Accomodating challenges of bringing measurement data into AM systems

ERPUG, Oct. 20th 2016

Vejdirektoratet
Danish Road Directorate

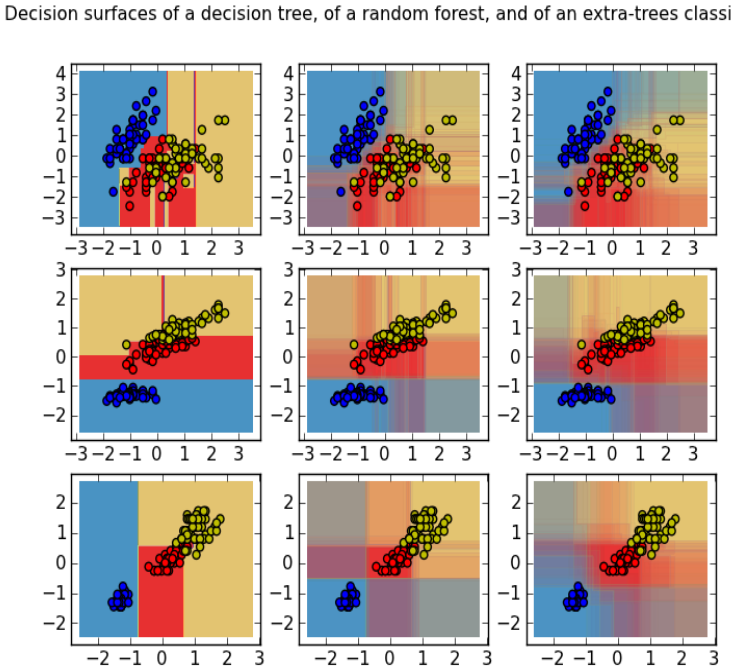
ERPUG

Motivation

Work on network-level data for Asset Management

Examples:

- PMS using Machine Learning (Big Road Data)
- Socio-economic analyses



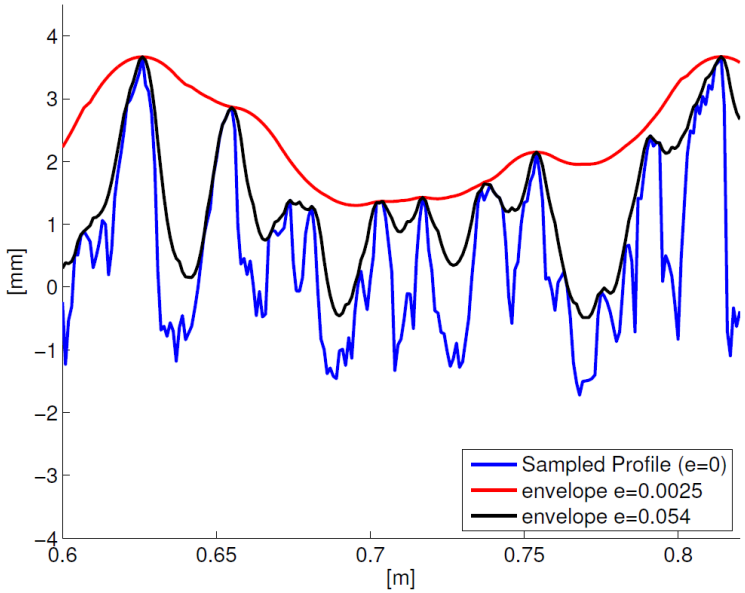
Vejdirektoratet
Danish Road Directorate

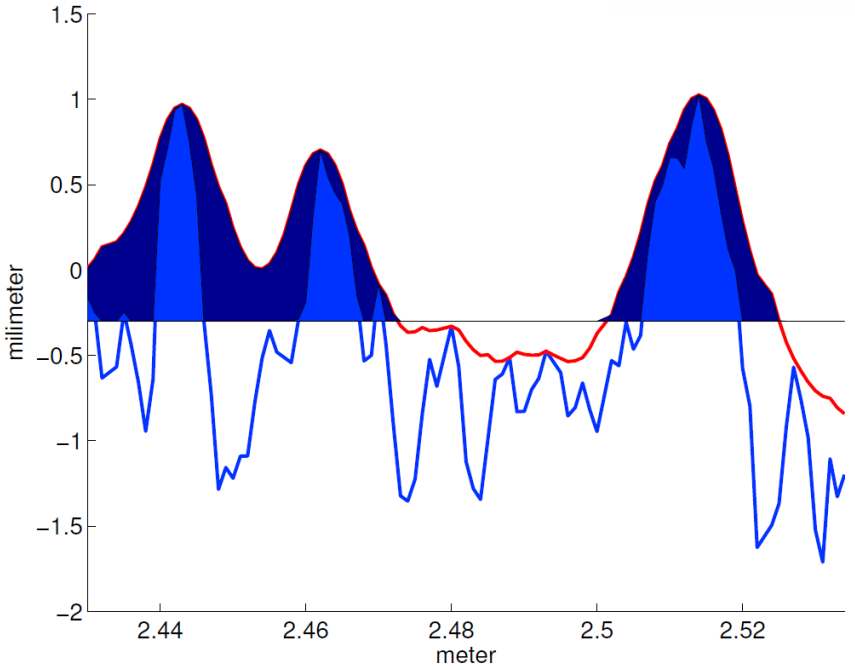
Motivation

Alternative road metrics for Asset Management

Examples:

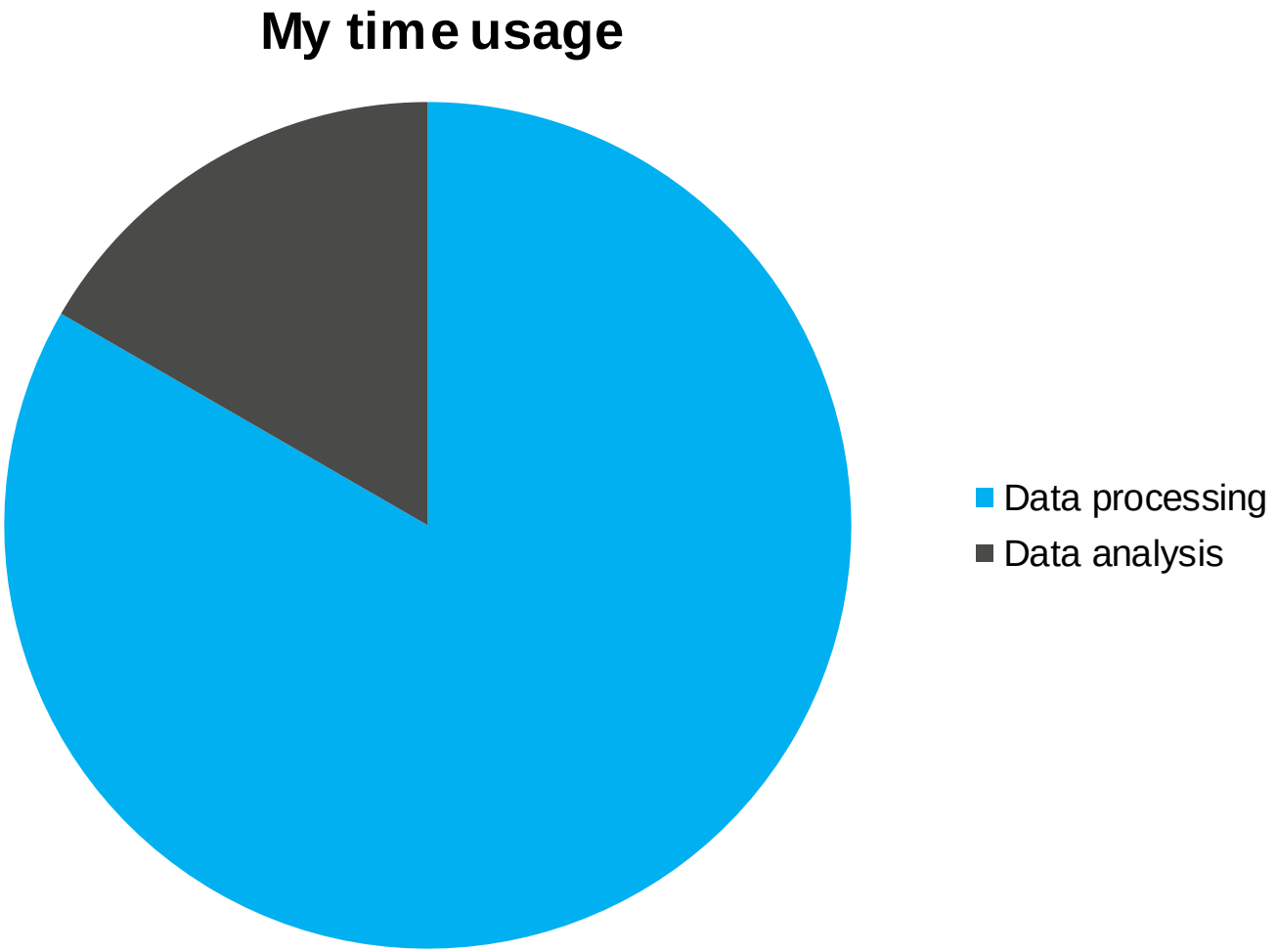
- Texture Penetration Area (TPA)
- Enveloping





Vejdirektoratet
Danish Road Directorate

Challenge



Vejdirektoratet
Danish Road Directorate

Example: Raw profiles

Data formats

- CSV / Almost CSV
- Non-standardized text-format
- TDMS (Proprietary format by Native Instruments)

Example: Raw profiles

Data formats

- CSV / Almost CSV
- Non-standardized text-format
- TDMS (Proprietary format by Native Instruments)

Writing a new parser for every data set!

Example: Raw profiles

Extracting data from equipment

- Binary format from vendor **1** – use their GUI software
- Binary format from vendor **2** – use their GUI software
- …
- Binary format from vendor **n** – use their GUI software

Example: Raw profiles

Extracting data from equipment

- Binary format from vendor **1** – use their GUI software
- Binary format from vendor **2** – use their GUI software
- …
- Binary format from vendor **n** – use their GUI software

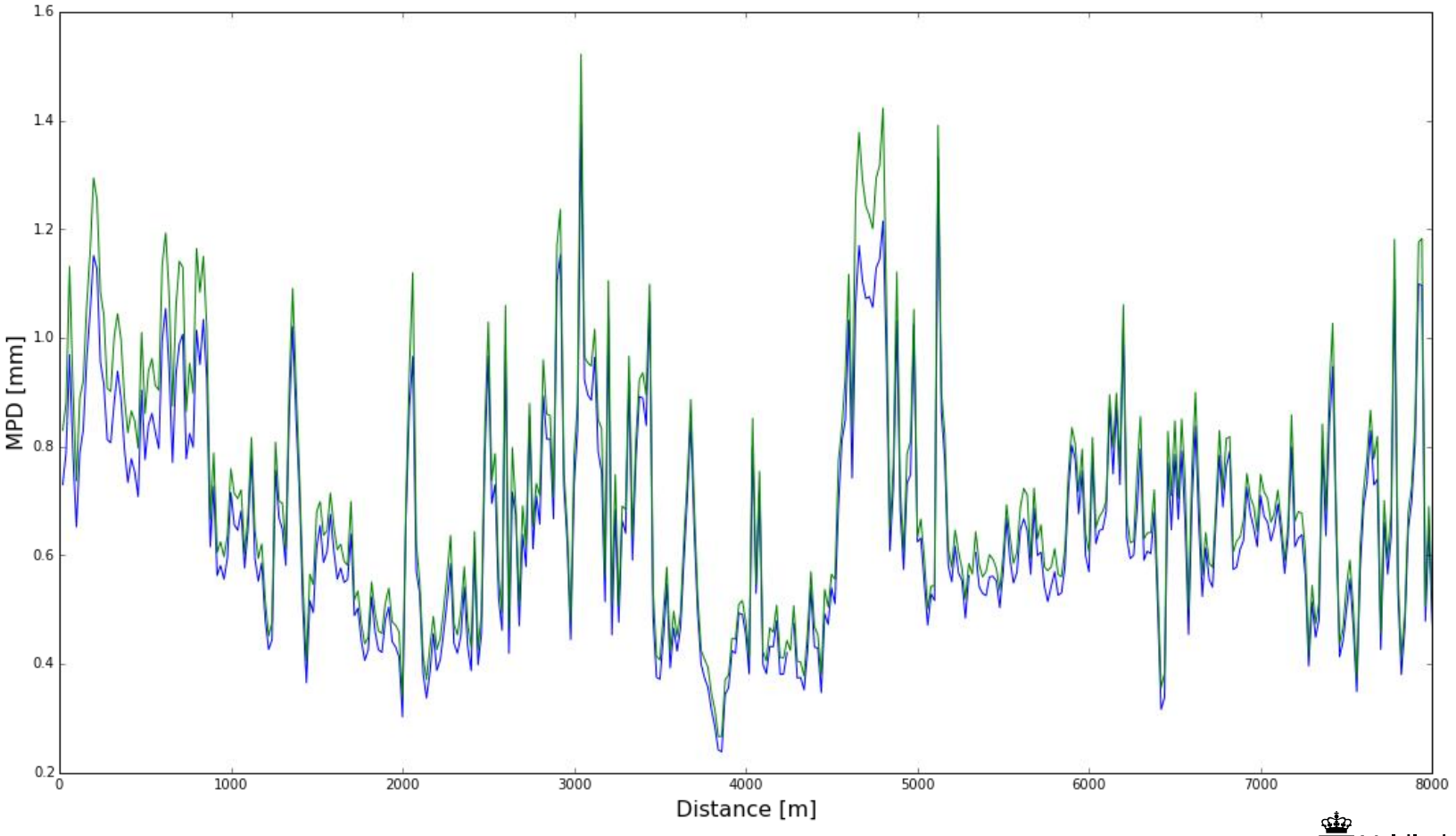
Obstacles:

- **Learning a new piece of software for each vendor**
- **Non-programmable**
- **Text output (verbose)**
- **Non-standard to deal with raw profiles**

Example: Raw profiles

Algorithm Differences

The devil is in the detail: Interpretation of standards (e.g. ISO13473-1)



Vejdirektoratet
Danish Road Directorate

Possible solutions

Standardized data storage

Possible solution

- CSV/JSON (+ gzip)
- SQLite (+ db schema)
- HDF5 (+ spec)

Pros

- API for many languages
- Platform independent storage
- MIT-like licenses

Possible solutions

Standardized data storage

Possible solution

- **CSV/JSON (+ gzip)**
- HDF5 (+ spec)
- SQLite (+ db schema)

Pros

- Human-readable (esp. CSV)

Cons

- Space inefficient
- Rigid storage (esp. CSV)
- Inefficient random access

Possible solutions

Standardized data storage

Possible solution

- CSV/JSON (+ gzip)
- **HDF5 (+ spec)**
- SQLite (+ db schema)

Pros

- Compact (binary) storage
- Flexible

Cons

- Single implementation
- Spec to describe structure

Possible solutions

Standardized data storage

Possible solution

- CSV/JSON (+ gzip)
- HDF5 (+ spec)
- **SQLite (+ db schema)**

Pros:

- Compact (binary) storage
- High flexibility of SQL

Cons:

- Single implementation
- Spec to describe structure
- One large file

Standardization

Should we try standardizing data storage?

Vejdirektoratet
Danish Road Directorate

Thank you for your attention!