CAN YOU ACTUALLY PREDICT A FUTURE FAILURE OF A VEHICLE?

RUNE PRYTZ – HEAD OF R&D
STRATIO AUTOMOTIVE

stratio AUTOMOTIVE

DSPT DATA SCIENCE PORTUGAL

history of connected vehicles

1991 - On board diagnostic (OBD)

1996 - GM OnStar

2000 - GM OnStar + GPS

2001 - Remote OBD scan

2003 - Vehicle Health reports

2014 - CarPlay, Android Auto, Apps, WIFI...

Add pictures of vehicles and data aqusition devices

1991 - OBD gets adopted
1996 - GM OnStar

2000 - GM OnStar + GPS, accident call with position

2001 - Remote OBD scan

2003 - Vehicle status reports

2014 - CarPlay, Android Auto, Apps, WIFI, Predictive maintenance...

Who drives the predictive maintenance efforts today?

Any guesses?

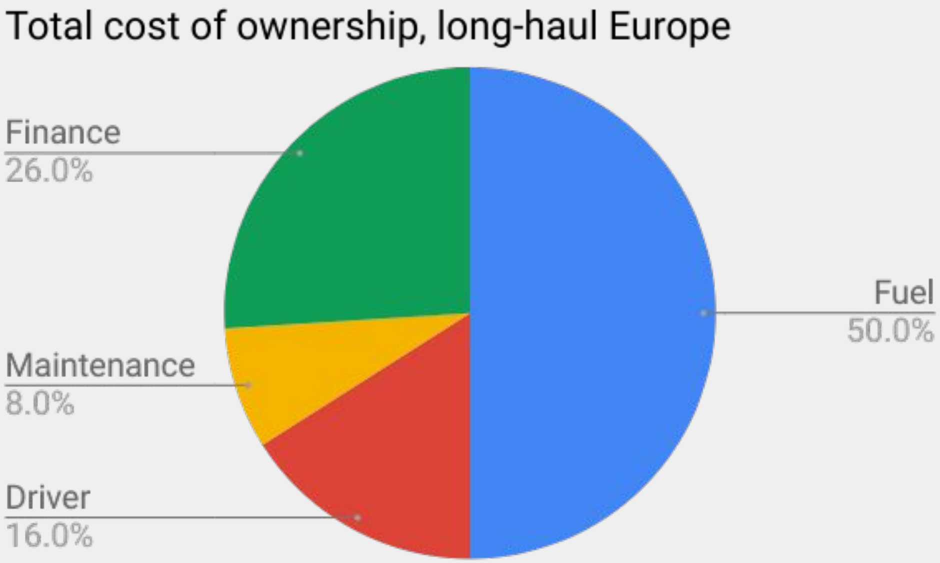
DRIVERS

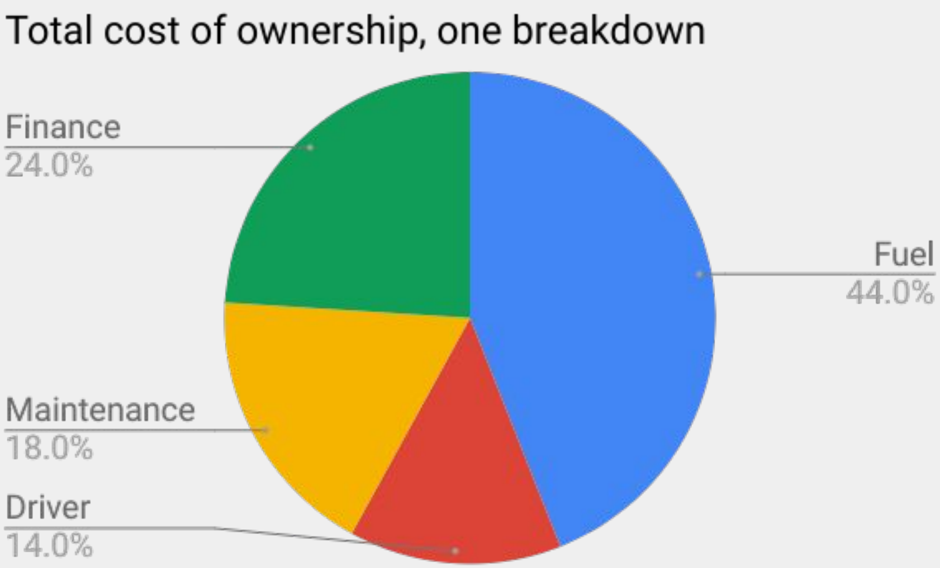
- The Operational costs
  - Downtime cost money

- Sales of spare parts at workshops

- Road system efficiency
  - congestions

Downtime cost money





TURBO BREAKDOWN
IN NUMBERS

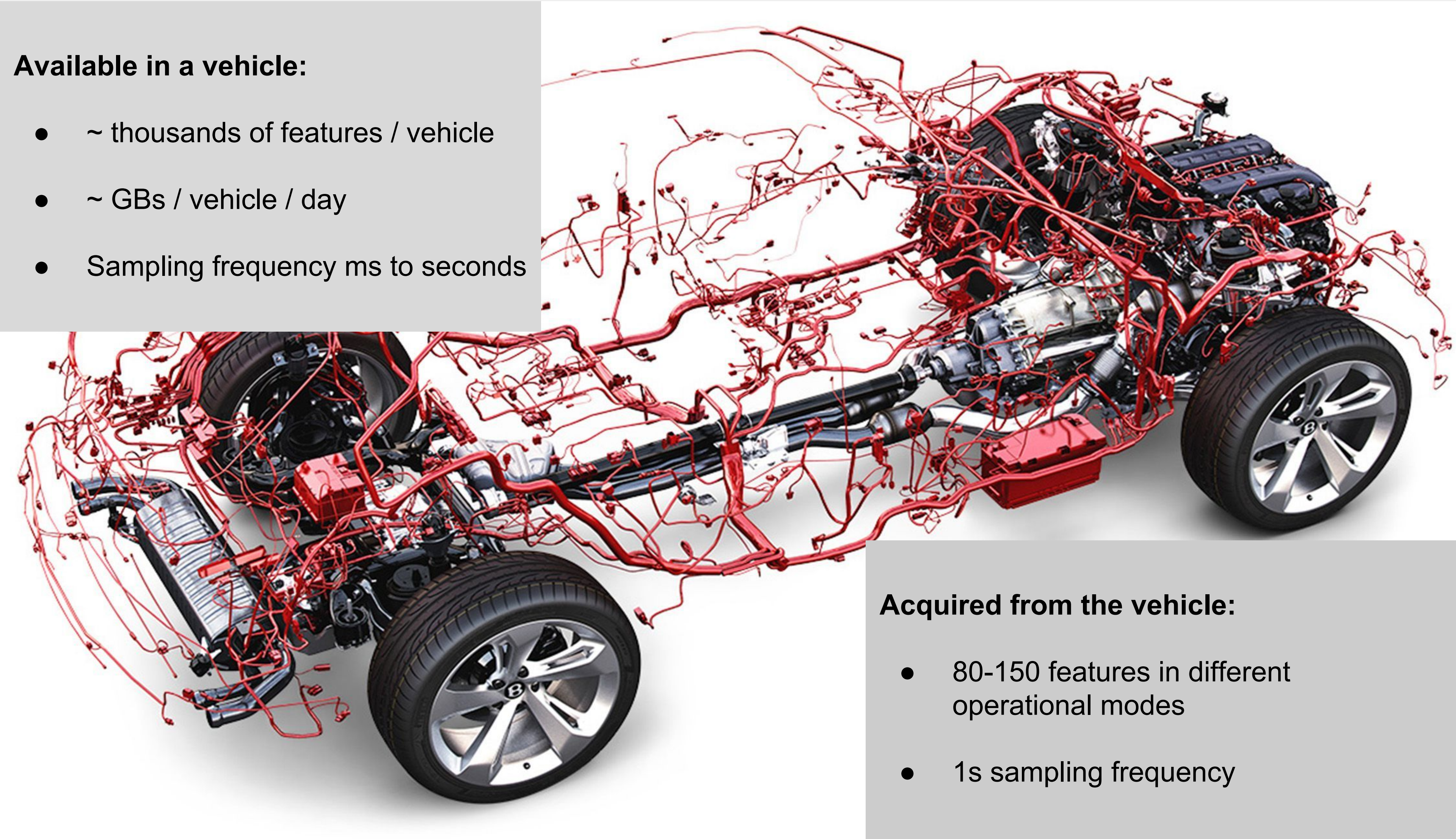
PREVENTIVE REPLACEMENT	
SPARE PART (TURBO)	€1500
LABOR	€250
TOTAL:	€1750

CORRECTIVE REPAIR	
TOWING	€500
SPARE PARTS	€5000
ENGINE CLEANING	€500
LABOR	€1000
LOST REVENUE	N/A
TOTAL:	>€7000

DATA ACQUISITION

Available in a vehicle:

- ~ thousands of features / vehicle
- ~ GBs / vehicle / day
- Sampling frequency ms to seconds



Acquired from the vehicle:

- 80-150 features in different operational modes
- 1s sampling frequency

TWO WAYS OF ACHIEVING P.M

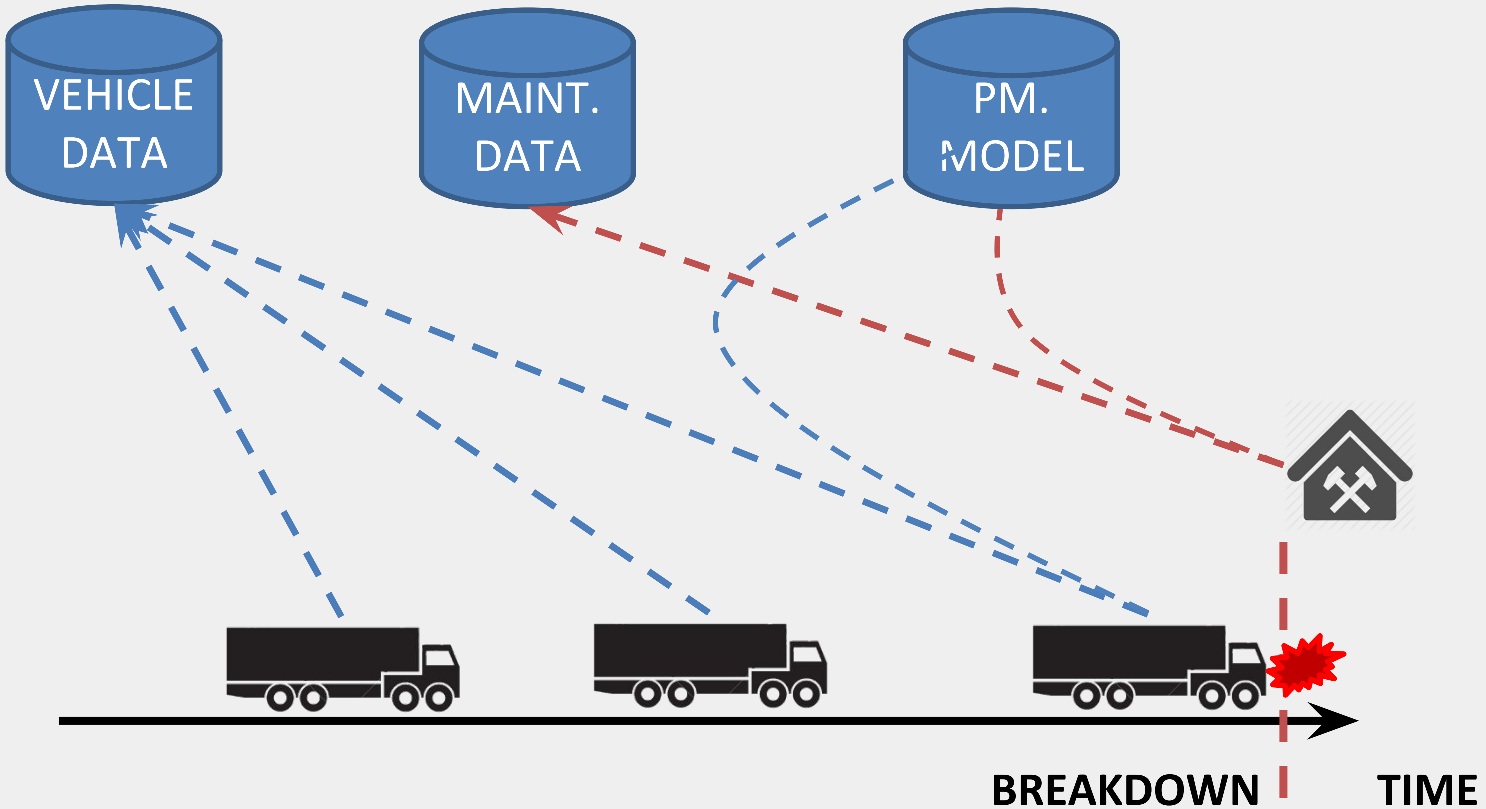
- Classic supervised machine learning

Interesting Faults → Vehicle Data → Predictive Maintenance

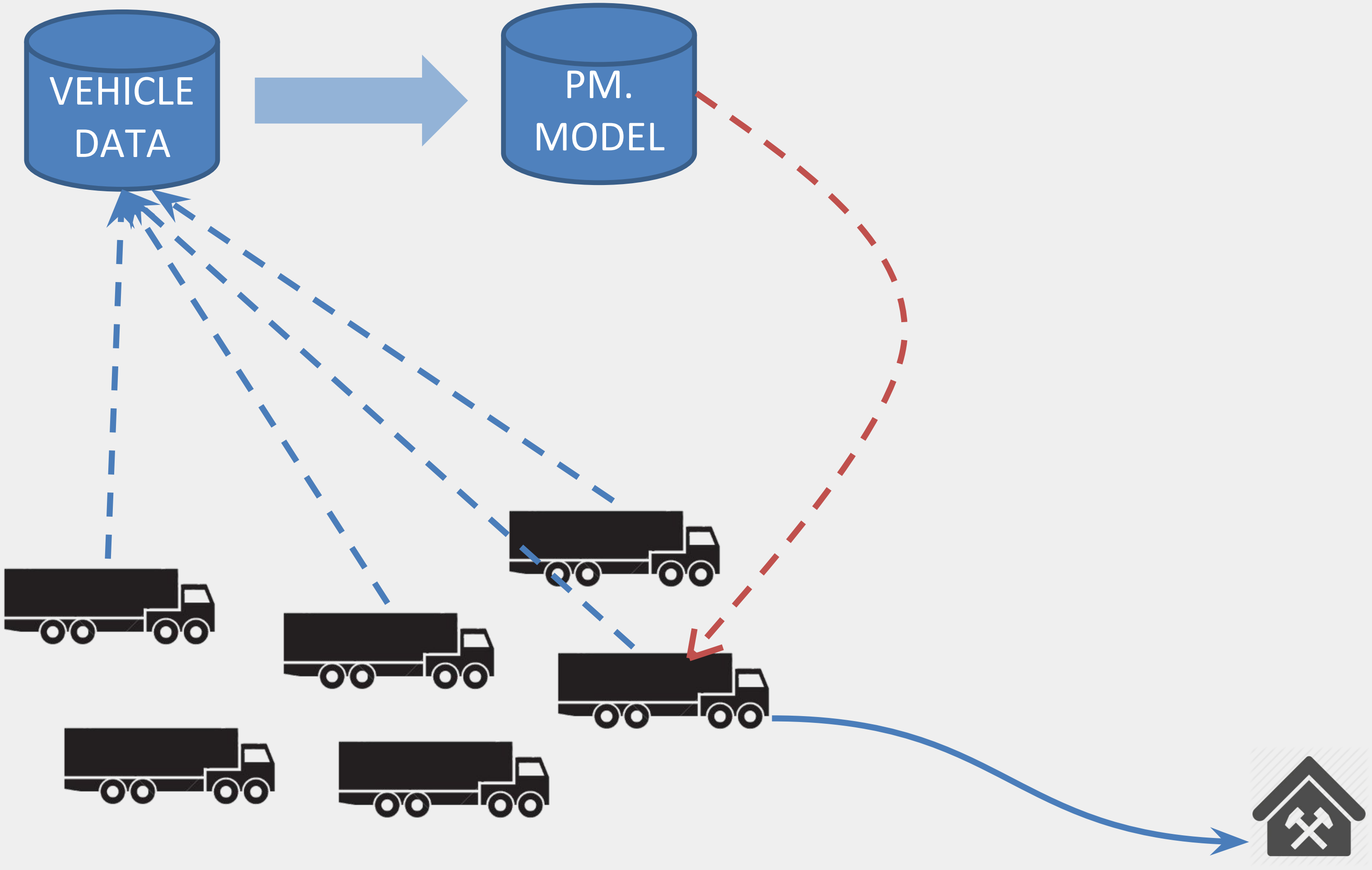
- Unsupervised knowledge discovery

Vehicle Data → Explaining Deviations → Predictive Maintenance

SUPERVISED LEARNING



SUPERVISED LEARNING



THE DATASET

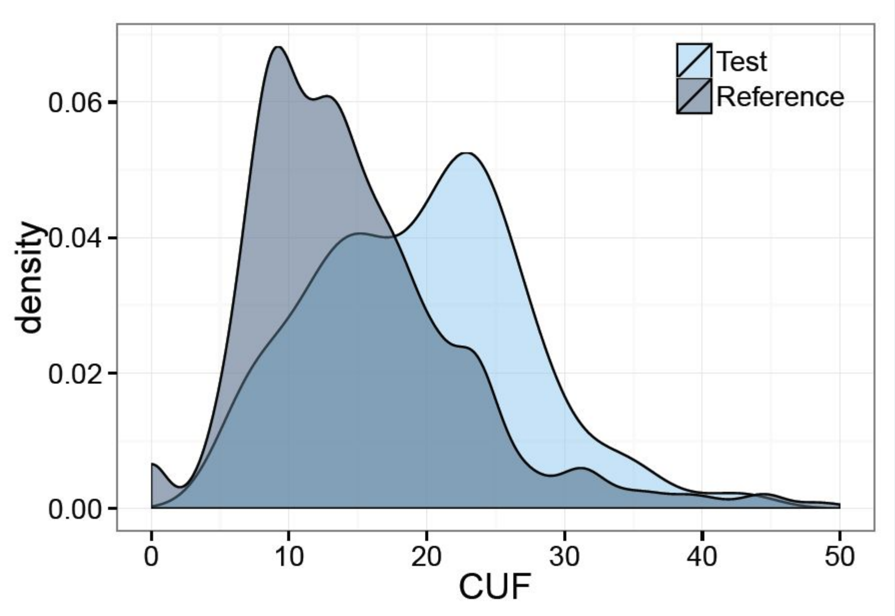
- The source of data is not Independent and Identically Distributed
  - Selecting train and test dataset is difficult

Date	Chassis ID	LVD CUF	LVD BIY	LVD OO	VDA IHF	TTF	Label
130101	A-12345	10	1240	10002	12	200	**Faulty**
130525	A-12345	15	3450	15200	12	75	**Faulty**
130430	A-23456	8	800	34423	13	Inf	Normal
131101	A-23456	6	1600	34555	14	Inf	Normal

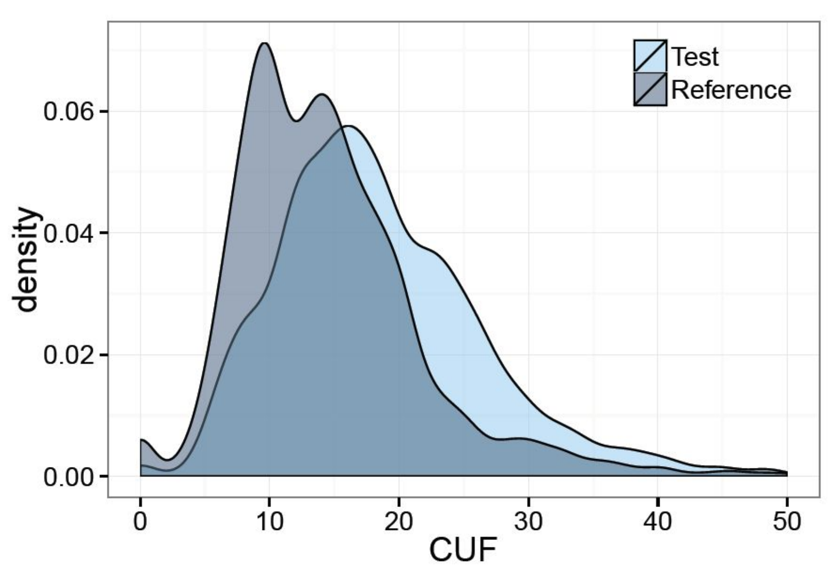
- Labelling data is hard
  - Prediction Horizon

FEATURE SELECTION

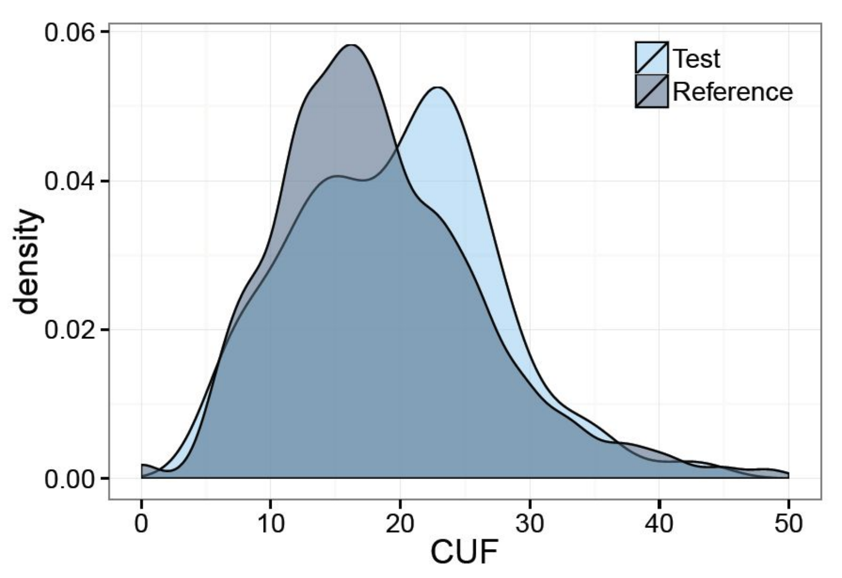
- Heterogeneous dataset across vehicles
  - No traditional method for feature selection
  - Differences in distribution based on usage and wear



FEATURE SELECTION
WEAR

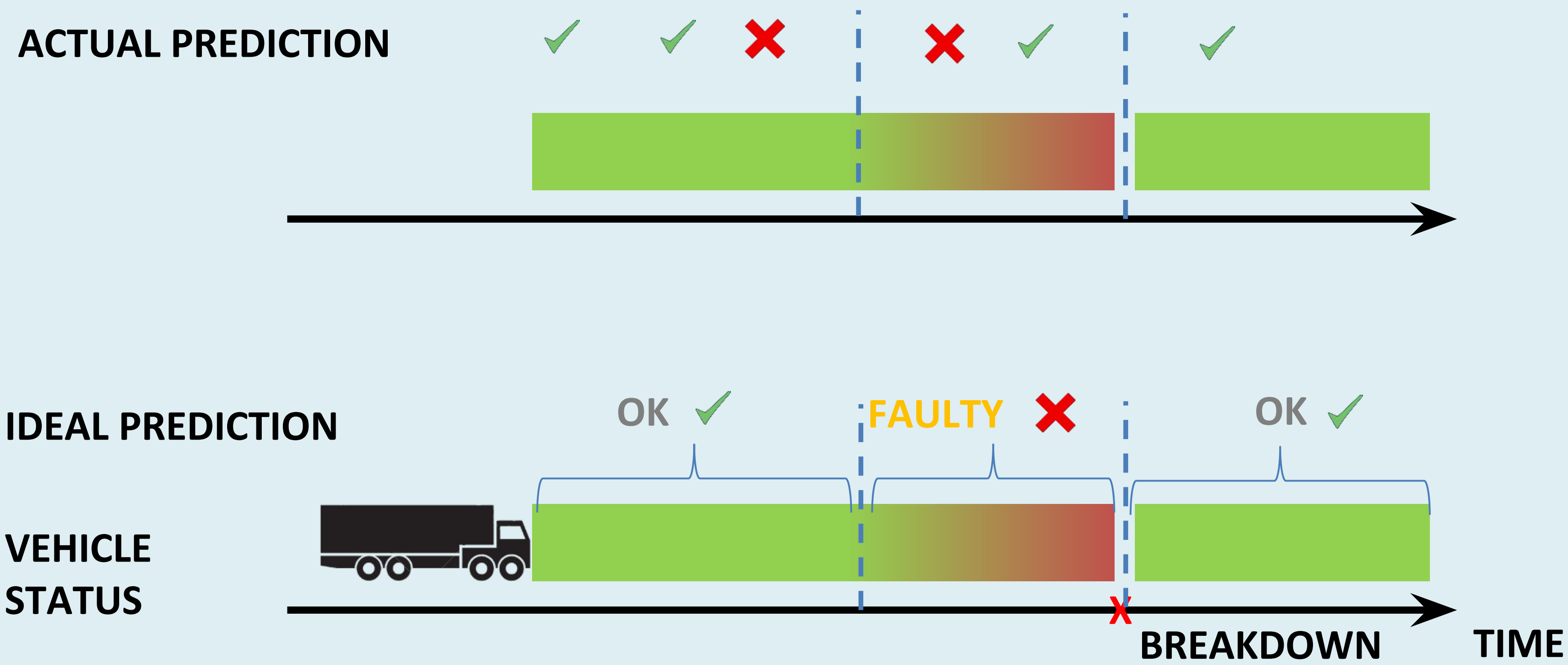


Data from 25 weeks prior failure



Data from 5 weeks prior failure

PREDICTING A SINGLE VEHICLE



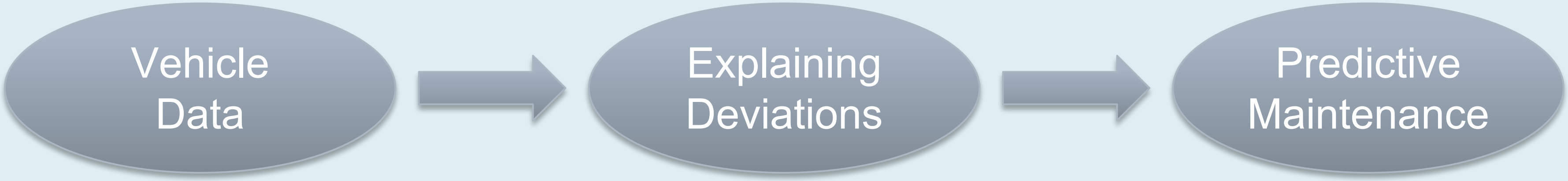
EVALUATING RESULT

	Predicted Faulty	Predicted Normal
Observed Faulty	True positive	False Negative
Observed Normal	False Positive	True Negative

EVALUATING RESULT

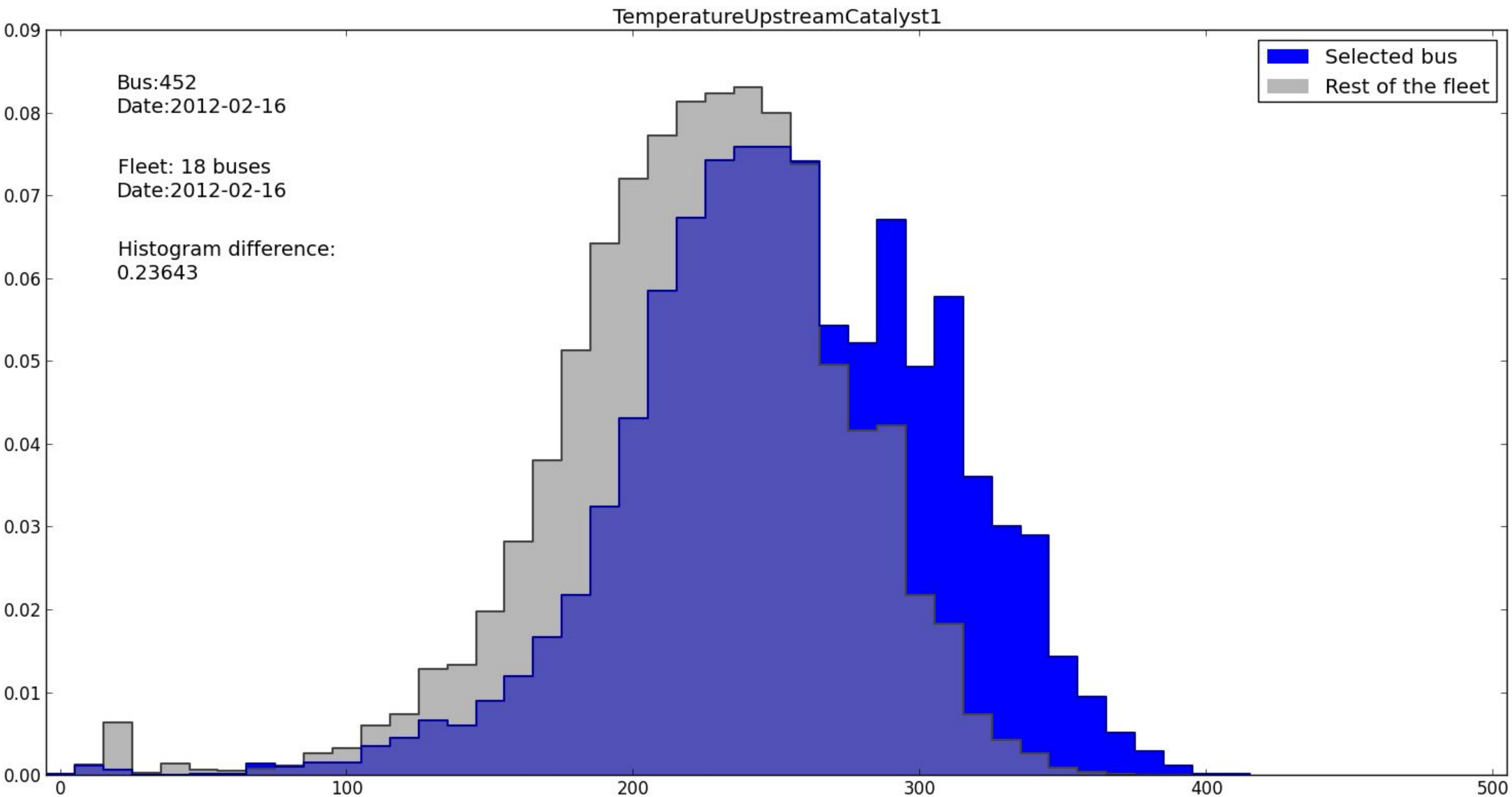
	Predicted Faulty	Predicted Normal
Observed Faulty	True positive Warned in advanced	False Negative No warning / vehicle failure
Observed Normal	False Positive Unnecessary warning / repair shop visit.	True Negative No action

UNSUPERVISED FLEET BASED APPROACH

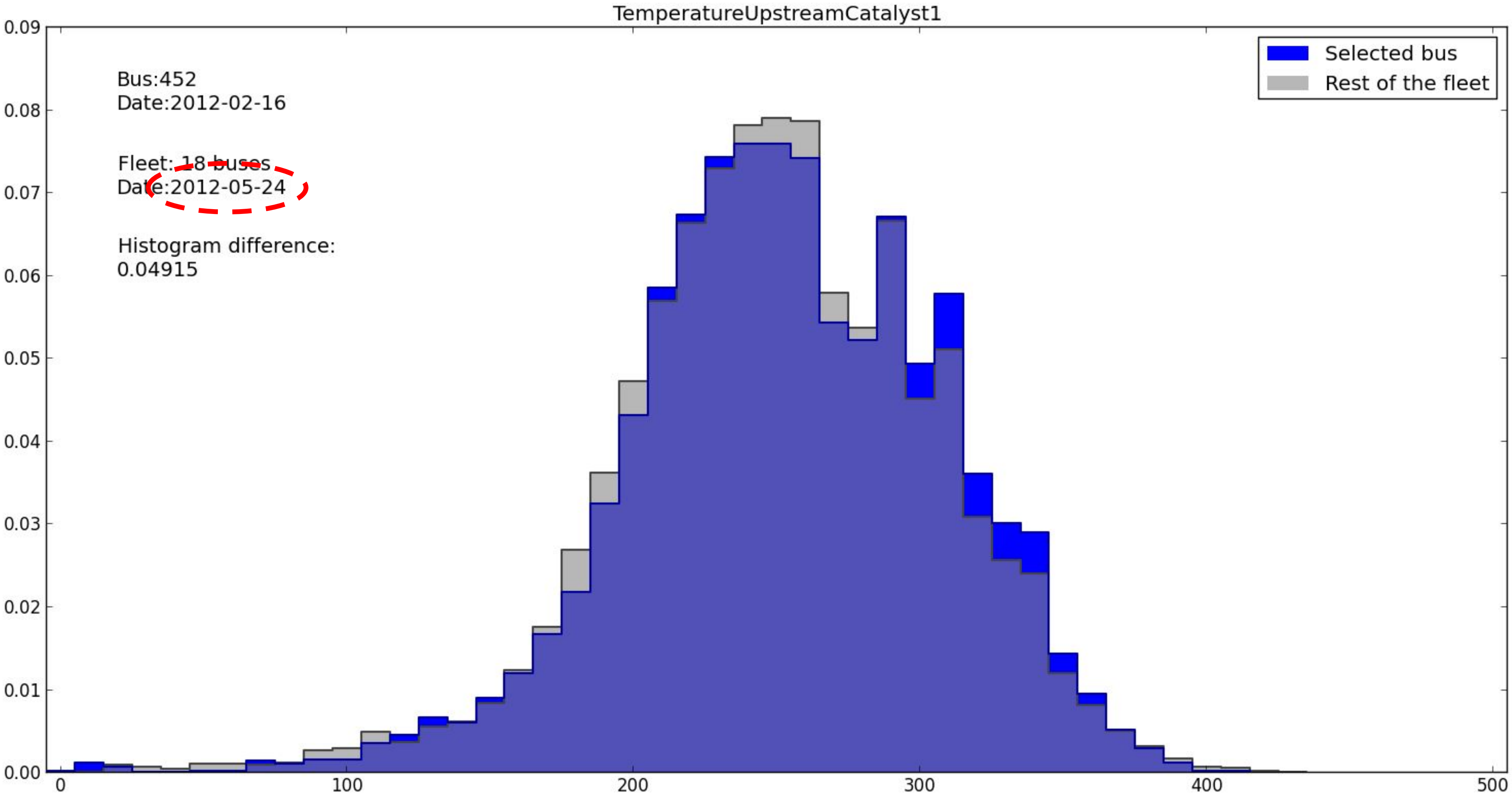


T. Rögnvaldsson, S. Nowaczyk, S. Byttner, R. Prytz, M. Svensson,
Self- monitoring for maintenance of vehicle fleets, Data Mining and Knowledge Discovery

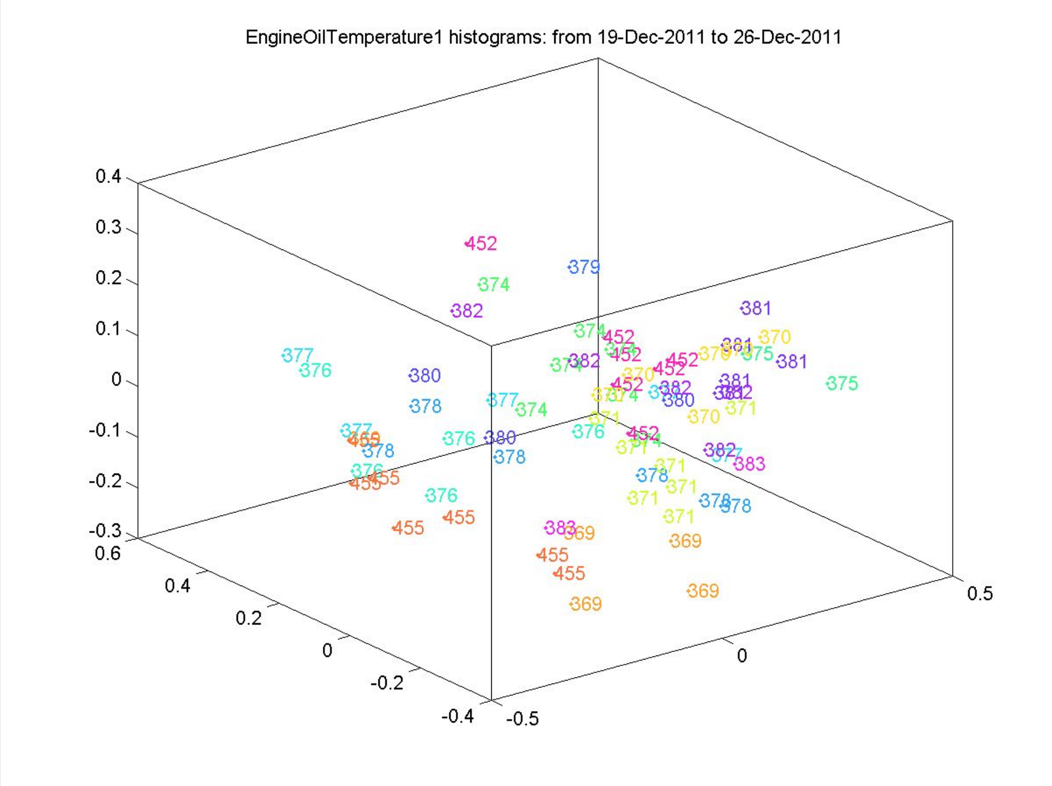
FLEET BASED APPROACH



FLEET BASED APPROACH



HISTOGRAMS OF VEHICLES

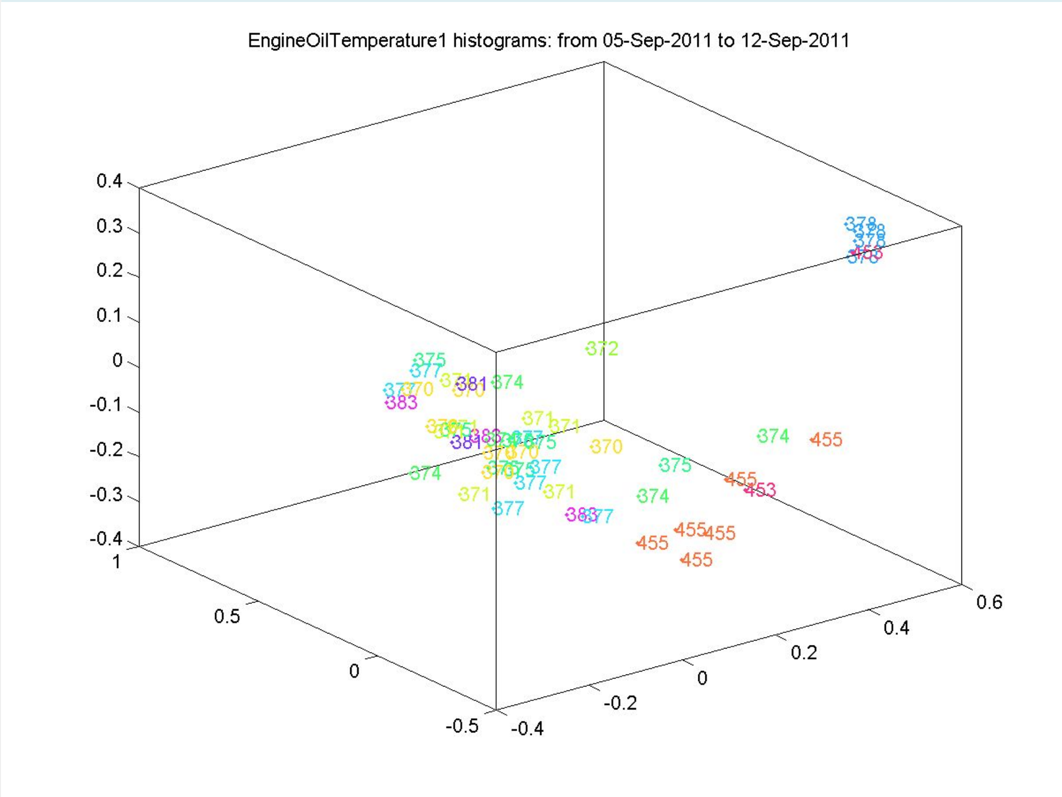


Each point corresponds to a bus observation (illustrated using multidimensional scaling)

- Distance between histograms
- 19 vehicles
- One week data

Evenly distributed - No significant pattern

WHEN SOMETHING IS ODD...



We are looking for a vehicle that is different from rest of the group

Short circuit in ECU lead to coolant fan ran at 100% all the time

WHY THIS IS IMPORTANT TO ME

- A strong business case with a clear objective
- Easy to relate to cars and busses. You see your product every day in Lisbon for instance.
- A tough problem with a potential to change people's life

QUESTIONS

notes

Outline:

1. History
2. Why Predictive Maintenance
3. How Predictive Maintenance
   a. Supervised
   b. Unsupervised
4. Feature engineering & feature selection
5. Problems & result
   a. Question about
6. Why this is important to me:
   a. Clear link between data - >model -> prediction business implication
   b. Vehicles you see in every day life (Carris, Lisbon trash trucks)
   c. Tough ML problems, get to work with great minds.