Research Stay: Predictive Maintenance of Industrial Equipment

Juan Echeagaray

Tec de Monterrey

September 11, 2023

2023-09-11

Research Stay: Predictive Maintenance of Industrial Equipment

Research Stay: Predictive Maintenance of Industrial Equipment

Juan Echeagaray

Tec de Monterrey

September 11, 2023

Agenda

Agenda
1. Introduction
2. Problem Statement
3. Objectives
4. Scope
5. Hypothesis
6. Exploratory Data Analysis
7. Methodology
8. Experimental Results
9. Conclusions
10. References

Maintenance scheduled by data analytics on historical data

- Enhance operational efficiency
- Sustainability and cost reduction
- Competitive advantage
- **Safety at the forefront**

Figure: Turbine failure mid-flight [1]

2023-09-11

Research Stay: Predictive Maintenance of Industrial Equipment
└Introduction
└Introduction

Introduction

Maintenance scheduled by data analytics on historical data

- Enhance operational efficiency
- Sustainability and cost reduction
- Competitive advantage
- **Safety at the forefront**

Figure: Turbine failure mid-flight [1]

Problem Statement

Predictive maintenance faces 2 main challenges:

Implementation

Need a model to estimate the RUL of a set of machinery, in an efficient and reliable manner

Interpretability

Need to *understand* why a model produces a given RUL estimate, becomes more important in critical environments

Problem Statement

Predictive maintenance faces 2 main challenges:

Implementation

Need a model to estimate the RUL of a set of machinery, in an efficient and reliable manner

Interpretability

Need to *understand* why a model produces a given RUL estimate, becomes more important in critical environments

Objectives

This research project aims to develop:

- PdM framework to predict RUL for a designated fleet of machinery
- Model which ensures reproducibility, stability, robustness and confidence Modelo estable, robusto, reproducible y confiable
- Tools to interpret and visualize the model's predictions

└─Objectives

Objectives

This research project aims to develop:

- PdM framework to predict RUL for a designated fleet of machinery
- Model which ensures reproducibility, stability, robustness and confidence Modelo estable, robusto, reproducible y confiable
- Tools to interpret and visualize the model's predictions

Scope

The previous objectives are to be accomplished subject to the following constraints and assumptions:

- RUL prediction of an uniform fleet of machines
- Availability of a labeled dataset with run to failure sequences of each machine

Scope

The previous objectives are to be accomplished subject to the following constraints and assumptions:

- RUL prediction of an uniform fleet of machines
- Availability of a labeled dataset with run to failure sequences of each machine

Throughout this research stay it is postulated:

Hypothesis

Using a set of sensor readings from a uniform fleet of machines in conjunction with environmental descriptors, it is possible to train an integrated predictive model for the Remaining Useful Life of an equipment, which is stable, robust, reproducible and reliable

Hypothesis

Throughout this research stay it is postulated:

Hypothesis

Using a set of sensor readings from a uniform fleet of machines in conjunction with environmental descriptors, it is possible to train an integrated predictive model for the Remaining Useful Life of an equipment, which is stable, robust, reproducible and reliable

Exploratory Data Analysis

NCMAPSS Dataset

Flight conditions and readings from a fleet of turbofan engines, derived from NASA's CMAPSS model, including real flight conditions and relates the degradation process to the operating history of the machine. [2]

- Used in the PHMAP 2021 Data Challenge [3]
- State of the art prognosis dataset (akin to MNIST and CIFAR for CV)

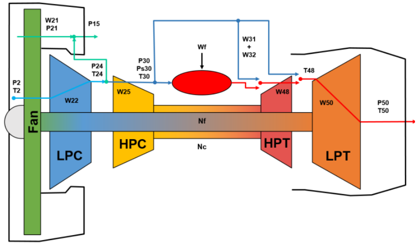
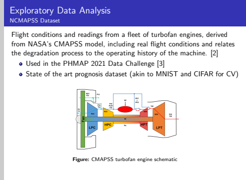


Figure: CMAPSS turbofan engine schematic

2023-09-11

Research Stay: Predictive Maintenance of Industrial Equipment
└Exploratory Data Analysis
└Exploratory Data Analysis

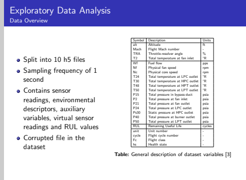


- Split into 10 h5 files
- Sampling frequency of 1 second
- Contains sensor readings, environmental descriptors, auxiliary variables, virtual sensor readings and RUL values
- Corrupted file in the dataset

Symbol	Description	Units
alt	Altitude	ft
Mach	Flight Mach number	-
TRA	Throttle-resolver angle	%
T2	Total temperature at fan inlet	°R
Wf	Fuel flow	pps
Nf	Physical fan speed	rpm
Nc	Physical core speed	rpm
T24	Total temperature at LPC outlet	°R
T30	Total temperature at HPC outlet	°R
T48	Total temperature at HPT outlet	°R
T50	Total temperature at LPT outlet	°R
P15	Total pressure in bypass-duct	psia
P2	Total pressure at fan inlet	psia
P21	Total pressure at fan outlet	psia
P24	Total pressure at LPC outlet	psia
Ps30	Static pressure at HPC outlet	psia
P40	Total pressure at burner outlet	psia
P50	Total pressure at LPT outlet	psia
RUL	Remaining Useful Life	cycles
unit	Unit number	-
cycle	Flight cycle number	-
Fc	Flight class	-
hs	Health state	-

Table: General description of dataset variables [3]

└Exploratory Data Analysis



Exploratory Data Analysis

Flight Class Distribution

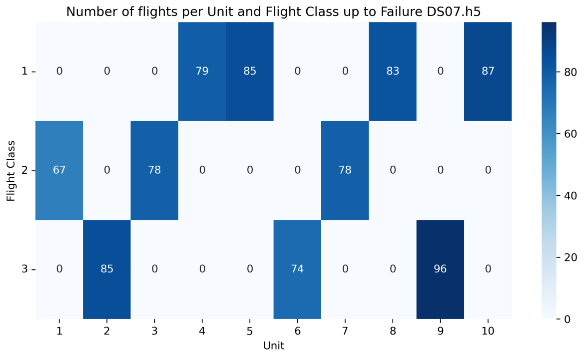
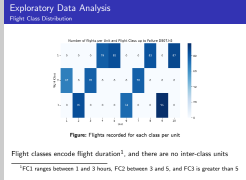


Figure: Flights recorded for each class per unit

Flight classes encode flight duration[1], and there are no inter-class units

[1] FC1 ranges between 1 and 3 hours, FC2 between 3 and 5, and FC3 is greater than 5



Exploratory Data Analysis

Sample Operating Conditions

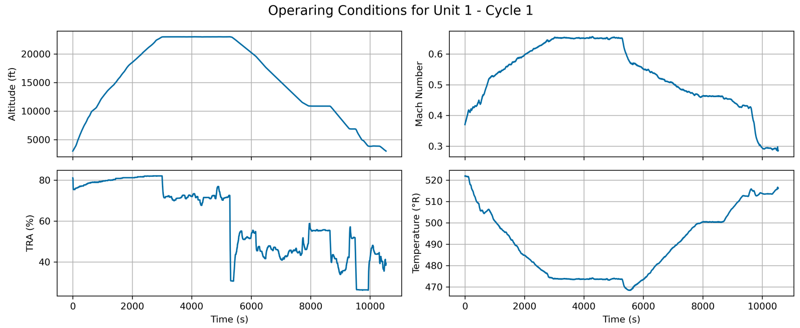
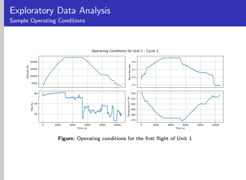


Figure: Operating conditions for the first flight of Unit 1

2023-09-11

Research Stay: Predictive Maintenance of Industrial Equipment
└Exploratory Data Analysis
└Exploratory Data Analysis



Exploratory Data Analysis

Sample Operating Conditions

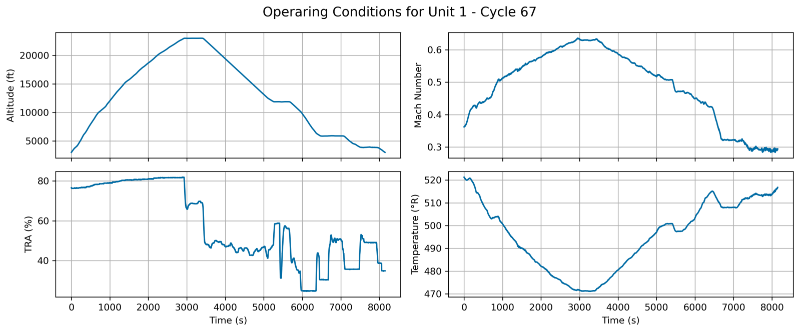


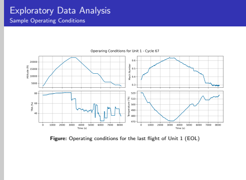
Figure: Operating conditions for the last flight of Unit 1 (EOL)

2023-09-11

Research Stay: Predictive Maintenance of Industrial Equipment

└Exploratory Data Analysis

└Exploratory Data Analysis



Exploratory Data Analysis

Operating Conditions per Flight Class

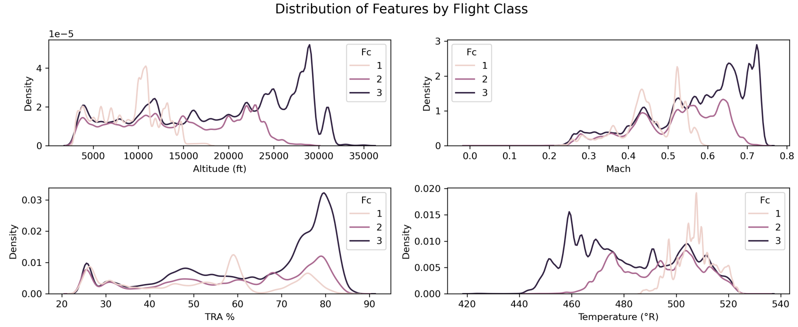
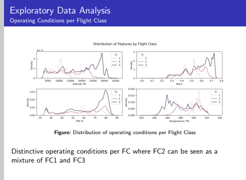


Figure: Distribution of operating conditions per Flight Class

Distinctive operating conditions per FC where FC2 can be seen as a mixture of FC1 and FC3



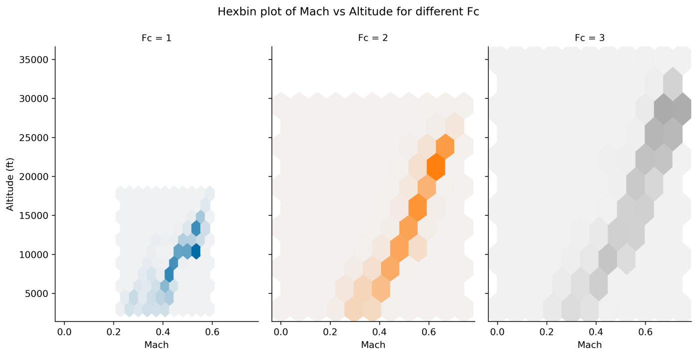


Figure: General relationship between Altitude and Mach Number

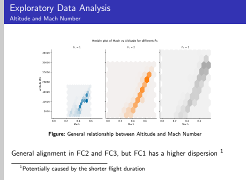
General alignment in FC2 and FC3, but FC1 has a higher dispersion [1]

[1] Potentially caused by the shorter flight duration

2023-09-11

Research Stay: Predictive Maintenance of Industrial Equipment
└Exploratory Data Analysis

└Exploratory Data Analysis



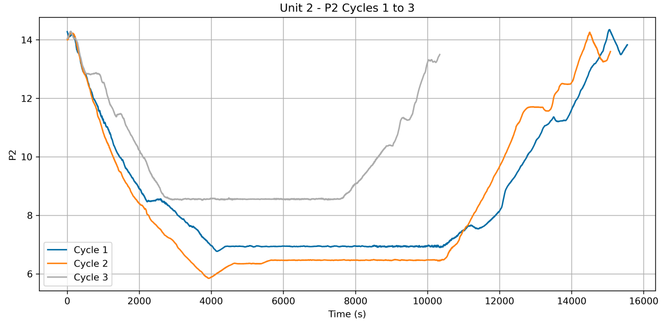
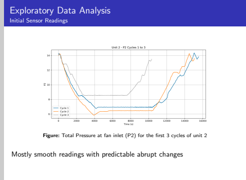


Figure: Total Pressure at fan inlet (P2) for the first 3 cycles of unit 2

Mostly smooth readings with predictable abrupt changes



Exploratory Data Analysis

Final Sensor Readings

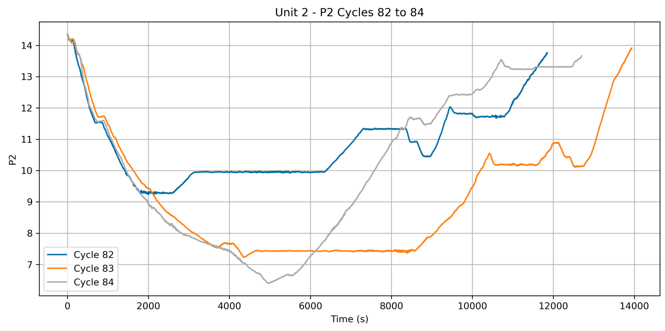
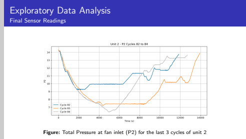


Figure: Total Pressure at fan inlet (P2) for the last 3 cycles of unit 2

Smoothness degradation with a presence of new plateaus near the end of the flight



Smoothness degradation with a presence of new plateaus near the end of the flight

Time Series Segmentation

Problem

Dataset consists of a set of multivariate time series of varying length with a distinct number of events per series

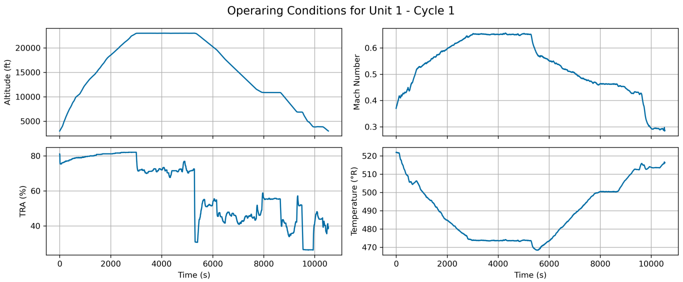
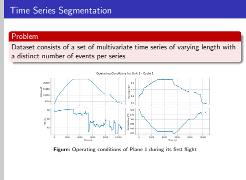


Figure: Operating conditions of Plane 1 during its first flight



1. Altitude almost always seems to have 3 states, ascent, plateau, descent
2. Mechanical operations vary throughout the flight, there's no smooth increase or decrease, rather multiple operating stages
3. Take advantage of the syncrhonization between time series, look at plateau

Pros & Cons

Pros

- Large data reduction **50x**
- Lower computational load
- Ease feature interpretation
- Reproducible

Cons

- Mixed or uniform MVT segmentation?
- Not a fixed number of stages in each flight
- Which descriptors to use?

- The negatives can be countered by optimizing a model conditioned on the segmentation parameters

Pros & Cons
Pros
- Large data reduction 50x
- Lower computational load
- Ease feature interpretation
- Reproducible
Cons
- Mixed or uniform MVT segmentation?
- Not a fixed number of stages in each flight
- Which descriptors to use?
- The negatives can be countered by optimizing a model conditioned on the segmentation parameters

Time Series Segmentation Basics

For a given signal $Y = \{y_t\}_{t=1}^{t=T}$ of T samples, where $y_t \in \mathbb{R}^d$, we assume there exists a set $\mathcal{T} = \{t_1^*, t_2^*, \ldots, t_n^*\}$ coding the $n-1$ stages of Y

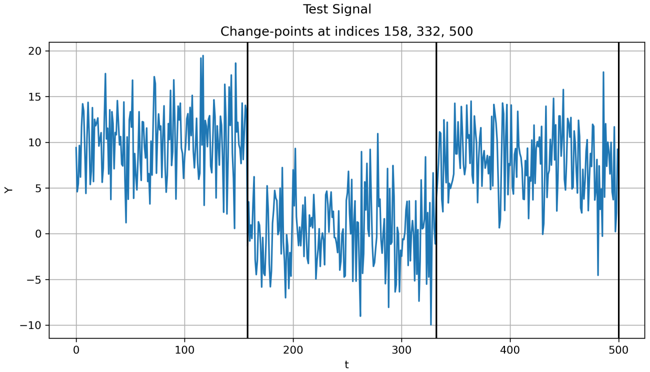
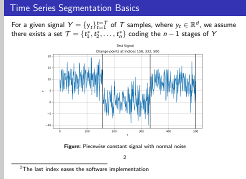


Figure: Piecewise constant signal with normal noise

2

[2] The last index eases the software implementation



Search Algorithms: Binary Segmentation

Greedy sequential algorithm that estimates 1st change-point as:

$$\hat{t}_1 := \underset{1 \le t < T-1}{\arg\min}\ C(y_{0\ldots t}) + C(y_{t\ldots T}) \tag{1}$$

It then repeats the operation on both left and right segments until the specified number of splits is achieved.

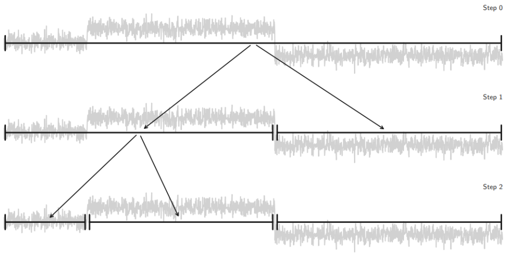


Figure: Overview of Binary Segmentation retrieved from [4]

Search Algorithms: Binary Segmentation
Greedy sequential algorithm that estimates 1st change-point as:
It then repeats the operation on both left and right segments until the specified number of splits is achieved.
Figure: Overview of Binary Segmentation retrieved from [4]

Cost Function: L1 Norm

A robust estimator of shifts in the central point of a distribution. Given a series $\{y_t\}_{t\in\mathcal{I}}$ where $y_t \in \mathbb{R}^d$:

$$C(y_{\mathcal{I}}) = \sum_{d} \sum_{t\in\mathcal{I}} ||y_t - \bar{y}||_1 \tag{2}$$

where $\bar{y}$ is the component wise median of $y_{\mathcal{I}}$

- Needs scaling, considering not to penalize on segment length

Cost Function: L1 Norm
A robust estimator of shifts in the central point of a distribution. Given a series $\{y_t\}_{t\in\mathcal{I}}$ where $y_t \in \mathbb{R}^d$:
$$C(y_{\mathcal{I}}) = \sum_{d} \sum_{t\in\mathcal{I}} ||y_t - \bar{y}||_1 \tag{2}$$
where $\bar{y}$ is the component wise median of $y_{\mathcal{I}}$
- Needs scaling, considering not to penalize on segment length

Segmentation of Environmental Descriptors

For a test case we apply Binary Segmentation with L1 norm and min. size for segment of 20% the length of the array

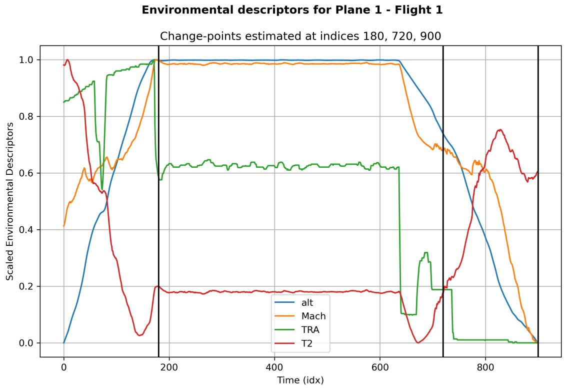
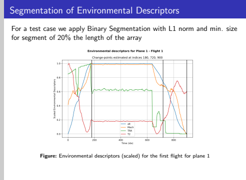


Figure: Environmental descriptors (scaled) for the first flight for plane 1

2023-09-11

Research Stay: Predictive Maintenance of Industrial Equipment
└Experimental Results
└Segmentation of Environmental Descriptors



Segmentation of Environmental Descriptors

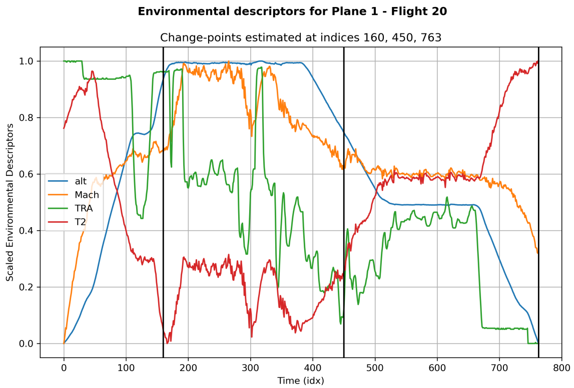
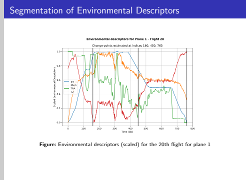


Figure: Environmental descriptors (scaled) for the 20th flight for plane 1

2023-09-11

Research Stay: Predictive Maintenance of Industrial Equipment
└Experimental Results
└Segmentation of Environmental Descriptors



References

BBC News, "Boeing 777: Dozens grounded after denver engine failure,", Feb. 22, 2021. [Online]. Available: https://www.bbc.com/news/world-us-canada-56149894.

M. Arias Chao, C. Kulkarni, K. Goebel, and O. Fink, "Aircraft engine run-to-failure dataset under real flight conditions for prognostics and diagnostics," *Data*, vol. 6, no. 1, p. 5, 2021.

PHM Society. (Oct. 21, 2021), 2021 phm conference data challenge - phm society data repository, [Online]. Available: https://data.phmsociety.org/2021-phm-conference-data-challenge/.

C. Truong, L. Oudre, and N. Vayatis, "Selective review of offline change point detection methods," *Signal Processing*, vol. 167, p. 107 299, 2020.

References

BBC News, "Boeing 777: Dozens grounded after denver engine failure,", Feb. 22, 2021. [Online]. Available: https://www.bbc.com/news/world-us-canada-56149894.

M. Arias Chao, C. Kulkarni, K. Goebel, and O. Fink, "Aircraft engine run-to-failure dataset under real flight conditions for prognostics and diagnostics," *Data*, vol. 6, no. 1, p. 5, 2021.

PHM Society. (Oct. 21, 2021), 2021 phm conference data challenge - phm society data repository, [Online]. Available: https://data.phmsociety.org/2021-phm-conference-data-challenge/.

C. Truong, L. Oudre, and N. Vayatis, "Selective review of offline change point detection methods," *Signal Processing*, vol. 167, p. 107 299, 2020.

False frame

False frame