Probabilistic Modeling of Workload Deviations in Aircraft Light Maintenance Using Bayesian Networks

Júlia Dâmaso de Aguiar Viveiros
julia.viveiros@tecnico.ulisboa.pt

Instituto Superior Técnico, Universidade de Lisboa, Portugal

January 2021

Abstract

This dissertation, developed in the context of preventive maintenance on commercial aircraft, aims at analyzing probabilistically the discrepancies between workloads predicted by aircraft manufacturers for light periodic inspections tasks and the ones registered upon their execution. The dissertation is based on a sample of historical data of ”A-checks” inspections conducted on Airbus A330 and A340, between 2013 and 2020, provided by *MESA – HiFly*'s Maintenance and Engineering services provider. A Bayesian networks framework is used to model probabilistically the deviations from the workloads predicted by the aircraft's manufacturer, from the available data. The adopted methodology requires a review of the evolution, planning and registration of aircraft maintenance, along with a detailed analysis of workcards and the *Maintenance Planning Document*. The basic principles and processes of the development of Bayesian networks models through data and validation through sensitivity analyses are presented. Two Bayesian networks are developed from the data: one for the modeling of checks, where the sensitivity analysis identifies that $2A$ items and maintenance performed at the base station are likelier to present high deviations, and one for tasks, where it is evaluated that *General Visual Inspections*, zones 400 and 700, and *Powerplant* and *Airframe* skills are the variables with higher impacts in the deviations. Two practical examples of application of the models for maintenance capacity planning are presented. It is expected that the dissertation will bring benefits in the planning of these inspections, given that the degree of uncertainty of this activity can be reduced through the developed models.

Keywords: Bayesian Networks; Aircraft Maintenance; A-checks; Workload Deviations; Capacity Planning.

1. Introduction

1.1. Motivation

A substantial amount of information is generated when performing aircraft maintenance (about vehicles, operators, interventions), and it is still yet to provide a decisive competitive advantage to Part 145 or Maintenance, Repair and Overhaul (MRO) organizations [1] due to the fact that little or no sensitivity and robustness analysis of aircraft maintenance data is performed by airlines [2].

An aircraft maintenance check contains several tasks, for which the workload suggested in the manufacturer's Maintenance Planning Document (MPD) [3, 4], in Man/Hours (M/H), does not always agree with the actual values registered in the workcards by the operator's maintenance technicians upon performing the task – some tasks require less manpower while others require significantly more than expected, which can be represented by a problem of an essentially probabilistic nature.

1.2. Topic Overview

The different types of aircraft maintenance events are briefly explained in Table 1. The M/H unit is the time required for a labor unit to finish a unit work amount [5].

The factors contributing to delays during A-check inspections are studied by Mofokeng and Marnewick [6], and according to the authors, these delays result in the loss of revenue because of potential penalties, and the identification of what aspects influence said delays can help airlines to identify the gap between best practice and current practice. By knowing the causes of delays, the maintenance company can adjust their operational strategies. An A-check is the most crucial requirement in scheduling because of the relatively short interval between the next required check.

Bayesian Networks (BNs) are frequently mentioned in the literature, as reviewed by Weber et al. [7], due to their ability to model complex sys-

Table 1: Main types of aircraft maintenance work.

Maintenance Type	Description
Line Maintenance	Routine tasks with low intervals, performed at line stations.
Base Maintenance	Performed at airline's base maintenance station that has the manpower to do all kinds of work.
Letter Checks	A-, C- or D- checks, ranging from visual inspections to exhaustive overhauling actions.
Light Maintenance	A-checks, executed in intervals of 800FH, requiring $50-70$M/H to be completed.
Intermediate Maintenance	C-checks, performed every $20-24$ months, taking up to 7 days.
Heavy Maintenance	D-checks, done every $6-10$ years, requiring downtimes of over 7 days.

tems and make predictions regarding the occurrence probability of events, along with the possibility to update probabilities according to evidences [8], making them an adequate and powerful tool to address problems regarding uncertainties.

The aeronautics industry aims to come up with important changes in its maintenance strategies, because despite the arising number of solutions, it is still a highly unpredictable field. Ferreiro et al. [9] develop a Bayesian network to model the case of predicting brake wear, in a study that explains the use of BNs as a prognostic technique applied to aircraft maintenance. Dinis et al. [1] address the aircraft maintenance capacity planning problem, and the applicability of BNs as a Big Data and Predictive Analytics (BDPA) tool is studied – given their probabilistic nature, BNs are a reliable technique to address the uncertainty of maintenance workload estimations, therefore improving the MRO's capacity planning decision-making process. If the available capacity is higher than required, there is underutilization of resources and financial inefficiency occurs; if, on the contrary, the available capacity is lower than required, delays will happen with potential financial penalties and damages to the reputation of the maintenance organization.

1.3. Objectives

This dissertation aims to develop a probabilistic model for the workload of a maintenance A-check, as an attempt to reduce the unpredictability associated to the maintenance planning process through the identification of the variables that could have an impact on a task's workload. Bayesian networks present several advantages representing problems of probabilistic nature; therefore, BNs are developed to model the causal relationships between variables such as the aircraft's model, Flight Hours (FH) and tail $\#$ or even the task's zone and skill codes, and the check or task's total workload, from real maintenance data. A sensitivity analysis is then performed to quantify the influence of each parameter contributing to the output – the workload deviation of a single task or a check.

The data used is provided by a EASA Part 145 regarding the light maintenance work done on a portuguese wet lease and charter airline's fleet, more specifically, A-check inspections for the Airbus A330 and A340 (pictured in Figures 1 and 2).

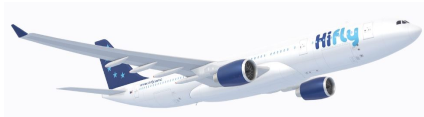

Figure 1: Airbus A330-200 [10].

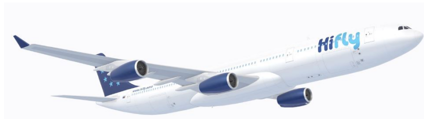

Figure 2: Airbus A340-300 [11].

2. Methodology

2.1. Maintenance Tasks and Checks

The generic list of tasks, required skills and maintenance zones, as stated by EASA [12], is presented in Table 2.

An A-check consists of a general inspection of the airplane with specific target areas opened, requiring about 20 to 60 M/H to be completed. This check's periodicity varies by aircraft type, cycle count, or even number of hours flown since the last check [13], though it is typically performed every 800 FH.

Not all A-checks are the same – items are numbered so that $2A$ items are carried out in every other A-check, in an $A2$ inspection, and both $2A$ and $4A$ items are performed in every fourth inspection, called an $A4$. $1A$ items are performed in every A inspection, and after every $A4$ inspection, a new cycle of four begins.

2.2. Bayesian Networks

Causality can be graphically represented in BNs [14], which are a type of quantitative causal model structure based on the Bayes' theorem that represents and processes knowledge in a probabilistic way, making it an excellent tool for reasoning under uncertainty [15].

The Bayes' theorem works by taking old probabilities along with new data as inputs, and delivering new updated probabilities as outputs. Dividing a domain Ω into n mutually exclusive sets $A_1, A_2, ...A_n$, and for a certain random variable B, then (notice equation (1)):

Table 2: Task, Zones and Skill Codes List.

Task	Definition	Skill	Definition	Zone	Definition
DET	Detailed Inspection	AF	Airframe	100	Fuselage Lower
DIS	Discard	AV	Instrument	200	Fuselage Top
FNC	Functional Check	CA	Cabin Utility	300	Stabilizers/Empennage
GVI	General Visual Inspection	EL	Electrical	400	Nacelles/Pylons
LUB	Lubrication	EN	Powerplant	500	Left Wing
OPC	Operational Check	NDT	Non-Destructive Test	600	Right Wing
RST	Restoration	RA	Radio	700	Landing Gear Compartment
SDI	Special Detailed Inspection	UT	Utility	800	Doors
SVC	Drain, Servicing, Replenishment			900	Lavatories & Galleys
VCK	Visual Check				

$$P(A_k|B) = \frac{P(A_k) * P(B|A_k)}{\sum_j P(A_j) * P(B|A_j)} \quad (1)$$

For this instance, given a posterior probability P_m, the probability that the next observation will be C is given by equation (2):

$$P_m(C) = \sum_j P_m(A_j|C) = \sum_j P_m(C|A_j) * P_m(A_j) \quad (2)$$

A BN consists of a qualitative part, a Directed Acyclic Graph (DAG) along with a quantitative part, a Conditional Probability Table (CPT) (or a set of them). The DAG (in Figure 3) contains nodes representing random variables and directed arcs representing dependencies or causal relationships between variables; then, a joint probability distribution is defined over the variables depending on the directed arcs, which makes inference through conditional probabilities possible. The directions of links between variables (directed arcs) represent the parent-child relationships, with the arrow head pointing in the direction of causality, i.e. the child [16].

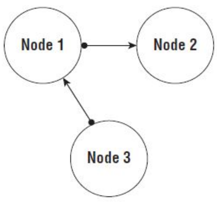

Figure 3: Example of a directed acyclic graph [17].

The qualitative and quantitative parts of a BN can be defined manually or through computational methods capable of inferring the network's structure and CPT from the data, and there is a five-step process often mentioned in the literature [8,15] for the development of a BN structure: 1. Delineating the objectives of the model; 2. Defining the variables; 3. Designing the network's graphical structure; 4. Building the network's CPT; 5. Validating the model.

The chosen structure learning algorithm is *Bayesian Search*, the most popular one [18], and the parameter estimation algorithm is the Expectation-Maximization (EM) algorithm, which computes maximum-likelihood estimates for the parameters from datasets that may contain missing values [19,20].

2.3. Sensitivity Analysis

Dinis et al. [15] propose a sensitivity measure formulated in terms of variation in the posterior distributions of the model variables resulting from introducing an evidence on a state of the model's output, as follows.

Let $X_{i,j}$ be the variable i of the BN model with $j = 1, ..., m_i$ states and $P(X_{i,j=1,...,m_i}|Y = e)$ its posterior probability distribution when providing the evidence e to a particular state of the *Y* variable. The variation in the posterior probability distribution of the variable $X_{i,j}$ when Y changes from state e to f is given by equation (3):

$$\Delta P(X_{i,j=1..m}|Y) = P(X_{i,j=1..m}|Y = f) - P(X_{i,j=1..m}|Y = e) \quad (3)$$

A global measure of the importance of the variable X_i on the variable Y is then defined based on ΔP as equation (4) writes it:

$$S_{X_i} = \sqrt{\sum_{j=1}^{m} \frac{(\Delta P(X_{i,j}|Y))^2}{2}} \quad (4)$$

In addition to the global sensitivity measure S_{X_i}, a state j sensitivity measure $S_{X_{i,j}}$ of variable X_i can also be derived as visible in equation (5):

$$S_{X_{i,j}} = \frac{\Delta P(X_{i,j}|Y)/2}{S_{X_i}} \quad (5)$$

In which $\Delta P(X_{i,j=1,...,m}|Y)$ is given by equation (3) and S_{X_i} is the global sensitivity measure of the variable X_i, given by equation (4), used to normalise the state variation of posterior probabilities.

The state sensitivity measures the relative variation of the state's posterior distribution.

2.4. Description of Maintenance Dataset

A total of 127 A-checks were analyzed: 67 for the Airbus A330, and 60 for the Airbus A340.

Table 3: Average values for checks.

Aircraft Model	A330			A340		
Check Item	**1A**	**2A**	**4A**	**1A**	**2A**	**4A**
# of Samples	38	21	8	36	17	7
Average # of MPD Tasks	32	13	9	37	21	10
Average MPD Tasks Workload [M/H]	31,68	13,30	11,90	38,63	18,31	15,75
Average Estimated MPD Tasks Workload [M/H]	11,15	3,46	10,44	14,54	5,07	5,90
Average Total Workload Deviation	184,8%	298,8%	13,8%	166,8%	260,4%	165,4%
Average Per Task Workload Deviation	412,6%	490,9%	102,8%	327,2%	585,1%	323,5%

Breaking down the dataset in terms of variables, every check is classified according to the aircraft's model, tail number, age, in FH and the location where the maintenance event takes place.

The location of the maintenance event is relevant because, as Rosales [13] explains, the manpower and facilities at line stations are usually more limited, which is why it is relevant to distinguish which checks were done at a hangar or maintenance base (where all the necessary tools and equipment are nearby) and which were performed at the airport apron (where the technicians must get the tools and equipment from the line maintenance station, which can cause unexpected delays), especially because the studied inspections do not require a high number of tools and consumables to be fetched. Furthermore, because the airline only operates one maintenance base station, it is likely that it will often be more congested in terms of workload, and inconveniences such as having to wait for units or spare parts to arrive makes this situation prone to delaying the maintenance process.

On the other hand, and because every check is composed by tasks, there is a need to identify variables that assess tasks individually. A task is then defined by its task code, zone code, and skill code.

3. Preliminary Analysis of Maintenance Workloads

Table 3 presents the obtained statistics for all three $1A$, $2A$ and $4A$ items, that result from the averages of the values computed for each individual check.

The *# of MPD Tasks* accounts for the number of tasks in the check sourced from the MPD; the *MPD Tasks Workload* is the sum of the registered workloads of said tasks, as stated on the workcard; the *Estimated MPD Tasks Workload* is the sum of the workloads of the aforementioned tasks, as stated on the MPD (i.e. their expected required workload); the *Workload Deviation* (WL_Dev) variables (*Total* and *Per Task*) refer to the ratio between the registered deviations (in M/H) and the suggested workload from the MPD (also in M/H), applied to both check and task values. This is formulated in equation 6:

$$WL\_Dev = \frac{Actual\,WL - Planned\,WL}{Planned\,WL} \tag{6}$$

In which the $Actual\,WL$ is the one registered in the workcards, while the $Planned\,WL$ is the one stated on the MPD.

The above-mentioned statistics for both aircraft indicate that the MPD is very optimistic with regard to the required workload for aircraft maintenance tasks given that, in average, the observable discrepancies don't have a negligible order of magnitude.

From the table, it can be stated that generally $2A$ items register the largest discrepancies with the highest workload deviations ($298,8\%$ and $260,4\%$ for the A330 and A340, respectively), while on the opposite end, $4A$ present the smallest ($13,8\%$ and $165,4\%$).

4. Probabilistic Modeling of Maintenance Workload Deviations by Bayesian Networks

4.1. BN Modeling

The Bayesian networks to model the A-checks were developed using the computer software *GeNIe* [21], with the purpose of obtaining workload deviation predictions for maintenance checks.

Following the steps mentioned in 2.2, the BN model is developed to get a prediction of the workload deviation for light maintenance inspections. For this reason, there is a need to build two distinct BN models: one for assessing the workload deviation of a check (weighing in parameters such as the type of check being performed, the location of the station, the aircraft's model, tail number and age in FH) and another one for evaluating the workload deviation for a single task (taking into account the zone where maintenance is required, the aircraft's model, the job's skill and task codes).

It is important to refer that in order to get models with discrete variables only, some states are grouped into classes, namely from the variables FH and Task/Check Workload Deviation. The intervals of the FH classes are chosen in a manner that the data is evenly distributed, for both aircraft ($30\% \pm 13\%$ of the samples in each class).

The BN model's graphical structure is both forced and also assumed by the software: on one hand, some causal relationships make theoretical sense; on the other hand, the software infers the parent-child relationships between the remaining parameters.

The CPTs assigned to each variable are obtained through learning techniques from maintenance data provided to the software. The BN models are fully quantified (in terms of the *a priori* knowledge), as depicted in Figures 4 and 5 that also present the characterization of the dataset.

With this framework it is possible to simulate scenarios with respect to future work generated by providing evidences to specific states of model variables. This ability to experiment possible scenarios can be a valuable tool for the airline, provided it becomes possible to see the category that the planned work falls into, and therefore get reliable information on the most likely output event.

Another characteristic of this framework is the possibility of updating it as new knowledge becomes available. If information on future checks and tasks is to be introduced in the BN model, its accuracy (regarding the delivered results) will be continuously improved.

4.2. Sensitivity Analyses

For this validation, the several states of the *Workload Deviation* variable are quantitatively described as presented in Table 4. The *Very High* classification only applies to the *Task Workload Deviation* variable.

Table 4: Qualitative classification of workload deviations.

Deviation Range	Classification
$< 0\%$	Negative
$0 - 100\%$	Low
$100 - 500\%$	Moderate
$500 - 1000\%$	High
$> 1000\%$	Very High

This classification provides an easier understanding and representation of the possible consequences (negative to moderate deviations are not considered as critical as moderate to high), and it makes it possible to evaluate the ideal settings that take place in each chosen range.

A global sensitivity analysis allows for a quantification of how each variable affects the model's outcome. This analysis is conducted in both models in order to find the most relevant variables in the model, and as a criteria for selecting the most appropriate ones to conduct local sensitivity analysis.

Figure 6 presents the global sensitivities S_{X_i} of the check BN model variables calculated through equation 4, when changing the evidence in the *Workload Deviation* from negative to low, low to moderate and moderate to high.

It is quite evident that the variable with a consistently strong impact on the total workload deviation is the *Check* being performed. This finding makes sense due to the fact that different checks require different sets of tasks, thus the type of workload is the most relevant variable in the model.

The *Location* proves to be a somewhat relevant input when assessing low to moderate deviations. This is a parameter that definitely requires further investigation (namely, a local sensitivity analysis) in order to figure out if this global sensitivity is more reactive to the evidence being on the line (NO_BRU) or base (BRU) maintenance station.

The influence of the *Tail Number* is approximately constant throughout the classes provided it is only a measure of how wide the sample is with respect to different aircraft.

Regarding the aircraft's age, the *FH* appears to gain impact as the workload deviations increase, which favors the idea that delays can in fact be potentiated by the aircraft's usage parameter.

Figure 7 presents the results of the global sensitivities S_{X_i} of the task BN model variables calculated with equation 4, when changing the evidence in the *Workload Deviation* from negative to low, low to moderate, moderate to high, and finally, from high to very high. Note that the *Task* variable refers to the Task Code.

Starting with the *Model* variable, its influence is consistently low throughout the deviation classes except for the last one: when registered workload deviations are high to very high, the two aircraft models present a different pattern. This is corroborated by the evidence from Table 3. For example, a 2A item for the A330 presents an average per task workload deviation of $490,9\%$, categorizing this as a moderate deviation, while the same item for the A340 averages a $585,1\%$ of deviation per task, which belongs to the superior category of high deviation.

Because this BN model is focused on all the tasks and not on checks as packages, the *Check* variable is not a very important, so its inconsistent pattern is disregarded.

The *Zone* where the maintenance work is being performed is one of the most important parameters, as some zones require technicians to be more careful and thorough than others (one can suppose that engine maintenance tasks will be the case), meaning these tasks are more prone to delays.

The required *Skill* also has a relevant impact on the output of the model. This is another variable that demands a local sensitivity analysis, in order to identify which states potentiate this contribution.

The most dominant parameter in the model is the *Task* code, which classifies the type of work that must be done, hence it is more than acceptable that this is the input with the largest influence on the task's workload deviation.

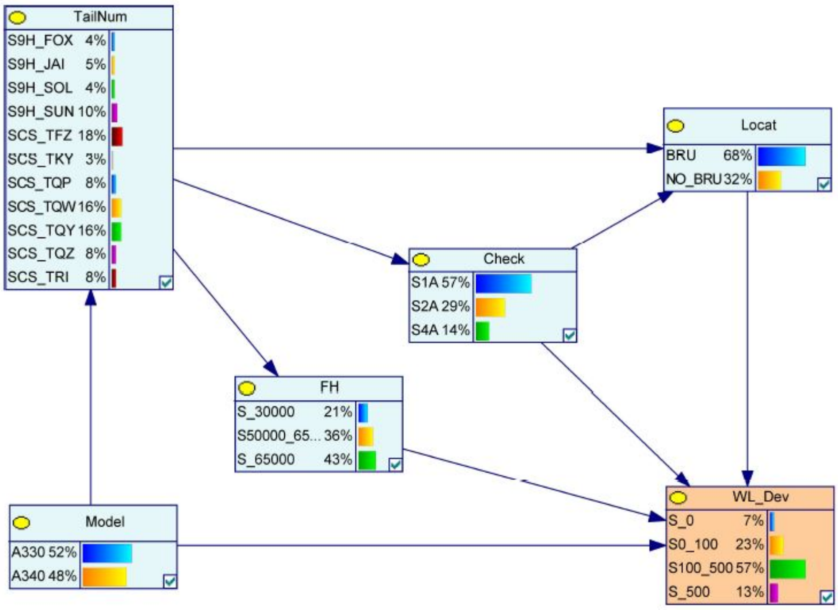

Figure 4: Bayesian network for the checks.

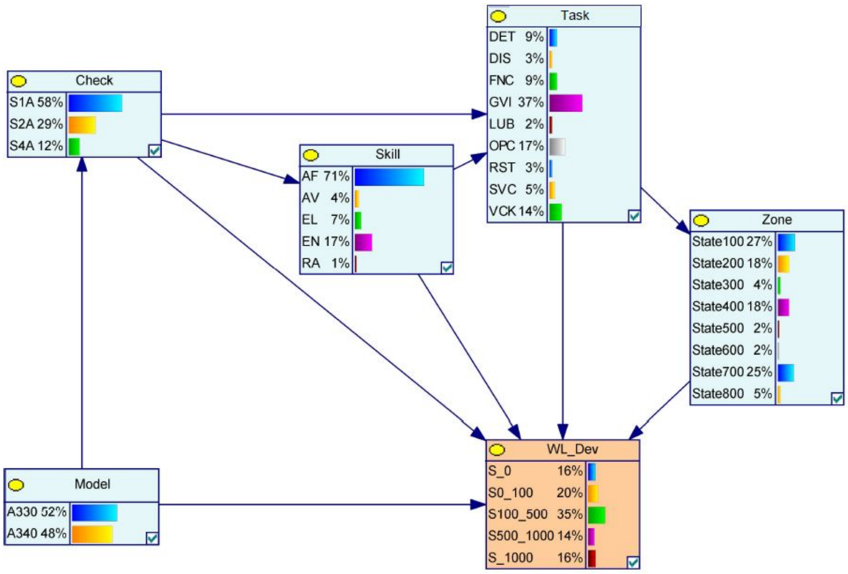

Figure 5: Bayesian network for the tasks.

The local sensitivity analysis, performed with equation 5, allows for a quantitative understanding, on a deeper level, of how each variable's state influences the model outcome.

Regarding the check BN model, the parameters chosen to perform a local sensitivity analysis are the *Check* and *Location*, which are the ones that the Check *Workload Deviation* proved to be more sensitive to.

The analysis confirms that $2A$ items have the highest weight when assessing moderate and high deviations, and a negative contribution on the negative to low range – this check is undoubtedly the most critical one in terms of overtimes. The fact that $4A$ items present only negative or null (for the moderate to high category) sensitivities is in accordance with what the previous data analyses had been pointing out: that these tasks are the least prone to incur in duration discrepancies.

It is also confirmed that this maintenance base station tends to be more prone to deviations of higher magnitude: there is a positive sensitivity on the low to moderate and moderate to high states, while line stations show a positive sensitivity for the negative to low range of deviation.

Shifting the focus to the task BN model, the parameters submitted to a local sensitivity analysis are the ones that provided the largest impacts on the *Task Workload Deviation* upon performing the global sensitivity analysis: *Task* (Code), *Zone* and *Skill*.

Starting with the *Task*, Lubrication (LUB), Dis-

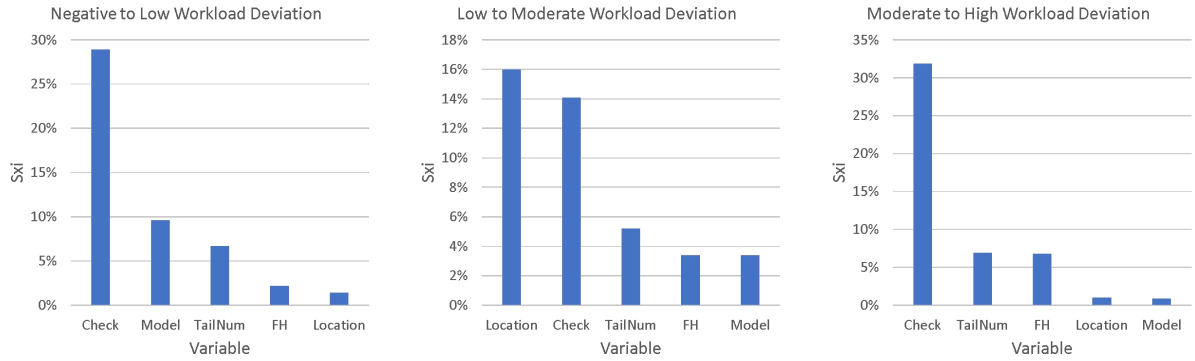

Figure 6: Global sensitivities S_{x_i} of model variables when the check workload deviation changes from negative to low, low to moderate and moderate to high.

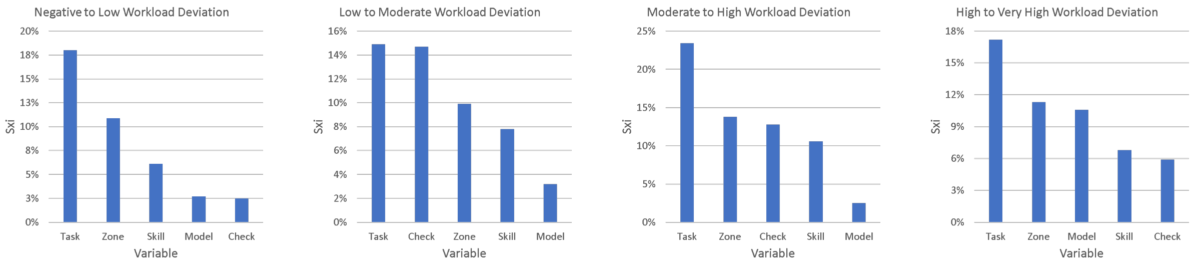

Figure 7: Global sensitivities S_{x_i} of task model variables when the task workload deviation changes from negative to low, low to moderate, moderate to high and high to very high.

card (DIS) and Restore (RST) tasks do not tend to vary much with respect to their suggested workloads – these state present nearly negligible sensitivities.

Servicing (SVC), Functional Check (FNC) and Detailed Inspection (DET) are also not very critical states – these tasks have the purpose of maintaining inherent design capabilities or determining if an item performs within specified limits, and because inspections have such low intervals it is not often for the components to require extra work.

An Operational Check (OPC) consists of operating the aircraft to make sure all systems function accordingly – this task can be delayed because although each system has its own function, said function is not independent from other systems of the aircraft, hence the strong sensitivity of this state on the moderate to high class.

On the other hand, the General Visual Inspection (GVI) state is undoubtedly the one that presents the greatest impacts on all classes of deviation, except for moderate to high – although the probability of performing a GVI with a high deviation is still higher than for the rest of the states, this range covers most of the Task Code possible states in an almost uniform manner, while other ranges are more focused on specific states.

GVI tasks are supposed to be performed at a maximum of an arm's length of distance when examining the components, which means that the inspection can be rather quick if the component is visible (which explains the low deviations) or very long if the component requires the opening (and closing) of some areas to get to it – hence the strong sensitivity registered on the high to very high deviations.

The *Zones* that present the largest impacts on the highest class of discrepancy (High to Very High) are 400 – Nacelle/Pylons, and 700 – Landing Gear Compartment. This is definitely an expected result, as the engines and landing gears require the technicians to be more thorough with their work in these zones due to their fundamental role in the aircraft. The engines are crucial elements of an aircraft, and the landing gears are components that get worn off between cycles, unlike many other aircraft zones that require a less detailed inspection.

Zone 100, the lower part of the aircraft's fuselage, has a higher probability of presenting a negative workload deviation rather than low, and this likelihood reduces throughout the output states. Maintenance tasks in this zone should not be very critical.

Zone 800, the vehicle's doors, presents a high state sensitivity in the low to moderate range, which is acceptable because inspections in this zone do not require the same level of detail as, per example, zones 200 and 300 (Fuselage Top and Stabilizers/Empennage, respectively) that are settled in the moderate to high category, meaning they might be prone to deviations (although still not as intensely as zones 400 and 700). This has to do with the fact that doors do not suffer from tearing or wear off between cycles as much as the remaining zones.

Without a doubt, Airframe (AF) and Powerplant (EN) are the most relevant *Skills*. This is an expected result, as AF skills comprise flaps/slats and landing gear, and EN skills include engines and Auxiliary Power Unit (APU) accessories (which is in line with the results obtained for the zones).

Because Radio (RA) skills are only required in 4A items, and seeing as these items present the lowest workload deviations in the model, it makes sense that the influence of this state is practically negligible.

4.3. Capacity Planning

The examples provided below serve the purpose of demonstrating the practical benefits of applying BNs to aircraft maintenance capacity planning. As Dinis et al. [1] state, capacity planning balances the expected workload with the available manpower, thus being responsible for the management of uncertainty between the tactical and operational decision levels.

The maintenance services provider receives a request to perform an intervention at its maintenance base station, specifically a $2A$ check, on an Airbus A330 (tail number CS-TFZ) with $70\,000$ FH, for which the MPD predicts a total required workload of 3 M/H.

As presented in Figure 8, by instantiating the aforementioned states for each variable on the check BN, the operator gains knowledge on the probabilities of incurring in each of the states of workload deviation, as represented in Table 5.

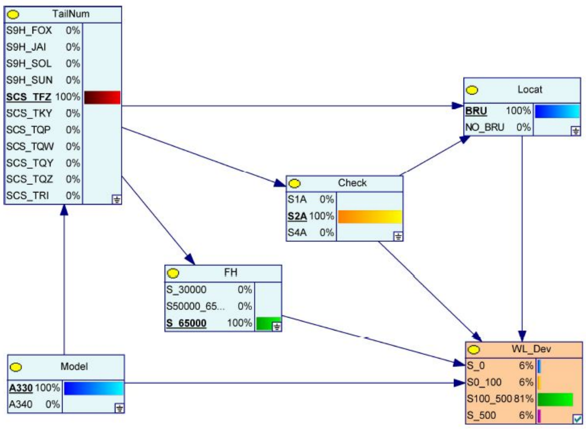

Figure 8: Capacity planning example 1.

The maintenance services provider is now planning the same check from the previous example, and desires to gain knowledge on the distribution of workload deviations regarding the required skills, in order to allocate the technicians and plan their schedules accordingly (at a maintenance base, technicians are differentiated by teams of skills).

Table 5: Probabilities of example 1 workload deviations.

WL_Dev	WL [M/H]	Prob.	Cumulative Prob.
<0%	[0, 3]	6,25%	6,25%
0-100%	]3, 6]	6,25%	12,50%
100-500%	]6, 18]	81,25%	93,75%
>500%	]18, +∞)	6,25%	100,00%

Table 6 presents the distribution of skills per check.

Table 6: Skills distribution per check.

	AF	AV	EL	EN	RA
1A	82,70%	2,97%	2,97%	11,08%	0,28%
2A	46,49%	3,24%	11,35%	35,68%	3,24%
4A	76,25%	7,50%	13,75%	1,25%	1,25%

From the BN model it is also possible to build Table 7 with the probabilities of occurrence of each class of workload deviation for each skill in A330 $2A$ checks. Although Figure 9 only presents one case, with the toggling of the AF state, it is required to instantiate each skill separately and register the posterior probability distribution for the classes of workload deviation, keeping the other known states for the other variables toggled (A330 model and $2A$ check).

Table 7: Skills workload deviations in A330 2A checks.

	<0%	0-100%	100-500%	>500%
AF	18,86%	16,42%	30,47%	34,25%
AV	20,00%	30,00%	25,00%	25,00%
EL	11,51%	11,51%	53,96%	23,02%
EN	12,61%	49,56%	12,61%	25,22%
RA	15,70%	15,70%	15,70%	52,90%

In practical terms, when the operator computes the predicted M/H for the group of tasks of a certain skill in the check to allocate the manpower to the available teams, the BN model presents the probabilities for each class of deviation for this value (regardless of the states of the other variables).

As the table lists, for this example the most critical skill would be RA, due to the fact that it registers the highest probability of presenting a high workload deviation $P(WL\_Dev{>}500\%) = 52,90\%$.

5. Conclusions

The goal of this research is to find out if the predictions for the workloads of light periodic maintenance tasks in the aircraft's MPD are in line with those observed in practice, and to get an understanding of what variables are relevant for predict-

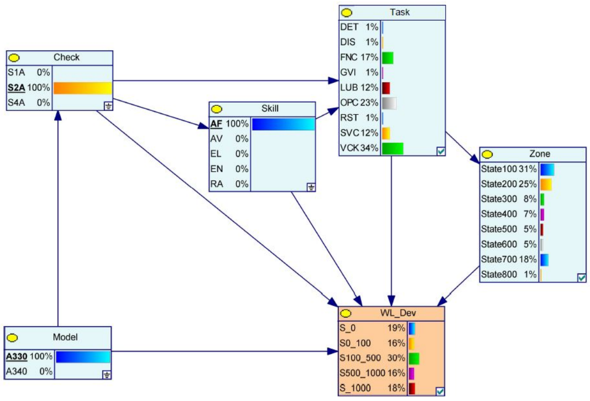

Figure 9: Capacity planning example 2.

ing workload deviations in maintenance checks, due to the high uncertainty on the maintenance reports.

For this, a framework able to analyze a problem of probabilistic nature – workload deviations in light maintenance checks – and reduce the level of uncertainty on the workload predictions is developed. Data on maintenance light inspections (A-checks) from an EASA certified Part 145 is gathered and several variables taking part in the checks are grouped, with the main objective of evaluating how they influenced the overload, in M/H, to complete the checks' tasks, and how that time deviates from the value suggested in the aircraft's MPD.

In order to assess the variables and their importance in the workload deviations, two Bayesian networks are developed: one for analyzing checks as a whole, and one for treating more specific data about tasks. The validation of the models is made through global and local sensitivity analyses, that aim respectively at identifying which parameters are of greater importance, and which of their states provide the greatest changes in the outputs.

For this specific maintenance operator the MPD manpower recommendations fail to deliver precise values, in M/H, for most of the reported tasks, with discrepancies that have a non-negligible order of magnitude. For the Airbus A330 and A340, respectively, tasks performed in 1A checks present average deviations of $412,6\%$ and $327,2\%$, tasks from 2A checks present average deviations of $490,9\%$ and $585,1\%$, and for 4A checks the average task deviations are of $102,8\%$ and $323,5\%$.

Overall, the item of check being performed has a decisive role over the check's workload deviation, with 2A items presenting the higher probabilities of being severely delayed, and on the opposite end, with 4A items registering high probabilities of taking less time than expected to be completed. Still under the check's workload deviation umbrella, inspections performed at the maintenance base (in Brussels) have a higher tendency of taking longer than those performed at line stations, due to the possibility of the base being overflowed in terms of workload, requiring longer waiting times for the arrival of units or spare parts.

Shifting to the factors that affect the duration of individual tasks, the task code is predominantly the one of greater importance, with the *General Visual Inspection* state presenting the higher state sensitivities for most ranges of delay. Regarding the maintenance zone, nacelle/pylons and the landing gear compartment (zones 400 and 700, respectively) prove to be the ones that require more thorough work, being accountable for the data in the high to very high category of task workload deviations. At last, assessing the required skills, negative to low and moderate to high workload deviations are sensitive to the *Powerplant* state that stands for engines and Auxiliary Power Unit accessories, while the low to moderate and high to very high ranges are vulnerable to the *Airframe* skill, that includes the landing gears. Engines and landing gears are components that can get very worn off between checks, hence requiring longer inspections to ensure the equipment is working accordingly.

The practical examples of the application of the developed BN models are presented to demonstrate their benefits for maintenance capacity planning – on the one hand, the check BN allows for estimating the check's workload while on the other hand, the tasks BN allows for allocating teams and scheduling shifts accordingly.

5.1. Suggestions

The use of the proposed BN models could be implemented in the maintenance planning process by the maintenance services provider whose data was analyzed – compared to the traditional estimation methods, the proposed BNs weigh in information about the skill, task and zone codes, as well as FH, location and tail number, which increases the accuracy of the workload estimations.

Concerning the actual execution of the task, a digital platform could be implemented such that technicians could sign in/out of tasks and fill out workcards in an electronic device, which would eliminate the need for paper and enhance the accuracy of the registering of the actual tasks' lengths (reducing the influence of human factors in the registration of the actual workloads). This could contribute to an improvement of the company's reliability levels if further studies were to be made using data of such sort.

5.2. Future Work

It would be extremely relevant to perform a study to investigate the economical consequences of workload deviations to maintenance companies (in regards to loss of revenue due to extra ground times or even risk analysis of possible losses of clients).

A project focused on the airline's maintenance base capacity planning would also be important, because it could allow for a reduction of the delays that light inspections tend to incur.

References

[1] D. Dinis, A. Barbosa-Povoa, and A. Teixeira, "Valuing data in aircraft maintenance through big data analytics: A probabilistic approach for capacity planning using bayesian networks," *Computers & Industrial Engineering*, vol. 128, pp. 920–936, 10 2018.

[2] W. Moudani and F. Mora-Camino, "A dynamic approach for aircraft assignment and maintenance scheduling by airlines," *Journal of Air Transport Management*, vol. 6, pp. 233–237, 10 2000.

[3] Airbus, "A330 Maintenance Planning Document (MPD)," Tech. Rep. MAR 01/20, 2010.

[4] Airbus, "A340 Maintenance Planning Document (MPD)," Tech. Rep. MAR 01/20, 2010.

[5] A. Kazaz, T. Acıkara, and B. Er, "Comparison of theoretical man-hour values and practical man-hour values in turkey," *International Journal of Advances in Mechanical and Civil Engineering*, vol. 3, pp. 2394–2827, 08 2016.

[6] T. J. Mofokeng and A. Marnewick, "Factors Contributing to Delays Regarding Aircraft During A-Check Maintenance," *2017 IEEE Technology and Engineering Management Society Conference, TEMSCON 2017*, pp. 185–190, 2017.

[7] P. Weber, G. Medina-Oliva, C. Simon, and B. Iung, "Overview on Bayesian Networks Applications for Dependability, Risk Analysis and Maintenance Areas," *Engineering Applications of Artificial Intelligence*, vol. 25, no. 4, pp. 671–682, 2012.

[8] H. Langseth and L. Portinale, "Bayesian Networks in Reliability," *Reliability Engineering and System Safety*, vol. 92, no. 1, pp. 92–108, 2007.

[9] S. Ferreiro and A. Arnaiz, "Prognostics Applied to Aircraft Line Maintenance: Brake Wear Prediction Based on Bayesian Networks," *IFAC Proceedings Volumes*, vol. 1, no. PART 1, pp. 146–151, 2010.

[10] HiFly, "Airbus A330-200 Specifications," Lisbon, Tech. Rep., 2012.

[11] HiFly, "Airbus A340-300 Specifications," Lisbon, Tech. Rep., 2012.

[12] European Aviation Safety Agency, "Development of Scheduled Maintenance," Structural Health Monitoring Working Group, Tech. Rep. 1, 2009. [Online]. Available: https://easa.europa.eu/system/files/dfu/IP 105 attachment .pdf

[13] L. Rosales, "Analysing Uncertainty and Delays in Aircraft Heavy Maintenance," Ph.D. dissertation, University of Manchester, 2015.

[14] F. V. Jensen, "Bayesian Networks," *WIREs Comp Stat*, vol. 1, no. November/December, pp. 307–315, 2009.

[15] D. Dinis, A. P. Teixeira, and C. Guedes Soares, "Probabilistic Approach for Characterising the Static Risk of Ships Using Bayesian Networks," *Reliability Engineering and System Safety*, vol. 203, no. May, p. 107073, 2020. [Online]. Available: https://doi.org/10.1016/j.ress.2020.107073

[16] S. Fu, D. Zhang, J. Montewka, X. Yan, and E. Zio, "Towards a Probabilistic Model for Predicting Ship Besetting in Ice in Arctic Waters," *Reliability Engineering and System Safety*, vol. 155, no. 2016, pp. 124–136, 2016. [Online]. Available: http://dx.doi.org/10.1016/j.ress.2016.06.010

[17] J. Bell, "Bayesian Networks," in *Machine Learning: Hands-On for Developers and Technical Professionals*. John Wiley & Sons, Inc., 2015, ch. 4, pp. 69–90.

[18] G. F. Cooper and E. Herskovits, "A Bayesian Method for the Induction of Probabilistic Networks from Data," *Machine Learning*, vol. 9, no. 4, pp. 309–347, 1992.

[19] A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum Likelihood from Incomplete Data via the EM Algorithm," *Journal of the Royal Statistical Society. Series B (Methodological)*, vol. 39, no. 1, pp. 1–38, 1977.

[20] S. L. Lauritzen, "The EM Algorithm for Graphical Association Models With Missing Data," *Computational Statistics and Data Analysis*, vol. 19, no. 2, pp. 191–201, 1995.

[21] M. Druzdzel, "SMILE: Structural Modelling, Inference, and Learning Engine and GeNIe: A Development Environment for Graphical Decision-Theoretic Models," in *Proc. Sixt. Natl. Conf. Artif. Intell*, Pittsburgh, PA 15260, 1999, pp. 342–343.