Data analytics case studies in the maintenance, repair and overhaul (MRO) industry

Maurice Pelt^{1,*}, Konstantinos Stamoulis¹, and Asteris Apostolidis²

¹Aviation Academy, Faculty of Technology, Amsterdam University of Applied Sciences (AUAS)
Amsterdam, Netherlands

²KLM Royal Dutch Airlines, Netherlands

Abstract. Data analytics seems a promising approach to address the problem of unpredictability in MRO organizations. The Amsterdam University of Applied Sciences in cooperation with the aviation industry has initiated a two-year applied research project to explore the possibilities of data mining. More than 25 cases have been studied at eight different MRO enterprises. The CRISP-DM methodology is applied to have a structural guideline throughout the project. The data within MROs were explored and prepared. Individual case studies conducted with statistical and machine learning methods, were successfully to predict among others, the duration of planned maintenance tasks as well as the optimal maintenance intervals, the probability of the occurrence of findings during maintenance tasks.

1 Introduction

The aircraft maintenance process is often characterized by unpredictable process times and material requirements. This problem is compensated for by large buffers in terms of time, personnel and parts, resulting in higher cost. Furthermore, traditional preventive maintenance policies often result in components replacements before the end of life. This increases part costs and conflicts with the growing need for sustainable operations.

Data analytics seems to be a promising way to tackle the problem of unpredictability in MRO by unleashing valuable information from the massively growing amount of data (Sahay, 2012)

An applied research project was organized across 25 case studies for eight different aviation MRO companies (Pelt, 2019). Medium sized MRO companies see a lot of data science activity at the larger MRO players and OEMs. The amount of data and applications is growing fast. However, the smaller companies lack the resources and expertise. The main research objective was to decrease maintenance costs and aircraft downtime at medium sized MROs using fragmented historical maintenance (and other) data.

The current paper presents the results from representative case studies that reflect the typical challenges of MRO companies and show the tested data analytics methods.

2 Methodology

2.1 Scope of predictive maintenance

Predictive maintenance is nowadays often considered superior compared other maintenance strategies. With this strategy maintenance is performed timely but not too early and can be planned in advance. Data analytics for predictive maintenance can be performed at different aggregation levels:

- Operation: On a fleet or MRO operation level typical predicted parameters are: The maintenance resources, fleet availability, utilization, working capital and lead times.
- Aircraft: On the aircraft level turnaround time and maintenance costs are of interest.
- System: For the system level, the probability that maintenance is required, delivery speed and accuracy and specific resources, such as manhours and spare parts, are important parameters to predict.
- Part: The predictions on part level are often directly related to failure mechanisms. To make better products OEMs need to be understand and address failure mechanisms. Related parameters are mean time between failure (MTBF), mean time to repair (MTTR), failure pattern and remaining useful life (RUL).

Models process data from many sources into maintenance predictions on different aggregation levels: Each level improves predictions of the level above. For example, the prediction of the maintenance duration of a whole aircraft will become more accurate if detailed predictions of the remaining useful life (RUL) of its parts are present.

For the MRO companies in the present study the medium levels of maintenance prediction were the most important.

2.2 CRISP-DM approach

Data mining is a logical process that helps researchers search through large amounts of data in order to find interesting insights hidden within. The goal of this technique is to use a sequence of phases to find previously unknown patterns.

Our approach for this project is based on the Cross Industry Standard Process for Data Mining methodology, commonly known by its acronym CRISP-DM (Chapman, et al., 2000). A 2014 survey in the Knowledge Discovery community (Piatetsky G. , 2014) showed that CRISP-DM is the most used data mining method. This approach was used to investigate all of this project's research questions.

Fig. 1. CRISP-DM approach

2.3 Data sources in Aviation

The MRO industry is characterized by a variety of data sources. The data sets selected in each case depend on the initially-defined data mining goals. There must be a plausible connection between the data sets and the data mining case. The data understanding phase proceeds with activities that help researchers become familiar with the data, identify data quality problems, discover first insights into the data, and detect interesting subsets to form hypotheses for hidden information. This task is performed in principal by visualizing the data and examining trends and patterns.

Data from many sources are input for maintenance predictions, see Fig. 2. Most MROs still use few of these data

Fig. 2. Data from many sources are input for maintenance predictions

This study used three main categories of data sources:

1. Maintenance data from MRO
2. Flight recorder data
 - Operational data from the Flight Data Recorder and Quick Access Recorder
 - Sensor data and messages from the Aircraft Health Management system
3. External data
 - Benchmarking data gathered from a large group of similar systems or processes
 - Weather data
 - Aircraft position data (such as ADS-B)

2.4 Access to data sources

A number of technical and non-technical obstacles can present themselves while researchers are assembling reliable data sets. Standardized data availability is a basic requirement for data analysis. MRO companies often rely on multiple (IT) systems for data collection and storage which results in fragmented and non-comparative data sets. At the same time, as data becomes increasingly valuable for all parties involved, medium sized MROs have lower bargaining power compared to OEMs and other important contributors in the data pipeline and supply chain. Finally, MRO SMEs have limited resources for personnel specialized in data science.

Often the following data quality problems are present: Missing values, outliers, datasets not accessible or not available, datasets incomplete, data mis- interpretation (metadata), and errors in values.

It takes a lot of tedious data preparation work to bring the data quality to a sufficient level for modelling.

Another issue is the access to data and/or the rights to use the data. Many creators, users, and owners of data were found in the case studies. An example of a data distribution in Aviation MRO is shown in Table 1.

Table 1. Example of Creators, Users and Owners of data sources

Stakeholder	Available data							
	Flight data		Maintenance data				External data	
	Operations data	Aircraft Health Monit	ERP	MPD	Jobcard	Form 1	OEM maintenance documentation	External sources
Airline	C U O	U?	C U O					U?
Aircraft owner	U O		U?				U?	U?
Airworthiness manager (CAMO)	U?		C U O	C U O			U	U?
OEM of aircraft, engine or other		U O					C O	U?
MRO company (Part-145)	U?	U?	C U O	U O	C U O	C U O	U	U?
MRO Support /tooling		U?	C U O	U O	C U O	C U O	U	U?

2.5 Data analytics in MRO

Maintenance used to start with expert knowledge. Experienced operators could tell whether a machine needed maintenance from a variety of factors – noise, vibration, heat, moisture, and so on. Then came physical models, which explained why degradation happened and when to expect a part’s end of life. Statistical process control was then added to evaluate whether a given machine was functioning within pre-defined limits. In recent years data-driven methods have been added, using machine learning.

Physical models can help achieve higher prediction accuracy. Surprisingly, just a few case studies tested this approach. In general, the most accurate predictions will be achieved if the three models are combined: expert models, physical models, and data driven models.

The MRO case studies did highlight a major challenge in terms of prediction. An aircraft has many different parts with different failure mechanisms. Some are caused by high forces, others to wear, metal fatigue, temperature changes, or electricity-related issues. This requires the measurement and analysis of many different parameters. As a result, MROs will need many data mining applications, each with their own specialized applicability to different areas of aircraft maintenance.

Based on the cases studies, associated with the current project, a framework was designed, as illustrated in

Fig. 3. This MRO Analytics framework is not completely aligned with other structures described in literature (Sayad, 2017), but it proved to be useful to explain the different groups that were found in the case studies.

Fig. 3. MRO Data Analytics framework

3 Case studies with visualization and statistics

In most cases, data analytics in MRO starts with the more traditional methods such as visualization and statistics. These methods increase the understanding of the data and deliver already much value for maintenance improvements.

Almost all described cases in this and next chapter are programmed in R or Python.

Case: Optimal aircraft tires replacement

Commercial airlines visually inspect the condition of aircraft tires after each landing. The decision to replace them is based on the observed condition of the tires. This case concerns the estimation of the remaining lifetime of tires of various types. The available data sets consisted of independent variables such as total air temperature, reverse thrust settings, deceleration rate and landing weight per landing location (airport), as well as a dependent (target) variable that indicated the wear level of the tires. A linear regression model was applied to predict the remaining useful flight cycles of the tires. This allowed a specification of the optimal interval for tire replacement.

Case: Aircraft maintenance duration prediction

MRO planners commonly estimate the required manhours for either planned or unplanned MRO tasks caused by inspection findings. Accurate prediction can improve the efficiency and the profitability of MRO processes and shorten turnaround time for aircraft operators. In this study, task duration prediction was enhanced by introducing forecasting methods and distribution functions:

1. Best performing time series forecasting model e.g. Autoregressive Integrated Moving Average (ARIMA) and Exponential Smoothing models.
2. Best fitting statistical distribution to predict maintenance duration. After comparison of methods the researchers found the K-S method to be most accurate at finding the correct statistical distribution. The target statistical distribution functions were: Beta, Exponential, Gamma, Gaussian and Lognormal.

Fig. 4. Interactive dashboard for duration prediction

Results obtained from the models were presented in an interactive dashboard (see Fig. 4).

4 Case studies with machine learning

Machine learning in MRO is an application of artificial intelligence where historical maintenance data and other data are used to recognize patterns and to make predictions about maintenance without being explicitly programmed. The predictions become more accurate over time when more MRO data become available. Commonly used techniques are classification and clustering.

Case: Text mining to analyze maintenance reports

Maintenance records often contain textual explanations written by mechanics concerning findings and repairs – information that can potentially be used to predict failures and propose solutions. A systematic analysis of these text records is time-consuming. However, this task

can be performed through automated natural language processing (NLP) systems. The researcher extracted text records from the AMOS maintenance management system and then processed them to find interesting patterns. He assigned pre-defined thematic categories to the text using the K-nearest neighbour text classification Machine Learning algorithm, and calculated similarity. This procedure found relevant information concerning failures and solutions with an accuracy score of more than 75%.

Case: Causes of low fleet availability in high season

The MRO challenge for an MRO provider was a significant and unexplained decrease in fleet availability during the high season. The researcher visualized the number of Unscheduled Ground Time (UGT) events in the high and low seasons per ATA chapter. Six of the ATA chapters were selected for further analysis.

To find the exact reason for the difference between high and low seasons, the researcher used Support Vector Machine (SVM) analysis to see whether an operation-disturbing ATA chapter could be predicted based on parameters such as air temperature, cycles and humidity. It was found that a major cause of unplanned ground time – the replacement of coalescer sacs – was related to humidity and temperature. Based on these insights, the researcher proposed a new maintenance schedule for coalescer sac replacements.

Case: Choose the best performing machine learning algorithm to predict unplanned maintenance tasks

An aviation analytics software provider wants to increase the efficiency of resource planning for base maintenance tasks. From a research point of view, this efficiency can be improved by predicting whether a certain task card will become a finding or not. During the modelling phase, the performance of a wide variety of classification algorithms are evaluated and thereafter, the algorithm with the best performance is selected as the final model. Fig. 5 demonstrates the performance of multiple learning algorithms on a given maintenance task card. The Receiver Operating Characteristic (ROC) curve is used to compare the effectiveness of the individual models.

Fig. 5. ROC curves best performing algorithms to predict unplanned maintenance

Case : Causes of a reduced delivery reliability in aircraft component maintenance

This case was commissioned by the RNAF. Their MRO challenge was to identify the causes of low delivery reliability in component maintenance and predict the situations in which this would happen. The researcher analysed the dependence of multiple variables within delivery reliability. This dependency analysis included statistics such as Chi-squares, and data visualizations such as mosaic plots. Finally, the researcher used the dependent variables to develop a decision tree by to predict if an order was more likely to be on time or too late.

Case : Predict aircraft component failures model using external data sources for flight path and weather

An aviation analytics software provider aims to forecast unplanned maintenance for legacy aircraft by using a combination of flight data, maintenance records and/or airworthiness records of legacy aircraft. In this specific case, sensitive flight data had to be excluded from the analysis and be replaced by data deriving from public domain sources. This challenge was tackled by focusing on ADS-B and weather data, obtained from open domain flight-tracking platforms. Predicted aircraft failures were based on anomalous flights. For example, a hard landing might eventually result in landing gear failure and can be considered as an anomalous case. Two different clustering algorithms were used: Density-Based Spatial Clustering of Applications with Noise (DBSCAN) and k-means clustering. Fig. 6 illustrates a clustering example of the take-off phase of a legacy aircraft

Fig. 6. DBSCAN clustering of take-off phase

5 Conclusions and recommendations

The case studies proved the value of statistical and machine learning methods. The proof of concepts showed potential improvements in:

- Aircraft uptime: Optimal and accurate planning
- MRO costs: Reduction of idle time and over-processing
- Part costs: Optimal use of components remaining life

Full implementation of data analytics is a long-term development that definitely requires improvements of processes, infrastructure, knowledge, and data access at medium sized MROs. There exists not a standard data analytics solution for Aviation MRO companies. In general data analytics at MROs is in an early maturity stage. The data analytics strategy should be adjusted to the maturity level and the role in the supply chain. Start with applications that target failures with significant impact.

Clearly, aviation maintenance companies are underutilizing the potential of data, due mainly to a focus on compliance rather than prediction. The availability of external data from airline operators, suppliers and OEMs is hampered by confidentiality and ownership issues. Time-consuming data preparation work was often needed to make the data quality acceptable. MRO companies should negotiate with data owners such as OEMs and airlines and search other sources. On an industry level more attention should be paid to procedures and technical solutions for secure data sharing.

In many case studies the prediction accuracy is compromised by the lack of relevant inputs. For example, sensors are missing to measure certain characteristics that are related to the failures of the components in an aircraft. Another explanation is the relatively small

sample size of similar events due to the wide variation in components and maintenance tasks. Despite the growing number of sensors, inspections by maintenance engineers remain indispensable. Data driven predictions can guide inspection activities and improve efficiency.

Correlations were often investigated. However, it is more powerful to measure parameters that have a causal relationship with defects. Companies should combine data-driven models with expert and failure models to create higher prediction accuracy. Data driven predictions describe probabilities, and this requires a different way of thinking compared to the unambiguous outcomes of inspections. Machine learning algorithms are often not transparent, therefore it takes time to rely on them.

Data visualization is a natural starting point in data analytics and has proven to be very useful for MRO companies as they start data mining. Next is prediction and machine learning. Focused applications that target real problems obtain the best results.

The human factor is very important in data analytics in MROs. Companies should introduce data scientists into the organization. It is important to train operational management and mechanics, because they generate the data and use the new information sources to improve their work. Interaction between data scientists and shop floor should be organized.

All these DM case studies highlight the prospects for optimized and sustainable MRO processes. Overall, the 'Data Analytics in MRO' process optimization research project delivered promising proofs of concept and pilot implementations. It created valuable insights and recommendations about the feasibility and effectiveness of modern data science techniques at medium-sized aviation maintenance companies.

References

1. Sahay, A. (2012). Leveraging information technology for optimal aircraft maintenance, repair and overhaul (MRO). Oxford: Woodhead Publishing
2. Chapman, P., Clinton, J., Kerber, R., Khabaza, T., Reinartz, T., Shearer, C., & Wirth, R. (2000). CRISP-DM 1.0 Step-by-step data mining guide. The CRISP-DM consortium.
3. Boer, R. et al. (2015). Maintaining your competitive edge. Amsterdam: Centre for Applied Research Technology of the Amsterdam University of Applied Sciences.
4. Pelt, M. et al. (2019). Data Mining in MRO. Amsterdam: Aviation Academy Research Programme, Amsterdam University of Applied Sciences. ISBN: 9789492644114
5. Piatetsky, G. (2014). CRISP-DM, still the top methodology for analytics, data mining, or data science projects. Retrieved from www.kdnuggets.com.
6. Sayad, S. (2017). An introduction to data mining. Retrieved from www.saedsayad.com.
7. Tinga T., Z. E. (2013). Predictive Maintenance of Military Systems Based on Physical Failure Models. Chemical Engineering Transactions. Vol. 33.
8. Pelt, M., Stamoulis, K. (2018). Data Mining for aircraft Maintenance Repair and Overhaul (MRO) process optimization. Istanbul: Proceedings Book of ISATECH – 2018, ISBN: 9786056864049
9. Apostolidis, A., Stamoulis, K., Pelt, M. (2020). Aviation Data Analytics in MRO Operations: Prospects and Pitfalls. Proceedings, 66th Annual Reliability & Maintainability Symposium (RAMS 2020)