. Figure 3.14 shows its representation and the code performance using it as activation

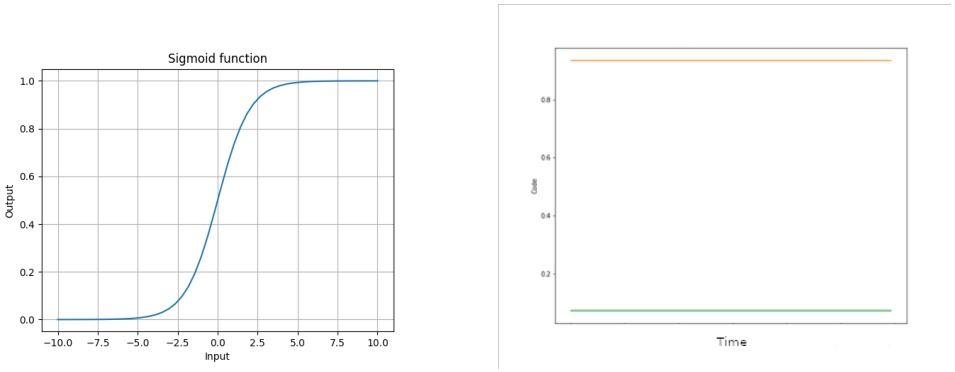


(a) ELU function representation

(b) Code time-series representation

Figure 3.13: Auto-encoder code performance using ELU as activation function

function.



(a) sigmoid function representation

(b) Code time-series representation

Figure 3.14: Auto-encoder code performance using Sigmoid as activation function

Even though the evaluation of the code value to each activation function are quite different, the patterns obtained in the training of the models are quite similar, except for the sigmoid function which presents flat lines. So, rejecting this last activation function, any of the other ones can be considered. The ReLU activation function has been selected, wherein one of its dimensions is constrained to zero. This noteworthy attribute highlights that a variable may be deemed superfluous, however, the performance of the model is adversely affected if a single dimension in the latent space is discarded. Hence, it can be inferred that although the variable may present a zero value

result, it remains essential for the interconnections within the neural network.

3.5 Model Analysis

Several strategies were employed to gain a deeper understanding of the nature and patterns arising from the model's performance. The first strategy, referred to as Model Distance, involved comparing the model output to the actual plant performance. The second strategy involved a temporal analysis of the time-series data. Then, a topological analysis of the deviation was conducted. Finally, some clustering methods were also applied together on the post-maintenance set. Next, some valuable insights are presented about the methods and metrics utilized.

3.5.1 Model Distance

Model distance aims to measure the deviation between the model and the real machine performance. This calculation is computed using two measurements and two time-window samples processing. Indeed, the measurements are the Absolute Difference (AD), the Mean Squared Error (MSE), and Fréchet Distance (FD), while the time-window samples are the Moving Average (MA) and the Incremental Window Average (IWA).

- AD is the absolute difference between the real plant values $\mathbf{x}$ and the inferred values of the AE model $\mathbf{x'}$ at each time step. Therefore,

$$AD_i = ||\mathbf{x}_i - \mathbf{x'}_i||_1 \tag{3.15}$$

- *MSE* is an estimator measure of the average of square difference of the magnitudes of vectors in a space. It is expressed as:

$$MSE = \frac{1}{N}\sum_{i=1}^{N}||\mathbf{x}_i - \mathbf{x'}_i||_2^2 \tag{3.16}$$

- The Fréchet distance *FD* is defined as a measurement of similarity between two curves. Moreover, it can be used to measure the similarity of two distributions. In this case, the comparison is performed between two multivariate distributions —the real plant distribution and the inferred one— and is composed by the same number of features each and a specific time window. Assuming them as two multivariate Gaussian distributions, the distance is calculated as

$$FD = |\mu_X - \mu_Y|^2 + tr(\Sigma_X + \Sigma_Y - 2(\Sigma_X\Sigma_Y)^{1/2}) \tag{3.17}$$

  where μ_X and μ_Y are the means of both distributions, and Σ_X and Σ_Y are their covariance matrices, respectively.

- Moving average, *MA* is the calculation of the average of consecutive subsets of data obtained by sliding a fixed time window along the whole data set. It is defined as

$$MA = \frac{1}{n}\sum_{i=0}^{n-1}\mathbf{x}_{m-i} \tag{3.18}$$

  where m is the current time step, and n is the moving time window size.

- Incremental window average, *IWA* is a calculation of the average of cumulative subsets of data obtained by adding a fixed time

window of the whole data set at each step. It is defined as

$$IWA_n = \frac{1}{n}\left(\sum_{i=0}^{t} \mathbf{x}_t + (n-t)IWA_{m-1}\right) \tag{3.19}$$

where m is the current time step of the grouping subset, n is the length of the cumulative subsets, $\mathbf{x}$ are the instances of the current subset, and t is the length of the current subset.

The results obtained from the absolute difference, AD, and from the Mean squared error, MSE, were very similar. Thereby, regarding this two metrics, only the AD is used in the combination of the measurements with the time window samples. Therefore, this lead to four indicators as it is illustrated in Table 3.1.

Table 3.1: The combinations of the two considered distances and the two smoothing time window calculations

	Moving Average	**Incremental Window Average**
Absolute Error	ADMA	FDMA
Fréchet distance	ADIWA	FDIWA

3.5.2 Temporal Analysis

The data used to perform the analysis is temporal data, meaning that there is a representation of the state of the system at each timestamp. In order to investigate the variability of the parameters, two methods are proposed: the decomposition of time series and Wavelet Transform (WT).

3.5.2.1 Decomposition of Time Series

The decomposition of time series is a method used to analyze a time series by separating it into its underlying components. These

components typically include trend, seasonality, and residuals [110].

Trend T_t refers to the gradual deviation of a time series to higher or lower values over time. The time series is considered stationary if the trend does not exist.

Cyclic C_t is defined as a repeated pattern, but non-periodic fluctuations. The period of these fluctuations depends on the nature of the time series.

Seasonality S_t corresponds to a repeating pattern over a successive period of time. The pattern must be stable over time in magnitude and direction.

Residuals R_t is the unexplained or random variation of the time series. It refers to the remainder of the time series after the other components have been removed.

Once the components are identified, they can be analyzed separately to gain a better understanding of the underlying patterns and trends in the data. For example, the trend component can be used to identify whether the time series is increasing, decreasing, or staying constant over time. The seasonality component can be used to identify whether there are any regular patterns or cycles in the data, such as monthly or yearly fluctuations.

The decomposition can be perform using two different models, additive and multiplicative models. The additive model can be expressed as:

$$y_t = T_t + C_t + S_t + R_t \tag{3.20}$$

whereas the multiplicative models is expressed as:

$$y_t = T_t \cdot C_t \cdot S_t \cdot R_t \tag{3.21}$$

The additive model would be used when the variations on the trend do not vary with the level of the time series, while a multiplicative models is appropriate whether the trend is proportional to the level of the time series [111].

3.5.2.2 Wavelet Transform

Wavelet transform, WT is a mathematical technique used to analyze temporal data by decomposing them into different frequency components. It is based on a family of functions called wavelets that are well-localized in both time and frequency domains. A wavelet is a mathematical function that can be scaled and translated to analyze different frequency components of a signal. The wavelet function has a finite duration, which makes it suitable for analyzing signals with localized features.

The WT works by breaking down a signal into a set of wavelet functions at different scales and positions. These wavelets are generated by dilating and translating a mother wavelet function. The dilation and translation parameters are referred to as the scale and position, respectively. By adjusting the scale and position of the wavelet function, different frequency components of the signal can be analyzed [112]. The wavelet function can be represented as,

$$\psi_{a,b} = \frac{1}{\sqrt{a}}\psi\left(\frac{t-b}{a}\right) \tag{3.22}$$

where a determines the scale parameter, and b refers to the location of the wavelet.

Several representations of WT exist. In this study, the Discrete Wavelet Transform (DWT) is used [113]. This one is chosen since

this type of wavelet is specially adaptable for sampled values. The DWT employs a dyadic grid where the wavelet is scaled by power of two, indeed $a = 2^j$ and translated by an integer, indeed $b = k \cdot 2^j$, where j runs from 0 to J being J the total number of scales, and k is the location index running from 1 to $2^j \cdot N$ being N the total number of observations. Therefore, the DWT is expressed by the following equation:

$$\psi_{j,k} = 2^{-j/2}\psi(2^{-j}t - k) \tag{3.23}$$

DWT components are obtained from the following expression:

$$WC_{j,k} = WC(2^j, k2^j) = 2^{-j/2} \int_{-\infty}^{\infty} f(t)\overline{\psi(2^{-j}t - k)}\, dt \tag{3.24}$$

where $f(t)$ is the time-frequency representation of the signal.

3.5.3 Topological Analysis

The motivation of this analysis is to study the shape of data. This approach uses techniques from topology, which include analyzing the data in a way that is insensitive to the particular metric chosen, and provides dimensionality reduction and robustness to noise [114]. The analysis is performed using two methods: t-distributed Stochastic Neighbor Embedding (t-SNE) and Uniform Manifold Approximation and Projection (UMAP).

3.5.3.1 T-distributed Stochastic Neighbor Embedding

This method consists of visualizing the resulting similarity data from mapping the high-dimensional state-vectors onto a pairwise similarities matrix. t-SNE is capable of capturing much of the local structure of the high-dimensional data such as the presence of clusters at

several scales [115].

t-SNE comprises two main stages. On the one hand, the method computes a probability distribution over pairs of high-dimensional objects in a manner that similar points are assigned with higher probability, and vice versa. The probability $p_{ij} \in \mathcal{P}$ that are proportional to the similarity of two objects $\mathbf{x}_i$ and $\mathbf{x}_j$ is computed as,

$$p_{j|i} = \frac{exp(-||\mathbf{x}_i - \mathbf{x}_j||^2/2\,\sigma_i^2)}{\sum_{k \neq i} exp(-||\mathbf{x}_i - \mathbf{x}_k||^2/2\,\sigma_i^2)} \tag{3.25}$$

where $\mathbf{x}_i$ and $\mathbf{x}_i \neq \mathbf{x}_j$.

The remaining parameter to select is the variance σ_i of the Gaussian distribution that is centered over each high-dimensional point $\mathbf{x}_i$. To determine this parameter, a value search of σ_i is performed that produces a probability distribution P_i over all other data points. This distribution is produced with a fixed perplexity that is specified by the user. The perplexity can be interpreted as a measure of effective number of neighbor. It is expressed as:

$$Perp(P_i) = 2^{H(P_i)} \tag{3.26}$$

where $H(P_i)$ is the Shannon entropy. It is defined as:

$$H(P_i) = -\sum_j p_{j|i}\, log_2 p_{j|i} \tag{3.27}$$

Typical values of perplexity are comprises between 5 and 50.

On the other hand, t-SNE aims to reflect the similarities p_{ij} over the points in the low-dimensional map $\mathbf{y}_1, \ldots, \mathbf{y}_N$, with $\mathbf{y}_i \in \mathbb{R}^d$, and d usually chose as 2 or 3, strictly. The similarities between two points

in low-dimensional map uses the T-Student distribution with a single degree of freedom to allow dissimilar object to be modelled far apart in the map. Therefore, the measure similarities $q_{ij} \in \mathcal{Q}$ is expressed as

$$q_{ij} = \frac{(1 + ||\mathbf{y}_i - \mathbf{y}_j||^2)^{-1}}{\sum_k \sum_{l \neq k} (1 + ||\mathbf{y}_k - \mathbf{y}_l||^2)^{-1})} \tag{3.28}$$

The locations of the points $\mathbf{y}_i$ are determined by minimizing the distance of the two distributions, indeed $\mathcal{P}$ and $\mathcal{Q}$, the KL divergence, that is:

$$KL\left(\mathcal{P}||\mathcal{Q}\right) = \sum_{i \neq j} p_{ij} \log \frac{p_{ij}}{q_{ij}} \tag{3.29}$$

The Gradient Descent (GD) method is used to minize the KL divergence according to the points $\mathbf{y}_i$.

3.5.3.2 Uniform Manifold Approximation and Projection

UMAP is a non-linear dimensionality reduction technique that uses local manifold approximations and patches together their fuzzy simplicial set representations to construct a topological representation of the high dimensional data. A similar process can be used to identify an equivalent topological representation given low dimensional representation of the data. This methods was presented in [116].

Similarly as t-SNE, it mainly consists of two stages. Firstly, a graph is built in a high dimensional data representation where the exponential probability using binary search and the number of nearest neighbors to consider are computed. And secondly, an optimization is perform to build a low-dimensional representation.

Even though the processes concept is quite similar, the way how they are computed is very different.

- UMAP does not necessary use Eucliden distance to compute the exponential probability distribution in high dimensional representation. Otherwise it uses:

$$p_{j|i} = exp\left(\frac{-d(x_i, x_j) - \rho_i}{\sigma_i}\right) \tag{3.30}$$

where $d(x_i, x_j)$ is a user defined distance metric, ρ_i represents the distance from each i-th data point to its first neigbor, and σ_i corresponds to a normalization factor to the point x_i

- The k-nearest neighbors algorithm is used instead of perplexity. This is defined as:

$$k = 2^{\sum_j p_{ij}} \tag{3.31}$$

- Instead of using the t-Student distribution for modelling distance probabilities in low dimensional manifold, UMAP uses:

$$q_{ij} = (1 + a(y_i - y_j)^{2b})^{-1} \tag{3.32}$$

where a and b are determined from non-linear least squares fitting against the curve $\Psi : \mathbb{R}^d x \mathbb{R}^d \rightarrow [0, 1]$ where

$$\Psi(x, y) = \begin{cases} 1 & \text{if } ||y_i - y_j||_2 \leq min\_dist \\ exp(-(||y_i - y_j||_2 - min\_dist & \text{otherwise} \end{cases} \tag{3.33}$$

where min_dist is the desired separation between close points hyperparameter.

- UMAP uses *binary-Cross Entropy (CE)* as a cost function instead of KL-divergence.

$$\begin{aligned} CE(X,Y) \ &= \sum_i \sum_j \left[p_{ij}(X) log\left(\frac{p_{ij}(X)}{q_{ij}(Y)}\right) \right. \\ &\left. +(1-p_{ij}(X)) log\left(\frac{1-p_{ij}(X)}{1-q_{ij}(Y)}\right) \right] \end{aligned} \tag{3.34}$$

- UMAP uses the Stochastic Gradient Descent (SGD), which speeds up the computations and consumes less memory, instead of *Gradient Descent (GD)*.

These kind of methods are constructed based on many significant definitions and assumptions that are beyond this thesis. However, a short explanation is given in this section to have an intuition how these models work.

3.5.4 Clustering Methods

The goal of clustering is to identify significant patterns and relationships that exist within a given dataset, which can help to shed light on the underlying processes, descriptive characteristics, and groupings that are present. Depending on the nature of the data and the desired outcomes, there are various techniques and algorithms that can be used to categorize the data into meaningful groups. In this study mainly two well-known clustering alogirhtms are used, which are k-Means and Density Based Spatial Clustering of Applications with Noise (DBSCAN).

3.5.4.1 K-Means

It is a method of vector quantization that aims to divide a given set of points into k clusters, so that the within-cluster sum of squares is minimized [117]. The number of cluster are set at the beginning as a parameter. Formally, given $(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n)$ observations of dimension d, it is defined by

$$\underset{S}{arg\,min} \sum_{i=0}^{k} \sum_{\mathbf{x}_i \in S_i} ||\mathbf{x}_i - \mu_i||^2 \tag{3.35}$$

where μ_i is the mean of points in S_i.

3.5.4.2 Density Based Spatial Clustering of Applications with Noise

DBSCAN aims to discover clusters in a database of arbitrary shape based on its density [118]. Given certain number of points $\mathcal{B}$ in some space, the closest are packed together making as noise points the ones that lie apart in small-density region. Two parameters are used to compute the cluster, ε and the minimum number of points required to shape the cluster, m. ε is a common parameter in density based clustering methods that derive from the ε-neighbourhood of a point (N_ε). Given a database $\mathcal{D}$, N_ε is defined as,

$$N_\varepsilon(p_i) = \{q \in \mathcal{D} | dist(q, p) \leq \varepsilon\} \tag{3.36}$$

Chapter 4

Results

The main goal in this research is to develop support tools to capture the condition of an industrial gas turbine by using only a specific component, indeed the compressor data. Several data set were create to properly assess the equipment behavior from sensor to operational data that are presented in the last chapter. The main strategy followed, as explained in Section 1.10 is to create a model that is able to capture the optimal behavior of a IGT and perform an extensive analysis on understanding the nature of its operation as well as seeking for hidden patterns using the presented methods. In the following, the most relevant results obtained during the research are presented. These are organized as follows. Firstly, a brief study on the thermodynamics behavior is presented in order to gain further insights of IGT operation. Next, from the starting theoretical study, validation results supporting our main assumptions are obtained. Then, results using the differential model $\mathcal{DM}$ for the autoencoder with its corresponding analysis are presented. Finally, results associated to the expansion model $\mathcal{EM}$ are also shown.

4.1 Thermodynamic based Analysis

The first research line in our study corresponds to search for relevant patterns in the system using the thermodynamic theory. Firstly, an study is conducted looking for visual patterns when comparing the power generated and the consumed fuel, as well as considering an isentropic exploration. These analysis uses SE data set and the results are presented next.

4.1.1 Power Generation Exploration

The first research approach is a general IGT's analysis to look for patterns relating the amount of fuel usage versus the amount of power generated, also known as active load. When equipment degrades, the generated power should decrease along time for the same amount of consumed fuel. Ambient conditions could also affect this result because air flow is altered according to air temperature and pressure, so directly impacting power generation. Therefore, these parameters are considered in the analysis. This is the only study in our research where variables that do not belongs exclusively to the compressor, amount of power and fuel, are used.

Figure 4.1(up) shows the output power per fuel mass flow measured in kg/s. A clear decreasing tendency would be expected along time, however no significant pattern is shown. As mentioned before, also the ambient conditions could have an important impact in the power generation. Figure 4.1(down) shows the influence of ambient conditions in power generation. It seems that with lower inlet temperature, higher is the amount of power. However, similarly to the previous case, no clear pattern is defined.

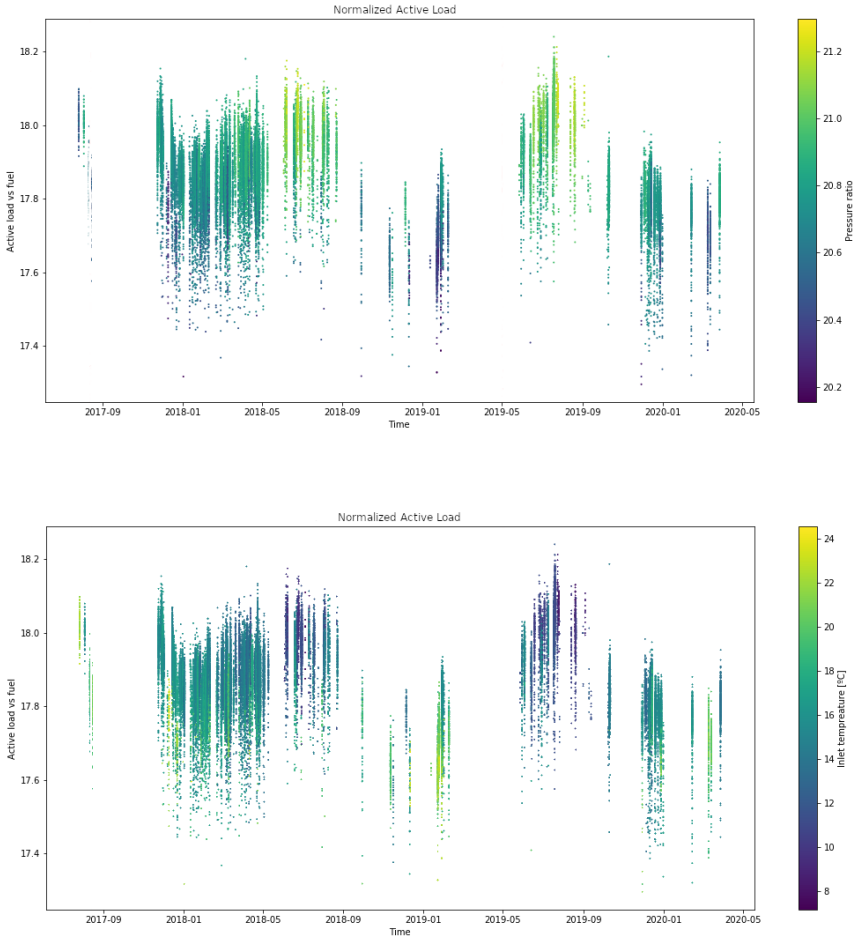


Figure 4.1: (Up) Time series for the active load variable normalized by fuel, coloured according to pressure ratio rP_c. (Down) Time series of the active load variable normalized by fuel, coloured according to inlet temperature $T_{c,i}$

Different deterioration causes for IGTs were introduced in Table 1.1, as well as its prevention methods, whether it is recoverable. The only recoverable deterioration without a major maintenance is fouling, recovered by cleaning the equipment. Thus, IGT performance could be affected whether any compressor washing maintenance is completed for the period of time shown in both figures. This information could provide some insights about how was the performance before the maintenance, in between and after washing. However, washing information does not give either any clue about any deviation pattern.

4.1.2 Isentropic Exploration

A second attempt to discover patterns from a theoretical perspective was tried. This second approach is focused on an important measure of the compressor, indeed the isentropic efficiency. It is a measurement that provides the degree of degradation of energy in steady flow. It is computed according to the pressure ratio influenced by the temperature. Therefore, a general behavior of the compressor can be captured by controlling these parameters. Due to the deterioration factors, the efficiency would decrease along time. Consequently, a decreasing trend in the pressure ratio normalized by the temperature factor would also be expected.

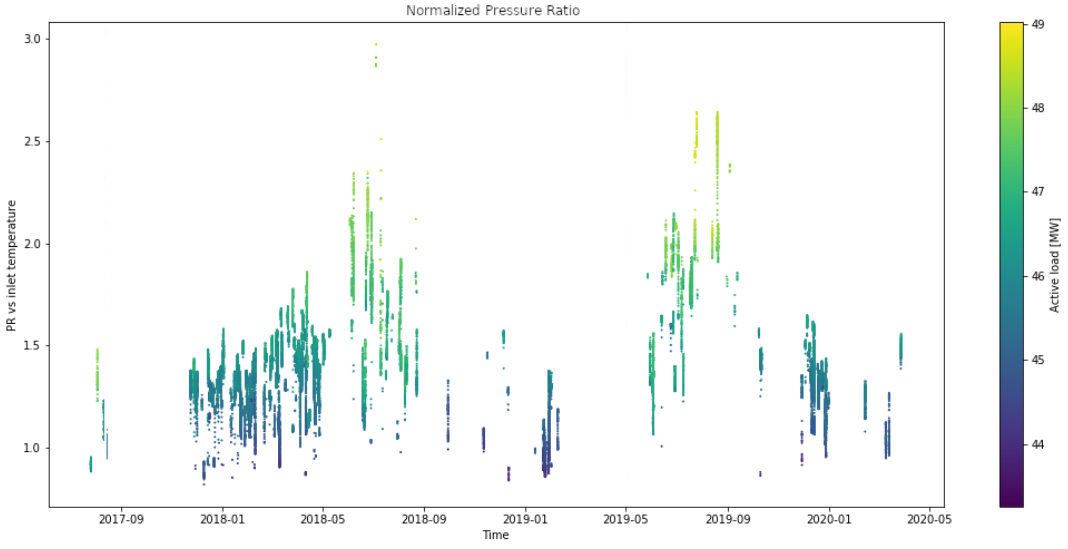


Figure 4.2: Time Series of the pressure ratio rP_c normalized by inlet temperature $T_{c,i}$, coloured according to active load

Figure 4.2 shows the pressure ratio evolution normalized by the inlet temperature with the corresponding washing maintenance actions. As in the previous case, no clear interpretation can be drawn. The plot verifies the relationship between either the pressure ratio or the

inlet temperature with the active load. Higher is the pressure ratio, higher is the amount of power as well as lower are temperature values.

These two physical based approaches have given some insights about the behavior of the machine and make us to understand better the problem in hands. Thus, they help to focus more precisely when moving to data-driven approaches.

4.2 Fresh Reconstruction Discrepancy

As mentioned in Section 1.10, the main data-driven research strategy is to generate a model using fresh operational data, indeed the first year of data, as a reference to compare the IGT short-term performance with long-term. The goal of this model is to capture the optimal operation of IGT and assess if there is any drift along its life. Drift is defined as a continuous slow movement from one working condition to another, mainly due to machine life deterioration, unexpected events, workload or maintenance interventions, like washing, to alleviate the fouling deterioration [119].

An indicator is defined, named *Fresh Reconstruction discrepancy* ($\mathcal{RD}$), based on the deviation between the model output and the real plant values. It is defined as,

$$\mathcal{RD}(\mathbf{x}) = d(\mathbf{x}_{\mathcal{F}}, AE_{\mathcal{M}}(\mathbf{x})) \tag{4.1}$$

It refers to the reconstruction distance when the autoencoder is trained with modeling data $\mathcal{M} = \{\mathbf{m}_k\}_{k=1}^{N_m}$, $AE_{\mathcal{M}} = (G_{\mathcal{M}}, H_{\mathcal{M}})$. In this situation, it is expected that a drift $\epsilon_{\mathcal{M}}(\mathbf{x}_k)$ appear for $\mathbf{x}_k$ not in the training set, as data are reconstructed in long-term performance conditions.

This indicator is evaluated using the AE models mentioned before, trained using features sets presented in Section 3.2, and further analysed using the techniques explained in Section 3.5. In the following sections, a thoroughly explanation is given.

4.3 Nasa Turbofan Jet Engine Data Set Experimentation

Firstly, a set of experiments are performed using the NASA Turbofan Jet Engine Data Set, where the data is already curated and well-prepared. Assuming that the Turbofan Jet Engine was working at fresh conditions during its first period of operation, a simple AE model is developed to capture its behavior. The proposed model was set to 6 compressor's input features, 4 units in single hidden layer for encoder and decoder, and a code size of 2, leading to a $6-4-2-4-6$ structure. The AE representation is shown in Figure 4.3

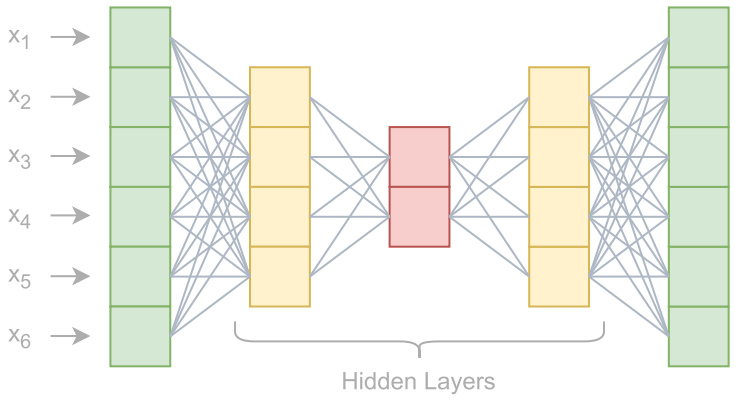


Figure 4.3: Representation of the autoencoder used to validate the drift detection in Nasa Turbojet Data Set

In Figure 4.4, the presented results show positive validation. The initial period of operation is accurately captured with a small error performance, while long-term performance exhibits a clear drift when the GT operates outside normal conditions. These favorable initial results led to further experimentation with data sets provided by Siemens Energy.

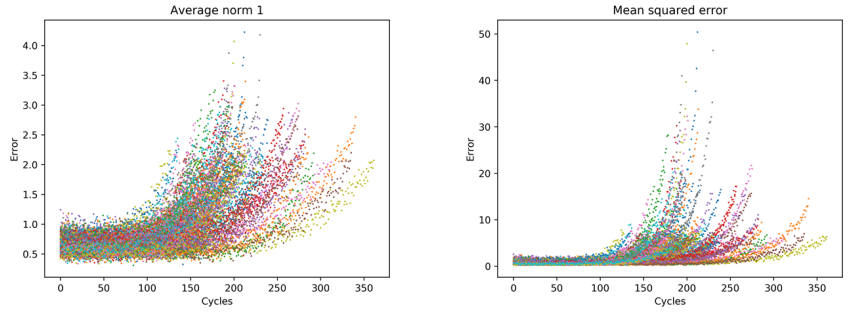


Figure 4.4: Deviation of multiple Turbofan Jet Engine operation using AE architecture based model

4.4 Fresh Reconstruction Discrepancy with a Differential Model Autoencoder

Several autoencoder models were used to proves the drift in fresh reconstruction discrepancy, $\mathcal{RD}$. In this section, the $\mathcal{RD}_{\mathcal{DM}}$ for the Differential Model is presented, the Expansion Models being presented in the next section. The first autoencoder model that was created was the differential model, $\mathcal{DM}$. For assessing it, the combination of two mentioned distances in Section 3.5.1, indeed absolute distance, AD and Fréchet distance, FD, and the two smoothing time window calculations, indeed moving average, MA and incremental window average, IWA, are computed for the deviation between the output for the autoencoder and the real plant values. The results are grouped according to the employed distances to make the results more understandable and easier to be compared. Therefore, AD on MA sampling, ADMA and AD on IWA sampling, ADIWA are first presented. Afterwards, using the Fréchet distance, the FDMA and FDIWA are introduced.

Moving average is computed over several time windows, namely daily, weekly, monthly, and yearly. These multiple representations of

the error reveal various events and trends, with the shape of the discrepancy appearing smoother when evaluated over longer time windows, such as the yearly window. In contrast, shorter time frames, such as the daily and weekly windows, display significant changes in the data. This enables the analysis to be approached from multiple perspectives, providing a more comprehensive understanding of the data.

In Figure 4.5, ADMA values for a single IGT's compressor are presented, with multiple time windows displayed. The plot highlights that the larger the time window, the smoother the representation of the data. Additionally, the same IGT component is displayed using the IWA calculation, resulting in a smoother function for all time windows compared to the ADMA plots. Figure 4.6 showcases the FD represented using both averaging calculations, i.e., MA and IWA.

Both figures contain five subfigures using the time-window moving average, which are ordered from less to more granularity, plus a last subfigure with the IWA time-window. Therefore, in the top, raw data are displayed (1 minute sampled data); next, several time windows linked to the moving average are considered until the largest time window (yearly); finally, a last sixth subfigure plots several IWA results together.

Based on Figure 4.5, a clear ascending trend (drift) can be observed in the reconstruction error along time. There are two main points of interest due to a shift in the reconstruction error: the end of 2015 and between 2017 and 2018. The incremental slope in these two periods is clearly visible in the yearly time window and the ADIWA representations. However, these increments are also observable in the

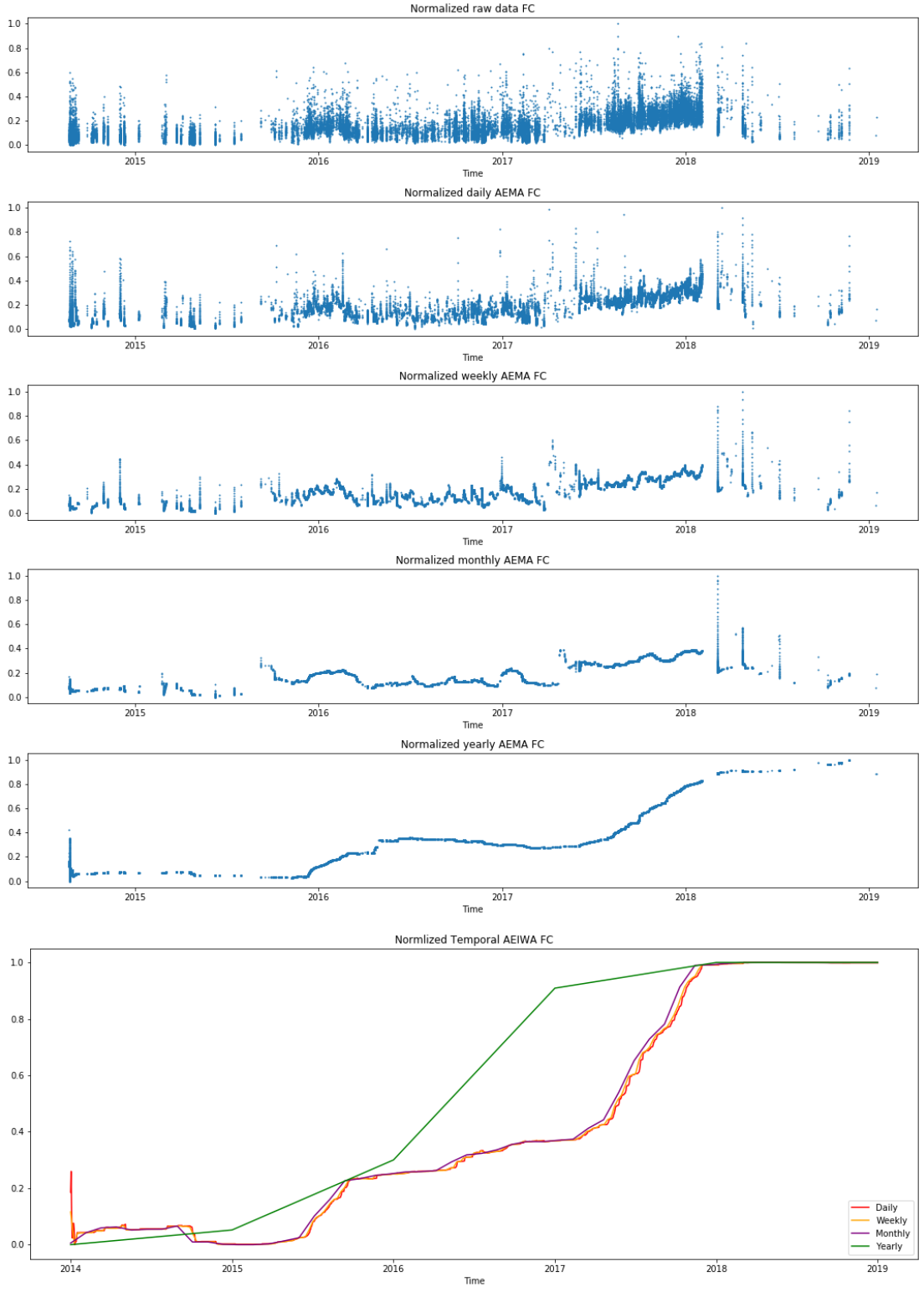


Figure 4.5: Absolute difference, AD metrics of a single IGT using a $\mathcal{DM}$ autoencoder model with multiple time window representations of segmented moving average on the top and grouped incremental window on the bottom

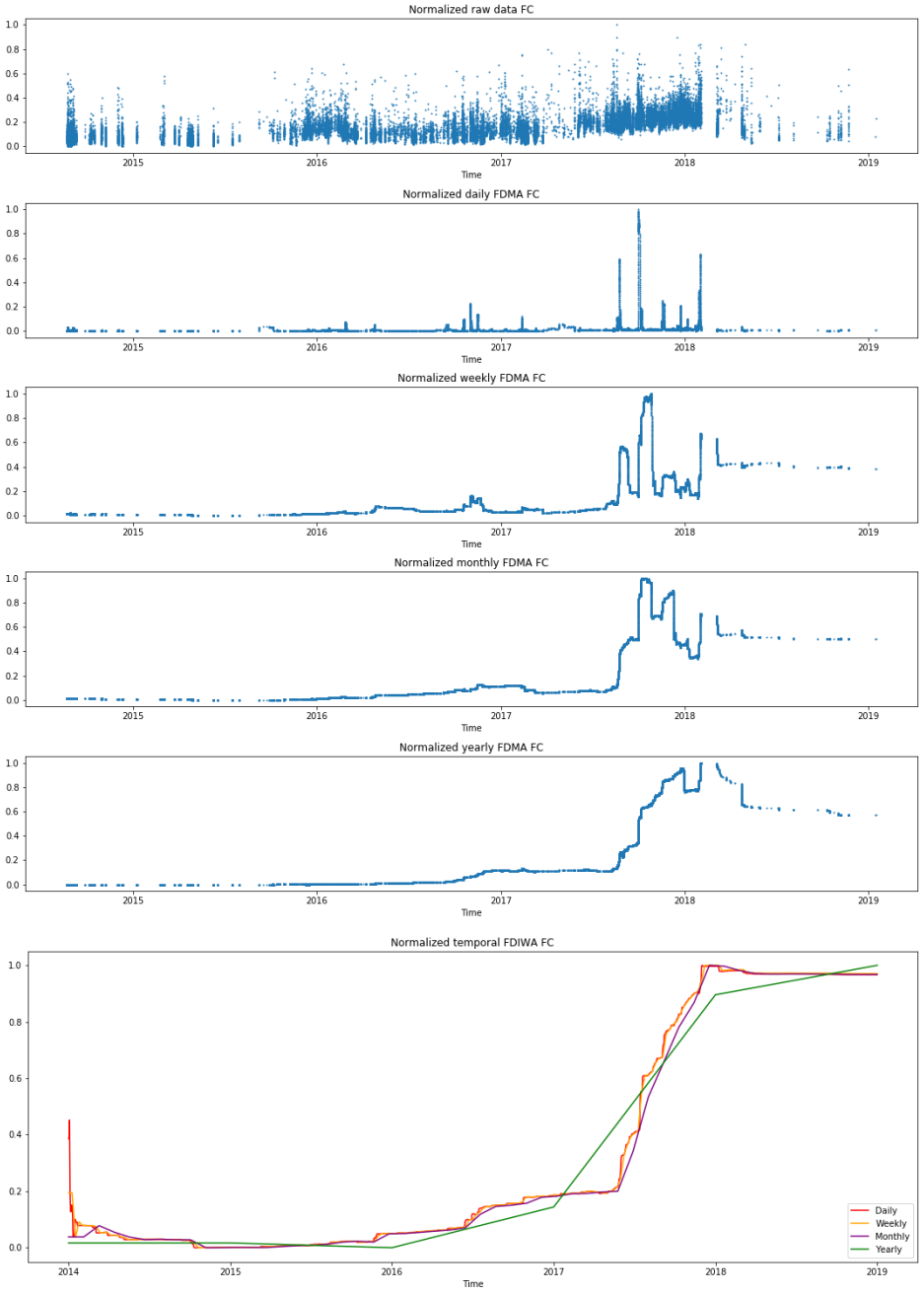


Figure 4.6: Fréchet distance, FD metrics of a single IGT using a $\mathcal{DM}$ autoencoder model with multiple time window representations of segmented moving average on the top and grouped incremental window on the bottom

other representations as a slight shift. For daily, weekly, and monthly time windows, the increments are from 0.0 to 0.2 and from 0.2 to 0.4, respectively. For the yearly time window, the increments are from 0.0 to 0.4 and from 0.4 to 0.8, respectively. The ADIWA representation shows the same shift location in all figures, with increments from 0.0 to 0.2 and from 0.4 to 1.0, respectively.

In Figure 4.6, it can be seen that most of the time-series exhibit an ascending trend, which is similar to those observed in the AD distance. However, the moving average calculation for daily, weekly, and monthly time windows only displays a few significant shifts. Upon closer inspection, the most significant shifts are observed at the end of 2017 and the beginning of 2018, with a small shift also seen at the end of 2016. These shifts can be ordered chronologically as follows: from 0.0 to 0.2, from 0.2 to 0.5, reaching a maximum of 1.0, and then from 0.3 to 0.6. When looking at the yearly time window for FDIWA, the increase is gradual but higher slopes are observed during the same periods of time as those seen for FDMA.

As shown in both figures, shifts have different slopes as well as amplitude; indeed, shift increments do not have the same dimension, but the drift is observed in all representations.

A further observation was performed by mapping the current obtained indicator with the one used in the company to assess the IGT. Figure 4.7 illustrates visually the correlation between the feature in a pair grid plot. The desired features are presented in the right-most column and the row at the bottom. These shown the correlation between the thermodynamic indicators together with the $\mathcal{RD}_{MD}$.

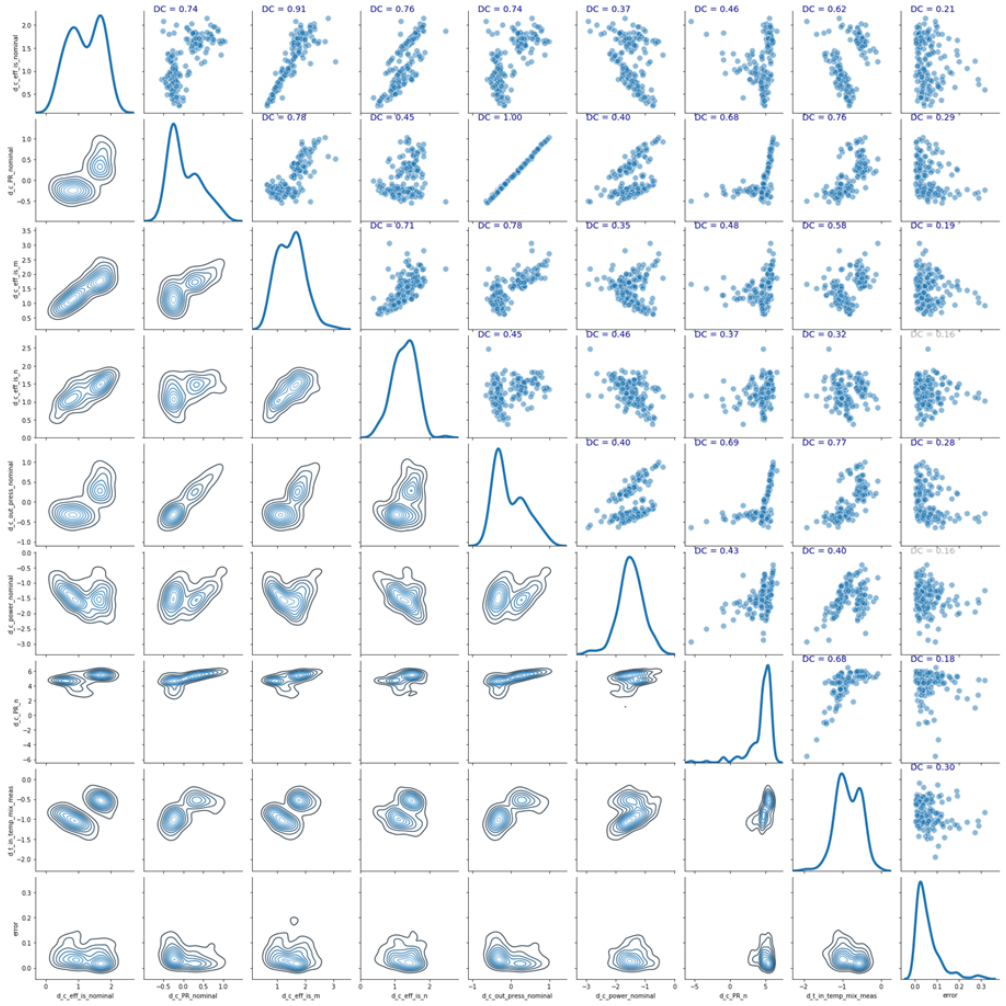


Figure 4.7: Representation of correlations amongst thermodynamic indicators and the Fresh Reconstruction Discrepancy

Considering each plot individually, any pair grid representation of thermodynamics indicators does not exhibit a distinct correlation. Thereby, this observation provides additional insights into the behavior and performance of IGTs.

4.5 Topological Analysis

At this stage, the first results were obtained showing a drift on IGTs performance. Then, an enhancement of the model was proposed by reducing the input features to raw features as,

- *Inlet Pressure*, $P_{c,i}$. Input air pressure into the compressor.
- *Pressure Ratio*, rP_c. Pressure Ratio.
- *Inlet Temperature*, $T_{c,i}$. Input air temperature into the compressor.
- *Outlet Temperature*, $T_{c,o}$. Output air temperature from the compressor.
- *Ambient Humidity*, H_a. Air humidity in the external.

Furthermore, an additional enhancement was suggested by imposing sparsity constraints on the hidden and latent space neurons. This approach enabled the formulation of an expanded model structure, resulting in an increased number of variables in both components. As a consequence, the structure was established as 5 - 8 - 8 - 8 - 5 and an illustrations is presented in Figure 4.8.

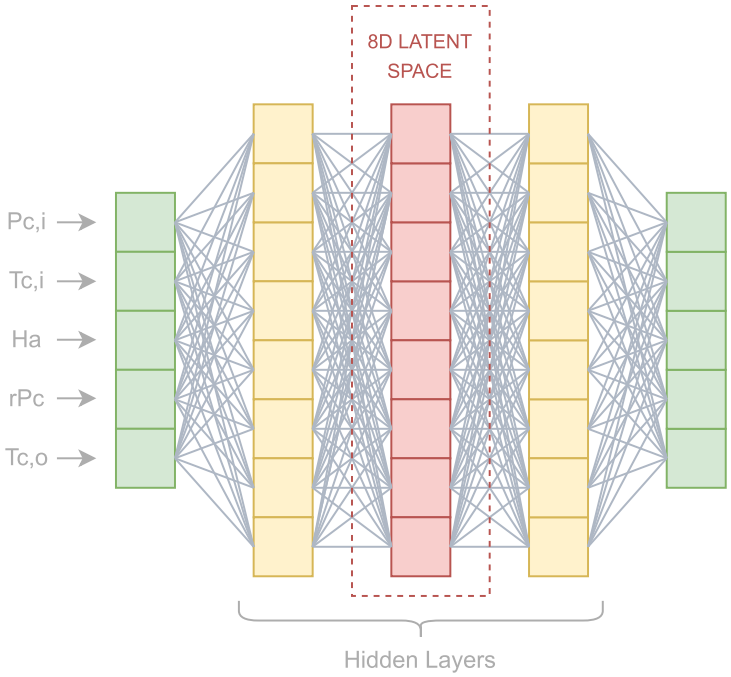
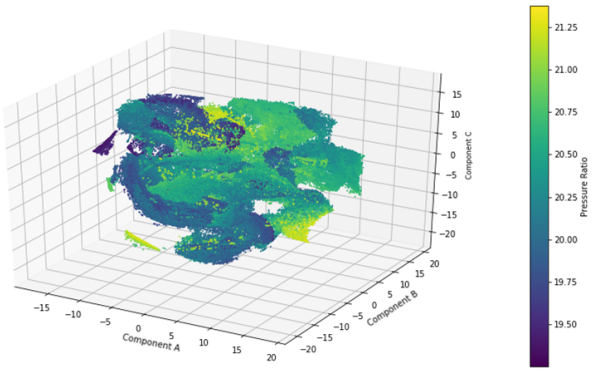


Figure 4.8: Representation of the autoencoder used to perform latent space topological analysis

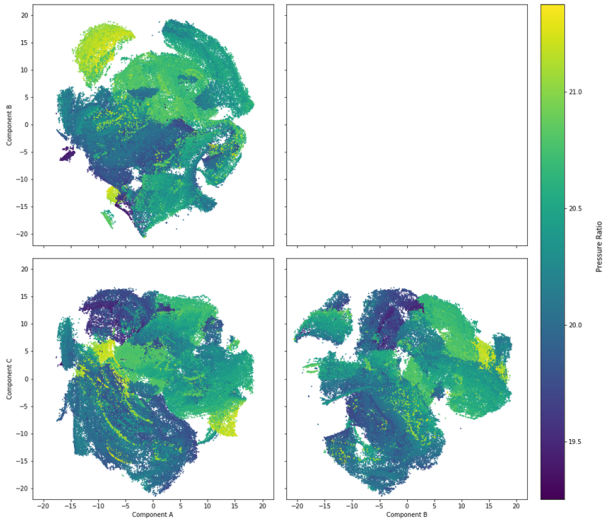
The perspective of the investigation was also shifted to concentrate on scrutinizing the latent space of the model. In order to gain a

better comprehension and visualization of the code variables, dimensionality reduction techniques, i.e. t-SNE and UMAP, were proposed to acquire further insights.

Figure 4.9 illustrates the results of t-SNE in 3D as well as in the three perspectives. It did not present any discernible results. It presents a huge mass where neither a perturbation nor any identifiable pattern can be observed.



(a) 3D representation



(b) Perspective representations

Figure 4.9: $\mathcal{DM}$ latent space using t-SNE technique

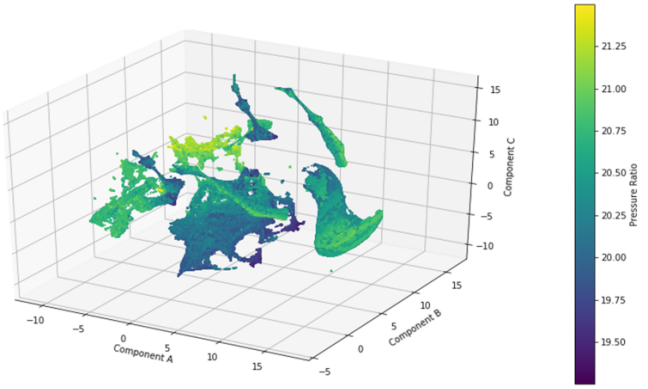
Similarly, Figure 4.10 depicts the UMAP results displayed in the same manner than previously. It reveals a clear drift in the data as a separate section in the data points appeard. This drift was identified

as the "ghost shape" and, after some analysis, it is observed to have correlation with major maintenance. Therefore, the drift that was previously observed in the $\mathcal{RD}$ indicator, is also observable in the latent space of the new expanded model. This finding is significant as it changes the drift perspective to a new indicator where the major maintenance can be properly captured. Additionally, it also shows the potential of UMAP to identify subtle patterns in the data that may not be apparent through other techniques. Further analysis was also conducted to determine the significance of any particular part of the ghost shape concerning a specific operation, but unsatisfactory results were obtained.

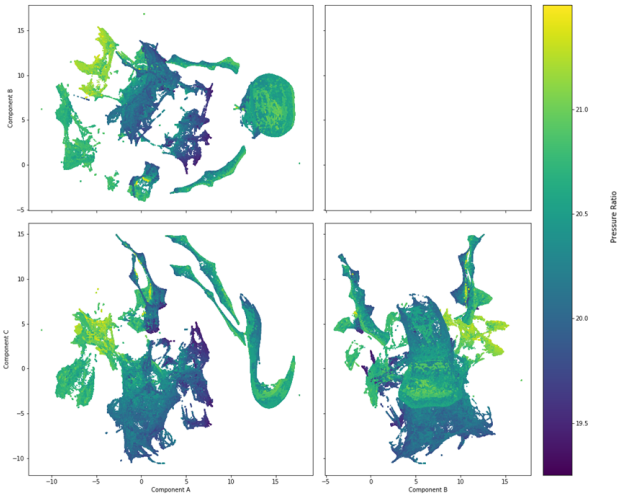
4.6 Fresh Reconstruction Discrepancy with an Expanding Model Autoencoder

In this phase, the model was adapted to prioritize explainability and reproducibility. The aim was to design a model that could be robustly applied to multiple IGTs, only using raw features. Once the impact of major maintenance on IGT performance was identified, these new criteria were imposed, which needed changes in the model. As a result, the focus shifted towards a low-dimensional code, utilizing raw features as input.

Four different kinds of autoencoders were used to perform the condition analysis, which differed in their structure and latent dimensions. In terms of structure, both a FC and sparse structures were selected. Table 4.1 presents the model combinations of $\mathcal{EM}$ and its corresponding latent space dimension.



(a) 3D representation



(b) Perspective representations

Figure 4.10: $\mathcal{DM}$ latent space using UMAP technique

Table 4.1: The combinations of the two considered autoencoder types and the two possible latent dimensions

		Latent Dimension	
		2D	**3D**
AE structure	Fully Connected	$\mathcal{RD}_{\mathcal{EM}}^{FC2D}$	$\mathcal{RD}_{\mathcal{EM}}^{FC3D}$
	Sparse	$\mathcal{RD}_{\mathcal{EM}}^{Sparse2D}$	$\mathcal{RD}_{\mathcal{EM}}^{Sparse3D}$

After several attempts, it was determined that MSE was the most suitable metric for computing the $\mathcal{RD}_{\mathcal{EM}}$ in these methods, in contrast to the metric used for drift evaluation in the $\mathcal{DM}$ evaluation. Similarly, the smoothing time window calculations were no longer utilized, and instead, a grouping frequency of hourly data points (i.e., $N = 60$ 1-minute samples) was used. These changes resulted in markedly different outcomes.

Furthermore, to gain additional insights into the $\mathcal{RD}_{\mathcal{EM}}$ indicator, both maintenance replacement events and inspections are mapped alongside it. As described in Section 3.3.1, three distinct datasets are utilized to create and evaluate the IGT model. The first dataset consists of 1 year of data, spanning from January 3rd, 2017 to January 3rd, 2018, and is used for training purposes. The second dataset, which covers the period from January 4th, 2018 to January 1st, 2019, is utilized for testing, before the occurrence of major maintenance. Finally, the post-maintenance dataset, which runs from January 22nd, 2019, after completion of the replacement tasks (represented by the wider green line in Figure 4.11), to January 1st, 2020, is used. In addition, minor maintenance and inspection events are also taken into account (and marked with a green line), in order to determine their impact on equipment operation.

Figure 4.11 displays the $\mathcal{RD}_{\mathcal{EM}}^{FC3D}$ model performance. From the results, it is evident that inspections (represented by thicker green lines) did not result in any changes to the operation of the IGT. However, significant maintenance work (represented by the widest green line) produced a significant deviation. Based on these initial findings, the inspections were deemed unnecessary for this study, as they did not contribute meaningfully to the results.

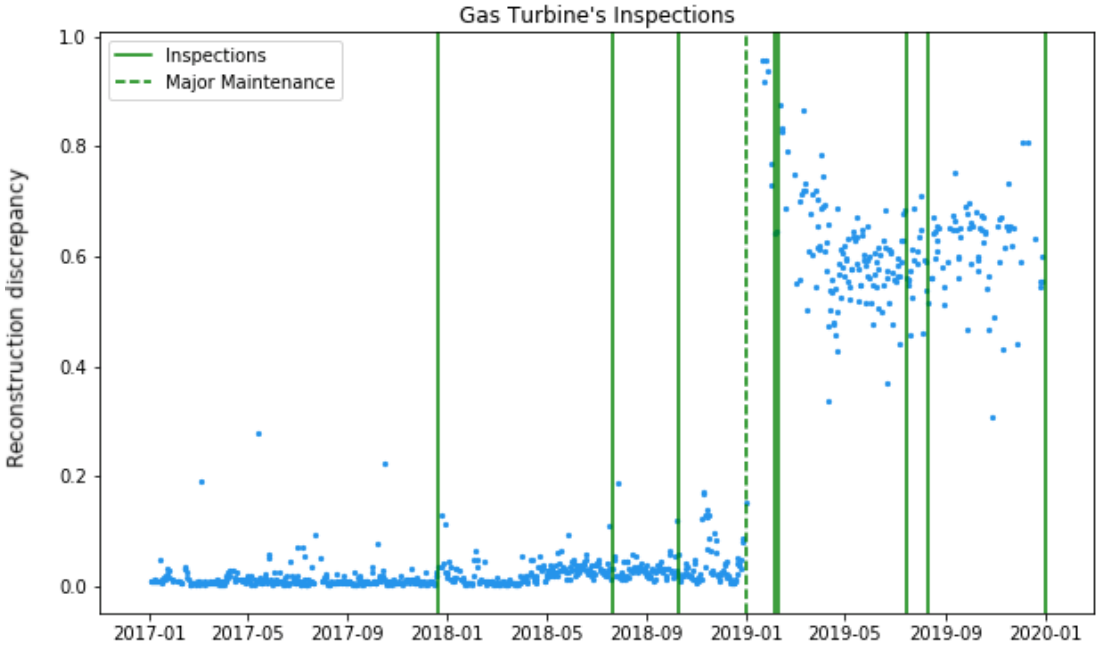
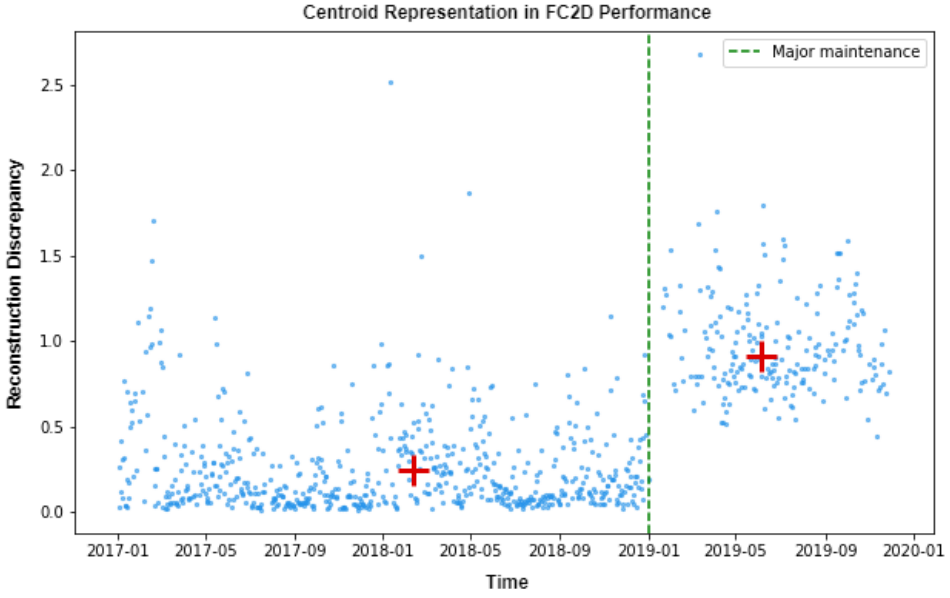


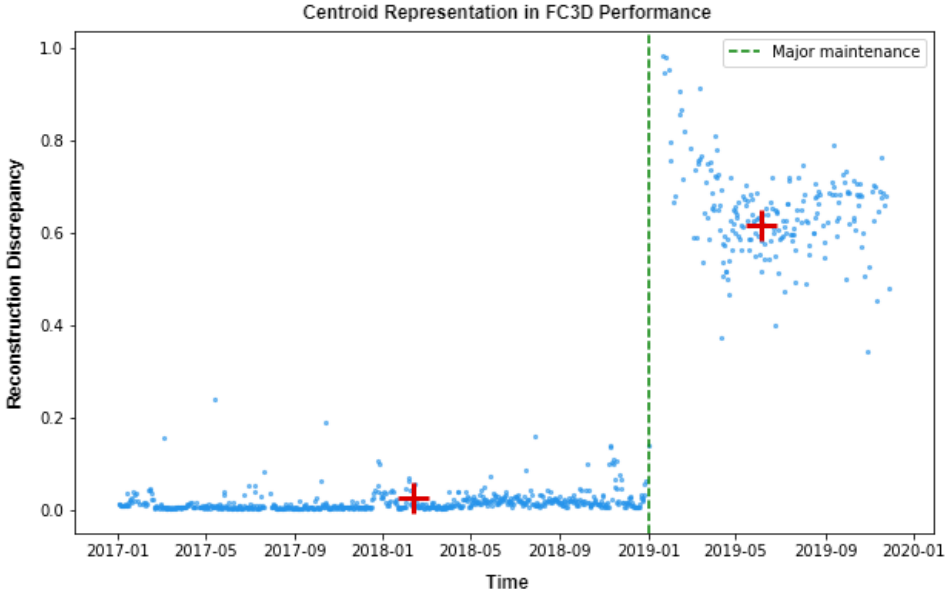
Figure 4.11: Maintenance dataset together with Fresh Reconstruction Discrepancy of $\mathcal{EM}^{FC3D}$ model

Next, several experiments were conducted focused on major maintenance events and their impact on the performance of its operation. These experiments provided valuable insights into the system's behavior under different conditions.

Figure 4.12 and Figure 4.13 summarise the effect of the replacements $\mathcal{R}$ during major maintenance in $\mathcal{RD}_{EM}$ for the four tested autoencoder structures. $\mathcal{R}$ is displayed as a dashed vertical green line, while the $\mathcal{RD}_{\mathcal{EM}}$ values are displayed in blue. Furthermore, the centroid of each distribution, i.e., before and after maintenance, is displayed as a red cross. According to the obtained results, the fully connected, FC models showed a clearer distinction before and after maintenance operation than sparse ones (see Figure 4.12). Moreover, the best results were obtained using the $\mathcal{RD}_{\mathcal{EM}}^{FC3D}$ autoencoder model, in which a clear shift can be observed when a replacement procedure was carried out.

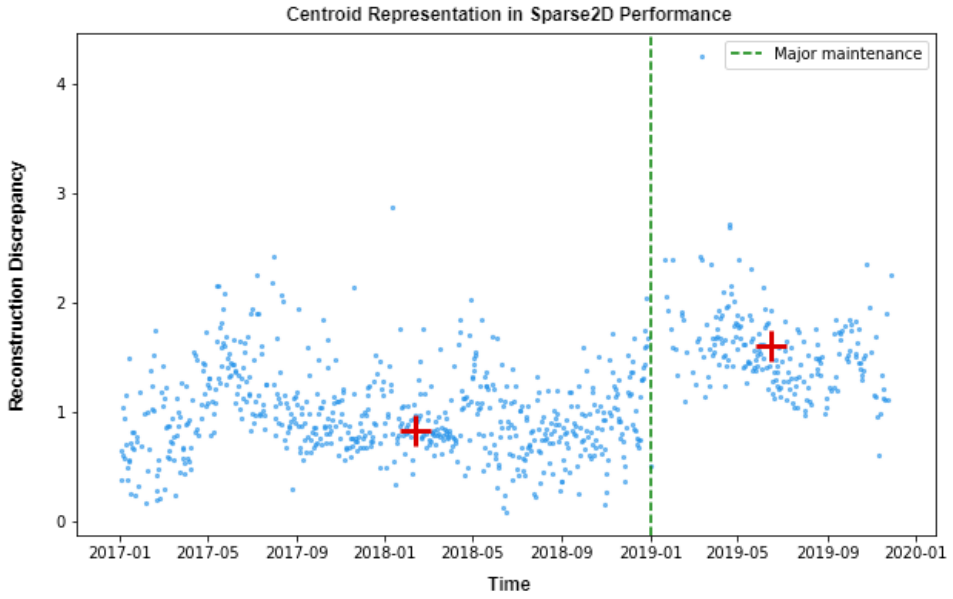


(a) $\mathcal{RD}_{\mathcal{EM}}^{FC2D}$ performance

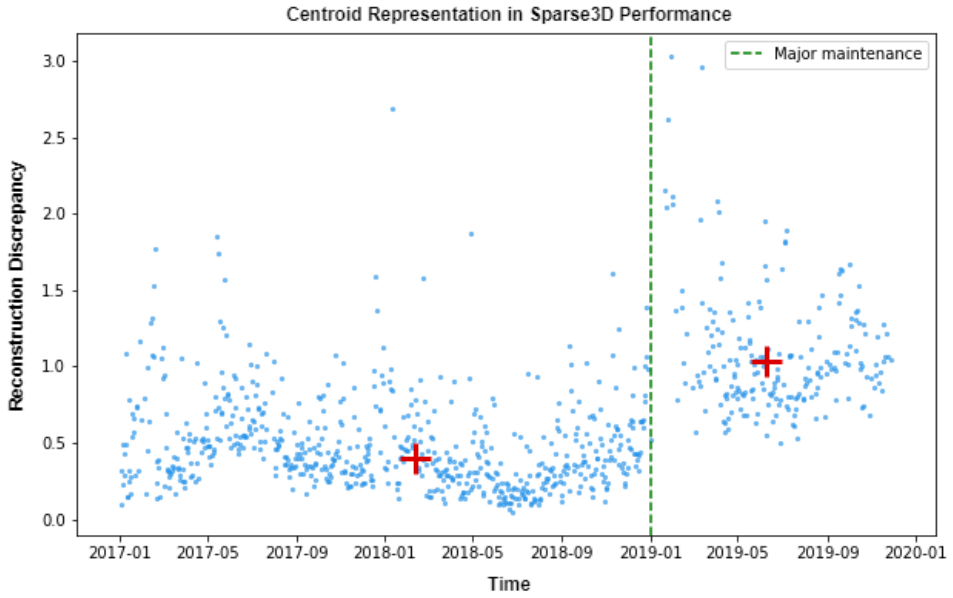


(b) $\mathcal{RD}_{\mathcal{EM}}^{FC3D}$ performance

Figure 4.12: Major maintenance visualization through Fresh Reconstruction Discrepancy in combination with fully connected $\mathcal{EM}$ autoencoder structures



(a) $\mathcal{RD}_{\mathcal{EM}}^{Sparse2D}$ performance



(b) $\mathcal{RD}_{\mathcal{EM}}^{Sparse3D}$ performance

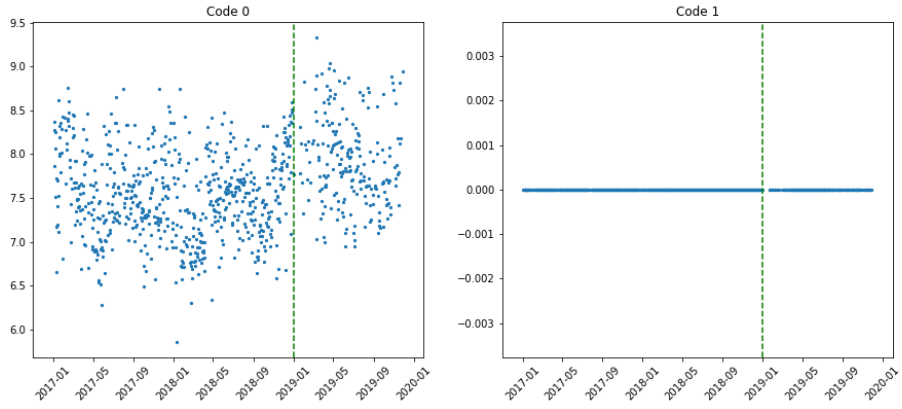
Figure 4.13: Major maintenance visualization through Fresh Reconstruction Discrepancy in combination with sparse $\mathcal{EM}$ autoencoder structures

4.7 Latent Space in an Expanding Model Autoencoder

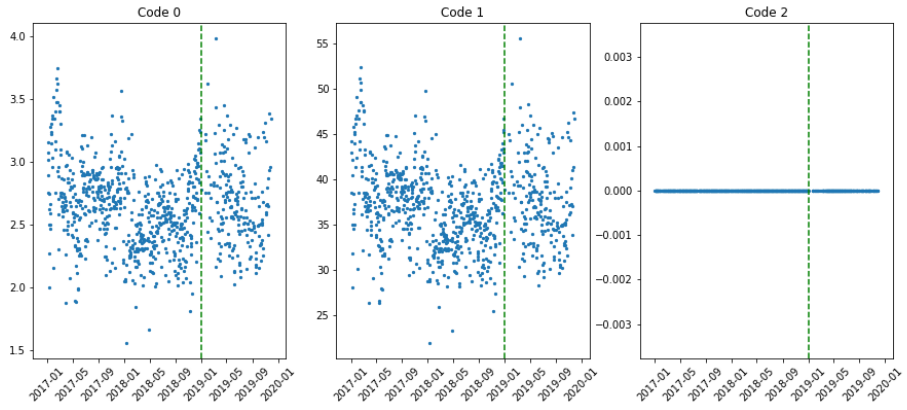
Latent space refers to the central part of the autoencoder, also named *code*. It is defined as the most characteristic part of these kind of models. Code aims to compress data in order to store the most relevant features. Therefore, it can be defined as a compressed representation of all significant data distribution. Moreover, it allows to get rid of any irrelevant information, indeed noise, to only put the focus on the most important characteristics. This part of the algorithms is crucial in the learning step because it is responsible to feed the decoder with significant data and therefore, to ensure a proper reconstruction.

$\mathcal{EM}$ autoencoders were constructed to clear visualize the code for time series. Results can be observed in Figure 4.14 and Figure 4.15. A similar analysis like the one developed in the previous section is completed by using the latent space of the time series together with the replacements, $\mathcal{R}$ events.

For the latent representations, the slope become more evident in the $\mathcal{EM}^{Sparse}$ representations (see Figure 4.15) while it is not differentiated in the $\mathcal{EM}^{FC}$ autoencoders (see Figure 4.14). Delving into representations, a flat line is presented in almost every model. This can be attributed to the use of the ReLU activation function, as explained in Section 3.4.2.3. ReLU only activates when positive numbers are passed, and produces 0 for negative numbers. Despite this flattening effect, the dimension cannot be removed, as the computation of the cells before and after compression is necessary for pattern learning. Tests were also conducted using 1D representations in the latent space, but unsatisfactory results were obtained.



(a) $\mathcal{EM}^{FC2D}$ Latent space representation

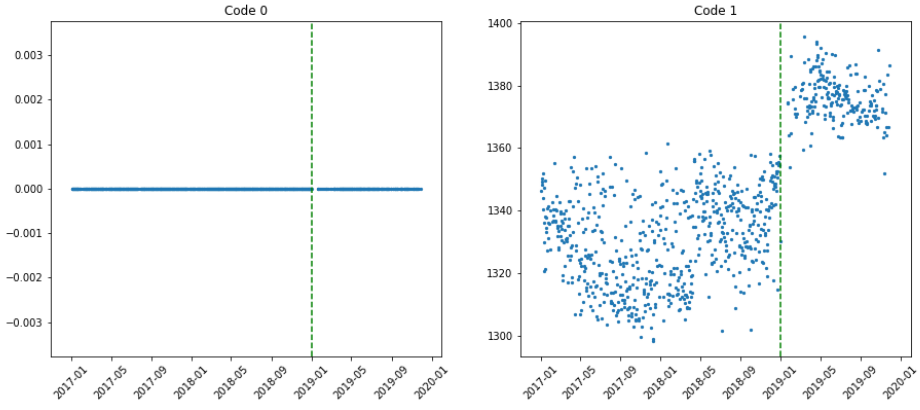


(b) $\mathcal{EM}^{FC3D}$ Latent space representation

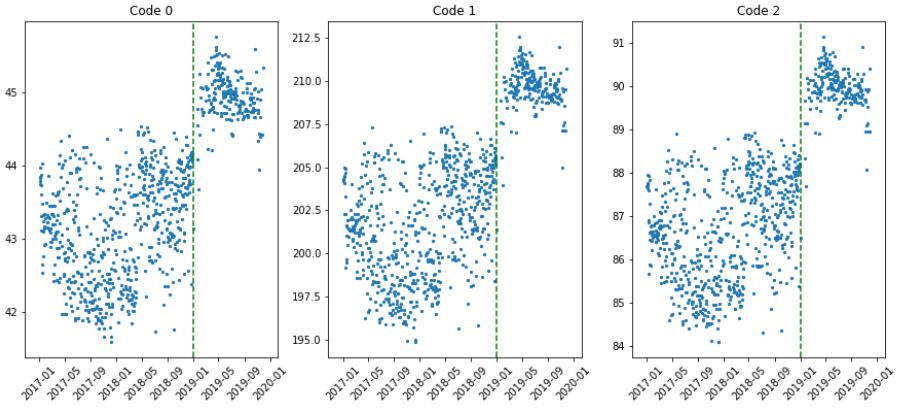
Figure 4.14: Major maintenance together with latent space representation in combination with fully connected $\mathcal{EM}$ autoencoder structures

4.8 Clustering Analysis

Clustering methods are a powerful tool that can be applied to different types of data to identify internal patterns and structures. In this case, two popular clustering algorithms, k-means and DBSCAN, are used to both, $\mathcal{RD}_{\mathcal{EM}}$ models and their latent representations. The aim is to find an automatic way to identify when major maintenance is performed, without human visual inspection, as well as to discover other potential patterns that may provide insights into the performance of the IGTs.



(a) $\mathcal{EM}^{Sparse2D}$ Latent space representation



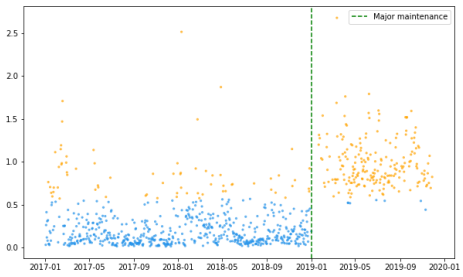
(b) $\mathcal{EM}^{Sparse3D}$ Latent space representation

Figure 4.15: Major maintenance together with latent space representation in combination with sparse $\mathcal{EM}$ autoencoder structures

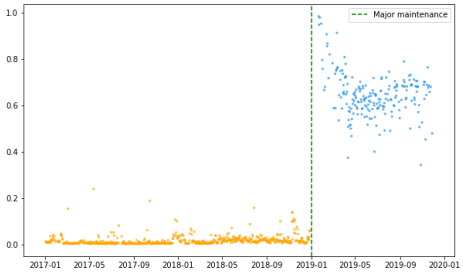
4.8.1 Clustering Evaluation of Expanding Model Autoencoders

The combination of the $\mathcal{RD}$ applied in the $\mathcal{EM}$ models together with the two clustering measurements presented in Section 3.5.4 provides as results those visualized in Figure 4.16 and Figure 4.17, where clusters aim to better describe the consequence of a major maintenance in the IGT operation.

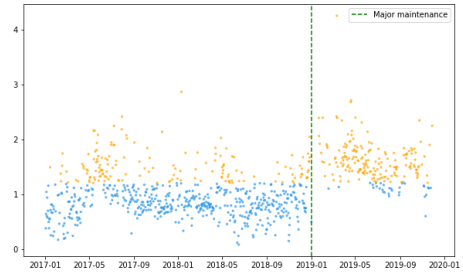
These figures show that clustering is clearly identifying a major maintenance, for k-means and DBSCAN algorithms, for the $\mathcal{EM}^{FC3D}$ representation. After several attempts and hyperparameter tuning, k-means method present several data clusters which do not correspond to any operation changes for the rest of representations; whereas,



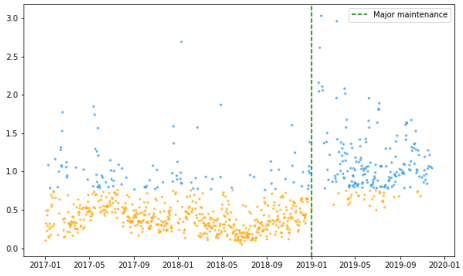
(a) $\mathcal{RD}_{\mathcal{EM}}^{FC2D}$ k-means representation of the latent space



(b) $\mathcal{RD}_{\mathcal{EM}}^{FC3D}$ k-means representation of the latent space

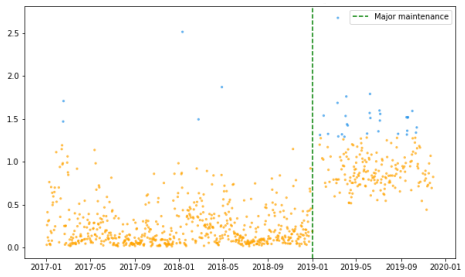


(c) $\mathcal{RD}_{\mathcal{EM}}^{Sparse2D}$ k-means representation of the latent space

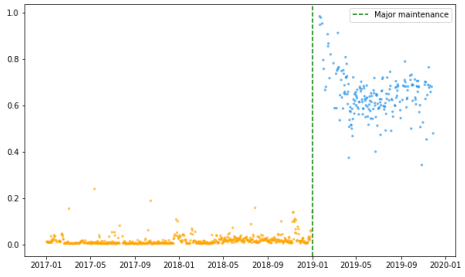


(d) $\mathcal{RD}_{\mathcal{EM}}^{Sparse3D}$ k-means representation of the latent space

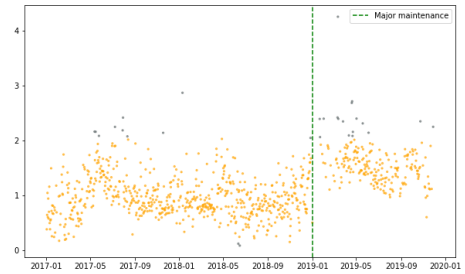
Figure 4.16: k-means method applied to Reconstruction Discrepancy metric using the $\mathcal{EM}$ autoencoder structure and the two latent space dimensions



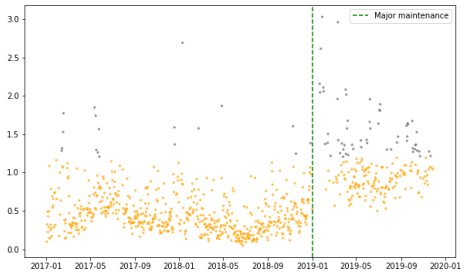
(a) $\mathcal{RD}_{\mathcal{EM}}^{FC2D}$ DBSCAN representation of the latent space



(b) $\mathcal{RD}_{\mathcal{EM}}^{FC3D}$ DBSCAN representation of the latent space



(c) $\mathcal{RD}_{\mathcal{EM}}^{Sparse2D}$ DBSCAN representation of the latent space



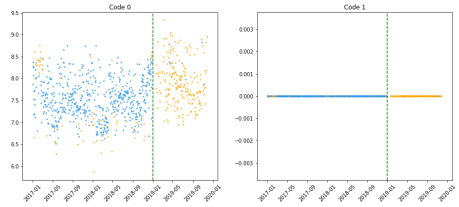
(d) $\mathcal{RD}_{\mathcal{EM}}^{Sparse3D}$ DBSCAN representation of the latent space

Figure 4.17: DBSCAN method applied to Reconstruction Discrepancy metric using the $\mathcal{EM}$ autoencoder structure and the two latent space dimensions

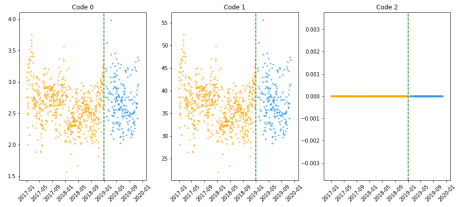
DBSCAN representations only detected a unique density based cluster.

4.8.2 Clustering Evaluation for Latent Space of Expanding Model Autoencoders

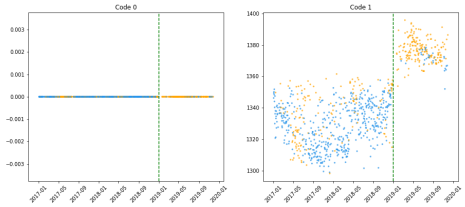
Next, the use of clustering methods in the $\mathcal{EM}$ autoencoders' latent space representations lead to the results visualized in Figure 4.18 and Figure 4.19, where the clustering results, indeed k-means and DBSCAN results, are displayed respectively.



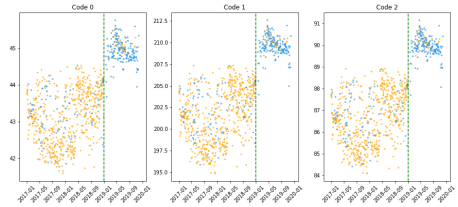
(a) $\mathcal{EM}^{FC2D}$ k-means latent space representation



(b) $\mathcal{EM}^{FC3D}$ k-means latent space representation



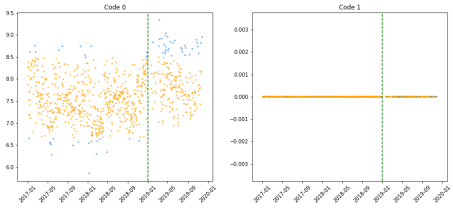
(c) $\mathcal{EM}^{Sparse2D}$ k-means latent space representation



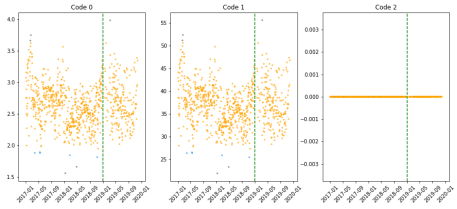
(d) $\mathcal{EM}^{Sparse3D}$ k-means latent space representation

Figure 4.18: k-means method applied to latent space representation using the $\mathcal{EM}$ structure and the two latent space dimensions

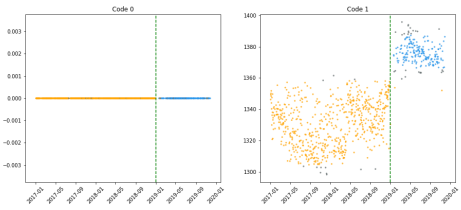
Figure 4.18 and Figure 4.19 display the clusters in the latent representation where more obvious patterns are found in comparison with $\mathcal{RD}_{\mathcal{EM}}$ grouping. In this case, k-Mean method present more accurate clusters even though some false positive can still be seen, where the most reliable results are in 3D latent representations, indeed $\mathcal{EM}^{FC3D}$ and $\mathcal{EM}^{Sparse3D}$ models. Regarding the DBSCAN methods, a clear distinction is shown in the sparse models while any



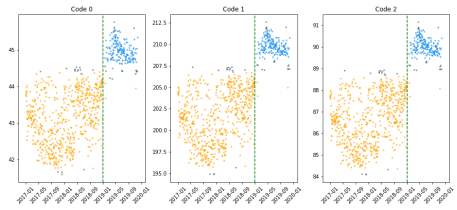
(a) $\mathcal{EM}^{FC2D}$ DBSCAN latent space representation



(b) $\mathcal{EM}^{FC3D}$ DBSCAN latent space representation



(c) $\mathcal{EM}^{Sparse2D}$ DBSCAN latent space representation



(d) $\mathcal{EM}^{Sparse3D}$ DBSCAN latent space representation

Figure 4.19: DBSCAN method applied to latent space representation using the $\mathcal{EM}$ structure and the two latent space dimensions

clear pattern is not shown in the FC models. It is due to the nature of latent space construction and the sparsity constraint in neurons.

4.9 Temporal analysis

A temporal analysis is conducted by deconstructing the values in $\mathcal{RD}_{\mathcal{EM}}$ looking for new hidden patterns or fluctuations in data. The methods proposed to decompose the indicator are tDWT and Time Series Decomposition, which were further explained in Section 3.5.2.

4.9.1 Time Series Decomposition

The decomposition process involves separating time series into underlying systematic components, which include trend and seasonality, and random or irregular component, which represents the noise or fluctuations in data. By isolating these individual components, analysts can gain a better understanding of the underlying patterns and make more informed decisions based on data. Overall, time

series decomposition provides a useful tool for understanding the behavior of the system and identifying potential issues that may impact its performance. Figure 4.20 illustrates the decomposition with daily period of decomposition.

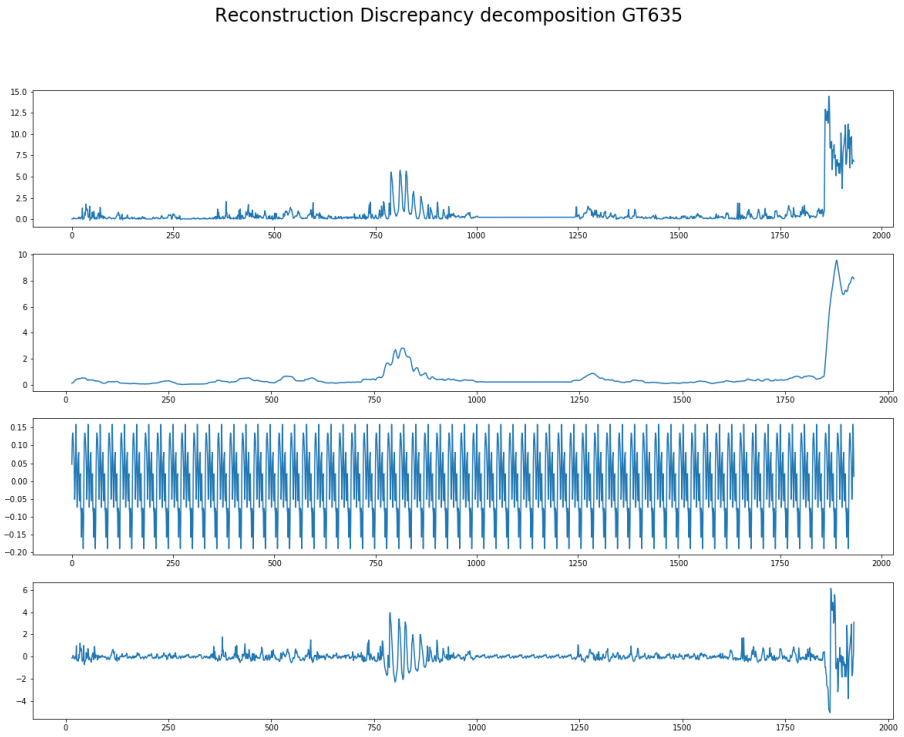


Figure 4.20: (First) Raw values of the RD time-series. (Second) Trend of RD after deleting seasonality and residuals. (Third) Seasonality component. (Fourth) Residuals component.

The decomposition analysis indicates a distinct trend shift caused by a significant maintenance operation. Additionally, an unidentifiable perturbation is observed. The examination of the seasonal component exposes the cyclical pattern of daily operation, while the residual component of the time series indicates the existence of unexplained variability in the data, which implies the presence of other factors that could impact the performance of the IGT. However, due to an inadequate comprehension of fault events, this analysis cannot yield further insights.

4.9.2 Discrete Wavelet Transform

DWT are mathematical tools for analyzing data where feature vary over different scales. These features can be frequencies, transients or slowly varying trends. In this case, the transformation is used to further analyse the $\mathcal{RD}_{\mathcal{EM}}$ autoencoder using the symlet wavelet function.

Figure 4.21 presents the results when applying the discrete wavelet transformation. The coefficients display unexplained variability, indicating the possibility of various factors that could potentially impact IGT performance. However, this analysis does not yield any additional insights beyond the primary maintenance shift. Similarly, the perturbations evident in the coefficients could not be thoroughly analyzed due to insufficient data at this stage of the study.

4.10 Long-term Performance of Fresh Reconstruction Discrepancy

Finally, at the end of the thesis, a last analysis is conducted to demonstrate the robustness and consistency of both, methodology and models. These experiments consist of using all available data, indeed five years of data, to evaluate the $\mathcal{RD}_{\mathcal{EM}}$ performance. Firstly, the long-term $\mathcal{RD}_{\mathcal{EM}}$ is illustrated in Figure 4.22, then the application of clustering method is presented in Figure 4.23.

The results of the study indicate the presence of two consecutive slopes observed at the beginning of the year 2019 and 2021, respectively. As previously defined, these shifts correspond to major maintenance events that have impacted the performance of the equipment. In order to provide a clear visualization of the results,

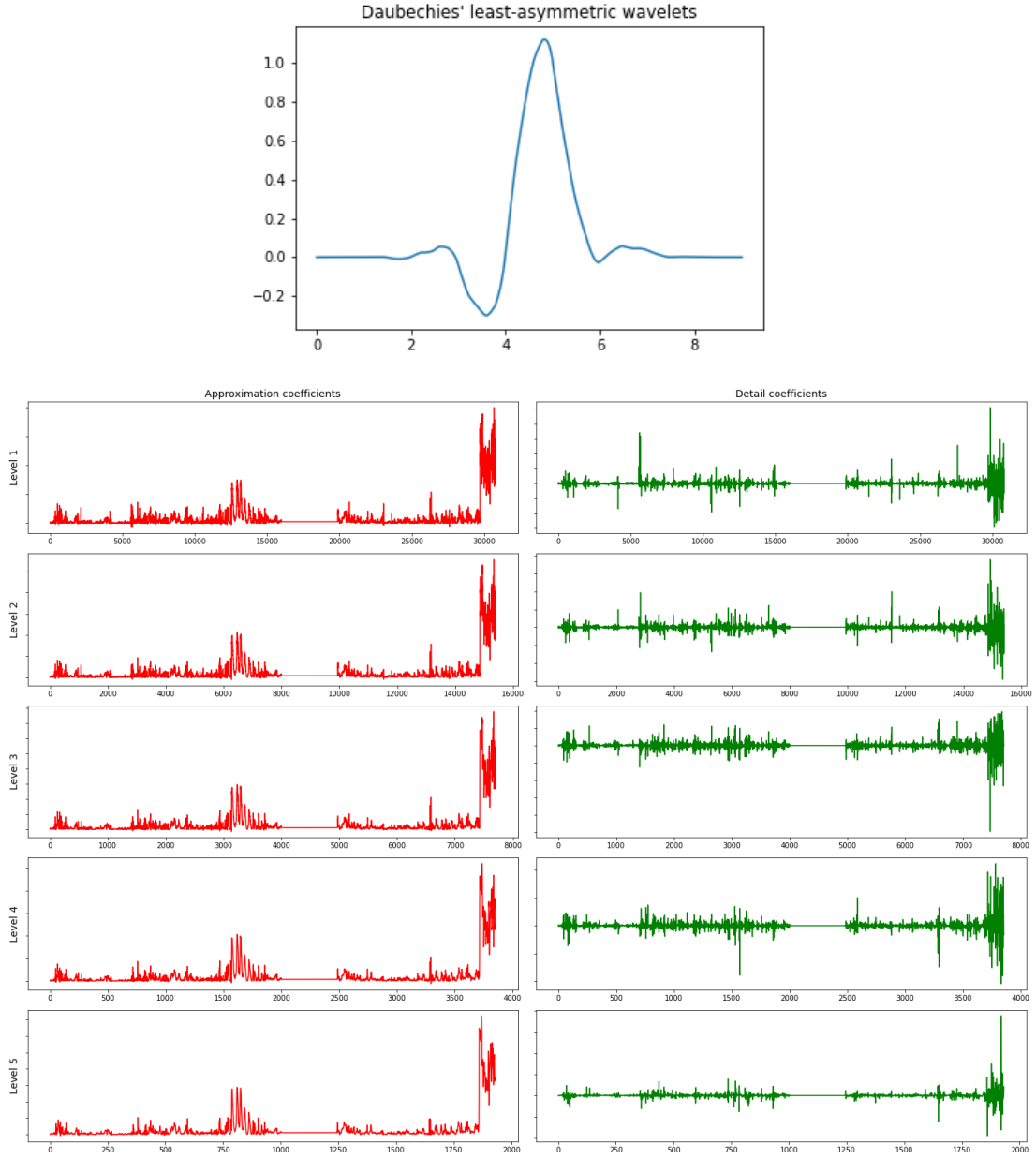


Figure 4.21: (Up) The form of the used wavelet, Symlets function. (Down - left) The fifth harmonics filtered using the Symlets function. (Down - right) The filtered signal using the Symlets function.

the DBSCAN clustering method was employed to identify and capture the shifts more accurately. However, the information regarding the second major maintenance event was not available to the researchers, needing the verification of the results with the company. The obtained results were consistent with the anticipated outcomes,

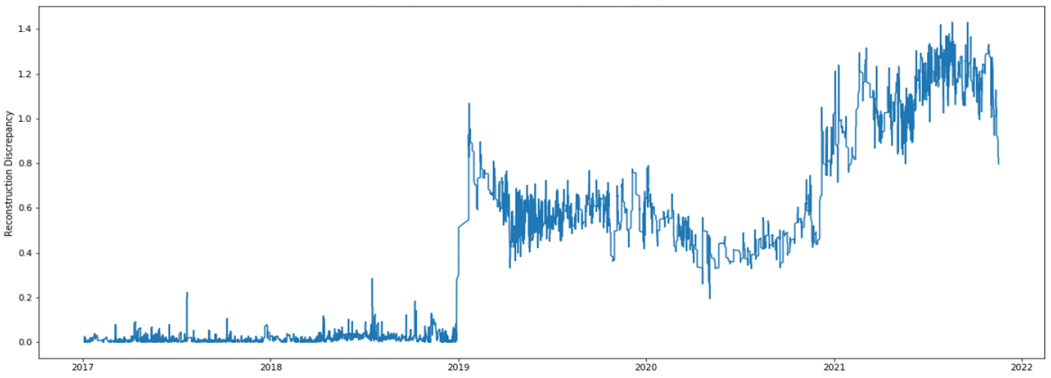


Figure 4.22: Long-term performance of $\mathcal{RD}_{\mathcal{EM}}$ indicator using five years of data

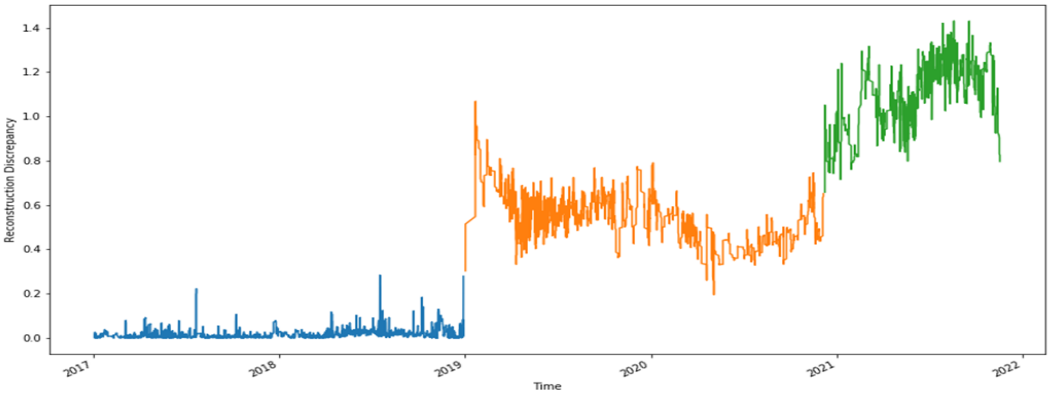


Figure 4.23: Long-term performance of $\mathcal{RD}_{\mathcal{EM}}$ indicator using the DBSCAN clustering method

confirming that the second shift also corresponded to a major maintenance event. This verification step helped to strengthen the validity and reliability of the study findings.

Chapter 5

Discussion

This research study presents several methods for assessing industrial gas turbines performance using solely compressor data and ML based algorithms. This section provides a detailed account of the entire process, from data analysis to the presentation of final results. Each stage of the process has played a significant role in achieving the desired outcomes, and therefore, the methodology used to obtain the results is equally important. The section is structured to begin with a discussion of data treatment, including the selection of IGTs and data preprocessing techniques. Next, a detailed examination of the model architecture and its variants is offered. Finally, two main sections are derived from the results section: IGT drift evaluation and IGT model analysis. Within each section, a more in-depth explanation is provided, focusing on the various techniques employed in the study.

5.1 Data Evaluation

The evaluation of the model's performance relied on two different data sources. Firstly, the NASA Turbofan Jet Engine data was utilized to assess the efficacy of the proposed model in analyzing the SE equipment. Secondly, the Siemens Energy's provided data was employed to evaluate the performance of the selected equipment. The utilization of the NASA Turbofan Jet data allowed for preliminary insights into the potential outcomes of implementing the model in the analysis of IGT equipment. Unlike NASA Turbofan Jet Engine dataset, which had already undergone curation for research purposes, significant efforts were required to properly curate the data provided by Siemens Energy in order to yield optimal results. By employing both datasets, the study aimed to provide a comprehensive evaluation of the model's performance in analyzing IGT equipment

Before proceeding with data curation, the selection of the most appropriate IGTs based on the data from SE repositories was necessary. The primary objective was to ensure that the chosen equipment package type was the most useful for the analysis since different machine types may conduct to different types of operations, thus making it challenging to reproduce the results for other IGTs. The meta-analysis was conducted to determine the package type that had the most instances and total IGT and sensor data available. The results revealed that the most significant IGT type to use was the 'Package Type 11', which enclosed the highest number of instances and all the required sensors. It was also found that GTs operate differently based on ambient conditions, mostly the inlet temperature. Hence, analyzing sensor data could help refine the comprehension of which

the equipment is suitable to work with under similar ambient conditions while avoiding uncommon operational modes. Additionally, the analysis confirmed that all the necessary sensors were available in the working IGTs. Consequently, out of the 58 studied instances, 19 were identified to be suitable for the analysis under proper ambient conditions.

Extensive data curation was carried out to ensure that the sensor data were appropriate for analysis under the selected assumptions. The operating conditions of the equipment were determined to operate under full load conditions to enable the detection of malfunctioning points. Multiple studies were conducted to determine the most suitable filtering method, and it was established that setting thresholds using the median absolute deviation (MAD) instead of standard deviation (STD) metrics was the most effective approach. Moreover, the grid regulation code was taken into account to determine whether some of the IGTs were operating at their limit. The VGV sensor was also utilized to confirm that the operation points were included in normal operating conditions. As a result of the filtering process, a significant amount of data was removed, and in some instances, certain IGTs had to be excluded due to insufficient data. As a consequence, the number of useful IGTs was ultimately reduced to twelve (12).

To assess the performance of the proposed model, three different datasets were utilized: Training/Validation set, Testing set, and Post-Maintenance set. The Training/Validation set was used to train the model. It was essential to determine whether the available data was sufficient to capture the GT's behavior in its first year of operation and included enough representative points. This assessment was performed using distribution plots, which were used to determine

the suitability of the feature for training a reliable model that could generate accurate results. Based on this analysis, it was found that two pieces of equipment lacked the minimum representation points across all distributions, which could result in training bias. Therefore, the number of IGTs was reduced from 12 to 10. Additionally, a further analysis was conducted to evaluate the correlation between the proposed input features for the initial model. As a result, it was discovered that two of the suggested features were highly correlated with others and were subsequently removed.

Next, a Testing data set was defined using all available data to evaluate the model. Additionally, the Post-Maintenance set was proposed based on the results obtained from the $\mathcal{RD}_{\mathcal{EM}}$ model. The maintenance events provided by the company during the mid-stage of the thesis were considered from that point forward.

The principal concern regarding the data pertained to the imperative of procuring high-quality data for the thorough training of the model. Additionally, it was essential to guarantee an adequate representation of instances to mitigate the potential introduction of bias during the training process. This task was the primary endeavor, and its successful accomplishment was greatly facilitated by several trial-and-error tasks and Siemens' specialists assessment.

5.2 Autoencoder Based Architecture Evaluation

The main goal of this thesis is to apply advanced data analytic and machine learning techniques to evaluate the operational status of IGTs. Several alternatives were contemplated, and an autoencoder

model was ultimately chosen due to its versatility in a range of applications such as anomaly detection, dimensionality reduction, and system performance modeling [120]. The model's effectiveness was initially evaluated on an appropriate dataset to gauge its performance in achieving the research objectives. However, the employment of such a model poses certain challenges needing special attention and exploration, leading to valuable insights into the operation of IGTs.

The first challenge is to select appropriate training data for autoencoder models. As previously mentioned, autoencoders are designed to learn and represent a manifold based on the training data. It is crucial to ensure that the training data accurately represents the testing data to prevent the model from concealing rather than revealing useful information. Thus, it is a keypoint to establish effective validation techniques to identify representative data for training purposes

An additional obstacle involves the considerable amount of data that is necessary to train an autoencoder. Training requires a significant amount of time for processing, tuning hyperparameters, and validating the model before constructing the actual model. Additionally, it is essential to identify an appropriate feature set that can effectively capture the desired behavior. Autoencoders are designed to acquire significant representations that are effective for reconstruction purposes. Hence, if the input data consists of only a small portion of relevant information for IGT condition assessment, the autoencoder may lose a substantial amount of data.

The selection of the autoencoder structure and type can significantly impact both the results and the internal analysis. Consequently, this

aspect has required substantial effort, necessitating a delicate balance between achieving a streamlined and comprehensible model for ease of study and understanding, while simultaneously maintaining robust results and accuracy. The process of arriving at the desired autoencoder structure and type has entailed a protracted period of trial-and-error.

Despite the challenges associated with using this type of model, the benefits of using an AE for IGT analysis outweigh the challenges. It provides a comprehensive framework for further analysis and insight generation from the data. Additionally, the two main results obtained from the model, namely the reconstruction error and the latent space, provide a foundation for further studies aimed at improving the understanding of the nature of the IGT based on compressed input features and model reconstruction performance. The reconstruction error allowed the capture of a comparable representation of the equipment based on the most noteworthy patterns in the data. The latent space, on the other hand, offered insight into the data's structure and the existence of any hidden patterns. Furthermore, given that the study involved real operational data, a robust model was required to handle non-linearities.

Furthermore, it is worth noting that there exist several useful variants of this type of model that could be employed in the study. In this particular case, only the sparse autoencoder was utilized, which enables the utilization of more hidden layers than the input. The inclusion of hidden layer constraints enables further investigation of the model's performance and provides an opportunity to identify new patterns in the data

In summary, it can be concluded that autoencoders are powerful

models with the essential capabilities to learn a meaningful representation of the input and model processes, providing a basis for further analysis to gain insights into the nature of system performance and significant data pattern representations.

5.3 Thermodynamic Evaluation

Prior to delving into the details of DA and ML tools, a thermodynamic analysis was conducted as the initial step. This analysis aimed to provide insights into the IGT's performance from a thermodynamic perspective, and to examine whether any patterns could be identified based on factors such as fuel consumption and pressure ratio. The pressure ratio, which is a key characteristic that directly impacts the isentropic efficiency, was given significant importance in this analysis.

Figure 4.1 on page 93 presented an analysis of the fuel consumption and the total power generated by the IGT. As it is well known, the ambient temperature can also affect the performance of the system. Therefore, both the pressure ratio and the inlet temperature are also included in the analysis to draw further insights from the data. However, the results of the analysis do not reveal any clear trend that could be linked to either the inlet temperature or the pressure ratio.

Another approach was taken in an attempt to gain insight into compressor performance by focusing on the pressure ratio. However, like the previous analysis, no significant results were obtained beyond confirming the expected thermodynamic process: a higher pressure

ratio corresponds to a higher amount of power generated and lower temperature values.

This lack of a clear trend could be due to the filtering methods used to ensure a high-quality dataset. These filtering methods may have been too restrictive, leading to a reduction in the number of instances in the dataset, and consequently, it may be difficult to obtain any visual pattern that could give some insights about the IGT's performance degradation, faults, and so on. Despite this limitation, the analysis provides a baseline understanding of the IGT's thermodynamic performance, which can serve as a starting point for more advanced analyses using DA and ML tools.

5.4 Topological Analysis

The purpose of the topological analysis went beyond the $\mathcal{RD}$ metrics and aimed to uncover any latent space patterns that could prove useful for examination, as well as to provide initial insights into reducing the number of variables in the code. This analysis was conducted using two well-known dimensionality reduction algorithms, namely t-SNE and UMAP. The primary goal of these algorithms is to reduce the number of features in a given set while preserving its internal structure.

Based on the analysis conducted in this study, several insights can be derived regarding both the methodology and the distribution in the latent space. Firstly, UMAP reveals a structured latent space where a distinct drift can be observed in the form of a 'ghost-like' shape. Upon further analysis, it was discovered that this drift corresponds to a major maintenance performed on the compressor. This suggests

that hidden patterns still exist in the latent space and by reducing the number of variables in the model, we could potentially obtain further insights.

Furthermore to the previous analysis, it should be noted that the UMAP algorithm has shown a higher capability to detect hidden patterns in the data compared to t-SNE. While t-SNE only produced a dense mass without any clear structure, UMAP was able to identify the ghost-like shape in the structured latent space. This suggests that UMAP may have a potential to identify subtle patterns in the data that may not be apparent through other dimensionality reduction techniques.

The topological analysis served as an intermediate step towards modifying the autoencoder model and gaining deeper insights into the hidden patterns within the data, as well as comprehensively examining the manifold of the autoencoder's latent space. In summary, it played a crucial role in advancing the study of the data and the model.

5.5 Model Evaluation

The primary contribution of this thesis is the development of an AI based tool that provides valuable insights into the performance of IGTs, particularly in relation to major maintenance scenarios. This tool is based on ML models that are capable of addressing issues such as drift in IGT operation and identifying changes in operational behavior resulting from major maintenance operations. The

effectiveness of the models is demonstrated by analyzing the reconstruction error and latent manifold generated by the autoencoder based model.

The principal challenge encountered during the comprehensive evaluation was the necessity for additional data to thoroughly assess certain trends within the results. While efforts were made to address this limitation to the fullest extent possible, the inclusion of more Siemens indicators would have facilitated a more seamless correlation of our findings with valuable insights into these trends. To mitigate this limitation, we approached the evaluation in a more qualitative manner, refraining from categorizing changes as either positive or negative events.

The results obtained from these analyses are presented and discussed in detail, highlighting the capabilities of each method. The significance of the findings lies in the fact that they provide a deeper understanding of the GT performance, allowing for more informed decision-making regarding maintenance and operations.

5.5.1 Drift Evaluation

Main practical result of this thesis is the identification of a drift in IGT operation from two distinct perspectives that resulted in two different outcomes. These outcomes are referred to as the $\mathcal{RD}$ metrics and are derived from two distinct models, namely the Differential Model and the Expanding Model, which have 9 and 5 input features, respectively. The initial outcome aims to detect the GT's behavior during its operation, utilizing the features described in the Section 3.2. A clear drift is revealed by combining two different metrics, namely absolute difference and Fréchet distance, along with

two calculations based on time-windows, namely moving average and incremental window average.

The second outcome detects a significant shift in the GT's operation following a major maintenance event. In this case, the deviation is detected using raw features and has a completely different model structure than the first outcome. Further outcomes are derived from this structure, which will be presented in this chapter.

Next, the two models, $\mathcal{DM}$ and $\mathcal{EM}$, are meticulously analyzed. The $\mathcal{DM}$ model is presented first, followed by the $\mathcal{EM}$ model.

5.5.1.1 Autoencoder Differential Model Insights

The autoencoder model adopts a compression based architecture to monitor significant changes in the condition of the compressor within an IGT. As previously stated, changes have been examined by utilizing a combination of two distinct distances and time-window calculations, while also analyzing the graphical progression over time.

The analysis using the absolute distance AD, as shown in Figure 4.5 on page 99, reveals significant drifts in the moving average representation in November 2015 and May 2017. In addition, a significant drift jump is observed in March 2018. The yearly time window captures a general shift in the compressor performance.

Next, in the incremental representation, similar patterns are detected where events are revealed as a change of slope. Furthermore, the compressor's overall behavior can also be assessed as in the yearly time window, thanks to the nature of its calculation, the cumulative function shown in Equation (3.19). The same ascending trend

is displayed, thus indicating the drift in the compressor's behavior compared to its initial performance.

The Fréchet distance, as presented in Figure 4.6 on page 100, reveals the most significant changes in the IGT's operation. Notably, the moving average representation indicates substantial drifts in 2016/10, 2017/08, 2017/10, and 2018/03. However, the shifts are slightly shifted when considering multiple time windows, likely due to the number of instances used to compute the distance for each case. The primary shift occurs in 2017/10, but the deviations in 2016/10 and 2018/03 are also noteworthy. Additionally, the yearly time window displays changes in the general behavior of the equipment.

Unlike AD, the incremental representation does not exhibit the same pattern. Only one notable change of slope is detected in 2017/08, but a general drift in behavior is still evident.

The only event that can be observed using both distances is the one that occurred in 2018/03. This is because the AE distance calculates the average error in a single time step, while the FD compares two multivariate distributions over a fixed time window, resulting in different representations. However, both distances exhibit an ascending trend over time, indicating a drift in the compressor's behavior relative to its initial performance.

5.5.1.2 Autoencoder Expanding Model Insights

At this stage of the study, there is a shift in the approach, where a different set of input features are utilized to evaluate the behavior of the compressor, and a novel model structure is proposed using expansion in hidden layers. When considering the training period,

it was found that for the first-year data, the fully connected FC architecture outperformed the sparse autoencoder, as evidenced by the lower values in $\mathcal{RD}_{\mathcal{EM}}$, as shown in Figure 4.12 on page 109 and Figure 4.13 on page 110. Additionally, among the models, the $\mathcal{RD}_{\mathcal{EM}}^{FC3D}$ had the best performance, with the training and validation data exhibiting the least amount of variability.

Although sparse architecture did not benefit from feature extraction, the study produced a clear result: a deviation in the operation point after maintenance performance was observed, even though the performance could still be maintained. In terms of $\mathcal{RD}_{\mathcal{EM}}$, all four tested models showed a positive slope after the maintenance operation regarding the data distribution centroid. The most evident distinction was found in the $\mathcal{RD}_{\mathcal{EM}}^{FC}$ models, particularly in the $\mathcal{RD}_{\mathcal{EM}}^{FC3D}$ model, where a clear shift was present. A more subtle shift was also observed in sparse models. Detecting this deviation is key to understanding how maintenance affects IGT performance and when an update of the internal operating models is necessary.

Furthermore, the modified model structure allows for a deeper analysis of the latent space distribution. As mentioned earlier, one of the primary objectives of altering the model architecture was to enhance interpretability using raw features and enable visualization of the code feature. Consequently, it becomes possible to conduct a more detailed assessment of how the data is distributed in the model's latent space.

As depicted in Figure 4.14 on page 112 and Figure 4.15 on page 113, a distinct slope can be observed, similar to the one observed in the $\mathcal{RD}$ metrics, indicating a major maintenance event. This same phenomenon can be observed in both the reconstruction discrepancy

and the latent distribution. Moreover, it is evident that the slope observed in the sparse representations is steeper than in the fully connected FC models. The FC model is designed based on the interaction of all cells, without any restrictions, which results in a chaotic distribution in the center of the model, i.e., the code. In contrast, the sparse model incorporates a sparsity constraint during training, leading to the construction of a structured latent space.

In summary, this type of model provides additional insights about how the operation of the GT shifts in response to maintenance events and how it is represented in each model. The $\mathcal{RD}$ metric shows a higher deviation in the fully-connected models, while the latent distribution analysis reveals that the sparse based architecture has a more structured manifold and better distinguishes operating conditions before and after maintenance.

The final result obtained using this methodology is related to the long-term performance. Despite the limited maintenance data, the model was able to detect the major maintenance event and the subsequent change in the operation. This outcome serves as evidence of the efficacy of the developed methodology and the indicator.

5.5.2 Temporal Analysis

Further analysis was conducted to delve deeper into the $\mathcal{RD}_{\mathcal{EM}}$ metrics proposed in this study and the nature of the autoencoder based structures. As the primary resource utilized in this study was time-series data, a temporal analysis was the initial approach taken to gain further insights into the results. This temporal analysis involved two distinct methodologies to extract more information from the data. Firstly, a decomposition analysis was carried out on the

proposed metrics to identify possible new insights. This method aimed to break down the metrics into simpler components, with the goal of revealing previously unseen patterns or trends in the data. Secondly, a wavelet transform analysis was conducted, owing to its ability to represent data across multiple scales and suitability in detecting transient behavior, changing frequencies, and slowly varying behaviors. This approach is particularly useful when investigating time-varying signals with both short-term and long-term patterns that may be difficult to capture through other methods.

The results of both, the decomposition analysis and the wavelet transform analysis reveal similar behaviors, primarily detecting a shift after major maintenance operations, as well as some other perturbations and noisy features that cannot be identified. It is worth notint that the nature of the provided data used for this analysis is non-structured data with natural language derived from customers and operation workers. Most of the data belongs to punctual fault events occurring along the IGT that could not be linked to any operational defects, but rather were malfunctioning events resulting from setup errors. Consequently, it has been difficult to process this data and no correlation can be found that matches with the perturbations or noisy features obtained from either the decomposition process or the wavelet transform.

5.5.3 Clustering Analysis

The final analysis was conducted in order to assess the performance of the $\mathcal{RD}$ metric as well as the latent manifold. It involved a clustering analysis. The primary objective of the clustering analysis was to identify the different operational points using an unsupervised

method, while also attempting to derive additional insights from this classification. By leveraging clustering techniques, the aims was to uncover any underlying patterns in the data that may not be immediately apparent through other methods.

The clustering analysis performed to evaluate the performance of the $\mathcal{RD}$ metric involved two main methods: k-means and DBSCAN. These methods differ in nature, with k-means being a centroid based clustering algorithm and DBSCAN being a density based algorithm. The k-means method aims to group data into k clusters based on their similarity, while DBSCAN creates a neighborhood of each point in a cluster based on a given radius and a minimum number of points. The analysis was performed on both the reconstruction discrepancy and the latent space time-series, with the goal of identifying different operational points using an unsupervised method and gaining further insights from the resulting classifications.

In regard to the $\mathcal{RD}_{\mathcal{EM}}$ metric, the analysis presented in Figure 4.16 on page 114 and Figure 4.17 on page 115 indicate that the distinctive operation is solely evident in the $\mathcal{RD}_{EM}^{FC3D}$ representation. The k-means method reveals several data clusters that do not correspond to any operation changes, whereas the DBSCAN representation only identifies a single density based cluster. As with the visual inspection of the $\mathcal{RD}_{EM}$ discussed earlier, the most evident shift is detected in the same metrics. Thereby, although the centroid is displaced, it is not distinct enough to be automatically separated using these clustering methods.

Figure 4.18 on page 116 and Figure 4.19 on page 117 illustrate the clusters obtained in the latent space, which exhibit more discernible patterns compared to the $\mathcal{RD}_{\mathcal{EM}}$ grouping. In this case, the k-means

method produces more precise clusters, although some false positives are still apparent. The most reliable results are observed in 3D latent representations, specifically in $\mathcal{EM}^{FC3D}$ and $\mathcal{EM}^{Sparse3D}$ models. On the other hand, the DBSCAN methods reveal a clear distinction in the sparse models, while no distinct pattern is evident in the FC models. The constraints imposed by the nature of the sparse model provide a more structured way of representing the latent space and make the shift after the major maintenance more apparent than in the Reconstruction Discrepancy metrics.

In conclusion, the clustering techniques were used to evaluate the performance of the $\mathcal{RD}$ metric in detecting the two different operations before and after maintenance. The analysis indicates that while these operations can be detected in most cases, only a few instances show that they can be automatically identified using the clustering methods. Furthermore, it is worth noting that the false positive or noise detected by the DBSCAN clustering method do not yield any significant results when compared with the available inspections and fault information. Hence, no other noteworthy insights can be drawn from these results.

Chapter 6

Conclusions and Future Work

The digitalization and the use of machine learning techniques in industrial gas turbines has enhanced the accuracy and effectiveness of algorithms, ultimately improving system performance. The field has made significant strides due to increased data availability and emerging techniques. This thesis was developed in collaboration with Siemens Energy A.G. to design and develop DA&AI tools for evaluating IGT performance using data solely from the compressor component. The data used for analysis was entirely from real plants with strategically placed sensors capturing equipment behavior. This study provides novel and custom-developed techniques for assessing IGT performance, specifically focusing on the compressor component, which is one of the main components of the IGT. By studying the best methods for this assessment, the thesis offers insights on IGT performance using real plant data.

One interesting point to note from this analysis is that the data used in the study was obtained from real plants operating under various modes for several years. This situation needed a thorough examination of the data to determine the most appropriate filters and data points to properly train each model. A comprehensive study

was conducted to establish a framework that could meet the data quality requirements for the ML method's training. Sufficient representative data points for the training process and the development of a more useful feature set based on the raw data were also critical to achieving the suggested outcomes. However, the restrictive methods applied to filter the data also resulted in the removal of some of the proposed IGTs from the initial analysis.

The main outcome of the study is the development of a discrepancy indicator that can evaluate the equipment's performance based on its operation. Two different distance measurements and two distinct calculation methods supported the findings. Using the reconstruction discrepancy indicator on a differential model of autoencoder, $\mathcal{RD}_{\mathcal{DM}}$, it clearly shows how drift appears in the IGT's operation. However, analyzing this metric, it did not provide any additional results beyond identifying a drift in the equipment's operation. Attempts were made to obtain insights using some thermodynamic indicators used in the company, but no significant outcomes were achieved. The causes of the drift may vary, but possible explanations include the degradation of the machine, maintenance activities carried out on the system, or equipment malfunction events.

A new indicator was developed by modifying the model to enhance explainability and visualization. In this case, a new type of autoencoder called sparse autoencoder was introduced, and further results were obtained using replacement events in conjunction with the fully-connected autoencoder. The primary outcome is a clear slope in the reconstruction discrepancy resulting from a significant maintenance operation. Further analysis was conducted using temporal analysis and clustering analysis to obtain additional insights. In temporal

analysis, a steeper detection was observed after significant maintenance, and some other perturbations were identified, but could not be further analyzed. Concerning clustering methods, one of the studied structures, specifically the $\mathcal{RD}_{\mathcal{EM}}^{FC3D}$, demonstrated clear outcomes using both, k-means and DBSCAN methods. However, the other structures did not show any significant results beyond centroid differentiation. The latent space analysis yielded better results, with clear insights obtained primarily from the 3D latent manifold of the sparse autoencoder.

Throughout the thesis, several gas turbine units were utilized to conduct several analyses, and the most notable results have been presented. In some cases, the outcomes may not correspond with the prior analysis due to this reason. However, it has been observed that the results are reproducible for any of the units used. Consequently, the methodology and procedures presented for curating the data, developing the models, and evaluating their performance are sufficiently robust to be implemented across a broad range of IGTs.

This thesis is a part of a broader project within the company, aimed at developing a general approach for IGT condition monitoring. It offers a solution for detecting changes in operational modes of IGTs and creates future possibilities for diagnosis systems and health management systems, with multiple exports for downstream systems to use. The tool developed in this thesis serves as a valuable aid and starting point for the company in ensuring optimal performance and longevity of IGTs. This comprehensive study can pave the way for future research by highlighting common faults, algorithms, and parameters considered in previous studies. Therefore, it simplifies identifying other vital areas for further investigation in respective studies.

6.1 Key Development Processes and Findings

This section describes the key processes that lead to the development of the whole study, as well as the key findings derived from the development section. Firstly, the key development processes are listed and then the findings are mentioned.

1. The study of the thermodynamic system of IGTs led to the development of a strategy for gaining further insights into their operation, which could aid in the development of effective condition monitoring systems.
2. The verification of the model developed in the study using curated data from NASA Turbojet Engines, which allowed for the identification of key assumptions and potential areas for improvement.
3. The use of updated data to assess the performance of the GTs in real-time enabled long-term performance monitoring, even with limited maintenance data.
4. The utilization of autoencoder models provided a powerful tool for assessing the performance of GTs by encapsulating the most important hidden patterns of data.
5. The experimentation with several model structures, utilizing both the reconstruction error and the latent manifold generated from the autoencoder models, allowed for the optimization of the models and the identification of critical features for GT performance monitoring.

From there, the following findings are derived:

1. An AI based indicator named $\mathcal{RD}_{\mathcal{DM}}$ that is able to capture the drift during the GT operation. It provides a means to monitor the equipment operation and detect any deviations from the expected performance. The $\mathcal{RD}_{\mathcal{DM}}$ model is designed to capture any changes or drifts in the system, which could be indicative of an impending failure or a change in operating conditions. This model was developed using a vast study of the thermodynamic system of GTs, and it has been shown to be effective in detecting drifts in equipment operation.

2. An AI based indicator named $\mathcal{RD}_{\mathcal{EM}}$ that is able to capture the slope after a major maintenance operation, indeed a replacement. This indicator provides a way to monitor the performance of IGTs after a major maintenance operation.

3. An automatic way to detect the maintenance operation based on clustering methods. The clustering methods were used to analyze the data and identify any patterns that might indicate a maintenance operation or malfunctioning events. Therefore, it provides a baseline to detect maintenance operations without manual intervention and the implications of it. This can be derived to reduce downtime and increase the efficiency of machine maintenance.

4. A baseline to develop a prognostic indicator for GT assessment based on compressor data. The compressor data was used to monitor the performance of the system, and the findings provide a foundation for developing a prognostic indicator that could predict the remaining useful life of the GT. This has significant implications for the maintenance of IGTs as it could help to

identify potential issues before they become critical and cause downtime.

5. A monitoring system of the GT performance based on the compressor data. The monitoring system uses autoencoder models to monitor the performance of the IGT and detect any deviations from expected performance. The use of the two models together could lead to monitor the performance in real-time and detect any anomalies or deviations from expected performance.

6.2 Future Research Lines

The framework presented in this thesis provides a comprehensive analysis of the operation of IGT and can serve as a foundation for several different future research directions.

While the proposed indicator provides valuable insights into the deviation of GT performance during operation and after a compressor replacement, it does not take into account the ratio of change. This limitation may result in different interpretations of the data. A possible avenue for future research is to develop a quantitative indicator that incorporates fault and maintenance data to gain a deeper understanding of their nature. Such an indicator could enhance the accuracy and usefulness of the proposed framework.

Another potential avenue for future research is to expand the proposed framework beyond solely using compressor data to assess the performance of the IGT. Including data from other components, such as the combustor or turbine, could provide additional information and potentially improve the accuracy of the performance assessment. Additionally, integrating data from multiple components with the

compressor data could enable a more comprehensive understanding of fault events and other anomalies in the IGT system. This could lead to the development of more robust condition monitoring and prognostic systems for IGTs.

This study proposes a baseline for monitoring and diagnosing faults and maintenance events in IGTs. By building on this work, it is possible to develop a prognostic indicator that predicts how the system will modify its operation based on its current behavior. To gain a more complete understanding of the IGT's components, it is necessary to incorporate information from other parts of the system, such as the combustion chamber and turbine. By doing so, additional insights into maintenance, faults, and unexpected events can be obtained.

The thermodynamic analysis conducted in this study provided valuable insights into the performance of the IGT. However, it was observed that the restrictive filtering methods employed may have led to the loss of some critical information. Hence, there is a potential research avenue to investigate alternative filtering methods or to adjust the thresholds of the current filtering methods to minimize the loss of relevant data. This could further enhance the accuracy and reliability of the performance analysis of the IGT.

The dataset used for model evaluation contains sensitive information that has been anonymized to protect company confidentiality. However, it is worth exploring whether there are any potentially sensitive data that could be related to the proposed indicators, such as the core-type of compressor. Further analysis within the company could shed light on this issue and further insights can be drawn from it.

Finally, the field of ML is constantly evolving and advancing at a rapid pace. As such, it would be valuable to continue exploring the potential of self-learning algorithms to develop more sophisticated models for predicting and diagnosing potential issues with IGTs. By leveraging the latest developments in ML, researchers can develop more accurate and targeted models that can identify potential issues earlier and more effectively than traditional monitoring and diagnostic techniques. This could ultimately lead to significant improvements in the reliability, efficiency, and safety of GTs across a wide range of industrial applications. Keeping up to date with the latest advancements in ML could be crucial for improving IGT performance and reducing maintenance costs.

References

[1] Jonas Schuett et al. A legal definition of ai. *arXiv preprint arXiv:1909.01095*, 2019.

[2] Alan M Turing. Computing machinery and intelligence (1950). *The Essential Turing: the Ideas That Gave Birth to the Computer Age*, pages 433–464, 2012.

[3] Stuart J Russell and Peter Norvig. Artificial intelligence a modern approach third edition, 2010.

[4] John McCarthy, Marvin L Minsky, Nathaniel Rochester, and Claude E Shannon. A proposal for the dartmouth summer research project on artificial intelligence, august 31, 1955. *AI magazine*, 27(4):12–12, 2006.

[5] Attila Benko and Cecília Sik Lányi. History of artificial intelligence. In *Encyclopedia of Information Science and Technology, Second Edition*, pages 1759–1762. IGI Global, 2009.

[6] Sumit Das, Aritra Dey, Akash Pal, and Nabamita Roy. Applications of artificial intelligence in machine learning: review and prospect. *International Journal of Computer Applications*, 115(9), 2015.

[7] Giuseppe Ciaburro and Gino Iannace. Machine learning-based algorithms to knowledge extraction from time series data: A review. *Data*, 6(6):55, 2021.

[8] Michael I Jordan and Tom M Mitchell. Machine learning: Trends, perspectives, and prospects. *Science*, 349(6245):255–260, 2015.

[9] Taeho Jo. Machine learning foundations. *Supervised, Unsupervised, and Advanced Learning. Cham: Springer International Publishing*, 2021.

[10] Pádraig Cunningham, Matthieu Cord, and Sarah Jane Delany. Supervised learning. *Machine learning techniques for multimedia: case studies on organization and retrieval*, pages 21–49, 2008.

[11] Justin Heinermann and Oliver Kramer. Machine learning ensembles for wind power prediction. *Renewable Energy*, 89:671–679, 2016.

[12] Salim Dridi. Supervised learning-a systematic literature review, 2021. https://osf.io/tysr4/.

[13] Sumana Ghosh, Shweta Shukla, and Deepti Mehrotra. Application of decision tree for understanding indian educational scenario. In *2016 International Conference on Electrical, Electronics, and Optimization Techniques (ICEEOT)*, pages 4367–4372. IEEE, 2016.

[14] M Emre Celebi and Kemal Aydin. *Unsupervised learning algorithms*, volume 9. Springer, 2016.

[15] Happiness Ugochi Dike, Yimin Zhou, Kranthi Kumar Deveerasetty, and Qingtian Wu. Unsupervised learning based on artificial neural network: A review. In *2018 IEEE International Conference on Cyborg and Bionic Systems (CBS)*, pages 322–327. IEEE, 2018.

[16] Xiaojin Jerry Zhu. Semi-supervised learning literature survey. *Computer Sciences TR*, 2005.

[17] Pavel Izmailov, Polina Kirichenko, Marc Finzi, and Andrew Gordon Wilson. Semi-supervised learning with normalizing flows. In *International Conference on Machine Learning*, pages 4615–4630. PMLR, 2020.

[18] Lars Schmarje, Monty Santarossa, Simon-Martin Schröder, and Reinhard Koch. A survey on semi-, self-and unsupervised learning for image classification. *IEEE Access*, 9:82146–82168, 2021.

[19] Xiangli Yang, Zixing Song, Irwin King, and Zenglin Xu. A survey on deep semi-supervised learning. *IEEE Transactions on Knowledge and Data Engineering*, 2022.

[20] Richard S Sutton and Andrew G Barto. *Reinforcement learning: An introduction*. MIT press, 2018.

[21] Pawel Ladosz, Lilian Weng, Minwoo Kim, and Hyondong Oh. Exploration in deep reinforcement learning: A survey. *Information Fusion*, 2022.

[22] Christopher JCH Watkins and Peter Dayan. Q-learning. *Machine learning*, 8:279–292, 1992.

[23] Gavin A Rummery and Mahesan Niranjan. *On-line Q-learning using connectionist systems*, volume 37. University of Cambridge, Department of Engineering Cambridge, UK, 1994.

[24] Vijay Konda and John Tsitsiklis. Actor-critic algorithms. *Advances in neural information processing systems*, 12, 1999.

[25] Kai Arulkumaran, Marc Peter Deisenroth, Miles Brundage, and Anil Anthony Bharath. Deep reinforcement learning: A

brief survey. *IEEE Signal Processing Magazine*, 34(6):26–38, 2017.

[26] IEA. World energy outlook 2022, 2022. https://www.iea.org/reports/world-energy-outlook-2022.

[27] Richard P Allan, Ed Hawkins, Nicolas Bellouin, and Bill Collins. Ipcc, 2021: Summary for policymakers. In V. Masson-Delmotte, P. Zhai, A. Pirani, S.L. Connors, C. Péan, S. Berger, N. Caud, Y. Chen, L. Goldfarb, M.I. Gomis, M. Huang, K. Leitzella, E. Lonnoy, J.B.R. Matthews, T.K. Maycock, T. Waterfield, O. Yelekçi, R. Yu, and B. Zhou, editors, *Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*, pages 3–32. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 2018.

[28] EIA. What is u.s. electricity generation by energy source?, 2022. https://www.eia.gov/outlooks/aeo/.

[29] Naser A Odeh and Timothy T Cockerill. Life cycle ghg assessment of fossil fuel power plants with carbon capture and storage. *Energy Policy*, 36(1):367–380, 2008.

[30] Uyioghosa Igie, Pablo Diez-Gonzalez, Antoine Giraud, and Orlando Minervino. Evaluating gas turbine performance using machine-generated data: quantifying degradation and impacts of compressor washing. *Journal of Engineering for Gas Turbines and Power*, 138(12), 2016.

[31] Sabyasachi Sahu, Dhirendranath Thatoi, and Alok Mohapatra. Analysis of techniques to improve sustainable performance of

gas-turbine based combined cycle system. Technical report, SAE Technical Paper, 2023.

[32] Shuo Wu, Yuantao Zhao, Wenge Li, Weilai Liu, Yanpeng Wu, and Fukang Liu. Research progresses on ceramic materials of thermal barrier coatings on gas turbine. *Coatings*, 11(1):79, 2021.

[33] IEA. Natural gas-fired electricity, 2022. https://www.iea.org/reports/natural-gas-fired-power.

[34] Cyrus B Meher-Homji, Andrew Bromley, et al. Gas turbine axial compressor fouling and washing. In *Proceedings of the 33rd turbomachinery symposium*. Texas A&M University. Turbomachinery Laboratories, 2004.

[35] Hiyam Farhat and Coriolano Salvini. Novel gas turbine challenges to support the clean energy transition. *Energies*, 15(15): 5474, 2022.

[36] Ozgur Balli and Hakan Caliskan. Various thermoeconomic assessments of a heat and power system with a micro gas turbine engine used for industry. *Energy Conversion and Management*, 252:114984, 2022.

[37] Yaguo Lei, Bin Yang, Xinwei Jiang, Feng Jia, Naipeng Li, and Asoke K Nandi. Applications of machine learning to machine fault diagnosis: A review and roadmap. *Mechanical Systems and Signal Processing*, 138:106587, 2020.

[38] Thyago P Carvalho, Fabrízzio AAMN Soares, Roberto Vita, Roberto da P Francisco, João P Basto, and Symone GS Alcalá. A systematic literature review of machine learning methods

applied to predictive maintenance. *Computers & Industrial Engineering*, 137:106024, 2019.

[39] Allan J Volponi. Gas turbine engine health management: past, present, and future trends. *Journal of Engineering for Gas Turbines and Power*, 136(5), 2014.

[40] Ahad Ali and Abdelhakim Abdelhadi. Condition-based monitoring and maintenance: state of the art review. *Applied Sciences*, 12(2):688, 2022.

[41] Jim Daily and Jeff Peterson. Predictive maintenance: How big data analysis can improve maintenance. *Supply Chain Integration Challenges in Commercial Aerospace: A Comprehensive Perspective on the Aviation Value Chain*, pages 267–278, 2017.

[42] Yuxin Wen, Md Fashiar Rahman, Honglun Xu, and Tzu-Liang Bill Tseng. Recent advances and trends of predictive maintenance from data-driven machine prognostics perspective. *Measurement*, 187:110276, 2022.

[43] Tony Giampaolo. *Gas turbine handbook: principles and practice*. River Publishers, 2020.

[44] Michael J Moran, Howard N Shapiro, Daisie D Boettner, and Margaret B Bailey. *Fundamentals of engineering thermodynamics*. John Wiley & Sons, 2010.

[45] Peter Jansohn. *Modern gas turbine systems: High efficiency, low emission, fuel flexible power generation*. Elsevier, 2013.

[46] Mathaios Panteli and Pierluigi Mancarella. Influence of extreme weather and climate change on the resilience of power systems: Impacts and possible mitigation strategies. *Electric Power Systems Research*, 127:259–270, 2015.

[47] Amir Samimi. Risk management in oil and gas refineries. *Prog. Chem. Biochem. Res*, 3(2):140–146, 2020.

[48] Rainer Kurz, Cyrus Meher-Homji, Klaus Brun, J Jeffrey Moore, Francisco Gonzalez, et al. Gas turbine performance and maintenance. In *Proceedings of the 42nd Turbomachinery Symposium*. Texas A&M University. Turbomachinery Laboratories, 2013.

[49] Félix C Gómez de León Hijes and José Javier Ruiz Cartagena. Maintenance strategy based on a multicriterion classification of equipments. *Reliability Engineering & System Safety*, 91(4): 444–451, 2006.

[50] Yongjun Zhao, Vitali Volovoi, Mark Waters, and Dimitri Mavris. A Sequential Approach for Gas Turbine Power Plant Preventative Maintenance Scheduling. *Journal of Engineering for Gas Turbines and Power*, 128(4):796–805, 12 2005. ISSN 0742-4795. doi: 10.1115/1.2179470. URL https://doi.org/10.1115/1.2179470.

[51] Melinda Hodkiewicz, Sarah Lukens, Michael P Brundage, and Thurston Sexton. Rethinking maintenance terminology for an industry 4.0 future. *International Journal of Prognostics and Health Management*, 12(1), 2021.

[52] Paola Cappanera, Giampaolo Manfrida, Andrea Nicoletti, Leonardo Pacini, Sergio Romagnoli, and Roberta Rossi. Digital model of a gas turbine performance prediction and preventive maintenance. *AIP Conference Proceedings*, 2191(1), 12 2019. ISSN 0094-243X. doi: 10.1063/1.5138766. URL https://doi.org/10.1063/1.5138766. 020033.

[53] Tiedo Tinga. Application of physical failure models to enable usage and load based maintenance. *Reliability Engineering & System Safety*, 95(10):1061–1075, 2010.

[54] Farshid Nasrfard, Mohammad Mohammadi, and Mohammad Rastegar. Probabilistic optimization of preventive maintenance inspection rates by considering correlations among maintenance costs, duration, and states transition probabilities. *Computers & Industrial Engineering*, 173:108619, 2022.

[55] Tiedo Tinga and Richard Loendersloot. Aligning phm, shm and cbm by understanding the physical system failure behaviour. In *European Conference on the Prognostics and Health Management Society*, 2014.

[56] Yongyi Ran, Xin Zhou, Pengfeng Lin, Yonggang Wen, and Ruilong Deng. A survey of predictive maintenance: Systems, purposes and approaches. *arXiv preprint arXiv:1912.07383*, 2019.

[57] Carlos Quiterio Gómez Muñoz, Fausto Pedro García Márquez, Borja Hernández Crespo, and Kena Makaya. Structural health monitoring for delamination detection and location in wind turbine blades employing guided waves. *Wind Energy*, 22(5): 698–711, 2019.

[58] Yongming Liu and Kai Goebel. Information fusion for national airspace system prognostics: A nasa uli project. In *10th Annual Conference of the Prognostics and Health Management Society, PHM 2018*. Prognostics and Health Management Society, 2018.

[59] Tiago Zonta, Cristiano André da Costa, Rodrigo da Rosa Righi, Miromar José de Lima, Eduardo Silveira da Trindade, and Guann Pyng Li. Predictive maintenance in

the industry 4.0: A systematic literature review. *Computers & Industrial Engineering*, page 106889, 2020.

[60] Zhengru Ren, Amrit Shankar Verma, Ye Li, Julie JE Teuwen, and Zhiyu Jiang. Offshore wind turbine operations and maintenance: A state-of-the-art review. *Renewable and Sustainable Energy Reviews*, 144:110886, 2021.

[61] Marcus Bengtsson and Gunnar Lundström. On the importance of combining "the new" with "the old"—one important prerequisite for maintenance in industry 4.0. *Procedia Manufacturing*, 25:118–125, 2018.

[62] Zeki Murat Çınar, Abubakar Abdussalam Nuhu, Qasim Zeeshan, Orhan Korhan, Mohammed Asmael, and Babak Safaei. Machine learning in predictive maintenance towards sustainable smart manufacturing in industry 4.0. *Sustainability*, 12 (19):8211, 2020.

[63] Advanced Factories Expo & Congress. Industry 4.0 congress, 2021. https://www.advancedfactories.com/app/uploads/sites/3/2021/07/memoria_af2021.pdf.

[64] Gerardo Boto Varela, Laia Vicens Serra, et al. Vi conference of pre-doctoral researchers abstract book: Volume vi, 2022.

[65] Narahari Rath, RK Mishra, and Abhijit Kushari. Aero engine health monitoring, diagnostics and prognostics for condition-based maintenance: An overview. *International Journal of Turbo & Jet-Engines*, 2022.

[66] Giovanni Bechini. *Performance diagnostics and measurement selection for on-line monitoring of gas turbine engines.* PhD

thesis, School of Engineering. Cranfield University, Cranfield MK43 0AL, UK, 2007.

[67] Fei Zhao, Liang Chen, Tangbin Xia, Zikun Ye, and Yu Zheng. Gas turbine exhaust system health management based on recurrent neural networks. *Procedia CIRP*, 83:630–635, 2019.

[68] Akram Mubarak, Mebrahitom Asmelash, Azmir Azhari, Ftwi Yohannes Haggos, and Freselam Mulubrhan. Machine health management system using moving average feature with bidirectional long-short term memory. *Journal of Computing and Information Science in Engineering*, 23(3):031002, 2023.

[69] Hossein Shahabadi Farahani, Alireza Fatehi, and Mahdi Aliyari Shoorehdeli. On the Application of Domain Adversarial Neural Network to Fault Detection and Isolation in Power Plants. *Proceedings - 19th IEEE International Conference on Machine Learning and Applications, ICMLA 2020*, pages 1132–1138, 2020. doi: 10.1109/ICMLA51294.2020.00182.

[70] Zhezhe Han, Md Moinul Hossain, Yuwei Wang, Jian Li, and Chuanlong Xu. Combustion stability monitoring through flame imaging and stacked sparse autoencoder based deep neural network. *Applied Energy*, 259:114159, 2020.

[71] Tryambak Gangopadhyay, Vikram Ramanan, Adedotun Akintayo, Paige K Boor, Soumalya Sarkar, Satyanarayanan R. Chakravarthy, and Soumik Sarkar. 3D convolutional selective autoencoder for instability detection in combustion systems. *Energy and AI*, 4:100067, 2021. ISSN 26665468. doi: 10.1016/j.egyai.2021.100067. URL https://doi.org/10.1016/j.egyai.2021.100067.

[72] Yanyan Shen and Khashayar Khorasani. Hybrid multi-mode machine learning-based fault diagnosis strategies with application to aircraft gas turbine engines. *Neural Networks*, 130: 126–142, 2020.

[73] Yang Hu, Xuewen Miao, Yong Si, Ershun Pan, and Enrico Zio. Prognostics and health management: A review from the perspectives of design, development and decision. *Reliability Engineering & System Safety*, 217:108063, 2022.

[74] Miryam Elizabeth Villa-Pérez, Miguel A Alvarez-Carmona, Octavio Loyola-Gonzalez, Miguel Angel Medina-Pérez, Juan Carlos Velazco-Rossell, and Kim-Kwang Raymond Choo. Semi-supervised anomaly detection algorithms: A comparative summary and future research directions. *Knowledge-Based Systems*, 218:106878, 2021.

[75] Homam Nikpey Somehsaraei, Susmita Ghosh, Sayantan Maity, Payel Pramanik, Sudipta De, and Mohsen Assadi. Automated data filtering approach for ANN modeling of distributed energy systems: Exploring the application of machine learning. *Energies*, 13(14):1–15, 2020. ISSN 19961073. doi: 10.3390/en13143750.

[76] Geunbae Lee, Myungkyo Jung, Myoungwoo Song, and Jaegul Choo. Unsupervised anomaly detection of the gas turbine operation via convolutional auto-encoder. *Proceedings of the Annual Conference of the Prognostics and Health Management Society, PHM*, 2020-June, 2020. ISSN 23250178. doi: 10.1109/ICPHM49022.2020.9187054.

[77] Song Fu, Shisheng Zhong, Lin Lin, and Minghang Zhao. A re-optimized deep auto-encoder for gas turbine unsupervised

anomaly detection. *Engineering Applications of Artificial Intelligence*, 101(May 2020):104199, 2021. ISSN 09521976. doi: 10.1016/j.engappai.2021.104199. URL https://doi.org/10.1016/j.engappai.2021.104199.

[78] Zhouzheng Li, Kun Feng, and Binbin Yan. Dynamic gas turbine condition monitoring scheme with multi-part neural network. In *Turbo Expo: Power for Land, Sea, and Air*, volume 84140, page V005T05A007. American Society of Mechanical Engineers, 2020.

[79] CKM Lee, Yi Cao, and Kam Hung Ng. Big data analytics for predictive maintenance strategies. In *Supply Chain Management in the Big Data Era*, pages 50–74. IGI Global, 2017.

[80] Zhe Li, Yi Wang, and Ke-Sheng Wang. Intelligent predictive maintenance for fault diagnosis and prognosis in machine centers: Industry 4.0 scenario. *Advances in Manufacturing*, 5(4): 377–387, 2017.

[81] Ph Kamboukos and K Mathioudakis. Comparison of linear and nonlinear gas turbine performance diagnostics. *J. Eng. Gas Turbines Power*, 127(1):49–56, 2005.

[82] Louis A Urban. *Gas turbine engine parameter interrelationships*. Hamilton Standard Division of United Aircraft Corporation, 1969.

[83] David L Doel. Temper: A gas-path analysis tool for commercial jet engines. In *Turbo Expo: Power for Land, Sea, and Air*, volume 78972, page V005T15A013. American Society of Mechanical Engineers, 1992.

[84] PC Escher and R Singh. An object-oriented diagnostics computer program suitable for industrial gas turbines. In *21st (CIMAC) international congress of combustion engines, Switzerland*, pages 15–18, 1995.

[85] D. L. Doel. TEMPER—A Gas-Path Analysis Tool for Commercial Jet Engines. *Journal of Engineering for Gas Turbines and Power*, 116(1):82–89, 01 1994. ISSN 0742-4795. doi: 10.1115/1.2906813. URL https://doi.org/10.1115/1.2906813.

[86] MJ Barweli. Compass-ground based engine monitoring program for general application. Technical report, SAE technical paper, 1987.

[87] MJ Provost, CH Sieverding, and K Mathioudakis. Kalman filtering applied to gas turbine analysis. *Von Karman Institute Lecture Series: Gas Turbine Condition Monitoring and Fault Diagnosis,(2003-01)*, 2003.

[88] Takahisa Kobayashi and Donald L Simon. Application of a bank of kalman filters for aircraft engine fault diagnostics. In *Turbo Expo: Power for Land, Sea, and Air*, volume 36843, pages 461–470, 2003.

[89] Xiaofeng Liu, Jiaqi Zhu, Chenshuang Luo, Liuqi Xiong, and Qiang Pan. Aero-engine health degradation estimation based on an underdetermined extended kalman filter and convergence proof. *ISA transactions*, 125:528–538, 2022.

[90] Liping Yan, Hualiang Zhang, Xuezhi Dong, Qiao Zhou, Haisheng Chen, and Chunqing Tan. Unscented kalman-filter-based simultaneous diagnostic scheme for gas-turbine gas path

and sensor faults. *Measurement Science and Technology*, 32(9): 095905, 2021.

[91] A Stamatis, K Mathioudakis, M Smith, and K Papailiou. Gas turbine component fault identification by means of adaptive performance modeling. In *Turbo Expo: Power for Land, Sea, and Air*, volume 79085, page V005T15A015. American Society of Mechanical Engineers, 1990.

[92] Changduk Kong, Myoungcheol Kang, and Gwanglim Park. Study on condition monitoring of 2-spool turbofan engine using non-linear gas path analysis method and genetic algorithms. *International Journal of Materials, Mechanics and Manufacturing*, 1(2):214–220, 2013.

[93] TV Breikin, GG Kulikov, VY Arkov, and PJ Fleming. Dynamic modelling for condition monitoring of gas turbines: Genetic algorithms approach. *IFAC Proceedings Volumes*, 38(1): 739–742, 2005.

[94] Suresh Sampath, Ankush Gulati, and Riti Singh. Fault diagnostics using genetic algorithm for advanced cycle gas turbine. In *Turbo Expo: Power for Land, Sea, and Air*, volume 3607, pages 19–27, 2002.

[95] Sanjay G Barad, PV Ramaiah, RK Giridhar, and G Krishnaiah. Neural network approach for a combined performance and mechanical health monitoring of a gas turbine engine. *Mechanical Systems and Signal Processing*, 27:729–742, 2012.

[96] Omar Mohamed and Ashraf Khalil. Progress in modeling and control of gas turbine power generation systems: a survey. *Energies*, 13(9):2358, 2020.

[97] Jinwei Chen, Zhenchao Hu, Jinzhi Lu, Xiaochen Zheng, Huisheng Zhang, and Dimitris Kiritsis. A data-knowledge hybrid driven method for gas turbine gas path diagnosis. *Applied Sciences*, 12(12):5961, 2022.

[98] Ying Peng, Ming Dong, and Ming Jian Zuo. Current status of machine prognostics in condition-based maintenance: a review. *The International Journal of Advanced Manufacturing Technology*, 50(1-4):297–313, 2010.

[99] Dan Simon and Donald L Simon. Aircraft turbofan engine health estimation using constrained kalman filtering. *J. Eng. Gas Turbines Power*, 127(2):323–328, 2005.

[100] Najmeh Daroogheh, Nader Meskin, and Khashayar Khorasani. A novel particle filter parameter prediction scheme for failure prognosis. In *2014 American Control Conference*, pages 1735–1742. IEEE, 2014.

[101] Holger Lipowsky, Stephan Staudacher, Michael Bauer, and Klaus-Juergen Schmidt. Application of bayesian forecasting to change detection and prognosis of gas turbine performance. *Journal of engineering for gas turbines and power*, 132(3), 2010.

[102] Nicola Puggina and Mauro Venturini. Development of a statistical methodology for gas turbine prognostics. *Journal of engineering for gas turbines and power*, 134(2), 2012.

[103] Yanrong Li, Shizhe Peng, Yanting Li, and Wei Jiang. A review of condition-based maintenance: Its prognostic and operational aspects. *Frontiers of Engineering Management*, 7(3):323–334, 2020.

[104] Lalith Madhav Peddareddygari and Douglas L Allaire. Time to failure prognosis of a gas turbine engine using predictive analytics. In *AIAA Scitech 2021 Forum*, page 1355, 2021.

[105] Yiyang Liu and Xiaomo Jiang. Towards predictive maintenance of a heavy-duty gas turbine a new hybrid intelligent methodology for performance simulation. In *Annual Conference of the PHM Society*, volume 14(1), 2022.

[106] Mojtaba Kordestani, M Foad Samadi, and Mehrdad Saif. A new hybrid fault prognosis method for mfs systems based on distributed neural networks and recursive bayesian algorithm. *IEEE Systems Journal*, 14(4):5407–5416, 2020.

[107] Oussama Meski, Farouk Belkadi, Florent Laroche, Asma Ladj, and Benoit Furet. Integrated data and knowledge management as key factor for industry 4.0. *IEEE Engineering Management Review*, 47(4):94–100, 2019.

[108] Abhinav Saxena and Kai Goebel. Turbofan engine degradation simulation data set. *NASA Ames Prognostics Data Repository*, 18, 2008.

[109] Feng Jia, Yaguo Lei, Jing Lin, Xin Zhou, and Na Lu. Deep neural networks: A promising tool for fault characteristic mining and intelligent diagnosis of rotating machinery with massive data. *Mechanical Systems and Signal Processing*, 72:303–315, 2016.

[110] Peter J Brockwell and Richard A Davis. *Introduction to time series and forecasting*. Springer, 2002.

[111] Dejan Dragan, Tomaz Kramberger, and Marko Intihar. A comparison of methods for forecasting the container throughput in north adriatic ports. In *Conference IAME*, 2014.

[112] Manel Rhif, Ali Ben Abbes, Imed Riadh Farah, Beatriz Martínez, and Yanfang Sang. Wavelet transform application for/in non-stationary time-series analysis: a review. *Applied Sciences*, 9(7):1345, 2019.

[113] Christopher E Heil and David F Walnut. Continuous and discrete wavelet transforms. *SIAM review*, 31(4):628–666, 1989.

[114] Frédéric Chazal and Bertrand Michel. An introduction to topological data analysis: fundamental and practical aspects for data scientists. *Frontiers in artificial intelligence*, 4:667963, 2021.

[115] Laurens Van der Maaten and Geoffrey Hinton. Visualizing data using t-sne. *Journal of machine learning research*, 9(11), 2008.

[116] Leland McInnes, John Healy, and James Melville. Umap: Uniform manifold approximation and projection for dimension reduction. *arXiv preprint arXiv:1802.03426*, 2018.

[117] John A Hartigan and Manchek A Wong. Algorithm as 136: A k-means clustering algorithm. *Journal of the royal statistical society. series c (applied statistics)*, 28(1):100–108, 1979.

[118] Martin Ester, Hans-Peter Kriegel, Jörg Sander, and Xiaowei Xu. A density-based algorithm for discovering clusters in large spatial databases with noise. In *Proceedings of the Second International Conference on Knowledge Discovery and Data Mining*, KDD'96, page 226–231. AAAI Press, 1996.

[119] Zijian Sun, Jian Tang, Junfei Qiao, and Chengyu Cui. Review of concept drift detection method for industrial process modeling. In *2020 39th Chinese Control Conference (CCC)*, pages 5754–5759. IEEE, 2020.

[120] Dor Bank, Noam Koenigstein, and Raja Giryes. Autoencoders. *arXiv preprint arXiv:2003.05991*, 2020.

Appendix A

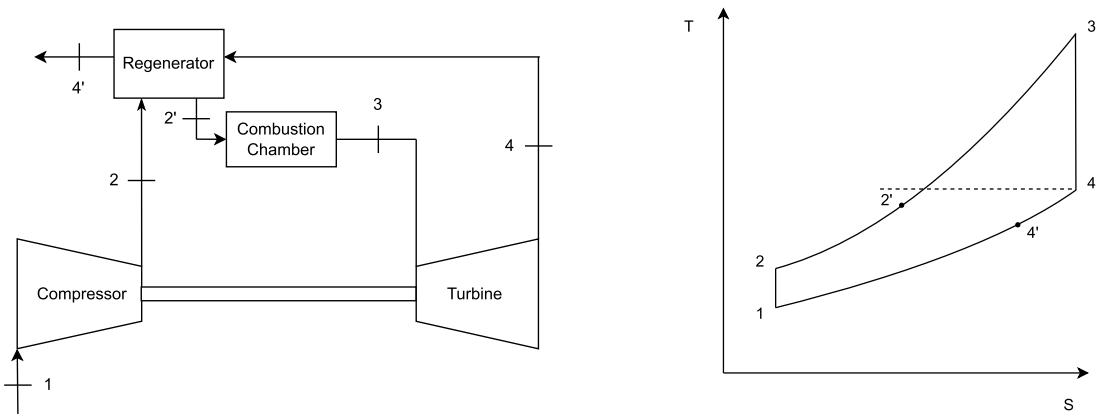
Gas Turbine Configurations

GT equipment can adopt several configurations according to the different performance of the brayton cycle in order to improve its operation. Next, the most common configurations based on the brayton cycle enhancements are presented.

A.1 Regenerative Gas Turbines

Regenerative Gas Turbines utilize the high-temperature exhaust gas of the turbine to elevate the temperature of the compressed air that enters the turbine. Typically, the exhaust gas from the turbine is hotter than the surrounding environment. To optimize this thermodynamic resource, known as exergy, a regenerator is situated before the air enters the combustor. This technique reduces the amount of energy, i.e. fuel, required to achieve the desired turbine inlet temperature. The Brayton cycle, incorporating the regenerator under ideal conditions, is thereby modified in the following manner:

Figure A.1 presents a representation of the GT configuration as well as the modified brayton cycle. The exhaust turbine temperature is



(a) Gas turbine Configuration (b) Regenerative Brayton cycle

Figure A.1: Regenerative Gas Turbine Representation

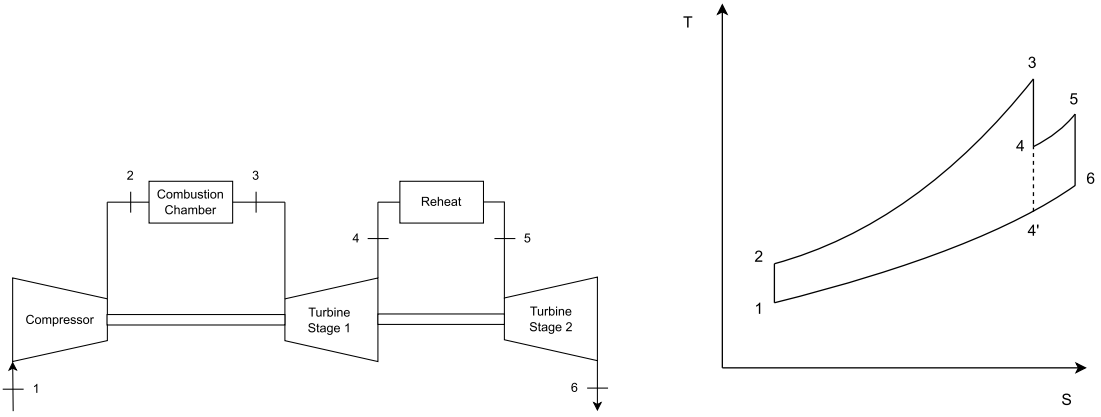
cooled from state 4 to state 4', while the compressor outlet temperature is increased from 2 to 2'. Hence, the heat added per unit of mass is given by

$$\frac{\dot{Q}_{in}}{\dot{m}} = h_3 - h_{2'} \tag{A.1}$$

In this configuration, the net work developed per unit of mass flow is not altered by the addition of a regenerator. Thus, thermal efficiency increases because of the heat need to increase the temperature to the working point is reduced.

A.2 Gas Turbines with Reheat

Controlling the temperature during combustion is critical to prevent any detrimental effects on the constituent materials. One approach to regulate the temperature involves supplying an excess amount of air to facilitate the combustion of fuel. As a result, the turbine's exhaust gases contain a sufficient amount of air to sustain an additional combustion stage, which is the opportune moment for the



(a) Gas turbine with Reheat Configuration

(b) Brayton cycle with Reheat

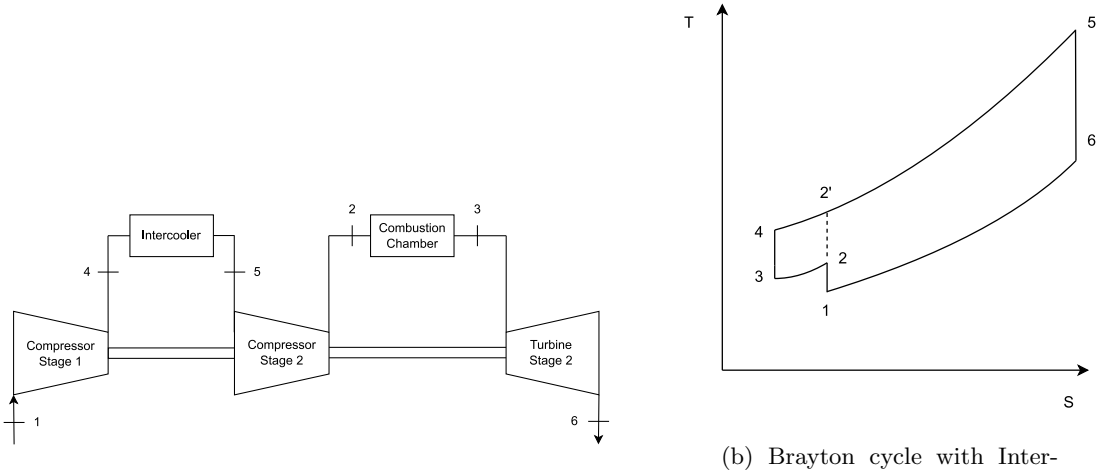
Figure A.2: Gas Turbine with Reheat Representation

reheat process to occur. By leveraging this available air, a multi-stage turbine process can be executed. Figure A.2 depicts the GT configuration along with the modified Brayton cycle.

As it is illustrated, the reheat cycle exhibits a larger area beneath its curve, resulting in a greater net specific work output. This enhancement, however, does not necessarily lead to a corresponding improvement in thermal efficiency, since an additional heat energy is required for the second combustion process. Nonetheless, the temperature at the outlet of the second-stage turbine, state 6, exceeds that of the corresponding state, state 4', in the cycle that lacks reheat. Consequently, the utilization of reheat amplifies the potential for regeneration. When both reheat and regeneration are implemented together, the thermal efficiency can increase considerably.

A.3 Gas Turbines with Intercooling

The intercooler aims to reduce the temperature in between the compression work. The goal behind this configuration is to reduce the



(a) Gas turbine with Intercooler Configuration

(b) Brayton cycle with Intercooler

Figure A.3: Gas Turbine with Intercooler Representation

work done by the compressor in order to increase the net work. Figure A.3 show the configuration of the GT and the specific Brayton cycle.

In order to make it more understandable, Figure A.4 illustrates the work done by the compressor using the pressure - volume diagram.

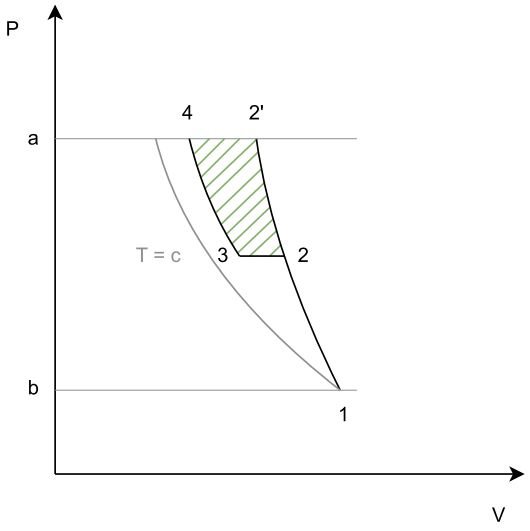


Figure A.4: P - V diagram of the Brayton Cycle with Intercooler

Without intercooling, the gas would be compressed isentropically in a single stage from state 1 to state 2'. This work is represented by states 1–2'–5–6 in T-s diagram and by the area generated amongst

1-2'-a-b in the P-V diagram. The same work done with intercooling it represented by states 1-2-3-4-5-6 in T-s diagram and by the area generated between 1-2-3-4-a-a in the P-V diagram. The work reduction effect of the cooling is showed in the P-V diagram with the crosshatched area.

The implementation of multi-stage compression with intercooling in a GT power plant elevates the net work output by diminishing the compression work. Nevertheless, compression with intercooling does not inevitably heighten the thermal efficiency of an equipment due to the corresponding reduction in the temperature of the air entering the combustor. Consequently, additional heat transfer would be necessary to attain the desired turbine inlet temperature. The lower temperature at the compressor exit, however, augments the potential for regeneration. Hence, when intercooling is employed together with regeneration, a significant improvement in thermal efficiency can be achieved.

A.4 Combined Cycle

One of the most prevalent approaches for using the exhaust heat in the regenerative cycle involves a CC power plant. In a CC power plant, two power cycles are integrated in such a way that the heat generated by heat transfer from one cycle, the Brayton cycle, is employed partially or entirely as the heat input for the other cycle, the Rankine cycle. The interconnection of the cycles is facilitated by a heat-recovery steam generator that acts as the boiler for the steam power cycle. The CC harnesses the high average temperature of heat

addition of the GT and the low average temperature of heat rejection of the steam power cycle. Thus, the thermal efficiency surpasses that of each cycle operating individually.

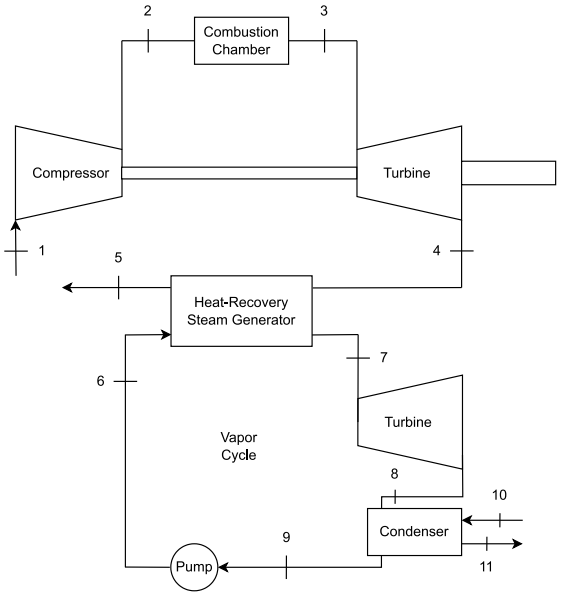


Figure A.5: Combined cycle representation

Figure A.5 shows an schematic of GT-vapor power plant. According to the thermodynamics, the efficiency of the power plant is expressed as:

$$\eta_{CC} = \frac{W_{GT} + W_{ST}}{Q_{in}} \tag{A.2}$$

Where W_{GT} is the net power generated by the machine, W_{ST} is the net power generated by the steam (vapor) turbine, and Q_{in} is the total rate of heat transfer to the CC.

These power plants are capable of achieving around 60% of thermal efficiency, thus being an interesting option for the power plant owner. Additionally, they offer the ability to generate further net output power while significantly saving fuel, reducing carbon dioxide emissions, and complying with low nitric oxide standards.