# Prediction of Non-Routine Tasks Workload for Aircraft Maintenance with Supervised Learning

Li, H.; Ribeiro, M.J.; Santos, Bruno F.; Tseremoglou, I.

**DOI**
10.2514/6.2024-2529

**Publication date**
2024

**Document Version**
Final published version

**Published in**
Proceedings of the AIAA SCITECH 2024 Forum

AIAA SciTech Forum
8-12 January 2024, Orlando, FL
AIAA SCITECH 2024 Forum
10.2514/6.2024-2529

# Prediction of Non-Routine Tasks Workload for Aircraft Maintenance with Supervised Learning

Haonan Li*, Marta Ribeiro†, Bruno Santos‡, Iordanis Tseremoglou§,
*Delft University of Technology, 2629 HS Delft, The Netherlands*

**Aircraft maintenance scheduling is a focus point for airlines. Maintenance is essential to ensure the airworthiness of aircraft, but it comes at the cost of rendering them unavailable for operations. In current operations, aircraft maintenance scheduling must often be updated to include time for non-routine and non-schedule tasks. These non-routine tasks can increase costs, maintenance workload, and uncertainty of the airlines' operations. This research introduces a supervised learning framework designed to forecast future non-routine task workloads accurately, improving the accuracy of the planned maintenance schedule. This framework consists of two random forest predictors which estimate the amount of non-routine tasks and the number of future work hours that should be allocated in advance for potential non-routine tasks. Our approach produces highly reliable predictions by leveraging a robust dataset obtained from an international airline. The results show an average of 20% improvement versus an existing on-site sampling method. Furthermore, our in-depth analysis of prediction distributions enables the identification of the underlying causes of significant prediction errors, shedding light on the unpredictabilities inherent to non-routine tasks.**



## I. Introduction

Determining the timing for aircraft maintenance is a critical aspect of an airline's day-to-day operations. Regular maintenance is essential to ensure the airworthiness of aircraft, but it comes at the cost of rendering them unavailable for operations, thereby impacting the commercial profitability of the airline. Recent research by Pater, Reijns, and Mitici [1], Lee and Mitici [2], and the Airline Maintenance Cost Executive Commentary [3] indicates that, on average, airlines allocate approximately 10% of their variable costs to cover maintenance expenses. An effectively implemented aircraft maintenance strategy has the potential to cut costs by addressing maintenance expenditures and potential failures, all while emphasizing safety and equipment reliability [4]. In the highly competitive global aviation industry, airlines feel compelled to optimize the flying hours and overall availability of their aircraft while simultaneously minimizing associated maintenance costs.

Maintenance tasks are often categorized into Routine tasks and Non-routine tasks. The latter is characterised by uncertainty, resulting in additional costs and disruptions to maintenance and flight operations. While routine tasks are executed periodically or predetermined within the maintenance program, non-routine tasks emerge unexpectedly from various sources: from crew members' observation of anomalies while performing routine tasks or from sensors' within the aircraft. Non-routine tasks often result in additional maintenance workload and failure to correctly plan maintenance activities or assess aircraft's turnaround time [5]. According to Aungst et al. [6], non-routine work can take up 50% of the total maintenance completed during an aircraft scheduled service.

The current estimation of the additional workload associated with non-routine tasks is based on averaging the work hours across multiple occurrences. However, this method has consistently proven to be inaccurate leaving a gap in research- there is high variability in the volume and hours of this unplanned work. Airlines are pursuing a more precise means of estimating the workload of non-routine tasks for scheduling and simulation. Unfortunately, limited research has been conducted on this problem. One major obstacle stems from the scarcity of accessible and valuable datasets. The present work attempts to cover these two gaps: (1) lack of methods that estimate the workload for each non-routine task not just the average, and (2) lack of methods build upon real data. The latter is commuted through directly cooperation with an international airline - this work has access to exceptionally valuable datasets, rich in comprehensive information, enabling us to advance research in this subject.

*PhD Candidate, Control & Operations Department, Faculty of Aerospace Engineering, George.Li@tudelft.nl
†Assistant Professor, Control & Operations Department, Faculty of Aerospace Engineering, M.J.Ribeiro@tudelft.nl
‡Associate Professor, Control & Operations Department, Faculty of Aerospace Engineering, B.F.Santos@tudelft.nl
§PhD Candidate, Control & Operations Department, Faculty of Aerospace Engineering, I.Tseremoglou@tudelft.nl

This paper builds a supervised learning tool to inform an airline of how many work hours should be allocated for future non-routine tasks. This tool allows airlines to better plan their maintenance schedule. As far as the authors know, this is the first work to use supervised learning methods to estimate future non-routine workload with complete airline data. In summary, this paper innovates by:

- Constructing supervised-learning models for predicting non-routine work hours: This paper will employ supervised-learning models to predict the number of future work hours that should be allocated in advance for potential non-routine tasks.
- Analysing the influence of different factors on the patterns of non-routine tasks: Non-routine tasks originate from various sources. A comprehensive examination of each category is needed to identify the impact of different influencing factors. Our objective is to establish relationships among different features to better understand the intricate connections and interdependencies within non-routine tasks. Thus, in this paper, the importance of each feature input is analysed in order to verify which characteristics highly influence the workload of non-routine tasks.
- Analysis of a probabilistic outcome: the employed supervised learning method outputs a probabilistic outcome enabling a direct evaluation of when this outcome may be trusted. This paper examines how probability and average quantity figures can be leveraged to identify forthcoming workloads of non-routine tasks and the requisite materials.

The subsequent sections of this paper are organized as follows. Section 2 undertakes a literature overview pertaining to the identified problem. Section 3 expounds upon the methodology employed in this study. Section 4 executes a case study based on authentic data procured from an international airline. The findings and analysis derived from the study are articulated in Section 5. Section 6 provides the concluding remarks for the paper.

## II. Literature Overview

In the realm of forecasting non-routine task requirements, prior research has framed the problem as a data analytics challenge, as evidenced by Pelt, Stamoulis, and Apostolidis [7]. This approach requires access to pertinent data. As proposed by Samaranayake and Kiridena [8], who introduced an integrated framework for resource and materials planning along with operational scheduling to mitigate the impact of non-routine tasks.

It is important to note that earlier studies, exemplified by Zorgdrager et al. [9], were primarily focused on the prediction of materials needed for future non-routine tasks. The prediction of work hours for such tasks is considerably more intricate, characterized by greater variability and a heightened need for precision. Recent investigations, including Georgiev and Vachev [10], have resorted to linear regression and supervised learning models for workload estimation based on historical data. However, the limited availability of data has resulted in inaccuracies and overfitting of supervised learning models. Surprisingly, the linear model yielded the highest accuracy. Nevertheless, it is the opinion of the authors of this paper that linear models cannot comprehensively capture the variability and complexity inherent in the workload of non-routine tasks.

To achieve accurate work hour estimations for non-routine tasks and extract meaningful features from the existing data, various supervised learning techniques are considered. Given the extensive datasets with well-defined quantifiable attributes, the application of Random Forest (RF) models, a well-established machine learning method, emerges as a viable choice for prediction and classification tasks. This algorithm is adept at handling large datasets, and its decision tree-based structure facilitates the creation of easily interpretable predictions. Several noteworthy studies have underscored the efficacy and efficiency of RF across different domains. Kishino et al. [11] demonstrated the high accuracy and efficiency of an RF model in predicting the fatigue life of bending polymer films. Mohamed et al. [12] introduced a hybrid mental health prediction model utilizing the RF algorithm, achieving an impressive accuracy rate of 98.13%. In the field of aviation research, RL models have shown to be stable and less susceptible to over-fitting [13], and have been effectively employed to predict future mechanical failures, as shown by Yan and Zhou [14].

Considering that our target end-users comprise technicians and schedulers, the emphasis is also placed on the explainability of the algorithm [15, 16]. We explore techniques to enhance the clarity and user-friendliness of these results, given the intended implementation of this model in real-world production environments. The paper delves into the means by which the final output of the model can be made.

# III. Methodology

## A. Non-Routine Task Workload Prediction Logic

In the anticipation of Non-routine Task workload, we address two key dimensions: the quantity of non-routine tasks and the corresponding labor hours. Although there is a conceivable approach to utilizing one random forest model for the direct prediction of the two previous elements, empirical findings from our experiments have proven to be less than satisfactory. Consequently, a decision was made to bifurcate the prediction process into two more tractable tasks: forecasting the total number of non-routine tasks and predicting the labor hours associated with these tasks. Two distinct random forest models were devised to obtain estimations.

The prediction logic employed in this work is presented in Figure 1. The first random forest model, forecasting the total number of non-routine tasks, furnishes valuable information for the prognostication of labor hours. It is, therefore, employed as a critical input for the subsequent prediction of non-routine labor hours. It is imperative to highlight that the prediction of total labor hours poses a more challenging task. In contrast, the prediction of task numbers represents a more straightforward undertaking, given its discrete nature and amenability to categorical classification.

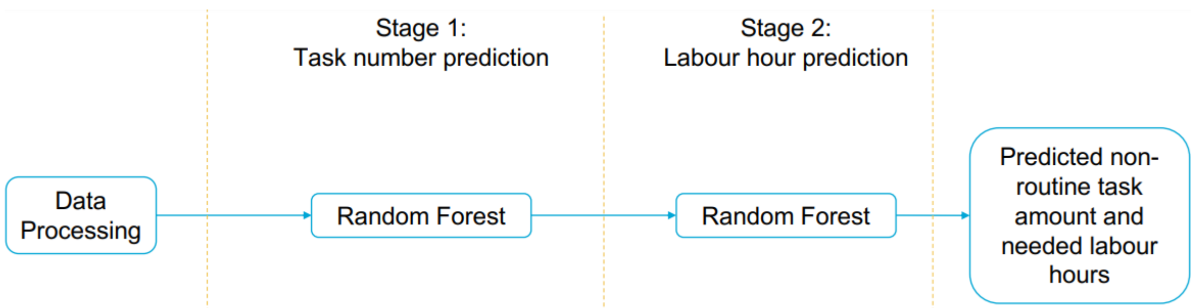



**Fig. 1 Predicting logic employed in this work.**

In resume, the methodology of this paper includes the following steps :

1) Data organization and preprocessing in a format which can used by the supervised learning method.
2) Prediction of the total number of tasks, by resorting to a Random Forest classification algorithm.
3) Labor Hour Estimation, with a second Random Forest regression algorithm, incorporating the output of the previous step.
4) Analysis of the predictive Results, including (1) an evaluation of the importance of each feature in the final prediction and (2) probability and average quantity figures.

## B. Random Forest

This section will define the employed supervised learning method. The Random Forest algorithm is an ensemble learning method commonly used for both classification and regression tasks in machine learning. It is based on decision tree algorithms and is designed to improve their performance, reduce overfitting, and enhance predictive accuracy. The pertaining elements are as follows:

1) Decision Trees: Random Forest starts with the concept of decision trees. Decision trees are a popular machine learning algorithm used for both classification and regression. They work by recursively splitting the dataset into subsets based on the most significant features, ultimately assigning a label or value to each subset. In this paper, we use 100 trees for fitting and there is no max depth for each tree.
2) Randomization: The key idea behind Random Forest is to introduce randomness into the tree-building process. We used the following two methods to introduce randomness:
   I) Bootstrapping: Random Forest generates multiple random subsets of the training data through a process known as bootstrapping. Each subset is used to train a separate decision tree. This helps create diversity among the individual trees.
   II) Feature Selection: At each node of the decision tree, only a random subset of features (variables) is considered for splitting. This prevents one dominant feature from skewing the entire model and promotes better generalization. The quantity of selected features at each node is auto-optimised by the model.

3) Voting or Averaging: Random Forest combines the predictions from each tree. For classification tasks, it often uses majority voting, where the class that receives the most votes across all trees is the final prediction. For regression tasks, it takes the average of the predictions. For a probabilistic outcome, each tree predicts a class and probabilities are calculated from these classes.

# IV. Case Study

### A. Data Profile

Empirical data was furnished by a prominent multinational airline. The dataset encompasses maintenance records pertaining to a fleet of 30 wide-body aircraft, spanning a chronological scope of five years. This dataset encompasses in excess of 50,000 work packages (WPs), coexisting with their respective Routine tasks (RTs) and Non-routine tasks (NRs). Herein, a work package (WP) signifies the cumulative workload associated with each instance of aircraft maintenance upon entry into the hangar. It is worth noting that multiple RTs and NRs constitute a singular work package.

### B. Source of Non-routine Tasks

Non-routine tasks can stem from different sources. All these different causes must be represented in the final model so that the final prediction of the model is accurate. In this work, we consider that non-routine tasks can originate from the following situations, in decreasing order of likelihood:

- Routine tasks: Non-routine tasks predominantly stem from repetitive tasks, exclusively associated with certain RTs. Certain RTs are more likely to encounter non-routine tasks or require additional work hours.
- Work packages: Technicians may spontaneously discover faults that do not correspond to any specific RT. This category of non-routine tasks exhibits a greater degree of randomness due to work packages being generated from several RTs in a non-deterministic manner.
- Reports by crew members or identified by sensors in the aircraft itself. However, it is worth noting that this source is of a relatively minor proportion compared to the previous categories.

### C. Feature Extraction

Our research endeavors are primarily focused on non-routine tasks originating from RTs. The presence of RTs that exhibit a greater propensity for non-routine tasks significantly enhances the probability of their manifestation. By employing advanced data analysis methodologies, we discern patterns that interconnect the quantity and typology of RTs. For example, the historical data pertaining to the same RT and the same aircraft proves invaluable in the process of pattern recognition.

Additionally, aircraft age emerges as a salient contributing factor, given that older components tend to be more susceptible to malfunctions and may require more frequent maintenance work. This, this element is added to the prediction method.

Finally, the duration of maintenance slots may exert an influence on the incidence of non-routine tasks. Technicians allocate more attention to the completion of RTs and conduct more comprehensive aircraft inspections during longer maintenance slots. Consequently, a higher frequency of non-routine tasks is anticipated in such cases.

Based on the above analysis, we included the following features in the supervised learning prediction model:

- Total number of labour hours spent on the RT.
- Registration of the aircraft (as to identify if the specific aircraft is prone to fault).
- Type of RT (i.e. A-check, C-check, P-check, PH-check, H-check).
- Fly cycle of the aircraft.
- Fly hour of the aircraft.
- Corresponding non-routine task number for the past 5 times (Time span may vary according to specific RTs noted as "NR task minus*" in Figure 2).
- Corresponding non-routine labour hours for the past 5 times (noted as "NR labour minus*" in Figure 2).

### D. Training & Testing

The dataset was split into training and testing set by 80% and 20%, respectively. Note that no outliers were removed before experiments since the latter confirmed that outlier removal did not affect prediction results.

# V. Results

Our research entails the experimental examination of numerous RTs. We have selected three distinct RTs to serve as exemplars, and our subsequent analysis will encompass a multifaceted exploration of these results. The chosen RT are represented in this paper as RT-1, RT-2, and RT-3. The results of predicting NR for these RTs are presented in this section.

## A. Prediction Accuracy

Within this section, we unveil the outcomes of our predictive modelling endeavors, which encompass the estimation of the total number of tasks and labor hours. The task number prediction is assessed based on accuracy, while labor hours are evaluated using the Mean Absolute Error (MAE). Concurrently, we benchmark our approach against an existing averaging method, currently utilized by the airline which made the underlining data available. The summary of these results is accessible in Table 1, offering a comprehensive overview of our predictive prowess. Note that our model is denoted as Random Forest + Random Forest (RF+RF), as two different RF models were used: one for determining the total number of tasks and a second one for predicting the total labour hours.

**Table 1 Final prediction accuracy of the model developed in this paper.**

| | Task Number (Accuracy [%]) | | Labour Hour [MAE/h] | |
|---|---|---|---|---|
| | Averaging Method | RF + RF | Averaging Method | RF + RF |
| RT-1 | 60 | 86 | 0.3364 | 0.08 |
| RT-2 | 98 | 99 | 0.02 | 0.002 |
| RT-3 | 97 | 98.59 | 0.03 | 0.01 |

The data within Table 1 reveals that, for all three RTs under investigation, the supervised learning model achieves improvements in both task number predictions and labour hour estimations compared to an averaging method. In general, the proposed model has demonstrated a noteworthy enhancement in the Mean Absolute Error (MAE) for labor hour predictions, with a substantial reduction of 20%. Similarly, the prediction accuracy for the number of tasks has exhibited a significant improvement, achieving a 21% increase in accuracy. It is noteworthy to mention that the magnitude of these enhancements varies among distinct RTs.

Of particular significance, RT-1 exhibits substantial enhancements, as its accuracy ascends from 60% to 86%, while the MAE reduced from 0.3364 hours to 0.08 hours. These improvements are attributed to the inherent nature of this RT, where non-routine occurrences tend to be both numerous and erratic, posing challenges for conventional sampling techniques. Our approach is able to capture these intricate patterns, thus improving the accuracy of the predictions.

In contrast, the other two RTs in focus exhibit marginal improvements. However, this outcome does not diminish the overall efficacy of our proposed method. RTs 2 and 3 predominantly represent scenarios with rare non-routine findings, obscured amid a profusion of zero occurrences. While traditional sampling methods lean on statistical inference, our approach excels in unearthing hidden patterns, resulting in a reduction of MAE from 0.02 to 0.002 and 0.03 to 0.01, as highlighted in Table 1. These enhancements are particularly consequential when extrapolated to a larger-scale context encompassing over 20,000 cases.

We delve into the feature importance of each RT in Figure 2, seeking to discern the key contributors to our predictive models' success. In Figure 2, the variables "registration" and "type" have been subjected to binary encoding to mitigate the challenge posed by a voluminous influx of feature inputs, serving as an indicator but not representing specific importance. The feature importance analysis in Figure 2 underscores the salient roles played by historical data from the preceding assessment and the actual labour hours in the RT. This finding is intuitive, as it suggests that the prior identification of non-routine tasks in this specific RT reduces the likelihood of their recurrence, while extended technician engagement with the RT augments the probability of detecting non-routine elements.

Moreover, this analysis led to the conclusion of how, for different routine tasks, different factors take precedence in the prediction of the consequent non-routine tasks. Such highlights the inherent differences to routine tasks. As per Figure 2, diverse RTs have exhibited pronounced distinctions in their respective feature importance outcomes. Notably, historical data underscores the significance of features for RT-1 and RT-3, with the Type of checks displaying marginal importance specifically for RT-1. This divergence is attributed to the inherent dissimilarities in the operational scope of each RT, encompassing distinct maintenance requisites, the targeted aircraft components, and the nature of the requisite tasks. The culmination of these multifaceted factors engenders distinctive feature characteristics for each RT. Consequently, a proactive identification of features influencing the prediction for each RT is rendered unattainable. Our methodology opts instead for the provision of a comprehensive selection of features for the model to assimilate. This elucidates the absence of values for certain features in Figure 2 corresponding to specific RTs.

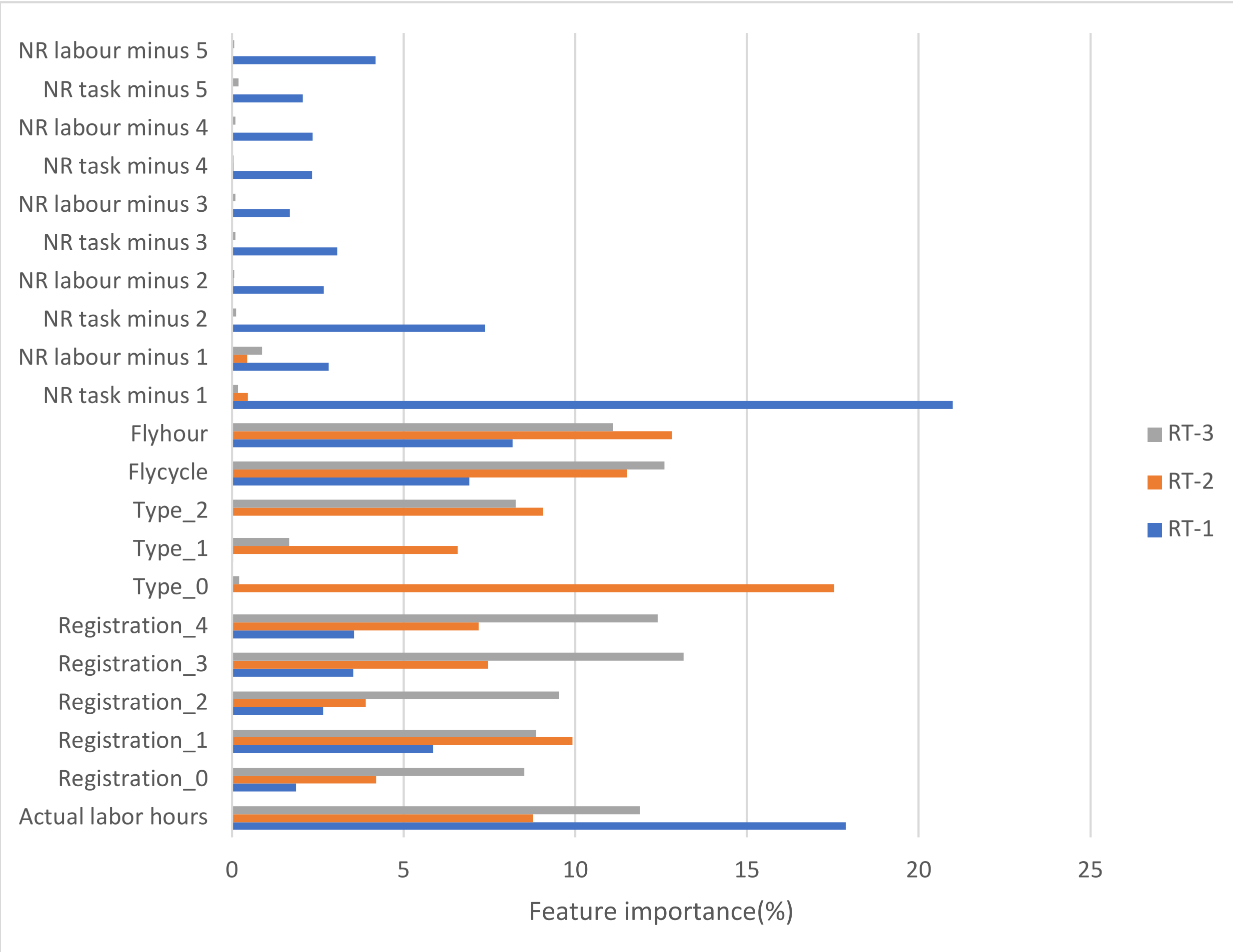
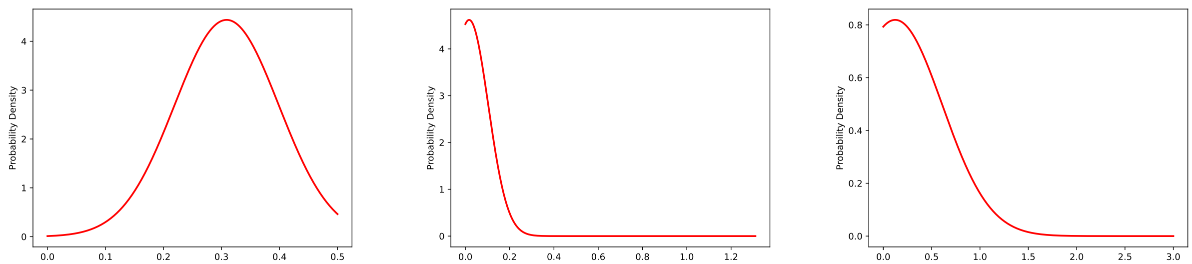



**Fig. 2 Feature importance for the three routine tasks examined in this work.**

### B. Prediction Distribution for Each of the Three Presented Routine Tasks

This approach allows us to delve into the individual predictions generated by each tree, thereby facilitating an in-depth analysis of prediction reliability and the factors contributing to significant errors.

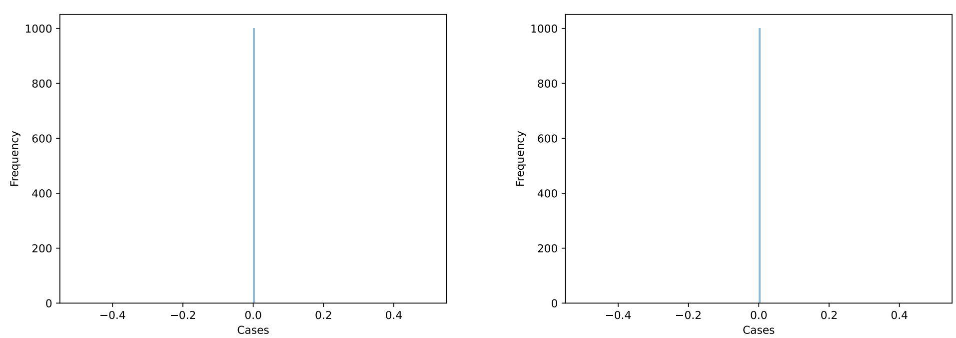
For RT-1, three distinct scenarios come to the forefront as shown in Figure 3. Figure 3a is characterised by a minimal variance in the prediction distribution, indicative of highly reliable predictions and a negligible margin of error. Conversely, Figure 3b reveals a peculiar case marked by a considerable error, albeit with a confined variance. In this instance, the substantial error behaves as an outlier that challenges the model's predictive capabilities. Figure 3c features both substantial error and elevated variance. This conveys a diminished level of prediction confidence, rendering it less reliable compared to other instances. The variations in error can be rationalized by the presence of unanticipated outliers and suboptimal predictions. Suboptimal predictions may stem from a lack of data, given that diverse aircraft exhibit disparities in the volume of available data. Furthermore, when historical records for a particular case display increased irregularity, it can lead to model disorientation, impeding its ability to discern underlying patterns.




**(a) Case:742-Error:0.159** **(b) Case:2423-Error:8.6** **(c) Case:3318-Error:2.41**

**Fig. 3 RT 1 - Distribution of the prediction results for three different cases.**

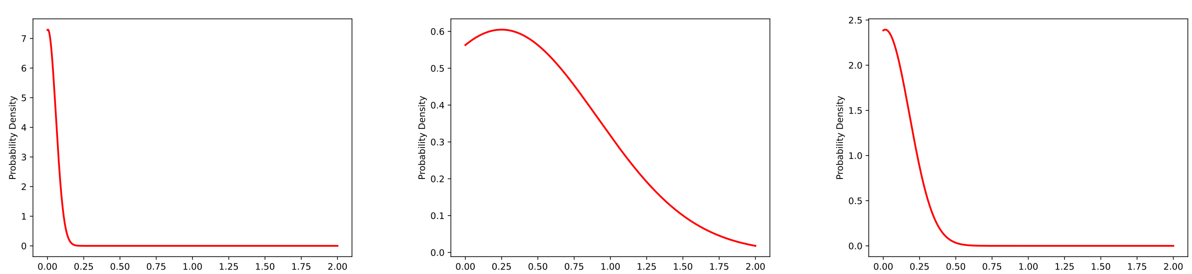
The observations within RT-2 (Figure 4) yield additional intriguing insights. The majority of cases within this JIC defy the establishment of a discernible distribution pattern. This is attributed to the model's robust recognition of underlying patterns, to the extent that most of decision trees uniformly converge on the same outcomes makes the results of other trees that disagree insignificant, as evidenced in Figure 4a and Figure 4b.




**(a) Case:821-Error:1.152** **(b) Case:1194-Error:0.012**

**Fig. 4 RT 2 - Distribution of the prediction results for three different cases.**

RT-3 has similar results to RT-1. Figure 5a displays a scenario where the model confronts an outlier, leading to a considerable error but with minimal variance. On the other hand, Figure 5b typifies a situation where the model exhibits a higher degree of uncertainty in its predictions, manifesting as a substantial error along with pronounced variance. Finally, Figure 5c reaffirms an ideal circumstance, characterized by both minimal error and variance.




**(a) Case:574-Error:1.41** **(b) Case:1910-Error:1.84** **(c) Case:940-Error:0.011**

**Fig. 5 RT 3 - Distribution of the prediction results for three different cases.**

The examination of distribution profiles across diverse cases and distinct RTs underscores the general success and reliability of our predictive methodology in a majority of instances. For those marked by substantial variance, we consider that the primary limitation may be attributed to data insufficiency, especially for those RTs that carried out less often. Despite the extensive dataset at our disposal, the fragmentation of data into distinct cases may yield only a limited dataset to draw upon, necessitating consistency in historical data with specific aircraft as an additional constraint.

## VI. Future Work

Nevertheless, the good accuracy achieved with the model, improvements are required towards turning into a tool that may be employed by airlines. To further enhance the scope of our research and refine its practical applications, we propose the following avenues for future exploration:

- **Work Package Prediction**: To optimize the utility of our research for end-users, there exists an opportunity to provide predictions for each work package they encounter. Therefore, it would be highly advantageous to develop a Work Package (WP) predictor that builds upon the existing RT predictor, as each work package consists of multiple RTs. This endeavour would significantly enhance end-user convenience and offer valuable insights.
- **Material usage prediction**: An additional consideration in aircraft maintenance pertains to the necessary materials. Owing to the inherent characteristics of certain materials, pre-storing them in inventory is neither economically viable nor feasible due to expiration concerns. Nonetheless, the temporal lag incurred in material procurement and awaiting delivery poses a potential disruption to the maintenance schedule. Consequently, a more advanced forecasting of material usage, coupled with a seamless integration with the inventory system, serves to preemptively mitigate material shortages, thereby safeguarding the integrity of the maintenance schedule.

## VII. Conclusion

This research introduces an innovative supervised learning model designed to forecast non-routine task workloads. Leveraging a rich and extensive dataset characterised by both scale and data quality, we have succeeded in attaining a commendable level of predictive accuracy and reliability. Furthermore, our comprehensive analysis offers the means to pinpoint the root causes behind substantial prediction errors, providing valuable insight.

Our analysis found that aircraft age is often in play for correct predictions on non-routine task workload. However, different RTs have shown quite distinctive importance. Historical records, in addition to actual labor hours, constitute essential factors in the prediction process. Furthermore, a limited dataset size serves as the primary driver of substantial prediction variance, owing to its heightened susceptibility to outlier influence. In exceptional cases, inadequate data quality can similarly result in significant prediction variances, particularly when on-site personnel fail to accurately document all relevant information. Regarding the level of confidence in the values output by the model, the average MAE improvement of labour hour prediction is around 20% and the average accuracy improvement of number of tasks prediction is around 21%.

This study serves as a valuable resource in the domain of non-routine task workload prediction, presenting airlines with a strategic avenue for optimizing maintenance task scheduling and reducing operational costs. Furthermore, we anticipate that our forthcoming initiatives will result in the creation of an algorithm that incorporates the prognostication of material utilization, seamlessly integrated with an inventory system.

## References

[1] Ingeborg de Pater, Arthur Reijns, and Mihaela Mitici. "Alarm-based predictive maintenance scheduling for aircraft engines with imperfect Remaining Useful Life prognostics". In: *Reliability Engineering and System Safety* 221 (2022). DOI: 10.1016/j.ress.2022.108341.

[2] Juseong Lee and Mihaela Mitici. "Multi-objective design of aircraft maintenance using gaussian process learning and adaptive sampling". In: *Reliability Engineering and System Safety* 218 (2022). DOI: 10.1016/j.ress.2021.108123.

[3] Markou C. 2022. URL: https://www.iata.org/contentassets/bf8ca67c8bcd4358b3d004b0d6d0916f/fy2021-mctg-report_public.pdf.

[4] Chellappan Sriram and Ali Haghani. “An optimization model for aircraft maintenance scheduling and re-assignment”. In: *Transportation Research Part A: Policy and Practice* 37.1 (2003), pp. 29–48. ISSN: 0965-8564. DOI: https://doi.org/10.1016/S0965-8564(02)00004-6.

[5] Al-Fares Hesham K and Duffuaa Salih O. *Maintenance Forecasting and Capacity Planning*. Ed. by Mohamed Ben-Daya et al. London: Springer London, 2009, pp. 157–190. DOI: 10.1007/978-1-84882-472-0{_}8.

[6] Josanne Aungst et al. “Planning of non-routine work for aircraft scheduled maintenance”. In: *Proceedings of the 2008 IAJC-IJME International Conference, Purdue University*. 2008.

[7] Maurice Pelt, Konstantinos Stamoulis, and Asteris Apostolidis. “Data analytics case studies in the maintenance, repair and overhaul (MRO) industry”. In: *MATEC Web of Conferences* 304 (2019). DOI: 10.1051/matecconf/201930404005.

[8] Premaratne Samaranayake and Senevi Kiridena. “Aircraft maintenance planning and scheduling: An integrated framework”. In: *Journal of Quality in Maintenance Engineering* 18.4 (2012). ISSN: 13552511. DOI: 10.1108/13552511211281598.

[9] M Zorgdrager et al. “A predictive method for the estimation of material demand for aircraft non-routine maintenance”. In: *20th ISPE International Conference on Concurrent Engineering, CE 2013 - Proceedings* (Oct. 2013), pp. 507–516. DOI: 10.3233/978-1-61499-302-5-507.

[10] Krasin Georgiev and Nikolay Vachev. “Predicting the unscheduled workload for an aircraft maintenance work package”. In: *AIP Conference Proceedings* 2557.1 (Oct. 2022). ISSN: 0094-243X. DOI: 10.1063/5.0104651.

[11] Masayuki Kishino et al. “Fatigue life prediction of bending polymer films using random forest”. In: *International Journal of Fatigue* 166 (2023). DOI: 10.1016/j.ijfatigue.2022.107230.

[12] E Syed Mohamed et al. “A hybrid mental health prediction model using support vector machine, Multilayer Perceptron, and Random Forest algorithms”. In: *Healthcare Analytics* 3 (2023). DOI: 10.1016/j.health.2023.100185.

[13] Apoorv Maheshwari, Navindran Davendralingam, and Daniel A. DeLaurentis. “A Comparative Study of Machine Learning Techniques for Aviation Applications”. In: *2018 Aviation Technology, Integration, and Operations Conference*. Reston, Virginia: American Institute of Aeronautics and Astronautics, June 2018. ISBN: 978-1-62410-556-2. DOI: 10.2514/6.2018-3980.

[14] Weili Yan and Jun-Hong Zhou. “Predictive modeling of aircraft systems failure using term frequency-inverse document frequency and random forest”. In: *2017 IEEE International Conference on Industrial Engineering and Engineering Management (IEEM)*. 2017, pp. 828–831. DOI: 10.1109/IEEM.2017.8290007.

[15] Augustin Degas et al. “A Survey on Artificial Intelligence (AI) and eXplainable AI in Air Traffic Management: Current Trends and Development with Future Research Trajectory”. In: *Applied Sciences* 12.3 (2022). ISSN: 2076-3417. DOI: 10.3390/app12031295.

[16] S G ten Zeldam. “Automated Failure Diagnosis in Aviation Maintenance Using eXplainable Artificial Intelligence (XAI)”. PhD thesis. Netherlands: University of Twente, July 2018.