NOVA IMS
INFORMATION MANAGEMENT SCHOOL

Master's Degree Program in
Statistics and Information Management

Specialization in
Data Analytics

STUDY OF PATTERNS IN AIRCRAFT AIRFRAME MRO USING A DATA DRIVEN APPROACH.

Maria Helena Abreu Luz

NOVA Information Management School
Instituto Superior de Estatística e Gestão de Informação
Universidade NOVA de Lisboa

NOVA Information Management School
Instituto Superior de Estatística e Gestão de Informação
Universidade NOVA de Lisboa

STUDY OF PATTERNS IN AIRCRAFT AIRFRAME MRO USING A DATA DRIVEN APPROACH

by

Maria Helena Abreu Luz

Dissertation presented as partial requirement for obtaining the Master's degree in Statistics and Information Management, with a specialization in Data Analytics

Supervised by:

Márcia Lourenço Baptista, NOVA Information Management School
Bruno Damásio, NOVA Information Management School

July, 2025

STATEMENT OF INTEGRITY

I, Helena Luz, hereby declare having conducted this academic work with integrity. I confirm that I have not used plagiarism or any form of undue use of information or falsification of results along the process leading to its elaboration. I further declare that I have fully acknowledge the Rules of Conduct and Code of Honor from the NOVA Information Management School.

Lisbon, 2025/07/15
Helena Luz

Study of patterns in aircraft airframe MRO using a data driven approach

This document was created with the (pdf/Xe/Lua)LaTeX processor and the NOVAthesis template (v7.2.1) (Lourenço, 2021).

Abstract

Air travel is widely recognized as one of the safest and most convenient modes of transport worldwide. The aircraft maintenance industry is an important field of operation, therefore it is paramount to find an optimal approach to estimate aircraft tasks durations. This study aims to examine the available literature on existing methods to process data and to estimate duration in aviation. The research comprises an exploratory analysis, with the goal of understanding the real-world maintenance data set and finding relevant hidden insights. The analysis encountered patterns in task efficiency by skill type, aircraft type, and location. The second part of the research conducts an experimental analysis that compares four predictive machine learning models with traditional method PERT (Program Evaluating and Review Technique), a project management technique used to estimate the duration of tasks by considering optimistic, pessimistic, and most likely time estimates. The results of the study demonstrate the value of data-driven approaches in improving accuracy in maintenance task planning and performance.

Keywords: Maintenance, Repair, and Overhaul (MRO), Data Analytics, Machine Learning, Task Duration Estimation, Aircraft Maintenance, Data-Driven Decision Making

Sustainable Development Goals (SDG):

Resumo

O transporte aéreo é reconhecido como um dos meios de transporte mais seguros e convenientes do mundo. A indústria de manutenção de aeronaves é um importante campo de operação, portanto, é fundamental encontrar uma abordagem ideal para estimar a duração das tarefas de operação nas aeronaves. Este estudo visa examinar a literatura disponível sobre métodos existentes para processar dados e estimar os tempos de duração das tarefas. Este estudo compreende uma análise exploratória, com o objetivo de compreender o conjunto de dados de manutenção do mundo real e encontrar percepções ocultas relevantes. A análise encontrou padrões na eficiência das tarefas por tipo de aeronave, de habilidade e localização. A segunda parte da pesquisa conduz uma análise experimental que compara quatro modelos preditivos de aprendizagem automática com o método tradicional PERT, uma técnica de gestão de projetos usada para estimar a duração das tarefas considerando estimativas de tempo otimistas, pessimistas e mais prováveis. Os resultados do estudo demonstram o valor das abordagens baseadas em dados para melhorar a precisão no planeamento e desempenho de tarefas de manutenção.

Palavras-chave: Manutenção, Reparação e Revisão (MRO), Manutenção de Aeronaves, Análise de Dados, Aprendizagem Automática, Estimação da Duração de Tarefas

Contents

List of Figures

1

Introduction

Air travel is widely recognized as one of the safest and most convenient modes of transport worldwide. The level of safety is attributed to the rigorous maintenance procedures and strict regulatory standards, commonly known as MRO - Aircraft Maintenance, Repair, and Overhaul, of the aviation industry. To ensure safety and the best performance of the airplane, the MRO staff must manage various challenges, such as equipment failures, team pressure, and strict time constraints.

The motivation for this study comes from the need to enhance MRO operational efficiency by improving the maintenance duration prediction and the identification of hidden patterns. The present maintenance planning usually relies on traditional estimation methods that may not capture the complexity and variability in aircraft maintenance operations (Chandola et al., 2023) . This gap between predicted and actual maintenance durations can lead to an inefficient use of resources, and increased operational costs.

This study will address three fundamental research questions:

1. Can machine learning algorithms provide more accurate maintenance duration predictions compared to traditional methods? This question examines whether modern data-driven approaches can outperform traditional estimation techniques in forecasting maintenance duration.

2. What patterns exist in maintenance duration times? This question aims to uncover trends and correlations within maintenance data that may not be visible through traditional analysis methods.

3. Which maintenance activities are being underestimated or overestimated? This analysis aims to identify specific maintenance tasks where the predicted time deviates from the actual performance.

In the literature review section, it highlights the evolving landscape of MRO, underlining the importance of systematic maintenance practices, the role of data analytics, and the application of statistical methods to enhance operational efficiency in the aviation industry. The methodology section provides an overview of the methods used in the study, beginning with the exploratory data analysis and progressing to the experimental procedures. The results and discussion chapter point out the final results obtained from these analysis, interpreting the findings, and how they contribute to answering the research questions.

2

Literature Review

2.1 Introduction

The aircraft maintenance business continues to undergo constant development, therefore it is natural to look for opportunities that help achieve the best performance. Depending on the type of aircraft, the tasks tend to vary by complexity and durations, which has became a challenge to calculate a proper estimation of the activity duration.

This review aims to examine the available literature on existing methods for predicting the precise duration of a task performed by MRO airlines.

The review is divided into two parts. The first part explores the concept of maintenance and the understanding of all terms related to the time and duration of activities. This will allow a better comprehension of the aspects necessary to accurately predict the task duration of the MRO aircraft. The second part focuses on methods and techniques for estimating time duration on real world data provided by a Hangar.

This section is structured into four subsections:

- **Aircraft Maintenance**: Presents the scope of the Maintenance, Repair and Overhaul (MRO) in aircraft, as well as the concepts necessary to contextualize about the reality of this world.

- **Human Factor**: Explores the interrelationship between the components of the system and human beings, centering on the human error and possible psychology causes, like motivation and peer pressure.

- **Workload Management**: Highlights the importance of task planning, communication, and teamwork to address unexpected issues during base maintenance.

- **Estimation of Task Duration**: Examines the three methods that will be discussed on the study. It compares the traditional methods, such as PERT and Delphi with machine learning methods, such as Artificial Neural Network.

  It also discusses the literature suggestions for data analysis in the MRO.

Finally, it briefly describes and discusses some of the studies on aircraft maintenance, as well as how data-driven methods and artificial neural networks can be applied in the MRO airline.

2.1.1 Background of the Hangar

The corporation in charge of providing the data was established in 2008. It is a privately owned Maintenance, Repair, and Overhaul (MRO) organization headquartered in Germany. It holds EASA Part 145 approval, authorizing it to perform maintenance on a variety of aircraft types.

The company has a wide range of services, including base maintenance, line maintenance, and tool loan services. The hangar can accommodate up to ten simultaneous lines of base maintenance for both narrow and wide-body aircraft. It has custom solutions to meet the unique requirements of its clients, with the aim of providing personalized services that align with customer expectations and budget constraints. Therefore, it has been recognized from around the world and reflects its commitment to maintaining high standards in aircraft maintenance services.

2.2 Aircraft Maintenance

2.2.1 Understanding the Maintenance, Repair and Overhaul (MRO)

According to Norma NP EN (13306:2021) (Instituto Português da Qualidade, 2021), maintenance is described as a "combination of all technical, administrative, and managerial actions during the life cycle of an item intended to retain it in, or restore it to, a state in which it can perform the required function", so maintenance is the process of preserving the state of an item ensuring that the item performs its function as best as possible.

In the same document, repair is defined as a "physical action taken to restore the required function of a faulty item" and overhaul as "comprehensive set of preventive maintenance actions carried out, in order to maintain the required level of performance of an item". These last two actions are maintenance activities, in other words, there is no maintenance without the combination of these technical actions.

Furthermore, combining all these factors creates a broad term that includes measures needed to ensure the full performance of equipment, also known as Maintenance, Repair and Overhaul. It is the joint of these technical actions that provides the necessary maintenance. For example, with the technological advances of the aviation industry, it became a complex and reliable engine, and maintenance became an important work that requires rigorous maintenance at several levels.

In the maintenance area, there are tasks. These tasks are the most distinct and range from routine inspection duties such as checking lubrication or oil level to more complex tasks such as installation or removal of an item. Depending on complexity, these can take months.

Based on the review of Pelt, Apostolidis, et al. (2019) the maintenance tasks can be divided into three categories:

- **Corrective maintenance tasks**: These are unscheduled maintenance tasks in response to casual and unpredictable failures.
- **Alternative maintenance actions**: It involves removing a design fault or upgrading the functionality of a component.
- **Preventive maintenance tasks**: In this category, most tasks are incorporated to avoid failures resulting from progressive degradation. It can be divided into two sub-categories:
  - Conditioned-based maintenance, where the components conditions determined the maintenance interval
  - Time-based maintenance which the lifetime of the item is based on calculations and experience.

2.2.2 Time-based Preventive Maintenance

The implementation of a sensor monitoring system allowed the aircraft industry to move from a time-based maintenance to a condition-based maintenance (Rengasamy et al., 2018). The main difference between these two maintenances is that conditioned-based maintenance depends on components conditions to determine the maintenance interval, while time-based maintenance uses calculation and experience to predict the lifetime of the item.

This work will focus on preventive maintenance tasks specifically time-based maintenance, which is a set of schedule maintenance activities that are performed at predetermined intervals, regardless of the aircraft's usage. As a result, it is easier to plan and manage maintenance activities. Unlike condition-based maintenance, which depends on the current state of the equipment to determine maintenance needs.

Many models and studies on time-based preventive maintenance assume that the life-time of the item is known. This assumption simplifies the calculations for the planning process, but cannot reflect the scenarios of the real world (De Jonge et al., 2015). In practice, the lifetime distribution is uncertainty due to numerous reasons, such as incomplete data or incorrectly recorded failure codes.

According to Chandola et al. (2023), this approach is essential for maintaining safety and compliance with aviation regulations, as it helps prevent potential failures before they occur. The aviation industry has established rigorous maintenance schedules to ensure aircraft reliability and safety throughout their operational life.

The figure 2.1 is a simple illustration of aircraft maintenance, as it is possible to observe aircraft maintenance breaks into two parts: the unscheduled maintenance that includes correction maintenance tasks and the schedule maintenance tasks where are the preventive maintenance tasks and time-based maintenance. In the latter category, there are two major maintenance checks that can be divided into line and base maintenance.

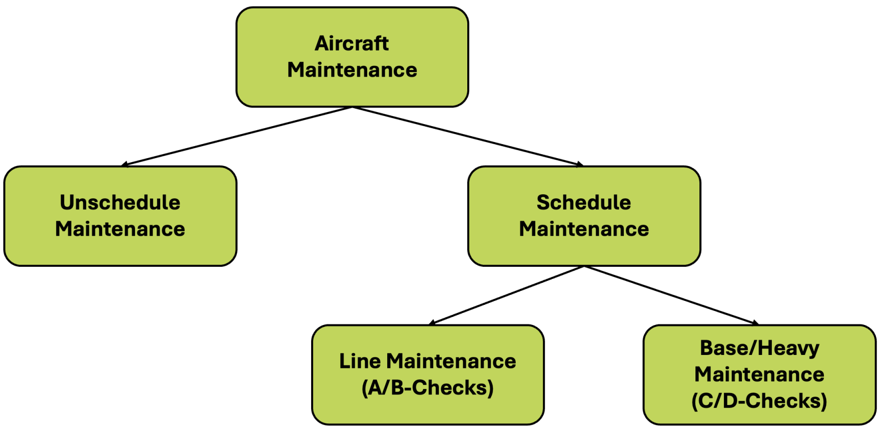


Figure 2.1: Aircraft Maintenance

2.2.3 Aircraft Maintenance Checks

The aircraft business usually requires various maintenance checks, which incorporate transit/ramp-checks, the A-checks and B-checks also known as line maintenance, which are lighter and simpler, and the C-checks and D-checks also call as heavy maintenance, which are more complex and extensive (Vencel' et al., 2022). The table 2.1 summarizes the maintenance checks of the aircraft industry.

Maintenance	Line Maintenance	A-Checks	It is the first major inspection, includes all major systems inspections, and is carried out every 8 to 10 weeks. It takes between 6 to 24 hours on a narrowbody aircraft.
		B-Checks	It is incorporated into consecutive A-checks. It includes detailed visual inspections and lubrication of all moving parts, and carries out every 6 to 8 months.
	Base Maintenance	C-Checks	It is more extensive than check A and B. The technicians perform specific tasks such as in-depth lubrication of the mounting and cable and examination of structures. It is carried out every 18 months to 2 years and can take more than 6000 maintenance hours.
		D-Checks	This is a special one, it is carried out every four years and requires the aircraft to be out of service for up to one month at a time. Once it is done, the aircraft looks like a new one, however, if the aircraft is inspected with two or three D checks, it can be considered a time for selling the aircraft or withdrawing it.

Table 2.1: Description of the aircraft maintenance checks

2.3 Human Factors in Maintenance

2.3.1 Human Factor

The human factor in maintenance is the interaction between human beings and other components of the system, which has a significant impact on safety and performance. Sheikhalishahi et al. (2016) quote that the role of humans in various stages of the product life cycle, including design, installation, production and maintenance, is considerable and cannot be fully replaced by the latest technologies and technology systems.

The paper (Sheikhalishahi et al., 2016) classifies human factors into three main areas:

- **Human Error/Reliability calculation** - This area focuses on understanding and quantifying human errors that may occur during maintenance activities. It involves developing models and methods for assessing the reliability of maintenance operations taking into account human factors. Factors such as time pressure, inadequate training, and reluctance to change can contribute to human errors.

- **Workplace/Organizations condition** - This section focuses on designing working environments that improve human performance and reduce the likelihood of errors. It considers the physical and cognitive aspects of the workplace, ensuring that the environment is conducive to efficient and safe maintenance practices.

- **Human resource management** - This includes effective staff management, including training, scheduling and crew resource management, to ensure that human factors are properly dealt with in maintenance operations. A culture that promotes open communication and incident reporting is essential for improving maintenance practices.

Human experts play an important role in determining the duration of tasks. However, their availability may be limited and they may find it difficult to make accurate estimates when faced with many activities (Li et al., 2023).

2.4 Workload Management and Efficacy

2.4.1 Maintenance Task Management

Base maintenance requires multiple tasks to be performed. As mentioned above, base maintenance tasks are carried out from time to time or after a certain amount of aircraft use. Also, it is massive and, depending on the type and size of aircraft, involves a larger number of inspections and maintenance procedures. Consequently, many resources, such as materials, tools, human labor, work spaces, and actions, are required.

An appropriate maintenance work management lies in planning, communication, and teamwork. Chandola et al. (2023) reported that each hangar has its own authorized personnel, technicians, engineers and operators who are responsible for planning and discussing how a task will be performed.

During the execution of the task, it can be discovered errors or damage items, which will increase not only the number of tasks, but also the estimated time of them. For that reason, the activity must be replanned, and the required new materials need to be ordered to address the issue, which leads to a higher performed time.

2.4.2 Time Management and time estimation

Francis-Smythe and Robertson (1999) stated that those who perceived themselves as good time managers are more likely to make a true prediction of the duration of future tasks. Time management and time estimation are different, although they are slightly correlated. Time management is about organizing and planning how to allocate time throughout activities to improve efficiency and effectiveness. It involves setting goals, prioritizing tasks, and scheduling to ensure optimal use of time. Effective time management helps individuals and teams achieve goals and meet deadlines. On the other hand, time estimation is the process of predicting the duration required to complete a specific task or project. This includes monitoring the effort and resources needed to perform a task within a scheduled timeframe. An accurate time estimate is important to set realistic deadlines and arrange workloads.

The relationship between time management and time estimation is significant. A precise time estimation informs effective time management by providing realistic timelines to complete the task. This allows for better schedules and priorities. In contrast, strong time management skills can improve the accuracy of time estimation by offering insights into the productivity levels of individuals or teams.

2.4.3 Work efficacy

Effective time management is vital in any project (Hameri & Heikkilä, 2002). Tasks delays can accumulate, causing bigger problems for the entire project. Good time management helps to avoid these mistakes. It is not only about working faster, but also ensuring that tasks are completed on schedule. This keeps the project on track and improves overall efficiency.

Work efficacy in aircraft maintenance is about doing the right task at the right time. Knowing how time is spent on each task allows for better control over the schedule. When teams are clear about what needs to be done and by when, it creates accountability. This clarity can save time and boost productivity. Another important factor is how tasks connect. Smooth transitions between tasks are crucial. If one task is delayed or poorly planned, it can disrupt everything that follows. Managing these handoffs carefully keeps the workflow steady and avoids unnecessary delays.

2.5 Estimation of Task Duration

2.5.1 Methods to Estimate Task Duration

Task duration estimation involves predicting the time required to complete a specific task. In this section, it will be explored the most popular methods to better predict task duration.

The traditional methods includes experts knowledge and experience program evaluating and review technique or three-point estimating technique (PERT). The last mentioned method is the most popular and according to Habibi et al. (2018) PERT has a problem relative to the precision of the estimation. To overcome this, the author suggest some improvements in the method, such as using fuzzy logic. However, applying an artificial neural network would be easier and more accurate to predict the time duration of the task (Lishner & Shtub, 2022).

The methods that will be compared in this work are the following:

- **Delphi Technique**: is a method used to gather expert opinions to reach a consensus on task duration estimates. It involves a group of experts who provide their estimates independently, and then these estimates are shared with the group for further discussion and refinement. This iterative process continues until a consensus is reached.

- **Three-Point Estimating Technique or Program Evaluating and Review Technique (PERT)**: calculates an average estimation of the duration using optimistic, pessimistic, and most likely. Considering Optimistic (O), Pessimistic (P) and Most likely (M) the PERT formula is $\frac{O+4M+P}{6}$.

  Both Delphi and PERT techniques have limitations. For example, both rely on experts estimation who tend to overestimate (Hill et al., 2000). Moreover, it is challenging to quantify all variables considered by the experts in the estimation process (Pospieszny et al., 2018).

- **Supervised learning**: is a machine learning approach where an algorithm learns from labeled data, with inputs mapped to known outputs.

2.5.2 MRO data analytics

Pelt, Stamoulis, and Apostolidis (2019) begin by highlighting the unpredictability of the process times and material requirements, as well as the traditional preventive maintenance often results in unnecessary component replacements, which can lead to inefficiencies and increased costs. So, the authors believe that implementing a data-driven approach on the datasets could help to enhance operational sustainability and efficiency in MRO companies.

The literature explored 25 case studies across eight different aviation MRO enterprises, the authors demonstrate how data mining techniques can uncover valuable insights from extensive datasets. It employs the CRISP-DM methodology, which is noted for its widespread use in data mining. This approach provides a systematic guideline for exploring and preparing data, ensuring that the research questions are effectively addressed.

Althought, there are cases where a simple data visualization is enough for MRO operator to take a decision, Apostolidis et al. (2020) reported that for specific complex cases a more sophisticated approach is needed. In the figure 2.2 it is illustrated a proposal for how to conduct an effective MRO analysis study.

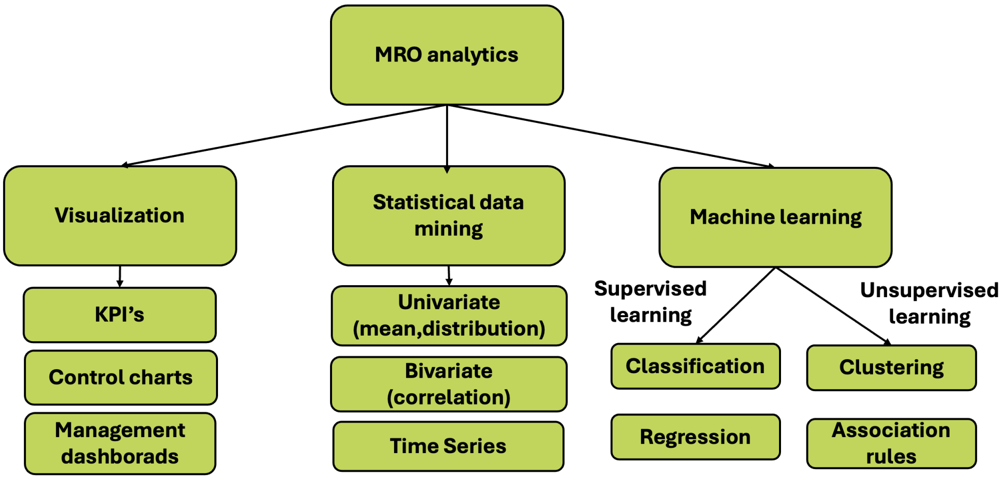


Figure 2.2: MRO analytics framework

2.5.3 Supervised Learning

The work of Rengasamy et al. (2018) aims to explore the current landscape of deep learning applications in MRO, identifying existing methodologies, their effectiveness, and research gaps. According to this reference, it was the “widespread of sensor in aircraft" (in other words, decision-making is made by collecting information through sensor monitoring) that allowed the development of model-based and data-driven methods in aerospace maintenance.

Pospieszny et al. (2018) only compared three ML methods (SVM, GLM and MLP) to estimate effort and duration of software projects. In contrast, Sousa et al. (2023) analyse and discuss eight machine learning methods to estimate the effort and duration of individual tasks in software project. Both use methods such as Support Vector Machine (SVM), and Neural Network(NN). Pospieszny et al. (2018) reported that the models that were used are the most accurate in the area. However, the author underlined that the quality of the input data will reflect on the performance of the process.

Support Vector Machine (SVM) is a supervised learning algorithm used for classification and regression tasks. It works by finding the optimal hyperplane that separates data into distinct classes with the largest margin. Meanwhile, Neural Networks (NN) are models inspired by the human brain. They consist of layers of interconnected nodes that process data. The nodes take inputs, apply weights, and pass results through an activation function to produce outputs. While processing, it learns patterns in data and adjusts weights to minimize errors.

Another important supervised learning is Linear Regression. It was the method chosen by Riahi et al. (2023) for its simplicity and interpretability. The simplicity allowed the authors to understand the relationship between various features and the response variable more naturaly. In the research, it was established a baseline with LR, so that the authors could evaluate the performance improvements offered by machine learning techniques like Extreme Gradient Boosting (XGBoost) and Random Forest (RF).

Pelt, Stamoulis, and Apostolidis (2019) criticized the limited exploration of the alternative deep learning architectures and the scarcity of large data sets, emphasizing the need for more comprehensive and lager data sets and the integration of multi-modal data to improve models accuracy.

2.5.4 Unsupervised Learning

Unsupervised learning is an area of machine learning that focuses on discovering patterns and structures in data without the need for labeled outputs. The main objective of an unsupervised learning algorithm is to group similar data points together based on their features, this is called clustering, or to determine density estimation, which is understanding the distribution of data points in the feature space, which helps in identifying areas of high and low data concentration, and or make a dimensional reduction, which is reducing the number of features in a dataset while retaining essential information (Tyagi et al., 2022).

Clustering methods are becoming increasingly popular in Maintenance, Repair, and Overhaul operations, according to Apostolidis et al. (2020). They are used to analyze and optimize various processes within the aviation industry. There are numerous applications of clustering methods in MRO, for instance, by analyzing historical data, clustering can help to predict when maintenance is likely to be needed, thus minimizing unexpected failures and operational disruptions.

The aviation industry generates vast amounts of data, and clustering helps manage this complexity by simplifying the data into understandable groups. Different clustering algorithms can be applied depending on the specific characteristics of the data and the problem at hand, which is why an algorithm that provides satisfactory results for a specific problem is not necessarily replicable for a similar one.

The effectiveness of clustering methods is mainly dependent on the quality of the data being analyzed. Poor data quality can lead to inaccurate clustering results, which in turn affects the reliability of maintenance predictions.

The authors (Apostolidis et al., 2020) mention that the development of a new clustering algorithm or the customization of existing algorithms to meet the specific needs of MRO operations is a good direction to take.

2.5.5 Statistical Knowledge

According to Graasa et al. (2018) the investigation of hidden partens in historical maintenance data is a key point for better predictive maintenance applications. The authors compare three goodness of fit tests: chi-square, Kolmogorov-Smirnov (K-S), and Maximum Likelihood Estimation (MLE) using Akaike Weights. The K-S test was found to achieve the highest accuracy in selecting the appropriate statistical distribution for time series datasets. The accuracy of these tests is significantly influenced by sample size, with larger samples yielding better results.

The statistical analysis involved estimating probability density functions (PDFs) from datasets using the Maximum Likelihood Estimator (MLE) technique. This process was repeated 2000 times for each sample size to ensure robust results while managing computational efficiency. The analysis also included a minimum sample size requirement to balance accuracy with the availability of maintenance processes, ensuring reliable probability calculations. It was found that smaller sample sizes did not significantly affect the accuracy of the selected statistical distribution, although this influence was not quantified.

The analysis demonstrated that the most suitable statistical distribution could differ from the reference distribution without significantly impacting probability calculations. This finding suggests flexibility in the choice of statistical models based on the data characteristics.

2.6 Evaluating Metrics

Accuracy metrics measure how well a model predicts. They compare predictions to actual results, and without proper metrics, predictions cannot be trusted or applied effectively in real-world scenarios (Hossin & Sulaiman, 2015).

A more detailed presentation of the evaluation metrics selected for this study is provided in Chapter 3.5. This chapter shows what metrics were selected and explains how they are applied within the context of the maintenance project analysis.

2.7 Conclusion

The estimation of the duration of the aircraft maintenance task is one of the most challenging areas to predict. Although there have been many exploratory reviews about data analysis in MRO airlines, few have showed an actual approach. However, outside of aircraft maintenance there are many studies on time management and estimation of the duration of projects (Francis-Smythe & Robertson, 1999; Pospieszny et al., 2018; Sousa et al., 2023), for example, which allowed a greater understanding of the subject. In order to calculate the duration estimation it was explored the application of traditional (expert estimation) and modern approaches (ML). The findings suggested that while traditional methods provide a structured framework for estimation, machine learning approaches showed good precision capacity (Pospieszny et al., 2018).

3

Methodology

3.1 Introduction

This chapter covers the research methodology. The objective is to present the research design selected for this work, a detailed data description, and the analytical methods utilized to achieve the main goal of the study.

This study employs a quantitative approach to investigate whether ML methods can deliver more accurate predictions of the duration of maintenance aircraft activities compared to traditional methods. Therefore, both ML techniques such as linear regression and traditional approach such as PERT will be applied to analyze the data.

The figure 3.1 represents the four main sections of this chapter, beginning with the designation of the research design for the study; the data analysis section it is presented the description and analysis methods to perform on the dataset; a description of the tools and methods used to estimate the duration of the aircraft activities can be found in the baseline model section, and in the last section the evaluation metrics to evaluate the results of the estimation times.

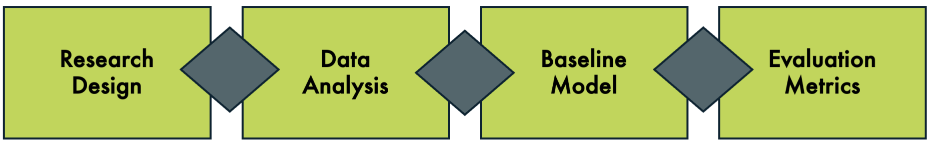


Figure 3.1: Chapter structure.

3.2 Research Design

This study will be divided into two parts. In the first part, an exploratory analysis will be performed, with the objective of finding relevant insights as well as trends. In the second part, to predict the duration activities by applying machine learning techniques, an experimental research was performed.

3.3 Data

The data set used in this study was provided by HANGAR 901 Aircraft Maintenance GmbH, and was collected by they evaluation team and employee records. In other words, it is a secondary research.

3.3.1 Data Description

In this subsection, the four aircraft types, and the features of the data set will be presented and described.

As stated before, the type of maintenance is heavy maintenance or known as C-checks. The Boeing 777, which are airplanes for long-haul flights, are in number in the data set, yet there is an Airbus 350 also used for long-haul flights. The following types of aircraft can be found in the data set:

1. **Airbus 350-941CJ** - It is derived from the Airbus 350-900, a long-haul flight. "CJ" stands for **Corporate Jet** indicating a private or government version. Hence, this airbus 350 version is designed for ultra-long-haul flights with luxurious interiors and advanced technology for high-profile people (VIP).

2. **Boeing 777-200** - It belongs to the Boeing 777 family, and has a wide body with twin-engine design to improve fuel efficiency in long-haul flights. The interiors are spacious and comfortable.

3. **Boeing 777-300ER** - It belongs to the Boeing 777 family, the "ER" stands for Extended Range and features higher fuel capacity, stronger landing gear, and improved aerodynamics for ultra-long-haul flights. It has bigger passenger capacity along all types and wider cabin.

4. **Boeing 777-328ER** - It is a customized version of Boeing 777-300 for a specific airline, in this case the 328 code refers to AIR FRANCE. Apart from customer-specific configurations, it retains the same overall capabilities and performance as a standard 777-300ER.

The data set analyzed in this study comprises records of maintenance activities performed in aircrafts. There are sixteen features in which two are quantitative data (Estimated Time and Actual Time), while the remaining are qualitative data. Each feature captures different aspects of aircraft maintenance. The following table summarizes the features of the data set.

Table 3.1: Description of the Data Set given

Column Name	Description
Project Number	It is an unique identifier (e.g H1-23-05-00092) assign to a specific project to identify tasks or works done on an aircraft and it helps to quickly locate on the database.
Aircraft Type	The different models/types of aircraft presented on the dataset.
Work Order Number	Unique sequential number that is assigned to each project for the purpose of allowing the hangar to identify and track the progress of the tasks.
Item Number	Unique number that refers to a specific part, component, or task within a work order. It is used to organize and track individual elements.
Work Order Type	If the order is the routine type (scheduled and planned maintenance) or non-routine type (unexpected and unscheduled maintenance).
Work in Progress Status (WIP Status)	Tells the status of the project, if is closed (actions are finished) or ready to go.
Work in Progress Reason (WIP Reason)	Tells the reason of the project, if is completed, canceled, not apply or to work.
Close Date	Represents the conclusion date of the project.
Milestone	Main steps of the checks incoming like inspections, opening the panels and testing to ensure the airwothiness and safety of the aircraft. If is found an abnormality it starts the non-routine work.
Aircraft Location	The part of the aircraft where the project/task is being carried out.
Description	Description of the intervention performed on the aircraft.
Primary Skill Code	Different types of skills that are necessary to perform the maintenance task.(e.g. Mechanic)
Actual Time	The actually time spent on the maintenance task.
Estimated Time	The calculated time required for the maintenance task.
Customer Work Card	It is the customer identification in the database.
ID	Unique number that identifies the employee in charge of the respective action.

3.3.2 Data Analysis

Python and Power BI desktop were the two main software tools utilized in this work. Python was chosen for its ability to handle large data sets, especially through Pandas package, while Power BI desktop was chosen for its ease of getting data insights.

In a Power BI desktop file, the data set (*MH_Table.xlsx*) was exported. Then it was made a reference of the original data called *clean data* were the data was cleaned, and filtered the following columns: “WORK_ORDER_TYPE” by choosing only the ROUTINE ones; “PRIMARY_SKILL_CODE” by removing the ADMIN, PLANNING, BM and STORE; “MILESTONE” by removing the ADMIN, BM and WAIT. In addition, the aircraft types Boeing 777-300ER and Boeing 777-328ER were merged because of maintenance similarities. Outliers were identified and consequently removed.

After the data cleaning process, an exploratory research was conducted to gain insights into the dataset. Scatter plots for most columns were performed, for instance, to examine how many "PRIMARY_SKILL_CODE" or "MILESTONE" there were and its frequency. Additionally, correlation analysis among the variables was done, using the Pearson Criteria coefficient, and a statistical description table was outputted.

Furthermore, for data reduction, the **K-means cluster algorithm** was used. According to Tyagi et al. (2022), this method is one of the most basic and the first go to clustering algorithm. It computes the Euclidean distance employing the following formula:

$$d(x, m_k) = \sum_{n=1}^{N} [x(n) - m_k(n)]^2 \tag{3.1}$$

where x is the data point , N is the dimensional space and m_k is the centroid (or the mean of the vector) of cluster k. The number of cluster is usually chosen by the user, but to select the best number of cluster an **Elbow Method** was done. This method selects the best number of clusters for the lowest error value using the following equation:

$$\frac{1}{N_v} \sum_{k=1}^{K} E_k \tag{3.2}$$

where N_v is the number of data points in the data set and E_k is the distance measure for each K cluster.

To demonstrate the clustering process, here is an example from (Tyagi et al., 2022) on random data. The following image (Figure 3.2) illustrates the Elbow Method, where it is possible to conclude that the ideal number of clusters is 2.

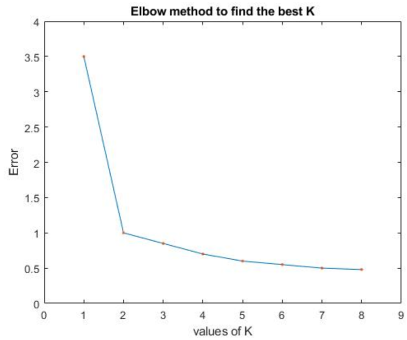


Figure 3.2: Example of the Elbow method.

Then, using *Scikit-learn* in Python, the k-means clustering was implemented. Based on the figure 3.3, it is confirmed that two groups were formed.

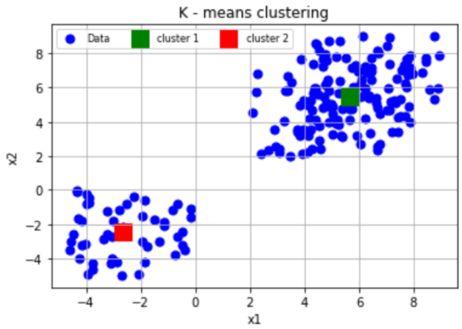


Figure 3.3: Example of the K-means clustering.

3.4 Baseline Model

3.4.1 Traditional Technique Framework

To address the experimental research and the main question of the study - "Which method provides more accurate predictive results for task durations?", the following methods were investigated.

In first place, the traditional method **Program Evaluating and Review Technique (PERT)** was computed using Python. The choice of using this well-known technique lies in its ability to estimate complex projects duration (Nafkha, 2016). It utilizes the expert evaluation for an activity shortest duration (Optimistic), most likely duration (Most-likely) and longest duration (Pessimistic). Having these three estimations it is possible to apply the following formula to obtain a new estimation of the activity duration.

$$\frac{O + 4M + P}{6} \tag{3.3}$$

where O is the Optimistic, P - Pessimistic and M - Most-likely.

3.4.2 Machine Learning Framework

In the second place, machine learning methods were tested. Based on some literature (Lishner & Shtub, 2022; Sousa et al., 2023) on the estimation of duration of tasks or projects four methods were selected for implementation: Linear Regression, Support Vector Regression, Random Forest and Neural Networks.

- **Linear Regression** - It is effective on large data sets and provides clear interpretable results. However, it is sensitive to outliers. It is based on the assumption of a linear relationship between the variables and the dependent variable (task duration) and it is represented by the following equation.

$$y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + ... + \beta_n X_n + \epsilon, \tag{3.4}$$

  where X_i are the features, β_i is the coefficient and ϵ is the error.

- **Support Vector Regression** - It is effective to model non-linear data and robust to outliers. Additionally, it is based on the Support Vector Machine, but it focuses on finding a hyperplane that best fits most data points within a margin of tolerance.

- **Random Forest** - It is effective on handling non-linear data. It builds multiple decision trees and average their predictions to reduce overfitting. However, it is less interpretable compared to Linear Regression.

- **Neural Networks** - It is effective for large data sets and model non-linear. However, it can be difficult to interpret the results and computationally consuming.

The reason for choosing these algorithms is to assess different modeling approaches where each offers distinct strengths and theoretical foundations.

The overall process is illustrated in figure 3.4. It starts with the raw data collected from the aircraft maintenance activities, this data undergoes a cleaning and selection process where outliers are detected, missing values are handled, and target variables selected. The cleaned data goes through an exploratory analysis in which a descriptive analysis, a correlation analysis, clustering, and data visualization are done. These steps aim to uncover patterns, and correlations to provide relevant insights. Then, a computational analysis is conducted using two different approaches: PERT algorithm and Machine Learning. The machine learning models includes linear regression, support vector regression, random forest, and neural network. The performance of both techniques is evaluated by key metrics such as MAE, RMSE, R^2, and MAPE.

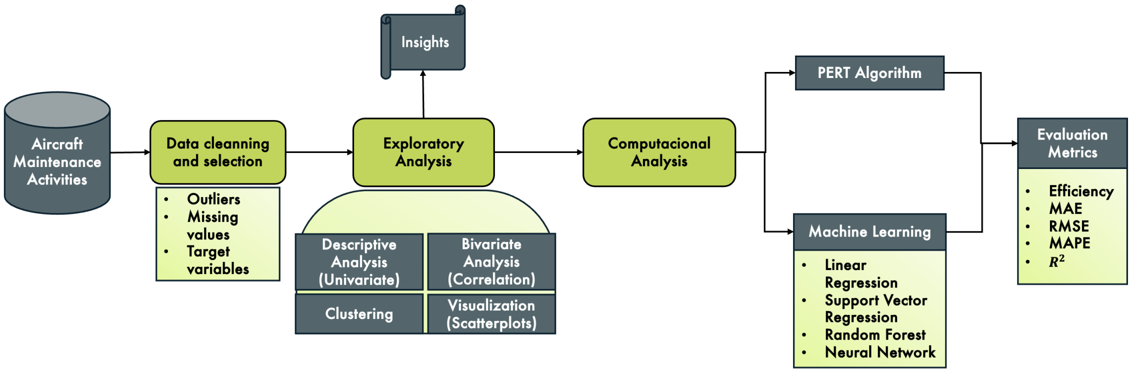


Figure 3.4: Processing workflow.

3.5 Evaluating Metrics

To assess the quality of the models trained it was used the following evaluation metrics: Efficiency ratio and the Regression metrics.

The Efficiency ratio was calculated to measure the efficacy of the estimated times given by the company as well to measure the efficiency of the PERT estimation. The formula used was the following:

$$\text{Efficiency Ratio} = \frac{\text{Estimated time}}{\text{Actual time}} \tag{3.5}$$

To evaluate the estimation achieved by the predictive models the following table summarizes the metrics utilized.

Table 3.2: Principal evaluating metrics to measure model performance.

	Metric	Formula	Description
Regression	Mean Absolute Error (MAE)	$\frac{\sum_{i=1}^{N}\lvert t_{\text{actual}_i} - t_{\text{estimated}_i}\rvert}{N}$	Average absolute difference between predicted and actual time.
	Root Mean Square Error (RMSE)	$\sqrt{\frac{\sum_{i=1}^{N}\left(t_{\text{actual}_i} - t_{\text{estimated}_i}\right)^2}{N}}$	Measures the square root of the average squared differences between predicted and actual time.
	Mean Absolute Percent Error (MAPE)	$\frac{\sum_{i=1}^{N}\lvert t_{\text{actual}_i} - t_{\text{estimated}_i}\rvert}{N \times \lvert t_{\text{actual}_i}\rvert} \times 100$	Average percentage error between predicted and actual time.
	R^2	$1 - \frac{\sum_{i=1}^{N}\left(t_{\text{actual}_i} - t_{\text{estimated}_i}\right)^2}{\sum_{i=1}^{N}\left(t_{\text{actual}_i} - \bar{t}_{\text{actual}}\right)^2}$	Measures the proportion of variance in actual time explained by the model.

Furthermore, a 30% tolerance for error was agreed with the company, which means that deviations between the estimates and the actual values are acceptable as long as they do not exceed 30%. The Figure 3.5 provides a schematic representation of how the tolerance is applied for each given scenario. For instance, a problematic scenario happens when the actual time significantly exceeds the estimated time, as is possible to observe in the second quadrant. These types of scenario are undesirable and should be avoided.

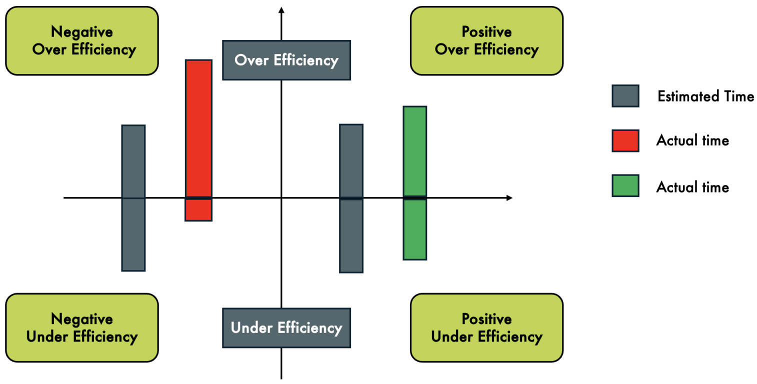


Figure 3.5: Processing workflow.

3.6 Conclusion

This study aims to explore whether machine learning methods can provide more accurate predictions for the duration of aircraft maintenance tasks compared to traditional approaches. A quantitative methodological approach was employed, and the research design discussed to achieve the main objective of the study. In the next chapter the planned methodology will be applied, and the results will be presented.

4

Results and discussion

In this section, the findings of the exploratory and experimental analysis will be presented. These results are organized by development order starting with data insights, followed by model comparison and discussion.

4.1 Data Insights

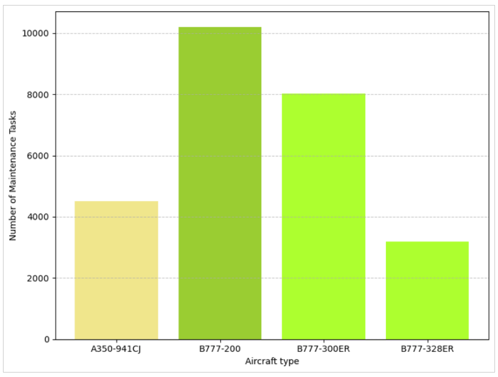
Table 4.1: Statistical Description of the data.

	Actual Time	Estimated Time
Count	17460	17460
Mean	5.226	4.108
Minimum	0.010	0
25 %	1.09	0.9
50 %	2.30	2
75 %	5.32	4.05
Maximum	455.15	280
Standard Deviation	11.829	9.09

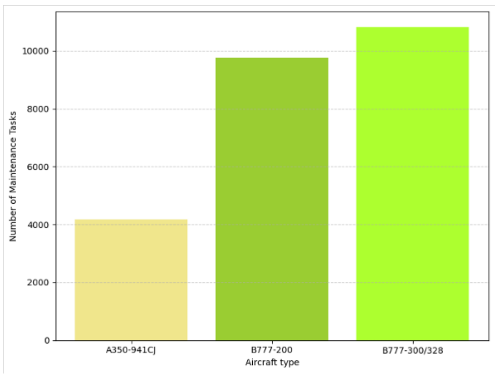
In the data insights section, it was performed an exploratory overview of the data. The data set is constituted in its majority by qualitative data and only two columns of quantitative data (Actual time and Estimated Time), see Table 3.1, for further details. Table 4.1 presents the statistical summary of these variables. A notable finding is the systematic underestimation in the "Estimated Time" column, where predicted values consistently fall below actual completion times across all statistical measures.

The mean actual time (5.226 hours) exceeds the mean estimated time (4.108 hours) by approximately 1.12 hours, representing a 21.4% ($\frac{1.1}{5.226} \times 100\%$) underestimation . This pattern persists across all percentiles, suggesting a systematic bias in the company's current methodology. The high standard deviations (11.829 for actual time, 9.09 for estimated time) indicate substantial variability in task durations, highlighting the complexity of maintenance operations.

Then some graphics were performed to gain insights from the data. The first graphic was a histogram with the number of maintenance interventions done in each plane; in this case, it represents each row of the data set. From the figure 4.1a it can be observed how many aircraft types there are in the data, which are they and their frequency. The data set covers four aircraft types: A350-941CJ, Boeing 777-200, Boeing 777-300ER, and Boeing 777-328ER. Due to maintenance procedure similarities between the Boeing 777-300ER and 777-328ER variants, these were consolidated into a single category (Figure 4.1b). The Boeing 777-200 and the combined Boeing 777-300/328ER categories account for the majority of maintenance interventions, reflecting either fleet composition or maintenance frequency differences.



(a) For each aircraft.



(b) After merging the BOEING 777-300.

Figure 4.1: Bar chart showing the number of interventions/maintenance tasks for each aircraft.

Another graphic included in the study was a pie chart with the percentage of the interventions carried out at each location of the airplane for all aircraft types (Figure I.3), and for each aircraft type which can be seen in Annex 1.

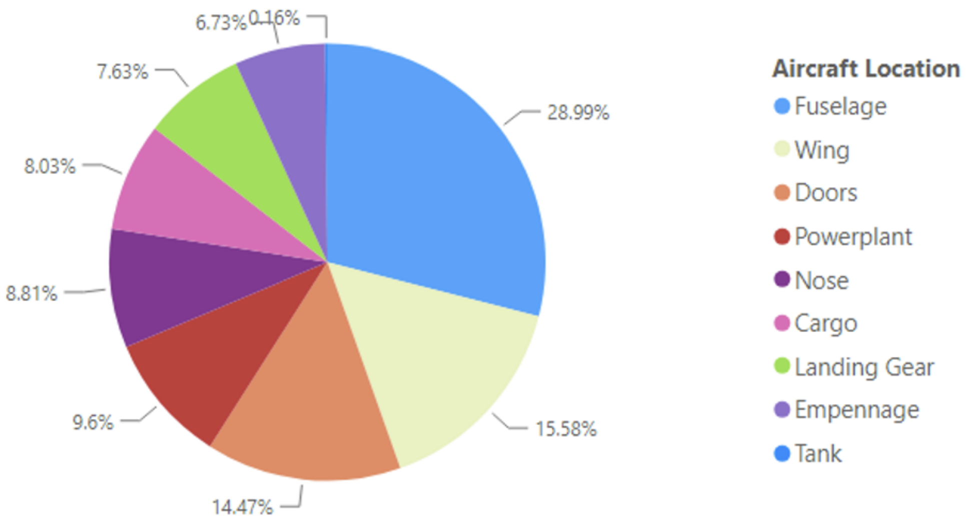


Figure 4.2: Pie Chart of the airplane maintenance interventions by aircraft location for all aircraft types.

In the figure I.3, it is possible to see that there are nine main locations in every airplane where maintenance tasks are performed. The area in the plane that has the most interventions from the maintenance staff is the fuselage (which is the main body of the plane), followed by the wings and doors. This happens for all types of Boeing 777 aircraft. The A350-941CJ exhibits a different maintenance pattern, with cargo and powerplant locations ranking second and third after fuselage maintenance. This divergence may reflect design differences, operational requirements, or maintenance philosophy variations between aircraft manufacturers. (see annex I)

Then a bar graph of the average and median time, for a variable named **Primary Skill Code**, spent to perform their task was plotted. It shows that the primary skill that takes longer average time is cleaning, but there is a large gap between the average (13.9h) and the median time (5.8h), this suggests significant outliers or a skewed distribution where some tasks took much longer than most. Sheetmetal and Cabin skills also show a large gap between the average and median time indicating variability and extreme outliers. Composite, Mechanical (MEC) and Avionics (AVI) skills have low median time suggesting that the tasks are quicker or simpler, and the distance between the average and median time is smaller. Non-Destructive Testing (NDT) and Paint have a small distance between the average and median time indicating less skewness and more consistent tasks durations. Engine Run shows a median of 0h and a low average time because these tasks are performed quickly.

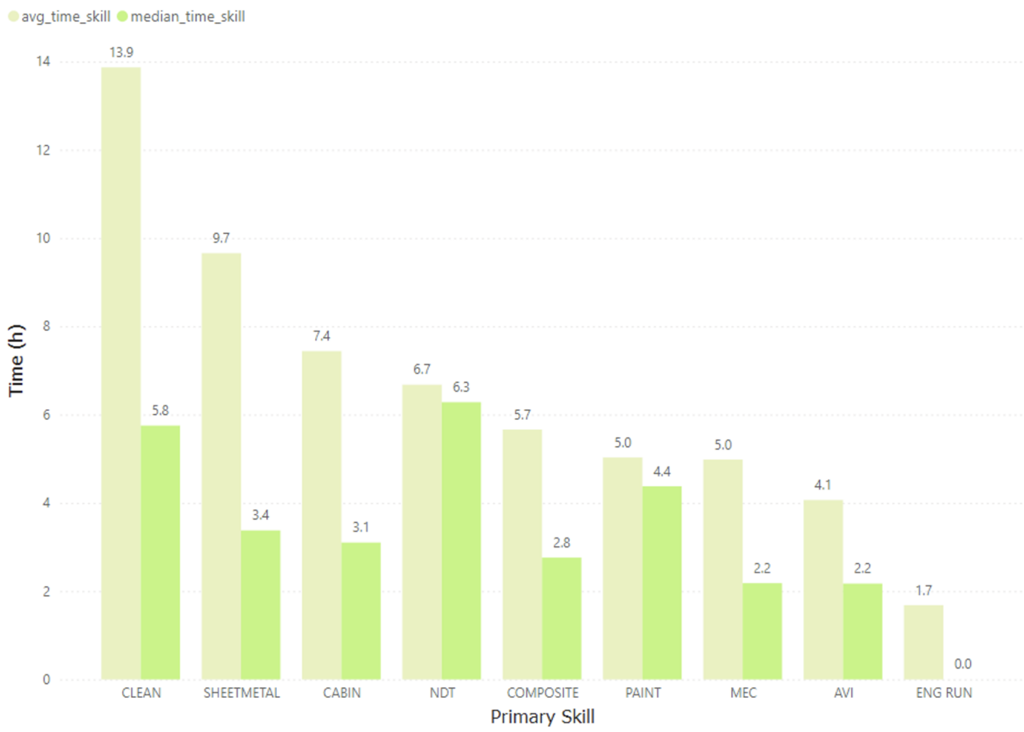


Figure 4.3: Average and median time spent to complete the tasks by primary skill code.

Since one of the goals of this project is to evaluate the efficiency of maintenance activity durations, the expression explained in chapter 3.5 was used to calculate the efficiency ratio.

To quantify maintenance performance, an efficiency ratio was calculated as the quotient of estimated time divided by actual time. Values above 1.0 indicate tasks completed faster than expected, while values below 1.0 suggest longer-than-expected completion times. The ideal efficiency ratio approaches 1.0, representing accurate estimation.

Data preprocessing revealed problematic values including extreme outliers and infinite values, likely indicating data entry errors or incomplete records. To allow meaningful analysis, efficiency ratios greater than 2.0 were capped at this value, and infinite values were removed, reducing the dataset from 34 434 to 17 460 records.

The most efficient primary skill is Engine and the least is Avionics (Fig. 4.4). The reason why the avionics skill is low should be investigated. Potential contributing factors may include a skill problem, systematic issues or resources constraints. Understanding these root causes is essential for proposing targeted improvements and enhancing overall maintenance performance.

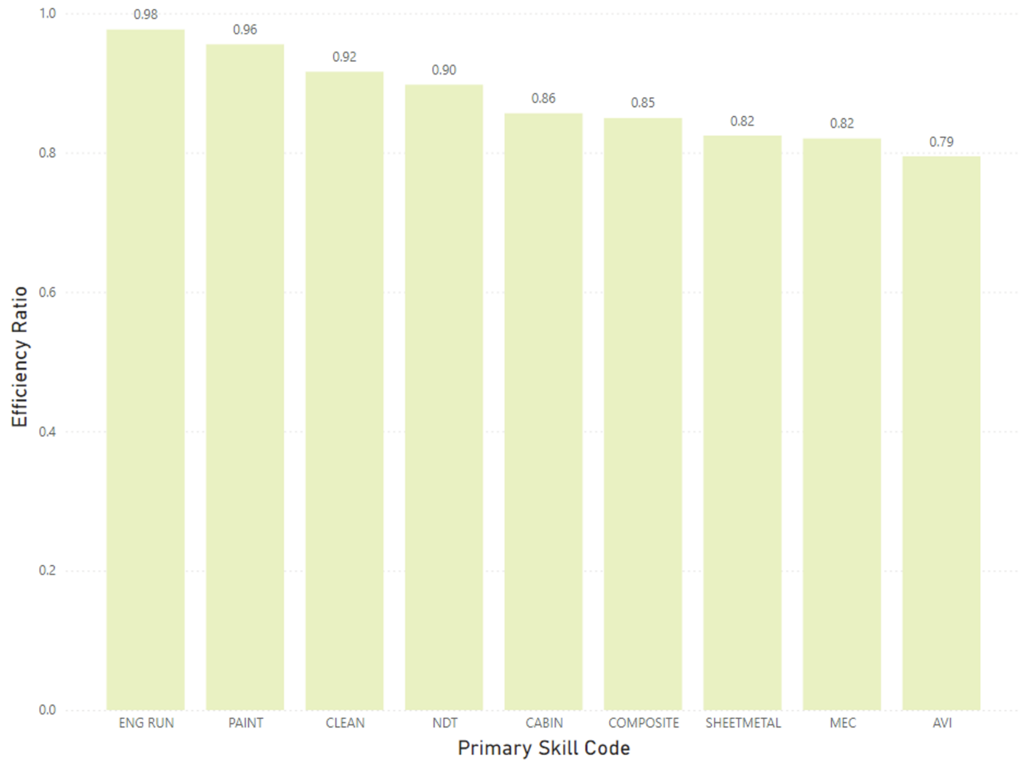


Figure 4.4: Average efficiency ratio by Primary Skill Code.

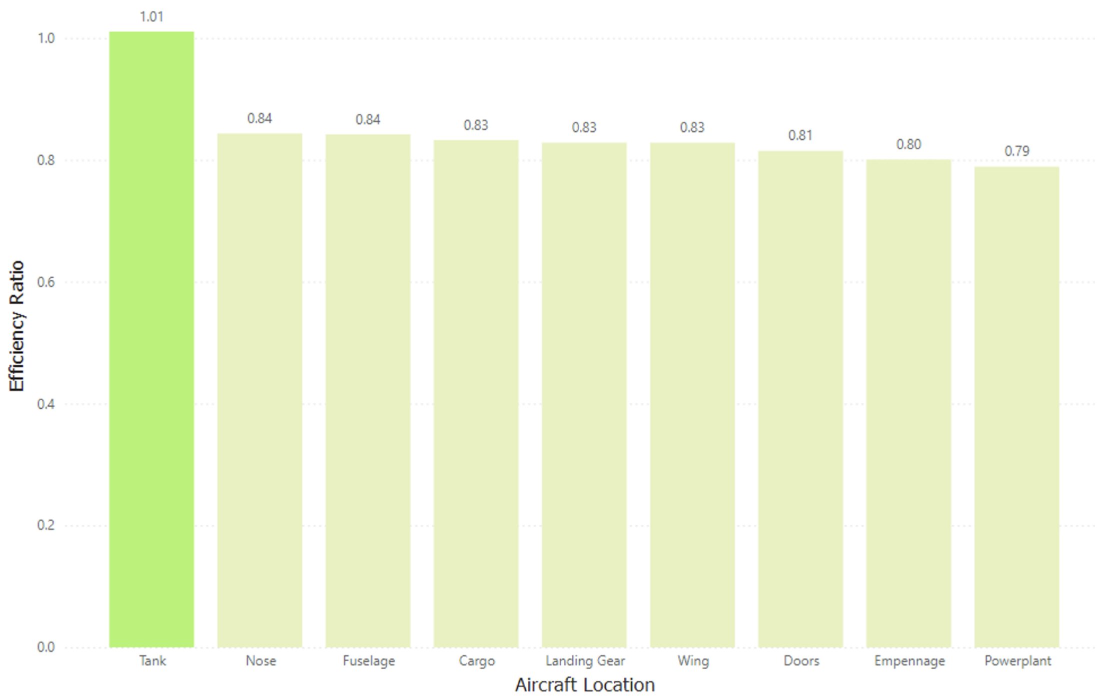


Figure 4.5: Average efficiency ratio by Aircraft Location.

In the case of aircraft location, the Tank is the most efficient location and the Powerplant is the least efficient. The Tank location is highlighted in a darker shade of green in the figure 4.5 due to its higher performance compared to other areas. A significant portion of the activities carried out at this location involve fuel tank operations, which are consistently executed with high efficiency. In contrast, the other aircraft locations show minor differences in performance, suggesting a more uniform level of efficiency in these locations.

4.1.0.1 Clustering Analysis

To address data reduction and uncover hidden patterns, a clustering analysis was performed. K-means was applied in the context of the efficiency ratio of the duration time. The elbow method determined optimal cluster numbers for the data set and individual aircraft types (Figures 4.9,4.7,4.8,4.10).

First, the data set containing all types of airplane was grouped, and the figure 4.6 illustrates the activities based on how long they take and how efficient they are. From the elbow method it was determined that the ideal number of clusters was 5, see figure 4.9. In the figure 4.6 it is shown the five clusters, each one highlighted with a different color and labeled accordingly.

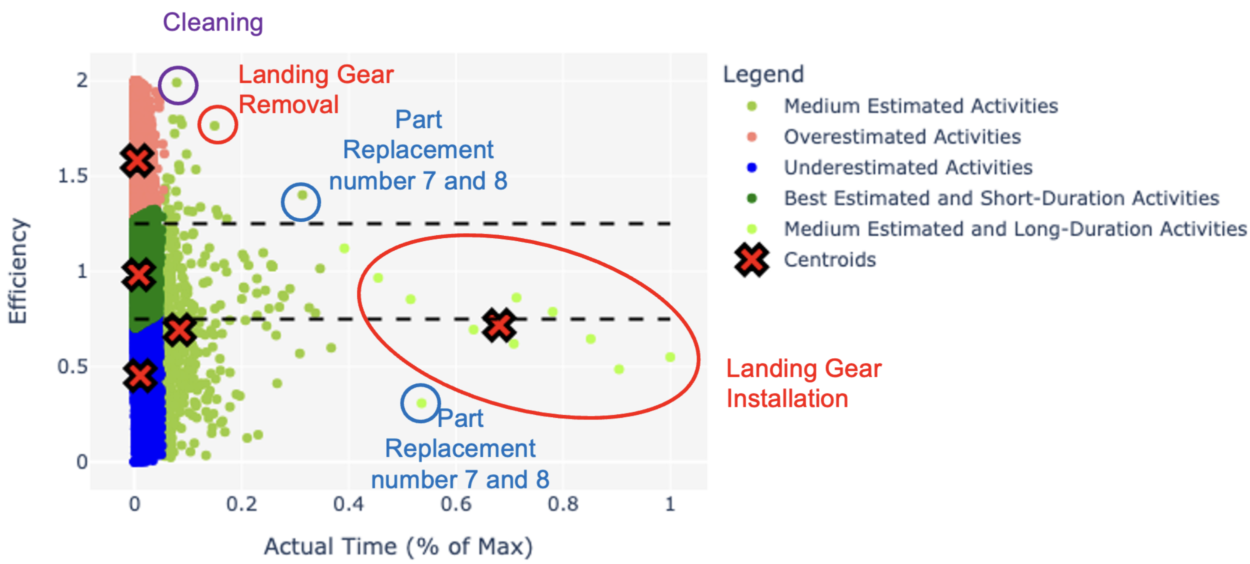


Figure 4.6: Clustering by efficiency for all aircraft types

There are five different colors to label the five clusters:

- Dark blue - Underestimated Activities - This cluster groups the activities with the lower efficiency and concluded in a short period of time.

- Dark green - Best Estimated and Short-Duration Activities - The activities in this cluster are the better fitted in the context of efficiency (between 0.75 and 1.25) and concluded in a short period of time.

- Medium green - Medium Estimated Activities - This group was the most difficult to label because it has activities in every range of efficiency, with no tendency to underestimated or overestimated, so it was decided to call it medium.

- Light green - Medium Estimated and Long-duration Activities - This group includes the long duration activities with acceptable efficiency.

- Salmon - Overestimated Activities - Upper side group with the overestimated activities.

The main observations about the figure are: Landing Gear Installation activity is the most time-consuming activity with good efficiency but with room for improvements; the Part Replacement Number 7 and 8 activity have both high and low efficiency, suggesting inconsistency in execution or estimation; Cleaning activity has high efficiency and short duration, suggesting that cleaning tasks are predictable and well managed.

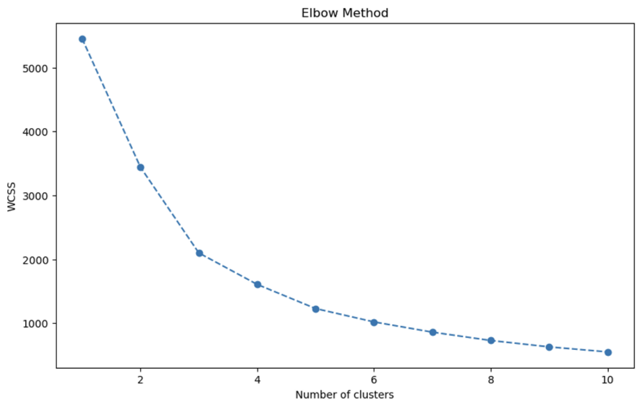


Figure 4.7: Aircraft Type A350

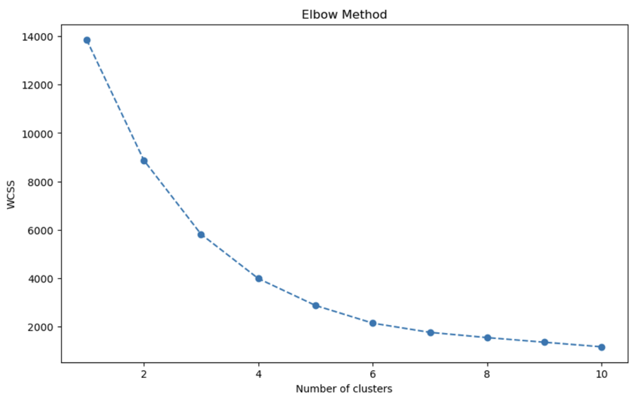


Figure 4.8: Aircraft Type B777-200

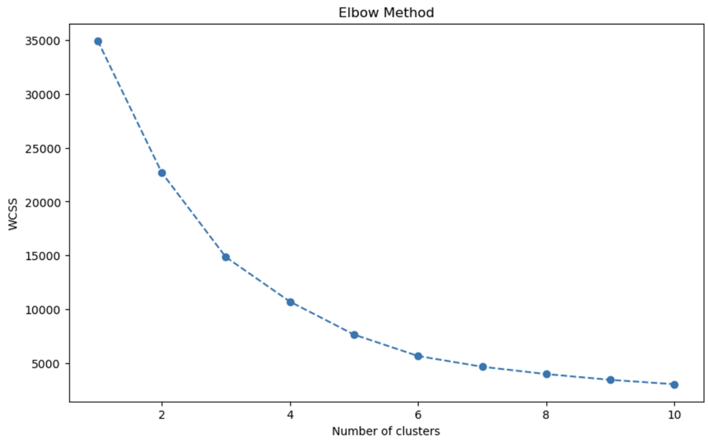


Figure 4.9: All Aircraft Types

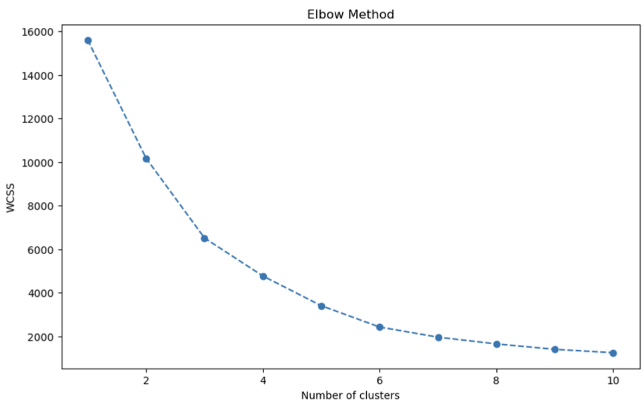


Figure 4.10: Aircraft Type B777-300/328ER

The figure 4.11 represents the clustering for the Airbus 350 type and shows that in general the activities performed take longer compared to the Boeing type. This can suggest that the managing time is better or that the team is well prepared for this aircraft type than for the Airbus. Regarding the Boeing type of aircraft, the activities that take longer are the Landing Gear Installation/Removal. In contrast, Cleaning activity in the Airbus type is the most time consuming. In the same figure(Fig. 4.11), Inspection and Lubrication show low efficiency, unlike Incoming. The Centroid of the Medium-Duration Activities is on the inferior dashed line suggesting overall moderate performance. The Cleaning activity and Inspection (C-checks and test) on the Boeing 777-200 (Fig. 4.12) shows to be contributing to hidden inefficiencies, even if they are not the most time-consuming.

Figure 4.11: Clustering by efficiency for aircraft type A350

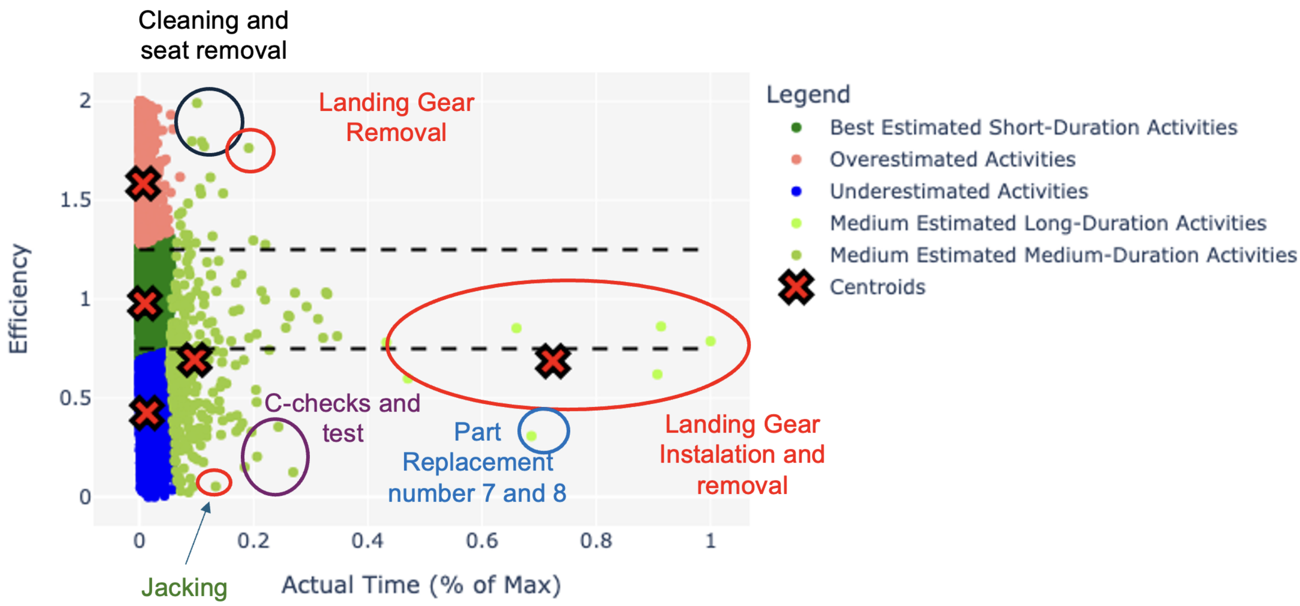


Figure 4.12: Clustering by efficiency for aircraft type B777-200

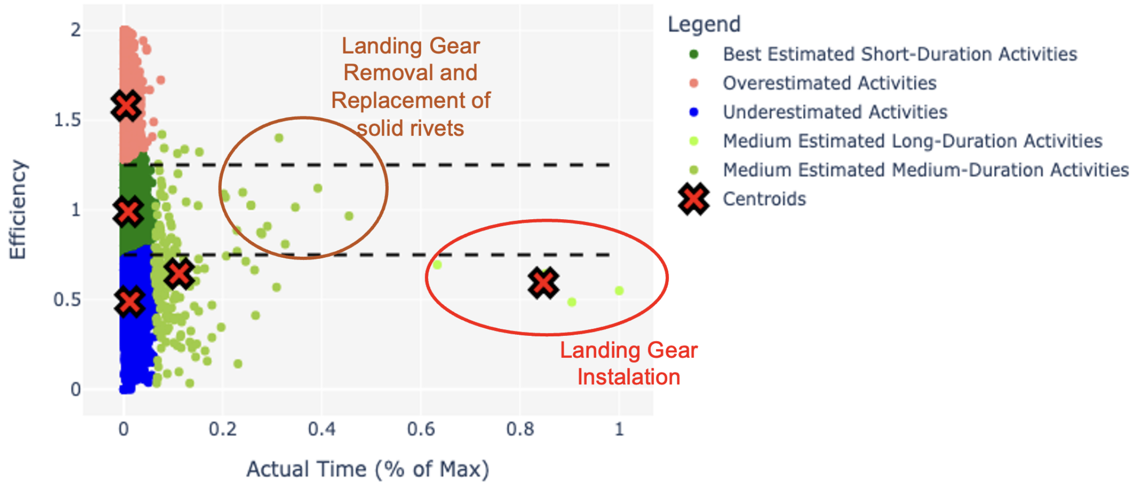


Figure 4.13: Clustering by efficiency for aircraft type B777-300/328ER

4.2 Model Comparison

To address the question of what model can provide more accurate predictions for the duration of aircraft maintenance tasks, five models were compared. The numerical results of this comparison are presented in Table 4.14.

From Table 4.14, the main conclusion is that the machine learning algorithms were able to provide better results than traditional methods. The Random Forest model achieved the highest coefficient of determination ($R^2 = 0.841$), explaining 84.1% of the variance in maintenance duration. The Artificial Neural Network demonstrated the lowest Mean Absolute Error (MAE = 1.139), indicating a higher prediction accuracy for individual tasks. Traditional methods showed poorer performance, with PERT achieving the lowest R^2 (0.802) and highest Mean Squared Error (MSE = 26.492). This substantial performance gap supports the adoption of machine learning approaches for predicting maintenance duration.

Metric / Algorithm	Mean Absolute Error (MAE)	Mean Squared Error (MSE)	Root Mean Squared Error	R^2
PERT	1.874	26.492	5.147	0.802
Linear Regression	1.546	12.825	3.581	0.831
Support Vector Regression (SVR)	1.492	13.749	3.708	0.817
Random Forest	1.222	12.079	3.476	0.841
Neural Network	1.139	12.338	3.513	0.837

Figure 4.14: Results of the model evaluation with different models.

Figure 4.15 displays the distribution of the signed, absolute, and squared errors for the different models.

From the signed error plot, the linear regression has the most variability, meaning that it makes over- or under-predictions. The support vector machine, random forest, and neural network shows more balanced estimations. The random forest and neural network have the lowest median and mean absolute errors suggesting that these models are more accurate. In conclusion, the RF and ANN models are the most stable and accurate, while linear regression is not.

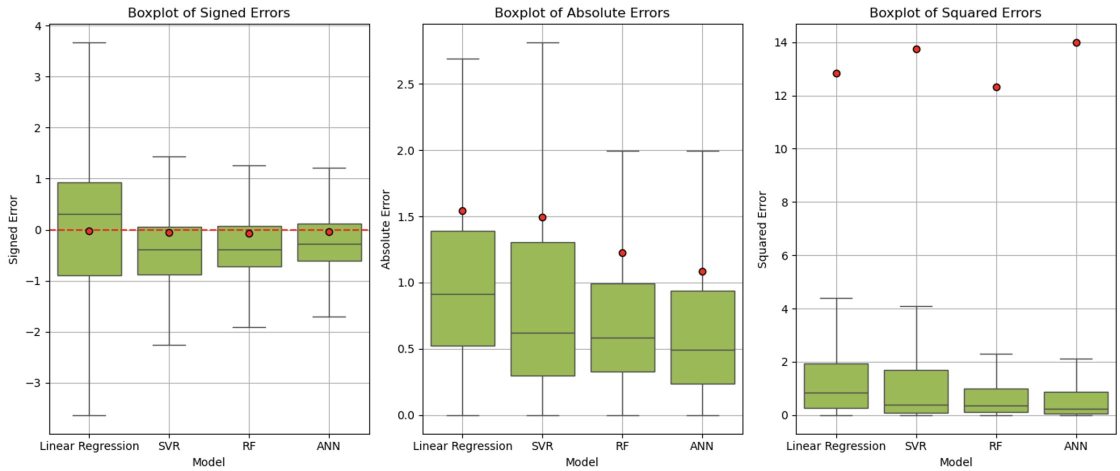


Figure 4.15: Box-plot graphs.

4.3 Discussion

The systematic underestimation identified in current company practices has significant operational implications. The 21.4% average underestimation can lead to resource allocation issues, scheduling conflicts, and customer service problems. Implementing machine learning based prediction models could significantly improve estimation accuracy and operational efficiency (Lishner & Shtub, 2022; Riahi et al., 2023; Sousa et al., 2023). The identification of skill-specific and location-specific efficiency patterns provides actionable insights for targeted improvements. The low efficiency in Avionics maintenance skill suggests the need for specialized training, resource reallocation, or process standardization, for example.

This study demonstrates the value of combining descriptive analytics (clustering, efficiency metrics) with predictive modeling for maintenance optimization. The clustering analysis provides interpretable insights into operational patterns, while machine learning models offer accurate prediction capabilities. This integrated approach supports both strategic decision making and operational planning (Hill et al., 2000; Pospieszny et al., 2018).

4.3.1 Limitations

This study has some limitations. Data quality issues (Apostolidis et al., 2020), evidenced by extreme outliers and missing values, required significant pre-processing that may have affected model performance. The scope of the study is limited to the prediction of duration and does not consider other important factors such as maintenance quality, safety implications, or cost.

4.4 Conclusion

This study demonstrates the value of integrating data-driven methods into aircraft maintenance operations (Rengasamy et al., 2018). The combination of efficiency metrics, clustering analysis, and predictive modeling provided insight into performance variability along tasks, skill types, and aircraft locations. There are three major contributions from this study:

- Underestimation calculations - Current company estimation consistently underestimate maintenance duration, indicating the need for improved estimation practices.
- Performance Variability - Significant differences in efficiency exist across maintenance skills and aircraft locations, with Avionics showing room for improvement while Engine and Tank operations demonstrate high efficiency.
- ML high predictive performance - Machine learning algorithms, particularly Random Forest and neural networks, surpassed traditional estimation methods, achieving values R^2 above 0.83 compared to 0.80 for traditional approaches.

Future work could extend this analysis by incorporating more variables (e.g., technician experience, environment conditions) and testing the models in operational environments to evaluate their real world impact.

5 Conclusion

This study main goal was to improve the estimation of the maintenance task durations within the aircraft maintenance operations by applying data analytics and machine learning techniques. The analysis was based on real operational data, that is comprised mostly of categorical variables, with two key continuous features: estimated time and actual time, where an underestimation of task durations was identified, revealing a significant discrepancy in the current estimation process.

Through the use of data analysis, visualizations, and clustering methods, the research featured patterns in task efficiency by skill type, aircraft type, and location, these insights revealed inefficiencies in Avionics and Inspection tasks, as well as good efficiencies in task areas such as Tank and Engine. Tasks such as Landing Gear installation and removal demonstrated to be more time-consuming in overall, while cleaning was the fastest, which was the case for the Boeing aircraft type. For the Airbus type interior cleaning was the most time-consuming task.

To address the prediction problem, machine learning models were compared against traditional estimation methods, where the Random Forest model achieved the highest R^2 (0.841), while the Artificial Neural Network model achieved the lowest MAE (1.139), but both outperforming the traditional PERT estimation technique. These results show the value of data-driven approaches in improving accuracy in maintenance task planning.

This thesis contributes to the MRO industry by providing two key frameworks: an analytical approach using cluster analysis to identify and categorize maintenance efficiency patterns, and a predictive modeling framework that integrates machine learning methods into MRO planning systems for an improved scheduling accuracy (Rengasamy et al., 2018).

In conclusion, this study demonstrates that the integration of data analytics and machine learning techniques can improve MRO task planning and performance.

Bibliography

Apostolidis, A., Pelt, M., & Stamoulis, K. P. (2020). Aviation data analytics in mro operations: Prospects and pitfalls. *2020 Annual Reliability and Maintainability Symposium (RAMS)*, 1–7. https://doi.org/https://doi.org/10.1109/RAMS48030.2020.9153694 (cit. on pp. 11–13, 34).

Chandola, D. C., Verma, S., Jaiswal, K., Chandola, P., Goyat, M., & Narvekar, N. (2023). An exploratory study on the significance and challenges of aircraft base maintenance engineering in the aviation industry. *2023 International conference on computational intelligence and knowledge economy (ICCIKE)*, 420–425. https://doi.org/https://doi.org/10.1109/ICCIKE58312.2023.10131844 (cit. on pp. 1, 6, 9).

De Jonge, B., Dijkstra, A. S., & Romeijnders, W. (2015). Cost benefits of postponing time-based maintenance under lifetime distribution uncertainty. *Reliability Engineering & System Safety*, *140*, 15–21. https://doi.org/https://doi.org/10.1016/j.ress.2015.03.027 (cit. on p. 6).

Francis-Smythe, J. A., & Robertson, I. T. (1999). On the relationship between time management and time estimation. *British Journal of Psychology*, *90*(3), 333–347. https://psycnet.apa.org/record/1999-11410-002 (cit. on pp. 9, 14).

Graasa, R., Knuyta, J., Zantmanb, M., Patróna, R. F., Pelta, M., & de Boera, R. (2018). Predicting aircraft maintenance durations using automatic selection of statistical distributions and time series forecasting models. In *Aegats2018*. https://www.researchgate.net/publication/338488850_Predicting_aircraft_maintenance_durations_using_automatic_selection_of_statistical_distributions_and_time_series_forecasting_models (cit. on p. 13).

Habibi, F., Birgani, O., Koppelaar, H., & Radenović, S. (2018). Using fuzzy logic to improve the project time and cost estimation based on project evaluation and review technique (pert). *Journal of Project Management*, *3*(4), 183–196. https://www.semanticscholar.org/paper/Using-fuzzy-logic-to-improve-the-project-time-and-Habibi-Birgani/27e7650c2251bf17d4e2ed9993648f0f744d8c1a (cit. on p. 10).

Hameri, A.-P., & Heikkilä, J. (2002). Improving efficiency: Time-critical interfacing of project tasks. *International Journal of Project Management*, *20*(2), 143–153. https://www.sciencedirect.com/science/article/abs/pii/S0263786300000442?via%3Dihub (cit. on p. 9).

Hill, J., Thomas, L., & Allen, D. (2000). Experts' estimates of task durations in software development projects. *International journal of project management*, *18*(1), 13–21. https://www.sciencedirect.com/science/article/abs/pii/S0263786398000623?via%3Dihub (cit. on pp. 10, 33).

Hossin, M., & Sulaiman, M. N. (2015). A review on evaluation metrics for data classification evaluations. *International journal of data mining & knowledge management process*, *5*(2), 1. https://www.aircconline.com/ijdkp/V5N2/5215ijdkp01.pdf (cit. on p. 14).

Instituto Português da Qualidade, P. (2021). *Np en 13306:2021 manutenção – terminologia da manutenção* [PDF]. Instituto Português da Qualidade. (Cit. on p. 5).

Li, J., Yin, Y., Lafond, D., Ghasemi, A., Diallo, C., & Bertrand, E. (2023). Integrated human-ai forecasting for preventive maintenance task duration estimation. *International Conference on Machine Learning, Optimization, and Data Science*, 3–18. https://doi.org/10.1007/978-3-031-53966-4_1?urlappend=%3Futm_source%3Dresearchgate.net%26utm_medium%3Darticle (cit. on p. 8).

Lishner, I., & Shtub, A. (2022). Using an artificial neural network for improving the prediction of project duration. *Mathematics*, *10*(22), 4189. https://doi.org/10.3390/math10224189?urlappend=%3Futm_source%3Dresearchgate.net%26utm_medium%3Darticle (cit. on pp. 10, 20, 33).

Lourenço, J. M. (2021). *The NOVAthesis LaTeX Template User's Manual*. NOVA University Lisbon. https://github.com/joaomlourenco/novathesis/raw/main/template.pdf (cit. on p. ii).

Nafkha, R. (2016). The pert method in estimating project duration. *Information Systems in Management*, *5*. https://www.semanticscholar.org/paper/The-pert-method-in-estimating-project-duration-Nafkha/55e2eb6f82d11b4dad7454e20faa4079d280c122 (cit. on p. 19).

Pelt, M., Apostolidis, A., de Boer, R. J., Borst, M., Broodbakker, J., Jansen, R., Helwani, L., Patron, R., & Stamoulis, K. (2019). *Data mining in mro: Centre for applied research technology*. Amsterdam University of Applied Sciences. https://doi.org/10.13140/RG.2.2.22325.29920 (cit. on p. 5).

Pelt, M., Stamoulis, K., & Apostolidis, A. (2019). Data analytics case studies in the maintenance, repair and overhaul (MRO) industry (S. Pantelakis & C. Charitidis, Eds.). *MATEC Web of Conferences*, *304*, 04005. https://doi.org/10.1051/matecconf/201930404005 (cit. on pp. 10, 12).

Pospieszny, P., Czarnacka-Chrobot, B., & Kobylinski, A. (2018). An effective approach for software project effort and duration estimation with machine learning algorithms. *Journal of Systems and Software*, *137*, 184–196. https://www.sciencedirect.com/science/article/abs/pii/S0164121217302947 (cit. on pp. 10, 11, 14, 33).

Rengasamy, D., Morvan, H. P., & Figueredo, G. P. (2018). Deep Learning Approaches to Aircraft Maintenance, Repair and Overhaul: A Review [ISSN: 2153-0017]. *2018 21st International Conference on Intelligent Transportation Systems (ITSC)*, 150–156. https://doi.org/10.1109/ITSC.2018.8569502 (cit. on pp. 6, 11, 34, 35).

Riahi, V., Hassanzadeh, H., Khanna, S., Boyle, J., Syed, F., Biki, B., Borkwood, E., & Sweeney, L. (2023). Improving preoperative prediction of surgery duration. *BMC Health Services Research*, *23*(1), 1343. https://doi.org/10.1186/s12913-023-10264-6 (cit. on pp. 12, 33).

Sheikhalishahi, M., Pintelon, L., & Azadeh, A. (2016). Human factors in maintenance: A review. *Journal of Quality in Maintenance Engineering*, *22*(3), 218–237. https://doi.org/10.1108/JQME-12-2015-0064?urlappend=%3Futm_source%3Dresearchgate.net%26utm_medium%3Darticle (cit. on p. 8).

Sousa, A. O., Veloso, D. T., Gonçalves, H. M., Faria, J. P., Mendes-Moreira, J., Graça, R., Gomes, D., Castro, R. N., & Henriques, P. C. (2023). Applying machine learning to estimate the effort and duration of individual tasks in software projects. *IEEE Access*. https://doi.org/10.1109/ACCESS.2023.3307310 (cit. on pp. 11, 14, 20, 33).

Tyagi, K., Rane, C., Sriram, R., & Manry, M. (2022). Unsupervised learning. In *Artificial intelligence and machine learning for edge computing* (pp. 33–52). Elsevier. https://www.scirp.org/reference/referencespapers?referenceid=3862208 (cit. on pp. 12, 18).

Vencel', M., Knap, M., Staňo, F., Jaroslav, M., & Vereš, M. (2022). Airline aircraft maintenance requirements, development of flight planning and maintenance planning models. https://jogsc.com/pdf/2023/1/airline.pdf (cit. on p. 7).

A

Appendix 1

Table A.1: Primary skill codes in the aircraft maintenance

Skill Types	Description
Aircraft Painting	Applies protective coatings and paint to aircraft surfaces, enhancing appearance and providing corrosion protection, requiring knowledge of various paint types, application techniques, and safety regulations.
Aircraft Cleaning	Ensures both the interior and exterior of the aircraft are clean and presentable, playing a role in passenger comfort, hygiene, and routine inspections.
Avionics	Focuses on maintaining and repairing electronic systems, including navigation, communication, and instrumentation, ensuring all components comply with aviation standards.
Cabin Maintenance	Ensures the upkeep of the aircraft's interior, maintaining seating, lighting, entertainment systems, and emergency equipment to ensure passenger comfort and safety.
Category B1 – Maintenance Certifying Technician (Mechanical)	Privileges: Permitted to issue certificates of release to service following maintenance on aircraft structures, powerplants, mechanical and electrical systems. This includes replacing avionic line replaceable units requiring simple tests to ensure serviceability.
Category B2 – Maintenance Certifying Technician (Avionic)	Privileges: Authorized to issue certificates of release to service following maintenance on avionic and electrical systems.
Composite Materials	Focuses on the repair and maintenance of composite structures like carbon fiber and fiberglass, ensuring they meet strength and durability requirements.
Mechanics	Responsible for inspecting, repairing, and maintaining mechanical components of an aircraft, including engines, landing gear, and hydraulic systems.
Non-Destructive Testing (NDT)	Utilizes techniques like ultrasonic, radiographic, and magnetic particle inspection to assess the integrity of aircraft components without causing damage, crucial for detecting flaws or defects.
Sheet Metal	Involves fabrication, repair, and modification of an aircraft's metal structures, including tasks like riveting, welding, and ensuring structural integrity.

I

Annex 1

Pie charts of airplane maintenance interventions by aircraft location for each aircraft type.

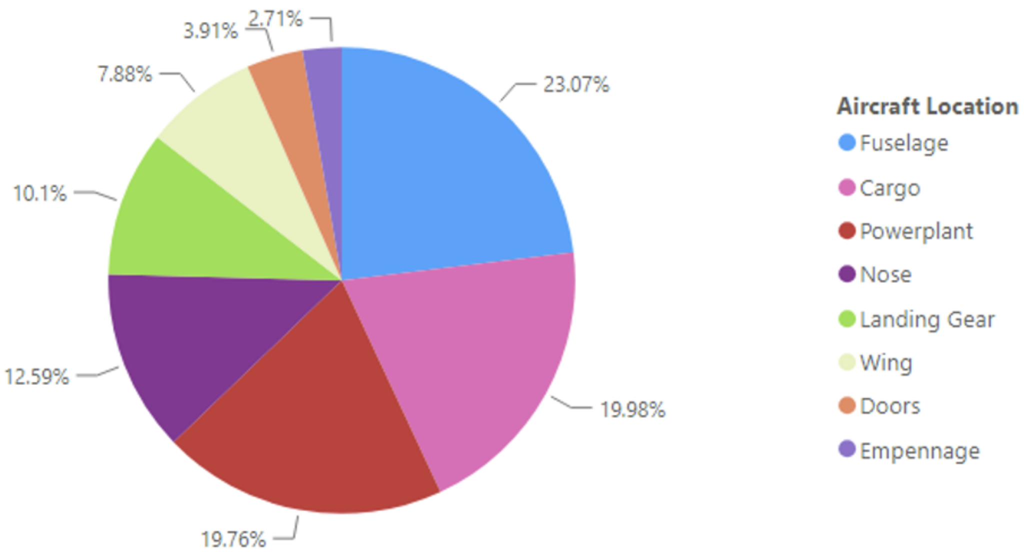


Figure I.1: Pie Chart for aircraft A350-941CJ type.

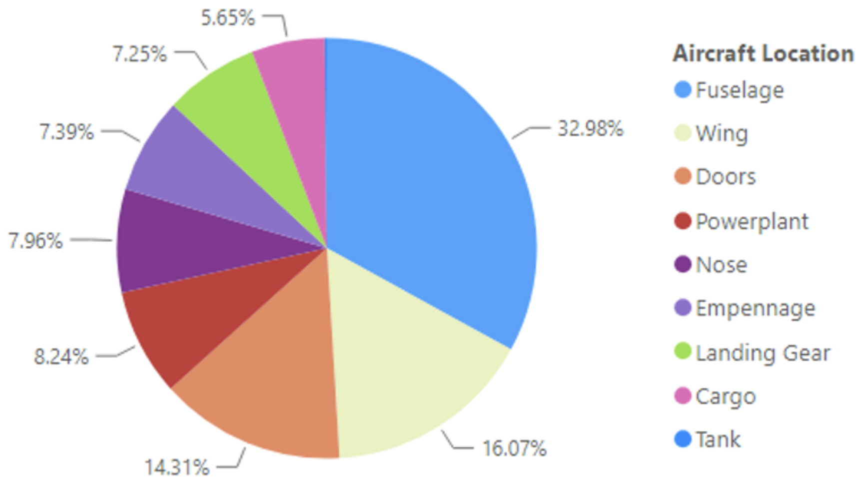


Figure I.2: Pie Chart for aircraft B777-200 type.

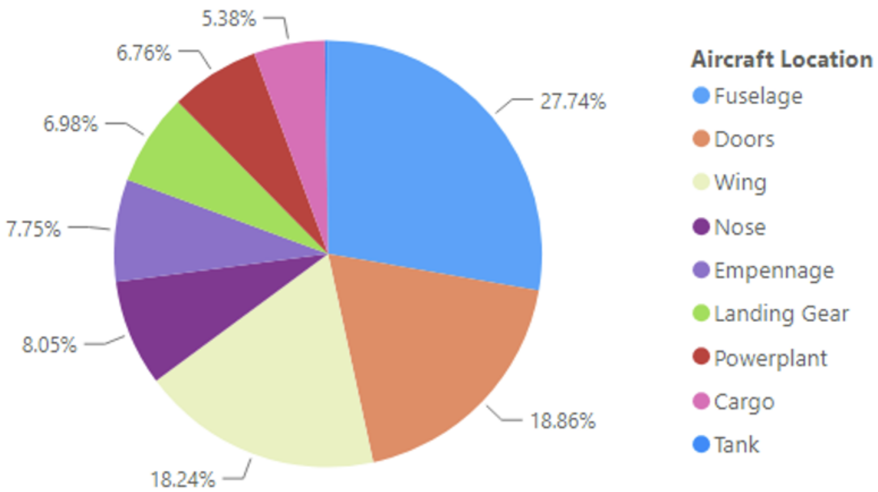


Figure I.3: Pie Chart for aircraft B777-300/328ER type.

This is to certify that:

Project No.: **STAT2025-3-127602**

Project Title: **Study of patterns in the aircraft Maintenance Repair and Overhaul using a data driven approach**

Principal Researcher: **Maria Helena Abreu Luz**

according to the regulations of the Ethics Committee of NOVA IMS and MagIC Research Center this project was considered to meet the requirements of the NOVA IMS Internal Review Board, being considered **APPROVED** on 14/03/2025.

It is the Principal Researcher's responsibility to ensure that all researchers and stakeholders associated with this project are aware of the conditions of approval and which documents have been approved.

The Principal Researcher is required to notify the Ethics Committee, via amendment or progress report, of

- Any significant change to the project and the reason for that change;
- Any unforeseen events or unexpected developments that merit notification;
- The inability of the Principal Researcher to continue in that role or any other change in research personnel involved in the project.

Lisbon, 14/03/2025
NOVA IMS Ethics Committee
ethicscommittee@novaims.unl.pt

Data with Purpose.

Campus de Campolide, 1070-312 Lisboa, Portugal
novaims.unl.pt