Predictive Maintenance for Aircraft Systems

Using textual elements as covariates

M.U. Oudkerk

Supervisor: dr.ir. W.J.C Verhagen

TU Delft

Predictive Maintenance for Aircraft Systems

Using textual elements as covariates

by

M.U. Oudkerk

to obtain the degree of Master of Science
at the Delft University of Technology,
to be defended publicly on Friday August 30, 2019 at 02:30 PM.

Student number:	4022815	
Project duration:	September 1, 2018 – August 30, 2019	
Thesis committee:	Prof. dr. R. Curran,	section chair,
	Dr. ir. W. J. C. Verhagen,	supervisor,
	Dr. R. M. Groves,	external committee member

An electronic version of this thesis is available at http://repository.tudelft.nl/.

TU Delft

Abstract

Unplanned maintenance is a costly factor in aircraft operations. Predictive Maintenance aims at reducing the surprise effect of unplanned maintenance and thereby its associated cost. A variety of statistical models are used to estimate the remaining life, as well as sensors to gauge component condition. The application to statistical models of sensory information coming from the pilot, in the form of pilot complaints, appears to be an overlooked option worth investigating. In other words: What is the effect of pilot complaints on the predictability of component removals? This question is answered by determining relevant words in the pilot complaints using a TF-IDF analysis and use the presence of these words as covariate in the well known Proportional Hazards Model. Left truncation and right censoring is applied to limit the time-invariant nature of these covariates. The results in the form of hazard ratios indicate a hazard increase of several orders of magnitude with respect to baseline hazard. These results are put into perspective when compared when compared to the known outcome of the pilot complaints, making their added predictability seem marginal. Another adverse indication is the violation of the proportionality assumption. The magnitude of the hazard ratios do suggest that additional measures in the from of a more in depth natural language processing and the application of time-varying covariates could bring the concept closer to practical application.

Keywords: *predictive maintenance, proportional hazards model, pilot complaint, covariate, hazard ratio, natural language processing, TF-IDF*

Contents

List of Abbreviations

ATA	ATA 100, (Air Transport Association of America)
DF	Document Frequency
FAA	Federal Aviation Administration
FFOP	Failure Free Operating Period
HR	Hazard Ratio
KME	Kaplan-Meier Estimator
MFOP	Maintenance Free Operating Period
NLP	Natural Language Processing
PA	Proportionality Assumption
PDF	Probability Density Function
PFD	Primary Flight Display
PHM	Proportional Hazards Model
PN	Part Number
PRSOV	Pressure Regulating Shut-Off Valve
RASP	Robust Accurate Statistical Parsing
TF	Term Frequency
TF-IDF	Term Frequency Inverse Document Frequency

List of Figures

List of Tables

1 Introduction

Whether an aircraft on the ground costs money or makes money depends on the point of view being that of an maintenance provider or that of an airline. Both parties however benefit greatly from insight into when an aircraft or one of its components will fail. Advanced maintenance programs have reduced the failures of crucial components. Failures can still severely disrupt airline operations and require maintenance providers to have ample capacity available for a timely return of the stricken aircraft to a flying state. The great cost associated with unplanned maintenance begs for methods to make the unplanned more ”plannable”.

In the absence of a crystal ball, engineers have been working for over a century to reduce the ”surprise” of a component failure. The field that studies the remaining lifetime of something or someone is called ”survival analysis”. Scientists in this field have come up with many parametric models such as the Weibull distribution [32] to give insight into remaining lifetime. Where more flexibility was required, non-parametric models such as the Kaplan-Meier Estimator [22] were devised. Some models allow for a multivariate approach to survival analysis, like the semi-parametric Proportional Hazards Model, or Cox model after its inventor Sir David Cox [5]. This model has proven itself over and over in the field of medicine, however, the recent work of Verhagen [31] has shown its relevance in aircraft maintenance. The Proportional Hazards Model in aircraft maintenance often makes use of operational parameters or physical parameters such as engine oil condition [18]. While pilot complaints are readily available due to regulatory requirements, the use of pilot complaint data as covariates in the Proportional Hazards model is yet to be investigated.

The purpose of this research is gauge the usability of the pilot complaints as an external source of data, and thereby use the pilot as a versatile sensor giving information on the condition of the aircraft. In other words:

What is the effect of pilot complaints on the predictability of component removals?

This question is to be answered on many levels. Useful information from the pilot complaints must first be extracted and put in a usable format. The outcome of the investigation should not just be presented in the form of common survival metrics such as hazard ratios. These results must also be judged as to their significance and relevance to the goal of providing enhanced predictability.

To meet these goals, a structured approach is laid out. This project is preceded by a review of the relevant literature, as presented in chapter 2. The data can be seen as the ”raw material” that is required to produce the results. From this raw material, a selection is made to achieve the most reliable results. Chapter 3 describes this process. The strategy used to take the data and use it to find answers to the research questions is presented in chapter 4. This strategy makes use of several models, these are elaborated on in chapter 5. The considerations taken into account while implementing the models described in chapter 5, as well as the way the data is fed into these models is explained in chapter 6. The results that follow from the strategy in chapter 4 are presented in chapter 7. Whilst these results are from the answer to the questions posed earlier, they require a scrutinous discussion with respect to their value and significance. This discussion is followed by recommendation as to how the results could be improved with respect to its goal, providing predictability of the component removals. This discussion and subsequent recommendations come together in chapter 8. Only now can this research be concluded in chapter 9 by providing an answer to the question ”What is the effect of pilot complaints on the predictability of component removals?”.

2

Literature Review

This chapter gives an overview of relevant literature for this research. Firstly, the topic of aircraft maintenance is generally introduced in section 2.1. The tree structure shown in figure 2.1 is followed on the path to preventive maintenance, discussed in section 2.2 and beyond. Section 2.3 gives an overview of survival analysis, the field that tries to determine the remaining lifetime. Sections 2.4 and 2.5 takes this one level further by going more in depth on two important models, being the Kaplan-Meier Estimator and the Proportional Hazards Model. Natural Language Processing, discussed in section 2.6, is required to extract the information from the pilot complaints and make them useful in the models mentioned earlier. The novelty of this combination is further stressed in section 2.7

2.1. Aircraft Maintenance

Aircraft operations always go hand in hand with failures of its components. A component failure can best be described by a part not being able to cope with the stresses it experiences in operation, something Findlay [13] argues to be predominantly caused by fatigue. Maintenance actions are performed in order to bring aircraft back into a state of functioning after a failure (reactive maintenance), or to prevent them from failing (proactive maintenance). These two important branches of maintenance are visualized in figure 2.1. The reactive branch of this tree corresponds to the use of the full life of a component, this advantage is offset by its unpredictable nature. Proactive maintenance returns the initiative to the operator at the expense of not utilizing the full life of components and taking an aircraft out of operation for servicing. These two types of maintenance define the failure free operating period (FFOP) and the maintenance free operating period (MFOP) respectively. These two are future metrics in aircraft reliability and maintenance, as stated by Dinesh Kumar [10].

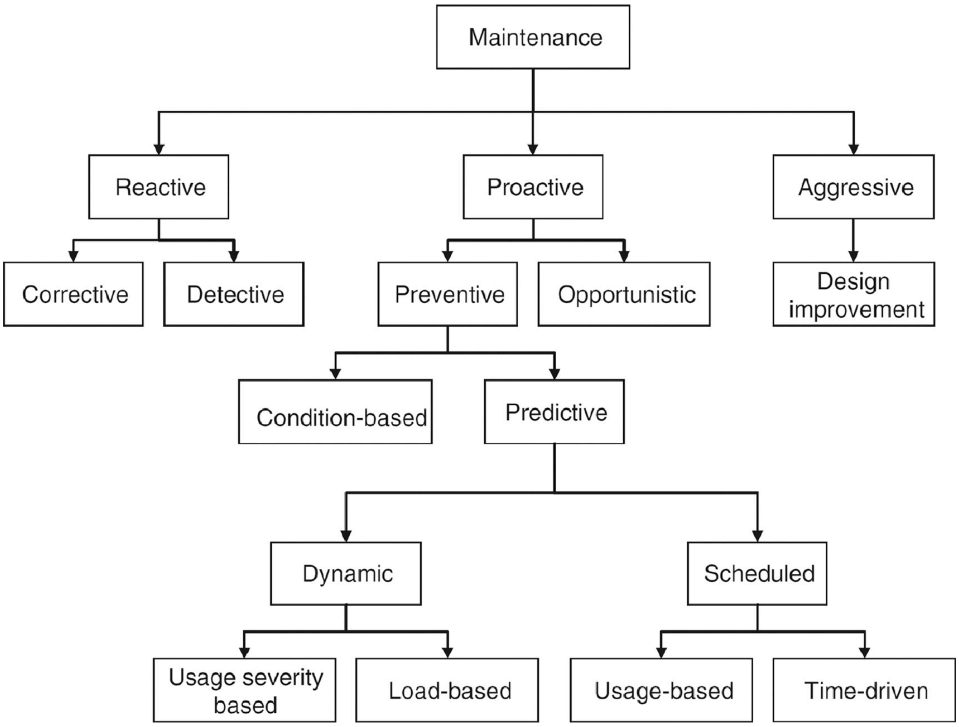


Figure 2.1: Tree structure of maintenance types as identified by Tinga [30].

2.2. Preventive Maintenance

The tree structure shown in figure 2.1 illustrated that the only option for an airline to eliminate unexpected maintenance events is to turn to proactive maintenance, as aggressive maintenance fall within the power of the aircraft manufacturer. This proactive maintenance can either be opportunistic, or occur at dedicated scheduling. The dedicated scheduling is performed either based on a measurable condition, condition based maintenance, or on models that predict the remaining lifetime. Statistical models are used to give insight into the remaining life of a component. Condition based maintenance hinges on the availability of data. Jardine [19] discusses a wide range of possible sensors to collect data, all being physical sensors. Instances of the pilot being used as a sensor in condition based maintenance are not found. More on this subject in section 2.6. Jardine [19] identifies three main steps in a condition based maintenance process:

1. Data acquisition
2. Data processing
3. Maintenance decision making

Without the availability, one can turn to the statistical analysis of the lifetime of components. The specific area of statistics dedicated to the analysis of remaining lifetime is called survival analysis and is further discussed in section 2.3.

2.3. Survival Analysis

The field of survival analysis, or time to event analysis, studies the time between a birth event and a death event. Despite its applicability being most obvious in medicine, other applications range from physics to economics [6]. When survival analysis involves the failure of components, it is often named reliability analysis [8].

2.3.1. Distributions

The lifetime of a component is often assumed to be described by random variables. The distribution is often dependent on the failure type. A distribution suited to a purely random failure process, as stated by Epstein [12] is the exponential distribution:

$$f(x,\theta) = \frac{1}{\theta} exp(-\frac{x}{\theta}) \tag{2.1}$$

where:

$f(x,\theta)$ = probability density function (pdf)
θ = rate parameter

Many options exist to model wear-out failures, or failure with increase failure rates, but one of them is the well known normal distribution[1]. Figure 2.2 shows increase failure rates.

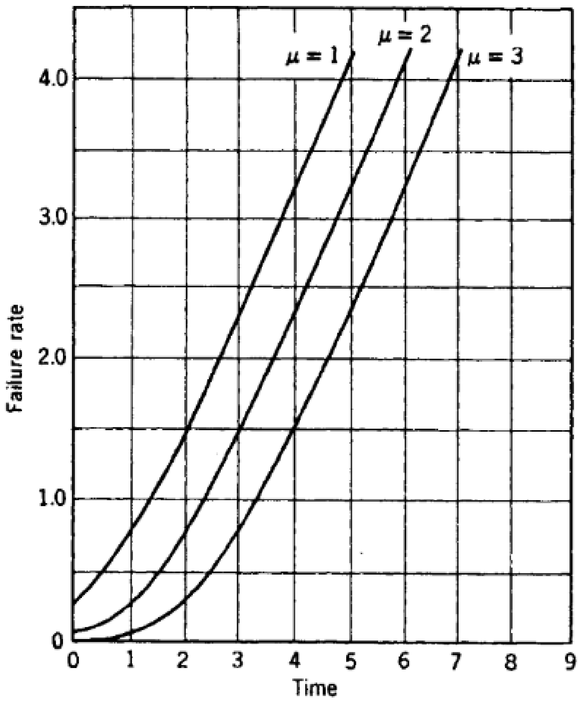


Figure 2.2: Failure rate curves of the normal distribution for σ = 1. [1]

A more flexible type of distribution that can cover both constant, wear-out and burn-in failures is the Weibull distribution, named after its inventor, mathematician Waloddi Weibull [32]. The Weibull distribution is defined as follows:

$$f(x,\alpha,\gamma) = \frac{\alpha}{\eta}\left(\frac{x-\gamma}{\eta}\right)^{\alpha-1} exp\left(-\left(\frac{x-\gamma}{\eta}\right)\right) \tag{2.2}$$

where:

$f(x,\beta,\gamma)$ = probability density function (pdf)
α = shape parameter
η = scale parameter
γ = location parameter

Constant hazard, as well as decreasing and increasing hazard can be modelled by changing the shape parameter, as can bee seen in figure 2.3.

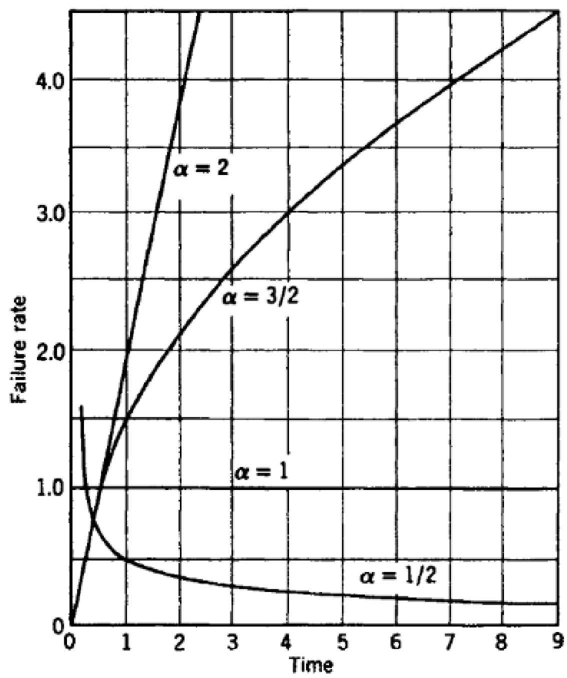


Figure 2.3: Failure rate curves of the Weibull distribution for $\eta = 1$. [1]

2.3.2. Censoring and Truncation

The use of the models mentioned above becomes less straightforward under the existence of incomplete data. Incomplete data can be caused by censoring and truncation. These two phenomenons are similar. The difference becomes more clear when they are properly defined. Cleves distinguishes the following types of censoring and truncation [4]:

- **Censoring:** Failure event does not occur during the observation time.
  - **Right Censoring:** Observation of subject stops without failure having occurred.
  - **Left Censoring:** Failure of subject occurs without observation having starting
  - **Interval Censoring:** Failure of subject has occurred withing a time interval without the subject being under observation.
- **Truncation:** Time in which the subject was not observed, but is know not to have failed in hindsight.
  - **Left Truncation:** Time at which the subject became at risk did not coincide with the start of observation, also know as late entry.
  - **Right Truncation:** Failure does not occur in observation period but is know to have occurred eventually.
  - **Interval Censoring:** Subject is not observed during a certain interval during the observation period but returns to be observed.

These types of censoring and truncation are visualized in figure 2.4, where the green cir circle represents the birth moment and the red symbol represents the death moment. The grey areas represent gaps in the observation. Each of these types of incomplete data requires additional methods to prevent a selection bias. Klein [23] demonstrates the existence of many options to handle censored data.

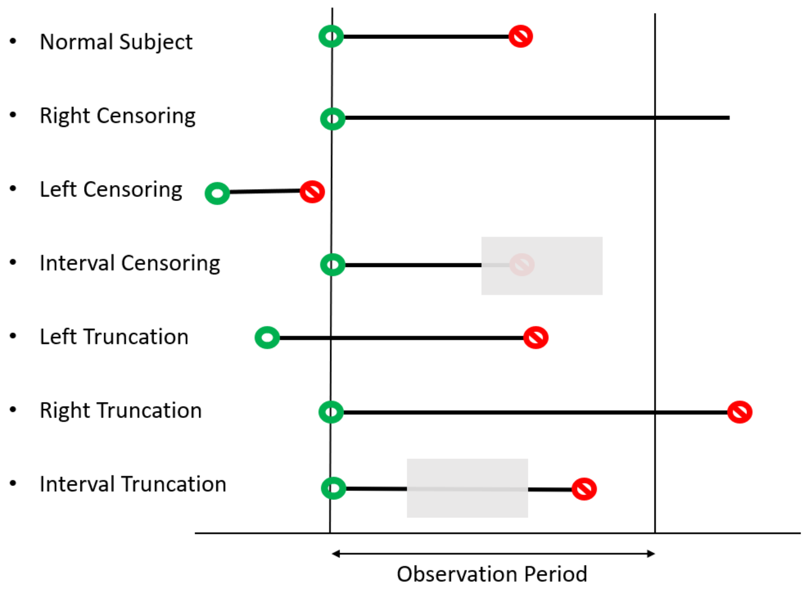


Figure 2.4: Visualization of different types of incomplete data.

2.4. Kaplan-Meier Estimator

In the pursuit of methods to handle incomplete data, the Kaplan-Meier Estimator was devised by Edward Kaplan and Paul Meier[22]. The Kaplan-Meier Estimator, frequently referred to as ”product limit estimator”, is a method to estimate the survival function bases on (censored) lifetime data. As opposed to the distributions mentioned in section 2.3, the Kaplan-Meier Estimator is non-parametric, meaning it does not adhere to a specific shape of distribution and therefore enjoys more flexibility. The Kaplan Meier Estimator can be enhanced with a confidence interval. A well know confidence interval is given by the Greenwood's confidence interval [14]. The Greenwood's confidence interval is symmetrical, whilst an asymmetrical confidence interval that does not exceed 1 nor 0 is desirable. The confidence interval that accomplishes this is given by the ”exponential Greenwood formula”[17] [21].

2.5. Proportional Hazards Model

Despite the flexibility of the Kaplan-Meier Estimator discussed in section 2.4, it is univariate. Other survival models exist that do possess that same limitation. One of the most well know regression models in the field of survival analysis is the proportional hazards model (PHM), or Cox model after its inventor, Sir David Cox [5]. The Cox model tries to model survival time while taking into account the effect of one or more explanatory variables, or covariates. The Cox Proportional Hazard Model is a semi-parametric model since its baseline hazard is free to take shape while the scaling of the baseline hazard function is parametrically defined. The popularity of the Proportional Hazards Model also gave rise to many adaptions, extensions or enhancements. The baseline hazard can be described by a parametric distribution. A Weibull distribution, as discussed in section 2.3, is often applied as baseline hazards, such as in the work of Love and Guo [26]. Therneau and Grambsch [29] describe extensions the Proportional Hazards Model to include time varying covariates. Despite its usage primarily in medicine, many instances of its use in aircraft reliability have been recorded through time. In 1987 for example, where engine parameters where used in the assessment of engine failures [18]. Modern analysis on the application to aircraft component reliability performed by Verhagen and De Boer [31] show its retained relevance.

2.5.1. Explanatory Variables

Each analysis requires a different choice of explanatory variables. Cox illustrates this with the following examples [7]:

1. A study into the severity of a respiratory disease among workers, where explanatory variable include age, working conditions, among others.

2. A study into the time to death among patients with a progressive and fatal disease, where explanatory variables are made up of treatment variables.

The explanatory variable ”age” might not be as relevant if the onset of the disease mentioned in the second example follows a similar pattern in both young and old patients. An important note made by Cox in regard to the example mentioned above that the time to death in the second example is measured from the moment of diagnosis instead of the literal ”birth” of the patient [7].

2.6. Natural Language Processing

Data in science is often comprised of numbers. Numerical data is something a computer can deal with easily. Some data comes as linguistic expressions by humans and require processing to be handled by computers. The case of using pilot complaints as covariates is not different. To be able to use text as a covariate, this text must first be converted to a numerical value somehow. This processing is called natural language processing. (NLP) Manning presents the foundations of NLP covering words, grammar and many applications [27]. Natural Language Processing is an essential step in facilitation the use of the pilot complaints as sensory output of the pilot, the latter functioning as a sensor with respect to the health of the aircraft components.

2.6.1. TF-IDF

A very important statistic is the term frequency-inverse document frequency (TF-IDF) as describes by Sparck Jones [28]. TF-IDF is a measure of relevance. It counts the frequency in a certain group of texts while correcting for the frequency across the entire spectrum of texts. When encountering smaller data sets, it might be required to apply smoothing as mentioned by Liu [25].

2.6.2. Synonyms

The pilot complaints contain many synonyms. Terms such as ”o2-bottle” ”oxygen-bottle” and ”oxygen bottle” mean the same thing. Methods exist to be able to identify, or ”acquire” synonyms. Hagiwara states that these methods often find their basis in the distributional hypothesis [15]. This distributional hypothesis, devised by Harris [16], means that textual elements with similar context possess semantic similarity. It has been shown by Kumari that a clustering of synonyms has positive effects on the efficacy of a TF-IDF analysis [24].

2.6.3. NLP Parsing

Further analysis going into the structure itself is called parsing. Identifying the structure of the text can be used to identify the subject and the descriptive elements possibly describing an adverse condition of a component. Figure 2.5 shows the dependency relations that can be identified. Many algorithms exist to parse text such as the RASP (robust accurate statistical parsing) system devised by Briscoe [3].

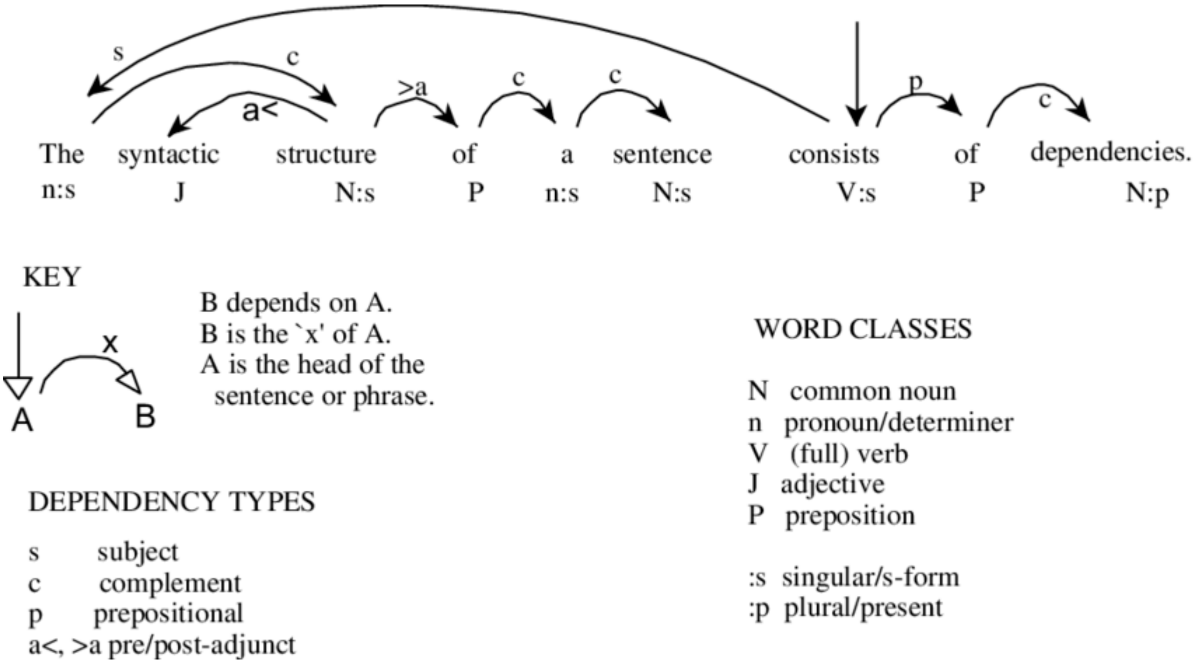


Figure 2.5: Dependency relations in Word Grammar [11].

2.7. Novelty

In a world where technologies like structural health monitoring, using a multitude of sensors, increasingly find their way into new aircraft designs, one could argue that the most comprehensive sensor of them all, the pilot, has been overlooked. Research such as that of Jardine [19] do not mention the pilot as a source of data. This versatile sensor produces information in the form of natural language. Using Natural Language Processing discusses in section 2.6, one can turn the textual information into numerical values to be used as covariates in survival analysis models such as the Kaplan-Meier Estimator and the Cox Proportional Hazards Model mentioned in sections 2.4 and 2.5 respectively. Using the information from this human sensor in such models might provide previously untapped insight into the predictability of component removals, positioning itself between predictive maintenance and condition based maintenance in the tree structure in figure 2.1.

3
Data

This chapter describes the data used in this research. The source of this data is a large European maintenance service provider that will remain unnamed for the purpose of confidentiality. This maintenance service provider has performed maintenance activities across multiple airlines, aircraft manufacturers and aircraft types. Records are kept of all parameters relevant to the technical state of the aircraft. Section 3.1 gives more information on the relevant tables within the data set by showing their relevant content and interrelations. Section 3.2 gives qualitative insight into the data, where sections 3.5 and 3.6 present the final size of the data after cleaning and sampling. These cleaning and sampling steps are described in sections 3.3 and 3.4 respectively.

3.1. Overview

This section gives an overview of the relevant tables within the data set. Sections 3.1.2 and 3.1.3 go into further detail on component removals and pilot complaints respectively, as they form the main source of information for further analysis. The relations between these two tables are depicted in figure 3.1. This figure shows that each data entry has a unique identifier, or ”primary key”, being the ”CompId” for a component removal and a ”PilotId” for pilot complaints. Both tables have the ”AircraftSerialNumber” as foreign key, being the entry used to link the data entry to a data entry in a foreign table. The most important information the data is the ”Date”, as insight into the date of a component removal is the definition of predictability. Both the date of the component removals itself as well external information from the pilot complaint could provide increased predictability. Section 3.1.1 is dedicated to the ATA-chapters as they require further explanation.

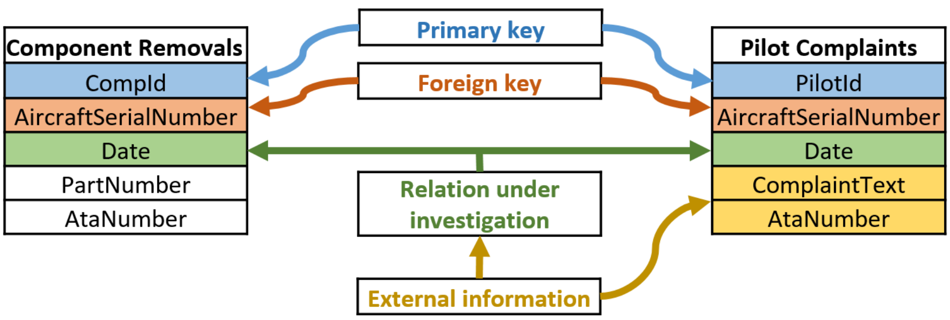


Figure 3.1: An overview of the two most important tables within the data set and their interrelations.

3.1.1. ATA Chapters

All systems in the aircraft are categorized by ”ATA 100” chapter. ATA 100 is a referencing standard originally created by the Air Transport Association of America (ATA), modified by the Federal Aviation Administration (FAA) [2]. Each chapter is identified by its two digit number, whilst sub chapters have multiple digits, spanning up to 6 digits. The categorization becomes more specific with each added digit. Table 3.1 shows a small part of ATA chapter 35, corresponding to the oxygen systems on board the aircraft. The ATA chapters mentioned in the component removal data entries often include the full range of applicable digits as that information is readily available. The ATA chapters in the pilot complaints are often truncated and posses only two digits.

Table 3.1: Small excerpt of the table containing the ATA chapter information belonging to the chapter for oxygen.

AtaNumber	AtaDescription
3521	GENERATION AND DISTRIBUTION
352100	GENERATION AND DISTRIBUTION
352101	OXYGEN DROP-OUT PANEL
352102	DROP-OUT PANEL LATCH
352103	OXYGEN GENERATOR

3.1.2. Component Removals

The component removals table is one of the two most important tables in the data set, together with the pilot complaints in section 3.1.3. An example of the relevant columns within this table is shown in table 3.2.

Table 3.2: Small excerpt from the component removals table showing only relevant columns.

CompId	PartNumber	Date	ExtOrgCode	AtaNumber	AircraftSerialNumber
426712	A42702-431	03/09/2011 00:00	1-HN0	334001	11567
426713	72015727	03/09/2011 00:00	1-HN0	253101	11567
426714	D07354-435	03/09/2011 00:00	1-HN0	324101	11566
426715	D07354-435	03/09/2011 00:00	1-HN0	324101	11566
426716	D07611-417	03/09/2011 00:00	1-HN0	324103	11578

3.1.3. Pilot Complaints

The pilot complaints provide information used to enhance the predictability of the component removals mentioned in section 3.1.2. Table 3.3 provides an example of this data. Note that pilot complaints often contain only the first two digits of the corresponding ATA chapter.

Table 3.3: Small example of relevant columns of pilot complaint data. The complaint text is withheld from the table but shown in figures 3.2 through 3.6.

PilotId	AtaNumber	ExtOrgCode	Date	AircraftSerialNumber	Complaint Text
261498	35	1-HN0	02/09/2011 00:00	11538	see figure 3.2
261499	35	1-HN0	02/09/2011 00:00	11574	see figure 3.3
261500	38	1-HN0	02/09/2011 00:00	11561	see figure 3.4
261501	49	1-HN0	02/09/2011 00:00	11536	see figure 3.5
261502	52	1-HN0	02/09/2011 00:00	11541	see figure 3.6

CREW OX BOTTLE @ 1500PSI
ACTION: CREW OX BOTTLE REPALCED IOAW AMM 35-11-01-400-814A TESTED STAIS SN IN ALT372-6271 SN OUT ALT372-3243

Figure 3.2: Complaint text corresponding to PilotID 261498 in table 3.3.

OXYGEN BOTTLE PRESS 1350 PSI
ACTION: REPLACED OXYGEN BOTTLE IAW AMM 35-11-01-400-814-A JUN 01/11 PN:176225 SN IN: ALT372-5229 SN OUT: ALT372-6272

Figure 3.3: Complaint text corresponding to PilotID 261499 in table 3.3.

WATER FILL CAP OFF CHAIN
ACTION: REINSTALLED CHAIN ON CAP, ALL SATIS

Figure 3.4: Complaint text corresponding to PilotID 261500 in table 3.3.

REF WO 2621308, LOW BLEED PRESS OF APU. REPLACED VGD ACTUATOR. PLS C/O OPERATIONAL TEST IAW AMM 49-00-00-710-815-A
ACTION: OP TEST C/O SATIS IAW AMM 49-53-00-710-815A, JUN01/07

Figure 3.5: Complaint text corresponding to PilotID 261501 in table 3.3.

WHEN OPENING THE SERVICE/EMERGENCY DOOR, THE HOOK DOENST ALWAYS STAY IN THE LOCK
ACTION: PARKING HOOK HOUSING ADJUSTED I.A.W. AMM 52-41-02-820-825-A REV JUN01/07

Figure 3.6: Complaint text corresponding to PilotID 261502 in table 3.3.

Please note that pilot complaints shown in figures 3.2 through 3.6 have the performed maintenance action retrospectively added to the pilot complaint. A very useful feature as it provides an upper bound to the predictability of a removal. Additionally, this feature is used in the TF-IDF analysis described in section 5.1.

3.2. Quality

It is assumed that all entries in the data-set represent reality and were not corrupted. Some factors negatively influence the quality of the data. The biggest negative influence on data quality are missing essential elements of some entries. Most of these missing entries are missing aircraft registrations, making it impossible to relate them. Some of the pilot complaints have their complaint text in another language than English. The sampling based on airline mentioned in section 3.4 resolves this as it excludes the airline using another language in their complaints. Typing mistakes are often observed within the pilot complaints. Figure 3.2 shows the word ”REPLACED” being misspelled as ”REPALCED” The effect of misspellings is deemed minor, due to its assumed random nature. Most pilot complaints have their corresponding action retrospectively added to the complaint text as can be seen in figures 3.2 through 3.6. This addition is a contamination of the complaint text. It is however possible to filter this action text out by disregarding all text after the word ”ACTION”. It must be noted however that some information is retrospectively added to the complaint text with the use of the ”ACTION” marker. An example of this can be seen in figure 3.5. The latter contamination is of detrimental effect on the quality of the data.

3.3. Cleaning

Data cleaning focuses on the removal of data entries that have no value to this research. It is essential to this research to be able to link the information on component removals with the information in the pilot complaints. This is done based on the aircraft serial number as shown in figure 3.1. The majority of the historical data entries in both the component removal table and the pilot complaint table have an absent aircraft serial number. This means these entries can not be linked to the same aircraft and are therefore discarded. Some of the entries in the pilot complaint are not of the English language. These entries should not be considered. Non-English entries do not require any specific action since the respective airline is not part of the sample described in section 3.4.

3.4. Sampling

This section describes the sampling that was applied to the whole data set. The meaningfulness of the pilot complaints as explanatory variables is reduced when influenced by factors that are not taken into consideration. The various subsections within this section elaborated further on what the sample is based on and why this decision is made.

Year

The data set spans a time period from as early as august 1987 up to the same month of 2016, not taking into account some date in the future that are erroneous. Quality of the data is not consistent over the course of these years, as mentioned in section 3.2. Some parameters might not have been recorded initially, or they might not have been digitized. Sparsity of data is another factor to be considered when sampling a certain time span within the data. Figure 3.7 gives a graphical representation of the activity of the fleet as recorded in the data.

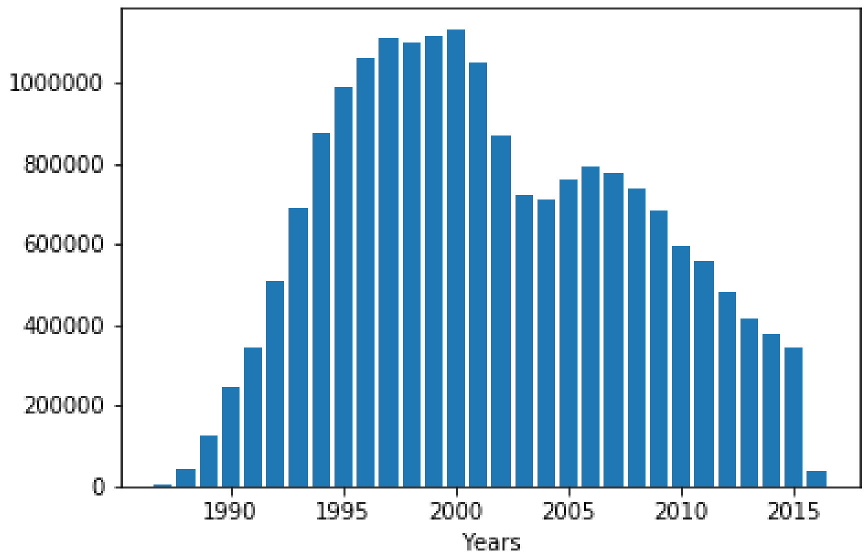


Figure 3.7: Flight hours throughout the year as performed by the aircraft in the data.

A noticeable peak between 1995 and 2000 can be observed. The gradual decline seen after 2006 corresponds with the phase-out of the aircraft in the data. However, this graph does not point to many sparse years with the exception of the years running up to 1990 and the year 2016. The latter still being ongoing in this data set. A further look into the availability of data in relevant tables, being component removals and pilot complaints, is required. Figure 3.8 shows the incidence of component removals normalized by flight hour.

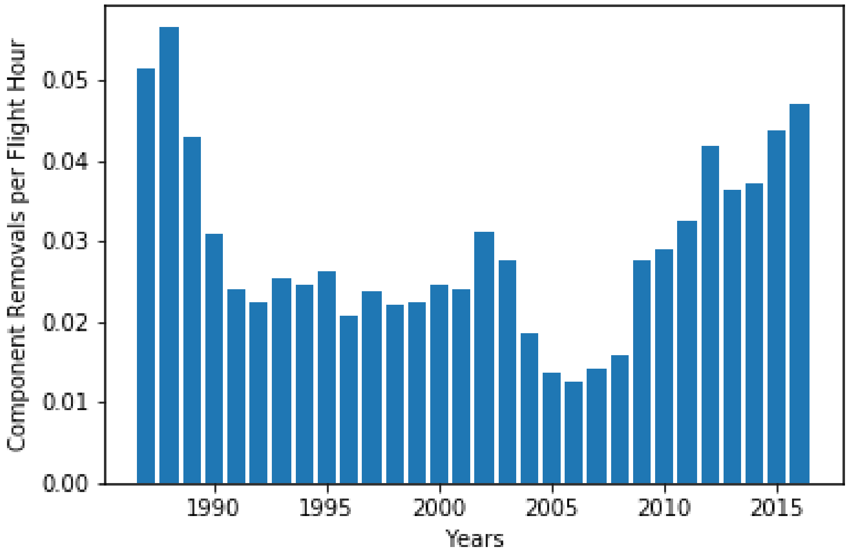


Figure 3.8: Amount of component removals per flight hour on a yearly basis.

Despite the high initial number of removals before 1990 and the small dip between 2005 and 2009, this graph is inconclusive. The same assessment performed on pilot complaints shows more conclusive evidence for found sampling range. Figure 3.9 shows a significantly high pilot complaint rate from 2010 onward. This information is the basis for the sampling of data between 2010 and 2015. The reason for the suppressed quantity of pilot complaints per flight hour are unknown, although it is likely to be related to the fact that pilot complaints before did not contain any complaint text.

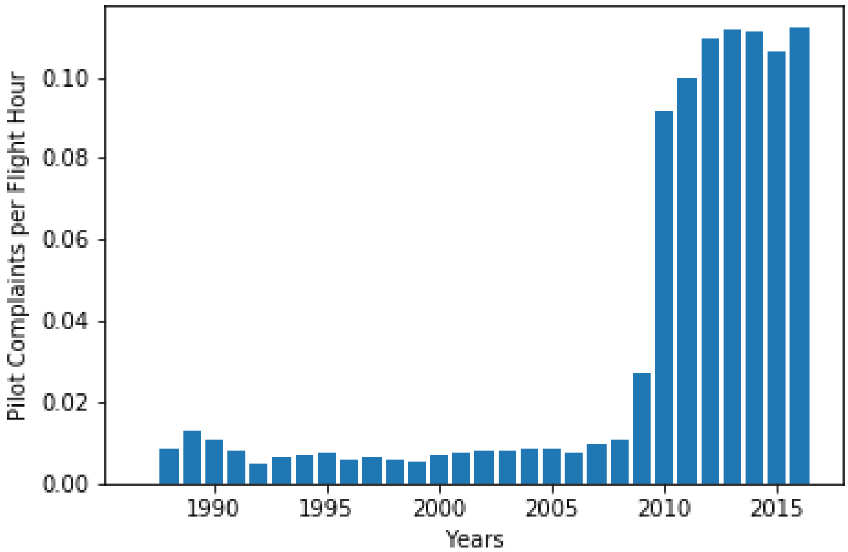


Figure 3.9: Amount of pilot complaints per flight hour on a yearly basis.

Airline

All airlines operating a certain aircraft type are bound by restrictions with respect to the maintenance, as outlined by their maintenance program. Within these restrictions, differences might still occur. One airline might chose to replace a component at the end of its life, another might do so preventively or opportunistically. This is the basis for the sampling of the data based on airline. The airline with the largest share of the data for pilot complaints and component removals is selected. For both tables, this airline is ”1-HN0”. The size of the data corresponding to airline ”1-HN0”, with respect to the total size of component removal and pilot complaint data is shown in figures 3.10 and 3.11 respectively.

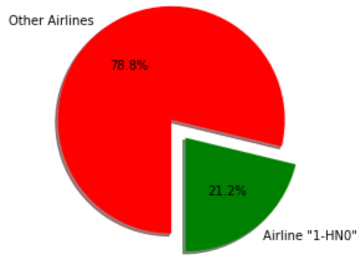


Figure 3.10: Sampling of the component removal data based on airline.

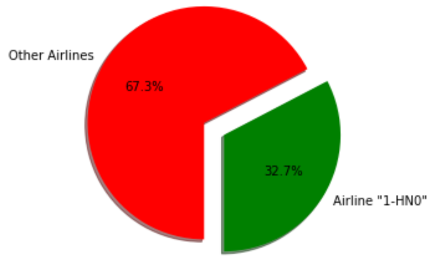


Figure 3.11: Sampling of the pilot complaint data based on airline.

Aircraft

Airline ”1-HN0” operated three types of aircraft in the time sample mentioned in section 3.4. All three aircraft types are made by the same manufacturer and two of them are sub types of the same aircraft type. To eliminate as many factors of influence on component survivability, only one aircraft type is selected to be part of the sample. In case of airline ”1-HN0” this is aircraft type 3. Aircraft type 3 makes up most of the data for both component removals and pilot complaints as can be seen in figures 3.12 and 3.13 respectively.

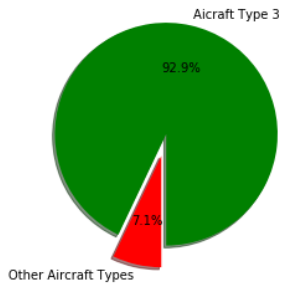


Figure 3.12: Sampling of the component removal data based on aircraft type.

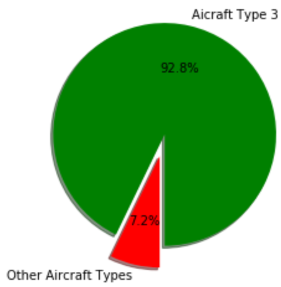


Figure 3.13: Sampling of the pilot complaint data based on aircraft type.

Component

An aircraft, being the complicated machine that it is, has a quite extensive list of components. Not all components are suitable for the type of statistical analysis as described in this report. The sample of components to be used in this research is based on component removal frequency. The five most frequently removed components are selected. Components with frequent removals provide more information as opposed to infrequently remove components. In the absence of information on component cost and failure impact, one could argue that insight in the most frequently removed components provides the largest value to the maintenance industry. Table 3.4 presents the top five most frequently removed components.

Table 3.4: Sampled components, their description and the amount of removals.

Name	Description	Removals
Component 0	Oxygen Bottle	2516
Component 1	Flow Control Valve	207
Component 2	Display Unit	196
Component 3	Pressure Regulating Shut-Off Valve	194
Component 4	Landing Light	176

The sampled components make up approximately a third of the total component removals as can be seen in figure 3.14. The frequency of the sampled components with respect to each other is presented graphically in figure 3.15. Note that the oxygen bottle makes up at least three quarters of the removal data.

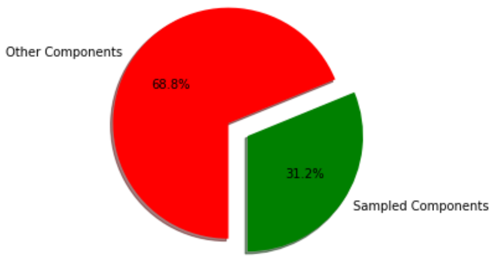


Figure 3.14: Share of component removals belonging to the selected five most frequently removed components.

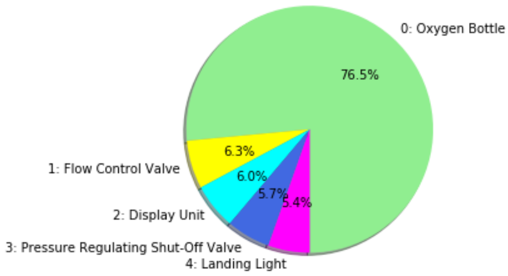


Figure 3.15: Frequency distribution of components removals withing the selection of five most frequently removed components.

3.5. Size

The size of the data is shown in tables 3.5 and 3.6. The relative size represents the share of the data with respect to the raw data size, while the relative size represents the data size with respect to the previous step. The final sizes for component removals and pilot complaints amount to 21,761 and 89,968. It must me noted that the size of the component removals is before sampling based on individual components.

Table 3.5: Size of the component removal data going through various filtering and sampling steps.

	Component Removals	**Relative**	**Absolute**
Raw	476,262	100%	100%
Filter Dates	132,351	28%	28%
Missing Registrations	102,451	77%	22%
Sample Airline	21,761	21%	5%
Sample Type	20,222	93%	4%
Sample Components	3,101	15%	0.65%

Table 3.6: Size of the pilot complaint data going through various filtering and sampling steps.

	Pilot Complaints	**Relative**	**Absolute**
Raw	428,737	100%	100%
Filter Dates	299,212	70%	70%
Missing Registrations	295,746	99%	69%
Sample Airline	96,951	33%	23%
Sample Type	89,968	93%	21%

3.6. Summary

This section summarizes the sampling steps and presents the final information on the data to be used in the analysis described in chapter 4.

- **Component Removals**
  - **Years:** 2010-2015
  - **Airlines:** Only ”1-HN0”
  - **Aircraft types:** Only ”type 3”
  - **Components:** The five most frequently removed components.
  - **Size:** 3,101 entries.
- **Pilot Complaints**
  - **Years:** 2010-2015.
  - **Airlines:** Only ”1-HN0”.
  - **Aircraft types:** Only ”type 3”.
  - **Size:** 89,968 entries.

4 Methodology

This chapter is dedicated to the description of the methodology used to find the answers to the main research question below:
What is the effect of pilot complaints on the predictability of component removals?
This main question is answered by answering the following auxiliary questions and their corresponding methodological steps:

- **What are the hazard ratios for endogenous covariates?:** Fit the Proportional Hazards Model using the covariates that originate from the component removal data itself, as mentioned in section 4.1
- **What are the most relevant words for each part?:** Run the TF-IDF analysis on the pilot complaints that have the part number mentioned in the action.
- **What are the hazard ratios for relevant exogenous covariates?:** Fit the Proportional Hazards model with the addition of each of the relevant exogenous covariates as mentioned in section 4.1, separately. Fitting the model for exogenous covariates separately is important as they are not assumed to be independent.
- **What are the hazard ratios for irrelevant exogenous covariates?:** Fit a Proportional Hazards Model with relevant words for other parts in order to confirm that they do not have a significant effect.
- **How do the results for the Kaplan-Meier Estimator and the Proportional Hazards Model compare?:** Fit a Proportional Hazards Model and a Kaplan-Meier Estimator and have the two included in a survival curve plot.
- **How do the predictive results compare to knowing the outcome?:** Use the Kaplan-Meier Estimator for mentioned part number data and compare it to the Proportional Hazard Model fit for the highest scoring word in a survival curve plot.
- **What is the sensitivity with respect to observation time?:** Show trends in hazard ratio and p-value for a Proportional Hazards Model fit for ATA chapter while varying the observation time.

The auxiliary questions states above make uses of several models, two being survival models, the Kaplan-Meier Estimator and the Proportional Hazards Model, the other being a natural language model, the TF-IDF model. The descriptions of these models can be found in chapter 5. The auxiliary question also mention endogenous and exogenous covariates as input to the survival models. These covariates are described in section 4.1. The choice made with regard to the implementation of the models and the covariates are discussed in chapter 6. A graphical summary of the methodology can be found in the methodology diagram in figure 4.1.

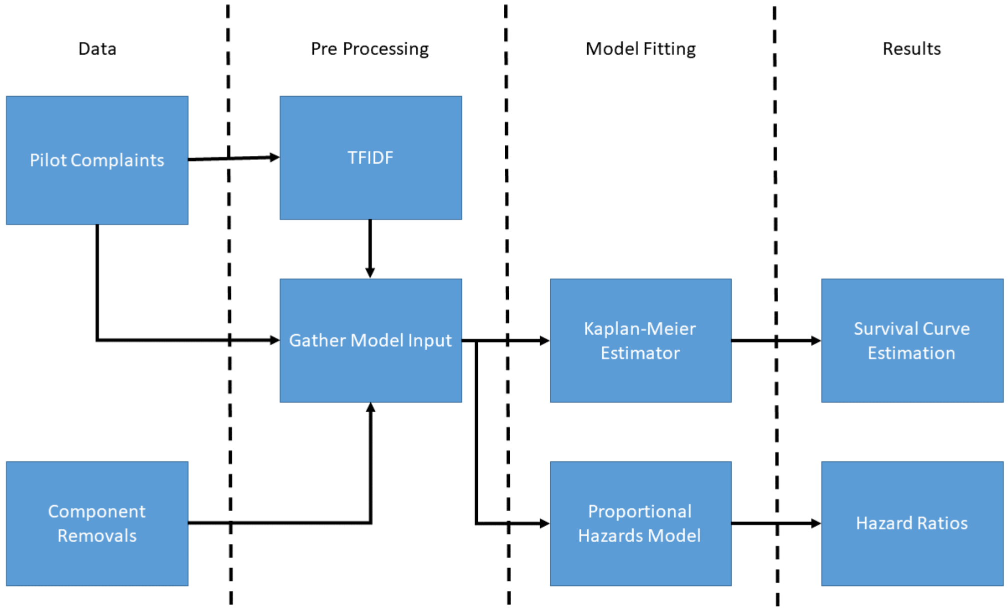


Figure 4.1: Methodology diagram.

4.1. Covariates

For each pilot complaint, the values of the following covariates are determined. The covariates consists of endogenous covariates and exogenous covariates. The former originating from the component removal data itself (Year, Fresh and Summer), while the latter originates externally from the pilot complaint data. The following list give an overview of the covariates used in this research:

- **Year:** This covariate corresponds to the year of the complaint. It is used to gauge the effects time has on the hazard.
- **Fresh:** This covariate has a value of 1 when the previous component removal was observed to be within two months of the pilot complaint under consideration. This covariate is used to judge whether a fresh installation has an effect on the hazard.
- **Summer:** This covariate has a value of 1 when the pilot complaint falls within the airline summer schedule. It is used to judge its effect on the hazard.
- **PN:** This covariate has a value of 1 if the part number is mentioned in the complaint text. This is the main use of the included actions as described in section 3.1.3. This is used to analyze the lifetime patterns when it is known in advance that a removal will occurs due to the pilot complaint in question.
- **ATA:** This covariate has a value of 1 when the first two digits of the mentioned ATA chapter, as described in section 3.1.1, match the first two digits of the component ATA chapter. This covariate is used to determine the effect that mentioning the specific subsystem group has on the hazard.

- **word:*:** This covariate has a value of 1 if the word represented by the asterisk is mentioned in the pilot complaint. This is used to measure the effect of certain words on the hazard. These words can either be words that are relevant to the component under investigation, or be relevant to any of the other components. In the latter case, the word is assumed to be irrelevant.

5
Models

This chapter described the models used to answer the questions in chapter 4, starting with the natural language model for word relevance, TF-IDF, in section 5.1. This is followed by the two survival models. Firstly the Kaplan-Meier Estimator, discussed in section 5.2, and secondly the Proportional Hazards Model, discussed in section 5.3.

5.1. Term Frequency Inverse Document Frequency (TF-IDF)

A selection of words is to be made due to the impracticality of using every word in the entire corpus of pilot complaints. This selection must be based on the predictive value of each word that is to be used as a covariate of the proportional hazards model described in section 5.3. The frequency of terms in the pilot complaints leading up to a removal or a certain part is not good metric on its own, since the same terms might also be of frequent occurrence in all other documents. To account for the latter, "Term Frequency-Inverse Document Frequency" (TF-IDF)is used [20]. TF-IDF is used in this research to score words based on how frequently they occur in the pilot complaints leading up to a removal, while correcting for its frequency in the entire corpus of pilot complaints. Whether or not a pilot complaint leads op to a removal is determined on having the part number mentioned in the retrospectively added action, as mentioned in section 3.1.3. The scoring scheme, as shown uses Laplace smoothing [25]. The scoring scheme is shown in equation 5.1:

$$a_{ij} = log(tf_{ij} + 1) * log\left(\frac{N + 1}{n_j}\right) \tag{5.1}$$

where:

a_{ij} = score of term j in document i
tf_{ij} = term frequency of term j in document i
N = total number of documents in corpus
n_j = number of documents that term j appears in [1]

Note that document i is the collection of documents that the relevant part number mentioned in the action, as discussed in section 3.1.3.

5.2. Kaplan-Meier Estimator (KME)

The Kaplan-Meier Estimator is selected in order to estimate the survival curve of the component. Kaplan and Meier have devised this Estimator as a non-parametric solution to survival data with censoring [22]. The Kaplan-Meier Estimator aims at making an estimation of the survival curve given by:

$$S(t) = P(T > t) \tag{5.2}$$

[1] Used j for term subscript. Liu [25] switches between i and j.

where S(t) gives to probability of a lifetime T longer than t. The definition of the Kaplan-Meier Estimator (KME) is as follows:

$$\hat{S}(t) = \prod_{i:\ t_i \leq t} \left(1 - \frac{d_i}{n_i}\right) \tag{5.3}$$

where:

$\hat{S}(t)$ = estimation of survival curve
t_i = time at which at least one removal occurs
d_i = number of component removals
n_i = number of components that have not yet been removed

The mathematical definition mentioned in equation 5.3 takes censored specimens into account but lacks information with regards to precision. A confidence interval is required to enhance the Kaplan-Meier Estimator with precision information. An example the Kaplan-Meier Estimator can be seen in figure 5.1. The ”Exponential Greenwood” confidence interval is added to the Kaplan-Meier Estimator to aid in its correct interpretation. The exponential Greenwood interval can take asymmetric values and is bounded by 0 and 1. The confidence interval is defined by:

$$exp(-exp(c_+(t))) < S(t) < exp(-exp(c_-(t))) \tag{5.4a}$$

with:

$$c_{\pm}(t) = log(-log\ \hat{S}(t)) \pm z_{a/2}\sqrt{\hat{V}} \tag{5.4b}$$

with:

$$\hat{V} = \frac{1}{(log\ \hat{S}(t))^2} \sum_{t_i \leq t} \frac{d_i}{n_i(n_i - d_i)} \tag{5.4c}$$

where:

$\hat{S}(t)$ = estimation of survival curve
z_α = α-th quantile of the normal distribution
t_i = time at which at least one removal occurs
d_i = number of component removals
n_i = number of components that have not yet been removed

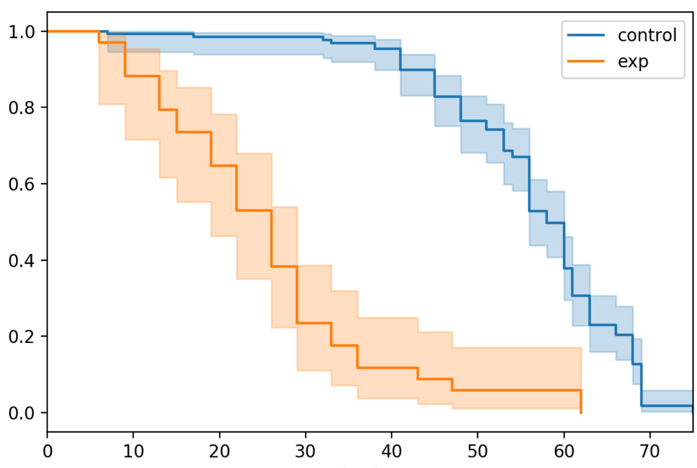


Figure 5.1: Generic example of an application of the Kaplan-Meier Estimator as described in section 5.2, showing the estimates for the survival of an exposed group with respect to a control group. [9]

5.3. Proportional Hazards Model (PHM)

This section presents the Proportional Hazards Model, or Cox Model after its inventor [5]. The Proportional Hazards model is widely used in the field of survival analysis. The reason for the application of the Proportional Hazards Model in this research is threefold:

- **Semi-parametric:** The baseline hazard function can take any shape, as opposed to parametric models such as the Weibull distribution [32].
- **Multivariate:** Multiple explanatory variables, or covariates are considered in the Proportional Hazards Model, allowing for better isolation of the specific effect of one of these covariates.
- **Censoring:** The Proportional Hazard Model takes censoring into account. The way this feature is utilized is discussed in section 6.2.

5.3.1. Definition

The Proportional Hazards Model (PHM) is defined by the following equation:

$$h(t|x) = b_0(t)\ exp(\sum_{i=1}^{n} b_i(x_i - \bar{x}_i)) \tag{5.5}$$

where:

$h(t|x)$ = hazard
$b_0(t)$ = baseline hazard
b_i = coefficients
x_i = covariates
$\overline{x}_i$ = lowest values of covariates

The Proportional Hazards Model is called a semi-parametric model as the only the partial hazard, $exp(\sum_{i=1}^{n} b_i(x_i - \bar{x}_i))$, has a parametric definition. The baseline hazard, $b_0(t)$, is non-parametric. It must be noted that only the baseline hazard varies through time. The partial hazard merely increases or decreases the baseline hazard. This effect is constant through time. Section 5.3.2 elaborates on this concept.

5.3.2. Assumptions

The main assumption to consider in the Proportional Hazards Model is presented in its name. The hazard is assumed to be proportional to the baseline hazard. Equation 5.5 shows that the partial hazard merely scales the baseline hazard. Another assumption that follows from the model definition and the proportionality assumption is the fact that the effect a covariate has on the baseline hazard is constant in time. This last assumption is challenging regarding the nature of this research, since information from pilot complaints is very time-variant. Information is presented at some moment in time while being unknown before, and this information might become less relevant in time. Measures taken to prevent these assumptions from being violated due to the pilot complaint information time-variance are described in chapter 6.

5.3.3. Inputs

Each entry of the input data for a proportional hazards model consist of a survival time, information on whether the event of interest was observed or censored and values for each of the covariates. An example of the data format can be seen in table 5.1. In this example, the time column represents the time the specimen was under study. The observed column indicates whether failure has been observed while the specimen was under study. A value of zero corresponds with a censored specimen. It can be noted that the values of the covariates can be binary, positive, negative, large and small.

Table 5.1: Example of data format as input for Proportional Hazards Model.

Time	**Observed**	X_1	X_2	X_3
4	1	0	0.4	230
15	0	1	-0.7	572
11	1	1	0.1	48
15	0	0	0.2	458
2	1	0	1.3	103

5.3.4. Outputs

While the Proportional Hazards Model is uniquely defined by the shape of its baseline hazards and the values of it coefficients mentioned in equation 5.5, its output usually consist of hazard ratios. These hazard ratios can be presented graphically in a forest plot, together with their corresponding confidence intervals, as can be seen in figure 5.2. This figure shows the hazard ratios on a logarithmic scale. The whiskers define the extend of the 95% confidence interval. The zero line is also marked. Any value right of zero increases the baseline hazard for in increase in the value of the covariate, while a value to the left of the zero line decreases the baseline hazard for an increase in the covariate.

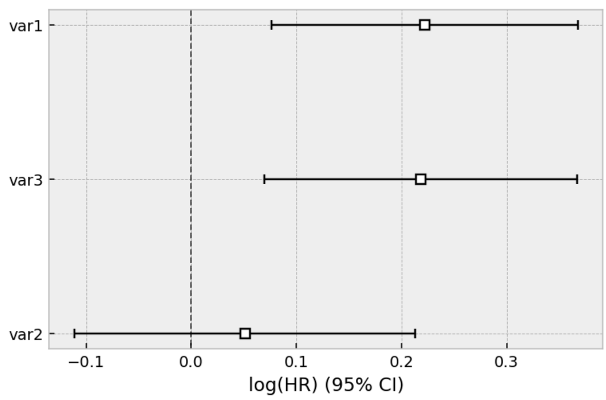


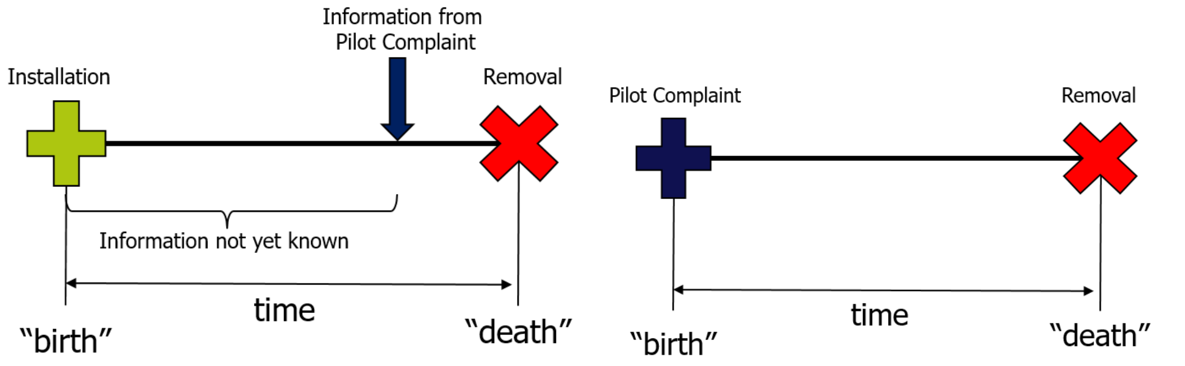
Figure 5.2: Generic example of a forest plot with hazard ratios as the output of a Proportional Hazards Model.[9]

6 Implementation

This chapter elaborates on how the models presented in chapter 5 are implemented in order to answer the questions posed in chapter 4. The survival models described in sections 5.2 and 5.3 are implemented using a readily available open source python package named ”Lifelines” [9]. The input format of this implementation coincides with the format mentioned table 5.1. The ”Time” and ”Observed” values for each entry depend on choices made regarding the definition of ”birth” and the reduction of the observation period mentioned in sections 6.1 and 6.2 respectively. The algorithm used to produce the required data format is described in section 6.3.

6.1. Left Truncation

In the Proportional Hazards Model, the covariates are defined as time-invariant. The case when the value of a covariate changes with time is considered to be in violation with the proportionality assumption discussed in section 5.3.2. Figure 6.1a depicts the situation where the birth is defined as the moment of a components installation. The information in the pilot complaints is added somewhere between birth and death, death being the moment of component removal. It is evident that this information was not yet know before the onset of the pilot complaint. The covariate representing the pilot complaint or its content is therefore time-variant. Figure 6.1b shows the situation where the birth moment coincides with the onset of the pilot complaint. The information presented in the pilot complaint is known during the entire time line and is not in violation of the proportionality assumption in the same way as the situation in figure 6.1a. The birth is therefore defined as the onset of each pilot complaint. The method described in this section is essence ”left truncation”, as shown in figure. 2.4 The truncated part corresponds with the part covered by the bracket, showing where the information is not yet known, in figure 6.1a



(a) "Situation where the component installation is selected as the birth moment."

(b) Situation where the pilot complaint is selected as the birth moment

Figure 6.1: Two options for birth moment.

6.2. Right Censoring

The ability of the Proportional Hazards Model to handle censored data allows for an artificial reduction of the observation period. Reducing the observation period give the partial hazards less opportunity to depart from the proportionality assumption as mentioned in section 5.3.2 by showing time-variant behaviour. A graphical representation of this concept can be seen in figure 6.2. This figure shows an observation period that continues until a time value of 25. The dashed line at 10 represents the artificial end of the observation period. Even though two of the three blue lines fall within the original observation period, they are "artificially right censored" in the new "reduced observation period.

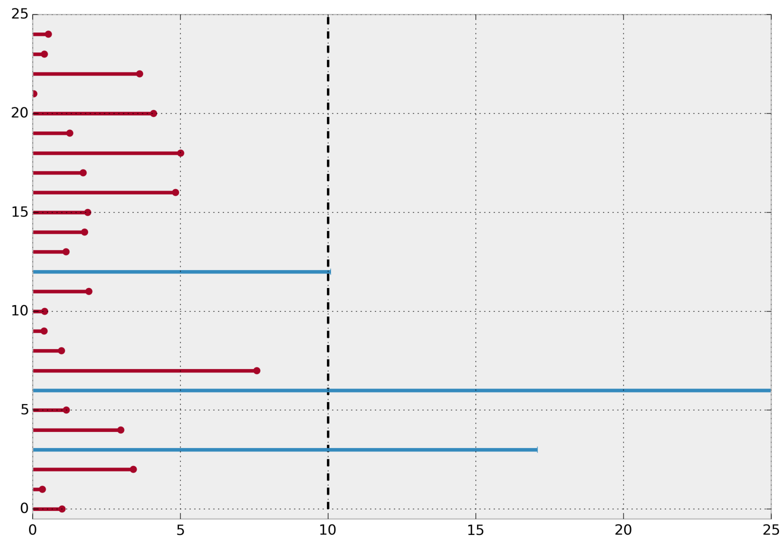


Figure 6.2: Example of survival data with (artificial) right censoring represented by the dashed line. [9]

6.3. Algorithm

In order to fit the Proportional Hazards Model and the Kaplan-Meier estimator, the data must fit the format required by the Lifelines package. This format is mentioned in table 5.1. Getting the data in the correct format consists of determining whether a subsequent component removal was observed and if so, within what time frame. Additionally, the values of the co-variates, mentioned in section 4.1, must be determined. The algorithm that performs these tasks is mentioned in algorithm 1.

Algorithm 1: Algorithm used to obtain data of the format described in table 5.1 as input to the Proportional Hazards Model.

```
for all sampled components do
    for all aircraft do
        for all pilot complaints do
            search for subsequent removal
            if subsequent removal exists then
                $\delta_t = t_{removal} - t_{complaint}$
                if $\delta_t > t_{max}$ then
                    time=$t_{max}$
                    observed=0
                else
                    time=$\delta_t$
                    observed=1
                end
            else
                time=$t_{max}$
                observed=0
            end
            determine values of covariates
            save column
        end
    end
end
```

7 Results

This chapter presents the results as described in chapter 4. The following results are presented for each component, if available:

- **TF-IDF:** Table with the eight highest scoring words in the TF-IDF analysis as described in section 5.1, including term frequency (TF) and document frequency (DF). The chosen words are marked in bold.
- **Proportional Hazards Model Fit ATA:** Forest plot of the hazard ratios for ATA chapter including a table with the summary for the fit.
- **Proportional Hazards Model Fit Words:** Four forest plot containing the hazard ratios for the fits for each of the words from the TF-IDF analysis. The fit summary tables corresponding to the latter four forest plot can be found in appendix A.
- **Proportional Hazards Model Fit Irrelevant Words:** Forest plot containing the hazard ratios of for each of the best scoring words from each components. The words that are relevant for one component are assumed to be irrelevant for other components. The fit summary tables can be found in appendix B.
- **Comparison Kaplan-Meier Estimator:** Comparison of the ATA results between Kaplan-Meier Estimator and Proportional Hazards Model.
- **Comparison Part Number Mentioned:** Comparison between Proportional Hazards Model fit of best scoring word (lower bound hazard of ratio furthest from zero) and Kaplan-Meier Estimator of having the part number mentioned in the complaint.

7.1. Component 0: Oxygen Bottle

TF-IDF

Table 7.1 shows the best scoring words for Component 0: Oxygen Bottle. It can be observed that most words identify the subject of the pilot complaint. Only three out of the eight words presented say something about the state of the system. The numbers ”1560” and ”1500” together with ”psi” say something about the oxygen levels in the oxygen bottle. The word ”below” indicates oxygen levels below the prescribed values, most likely being the two numbers mentioned earlier.

Table 7.1: Overview of TF-IDF scores for Component 0: Oxygen Bottle.

Word	TF	DF	Score
o2-bottle	154	281	52.68
1560	158	527	49.10
oxygen	304	1180	48.99
bottle	385	2157	46.52
1500	208	1313	45.54
psi	439	2987	45.09
o2	236	1604	44.41
below	236	2013	43.69

Proportional Hazards Model Fit ATA

Figure 7.1 presents the results of the Proportional Hazards Model fit of the three endogenous covariates as well as exogenous covariate ”ATA”. The hazard does not seem to increase nor decrease each year, as the variable ”Year” shows a perfect example of no measurable effect. The other three covariates in this fit show hazard reducing effects while its p-value indicates statistical significance. A fresh installation shows a reduced hazard for a removal, which does not sound off. The fact that the indication of the corresponding ATA chapter gives a hazard reducing effect is more surprising. This means that a pilot complaint about the oxygen system reduces the hazard for a removal of the oxygen bottle. This phenomenon is further discussed in section 8.3.4. It must be noted that the ATA covariate violates the proportionality assumption mentioned in 5.3.2, this problem is further discussed in section 8.2.1.

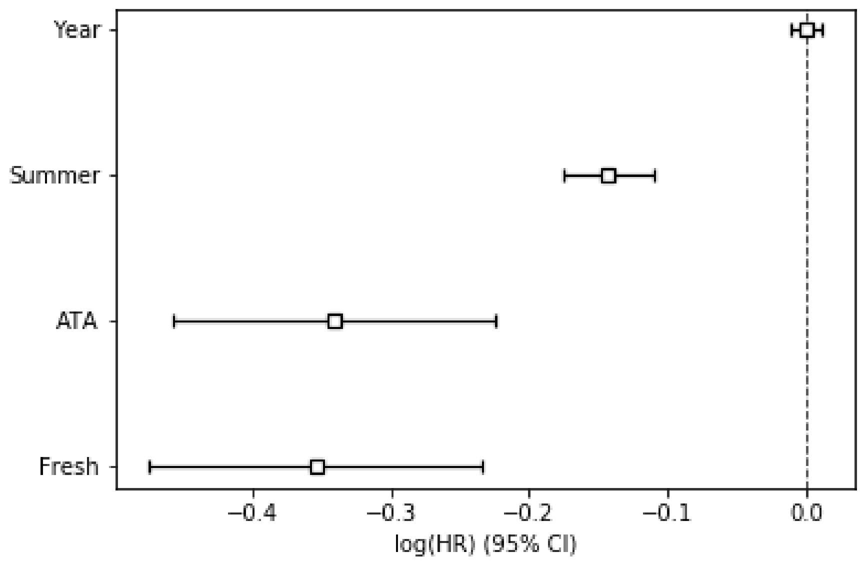


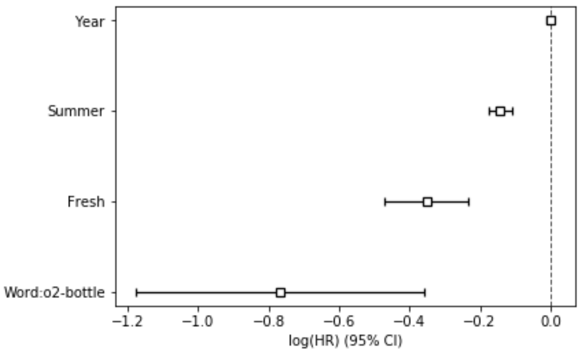
Figure 7.1: Forest plot of hazard ratios of Proportional Hazards Model fit of Component 0: Oxygen Bottle.

Table 7.2: Summary of Proportional Hazards Model fit for Component 0: Oxygen Bottle.

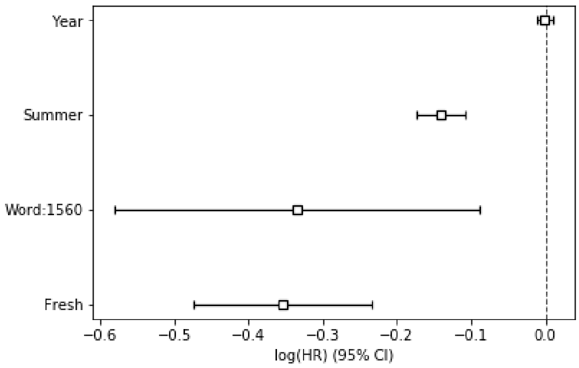
	coef	exp(coef)	se(coef)	z	p	-log2(p)	lower 0.95	upper 0.95	pa
ATA	-0.34	0.71	0.06	-5.74	<0.005	26.64	-0.46	-0.22	✗
Fresh	-0.35	0.70	0.06	-5.77	<0.005	26.87	-0.47	-0.23	✓
Summer	-0.14	0.87	0.02	-8.50	<0.005	55.52	-0.18	-0.11	✓
Year	-0.00	1.00	0.01	-0.06	0.95	0.07	-0.01	0.01	✓

Proportional Hazards Model Fit Words

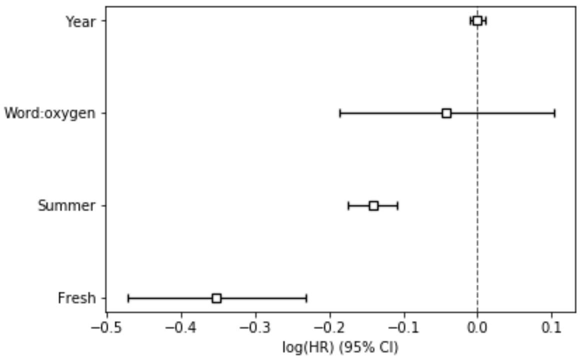
Figure 7.2 shows the fits for each of the four best scoring words in the TF-IDF analysis. As with ATA, all words show a reducing effect on the hazard of an oxygen bottle removal. The fit summaries containing the numerical values of this fit are presented in appendix A.1 to A.3. The numerical values confirm that the values for the endogenous covariates do not change with different words, indicating independence. The word ”o2-bottle” displays the largest scaling of the baseline hazard. It can be noted that mentioning ”o2-bottle” has a greater effect than just having the pilot complaint about the oxygen system, although this effect again the opposite of what is to be expected, as mentioned in section 8.3.4.



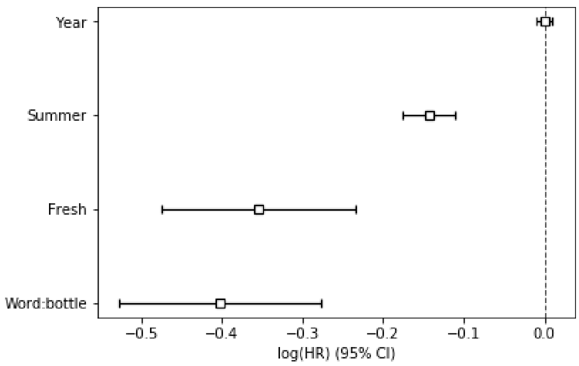
(a) ”o2-bottle”



(b) ”1560”



(c) ”oxygen”



(d) ”bottle”

Figure 7.2: Forest plots of hazard ratios of Proportional Hazards Model fit of Component 0: Oxygen Bottle for different words.

Proportional Hazards Model Fit Irrelevant Words

Figure 7.3 shows the results of the Proportional Hazards Model fit including the words deemed irrelevant. The p-values of the irrelevant words indicate statistical insignificance, as shown in appendix B.1.

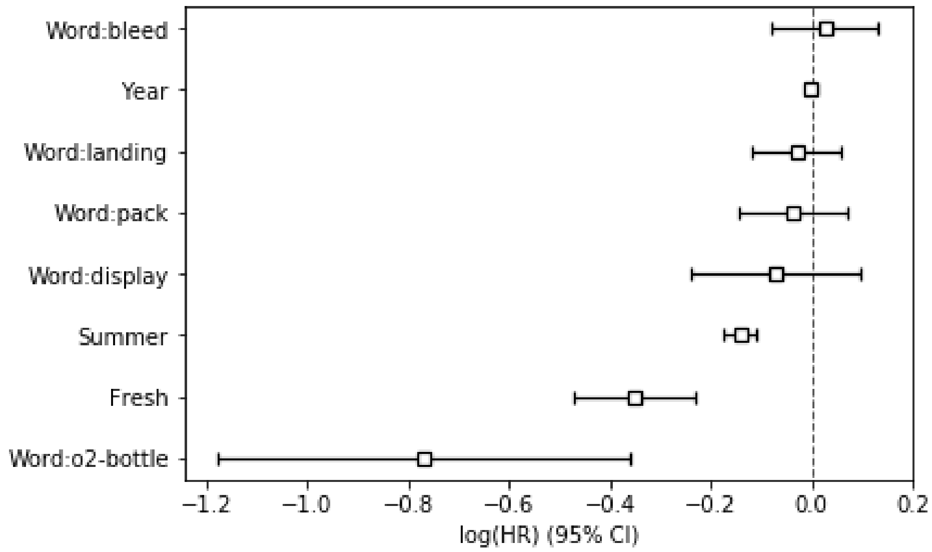


Figure 7.3: Forest plot showing hazard ratios of best scoring words of each component applied to Component 0: Oxygen Bottle.

Comparison Kaplan-Meier Estimator

Figure 7.4 compares the hazard function arising from the Proportional Hazards Model fit for ATA with the Kaplan-Meier Estimator. It can be noted that the baseline survival function for both models approximate each other. It can be noted that the survival estimation for ATA is underestimated by the Proportional Hazards Model with respect to the Kaplan-Meier Estimator.

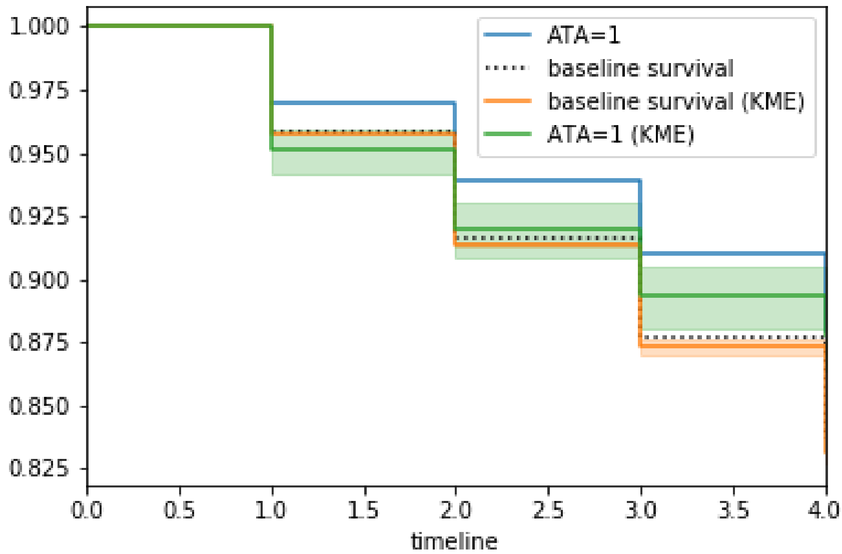


Figure 7.4: Comparison between Kaplan-Meier Estimator and Proportional Hazards Model for Component 0: Oxygen Bottle.

Comparison Part Number Mentioned

Figure 7.5 shows that having the part number mentioned in the respectively added actions has less of a reducing effect on the hazard that the mentioning of the word ”o2-bottle”. This effect is opposite of what is to be expected, this is further discussed in section 8.3.4.

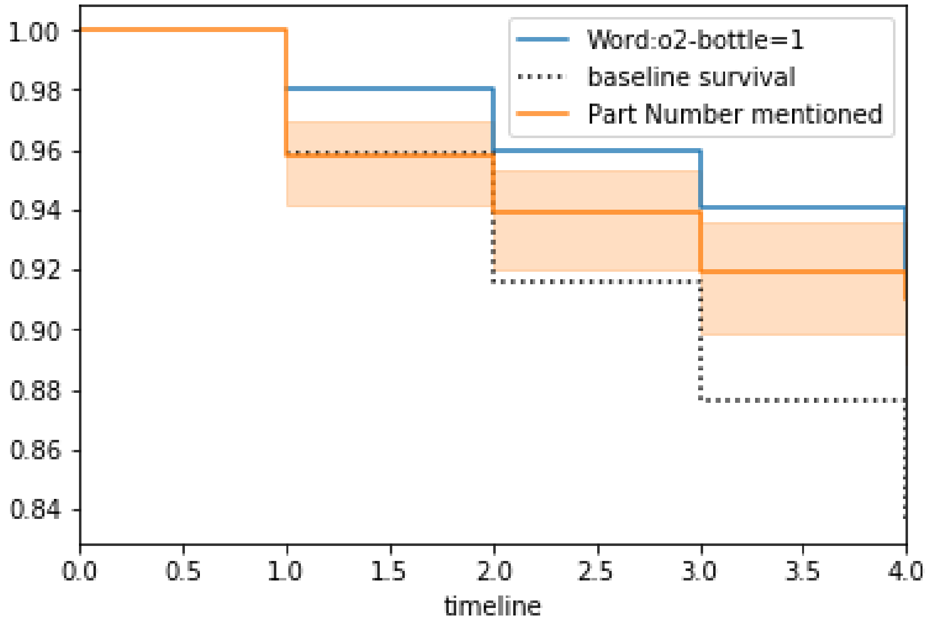


Figure 7.5: Comparison between Kaplan-Meier Estimator for having the part number mentioned and Proportional Hazards Model for Component 0: Oxygen Bottle for Word:o2-bottle.

Sensitivity

Figure 7.6 shows the results of the sensitivity analysis of the hazard ratio and the p-value with respect to the observation time. It can be noted that the hazard ration increases with a reduction in observation time while the p-value reduces. Figures 7.4 and 7.5 indicate an overestimation of the reducing effect, indicating a too low hazard ratio. The desirability of a high p-value makes a change in observation time desirable with respect to the hazard while being undesirable with respect to statistical significance.

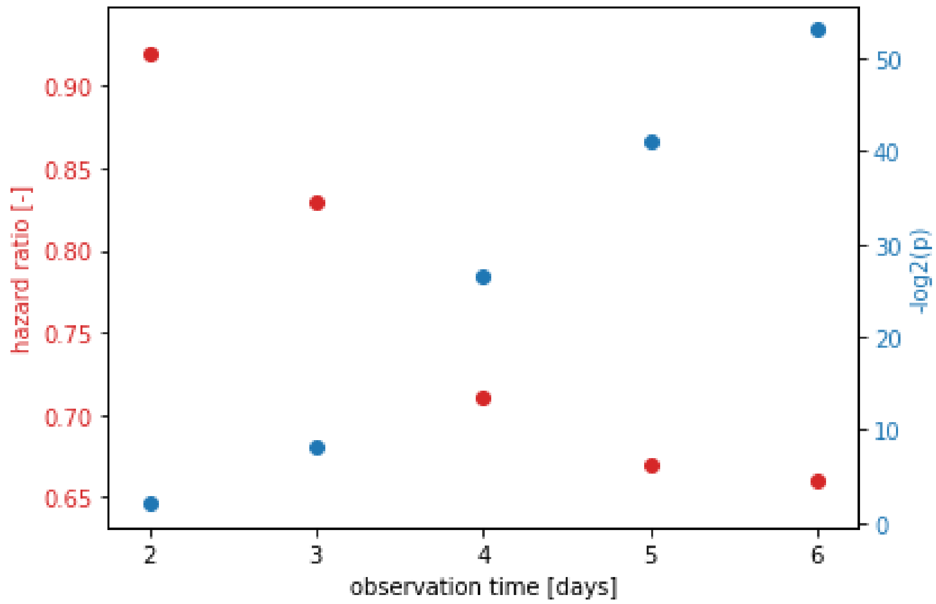


Figure 7.6: Sensitivity analysis of the effect of variation of observation time for Component 0: Oxygen Bottle.

7.2. Component 1: Flow Control Valve

TF-IDF

Table 7.3 shows the eight best scoring words for Component 1: Flow Control Valve in the TF-IDF analysis. The flow control valve is part of the air-conditioning packs that control the temperature inside the cabin. The TF-IDF results show that words regarding the general subject score high. No word regarding the state of the system have made it to the top eight.

Table 7.3: Overview of TF-IDF scores for Component 1: Flow Control Valve.

Word	TF	DF	Score
pack	11	2138	22.76
temp	8	2287	20.42
cabin	11	4457	20.27
control	8	3200	19.39
flow	4	906	18.03
valve	5	2685	16.88
+	7	6259	16.57
with	8	10308	15.80

Proportional Hazards Model Fit ATA

Figure 7.7 together with table 7.4 give the graphical and numerical results of the Proportional Hazards Model fit for ATA chapter, respectively. Exogenous covariate ”ATA” shows a large scaling of the baseline hazard, increasing it more than four fold, while its p-value indicates statistical significance, although the proportionality assumption is violated. The problem with violations of the proportionality assumption is further discussed in section 8.2.1. The only other statistically significant covariate is ”Year”. This means that with the passing of each year, the hazard for a component removal is reduced.

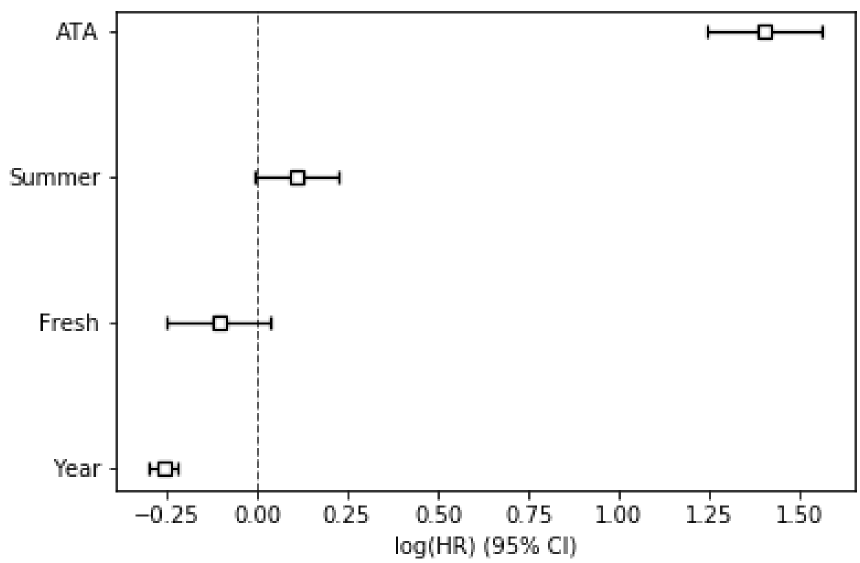


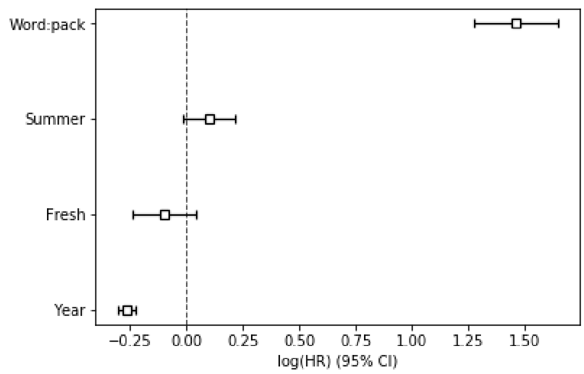
Figure 7.7: Forest plot of hazard ratios of Proportional Hazards Model fit of Component 1: Flow Control Valve.

Table 7.4: Summary of Proportional Hazards Model fit for Component 1: Flow Control Valve.

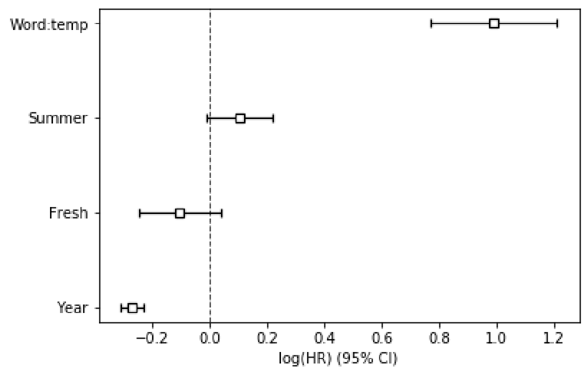
	coef	exp(coef)	se(coef)	z	p	-log2(p)	lower 0.95	upper 0.95	pa
ATA	1.40	4.06	0.08	17.63	<0.005	228.56	1.24	1.56	✗
Fresh	-0.10	0.90	0.07	-1.43	0.15	2.70	-0.25	0.04	✓
Summer	0.11	1.12	0.06	1.91	0.06	4.16	-0.00	0.23	✓
Year	-0.26	0.77	0.02	-12.62	<0.005	118.90	-0.30	-0.22	✓

Proportional Hazards Model Fit Words

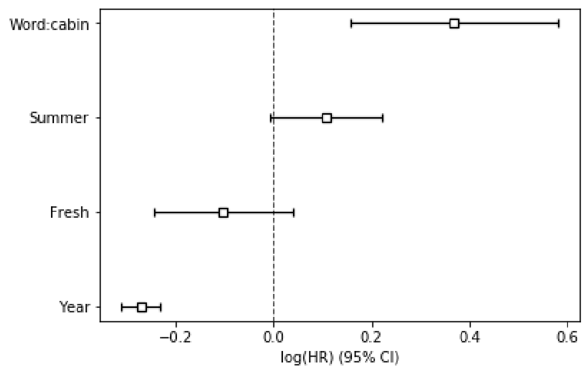
Figure 7.8 shows the graphical results of the Proportional Hazards Model for the four best scoring word in the TF-IDF analysis for this part. The numerical summaries can be found in appendix A. It is easily noted that all words have an increasing effect on the baseline hazard. At first glance, the word ”pack” shows the largest effect, however, this covariate violates the proportionality assumption. The only word that does respect the proportionality assumption is the word ”temp”. This covariate has a hazard ratio of less than three as opposed to more than four for the word ”pack”. The variation in the endogenous covariates only shows very minor variation between fits, implying reasonable independence.



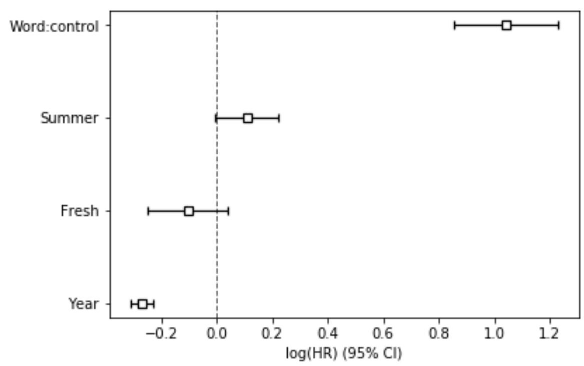
(a) ”pack”



(b) ”temp”



(c) ”cabin”



(d) ”control”

Figure 7.8: Forest plots of hazard ratios of Proportional Hazards Model fit of Component 1: Flow Control Valve for different words.

Proportional Hazards Model Fit Irrelevant Words

Figure 7.9 shows the results for irrelevant words. The numerical summary can be found in table B.2. None of the irrelevant words show any statistical significance demonstrated by their high p-values, as expected.

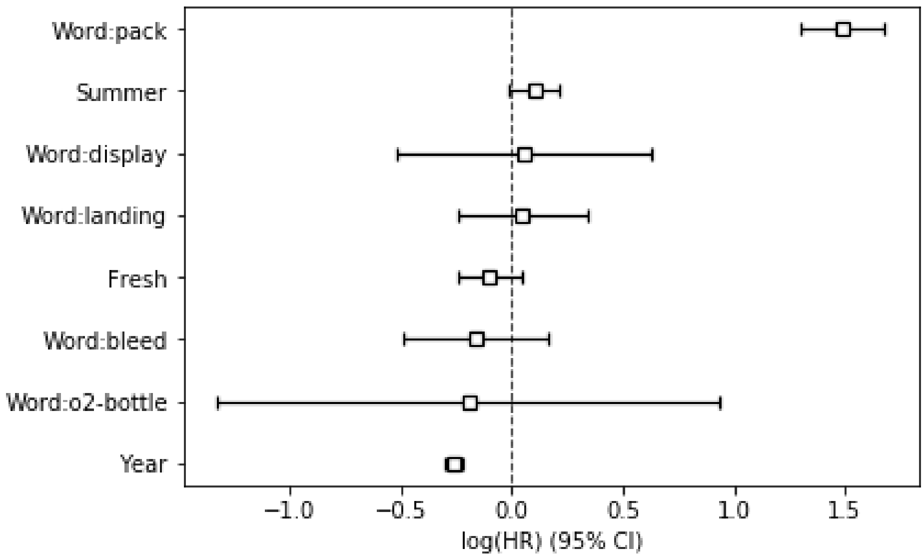


Figure 7.9: Forest plot showing hazard ratios of best scoring words of each component applied to Component 1: Flow Control Valve.

Comparison Kaplan-Meier Estimator

Figure 7.10 compares the hazard function arising from the Proportional Hazards Model fit for ATA with the Kaplan-Meier Estimator. It can be noted that the baseline hazard for the Proportional Hazards Model and the Kaplan-Meier Estimator coincide. However, the effect of ATA in the Proportional Hazards Model shows an underestimation with respect to the Kaplan-Meier Estimator.

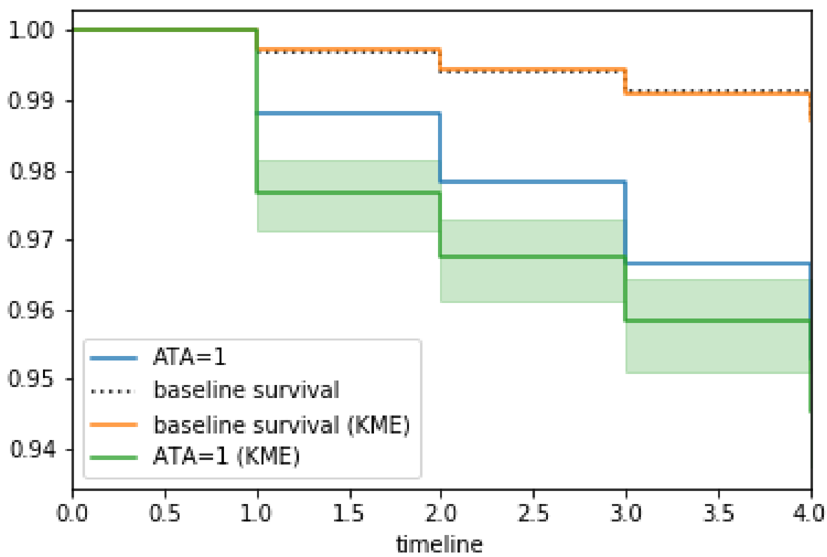


Figure 7.10: Comparison between Kaplan-Meier Estimator and Proportional Hazards Model for Component 1: Flow Control Valve.

Comparison Part Number Mentioned

Figure 7.11 shows the comparison between the Kaplan-Meier Estimator for the best scoring word, ”pack” in this case, and having the part number mentioned in the retrospectively added action. It is immediately obvious that having the part number mentioned reduces the chances of survival significantly. The reduction in survival due to the word ”pack” is much less pronounced. Having the component replaced right after the pilot complaint might suggest that it is very critical.

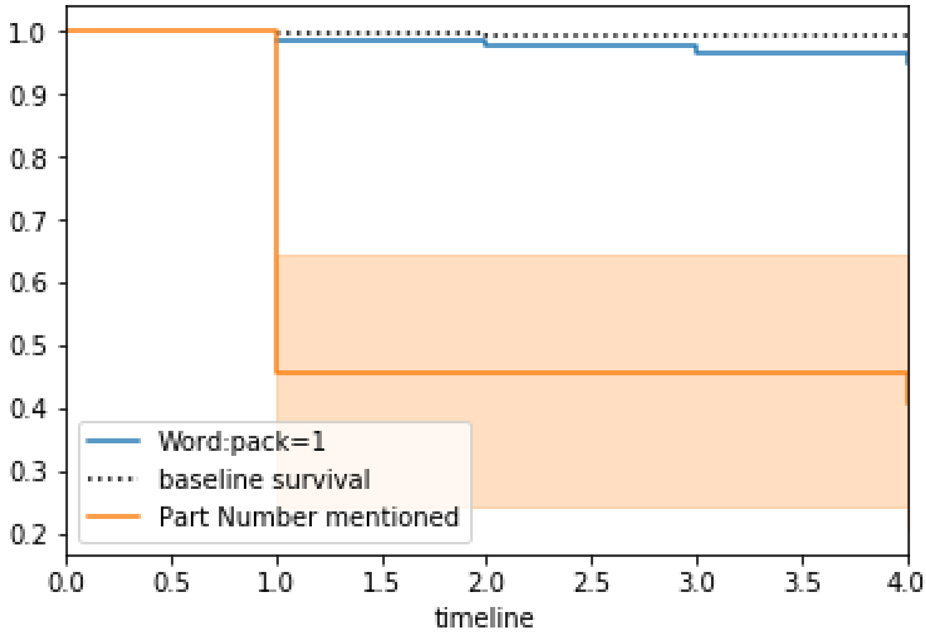


Figure 7.11: Comparison between Kaplan-Meier Estimator for having the part number mentioned and Proportional Hazards Model for Component 1: Flow Control Valve for Word:pack.

Sensitivity

Figure 7.12 shows the results of the sensitivity analysis of the hazard ratio and the p-value with respect to the observation time. While the hazard ratio increases for decreasing observation time, the p-value shows an optimum at three days of observation time. It must be noted that despite the observable optimum, all p-values indicate a high level of statistical significance.

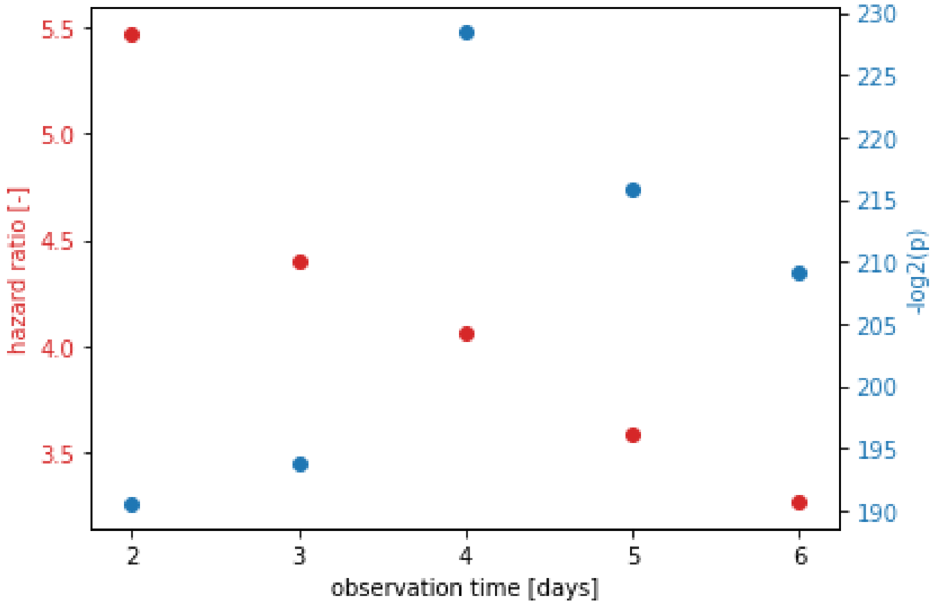


Figure 7.12: Sensitivity analysis of the effect of variation of observation time for Component 1: Flow Control Valve.

7.3. Component 2: Display Unit

TF-IDF

Table 7.5 shows the eight best scoring words for Component 2: Display Unit Valve in the TF-IDF analysis. All four highest scoring words either describe the display itself, of which of the two displays (per pilot). One being the pfd (primary flight display), the other being the nav-display (navigation display), the display where the weather radar is also visualized, hence the appearance of the word ”radar”. None of the four best scoring words give an adverse indication. Only the word ”dim” provides an adverse indication, however, this word was not in the top four best scoring words.

Table 7.5: Overview of TF-IDF scores for Component 2: Display Unit.

Word	TF	DF	Score
pfd	42	780	36.514
display	20	888	30.27
screen	12	391	29.26
radar	9	187	29.20
dim	11	443	28.10
f/o	14	2227	24.23
capt	9	2814	20.54
side	15	7094	20.40

Proportional Hazards Model Fit ATA

Figure 7.13 together with table 7.6 give the graphical and numerical results of the Proportional Hazards Model fit for ATA chapter, respectively. ATA seems to scale the baseline hazard by a factor of almost four while being statistically significant. However, the covariate ATA fails to respect the proportionality assumption. It can be noted that both a fresh installation and the summer season increase the removal hazard.

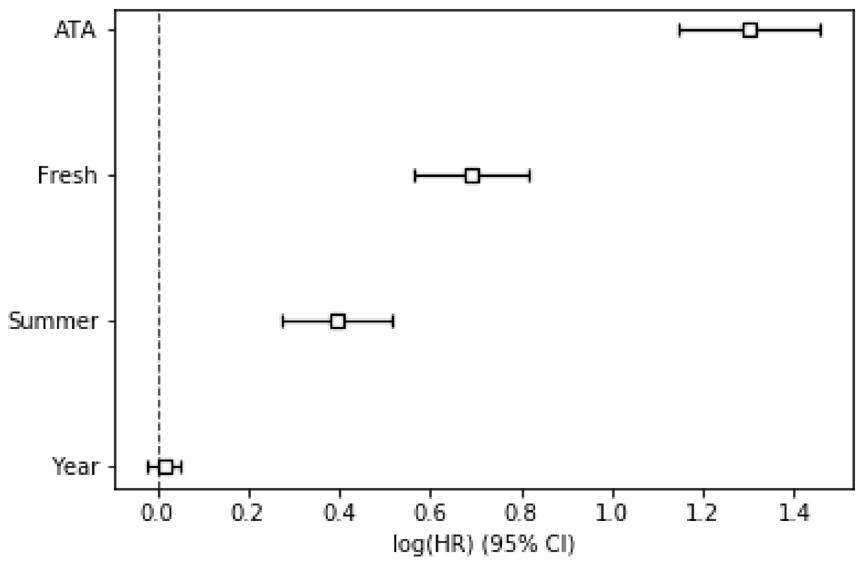


Figure 7.13: Forest plot of hazard ratios of Proportional Hazards Model fit of Component 2: Display Unit.

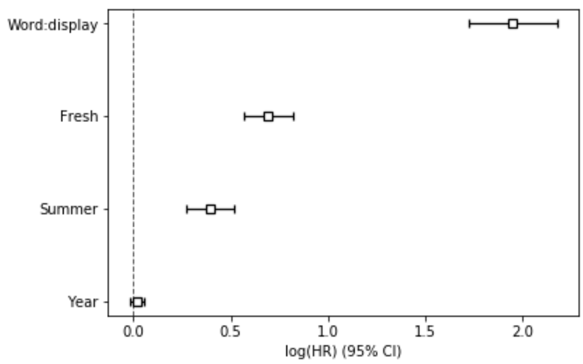
Table 7.6: Summary of Proportional Hazards Model fit for Component 2: Display Unit.

	coef	exp(coef)	se(coef)	z	p	-log2(p)	lower 0.95	upper 0.95	pa
ATA	1.30	3.67	0.08	16.50	<0.005	200.68	1.15	1.45	✗
Fresh	0.69	1.99	0.06	10.69	<0.005	86.16	0.56	0.82	✓
Summer	0.39	1.48	0.06	6.33	<0.005	31.89	0.27	0.52	✓
Year	0.01	1.01	0.02	0.81	0.42	1.26	-0.02	0.05	✓

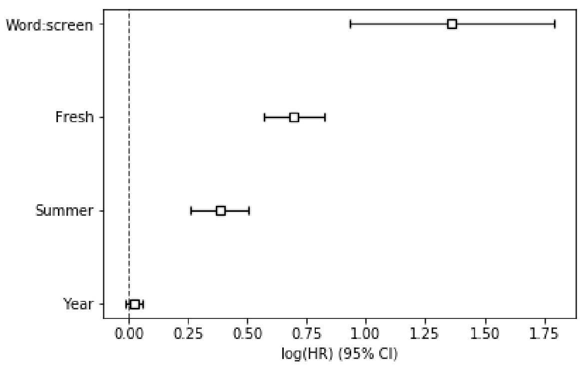
Proportional Hazards Model Fit Words

Figure 7.14 shows the graphical results of the Proportional Hazards Model for the four best scoring word in the TF-IDF analysis for this part. The numerical summaries can be found in appendix A. It can be noted that the results for the word ’pfd’ are missing for the reasons mentioned in section 7.6. The words ”display” and ”screen” both display a statistically significant increase in the baseline hazard. It must be noted that ”display”, while having the highest hazards ratio, does not respect the proportionality assumption. The endogenous covariates show very little variation between fits, indicative of their independence.

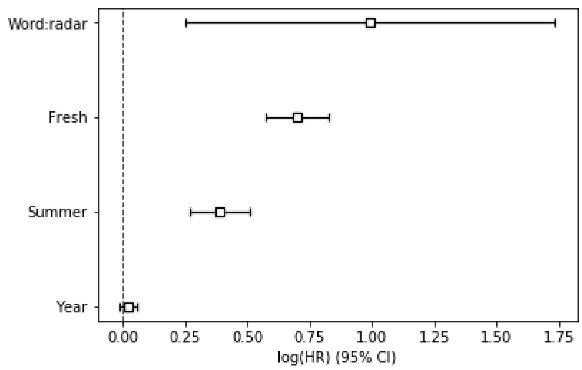
(a) ”pfd”



(b) ”display”



(c) ”screen”



(d) ”radar”

Figure 7.14: Forest plots of hazard ratios of Proportional Hazards Model fit of Component 2: Display Unit for different words.

Proportional Hazards Model Fit Irrelevant Words

Figure 7.15 shows the results for irrelevant words. The numerical summary can be found in table B.3. None of the irrelevant words show any statistical significance demonstrated by their high p-values, as expected.

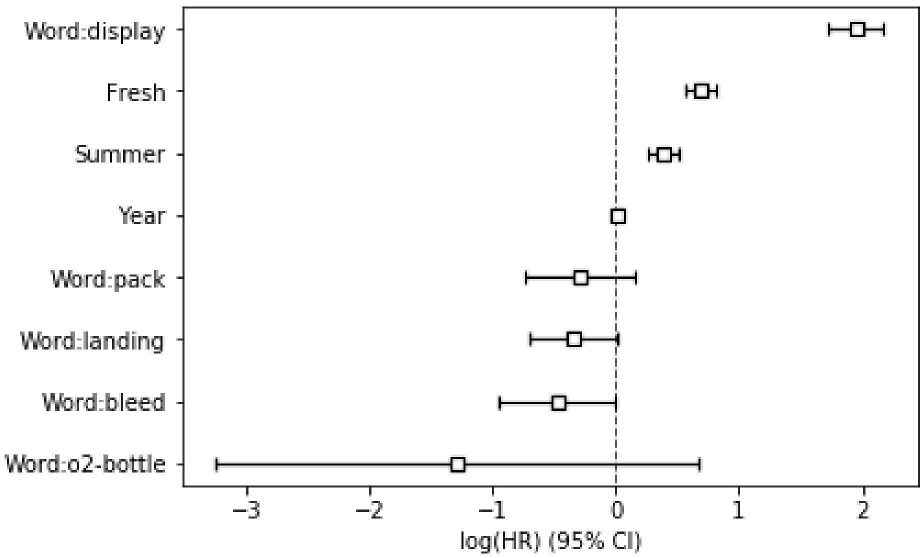


Figure 7.15: Forest plot showing hazard ratios of best scoring words of each component applied to Component 2: Display Unit.

Comparison Kaplan-Meier Estimator

Figure 7.16 compares the hazard function arising from the Proportional Hazards Model fit for ATA with the Kaplan-Meier Estimator. It can be noted that the baseline hazard for the Proportional Hazards Model and the Kaplan-Meier Estimator coincide. However, the effect of ATA in the Proportional Hazards Model shows an underestimation with respect to the Kaplan-Meier Estimator.

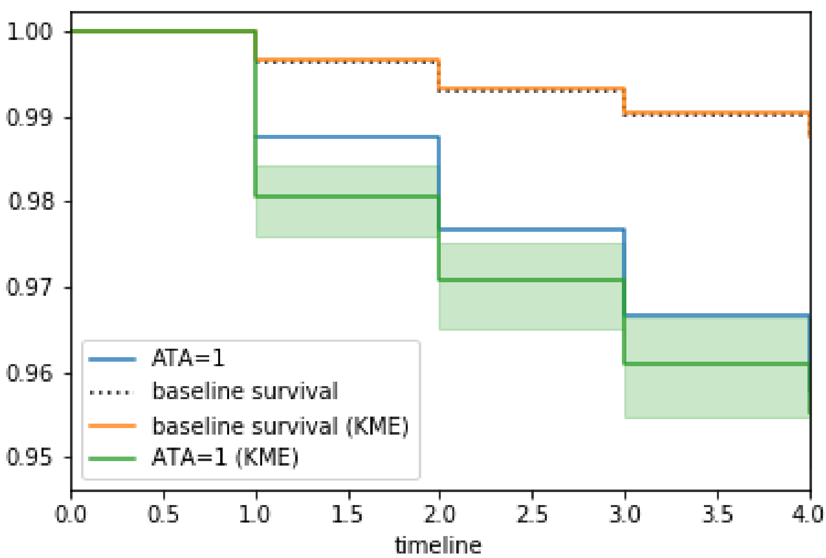


Figure 7.16: Comparison between Kaplan-Meier Estimator and Proportional Hazards Model for Component 2: Display Unit.

Comparison Part Number Mentioned

Figure 7.17 shows the comparison between the Kaplan-Meier Estimator for the best scoring word, ”display” in this case, and having the part number mentioned in the retrospectively added action. It is immediately obvious that having the part number mentioned reduces the chances of survival significantly. The reduction in survival due to the word ”display” is less pronounced. The survival curve for this component does not suggest the same level of criticality as the flow control valve mentioned in section 7.11.

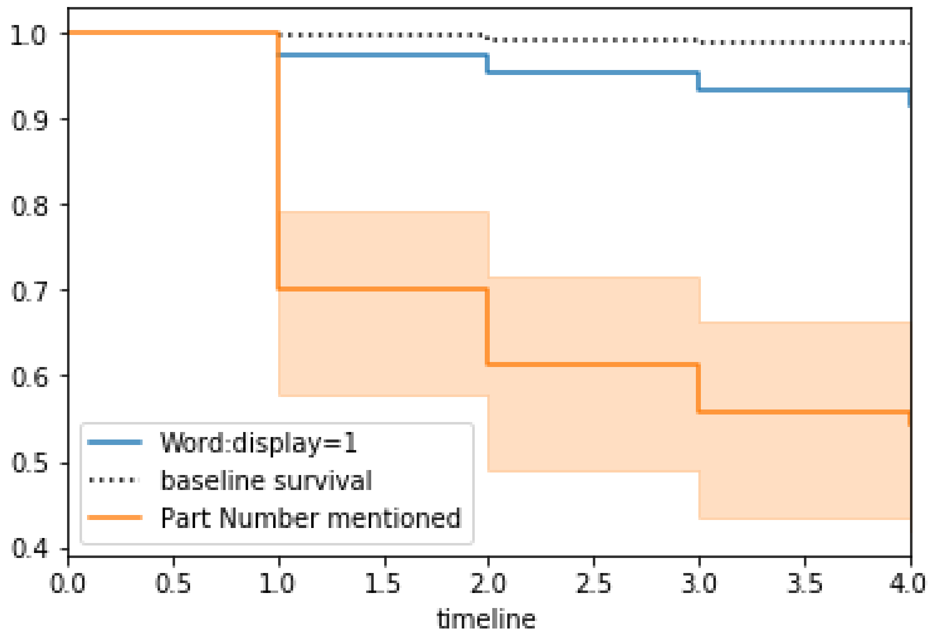


Figure 7.17: Comparison between Kaplan-Meier Estimator for having the part number mentioned and Proportional Hazards Model for Component 2: Display Unit for Word:display.

Sensitivity

Figure 7.18 shows the results of the sensitivity analysis of the hazard ratio and the p-value with respect to the observation time. While the hazard ratio increases for decreasing observation time, the p-value shows an optimum at three days of observation time. It must be noted that despite the observable optimum, all p-values indicate a high level of statistical significance.

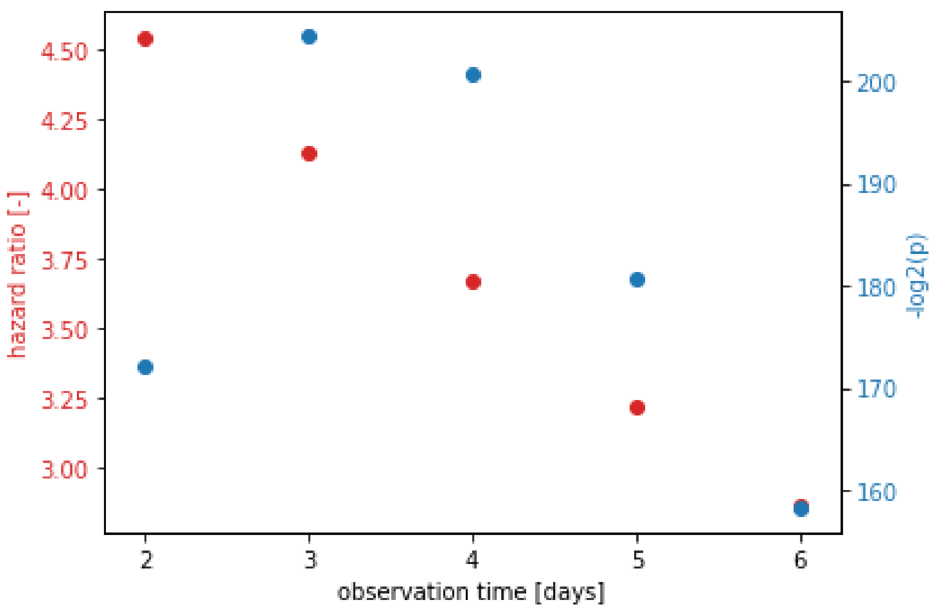


Figure 7.18: Sensitivity analysis of the effect of variation of observation time for Component 2: Display Unit.

7.4. Component 3: Pressure Regulating Shut-Off Valve

TF-IDF

Table 7.7 shows the eight best scoring words for Component 3: Pressure Regulating Shut-Off Valve in the TF-IDF analysis. The top three words are descriptive of the component itself or the system it belongs to. The fourth word is interesting, as it is descriptive of a location, ”rh” stands for right hand side. The only word descriptive of the state of the system is the word ”fault”, however, it is not among the four highest scoring words.

Table 7.7: Overview of TF-IDF scores for Component 3: Pressure Regulating Shut-Off Valve.

Word	TF	DF	Score
prsov	50	317	42.27
bleed	75	2158	35.57
valve	30	2685	30.18
rh	86	8930	30.18
and	38	26665	22.85
the	31	12959	22.74
fault	30	10333	22.74
with	26	10308	22.58

Proportional Hazards Model Fit ATA

No results for ATA. See section 7.6.

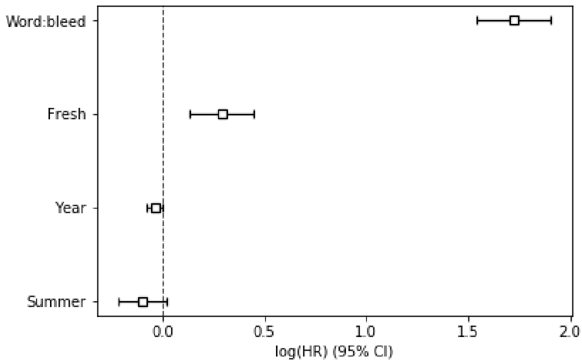
Proportional Hazards Model Fit Words

Figure 7.19 shows the graphical results of the Proportional Hazards Model for the four best scoring word in the TF-IDF analysis for this part. The numerical summaries can be found in appendix A. First thing to be noted is the missing results for the word ”prsov”. Is is due to the same reason as for the ATA chapter and is further discussed in section 7.6. A very high, at 5.60, and statistically significant hazard ratio is observed for the word ”bleed”, although this word violated the proportionality assumption. The word ”valve” has more moderate results while also violating the proportionality assumption. Interestingly, a fresh installation indicates in increased hazard with respect to the baseline hazard. The variable ”rh” is a borderline case with respect to statistical significance, which is surprising due to its generic description of location. The very small variance in the endogenous covariates is indicative of independence.

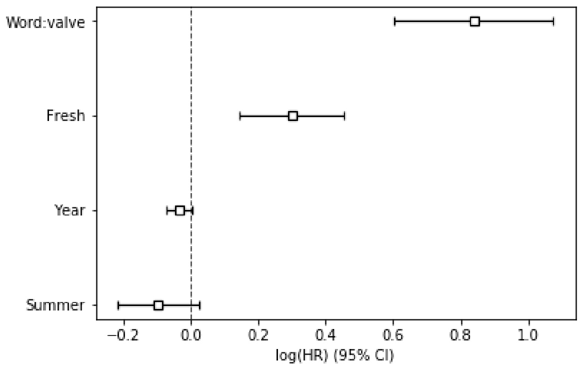
Proportional Hazards Model Fit Irrelevant Words

Figure 7.20 shows the results for irrelevant words. The numerical summary can be found in table B.4. None of the irrelevant words show any statistical significance demonstrated by their high p-values, as expected.

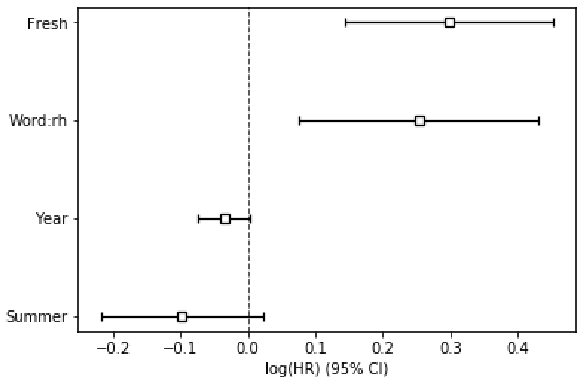
(a) "prsov"



(b) "bleed"



(c) "valve"



(d) "rh"

Figure 7.19: Forest plots of hazard ratios of Proportional Hazards Model fit of Component 3: Pressure Regulating Shut-Off Valve for different words. No results for 'prsov', see section 7.6.

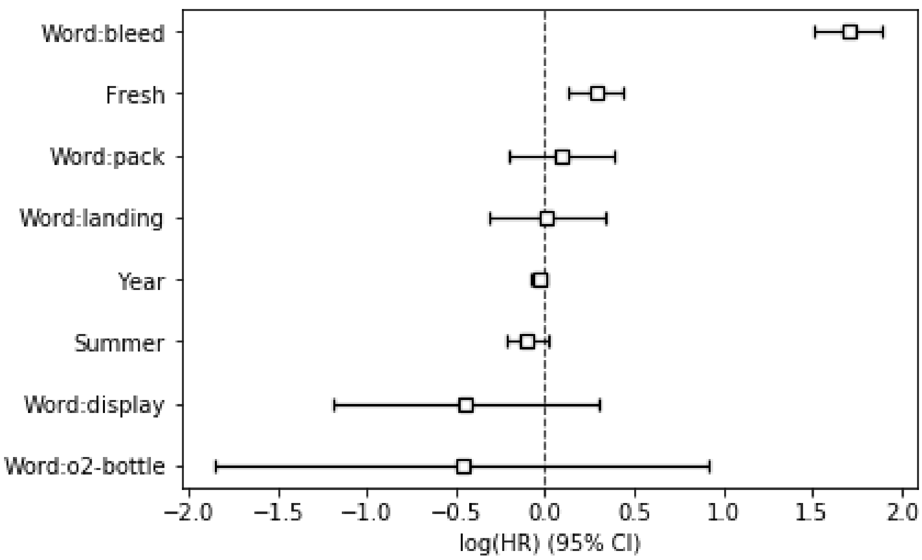


Figure 7.20: Forest plot showing hazard ratios of best scoring words of each component applied to Component 3: Pressure Regulating Shut-Off Valve.

Comparison Kaplan-Meier Estimator

No results for Proportional Hazards Model to compare to the Kaplan-Meier Estimator, see section 7.6.

Comparison Part Number Mentioned

Figure 7.21 shows the comparison between the Kaplan-Meier Estimator for the best scoring word, ”bleed” in this case, and having the part number mentioned in the retrospectively added action. It is immediately obvious that having the part number mentioned reduces the chances of survival signinificantly.

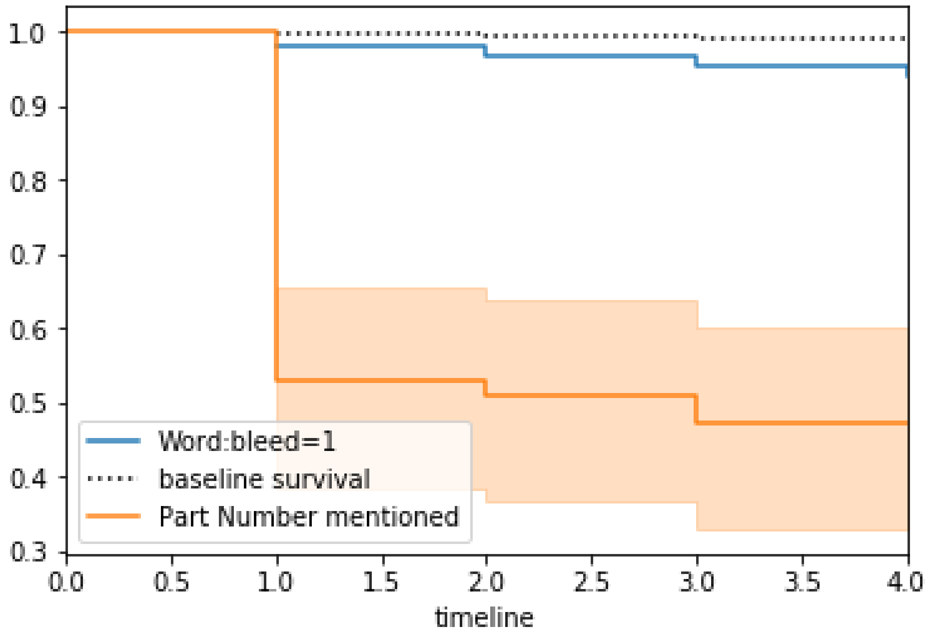


Figure 7.21: Comparison between Kaplan-Meier Estimator for having the part number mentioned and Proportional Hazards Model for Component 3: Pressure Regulating Shut-Off Valve for Word:bleed.

Sensitivity

No results for the sensitivity analysis, see section 7.6.

7.5. Component 4: Landing Light

TF-IDF

Table 7.8 shows the eight best scoring words for Component 4: Landing Light in the TF-IDF analysis. This component shows a higher variety in types of word that score high. The word ”landing” could both be part of the name, or indicate a problem during landing. The word ”extended” indicates a state of the system, albeit not an adverse one. The duplicate word ”extended.” containing the period is omitted from the top four. The word ”retract” is a verb indicating an action by the system. Finally, the word ”lh”, or left hand side, indicates a location.

Table 7.8: Overview of TF-IDF scores for Component 4: Landing Light.

Word	TF	DF	Score
landing	30	3147	27.78
extended	7	161	27.35
extended.	3	25	23.39
retract	16	548	21.03
lh	34	8061	20.28
light	5	20993	20.03
ldg	14	969	19.54
rh	15	8930	19.19

Proportional Hazards Model Fit ATA

Figure 7.22 together with table 7.9 give the graphical and numerical results of the Proportional Hazards Model fit for ATA chapter, respectively. ATA seems to scale the baseline hazard by 60% while being statistically significant. It must be noted that this covariate does respect the proportionality assumption. The only other covariate that is statistically significant is ”Year”, indicating a reduction in hazard for each more year.

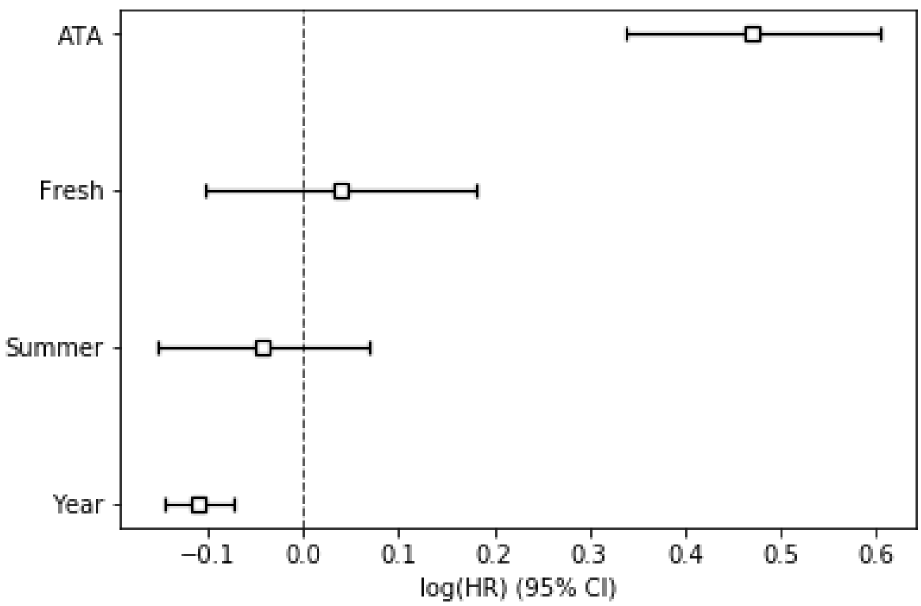


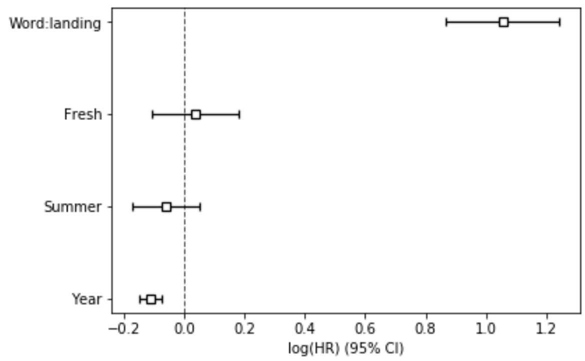
Figure 7.22: Forest plot of hazard ratios of Proportional Hazards Model fit of Component 4: Landing Light.

Table 7.9: Summary of Proportional Hazards Model fit for Component 4: Landing Light.

	coef	exp(coef)	se(coef)	z	p	-log2(p)	lower 0.95	upper 0.95	pa
ATA	0.47	1.60	0.07	6.91	<0.005	37.59	0.34	0.60	✓
Fresh	0.04	1.04	0.07	0.54	0.59	0.76	-0.10	0.18	✓
Summer	-0.04	0.96	0.06	-0.76	0.45	1.16	-0.15	0.07	✓
Year	-0.11	0.90	0.02	-5.93	<0.005	28.30	-0.15	-0.07	✓

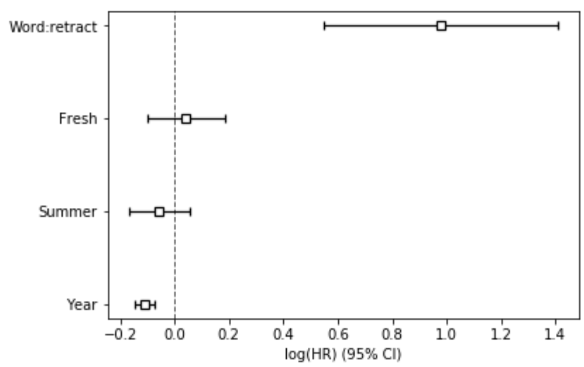
Proportional Hazards Model Fit Words

Figure 7.23 shows the graphical results of the Proportional Hazards Model for the four best scoring word in the TF-IDF analysis for this part. The numerical summaries can be found in appendix A. The analysis for the word ”extended” is missing, the reasons for this are described in section 7.6. Of the words, ”lh” misses statistical significance, as expected. The word ”retract”, although statistically significant, has a large standard deviation, as shows by the wide whiskers. The word ”landing” is the best performing word in this analysis, showing a hazard ratio of almost three, while being statistically significant and respecting the proportionality assumption.

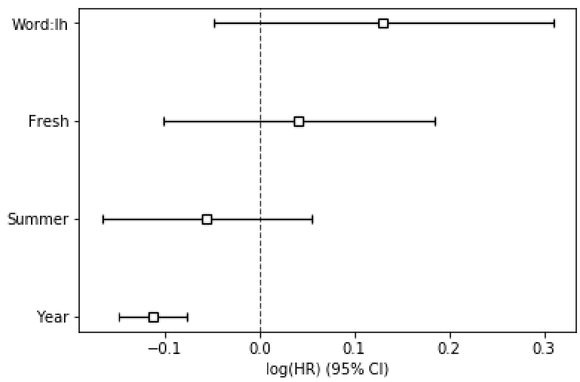


(a) ”landing”

(b) ”extended”



(c) ”retract”



(d) ”lh”

Figure 7.23: Forest plots of hazard ratios of Proportional Hazards Model fit of Component 4: Landing Light for different words. No results for ’extended’, see section 7.6.

Proportional Hazards Model Fit Irrelevant Words

Figure 7.24 shows the results for irrelevant words. The numerical summary can be found in table B.5. None of the irrelevant words show any statistical significance demonstrated by their high p-values, as expected.

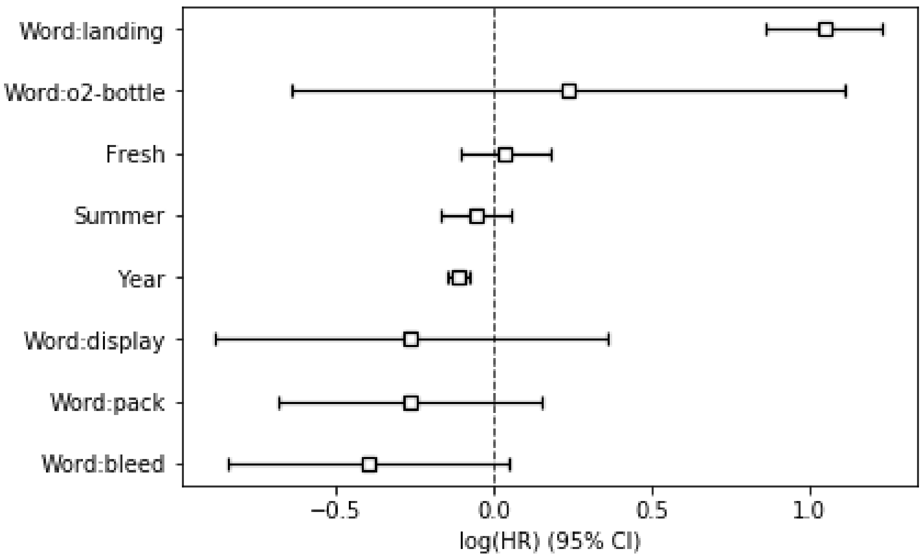


Figure 7.24: Forest plot showing hazard ratios of best scoring words of each component applied to Component 4: Landing Light.

Comparison Kaplan-Meier Estimator

Figure 7.25 compares the hazard function arising from the Proportional Hazards Model fit for ATA with the Kaplan-Meier Estimator. It can be noted that not just the baseline hazard, like in the other components, but also the values for ATA seem to coincide.

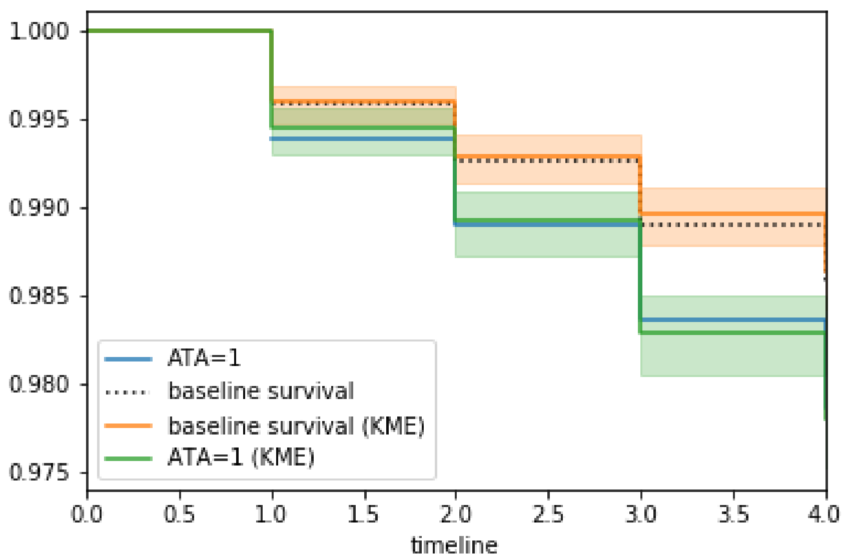


Figure 7.25: Comparison between Kaplan-Meier Estimator and Proportional Hazards Model for Component 4: Landing Light.

Comparison Part Number Mentioned

Figure 7.26 shows the comparison between the Kaplan-Meier Estimator for the best scoring word, "landing" in this case, and having the part number mentioned in the retrospectively added action. It is immediately obvious that having the part number mentioned reduces the chances of survival significantly, although not a significantly as the other components. This component appears to be the least "critical" of all components considered, something to be expected due to its nature.

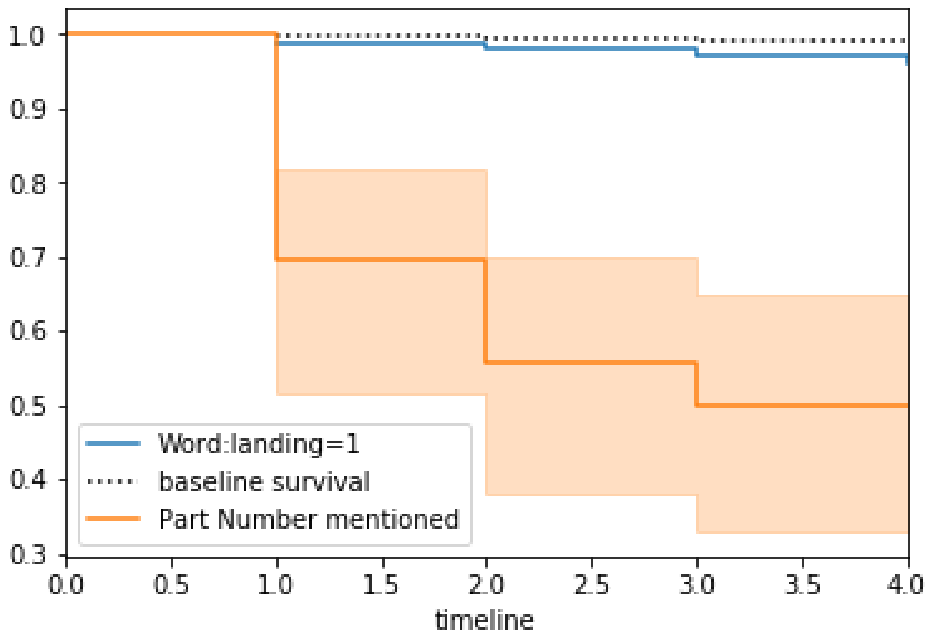


Figure 7.26: Comparison between Kaplan-Meier Estimator for having the part number mentioned and Proportional Hazards Model for Component 4: Landing Light for Word:landing.

Sensitivity

Figure 7.27 shows the results of the sensitivity analysis of the hazard ratio and the p-value with respect to the observation time. It can be noted that the highest hazard ratio is obtained using an observation period of three days. The p-value shows a local optimum at four days.

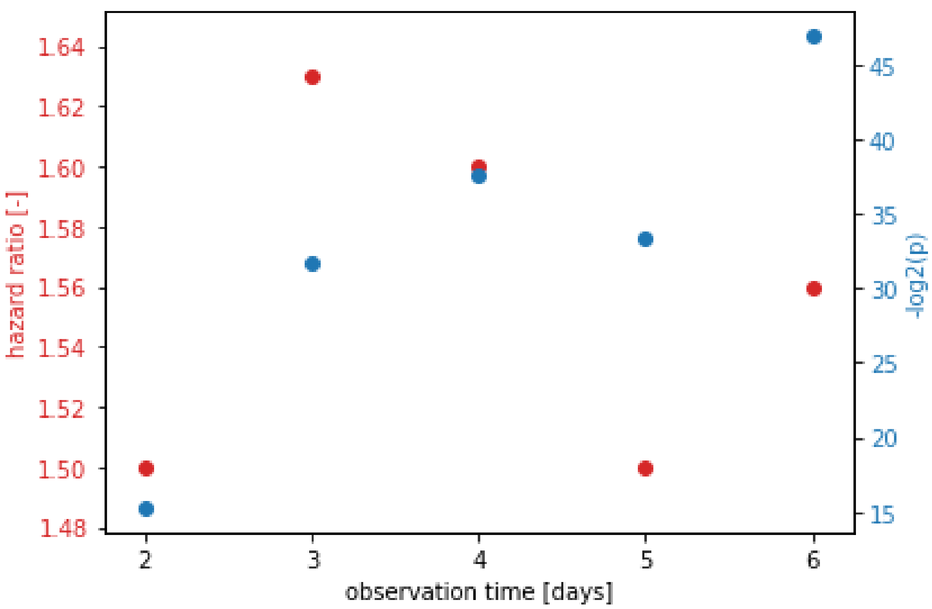


Figure 7.27: Sensitivity analysis of the effect of variation of observation time for Component 4: Landing Light.

7.6. Missing Results

Some of the results are missing. This is caused by high collinearity in the input data. Component 3: Pressure Regulating Shut-Off Valve is especially susceptible to this problem. More on this problem in section 8.2.2.

8
Discussion

This chapter provides a critical review of the results presented in chapter 7, starting with the positive indications with respect to the predictability of component removals in section 8.1. Subsequently, the positive indications are put into perspective by discussing the adverse indications in section 8.2. The chapter is concluded with an interpretation of the results in section 8.3, aiming to find possible explanations for the observed behavior of the components.

8.1. Positive Indications

8.1.1. Large Hazard Ratios

It can be noted from the results than the exogenous covariates provide much greater predictability increases than the endogenous covariates. Notable exception is the Component 0: Oxygen Bottle, which will be further discussed in section 8.3.4. For ATA, the lower end of the confidence interval of Component 4: Landing Light still scores a more than 40% increased hazard with respect to the baseline hazard, whereas the same measure for components 1 and 2 score an increase of 245% and 216% increased hazard. The hazard increase even goes as an increase of at least 458% when looking at textual content for component 2: Display Unit. However, these very high values can be placed into perspective when compared to a perfect predictor, as discussed in section 8.2.3.

8.1.2. Statistical Significance

The statistical significance in the case of the Proportional Hazards Model fit is measured by the p-value that is presented in the fit summary tables. This value represents the probability of the measured hazard ratio having occurred due to the null-hypothesis, or in other words, random chance. This value is lower than 0.005 for the best scoring word with respect to hazard ratio. This statistical significance confirms the measurable effect the content of the pilot complaints have on the predictability of the removals. However, it must be noted that these statistically significant results often occur under violation of the proportionality assumption. More on violations of the proportionality is discussed in section 8.2.1. Some of the endogenous covariates show statistical insignificance in the form of an elevated p-value. These cases of insignificance go hand in hand with very small hazard ratios, to be expected for immeasurable effects.

8.1.3. Increased Hazard Ratio from Content

One of the positive results presented in chapter 7 is the fact that some of the textual content scores better in terms of hazard ratios than the mere occurrence of an ATA related pilot complaint, despite the rudimentary methods used to extract textual content. This raises expectations for more sophisticated methods to yield even better results. Section 9.2.2 will go into further depth on possible extensions.

8.2. Adverse Indications

8.2.1. Non-Proportionality

One of the biggest adverse indications as to the value of the results is the non-proportionality of the partial hazards. The main assumption, the one that is even in the name of the model, is often violated. Only one of the components shows proportionality with respect to ATA. The efforts mentioned in sections 6.1 and 6.2 did not manage to counter the problem of non-proportionality. The exception to this problem seems to be Component 4: Landing Light. This component shows proportional behavior for all its covariates. This might very well correspond with the smallest drop in survivability after one day. Sudden drop appears in all components with the exception of Component 0: Oxygen Bottle. More on the behavior of this component in section 8.3.4. It can be argued that if a model has its assumption violated, results from said model lose their value. However, this violation occurs over a very short duration in time, and the hazard ratios discussed in section 8.1.1 are large. These two factors might save the relevance of the results for an engineering application. More on applicability is to be discussed in section 8.2.3.

8.2.2. Non-Convergence

Another problem that occurs in some case is that of non-convergence. This error is presented by Lifelines [9] in the form of the following message: "Convergence halted due to matrix inversion problems. Suspicion is high collinearity." The reason for this problem stems from the integer nature of the data. Time takes on integer values while most covariates are merely binary. Even the covariate 'Year' is reduced to integer values near zero inside the algorithm. In the case of Component 3, in many of the data entries, like in table 5.1, the values for Time, ATA, and Observed are all 1 simultaneously, causing this problem. A possible solution to this problem in the form of non-binary values for the textual covariates can be found in section 9.2.2.

8.2.3. Limited Practical Applicability

As with any research into engineering solutions, one must consider the practical applicability of the findings. While the positive indications mentioned in 8.1, consisting of large and statistically significant hazard ratios from textual content, might hint to a possible practical application, there are many reasons to doubt that. First of all, whilst statistically significant, section 8.2.1 mentions the violation of the proportionality assumption, weakening the credibility of the results. Even if these results would be backed by proportional behaviour of the covariates, one could assess the predictability increase as perceived by an aircraft operator and come to negative conclusions. Aircraft operators gauge the predictability of a component failure not by departure from a baseline hazard but by a more deterministic prediction of the failure event. The comparison to the Kaplan-Meier Estimation for when the part number is mentioned shows that the results compete very badly with this "perfect estimator". However, the unfairness of this comparison must be noted. This is further discussed in section 8.3.1. The problems mentioned above do not disqualify these findings from applicability completely. An aircraft with an undiagnosed failure could still benefit from having the order of inspections based on the highest scoring parts and could result in quicker diagnostics. Especially if the algorithm is improved by the steps discussed in section 9.2.

8.3. Interpretation

8.3.1. Part Number Mentioned

Having the action mentioned in the pilot complaint has proven to be important in the ability to value the predictability enhancements originating from the pilot complaints. Aside from providing a benchmark for the predictability, it is also a big red flag with respect to the proportionality assumption as discussed in section 8.2.1. It shows that some parts get replaced almost instantly after a pilot complaint. Behavior that is difficult to model using the regular Proportional Hazards Model. Suggested enhancement to this model are discussed in section 9.2.1.

8.3.2. Relevant Words

A pattern can be observed among the highest scoring words for components 1 to 4. These words all seem to be (part of) the name of system it belongs to. In some cases the component is the system itself. This information has a lot of overlap between the ATA chapter itself. Nearly all words used as covariates are nouns that do not describe any negative state of the component. The oxygen bottle is the exception to this rule, having '1560' as covariate, describing an oxygen level. Other components also have words more descriptive of an adverse state in the the form of 'dim' for the display unit, 'fault' for the pressure regulating shut-off valve. Section 9.2.2 discusses options for improvement for the textual part of this research.

8.3.3. Irrelevant Words

The irrelevant words have three important and reassuring patterns. Firstly, their hazard ratios are very close to zero, implying little effect on the hazard, as expected. Secondly, the p-values are high, implying the measured affect aside from small, also like to be caused by chance. Lastly, the irrelevant words nearly always are left of zero, implying a hazard ratio smaller than one. The latter is also to be expected, since if a pilot complaint is about something unrelated, it is not about the component of interest.

8.3.4. Oxygen Bottle Reversal

A frequent exception to most findings in this research is Component 0: Oxygen bottle. The most notable different with other components is the reversal of its behaviour. Whilst all other parts show a very sharp decrease in survival probability, the oxygen bottle exhibits the exact opposite effect: Having the part number mentioned in the action part of the pilot complaint increases the probability of survival. A very likely explanation for this behaviour is the repairable nature of this component. Pilot complaints often mention the reduce oxygen pressure in the bottle. Instead of removing the component, the levels of oxygen are merely replenished. The mechanic mentions the part number of the oxygen bottle without removing it, leaving an "refreshed" oxygen bottle that is less likely to be removed.

8.3.5. Year

Year has no effect for the hazard for the oxygen bottle nor the display unit, as can be seen by its its hazard ratio of 1 as well as high p-values. For the flow control valve and the landing light it does have a statistically significant effect show by its very small p-values. With each year, the hazard is reduced. Possible explanations for this include possible reduced utilization of the aircraft or changes in the maintenance program. The pressure regulating shut-off valve is a borderline case with respect to statistical significance as its p-values vary between 0.05 and 0.10.

8.3.6. Fresh

The the covariate 'Fresh', three components show statistically significant results, being the oxygen bottle, the display unit and the pressure regulating shut-off valve. The hazard ratios however show two different types of behavior. For the oxygen bottle, the hazard is reduced when freshly installed. The other components show the opposite effect, where a fresh installation increase the hazard. The last could be explained by so called 'burn-in failures'.

8.3.7. Summer

Only the oxygen bottle and the display unit show statistically significant results. They are however, just like the covariate for freshness, opposite. The summer season decreasing the hazard for the oxygen bottle while increasing the hazard for the display unit. These differences might stem from different maintenance regimes between summer and winter season. The increased hazard for the removal of the display unit might be cause by a dim screen being more bothersome in the summer sun.

9 Conclusions & Recommendations

9.1. Conclusions

The negative consequences of unexpected unplanned maintenance create need for more insight in these type of occurrences. The availability of previously untapped source of information in the form of pilot complaints gave rise to the following research question:

What is the effect of pilot complaints on the predictability of component removals?

The results presented in chapter 7 provided a comprehensive set of answers. The hazard ratios resulting from the Proportional Hazards Model provide very strong evidence for a statistically significant effect the information from the pilot complaint has on the hazard of a component removal. This effect however, is measure from the baseline hazard. While the hazard in some cases increases more than sevenfold, one must also consider the absolute effect this has on the expected ”mortality”. The comparison with preemptive knowledge puts the large hazard ratios into perspective by showing its weakness with respect to the ”crystal ball”. Many reasons for less that desirable results stem from the non-proportionality of the used covariates. Despite the application of left truncation and right censoring has resulted in a greatly reduced time window, the effect of the covariates is still not constant in time. The comparison with the Kaplan-Meier Estimator demonstrate an underestimation of the hazard after one day, this is corroborated by the sensitivity analysis, showing higher hazard rations for shorter time periods. Luckily, solutions exists that counter the non-proportionality problem, these are discussed in section 9.2.1. Additionally, more extensive processing of the pilot complaints could result in covariates with more predictive power, as discussed in section 9.2.2. Whilst not comparable to preemptive knowledge of a removal, the effect the pilots complaints have on the predictability is not insignificant. With the improvements mentioned in section 9.2, the usage of pilot complaints could be find a practical application into airline operations, providing a very low cost enhancement of the predictability of component removal they so desperately need.

9.2. Recommendations

9.2.1. Time Varying Covariates

The biggest negative effect on the credibility of the results is the frequent non-proportionality of the covariates, as mentioned in section 8.2.1. This problem undermines even the largest hazard ratios. The measures taken to limit the time-variant effects described in sections 6.1 and 6.2 have proven insufficient. What can be observed through the comparison with the Kaplan-Meier Estimator is that most parts show a very big drop in survival probability after just one day, while experiencing a reduction of this effect in the subsequent days. In order to correctly model the time-invariant effect, the model should be adapted to allow for such behaviour. Therneau and Grambsch describe extensions the Proportional Hazards Model to include time varying covariates.[29]

9.2.2. NLP Techniques

More advancements are to be made in the Natural Language Processing part of this research. Multiple words with the same meaning could be grouped into a single covariate. This could include multiple ways of writing some words. The pilot complaints for the oxygen bottle are a good example of this. Oxygen bottle and o2-bottle are synonyms. Synonyms very often occur in the pilot complaints, both as variation in spelling as well as different words with the same meaning. Clustering of synonyms has a positive effect on the efficacy of a TF-IDF analysis [24]. Additionally, as the ATA chapter usually defines the subject of the pilot complaint, the TF-IDF analysis could be performed with the document frequency being based only on related complaints instead of the entire corpus of pilot complaints.
The word ’bottle’ might have different meaning when it refers to the fire suppressant bottles in the engines for example. Bottle in combination with oxygen might correctly identify the subject of the pilot complaint to be the oxygen. This bi-gram, or higher level n-grams, might prove to be better covariates than same of the words on their own. After having confirmed the subject of a message, one could look for adjectives describing the state of the subject at hand. Parsing algorithms like the ”RASP” system [3] could extract adjectives to yield better performing covariates. A good example is the word ”dim” when the pilot complaint is about the display unit. The quality of textual covariates is expected to increase greatly with the application of more Natural Language Processing techniques.

Bibliography

[1] Richard E Barlow and Frank Proschan. *Mathematical theory of reliability*, volume 17. Siam, 1996.

[2] Flight Standards Service Regulatory Support Division Aviation Data System Branch. Joint Aircraft System/Component Code Table And Definitions. Technical report, Federal Aviation Administration, 10 2008.

[3] Ted Briscoe, John Carroll, and Rebecca Watson. The second release of the rasp system. In *Proceedings of the COLING/ACL on Interactive presentation sessions*, pages 77–80. Association for Computational Linguistics, 2006.

[4] Mario Cleves, William Gould, William W Gould, Roberto Gutierrez, and Yulia Marchenko. *An introduction to survival analysis using Stata*. Stata press, 2008.

[5] David R Cox. Regression models and life-tables. *Journal of the Royal Statistical Society: Series B (Methodological)*, 34(2):187–202, 1972.

[6] David R Cox. *Analysis of survival data*. Routledge, 2018.

[7] David R Cox and E Joyce Snell. The choice of variables in observational studies. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 23(1):51–59, 1974.

[8] Martin J Crowder. *Statistical analysis of reliability data*. Routledge, 2017.

[9] Cameron Davidson-Pilon. Camdavidsonpilon/lifelines: v0.21.0, April 2019. URL `https://zenodo.org/record/2638135`.

[10] U Dinesh Kumar. New trends in aircraft reliability and maintenance measures. *Journal of Quality in Maintenance Engineering*, 5(4):287–295, 1999.

[11] Eva Duran Eppler, Adrian Luescher, and Margaret Deuchar. Evaluating the predictions of three syntactic frameworks for mixed determiner–noun constructions. *Corpus Linguistics and Linguistic Theory*, 13(1):27–63, 2017.

[12] Benjamin Epstein. The exponential distribution and its role in life testing. Technical report, Wayne State University Detroit MI, 1958.

[13] SJ Findlay and ND Harrison. Why aircraft fail. *Materials today*, 5(11):18–25, 2002.

[14] Major Greenwood et al. A report on the natural duration of cancer. *A Report on the Natural Duration of Cancer.*, 1926.

[15] Masato Hagiwara, Yasuhiro Ogawa, and Katsuhiko Toyama. Selection of effective contextual information for automatic synonym acquisition. In *Proceedings of the 21st International Conference on Computational Linguistics and the 44th annual meeting of the Association for Computational Linguistics*, pages 353–360. Association for Computational Linguistics, 2006.

[16] Zellig S Harris. Distributional structure. *Word*, 10(2-3):146–162, 1954.

[17] David W Hosmer Jr, Stanley Lemeshow, and Susanne May. *Applied survival analysis: regression modeling of time-to-event data*, volume 618. Wiley-Interscience, 2008.

[18] AKS Jardine, PM Anderson, and DS Mann. Application of the weibull proportional hazards model to aircraft and marine engine failure data. *Quality and reliability engineering international*, 3(2):77–82, 1987.

[19] Andrew KS Jardine, Daming Lin, and Dragan Banjevic. A review on machinery diagnostics and prognostics implementing condition-based maintenance. *Mechanical systems and signal processing*, 20(7):1483–1510, 2006.

[20] Karen Spärck Jones. A statistical interpretation of term specificity and its application in retrieval. *Journal of documentation*, 2004.

[21] John D Kalbfleisch and Ross L Prentice. *The statistical analysis of failure time data*, volume 360. John Wiley & Sons, 2011.

[22] Edward L Kaplan and Paul Meier. Nonparametric estimation from incomplete observations. *Journal of the American statistical association*, 53(282):457–481, 1958.

[23] John P Klein and Melvin L Moeschberger. *Survival analysis: techniques for censored and truncated data*. Springer Science & Business Media, 2006.

[24] Madhu Kumari, Akshat Jain, and Ankit Bhatia. Synonyms based term weighting scheme: an extension to tf. idf. *Procedia Computer Science*, 89:555–561, 2016.

[25] Mingyoug Liu and Jiangang Yang. An improvement of tfidf weighting in text categorization. *International proceedings of computer science and information technology*, pages 44–47, 2012.

[26] CE Love and R Guo. Application of weibull proportional hazards modelling to bad-as-old failure data. *Quality and reliability engineering international*, 7(3):149–157, 1991.

[27] Christopher D Manning, Christopher D Manning, and Hinrich Schütze. *Foundations of statistical natural language processing*. MIT press, 1999.

[28] Karen Sparck Jones. A statistical interpretation of term specificity and its application in retrieval. *Journal of documentation*, 28(1):11–21, 1972.

[29] Terry M Therneau and Patricia M Grambsch. *Modeling survival data: extending the Cox model*. Springer Science & Business Media, 2013.

[30] Tiedo Tinga. Application of physical failure models to enable usage and load based maintenance. *Reliability Engineering & System Safety*, 95(10):1061–1075, 2010.

[31] Wim JC Verhagen and Lennaert WM De Boer. Predictive maintenance for aircraft components using proportional hazard models. *Journal of Industrial Information Integration*, 12:23–30, 2018.

[32] Waloddi Weibull et al. A statistical distribution function of wide applicability. *Journal of applied mechanics*, 18(3):293–297, 1951.

Appendices

A

Proportional Hazards Model Fit Summaries: Words

Component 0: Oxygen Bottle

Table A.1: Summary of Proportional Hazards Model fit for Component 0: Oxygen Bottle, Word:o2-bottle.

	coef	exp(coef)	se(coef)	z	p	-log2(p)	lower 0.95	upper 0.95	pa
Fresh	-0.35	0.70	0.06	-5.73	<0.005	26.57	-0.47	-0.23	✓
Summer	-0.14	0.87	0.02	-8.45	<0.005	54.89	-0.17	-0.11	✓
Year	-0.00	1.00	0.01	-0.08	0.94	0.09	-0.01	0.01	✓
Word:o2-bottle	-0.77	0.46	0.21	-3.67	<0.005	12.04	-1.18	-0.36	✓

Table A.2: Summary of Proportional Hazards Model fit for Component 0: Oxygen Bottle, Word:1560.

	coef	exp(coef)	se(coef)	z	p	-log2(p)	lower 0.95	upper 0.95	pa
Fresh	-0.35	0.70	0.06	-5.76	<0.005	26.80	-0.47	-0.23	✓
Summer	-0.14	0.87	0.02	-8.40	<0.005	54.37	-0.17	-0.11	✓
Year	-0.00	1.00	0.01	-0.08	0.94	0.10	-0.01	0.01	✓
Word:1560	-0.33	0.72	0.13	-2.67	0.01	7.04	-0.58	-0.09	✗

Table A.3: Summary of Proportional Hazards Model fit for Component 0: Oxygen Bottle, Word:oxygen.

	coef	exp(coef)	se(coef)	z	p	-log2(p)	lower 0.95	upper 0.95	pa
Fresh	-0.35	0.70	0.06	-5.74	<0.005	26.69	-0.47	-0.23	✓
Summer	-0.14	0.87	0.02	-8.40	<0.005	54.31	-0.17	-0.11	✓
Year	-0.00	1.00	0.01	-0.07	0.95	0.08	-0.01	0.01	✓
Word:oxygen	-0.04	0.96	0.07	-0.57	0.57	0.81	-0.19	0.10	✗

Table A.4: Summary of Proportional Hazards Model fit for Component 0: Oxygen Bottle, Word:bottle.

	coef	exp(coef)	se(coef)	z	p	-log2(p)	lower 0.95	upper 0.95	pa
Fresh	-0.35	0.70	0.06	-5.78	<0.005	27.00	-0.47	-0.23	✓
Summer	-0.14	0.87	0.02	-8.51	<0.005	55.68	-0.18	-0.11	✓
Year	-0.00	1.00	0.01	-0.02	0.98	0.03	-0.01	0.01	✓
Word:bottle	-0.40	0.67	0.06	-6.30	<0.005	31.61	-0.53	-0.28	✗

Component 1: Flow Control Valve

Table A.5: Summary of Proportional Hazards Model fit for Component 1: Flow Control Valve, Word:pack.

	coef	exp(coef)	se(coef)	z	p	-log2(p)	lower 0.95	upper 0.95	pa
Fresh	-0.10	0.90	0.07	-1.38	0.17	2.58	-0.24	0.04	✓
Summer	0.10	1.11	0.06	1.76	0.08	3.66	-0.01	0.22	✓
Year	-0.27	0.77	0.02	-13.08	<0.005	127.48	-0.31	-0.23	✓
Word:pack	1.46	4.32	0.09	15.51	<0.005	177.70	1.28	1.65	✗

Table A.6: Summary of Proportional Hazards Model fit for Component 1: Flow Control Valve, Word:cabin.

	coef	exp(coef)	se(coef)	z	p	-log2(p)	lower 0.95	upper 0.95	pa
Fresh	-0.10	0.90	0.07	-1.41	0.16	2.65	-0.24	0.04	✓
Summer	0.11	1.11	0.06	1.84	0.07	3.92	-0.01	0.22	✓
Year	-0.27	0.76	0.02	-13.33	<0.005	132.22	-0.31	-0.23	✓
Word:cabin	0.37	1.45	0.11	3.43	<0.005	10.68	0.16	0.58	✗

Table A.7: Summary of Proportional Hazards Model fit for Component 1: Flow Control Valve, Word:temp.

	coef	exp(coef)	se(coef)	z	p	-log2(p)	lower 0.95	upper 0.95	pa
Fresh	-0.10	0.90	0.07	-1.43	0.15	2.70	-0.25	0.04	✓
Summer	0.10	1.11	0.06	1.79	0.07	3.76	-0.01	0.22	✓
Year	-0.27	0.76	0.02	-13.33	<0.005	132.15	-0.31	-0.23	✓
Word:temp	0.99	2.70	0.11	8.82	<0.005	59.59	0.77	1.21	✓

Table A.8: Summary of Proportional Hazards Model fit for Component 1: Flow Control Valve, Word:control.

	coef	exp(coef)	se(coef)	z	p	-log2(p)	lower 0.95	upper 0.95	pa
Fresh	-0.10	0.90	0.07	-1.43	0.15	2.71	-0.25	0.04	✓
Summer	0.11	1.12	0.06	1.87	0.06	4.02	-0.01	0.22	✓
Year	-0.27	0.76	0.02	-13.28	<0.005	131.31	-0.31	-0.23	✓
Word:control	1.04	2.84	0.10	10.93	<0.005	89.95	0.86	1.23	✗

Component 2: Display Unit

Table A.9: Summary of Proportional Hazards Model fit for Component 2: Display Unit, Word:radar.

	coef	exp(coef)	se(coef)	z	p	-log2(p)	lower 0.95	upper 0.95	pa
Fresh	0.70	2.01	0.06	10.80	<0.005	87.89	0.57	0.82	✓
Summer	0.39	1.47	0.06	6.24	<0.005	31.11	0.27	0.51	✓
Year	0.02	1.02	0.02	1.17	0.24	2.05	-0.01	0.06	✓
Word:radar	0.99	2.70	0.38	2.62	0.01	6.83	0.25	1.74	✓

Table A.10: Summary of Proportional Hazards Model fit for Component 2: Display Unit, Word:screen.

	coef	exp(coef)	se(coef)	z	p	-log2(p)	lower 0.95	upper 0.95	pa
Fresh	0.70	2.01	0.06	10.80	<0.005	87.96	0.57	0.82	✓
Summer	0.39	1.47	0.06	6.20	<0.005	30.68	0.26	0.51	✓
Year	0.02	1.02	0.02	1.14	0.25	1.97	-0.01	0.06	✓
Word:screen	1.36	3.90	0.22	6.19	<0.005	30.59	0.93	1.79	✓

Table A.11: Summary of Proportional Hazards Model fit for Component 2: Display Unit, Word:display.

	coef	exp(coef)	se(coef)	z	p	-log2(p)	lower 0.95	upper 0.95	pa
Fresh	0.69	2.00	0.06	10.71	<0.005	86.54	0.57	0.82	✓
Summer	0.39	1.48	0.06	6.32	<0.005	31.84	0.27	0.51	✓
Year	0.02	1.02	0.02	0.88	0.38	1.39	-0.02	0.05	✓
Word:display	1.95	7.04	0.12	16.66	<0.005	204.54	1.72	2.18	✗

Component 3: Pressure Regulating Shut-Off Valve

Table A.12: Summary of Proportional Hazards Model fit for Component 3: Component 3: Pressure Regulating Shut-Off Valve, Word:bleed.

	coef	exp(coef)	se(coef)	z	p	-log2(p)	lower 0.95	upper 0.95	pa
Fresh	0.29	1.34	0.08	3.68	<0.005	12.06	0.14	0.44	✓
Summer	-0.10	0.91	0.06	-1.59	0.11	3.16	-0.22	0.02	✓
Year	-0.04	0.97	0.02	-1.81	0.07	3.83	-0.07	0.00	✓
Word:bleed	1.72	5.60	0.09	18.49	<0.005	251.13	1.54	1.91	✗

Table A.13: Summary of Proportional Hazards Model fit for Component 3: Component 3: Pressure Regulating Shut-Off Valve, Word:valve.

	coef	exp(coef)	se(coef)	z	p	-log2(p)	lower 0.95	upper 0.95	pa
Fresh	0.30	1.35	0.08	3.80	<0.005	12.74	0.14	0.45	✓
Summer	-0.10	0.91	0.06	-1.59	0.11	3.16	-0.22	0.02	✓
Year	-0.03	0.97	0.02	-1.70	0.09	3.50	-0.07	0.01	✓
Word:valve	0.84	2.31	0.12	6.96	<0.005	38.11	0.60	1.07	✗

Table A.14: Summary of Proportional Hazards Model fit for Component 3: Component 3: Pressure Regulating Shut-Off Valve, Word:rh.

	coef	exp(coef)	se(coef)	z	p	-log2(p)	lower 0.95	upper 0.95	pa
Fresh	0.30	1.35	0.08	3.78	<0.005	12.65	0.14	0.45	✓
Summer	-0.10	0.91	0.06	-1.60	0.11	3.20	-0.22	0.02	✓
Year	-0.03	0.97	0.02	-1.79	0.07	3.76	-0.07	0.00	✓
Word:rh	0.25	1.29	0.09	2.79	0.01	7.56	0.08	0.43	✗

Component 4: Landing Light

Table A.15: Summary of Proportional Hazards Model fit for Component 4: Landing Light, Word:landing.

	coef	exp(coef)	se(coef)	z	p	-log2(p)	lower 0.95	upper 0.95	pa
Fresh	0.04	1.04	0.07	0.51	0.61	0.72	-0.11	0.18	✓
Summer	-0.06	0.94	0.06	-1.02	0.31	1.69	-0.17	0.05	✓
Year	-0.11	0.89	0.02	-6.03	<0.005	29.16	-0.15	-0.08	✓
Word:landing	1.05	2.87	0.09	11.12	<0.005	92.97	0.87	1.24	✓

Table A.16: Summary of Proportional Hazards Model fit for Component 4: Landing Light, Word:retract.

	coef	exp(coef)	se(coef)	z	p	-log2(p)	lower 0.95	upper 0.95	pa
Fresh	0.04	1.04	0.07	0.57	0.57	0.81	-0.10	0.18	✓
Summer	-0.06	0.94	0.06	-1.00	0.32	1.66	-0.17	0.05	✓
Year	-0.11	0.89	0.02	-6.07	<0.005	29.57	-0.15	-0.08	✓
Word:retract	0.98	2.66	0.22	4.45	<0.005	16.80	0.55	1.41	✓

Table A.17: Summary of Proportional Hazards Model fit for Component 4: Landing Light, Word:lh.

	coef	exp(coef)	se(coef)	z	p	-log2(p)	lower 0.95	upper 0.95	pa
Fresh	0.04	1.04	0.07	0.57	0.57	0.81	-0.10	0.18	✓
Summer	-0.06	0.95	0.06	-0.98	0.33	1.62	-0.17	0.06	✓
Year	-0.11	0.89	0.02	-6.09	<0.005	29.76	-0.15	-0.08	✓
Word:lh	0.13	1.14	0.09	1.42	0.15	2.69	-0.05	0.31	✓

B

Proportional Hazards Model Fit Summaries: Irrelevant Words

Component 0: Oxygen Bottle

Table B.1: Summary of Proportional Hazards Model fit for Component 0: Oxygen Bottle, with irrelevant words.

	coef	exp(coef)	se(coef)	z	p	-log2(p)	lower 0.95	upper 0.95	pa
Fresh	-0.35	0.70	0.06	-5.73	<0.005	26.59	-0.47	-0.23	✓
Summer	-0.14	0.87	0.02	-8.45	<0.005	54.90	-0.17	-0.11	✓
Year	-0.00	1.00	0.01	-0.08	0.93	0.10	-0.01	0.01	✓
Word:o2-bottle	-0.77	0.46	0.21	-3.70	<0.005	12.17	-1.18	-0.36	✓
Word:pack	-0.03	0.97	0.05	-0.54	0.59	0.77	-0.14	0.08	✓
Word:display	-0.07	0.93	0.09	-0.82	0.41	1.29	-0.24	0.10	✗
Word:rh	-0.02	0.98	0.03	-0.87	0.38	1.38	-0.08	0.03	✓
Word:landing	-0.03	0.97	0.05	-0.59	0.56	0.84	-0.12	0.06	✓

Component 1: Flow Control Valve

Table B.2: Summary of Proportional Hazards Model fit for Component 1: Flow Control Valve, with irrelevant words.

	coef	exp(coef)	se(coef)	z	p	-log2(p)	lower 0.95	upper 0.95	pa
Fresh	-0.10	0.90	0.07	-1.38	0.17	2.57	-0.24	0.04	✓
Summer	0.10	1.11	0.06	1.75	0.08	3.65	-0.01	0.22	✓
Year	-0.27	0.77	0.02	-13.08	<0.005	127.48	-0.31	-0.23	✓
Word:o2-bottle	-0.19	0.83	0.58	-0.33	0.74	0.43	-1.32	0.94	✓
Word:pack	1.46	4.32	0.09	15.48	<0.005	177.14	1.28	1.65	✗
Word:display	0.05	1.05	0.29	0.18	0.86	0.23	-0.52	0.62	✓
Word:rh	0.01	1.01	0.09	0.12	0.91	0.14	-0.17	0.19	✗
Word:landing	0.05	1.05	0.15	0.32	0.75	0.42	-0.25	0.34	✓

Component 2: Display Unit

Table B.3: Summary of Proportional Hazards Model fit for Component 2: Display Unit, with irrelevant words.

	coef	exp(coef)	se(coef)	z	p	-log2(p)	lower 0.95	upper 0.95	pa
Fresh	0.69	2.00	0.06	10.71	<0.005	86.51	0.57	0.82	✓
Summer	0.39	1.48	0.06	6.31	<0.005	31.74	0.27	0.51	✓
Year	0.01	1.02	0.02	0.82	0.41	1.28	-0.02	0.05	✓
Word:o2-bottle	-1.28	0.28	1.00	-1.28	0.20	2.31	-3.24	0.68	✓
Word:pack	-0.35	0.70	0.23	-1.56	0.12	3.08	-0.79	0.09	✓
Word:display	1.95	7.00	0.12	16.59	<0.005	202.85	1.72	2.18	✗
Word:rh	-0.01	0.99	0.10	-0.13	0.90	0.16	-0.20	0.17	✓
Word:landing	-0.34	0.71	0.18	-1.88	0.06	4.06	-0.69	0.01	✓

Component 3: Pressure Regulating Shut-Off Valve

Table B.4: Summary of Proportional Hazards Model fit for Component 3: Pressure Regulating Shut-Off Valve, with irrelevant words.

	coef	exp(coef)	se(coef)	z	p	-log2(p)	lower 0.95	upper 0.95	pa
Fresh	0.30	1.35	0.08	3.80	<0.005	12.77	0.15	0.45	✓
Summer	-0.10	0.90	0.06	-1.64	0.10	3.29	-0.22	0.02	✓
Year	-0.03	0.97	0.02	-1.68	0.09	3.43	-0.07	0.01	✓
Word:o2-bottle	-0.52	0.59	0.71	-0.74	0.46	1.13	-1.91	0.86	✓
Word:pack	0.61	1.85	0.15	4.18	<0.005	15.09	0.33	0.90	✓
Word:display	-0.45	0.63	0.38	-1.20	0.23	2.11	-1.20	0.29	✓
Word:rh	0.25	1.28	0.09	2.72	0.01	7.27	0.07	0.43	✓
Word:landing	-0.07	0.93	0.17	-0.42	0.68	0.57	-0.40	0.26	✓

Component 4: Landing Light

Table B.5: Summary of Proportional Hazards Model fit for Component 4: Landing Light, with irrelevant words.

	coef	exp(coef)	se(coef)	z	p	-log2(p)	lower 0.95	upper 0.95	pa
Fresh	0.04	1.04	0.07	0.51	0.61	0.71	-0.11	0.18	✓
Summer	-0.06	0.94	0.06	-1.00	0.32	1.66	-0.17	0.05	✓
Year	-0.11	0.89	0.02	-6.05	<0.005	29.36	-0.15	-0.08	✓
Word:o2-bottle	0.25	1.29	0.45	0.56	0.57	0.81	-0.63	1.13	✓
Word:pack	-0.33	0.72	0.21	-1.56	0.12	3.08	-0.74	0.08	✓
Word:display	-0.27	0.77	0.32	-0.84	0.40	1.32	-0.89	0.36	✓
Word:rh	0.08	1.09	0.09	0.94	0.35	1.52	-0.09	0.25	✓
Word:landing	1.04	2.83	0.10	10.83	<0.005	88.33	0.85	1.23	✗

C
Summary Table Explanation

- **coef:** This is the value of the coefficient b_i that together with the value of the covariate, $(x_i - \bar{x_i})$, makes up the log-partial hazard that can be found in equation 5.5.
- **exp(coef):** The same value as above, but without the logarithm. This is the value described as the hazard ratio.
- **se(coef):** Standard error, or standard deviation of the coefficient mentioned above.
- **z:** Z-score, defined as the amount of standard errors away from the mean. The z-score multiplied by the standard error gives the value of the coefficient.
- **p:** P-value. This value states the probability of the results under the premise that the null-hypothesis is true. In other words, the probability of the results by chance, without there being an effect.
- **-log2(p):** Re-written form of the p-value. More useful as p-values are often very small number.
- **lower 0.95:** Lower end of the 95% confidence interval of the value of the coefficient.
- **upper 0.95** Upper end of the 95% confidence interval of the value of the coefficient.
- **pa:** Proportionality assumption as described in section 5.3.2. The check mark indicates that the covariate respects the proportionality assumption as tested within the Lifelines package [9]. The x-mark corresponds to a violation of said assumption.