Engine failure forecast model for saving the operating cost: a case of American Airlines

Nicolás Rosal Gómez
UCL School of Management, University College London, London, United Kingdom, zceinro@ucl.ac.uk

Abstract: We quantify the impact of applying an engine failure forecast model on the airlines' maintenance operating cost. We first compare the performances of existing engine failure forecast models and find that: (a) a stacking model with an XGBoost estimator performs best with 99.9% accuracy and 0% no-fault-found rate, and (b) engine pressure ratio (*epr*) is a key variable for predicting the engine failure. We then use American Airlines as a case study and quantify that applying a failure forecasting model can save up to $1.16 billion per year on maintenance expenses.

Key words: operation management, failure forecasting, maintenance operations, unscheduled events, American Airlines, machine learning

1. Introduction

In the last decade, airline's maintenance has been moving away from corrective and preventive towards predictive strategies. With 85% of Boeing's equipment failures happening at random regardless of inspection and service (Golightly, 2019), the use of preventive measures fall short to solve the problem of aircraft downtime and unexpected maintenance costs. Indeed, the ARC Advisory Group reported that traditional equipment monitoring fails to prevent 82% of equipment failures (Golightly, 2019).

On the contrary, predictive techniques can identify equipment bound to fail in the near future and provide enough time to schedule a required maintenance (Emorphis Technologies, 2019). Similar practices have been widely adopted in the aerospace industry, thanks to the advent of new IoT technologies (FT Maintenance, 2019), which enable the collection of real-time data from equipment. Robert Wright (2017) exemplified such a trend with Roll-Royce's recent adoption of engine digital sensors that monitor temperature, fuel flow, air flow and pressure, providing data about aero-engine health. Nonetheless, one challenge with such technology is analyzing the large amounts of real-time data and making timely predictions. Manual predictive maintenance techniques are impractical (Redd, 2020) and fail to correctly identify several malfunctions in aircrafts, leading to unwanted failures, unscheduled maintenance and unplanned downtime.

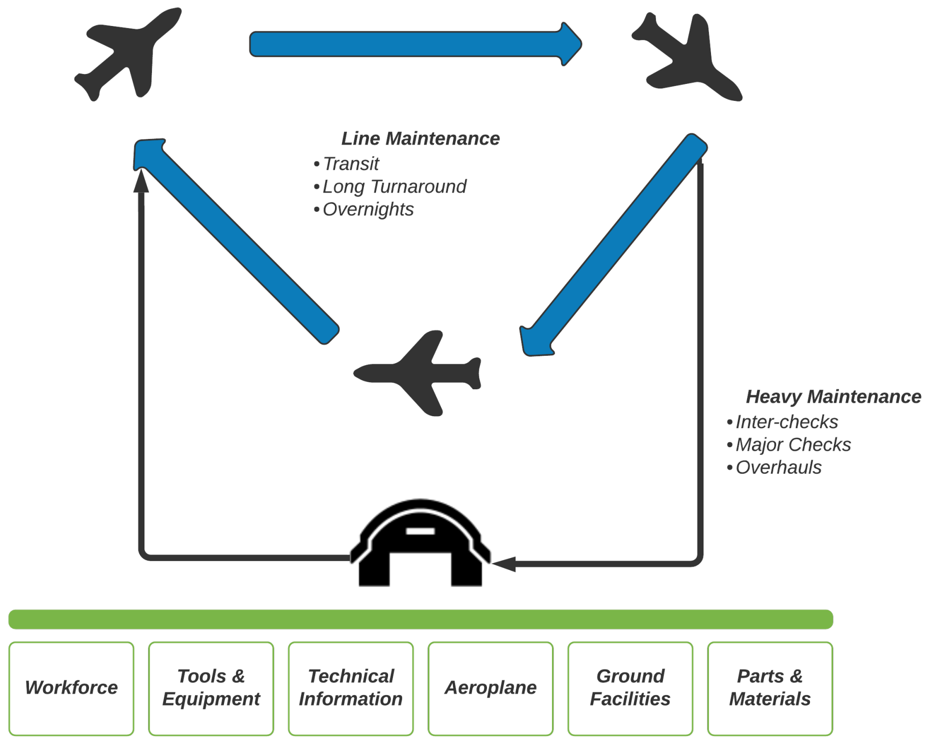

Figure 1 **Line and Heavy Maintenance (Salazar, 2015)**

Now, the introduction of new machine learning (ML) algorithms capable of dealing with ever-growing datasets present an interesting opportunity to improve these predictive maintenance strategies failure forecasting. As the amount of data generated by the global fleet continues to increase - from 2 million terabytes in 2016 to an estimated 98 million terabytes by 2026 (Hoyland et al., 2016) - the need for machine learning-based processes became more apparent. The increase in velocity of incoming data as a result of advancing information technologies (De Bree, n.d.) (e.g. IoT sensors) made manual analysis of data impractical (De Bree, n.d.) for maintenance teams. Thus, organisations identified ML algorithms' capacity of dealing with big data as an invaluable resource to improve their maintenance processes and help combat the existing problem with unscheduled maintenance and downtime.

Airline's maintenance is an important aspect of the companies' operational success, due to its impact on costs and service delivery. Maintenance, ground handling, aircraft acquisitions (influenced by downtime), in-flight services, and call centers represent up to 45 percent of an airline's cost structure (Doig et al. 2003). Hence, improving maintenance operations can help airlines significantly reduce an important percentage of its costs and, indirectly, ameliorate service delivery through reduced downtime.

Researchers have studied how ARMA (Li and Kang, 2008; Li et al., 2011; Singh, 1994), ARIMA (Ho and Xie, 1998), and even neural networks (Kozlowski, 2003) models can be used to predict remaining useful lifetime of various different aircraft parts. Nonetheless, few studies have explored how the classification power of various ML models can be combined to improve predictive power. Moreover, research often lack quantified impact figures and actionable recommendations that airlines can use to understand the value of ML-based technologies and best implementation strategies.

This paper contributes to the existing literature and informs practitioners by exploring multiple failure forecasting machine learning models with the aim of finding the best model that would help American Airlines (AA), the biggest airline in the world, overcome its challenges with unscheduled maintenance. Using AA's finances as a case study to quantify the impact of this technology we provide insight for other airlines on how ML can be used to improve maintenance operations. We identify a stacking model with an XGBoost estimator as the best performing model with an accuracy of 99.9%, F1 score of 99.8%, and 0% no-fault-found rate. We believe our research can further support future research to quantify the potential impact of the technology using company-collected data.

The remainder of the paper is organized in the following manner. In the *American Airlines and Its Unscheduled Maintenance and Downtime*, we describe the existing problem with unscheduled events faced by American Airlines, together with a quantification of its impact. Then, in the *The Data* section, we explore dataset used for developing the failure forecasting machine learning models, highlighting how the target variable was created. Next, in the *Data Preparation* section, we present the data preprocessing steps followed before modelling. In the *Methodology* section, we move on to introduce the machine learning models utilised for analysis and the metrics used to evaluate said models. Next, in the *Results* section and *Impact*, we present a summary of the results together with a quantification of the potential impact of this technology on American Airlines's maintenance operations and an optimal deployment process for the technology. Finally, the *Discussion and Conclusions* section explores the implementation challenges attributed to machine learning solutions and explore the optimal deployment process for American Airlines to adopt ML in their maintenance operations.

2. American Airlines and Its Unscheduled Maintenance and Downtime

American Airlines (AA) is the biggest airline in the world, offering a network of 6,700 per day to 350 destinations in more than 50 countries (Cleverism, 2020) with its fleet size of 1,317 airplanes (Mazareanu, 2019). Together with its regional partners [1], the company aims to provide high

[1] Envoy Air, Piedmont Airlines, and PSA Airlines, and OneWorld partner airlines (Centre for Aviation, 2020)

quality, reliable commercial, cargo, and business air travel services across domestic and international destinations (Cleverism, 2020). In 2019 alone, the airline carried 215.2 million passengers (Mazareanu, 2020), generating $31.6 billion in passenger revenues (American Airlines, 2020). This accounted for 91.9% of total revenues for the fiscal year, suggesting the company heavily relies on this sector. The company reported a net profit of $1.7 billion for the fiscal year ending in 2019 (American Airlines, 2020).

Following the travel restrictions imposed during the COVID-19 pandemic, the global passenger demand dropped 72.8% below September 2019 levels (International Airport Review, 2006), reaching an all-time low. Consequently, and keeping into consideration the dependence on passenger revenues, the company's operating revenue decreased from $10.9 billion in 2019 to $2.5 billion in 2020 (a 61.4% decrease) (American Airlines, 2020). Even with operating expenses also decreasing by 33.9% (American Airlines, 2020), mainly due to a decrease in fuel costs, selling expenses and other expenses, American Airlines reported a net loss of $6.7 billion. This means that, in just the 9 months ending on September 30 of 2020, the company lost 3.95 times the net profit of the previous year.

Once international and domestic travel go back to normal, American Airlines needs to improve their operations to maximise air travel while maintaining low costs. With efficient vaccines being distributed, pockets of international connectivity can be created as 'travel bubbles' open up (Harper, 2020), and this might be sooner than expected. To be capable of servicing these areas with a smaller fleet, the airline needs to develop strategies that reduce the downtime of their remaining 1,167 aircrafts, enabling the generation of constant passenger revenues. If reducing costs is the ultimate goal, the costs of maintenance and downtime needs to be reduced compared to previous years. Prolonged storage effectively impacts an aircraft airworthiness (Katoky et al., 2020), even with maintenance and protection plans in place. Hence, after months storing aircrafts, unscheduled maintenance can become a problem that affects profit margins in the near future if not addressed properly. Reducing unscheduled maintenance and downtime, together with the current plan to boost liquidity and conserve cash, will enable the company to capitalise on the upcoming passenger demand increase in 2021. The following subsections explore the tangible costs of unscheduled maintenance and downtime for American Airlines.

2.1. Costs of Unscheduled Maintenance

Unscheduled events tend to be most expensive in maintenance costs due to their unplanned nature (Pariyar et al., 2013). Indeed, it can cost 3 to 9 times more than scheduled maintenance due to overtime and call outs costs together with the expenses of buying parts hastily (Solved FM, 2018). Resto as cited by Salazar et al. (n.d.) claims that between forty and sixty percent of all maintenance activities are unscheduled events. Out of all maintenance completed during an aircraft scheduled service as much as 50 percent can be attributed to non-routine tasks (Aungst et al., 2008). Assuming that an average of 50 percent of American Airlines's maintenance activities are unscheduled, the cost of these activities up to Q3 of 2020 are estimated at approximately $1.3 billion. More detail on how the cost was calculated can be found in the Appendix A.1. This is clearly an issue for the company; 6.2 percent of operational expenses can be attributed to non-routine maintenance tasks, and the forecast for the future is not looking any better.

Mazareanu (2018) estimates a 37.2% increase in overall unscheduled maintenance (UM) costs between 2017 and 2025, and an additional 45.7% increase between 2025 and 2035 (Mazareanu, 2018). Assuming there was a 4.7% per annum increase in the first interval and a 4.6% increase in the second, it is possible to forecast AA's non-routine maintenance costs to be $1.6 billions by the third quarter of 2025 and $2.3 billions by the third quarter of 2035. Without a proper strategy to avoid this, American Airlines will witness a 79.5% increase in unscheduled maintenance costs in the next 15 years.

To break down the cost of unscheduled maintenance for AA, we focus on exploring the cost of unplanned A and C checks; these account for 40 to 50 percent of overall maintenance costs (University of Westminster, 2008). B checks are often merged in A checks, and D are rarely done since it is cheaper to replace the aircraft at that point. The latter, as part of heavy maintenance, are the most challenging due to the magnitude and sophistication of a major aircraft overhaul, and their activities can have flow-on effects on configuration management and component management (Aungst et al., 2008).

An A check should occur every 500 flight hours (University of Westminster, 2008), which translates to approximately 8 checks per aircraft per year; assuming a plane operates 12 hours per day on average (Plane Stats, 2017). Thus, AA should perform an estimate of 10,536 A-checks in a year. Without unscheduled maintenance tasks, the company should incur between $4.9 millions and $9.9 millions in costs for A checks. Nevertheless, referring back to Aungst et al.'s (2008) claim, 5,268 of these checks are likely to come from unscheduled events. Hence, the real cost of unscheduled maintenance A checks are estimated to lie between $17.3 millions and $34.6 millions. This allows to suggest that the company is currently paying approximately 3.5 times more for A checks due to unscheduled events. These calculations can be found in appendix A.2 and A.3, respectively.

Similarly, the company's C check costs were found to be significantly higher when unscheduled maintenance was considered. AA pays approximately 3.5 times more for C checks due to this type of event. These checks occur every 4000 to 6000 hours (University of Westminster, 2008), translating to approximately one check every 333 days to 500 days. This paper considers the lower bound in the following calculation, allowing to assume that AA performs 1,317 C checks per year. Once again, unscheduled maintenance is disregarded to find how much the airline would pay for scheduled C checks. This value amounts to $185.6 millions. Nonetheless, considering 658 of these checks will come from unscheduled events led to estimate the real cost of C checks at $649.2 millions. The previous calculations show that American Airlines might be losing $463.6 millions per year due to unscheduled A and C events. Considering this accounts for 35%+ of the total estimated cost of unscheduled maintenance shows the importance in creating strategies to more accurately predict failures to avoid unplanned events in aircrafts checks. These calculations can be found in appendix A.4 and A.5, respectively.

The probability of unscheduled maintenance per aircraft increases 0.8 percent with every one-year gain in age (RAND Corporation, 2006). Considering the average aircraft age in AA's fleet is 11.2 years (Plane Spotters, 2021), it can be estimated that the company's airplanes will be on average 9.8 percent more likely to experience unscheduled maintenance next year. Hence, the costs previously mentioned will also likely increase in the upcoming years due to aircraft aging. This further stresses the importance of creating better strategies to manage maintenance in order to avoid UM events.

2.2. Costs of Unscheduled Downtime

In addition to the costs of unscheduled maintenance, American Airlines has to deal with the costs of aircraft downtime. Airlines utilise a significant amount of resources to accurately plan and schedule maintenance checks beforehand to avoid unplanned disruption in service provision. This way, they are able to create an estimated budget for maintenance activities, including labour and material costs required to complete. Nevertheless, as previously mentioned, unscheduled maintenance occurs regularly. During heavy maintenance, an aircraft is temporarily taken out of revenue service and maintenance time is treated as downtime (Samaranayake and Kiridena, 2012). Thus, during unplanned maintenance tasks, aircrafts will face unscheduled downtime (UD) Similarly, even A and B checks, which are completed in line, are forced out of service and enter into unscheduled downtime until fixed.

UD can significantly affect flying schedules by delaying and/or even cancelling flights, if the time required to complete the unplanned activity is considerable (Salazar, 2015), i.e. if C and D checks are necessary. Delays and cancellations impact airlines' costs, creating additional operational expenses. On the one hand, a minute of the former can cost airlines $65, considering fuel & oil, crew, maintenance, ownership and other costs (Trefis, 2016). Basic salaries are constant, monthly leases are inevitable regardless of flight schedules (Saltoglu et al., 2016), and fuel cannot be re-purposed once loaded. From January 2019 to December 2019, 23 percent of American Airlines flights were delayed, with an average delay time of 67 minutes (Bureau of Transportation Statistics, 2020). Considering that AA operates approximately 2.45 million flights per year, nearly 6,700 flights per day (American Airlines Group, 2021), this translates to 562,500 flights delayed, 3.78 million minutes in delays, and $245.7 million in delay costs in just one year. On the other hand, cancellation costs are estimated at $5,770 per cancelled flight segment (Ausick, 2020). Taking into account the 2.1% flights cancelled between January 2019 and December 2019 (Bureau of Transportation Statistics, 2020), one can estimate the cancellation costs at $301.0 million.

Based on the latest complete yearly delay and cancellation data, and assuming that about 0.62% of delays to happen due to weather complications (Bureau of Transportation Statistics, 2020), American Airlines currently incurs in approximately $543.3 million in extra expenses due to delays and cancellations. This further evidences the impact of unscheduled maintenance and downtime on American Airlines finances.

The presence of tangible unplanned maintenance and downtime costs emphasises the importance of accurately predicting failure events in maintenance operations. In the following section, we present the data and machine learning algorithms that can help improve failure forecasting.

3. The Data

We used the Turbofan Engine Degradation Simulation Dataset, accessed through NASA's Prognostics Data Repository, which focuses on datasets that can be used for the development of prognostic algorithms (NASA, 2021). Considering the most common types of aircraft in American Airline's fleet - Boeing 737-800 and Airbus A321-200 - use turbofan engines (BAA Training, 2018; Skybrary, 2021), the data was selected as the optimal to develop failure forecasting technology for AA's aircraft engines. This data comprises four sets of engine degradation simulations carried out using C-MAPSS under different combinations of operational conditions and fault modes (NASA, 2021). This analysis focuses on the set FD003 which simulated 100 engine degradations under one operational condition, namely sea level, and two fault modes, namely high-pressure compressor (HPC) and fan degradation. Our study uses the FD003 train subset, with 24,720 observations, for the development of the model.

The chosen subset consists of 100 multivariate time series, one for each engine degradation simulation. Each series starts with the engine operating normally and ends when the faults led to

	T24	T30	T50	P15	P30	Nf	Nc	epr	Ps30	phi	NRf	NRc	BPR	htBleed	W31	W32
count	24720.0	24720.0	24720.0	24720.0	24720.0	24720.0	24720.0	24720.0	24720.0	24720.0	24720.0	24720.0	24720.0	24720.0	24720.0	24720.0
mean	642.5	1588.1	1404.5	21.6	555.1	2388.1	9064.1	1.3	47.4	523.1	2388.1	8144.2	8.4	392.6	39.0	23.4
std	0.5	6.8	9.8	0.0	3.4	0.2	20.0	0.0	0.3	3.3	0.2	16.5	0.1	1.8	0.2	0.1
min	640.8	1564.3	1377.1	21.4	549.6	2386.9	9018.0	1.3	46.7	517.8	2386.9	8099.7	8.2	388.0	38.2	22.9
25%	642.1	1583.3	1397.2	21.6	553.1	2388.0	9051.9	1.3	47.2	521.2	2388.0	8134.5	8.4	391.0	38.8	23.3
50%	642.4	1587.5	1402.9	21.6	554.0	2388.1	9060.0	1.3	47.4	522.0	2388.1	8141.2	8.4	392.0	39.0	23.4
75%	642.8	1592.4	1410.6	21.6	556.0	2388.1	9070.1	1.3	47.6	523.8	2388.1	8149.2	8.4	394.0	39.1	23.5
max	645.1	1615.4	1441.2	21.6	570.5	2388.6	9234.4	1.3	48.4	537.4	2388.6	8290.6	8.6	399.0	39.8	24.0

Table 1 Summary Statistics of NASA's Turbofan Engine Dataset (Variable Descriptions - Appendix B.1.)

system failure (Kaggle, 2019). Time in this analysis is measured in cycles; defined by segments such as start, idle, taxi, takeoff, climb, cruise, approach, landing, thrust reverse and shutdown (Federal Aviation Administration, 2009). For example, the first time series in the data starts in cycle 1 and ends in cycle 233 at failure. To mimic normal conditions, each engine starts with varying degrees of initial wear and manufacturing differences and the data is contaminated with sensor noise (Kaggle, 2019). Table 1 provides summary statistics of the dataset.

3.1. Target Variable

The model was developed to predict the time left to engine failure (future event). This time metric is known as remaining useful life, or RUL, in the maintenance industry. Procurement of materials for unplanned maintenance takes 20 days (Samaranayake and Kiridena, 2012). For this reason, the target variable was engineered to be a binary column that identified the engines with a remaining useful life (RUL) of 20 days or less (e.g. 1 equals risk of failure). This time frame will allow American Airlines to prepare for the maintenance activity at hand and perform repairs on time. Note that one cycle is assumed to equal one day for the purpose of this analysis.

To create the target binary column for each series, we first calculated the difference between the last cycle (cycle at failure) and the cycle of each observation (Kucherevskiy, 2020). This would show the RUL for the engine at that cycle. Then, if the calculated remaining useful life is smaller or equal to 20, the target value would equal 1. Otherwise, it equals 0. For example, for the first time series, the observation at cycle 200 has a remaining useful life of 33 cycles, which is larger than 20. Hence, the target variable would equal 0 for this observation. This eliminates the sequential nature of the data and creates a categorical variable to perform classification.

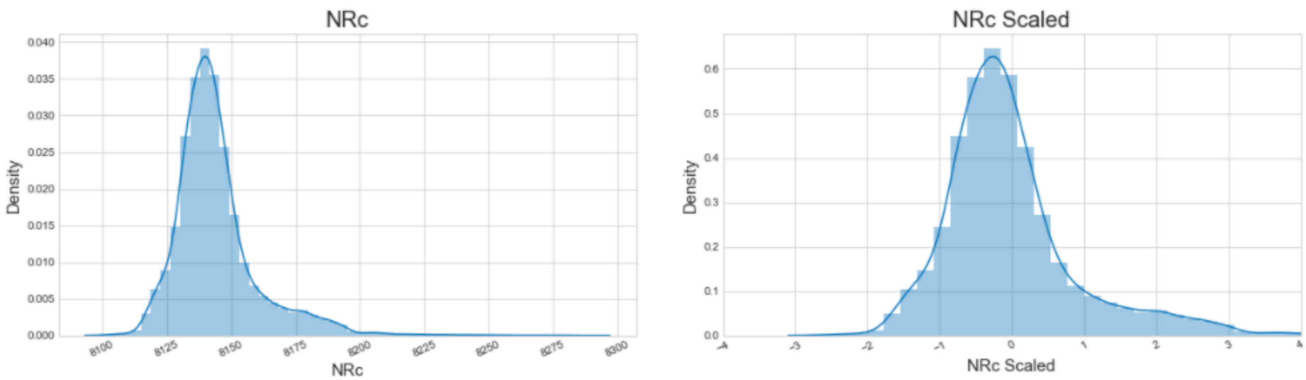

Figure 2 **Distributions of NRc (Left) and Scaled NRc (Right)**

4. Data Preparation

We normalised the data to have a mean of 0 and standard deviation of 1. Figure 2 shows an example distribution of scaled variables. After normalisation, we have addressed a significant data imbalance problem. There are about 14 times more data points for label 0 (healthy) than for label 1 (at risk of failure); 23,120 versus 1,600 respectively. In cases with extreme data imbalances, ML models will treat the minority class as outliers and will end up dropping them (Dangut et al., 2020), leading to a model that is only able to identify data points that belong to the majority class (e.g. healthy reading). This creates further problems as it leads to misleading accuracy scores. Thus, we have used upsampling and generative adversarial network (GAN)-generated synthetic data as preprocessing steps to deal with data imbalance prior to training ML models.

5. Methodology

After exploring and preparing the data, we have applied various machine learning models to predict engine failure. Each model has been trained and tested using the imbalanced data (original), balanced data with upsampling technique, and GAN-balanced data. Testing the combination of different models and different datasets would help identify the best algorithm and data balancing technique for the problem at hand.

Firstly, we have trained and tested the following models: logistic regression, decision tree, random forest, XGBoost, and Neural Network. Then, the F1 score, recall, precision, and accuracy metrics were used to identify the best performing model and data imbalance preprocessing technique combinations. Once the performances were evaluated, we used a feature importance analysis to determine what features would aid maintenance engineers at accurately identifying an engine failure developing without the need of machine learning models. Finally, we identified the best models and applied a stacking ensemble to explore their combined predictive capabilities.

6. Results

6.1. Model Selection

We tested the following models for the best fault predictor: logistic regression, decision trees, random forests, XGBoosts, and neural networks models. The best performing model would need to surpass a base benchmark of 60% accuracy and an optimal predictive accuracy of 99% and 1.5% false negative rate. The latter are based on Honeywell Forge's accuracy scores (Melin, 2021) (model available in market). The accuracy metric can be misleading by not recognising high false negative rates in models' predictions. Since this can be significantly harmful for the American Airlines, the F1 Score metric has been chosen for model comparison and optimisation. First, these provide evidence of how data imbalance can significantly affect a model's predictive performance. Indeed, models trained with the imbalanced dataset showed F1 scores of 86% and lower while models with balanced data showed scores of 97% and higher. Second, it shows how accuracy alone can be a misleading metric. The first models show accuracies of 97%+ when in reality the models return high false negative and positive rates; recall and precision scores are lower for these models than for the ones trained with balanced data. Third, neural networks performed significantly worse than all other models. Deep learning methods tend to overfit with small amounts of data (Candido de Oliveira, 2017). Hence, why this model might have underperformed. Finally, it allows to identify the decision tree, random forest, and XGBoost algorithms trained with the upsampled data as the best performing models with F1 scores of 99%. Thus, achieving the optimal predictive accuracy. Table 2 shows the accuracy, recall, precision, and F1 score for each of the tested models.

6.2. Feature Importance

Identifying the best performing models helped pinpoint the variables that can help American Airlines' maintenance engineers recently discover equipment at risk of failure. First, the decision tree algorithm trained with the upsampled data indicated that epr, followed by Ps30 and BPR, are the most significant variables to forecast engine failure. This suggests that engine pressure ratio, static pressure at HPC and bypass ratio can be used by maintenance engineers to identify an engine at risk of failure. Second, the random forest trained with the upsampled data identifies epr, NRf, and NRc, in that order, as the top three most significant variables. This further supports the finding

	Accuracy	Recall	Precision	F1 Score
Logistic Regression	0.983000	0.854000	0.877000	0.865000
Decision Tree	0.972000	0.775000	0.790000	0.783000
Random Forest	0.983000	0.835000	0.886000	0.860000
XGBoost	0.982000	0.826000	0.885000	0.854000
Neural Network	0.064000	1.000000	0.064000	0.120000
Logistic Regression with Upsampled	0.971000	0.980000	0.963000	0.971000
Decision Tree with Upsampled	0.992000	1.000000	0.985000	0.992000
Random Forest with Upsampled	0.993000	1.000000	0.987000	0.993000
XGBoost with Upsampled	0.991000	1.000000	0.982000	0.991000
Neural Network with Upsampled	0.510000	1.000000	0.510000	0.675000
Logistic Regression with Synthetic Data	0.891000	0.998000	0.825000	0.903000
Decision Tree with Synthetic Data	0.986000	0.985000	0.988000	0.986000
Random Forest with Synthetic Data	0.989000	0.986000	0.993000	0.989000
XGBoost with Synthetic Data	0.988000	0.986000	0.990000	0.988000
Neural Network with Synthetic Data	0.510000	1.000000	0.510000	0.675000

Table 2 Scores for Models

of the decision tree that epr is a very important predictor. Moreover, it suggests that corrected fan speed is more important than corrected core speed for forecasting. Finally, the XGBoost trained with the upsampled data identified epr as the most important to forecast engine failure. Also, findings suggest that static and total pressure at HPC can help engineers identify risk of failure better than static pressure at HPC but their feature importance is very low.

Undoubtedly, the engine pressure ratio (epr) is a key reading for accurately predicting engines at risk of failure in 20 days. Each model has identified this variable as the sole most significant when predicting failure. This suggests that American Airlines' maintenance engineers should strongly consider epr when analysing readings coming from IoT devices installed in engines. The outlier analysis helps suggest that values for engine pressure ratio of 1.31 and 1.32 should help these actors recently identify failures developing that put machinery at risk of failure in 20 days. Engine pressure ratio provides a measure of the amount of thrust produced by a jet engine (Skybrary, 2017). Thus, these levels might help indicate when an airplane's engine is not producing enough thrust, suggesting damage, or to much thrust, indicating that healthy levels are being exceeded.

6.3. Stacking Ensemble

Since both false positives and negatives can impact AA, we further optimized the models to reduce or even eliminate these misclassification errors. Figure 5 shows that the three models are able to predict observations at risk of failure in less than 20 days with 100% accuracy. Nevertheless, it is evident that these models can erroneously classify healthy observations as 'failures' at times. The decision tree, random forest, and XGBboost returned 74, 63, and 84 false positives, respectively, in their predictions. One could argue that this number is not significant given that this amounts to approximately 0.1% of the total predictions made. However, maintaining the same rate, if American Airlines's model evaluates 2 million observations, the model will wrongfully classify 200,000 as unhealthy. This, in turn, creates 200,000 events that will require inspection and analysis, leading to unnecessary resource expenditure and potential early part removal.

The high false positive rates incentivised the development of a stacking ensemble model composed of the best models identified. This type of model uses a combination of the predictions from the three models as inputs to train a new algorithm. At first, the stacked model was trained using a logistic regression as the final estimator. This would give us insight into whether further development of this type of model would be useful. As figure 5 shows, this model improved all scores and rates compared to the three previous best models.

Next, we utilised a grid search algorithm with cross validation to tune the stacking model created. The grid search performs an exhaustive search on a parameter of a model (Malik, 2020), allowing for model parameter optimisation. In this case, it helped test different values (models) for the final estimator of the stacking model. Additionally, cross validation will ensure that the best performing algorithm is not over-performing only due to the way the original data has been split into train and test sets, making the findings more robust.

6.4. Best Model

The best predictor was identified as the stacking model with a XGBoost estimator, which returned a F1 score of 99.8%. Moreover, it returned no false negatives and only 1 false positives. This model proved to be the best strategy to forecast engine failure, allowing it to surpass benchmarks of even the current best performing solutions; higher than 99% accuracy and lower than 1.5% no-failure-found rate. Table 3 shows the scores of all the stacking models tested.

	Accuracy	F1 Score	Recall	Precision	False Positives	False Negatives
Decision Tree with Upsampled Data	0.992000	0.992000	1.000000	0.985000	74	0
Random Forest with Upsampled Data	0.993000	0.993000	1.000000	0.987000	63	0
XGBoost with Upsampled Data	0.991000	0.991000	1.000000	0.982000	84	0
Stacking Model - Logistic Regression Estimator	0.998000	0.998000	1.000000	0.997000	16	0
Stacking Model - XGBoost Estimator	1.000000	0.998000	1.000000	1.000000	1	0

Table 3 Table of Scores for Best Models (Best Highlighted Green)

7. Impact

We estimated the impact of implementing a machine learning-based maintenance operation approach for failure forecasting at American Airlines. It was estimated that AA performs 10,536 A checks and 1,317 C checks a year; out of which 5,268 and 658, respectively, are likely to come from unscheduled events. Implementing the stacking model with a XGBoost estimator to forecast engine failures would potentially help reduce the number of unscheduled events to **five** A checks and **one** C check per year. This will lead to a significant reduction in maintenance expenses. On the one hand, the cost of one unplanned A check was calculated to lie between \$2,818.2 and \$5,636.4, generating from \$17.3 millions to \$34.6 millions in unscheduled costs (Appendix A.3). The 99.9% accuracy of the model will potentially help lower these to \$17,320.7 and \$34,641.3. On the other hand, the original cost of unscheduled C checks estimated at \$556.5 millions per year can be reduced to \$649,208.3 per year. The previous calculations show that American Airlines' unscheduled A and C checks' maintenance expenses will be reduced from up to \$683.8 millions to up to \$683,849.6; a reduction of \$683.1 millions.

Moreover, the implementation of this model will help reduce American Airlines' delays and cancellation expenses due to unscheduled events. It was estimated that the company incurs \$245.7 million and \$301.0 million in delay and cancellation costs, respectively, per year due to unplanned maintenance per year. Using this machine learning technology will potentially enable AA to reduce these costs to \$245,700.0 and \$301,000.0, respectively. Once again, considering that 0.62% of delays happen due to weather complications, the total savings are around \$542.8 millions. Together with the savings in maintenance checks, AA total savings are around \$1.2 billions when adopting a failure forecasting technology.

Finally, to calculate the real potential savings in the long term, Mazareanu's growth estimates for overall unscheduled maintenance costs is used once again. A model with a 99.9% accuracy, like the stacking model with XGBoost estimator, would also potentially reduce AA's long-term non-routine maintenance costs. Indeed, it is estimated that expenses can be reduced from \$1.6

billion to $1.6 million by the third quarter of 2025 and from $2.3 billion to $2.3 million by the third quarter of 2035. This could potentially free up $3.9 billion for investment by the year 2035.

8. Discussion and Conclusions

Implementation costs, data imbalance, and operational support are the main three challenges for the correct implementation of predictive maintenance technologies. We discuss how these would affect the development and adoption of failure forecasting technologies in American Airlines's engine maintenance operations and why these can be overcome. Finally, we recommend American Airlines the optimal four-step process to follow for deploying a machine learning-based maintenance operation.

8.1. Implementation Costs

First, the company will need to invest in IoT sensors to collect data from engines, data warehouses and ETL (extract, transform, and load) software to clean, transform data and store for the model, and the deployment of the forecasting model. Fortunately, the price of IoT sensors have been steadily decreasing. Indeed, the price decreased from $1.30 in 2004 to $0.38 in 2020 (Leonard, 2019)), making the large-scale acquisition and implementation of IoT devices less costly. American Airlines would be able to collect a significant percentage of the millions of gigabytes of data generated by their fleet's engines at a fraction of the cost of 15 years ago.

Moreover, the implementation of data warehouses and ETLs is said to generate high return on investment since it helps them save time through the standardisation of data storage, enhances data quality and improves data consistency (DWIC, n.d.). This is supported by an International Data Corporation's study which concluded that the implementation of ETLs generated a 112% median five-year ROI (return on investment) with a mean payback of 1.6 years (Glow Touch, n.d.). AA should benefit from their investment in just under 2 years after implementing data warehouses with ETLs.

Finally, predictive maintenance strategies cost $9 per hour per year compared to $13 for preventive strategies (Lavi, 2018). Hence, even when deploying a ML model can cost a company between $51,750 to $136,750 (including the cost of data, research, and production) (Incze, 2019), the company will reduce costs in the long-run with this implementation.

8.2. Data Imbalance

Data imbalances are common in aircraft maintenance operations. Indeed, the majority of the data collected from aircrafts is characterised by a healthy majority and faulty minority (Dangut et al., 2020). This creates problems as it leads to misleading accuracy scores. If more data points belong to the 'healthy' class and the model only predicts 'healthy', there is a higher chance that the model will be correct but present a high false negative rate. Failing to identify unhealthy readings with high accuracy can lead to unscheduled maintenance activities and millions in additional expenses per year. Similarly, false positives, even when not as harmful as false negatives, can still significantly impact American Airlines. These will lead to early replacement, creating unnecessary expenses for AA.

We have partially addressed this issue by balancing the sample through upsampling and GAN-generated synthetic data. Nevertheless, note that the data imbalance problem has not been eliminated, simply mitigated.

8.3. Operational Support

In order to correctly implement machine learning in maintenance operations, comprehensive support from members in the executive team is required. There are cases when successful implementers of ML have received support from their COO and not only maintenance leaders (Hirshman et al., 2020). Maintenance needs to be connected more effectively with supply chain, operations, and network peers to be able to modify standard maintenance practice (Hirshman et al., 2020). The benefits of this technology come from the gains in operational efficiency and the reduction in costs. If the maintenance process, system structure and staff training level have not been optimised, the benefits will be inhibited. Indeed, timeline and budget risks are introduced (Hirshman et al., 2020), contributing to the problems the technology aims to solve.

Moreover, the COO needs to support new training programs and expert hiring processes to prepare for when the technology has been adopted. It is difficult to unlock the potential of ML applications when an organisation does not have the talent pool required to fully understand and implement the algorithms (Maruti Techlabs, n.d.). Findings need to be analysed and translated into actionable recommendations for non-technical colleagues. Otherwise, ML models will not significantly improve maintenance processes and reduce unscheduled maintenance. Hence, current staff needs to be provided with the necessary skills to interpret statistical information and techniques to uncover insights from data through training programs.

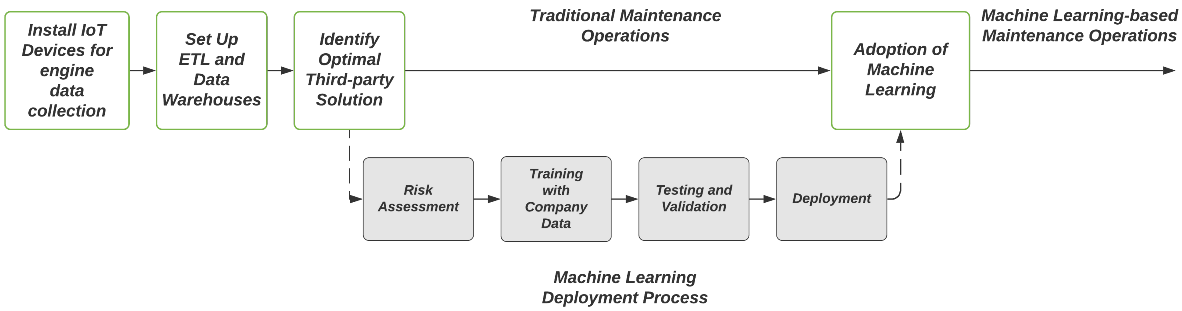

Figure 3 **Optimal Deployment Process**

8.4. Deployment Process

First, AA needs to ensure that their IoT infrastructure will allow them to collect the necessary information from their engines to forecast failure. Moreover, the data collected should be properly cleaned and stored. This will make the access of training data much simpler.

Second, once the third-party solution is identified, it is advised that AA should optimise the current normal maintenance operations while the ML deployment is taking place. Indeed, the study recommends that the company's current maintenance engineers focus on the engine's pressure ratio (epr) while analysing engine readings to effectively identify a failure. Once again, values of 1.31 and 1.32 tend to suggest equipment at risk of failure. Total temperature at LPC outlet and static pressure at HPC outlet can also aid engineers in this process.

Third, AA should verify that the solution will not affect the process's integrity and safety. Additionally, the company should ensure that the solution is trained, tested, and validated using company data to guarantee suitability for the problem at hand. The solution can then be deployed and adopted in the company's maintenance operations.

Finally, once the model has been adopted, the company can transform their maintenance operations to fit the structure of a machine learning-based maintenance operations. The implementation of these machine learning technologies throughout the process makes decision making more efficient and foments collaboration. A machine learning-based maintenance operation should involve all teams simultaneously analysing abnormal data, working to create work orders, planning maintenance dates, ensuring resources are available and placing work orders instead of all these being individual steps in a linear process. Figure 3 shows a process diagram for this four step adoption strategy.

A machine learning-based maintenance approach will significantly reduce the company's maintenance expenses and the number of unplanned events. This paper ends with two remarks: 1) **a stacking ensemble is the best algorithm for American Airlines to adopt in a machine**

learning-based maintenance operation and 2) **the best performing models were trained with a simple upsampling technique**. For future studies, it is suggested that the use of GANs for data balancing is further explored given that neural networks tend to perform better with larger datasets than the one used in this study. Moreover, since this study's impact was calculated using the findings of a model trained with simulated data, further research conducted using real American Airlines's engine data will provide further insight.

Appendix A: Calculations

A.1. Total Maintenance Costs

Tulsa Maintenance Base employs 5,000 maintenance technicians (KJRH Digital, 2020). Note that the company's furloughs due to the Covid-19 Pandemic did not heavily impact this base. By October, only 170 employees in the maintenance divisions were put on furlough (KJRH Digital, 2020). Average yearly salary of a maintenance technician is $51,482.0, equivalent to a $24.8 per hour salary (Zippia, 2021a) (2,080 working hours per year). We calculate total technicians' salary until Q3 of 2020:

$$\text{Average Work Hours by Q3 of 2020} = 1,560$$

$$\text{Salary per hour} = \$24.8$$

$$\text{Total Salary by Q3 of 2020 per technician} = \$38,688.0$$

$$\textbf{Total Technician Salary Expenses by Q3 of 2020} = 5,000 * \$38,688.0$$

$$= \$193,440,000.0$$

American Airlines entered into a contract with 30,000 mechanics, fleet workers and other union workers in early 2020, providing pay rises, bonuses and job security for the year (Arnold, 2020). The following calculations assume mechanics, fleet workers and other union workers have similar yearly salaries. Average yearly salary of a maintenance mechanic is $52,000.0, equivalent to a $25.2 per hour salary (Zippia, 2021b) (2,060 working hours per year). We calculate total mechanics' salary until Q3 of 2020:

With this information, together with the maintenance, materials, and repair costs reported in American Airlines's Q3 2020 Financial Report (American Airlines, Inc., 2020), we calculated the total maintenance costs

$$\text{Average Work Hours by Q3 of 2020} = 1,545$$

$$\text{Salary per hour} = \$25.2$$

$$\text{Total Salary by Q3 of 2020 per mechanic} = \$38,934.0$$

$$\textbf{Total Mechanic Salary Expenses by Q3 of 2020} = 30,000 * \$38,934.0$$

$$= \$1,168,020,000.0$$

$$\text{Maintenance, materials, and repairs cost} = \$1,253,000,000$$

$$\text{Maintenance Labour Expenses} = \$193,440,000.0 + \$1,168,020,000.0 = \$1,361,460,000$$

$$\textbf{Total maintenance costs} = \$1,253,000,000 + \$1,361,460,000$$

$$= \$2,614,460,000$$

A.2. Scheduled Maintenance A-check Cost

To estimate the cost of a scheduled maintenance A check, first labour costs are considered. Two technicians require between 10 to 20 hours of work to complete the check (Skybrary, n.d.; University of Westminster, 2008) and with an hourly salary of $24.8 (Appendix A.1.), this would cost between $248.0 and $496.0. Including labour expenses, American Airlines's total maintenance costs accounted for 12.5% of the operating expenses, or $2.6 billion (Appendix A.1), by the 9 Months Ending September 30. Thus, labour costs would account for 52.8 percent of maintenance costs. This allows us to estimate the total cost of a scheduled A check to be between $469.7 and $939.4.

An A check should occur every 500 flight hours (University of Westminster, 2008), which translates to approximately 8 checks per aircraft per year; assuming a plane operates 12 hours per day on average (Plane Stats, 2017). Thus, American Airlines should perform an estimate of 10,536 A-checks in a year.

A.3. Unscheduled Maintenance A-check Cost

Unscheduled events cost between 3 to 9 times more than scheduled events (Trefis, 2016). We used the middle of the range (6 times) to avoid extremes. We estimate that 50 per cent, or in this case 5,268, of A checks are likely to come from unscheduled events.

A.4. Scheduled Maintenance C-check Cost

Considering the Boeing-737 is the most common type of aircraft in American Airlines fleet, we decided to use the average cost per C-check for this aircraft as an estimation for the average cost of all aircrafts.

$$\text{Number of A checks per year} = 10,536$$

$$\text{Cost Scheduled A check per aircraft (a)} = \$469.7$$

$$\text{Cost Scheduled A check per aircraft (b)} = \$939.4$$

$$\textbf{Total Cost Scheduled A check (a)} = 10,536 * \$469.7$$

$$= \$4,948,759.2$$

$$\textbf{Total Cost Scheduled A check (b)} = 10,536 * \$939.4$$

$$= \$9,897,518.4$$

$$\text{Cost Scheduled A check per aircraft (a)} = \$469.7$$

$$\text{Cost Scheduled A check per aircraft (b)} = \$939.4$$

$$\text{Cost Unscheduled A check per aircraft (a)} = (\$469.7 * 6) = \$2,818.2$$

$$\text{Cost Unscheduled A check per aircraft (b)} = (\$939.4 * 6) = \$5,636.4$$

$$\textbf{Total Cost Unscheduled A check (a)} = (5,268 * \$469.7) + (5,268 * \$2,818.2) = \$17,320,657.2$$

$$\textbf{Total Cost Unscheduled A check (b)} = (5,268 * \$939.4) + (5,268 * \$5,636.4) = \$34,641,314.4$$

An average C check for a 737 costs \$32.18 pre flight hour (Skylink, 2018). Moreover, on average, a plane operates 12 hours per day, translating to 4,380 flight hours per year.

$$\text{Flight Hours per Year} = 4,380$$

$$\text{Average C check cost per flight hour} = \$32.18$$

$$\text{Number of C checks per year} = 1,317$$

$$\text{Scheduled C check cost per aircraft} = (4,380 * \$32.18) = \$140,948.4$$

$$\textbf{Total Cost Scheduled C check} = (\$140,948.4 * 1,317) = \$185,629,042.8$$

A.5. Unscheduled Maintenance C-check Cost

Unscheduled events cost between 3 to 9 times more than scheduled events (Trefis, 2016). We used the middle of the range (6 times) to avoid extremes. We estimate that 50 per cent, or in this case 658, of C checks are likely to come from unscheduled events.

$$\text{Cost Scheduled C check per aircraft} = \$140,948.4$$

$$\text{Cost Unscheduled C check per aircraft} = (\$140,948.4 * 6) = \$845,690.4$$

$$\textbf{Total Cost Unscheduled C check} = (658 * \$140,948.4) + (658 * \$845,690.4) = \$649,208,330$$

Appendix B: Further Information

B.1. Variable Descriptions

Symbol	*Description*	*Units*
Parameters available to participants as sensor data		
T2	Total temperature at fan inlet	°R
T24	Total temperature at LPC outlet	°R
T30	Total temperature at HPC outlet	°R
T50	Total temperature at LPT outlet	°R
P2	Pressure at fan inlet	psia
P15	Total pressure in bypass-duct	psia
P30	Total pressure at HPC outlet	psia
Nf	Physical fan speed	rpm
Nc	Physical core speed	rpm
epr	Engine pressure ratio (P50/P2)	--
Ps30	Static pressure at HPC outlet	psia
phi	Ratio of fuel flow to Ps30	pps/psi
NRf	Corrected fan speed	rpm
NRc	Corrected core speed	rpm
BPR	Bypass Ratio	--
farB	Burner fuel-air ratio	--
htBleed	Bleed Enthalpy	--
Nf_dmd	Demanded fan speed	rpm
PCNfR_dmd	Demanded corrected fan speed	rpm
W31	HPT coolant bleed	lbm/s
W32	LPT coolant bleed	lbm/s

Table 4 Table of Variable Descriptions (Saxena et al., 2018)

Acknowledgments

I would like to thank Lina Song for guiding and supporting me through the process of writing this paper.

References

American Airlines (2020). American Airlines reports third quarter 2020 financial results. *American Airlines Newsroom*; (January 4, 2021). https://americanairlines.gcs-web.com/news-releases/news-release-details/american-airlines-reports-third-quarter-2020-financial-results

American Airlines Group (2021). Parent Company of American Airlines *American Airlines*; (January 4, 2021). https://www.aa.com/i18n/customer-service/about-us/american-airlines-group.jsp

American Airlines, Inc. (2020) American Airlines Reports Third-Quarter 2020 Financial Results. *American Airlines Group*; (January 10, 2021). https://www.globenewswire.com/en/news-release/2020/10/22/2112668/0/en/American-Airlines-Reports-Third-Quarter-2020-Financial-Results.html

Arnold, K. (2020) American Airlines maintenance workers overwhelmingly approve contract after 5-year battle. *Dallas Morning News*; (March 24, 2021). https://www.dallasnews.com/business/airlines/2020/03/27/american-airlines-maintenanceworkers-overwhelmingly-approve-contract-after-5-year-battle/

Aungst, J. Johnson, M.E., Lee, S.S., Lopp, D. and Williams, M. (2008) Planning of Non-routine work for aircraft scheduled maintenance. *Purdue University*; paper 44, entry 203.

Ausick, P. (2020). The high cost of cancelled airline flights [Internet]. *24/7 Wall Street*; (January 12, 2021). https://247wallst.com/services/2014/10/29/the-high-cost-of-cancelled-airline-flights/

BAA Training (2018). Did You Know that Boeing Was the First to Use Turbofan Engines?. *BAA Training*; (March 14, 2021). https://www.baatraining.com/did-you-know-that-boeing-was-the-first-to-use-turbofanengines/

Bureau of Transportation Statistics (2020). Airline On-Time Statistics and Delay Causes. *Bureau of Transportation Statistics*; (May 1, 2021). https://www.transtats.bts.gov/OT_Delay/ot_delaycause1.asp?display=chart&pn=1

Candido de Oliveira, D. (2017). Is Deep Learning Not Useful If There is No Large Dataset Available?. *Research Gate*; (May 14, 2021). https://www.researchgate.net/post/Is_deep_learning_not_useful_if_there_is_no_large_dataset_available.

Centre for Aviation (2020). American Airlines Group Inc. *CAPA*; (January 20, 2021). https://centreforaviation.com/data/profiles/airline-groups/american-airlines-group-inc

Cleverism (2020). American Airlines. *Cleverism*; (January 10, 2021). https://www.cleverism.com/company/american-airlines/

Dangut, M.D., Zakwan, S., and Jennions, I. (2020). Aircraft Predictive Maintenance Modeling using a Hybrid Imbalance Learning Approach. *9th International Conference on Through-life Engineering Service*; (May 22, 2021). https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3718065

Data Warehouse Information Center (n.d.). Benefits of a Data Warehouse. *DWIC*; (February 13, 2021). https://datawarehouseinfo.com/data-warehouse/benefits-of-a-data-warehouse/

De Bree, S. (n.d). Big Data in Aviation - Reduce Costs Thru Predictive Maintenance. *Exsyn*; (February 13, 2021). https://www.exsyn.com/blog/big-data-in-aviation-predictive-maintenance

Doig, S.J., Howard, A., Ritter, R.C.,(2003) The hidden value in airline operations. *McKinsey & Company*; (May 10, 2021). https://www.mckinsey.com/industries/travel-logistics-and-infrastructure/our-insights/the-hidden-value-in-airline-operations

Emorphis Technologies (2019). How IoT & Machine learning changing the face of Predictive Maintenance. *Medium*; (May 10, 2021). https://medium.com/emorphis.technologies/how-iot-machine-learning-changing-the-face-of-predictive-maintenance-7cf927fca877

Federal Aviation Administration (2009). Guidance Material for Aircraft Engine Life-Limited Parts Requirements. *Federal Aviation Administration*; (March 14, 2021). https://www.faa.gov/documentLibrary/media/Advisory_Circular/AC_33_70-1.pdf

FT Maintenance (2019). How maintenance practices have changed — history of maintenance. *FT Maintenance*; (June 5, 2021). https://ftmaintenance.com/maintenance-management/how-maintenance-practices-have-changed-history-of-maintenance/

Glow Touch (n.d.). The Benefits of Data Warehousing and Extract, Transform and Load (ETL). *Glow Touch*; (February 15, 2021). https://www.glowtouch.com/the-benefits-of-data-warehousing-and-etl/

Golightly, R. (2019). Why time's up on preventive maintenance. *Processing Magazine*; (January 28, 2021). https://www.processingmagazine.com/maintenance-safety/article/15587836/why-times-up-on-preventive-maintenance

Harper, L. (2020). Will there be an international air travel resurgence in 2021?. *Flight Global*; (April 26, 2021). https://www.flightglobal.com/networks/will-there-be-an-international-air-travel-resurgence-in-2021/140629.article

Hirshman, B., Milon, T., Brimmer, A., Brinkopf, B., Rabson, M., and Smith, K.(2020). How value can take off with predictive aircraft maintenance. *Boston Consulting Group*; (February 13, 2021). https://www.bcg.com/publications/2020/building-value-with-predictive-aircraft-maintenance¿

Ho, S. and Xie, M. (1998).The use of ARIMA models for reliability forecasting and analysis. *Computers & Industrial Engineering*, 35(1–2): 213-216.

Hoyland, T., Spafford, C., and Medland, A. (2016). MRO Big Data - A Lion or A Lamb?. *Oliver Wyman*; (February 23, 2021). https://www.oliverwyman.com/content/dam/oliver-wyman/global/en/2016/apr/NYC-MKT9202-001MRO-Survey-2016_web.pdf

Incze, R. (2019). The Cost of Machine Learning Projects. *Medium*; (February 13, 2021). https://medium.com/cognifeed/the-cost-of-machine-learning-projects-7ca3aea03a5c

International Airport Review (2020). Passenger demand remains drastically low in September 2020. *IAR*; (February 15, 2021). https://www.internationalairportreview.com/news/143436/passenger-demand-drastically-low-september-2020/

Kaggle (2019). NASA Turbofan Jet Engine Data Set. *Kaggle*; (March 14, 2021). https://www.kaggle.com/behrad3d/nasa-cmaps

Katoky, A., Stringer, D. and Saxena, R. (2020). Here's what you do with two-thirds of the world's jets when they can't fly. *Bloomberg*; (April 12, 2021). https://www.bloomberg.com/news/features/2020-04-16/coronavirus-travel-what-happens-to-planes-grounded-by-covid-19.

KJRH Digital (2020) American Airlines to furlough 19,000 employees, Tulsa impact unknown. *KJRH*; (January 6, 2021), https://www.kjrh.com/news/local-news/american-airlines-to-furlough-19-000-employees-tulsa-impact-unknown

Kozlowski, J.D. (2003). Electrochemical cell prognostics using online impedance measurements and model-based data fusion techniques. Aerospace conference, (7):3257-3270

Kucherevskiy, I. (2020). RUL Prediction. *Kaggle*; (March 26, 2021). https://www.kaggle.com/kucherevskiy/rul-prediction

Lavi, Y. (2018). The rewards and challenges of predictive maintenance. *InfoQ*; (June 1, 2021). https://www.infoq.com/articles/predictive-maintenance-industrial-iot/

Leonard, M. (2019). Declining Price of IoT sensors Means Greater Use in Manufacturing. *Supply Chain Dive*; (March 13, 2021). https://www.supplychaindive.com/news/declining-price-iot-sensors-manufacturing/564980/

Li, B., Zhao, J., and Guo, J. (2011). Innovative metrics for equipment failure evaluation and prediction system based on ARMA model. *Systems Engineering and Electronics*, 33(1):98-101.

Li, R.Y. and Kang, R. (2008). Research on failure rate forecasting method based on ARMA model. *Systems Engineering and Electronics*, 8.

Malik. F. (2020). What is Grid Search?. *Medium*; (February 26, 2021) https://medium.com/fintechexplained/what-is-grid-search-c01fe886ef0a

Maruti Techlabs (n.d.). Challenges Faced by Businesses in Adopting Machine Learning. *Maruti Techlabs*; (March 23, 2021). https://marutitech.com/challenges-machine-learning/

Mazareanu, E. (2018). Estimated annual global costs from unscheduled maintenance from 2017 to 2035, by aircraft type. *Statista*; (January 15, 2021). https://www.statista.com/statistics/922666/estimated-costs-worldwide-from-unscheduled-maintenance-by-aircraft-type/

Mazareanu, E. (2020). Number of passengers enplaned by American Airlines from FY 2014 to FY 2019. *Statista*; (January 15, 2021). https://www.statista.com/statistics/603884/passenger-enplanements-american-airlines/

Mazareanu, E. (2019). Top global airlines in 2019, by fleet size. *Statista*; (January 15, 2021). https://www.statista.com/statistics/1013159/airlines-worldwide-fleet-size/

Melin, J. (2021). Moving Beyond the Hype of Predictive Maintenance [Internet]. *Honeywell* (March 3, 2021). https://aerospace.honeywell.com/en/learn/about-us/blogs/2019/04/moving-beyond-th e-hype-of-predictive-maintenance

National Aeronautics and Space Administration (2021). Prognostics Center of Excellence. *NASA*; (March 14, 2021). https://ti.arc.nasa.gov/tech/dash/groups/pcoe/prognostic-data-repositor/

Pariyar, U., Tamilselvan, T., Thilakan, S., and Shanmugaraja, M. (2013). Analysis on Costs for Aircraft Maintenance. *Advances in Aerospace Science and Applications*, 3(3):177 - 182

Plane Spotters (2021). American Airlines Fleet Details and History [Internet]. *Plane Spotters*; (January 11, 2021). https://www.planespotters.net/airline/American-Airlines

Plane Stats (2017). Aircraft Cost and Operations: Wide-bodies. *Plane Statistics*; (January 15, 2021). https://www.planestats.com/bhsw_2017dec

RAND Corporation (2006). The maintenance costs of aging aircraft: Insights from commercial aviation. *RAND Project Air Force*; (February 28, 2021). https://apps.dtic.mil/dtic/tr/fulltext/u2/a459880.pdf

Reddy, R. (2020). Machine learning 101 in predictive maintenance. *Industry Week*; (March 21, 2021). https://www.industryweek.com/technology-and-iiot/article/21134278/machine-learning-101-in-predictive-maintenance

Salazar, L.J. (2015). Analysing uncertainty and delays in aircraft heavy maintenance. *Alliance Manchester Business School*, Vol I.

Salazar, L.J., Yang, J., and Chen, Y. (n.d). Analysing delays and disruptions in Aircraft Heavy Maintenance. *The University of Manchester*; (January 22, 2021). https://proceedings.systemdynamics.org/2014/proceed/papers/P1326.pdf

Saltoglu, R., Humaira, N., and Inalhan, G. (2016). Scheduled maintenance and downtime cost in aircraft maintenance management. *World Academy of Science, Engineering and Technology*, 10(3).

Samaranayake, P. and Kiridena, S. (2012). Aircraft maintenance planning and scheduling: An integrated framework. *Journal of Quality in Maintenance Engineering*; 18(4).

Saxena, A., Goebel, K., Simon, D. and Eklund, N. (2018). Damage Propagation Modeling for Aircraft Engine Run-to-Failure Simulation. *NASA Ames Research Center* (March 27, 2021). https://www.researchgate.net/publication/251867156_Damage_propagation_modeling_for_aircraft_engine_run-to-failure_simulation

Singh, N. (1994). Forecasting time-dependent failure rates of systems operating in series and/or in parallel. *Microelectronics Reliability*, 34(3):391-403.

Skybrary (2021). Airbus A-321. *Skybrary*; (March 13, 2021). https://www.skybrary.aero/index.php/A321

Skybrary (2017). Engine Pressure Ratio (EPR). *Skybrary*; (June 25, 2021). https://www.skybrary.aero/index.php/Engine_Pressure_Ratio_(EPR)

Skybrary. (n.d) Maintenance Workload. *Skybrary*; (June 17, 2021). https://skybrary.aero/index.php/Maintenance_ Workload

Skylink (2018). Here's how much your B737 C-check costs per flight hour (plus how you can reduce the maintenance cost). *Skylink*; (May 16, 2021). https://www.skylinkintl.com/blog/heres-how-much-your-b737-c-check-costs-per-flight-hour-plus-how-you-can-reduce-the-maintenance-cost

Solved FM. (2018). Planned Maintenance vs Unplanned: Does it matter?. *Solved FM*; (January 5, 2021). https://www.solvedfm.co.uk/planned-maintenance-vs-unplanned-matter/

Trefis Team (2016). What is the impact of flight delays?. *Trefis*; (June 10, 2021). https://www.trefis.com/stock/dal/articles/375013/what-is-the-impact-of-flight-delays/2016-08-31

University of Westminster (2008). Dynamic Cost Indexing. *Transport Studies Group*; (June 10, 2021). https://www.eurocontrol.int/eec/gallery/content/public/documents/projects/CARE/CARE_INO_III/DCI_TDD9-0_Airline_maintenance_marginal_delay_costs.pdf

Wright, R. (2017). Sensors and sensibility keep today's jet planes working hard. *Financial Times*; (June 20, 2021). https://www.ft.com/content/c34736de-962e-11e7-8c5c-c8d8fa6961bb

Zippia (2021) How much does an aircraft maintenance technician earn in the united states?. *Zippia*; (June 17, 2021), https://www.zippia.com/aircraft-maintenance-technician-jobs/salary/

Zippia (2021). How much does an aircraft mechanic earn in the united states?. *Zippia* (June 17, 2021). https://www.zippia.com/aircraft-mechanic-jobs/salary/