Amsterdam University of Applied Sciences

Data mining for aircraft maintenance repair and overhaul (MRO)

Maurice Pelt

Aviation Academy, Faculty of Technology
Amsterdam University of Applied Science

17 April 2019

CREATING TOMORROW

Amsterdam University of Applied Sciences

Contents

Amsterdam University of Applied Sciences

Aircraft Maintenance and Unpredictability

Source: blog.klm.com

MRO benchmarks: TAT, reliability, cost

Challenges:

- Large variation in maintenance duration (and TAT)
- Uncertainty in inspection findings / spare parts needed
- Components replaced (long) before end of life

Opportunity:

- Data growth and powerful algorithms

TAT: Short and reliable MRO lead times

Costs: Reduction of MRO idle time and overprocessing

Costs: Optimal use of components remaining life

Amsterdam University of Applied Sciences

Research project Data Mining in MRO

HvA initiated applied research project, 2016 - 2018

28 case studies at 10+ companies

RAAK MKB program funded by SIA, Ministry of Economic Affairs

SiA Nationaal Regieorgaan Praktijkgericht Onderzoek

Research objective:
How can SME MRO's use fragmented historical maintenance data to decrease maintenance costs and aircraft downtime?

ABS JETS · EXSYN Digitizing Aviation · Tec4Jets · JetSupport Amsterdam · JetNetherlands · NEDAERO components · KVE Composites Group · NLG Netherlands Aerospace Group · mro air · TU Delft Delft University of Technology · KLM Engineering & Maintenance · AYA Aircraft Services · Koninklijke Luchtmacht · CHC · FLYINGGROUP · Lufthansa Technik

Amsterdam University of Applied Sciences

Timeline of project: Data Mining in MRO

2013	Research Ultrasonic Verification of Composites
2014	RAAK project: Maintain your competitive edge (Lean)
2015	July: First ideas about a Data Mining project Writing of proposal, workshops with industry partners Consortium NAG, Nedaero and JetSupport; Partner Exsyn, Novulo, JetNetherlands, KVE, Flying Service, Tec4Jets (TUI) ABS Jets, CHC Helicopters; Others: Royal Netherlands Air Force, TU Delft October: Proposal submission to SIA
2016	March: proposal v2; approved in July 1st wave of case studies; workshops, round table More companies: Nayak, Lufthansa, KLM, Transavia, NS, Fokker, NLR Some initial partners left the project
2017	2nd wave of case studies, expansion to machine learning; workshops Conferences 2017-18: RAeS UK IET UK, AEGATS FR, SLF NL, ISATECH TR
2018	3rd wave of case studies, deployment; workshops Integration of project results
2019	March: Project closing; final report Conferences: maintenance research day NL, EASN GR, RAMS USA and others
2020 and beyond	New RAAK project Aviation knowledge hub MRO

Amsterdam University of Applied Sciences

The final report

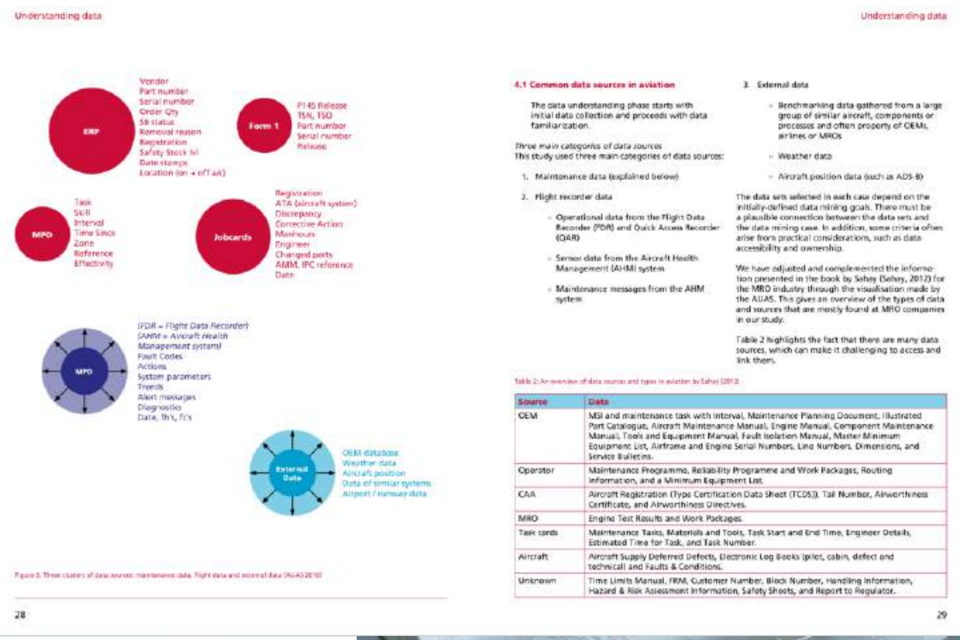


Figure 5. Three clusters of data sources: maintenance data, flight data and external data (AUAS 2018)

4.1 Common data sources in aviation

The data understanding phase starts with initial data collection and proceeds with data familiarization.

Three main categories of data sources
This study used three main categories of data sources:

1. Maintenance data (explained below)
2. Flight recorder data
   - Operational data from the Flight Data Recorder (FDR) and Quick Access Recorder (QAR)
   - Sensor data from the Aircraft Health Management (AHM) system
   - Maintenance messages from the AHM system
3. External data
   - Benchmarking data gathered from a large group of similar aircraft, components or processes and often property of OEMs, airlines or MROs
   - Weather data
   - Aircraft position data (such as ADS-B)

The data sets selected in each case depend on the initially-defined data mining goals. There must be a plausible connection between the data sets and the data mining case. In addition, some criteria often arise from practical considerations, such as data accessibility and ownership.

We have adjusted and complemented the information presented in the book by Sahay (Sahay, 2012) for the MRO industry through the visualisation made by the AUAS. This gives an overview of the types of data and sources that are mostly found at MRO companies in our study.

Table 2 highlights the fact that there are many data sources, which can make it challenging to access and link them.

Table 2: An overview of data sources and types in aviation by Sahay (2012)

Source	Data
OEM	MSI and maintenance task with interval, Maintenance Planning Document, Illustrated Part Catalogue, Aircraft Maintenance Manual, Engine Manual, Component Maintenance Manual, Tools and Equipment Manual, Fault Isolation Manual, Master Minimum Equipment List, Airframe and Engine Serial Numbers, Line Numbers, Dimensions, and Service Bulletins.
Operator	Maintenance Programme, Reliability Programme and Work Packages, Routing Information, and a Minimum Equipment List.
CAA	Aircraft Registration (Type Certification Data Sheet (TCDS)), Tail Number, Airworthiness Certificate, and Airworthiness Directives.
MRO	Engine Test Results and Work Packages.
Task cards	Maintenance Tasks, Materials and Tools, Task Start and End Time, Engineer Details, Estimated Time for Task, and Task Number.
Aircraft	Aircraft Supply Deferred Defects, Electronic Log Books (pilot, cabin, defect and technical) and Faults & Conditions.
Unknown	Time Limits Manual, FRM, Customer Number, Block Number, Handling Information, Hazard & Risk Assessment Information, Safety Sheets, and Report to Regulator.


4 UNDERSTANDING THE DATA

A variety of data sources
The MRO industry is characterized by a variety of data sources, from technical data recorded during a flight through a number of systems (e.g. FDR, QAR), to shop and maintenance data. External sources, such as airport, weather and ADB-S data, are also commonly used. At the same time, a number of technical and non-technical obstacles can present themselves while researchers are assembling reliable data sets. These obstacles can include sensor malfunctions, compatibility, and legal or contractual restrictions.

First insights
The data understanding phase starts with initial data collection and then proceeds with activities that help researchers become familiar with the data, identify data quality problems, discover first insights into the data, and detect interesting subsets to form hypotheses for hidden information. This task is performed in principal by visualizing the data and examining trends and patterns. Clearly, this process requires sufficient time as well as significant experience in the nature of data.

Connecting data to the business case
Once researchers have gained an understanding of the case, they then take a closer look into the data available for data mining. This is important because it connects the data to the business case to help retrieve the relevant parameters. This in turn requires an understanding of the business and the physical properties related to maintenance.

This data usually comes from existing, purchased and additional data – a variety of sources, in other words. The task of the researchers is collect this data, make relevant observations and identify variables. Then, they extract and describe relevant data as required. The resulting descriptive report should contain the amount of data, the value types and the coding schemes used. Using a univariate or bivariate method, the researchers can also explore the data and make premature conclusions for further data mining. This phase ends with a description of data quality, including missing data, data errors, measurement errors, coding inconsistencies and metadata mismatch.

Initiators / researchers / authors of Data Mining in MRO


Robert Jan de Boer


Jonno Broodbakker


Ruud Jansen


Konstantinos Stamoulis


Asteris Apostolides


Maurice Pelt

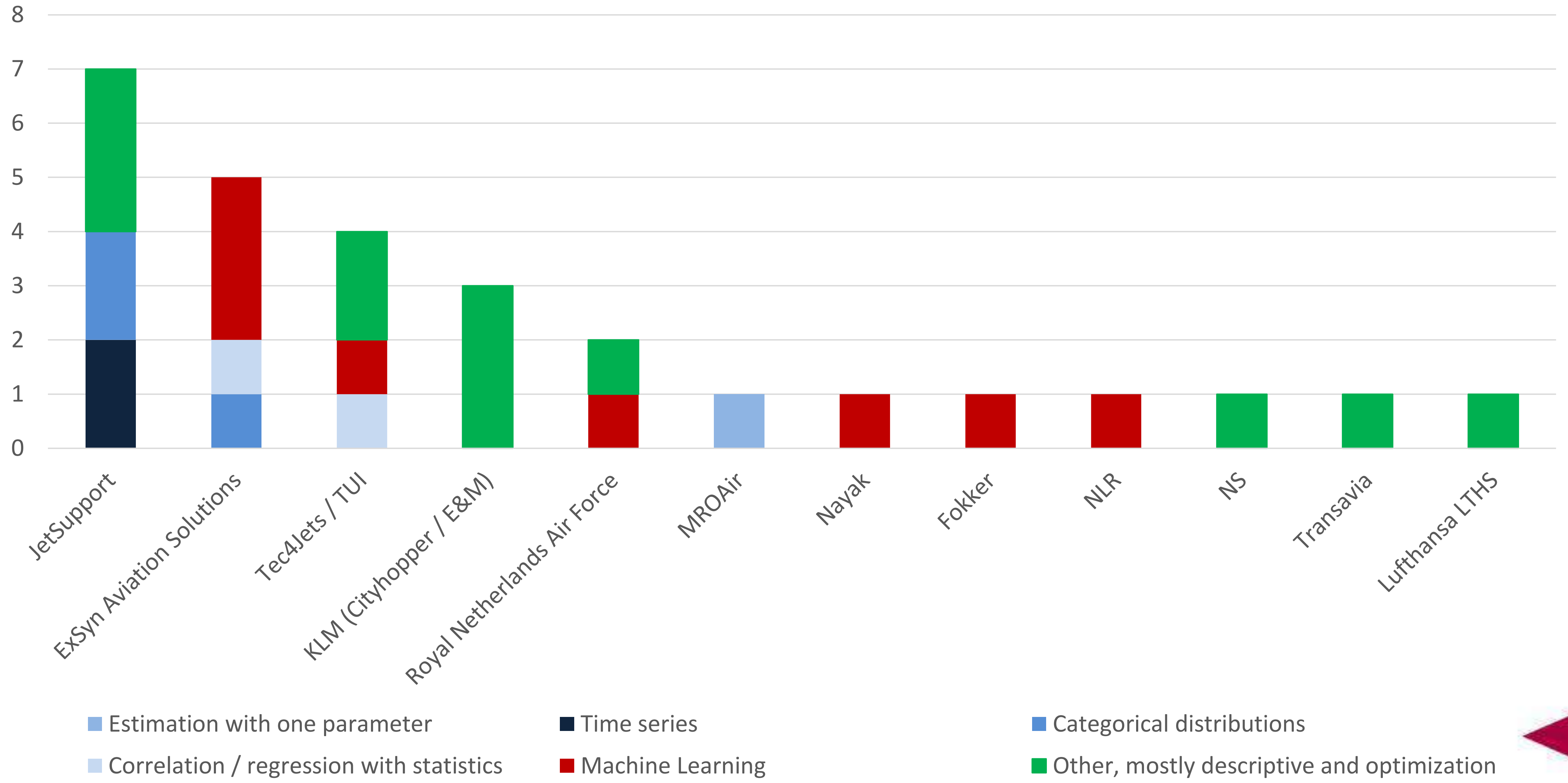

Maaik Borst


Roberto Felix Patron


Lorance Helwani

Timeline

Case studies per company and method; 28 in total



Amsterdam University of Applied Sciences

Case studies and student-researchers (1)

Group 1: Estimation with one parameter	
Raymond Molleman	Predicting findings on aviation maintenance task cards (MROAir, 2017).
Group 2: Time series	
Michael Killaars	Predictive maintenance in MRO with datamining techniques (JetSupport, 2017).
Rik Graas	Predicting maintenance durations using time series forecasting techniques (JetSupport, 2018).
Group 3: Categorical distributions	
Jerry Knuyt	Aircraft maintenance duration prediction using the most appropriate statistical distribution model (JetSupport, 2018).
André Koopman	Application of established reliability-based methods for predictive maintenance in a small to medium third-party maintenance organization (JetSupport, 2017).
Cheryl Zandvliet	Data mining in aviation: predictive component reliability (ExSyn Aviation Solutions, 2016).
Group 4: Correlation / regression with statistics	
Gerben de Jager	Potentie van datamining bij Tec4Jets (Tec4jets, 2018).
Bashir Amer	Engine Health Monitoring: Monitoring the heart of the aircraft (ExSyn Aviation Solutions, 2017).
Group 5: Machine Learning	
Jonno Broodbakker	Data mining applied to operational data from the Fokker 70 fleet of KLM Cityhopper (Nayak, 2016).
Sam van Brienen	Data potentials: Scheduling unplanned maintenance of legacy aircraft (ExSyn Aviation Solutions, 2018).
Arjan Francken	Aircraft component failure prediction using unsupervised data mining (ExSyn Aviation Solutions, 2018).
Manon Wientjes	Base maintenance findings risk predictor (ExSyn Aviation Solutions, 2018)
Laurens Scheipens	TUI’s aircraft reliability dashboard model (TUI, 2018).
Lorance Helwani	Machine learning and natural language processing in maintenance engineering (Fokker, 2018).
Ruud Jansen	Predicting aircraft speed and altitude profiles on departure (NLR, 2017 (not MRO related)).
Myrthe Dost	Causes of a reduced delivery reliability (RNAF, 2017).

Case studies and student-researchers (2)

Group 6: Other, mostly descriptive and optimization	
Martijn Bloothoofd	Manpower Planning of TUI Engineering and Maintenance (TUI, 2018).
Nino Mooren	Enhancing a predictive aircraft maintenance duration tool by improving the data fetching algorithm and the implementation of weather data (JetSupport, 2018).
Leon de Haan	Predictive maintenance in MRO calculation and analysis of Key Performance Indicator Manhours per Flight hour (Jetsupport, 2018).
Britt Bruyns	How A-checks can be improved (KLM Cityhopper, 2018).
Doris van der Meer	The first steps of the extension of the safety failure data analysis (Prorail, 2017).
Bob Laarman	Exploring expendables for repair development and cost reduction in an MRO environment (KLM, 2017).
Emiel van Maurik	Post production analysis (Transavia, 2017).
Thom van de Engel	Maintenance planning optimization (Tec4jets, 2017).
Ruby Weener	Quantification of the possible added value of the CFM56-7B's KLM customized workscope planning guide (KLM E&M, 2017).
Jeroen Verheugd	The potential of data mining techniques in avionics component maintenance (JetSupport, 2016).
Marc Hogerbrug & Julian Hiraki	Data mining in aviation maintenance, repair and overhaul (JetSupport, 2016).
Kylian Timmermans	Providing value added services from the digital shadow of MRO logistics providers (Lufthansa LTHS, 2016).
Bram Benda & Kaan Koc	Data mining in aviation: predictive component reliability (Koninklijke Luchtmacht, 2016.

Amsterdam University
of Applied Sciences

Contribution of this project to education

- Input for curriculum: Smart Maintenance modules
- Predictive Maintenance track part of the minor Data Science
- Contribution to Studios Predictive Maintenance en Data Science
- Input for Research Data Management function of Faculty of Technology

AVIATION ACADEMY

EDUCATION - RESEARCH - PARTNERSHIPS

Amsterdam University of Applied Sciences

We would like to thank SIA for funding this research project

SiA
Nationaal Regieorgaan Praktijkgericht Onderzoek

Amsterdam University of Applied Sciences

AVIATION ACADEMY
EDUCATION - RESEARCH - PARTNERSHIPS

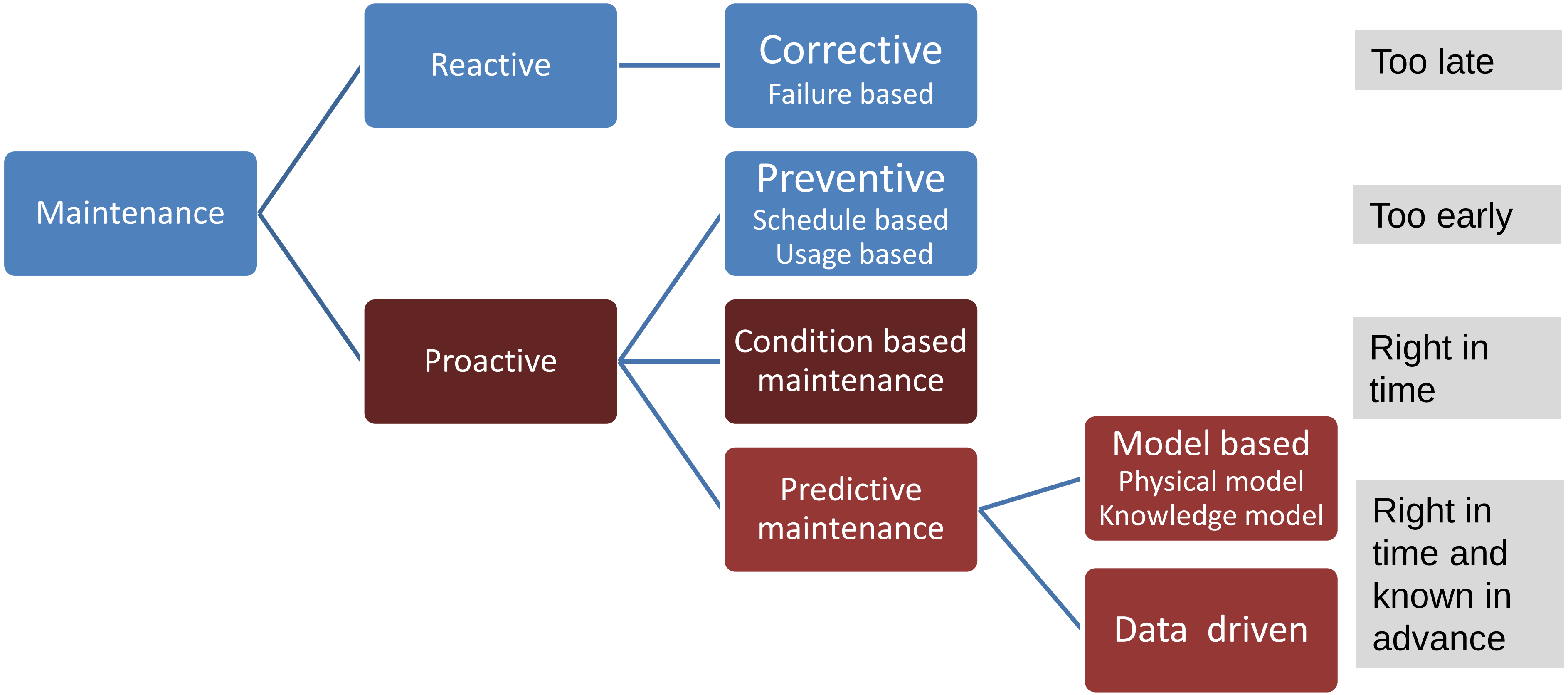
Amsterdam University of Applied Sciences

Contents

- RAAK project Data Mining in MRO
- Methodology
- Data sources and preparation
- Modelling
- Concluding remarks

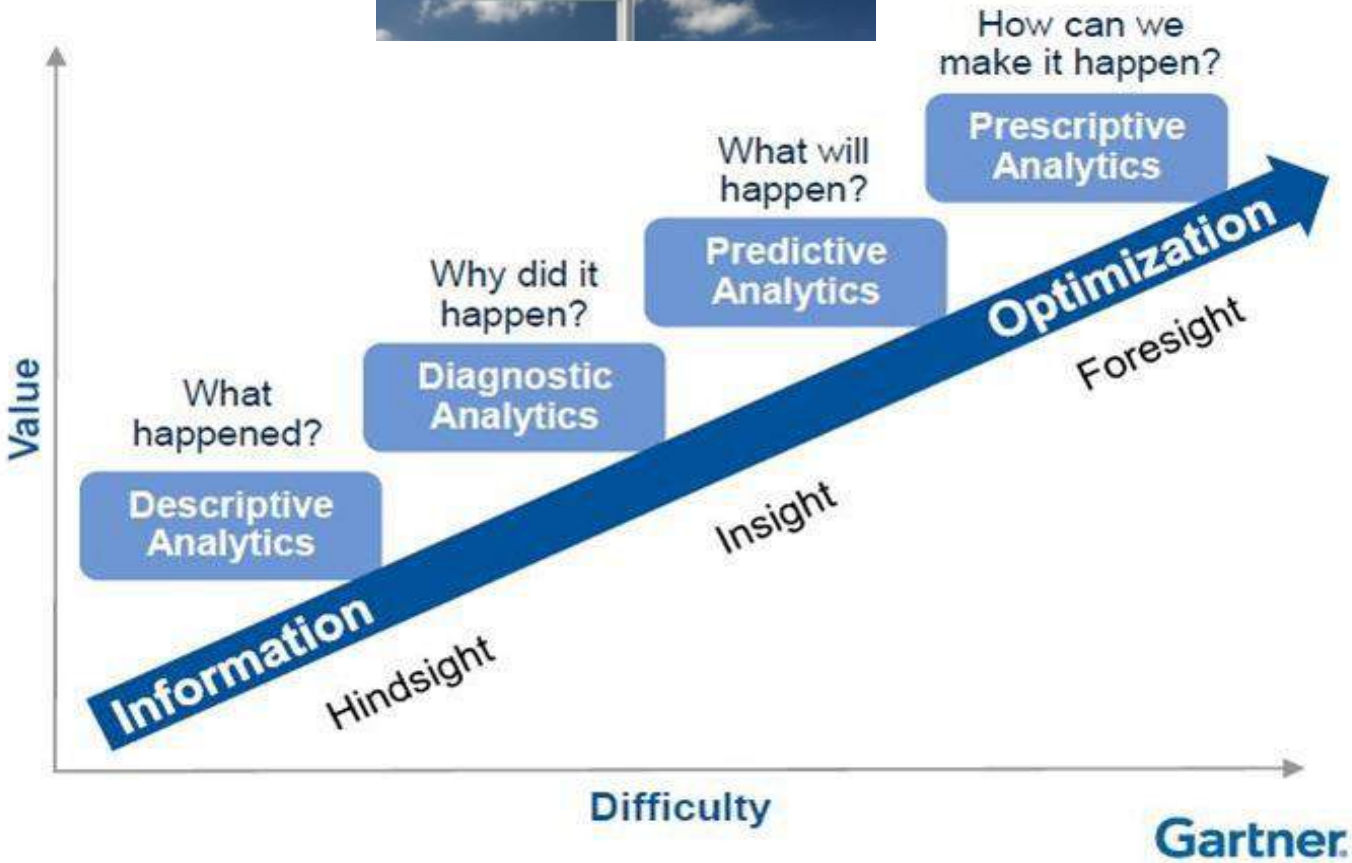
Amsterdam University
of Applied Sciences

Maintenance taxonomy



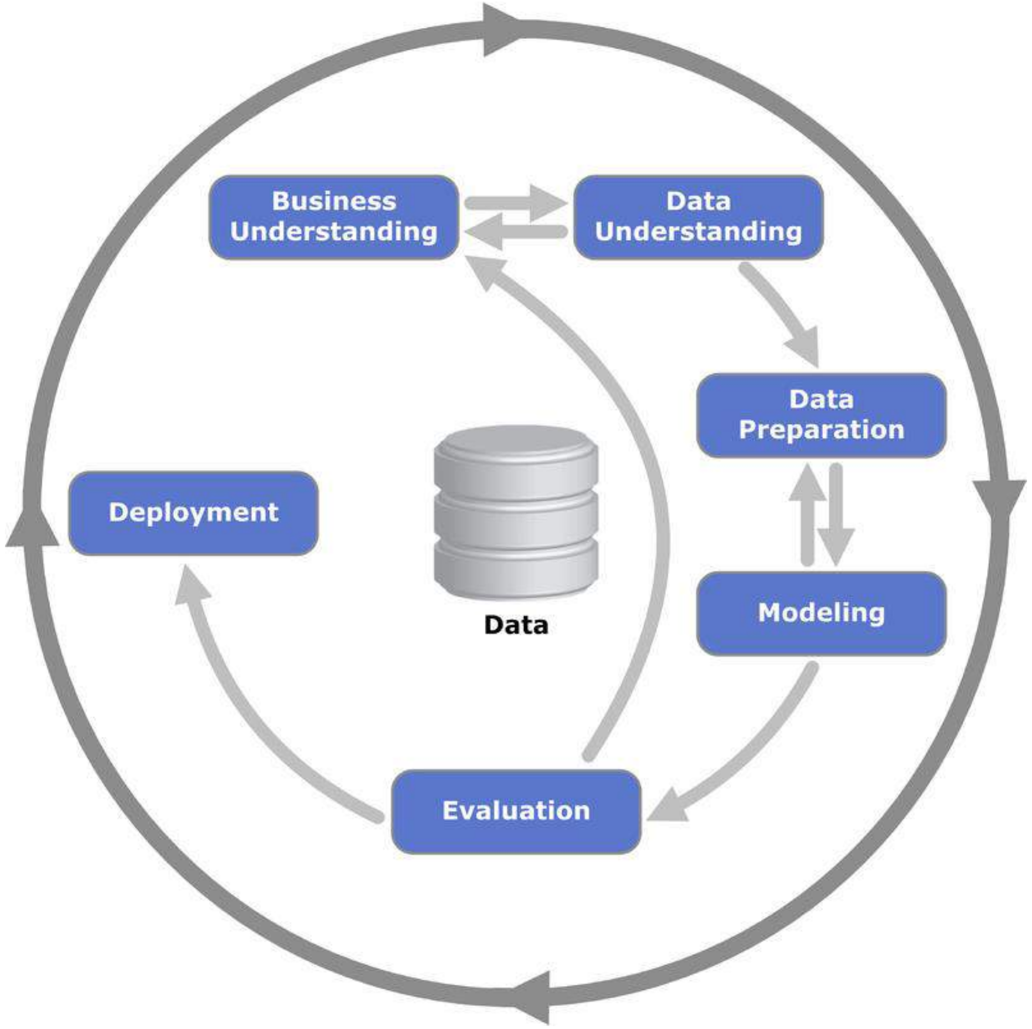
Amsterdam University of Applied Sciences

First describe and analyse the past, then predict the future and prescribe actions to be taken



CRISP-DM methodology for Data Mining in MRO

- Data mining: A sequence of steps
- Cross Industry Standard Process for Data Mining methodology: CRISP-DM
- Standard for data mining projects based on practical, real-world experience
- CRISP-DM is the most used data mining method (Piatetsky, 2014)



Source: Chapman, et al. (2000)

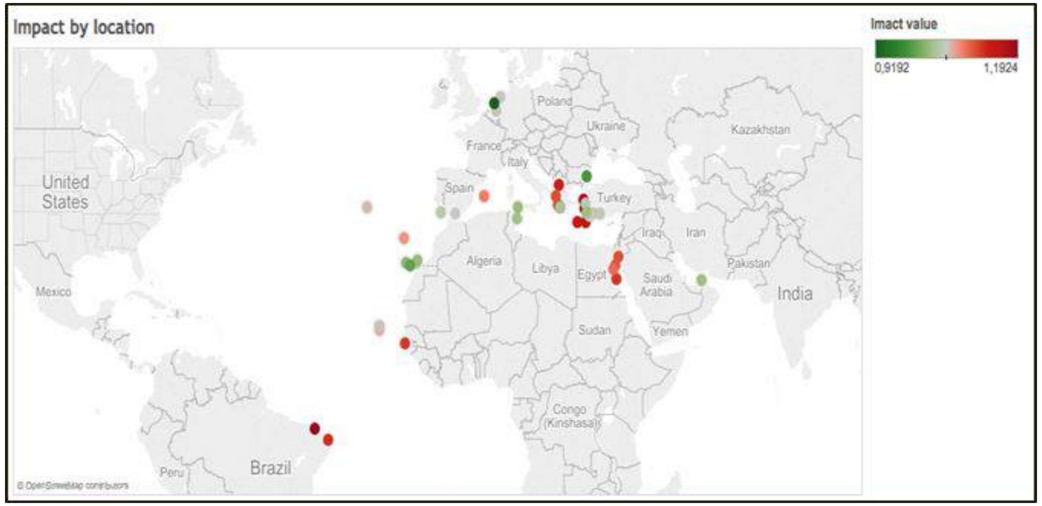
Amsterdam University of Applied Sciences

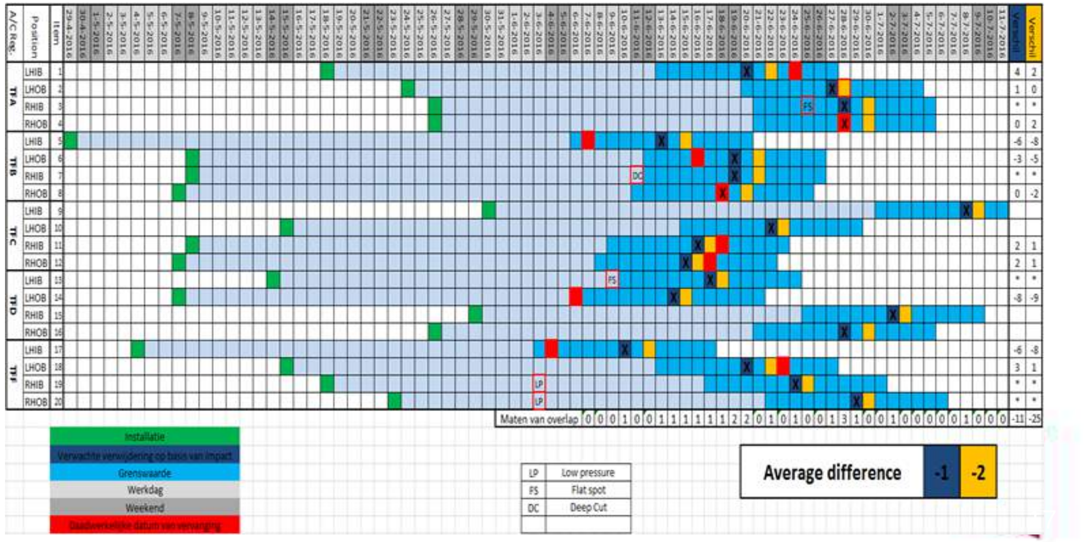
Case: Optimal aircraft tires replacement

Company: Line maintenance and A checks

→ Increase availability and lower maintenance costs

CRISP methodology	
Business understanding	Prediction of the remaining useful life time Optimal schedule for tire replacement
Data understanding	AMOS, FDM cycles, weight, braking action, location, runway length and temperature
Data preparation	Cleaning, integration into single dataset
Modelling	Linear regression
Evaluation Deployment	Highest correlation found: tire wear and airport Proof of concept: Prediction of optimal replacement moment





Contents

- RAAK project Data Mining in MRO
- Methodology
- Data sources and preparation
- Modelling
- Concluding remarks

Who has access to data and/or the rights to use?
Many formats, creators, users, owners of data were found in the case studies

Flight data

Maintenance data

External data

- Manuals, forms digital or on paper
- Structured tables in relational databases (e.g. ERP)
- Free text reports of findings and repair action
- External data sources in various formats
- Sensor data
- Pictures, samples

Available data

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

C: Creator
U: User
O: Owner

Example of a data distribution in Aviation MRO

Amsterdam University of Applied Sciences

Data preparation to clean and construct the final datasets from the initial raw data

- Deal with imperfect and incomplete data
- Clean, integrate, format and verify
- Often tedious, time consuming

Missing values

Outliers

Datasets not accessible, not available

Datasets incomplete

Data interpretation variability

Errors in values

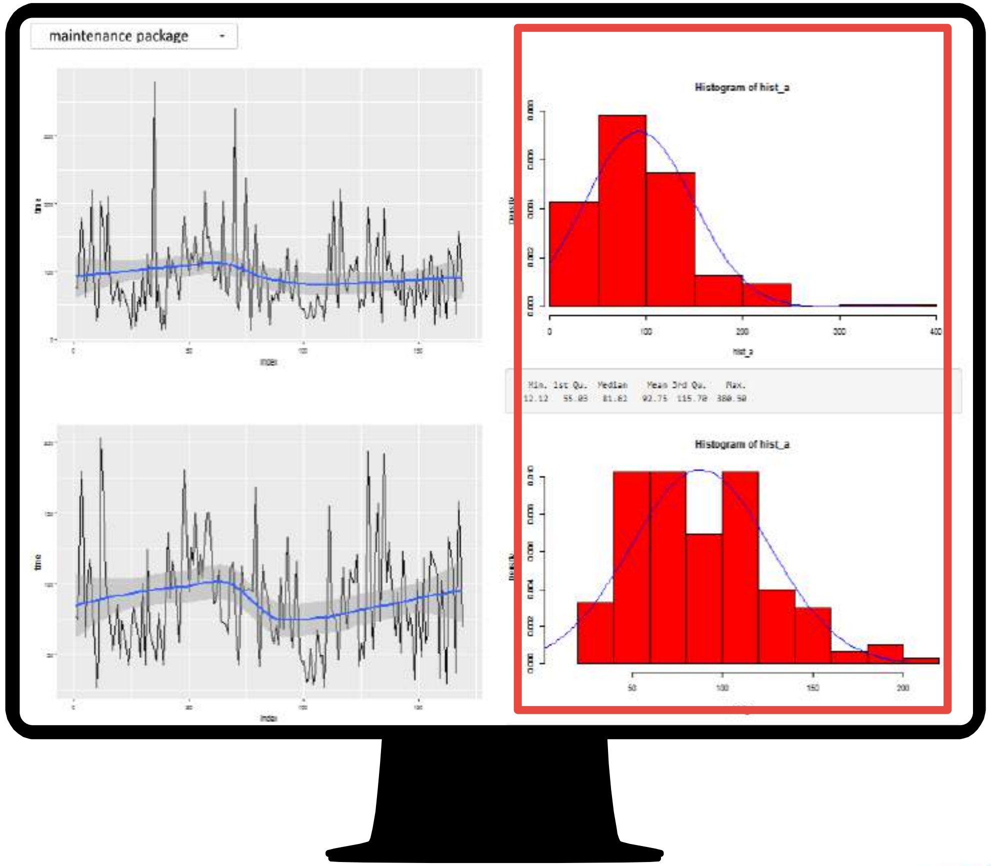
	Cleaning steps	Construct data	Integrate data	Transform data	Reduce data
Software developer	Remove duplicates; Remove false malfunctions	Yes	Yes	Yes	No
MRO company 1 a	Remove errors; Fill empty cells; Remove empty cells; Outliner removal; Remove irrelevant data	Yes	Yes	Yes	Yes
MRO company 1 b	Remove irrelevant data	Yes	Yes	Yes	No
MRO company 1 c	Correct errors; Fill empty cells; Remove empty cells	Yes	No	Yes	No
Airline MRO 2	-	Yes	No	Yes	Yes
MRO company 2	Correct errors; Fill empty cells; Outliner removal	Yes	Yes	Yes	No
In house MRO	Remove errors; Fill empty cells; Remove irrelevant data	Yes	Yes	Yes	No
MRO company 3	Remove errors; Fill empty cells; Remove empty cells	Yes	Yes	Yes	Yes

Case: Maintenance duration prediction

A predictive maintenance tool with reasonable accurate predicted maintenance tasks duration with automated selection of the:

1. Best fitting statistical distribution
2. Best performing time series forecasting model

For every maintenance package and/or job card of any aircraft type



Amsterdam University of Applied Sciences

Predictive Maintenance Tool dashboard

The probability of maintenance package 31051 requiring manpower for a duration between 26 and 84 minutes is 77.7 %

Amsterdam University of Applied Sciences

Contents

- RAAK project Data Mining in MRO
- Methodology
- Data sources and preparation
- Modelling
- Concluding remarks

The 28 case studies can be divided in 3 groups of data mining approaches

Visualization	• Descriptive analytics using established math and graphical methods, resulting in outputs such as KPI's control charts, management dashboards
Statistical data mining	• Descriptive and predictive analytics using established statistical methods, such as probability calculation, correlation and time series forcasting
Machine Learning	• Predictive analytics using machine learning methods such as regression, classification and clustering

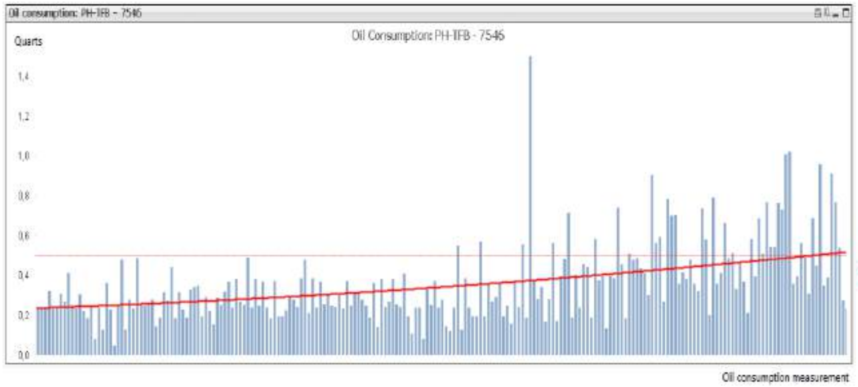
Amsterdam University of Applied Sciences

Case: Engine Health Monitoring with data that are available for Airlines

Inflight data from aircraft engines are sent to the manufacturer only

→ Improve maintenance efficiency using free available data

CRISP methodology	
Business understanding	Economic Replacement Point (ERP), Life Limiting Parts (LLP) and Exhaust Gas Temperature (EGT) define the optimal replacement time of engines
Data understanding	Available data: EGT, fuel consumption, oil pressure and oil consumption
Data preparation	Select engine type Clean and check data
Modelling	Develop Engine Health Monitoring model Forecast optimal engine replacement point
Evaluation Deployment	Aircraft uptime ↑, Part costs↓ EGT & LLP limits reached sooner than ERP



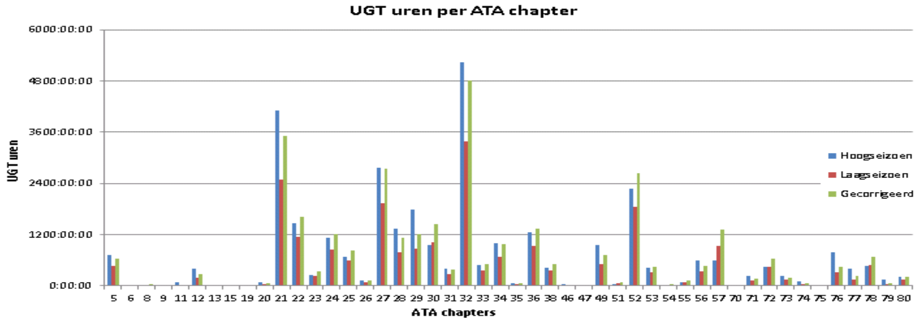
Amsterdam University
of Applied Sciences

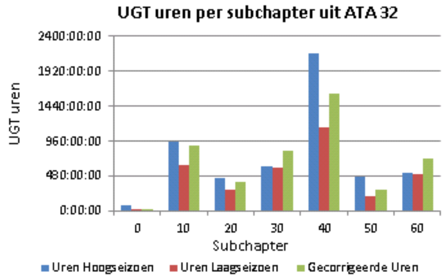
Case: Causes of low fleet availability in high season

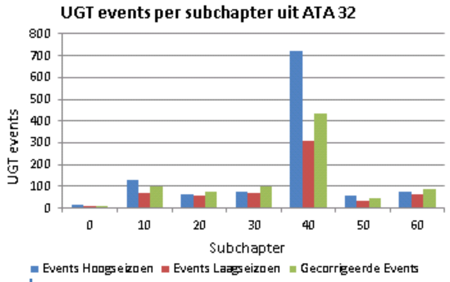
A/B-checks and line maintenance for Airline fleet

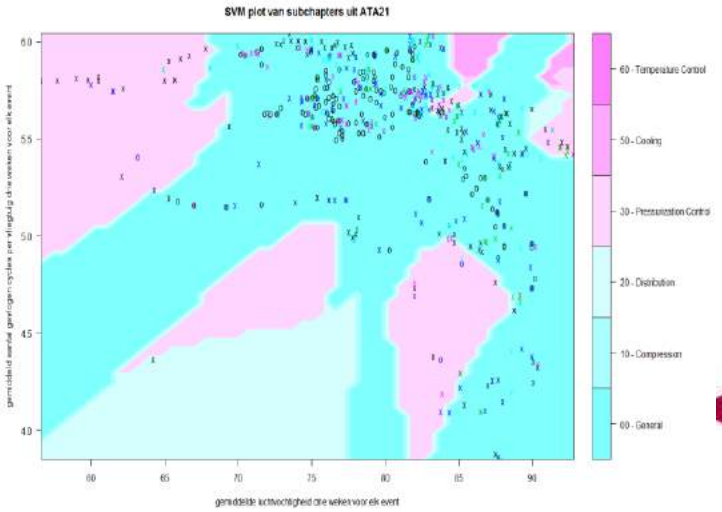
→ Causes of drop in Fleet Availability during high season

CRISP methodology	
Business understanding	Performance contract: aircraft uptime Correlate ATA (sub)chapter to problems
Data understanding	AMOS, weather data, flight data, unscheduled ground time events
Data preparation	Cleaned and integrated
Modelling	Descriptive analysis: highest unplanned ground time Support Vector Machine to predict problems related to weather
Evaluation Deployment	Aircraft uptime ↑, part costs↓ Performance drop correlated to ATA subchapter, e.g. tyres, brakes and cabin air quality

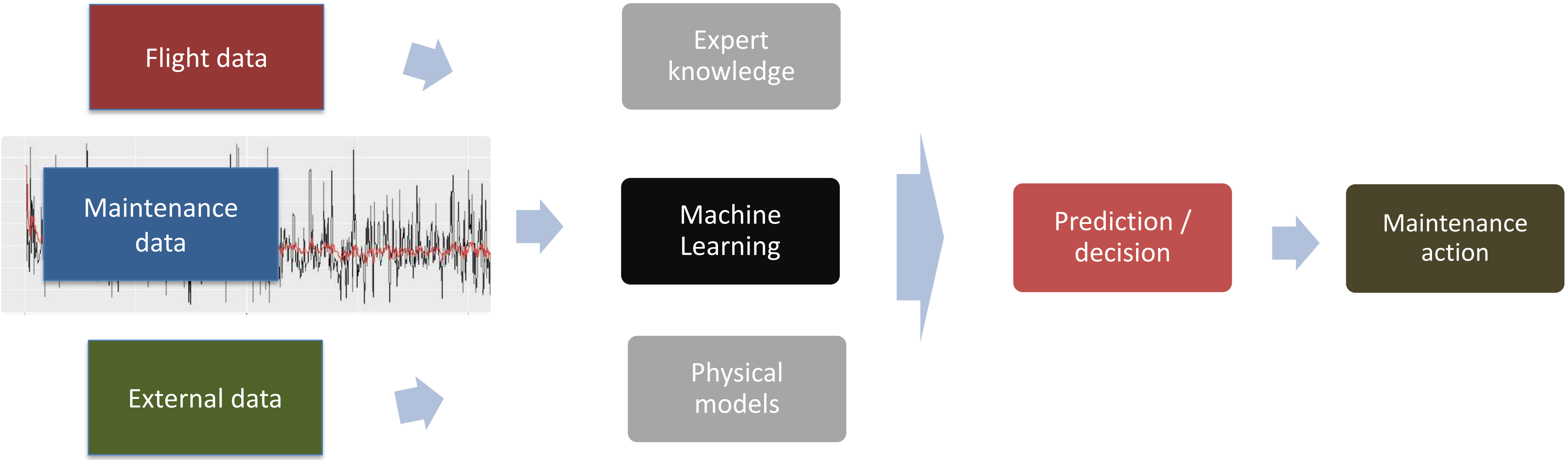








In this research other data sources and machine learning were added to overcome the prediction limitations of statistics on MRO datasets



Machine learning methods process many parameters and data types
Determine the parameters that strongly influence the output
Include the data of healthy systems

Amsterdam University of Applied Sciences

Case: Text mining to analyze maintenance reports

Use historical work order summary reports to trigger alerts if a failure or repair occurs more often than usual

Show similar failures or repairs from the past to support investigations

CRISP methodology	
Business understanding	Improve TAT and reduce maintenance costs if failures and solutions are known in an earlier stage
Data understanding	AMOS database: Work order summary reports and additional aircraft data
Data preparation	Retrieved and checked
Modelling	Chi Squared Distance Function and K-Nearest Neighbours method to classify report text Present results in Reliability Dashboard
Evaluation Deployment	Accuracy score 75,5%. With human control (reinforcement): 77,5%

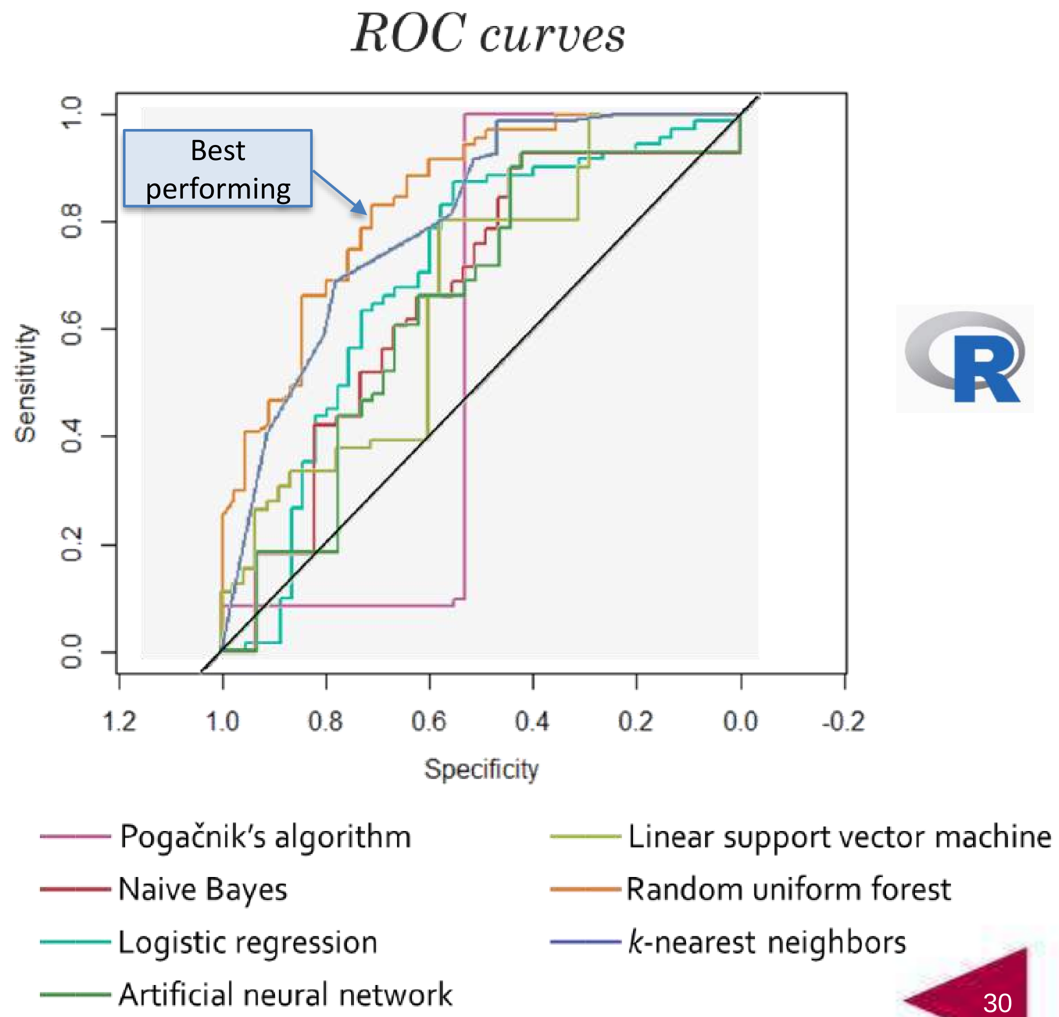
1. DATA INPUT PROCESS → **2. DATA PRE-PROCESSING MODEL** → **3. AIRCRAFT RELIABILITY DASHBOARD**

1. DATA INPUT PROCESS	2. DATA PRE-PROCESSING MODEL	3. AIRCRAFT RELIABILITY DASHBOARD
DATA MINING	*DATA MINING* *TEXT MINING*	*STATISTICAL PROCESS CONTROL*

python

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

CRISP methodology	
Business understanding	Predict unplanned maintenance tasks i.e. failures of components
Data understanding	Maintenance database: 600.000 task instances Parameters: task type, operator, aircraft type, age, flight hours, cycles, engine type, location, finding
Data preparation	Select test task type with 120 instances and 50/50 chance of failure
Modelling	Compare prediction accuracy of 7 machine learning algorithms Optimize parameters
Evaluation Deployment	Prediction accuracy too low for this task Additional data needed from e.g. weather, sensors, data sharing or synthetic data



Software applied in Data Mining in MRO

Open source software

Large user community, need to employ a data scientist

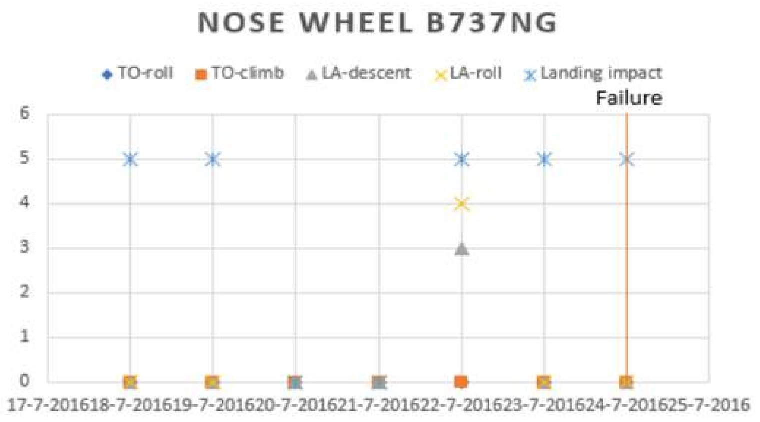
- R
- Python

Commercial software

- Matlab
- IBM - SPSS
- Tableau
- Microsoft - Azure
- Exsyn: Aviation Analytics

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

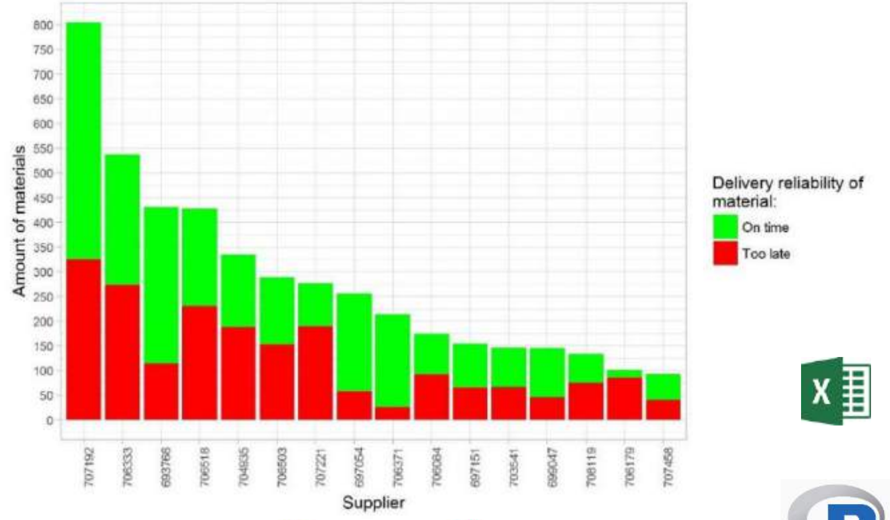
CRISP methodology	
Business understanding	Predict failures of components that possibly relate to flight path and/or weather anomalies
Data understanding	Maintenance data, ADS-B data (Flightradar24), weather data (NCEI)
Data preparation	Select from maintenance data a test component: Nose wheel Calculate acceleration forces from ADS-B data
Modelling	Dimensionality reduction Apply K-means and DBSCAN machine learning techniques to detect flight anomalies Correlate flight anomalies and nose wheel failures
Evaluation Deployment	Proof of concept: (Weak) correlation found

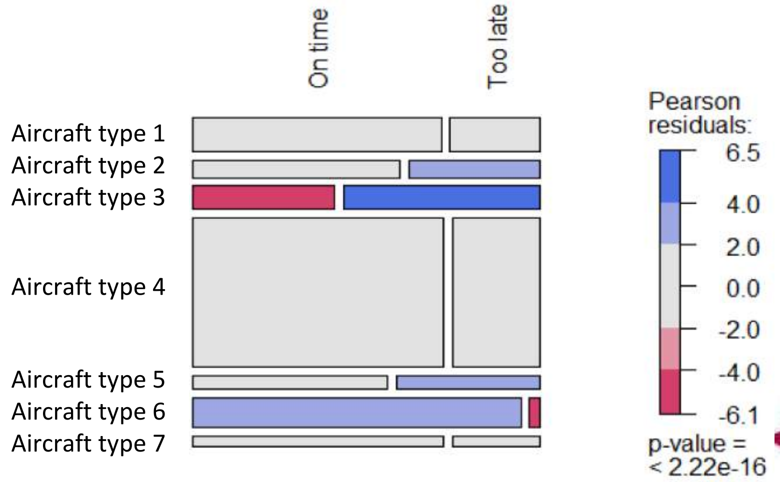


Amsterdam University of Applied Sciences

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

CRISP methodology	
Business understanding	Explain the causes of the low delivery reliability of component maintenance (between 49% and 97%)
Data understanding	Maintenance database, parameters: Delivery reliability, group, priority, maintenance type, order type, work centers, supplier and materials, execution status, actual costs, added value, planned and actual worked hours, planned and actual TAT
Data preparation	Retrieved and checked on year of data from SAP maintenance management system
Modelling	Examined the relationship between delivery reliability and 13 selected parameters. Data visualization e.g. mosaic plot. Statistics e.g. chi-squared. Machine learning (Decision tree) to predict delivery performance of parts.
Evaluation Deployment	Pilot project proved to successful. Main causes identified.





Contents

- RAAK project Data Mining in MRO
- Methodology
- Data sources and preparation
- Modelling
- Concluding remarks

Amsterdam University of Applied Sciences

Conclusions

Research objective: How can SME MRO's use fragmented historical maintenance data to decrease maintenance costs and aircraft downtime?

- Case studies proved the value of statistical and machine learning methods (proof of concept)
  - Aircraft uptime: optimal and accurate planning
  - MRO costs: efficiency, part costs
- CRISP-DM methodology useful
- Confidentiality and data ownership issues
- Visualization already proved to be very useful for companies
- Databases designed for compliance not analysis
- Data preparation much work
- Selection of appropriate algorithms need expert knowledge

Amsterdam University of Applied Sciences

Recommendations from the Data Mining in MRO research

Strategy

- Include data mining in the company's strategy
- Assess the current maturity level in data mining
- Start with focused applications that target real problems
- Set data mining performance goals

People

- Introduce data scientists
- Minimize the risk of unlawful or unwanted data sharing
- Provide on-the-job information to mechanics
- Organize close interaction between (academic) data scientists and shop floor mechanics

Process

- First visualization, then diagnostics, then prediction
- Combine data driven-models with expert and failure models
- Negotiate with OEMs and asset owners about access to data

ICT

- Increase data volume with (automated) maintenance reporting and sensors
- Modernize ICT to support data driven approach
- Investigate methods that deal with small datasets and open source data

Amsterdam University of Applied Sciences

Tip: Predictive Maintenance / Data Mining in MRO binnenkort te beluisteren in een podcast van de Dataloog
(Jurjen Helmus en Lex Knape interviewen Sander de Bree (Exsyn) en Maurice Pelt)

www.dedataloog.nl

Amsterdam University
of Applied Sciences

Thank you for your attention

Maurice Pelt
m.m.j.m.pelt@hva.nl