The use of AI for improved moisture measurement

Radek Strnad, ČMI

Prometec

TÉCNICO LISBOA

DANISH TECHNOLOGICAL INSTITUTE

VERDO

CETIAT
ensemble, innover et valider

PTB
Physikalisch-Technische Bundesanstalt
Braunschweig und Berlin

THE FRESNEL COMPANIES

TÜBİTAK UME

IMBIH

ČESKÝ METROLOGICKÝ INSTITUT

biofMET

EMPIR

EURAMET

The EMPIR initiative is co-funded by the European Union's Horizon 2020 research and innovation programme and the EMPIR Participating States

Stakeholder workshop
2.-3. June 2022

WP3 – Industrial validation

- Task 3.1: Solid biofuels
- Task 3.2: Liquid biofuels
- Task 3.3: Data mining and data-science techniques

The aim of this work package is therefore to **validate** the methods and devices developed in WP1 and WP2.

- **uncertainty propagation** models for industrial use will be prepared
- methods and devices will be tested in **real industrial applications** for solid and liquid biofuels measurement
- analyses will be refined by **applying data science methods** such as machine (deep) learnings and artificial intelligence on the now fully digitalised measurements (correlations arising from the industrial environment - *temperature, reflections, humidity, dust, vibrations* etc.)



WP3 – Industrial validation

- Task 3.1: Solid biofuels
- Task 3.2: Liquid biofuels
- **Task 3.3: Data mining and data-science techniques**

The aim of this task is to exploit **advanced data-science techniques** to improve reproducibility and repeatability of moisture content measurements. In addition, an **uncertainty model of real application** will be developed, and a **good practise guide** will be compiled.

- Development of a method for the analysis of results of moisture content (machine learning, artificial intelligence, deep learning), tested on data from the measurements campaign



Data mining and data-science techniques

- Data mining = process to crunch huge amount of data to extract new patterns, trends, correlations, making predictions etc.:
  - Classification
  - Segmentation
  - Prediction
  - Regression
  - Making associations
  - Text Mining
- How?

Statistical x propabilistic approach, etc

$$RSS = \mathbf{e}^T\mathbf{e} = \sum_{k=1}^{N}\left(y(k) - y_M(k)\right)^2$$

$$AIC = N\left(1 + \ln(2\pi) + 2\ln(RSS)\right) + 2m_p$$

$$BIC = N\ln(RSS) + m_p\ln(N)$$

$$AICFPE = \frac{1 + N/m_p}{1 - N/m_p}\frac{RSS}{2}$$

  - -> Understanding of the data -> Understanding of the goals of the project -> Cleaning and processing of the data -> Modelling of data (may ressult to reassess the processing of data) -> Select best suited modell to our case (ROC, Gains,… - **evalution criteria**) -> Deployment

Data mining and data-science techniques

- Heuristic approach
  - Artifical Neural networks / Neural networks:
    - collection of units collected together making network, similar to neuron cells in (human) brain
  - decision trees, k-nearest neighbor, support vector machine
  - Fuzzy systems
  - Expert system
  - Evolution and genetic algorithms
  - data – 50/25/25 (training/testing/validation)

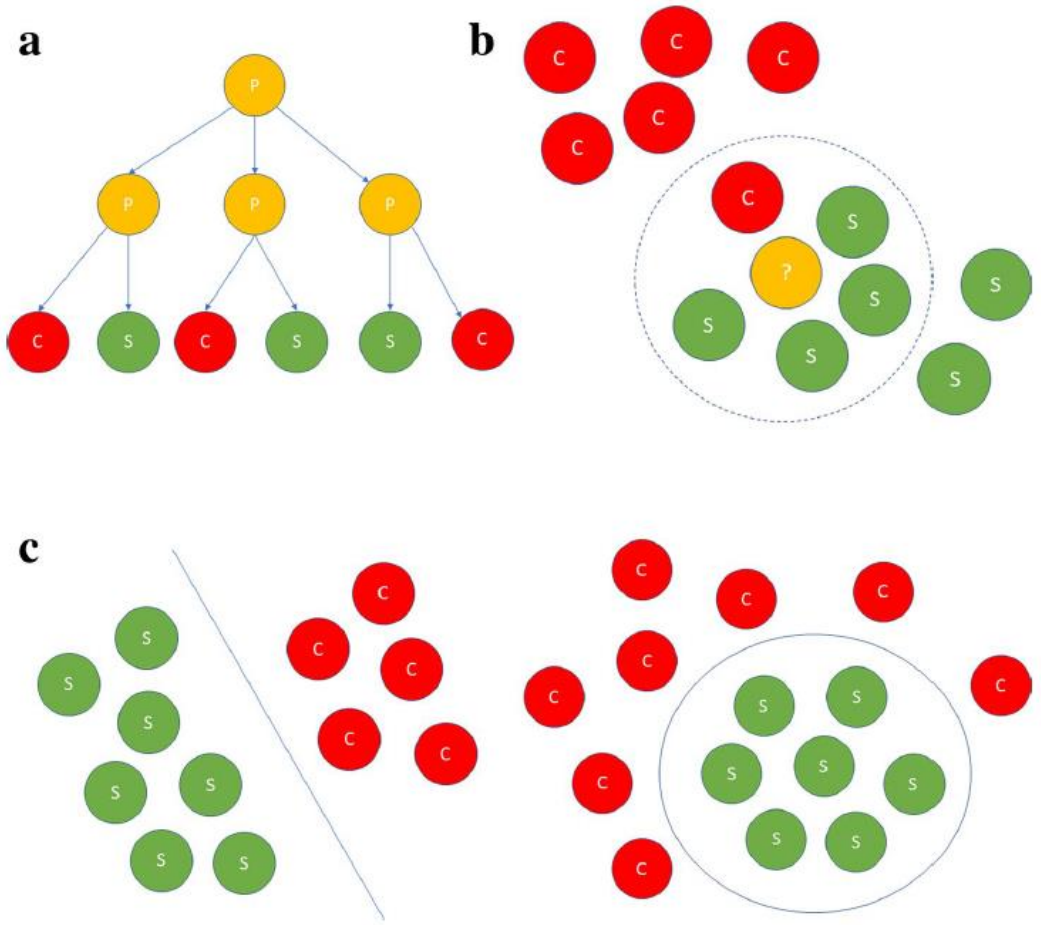


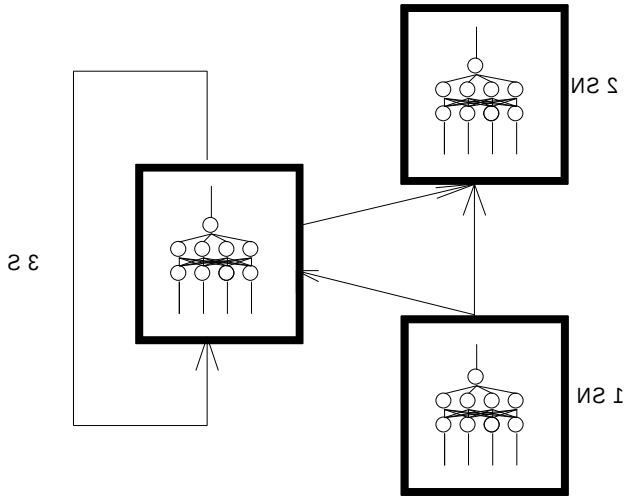
Fig. 2 Principle of the classifiers: **a** decision trees, **b** *k*-nearest neighbor, **c** support vector machine

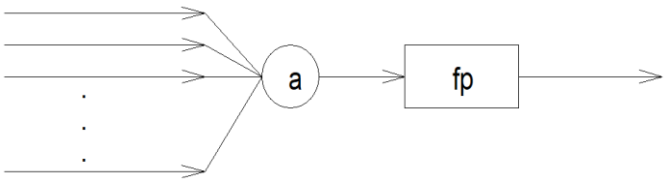
Picture: M. Tiitta et al. „Air-coupled ultrasound detection of natural defects in wood using ferroelectret and piezoelectric sensors," *Wood Sci. and Tech.*, sv. 54, no. 4, pp. 1-14, 2020.

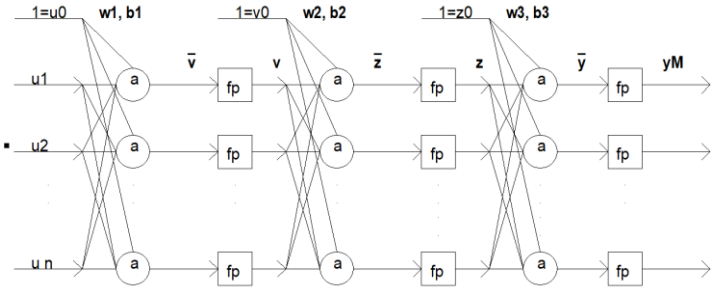
Data mining and data-science techniques

Decide what we would like to receive

- Calibration curves
  - Regression
  - Structure identification
  - Nonparametric model
- Neural networks
  - Standard
  - Recursive
  - type of perceptron model
- Influence factors
- Structure of the data
  - Who
  - When
  - What







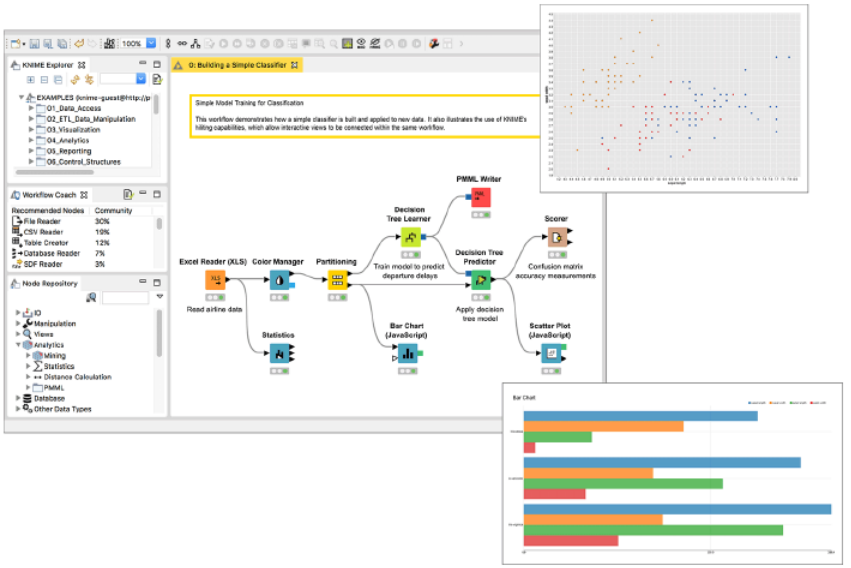
Data science tool for advance AI/ML models

Open for Innovation KNIME

Swiss army knife for data

What is KNIME Analytics Platform?

- A tool for data analysis, manipulation, visualization, and reporting
- Based on the graphical programming paradigm
- Provides a diverse array of extensions:
  - Text Mining
  - Network Mining
  - Cheminformatics
  - Many integrations, such as Java, R, Python, Weka, Keras, Plotly, H2O, etc.

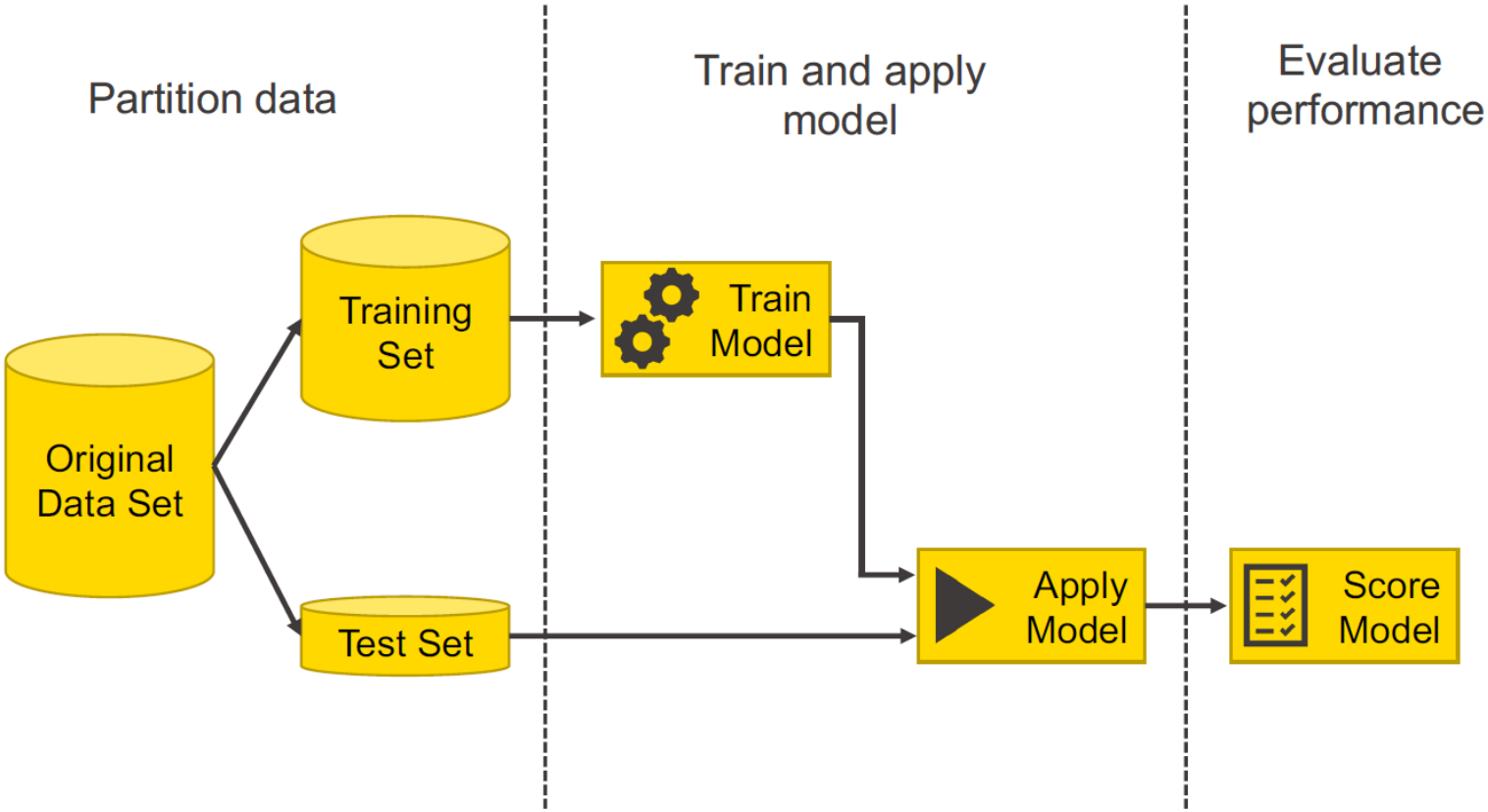


Data science tool for advance AI/ML models

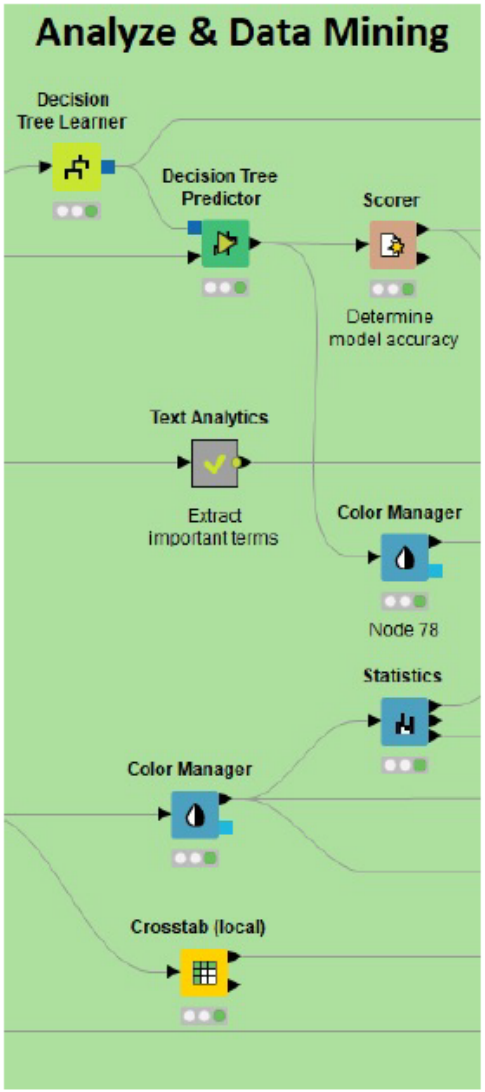
Open for Innovation KNIME

Swiss army knife for data

Data Mining: Process Overview



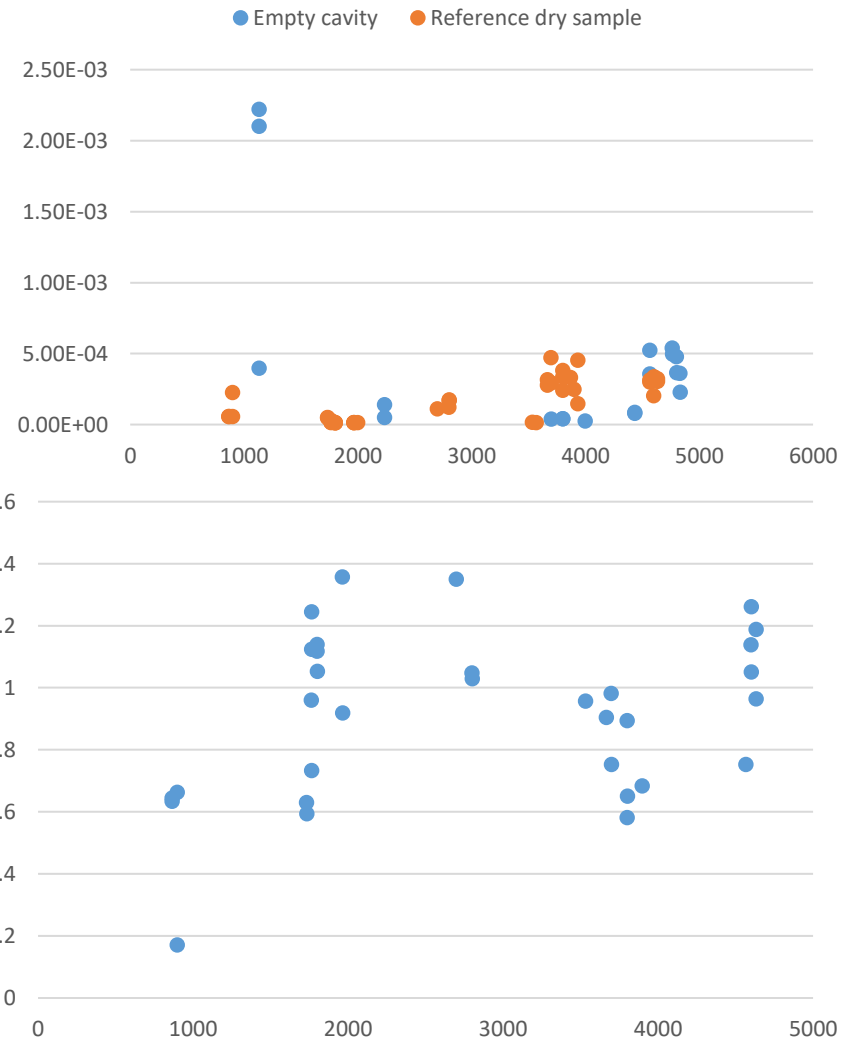
Data science tool for advance AI/ML models



- Regression
  - Linear, logistic
- Classification
  - Decision tree, ensembles, SVM, MLP, Naïve Bayes
- Clustering
  - k-means, DBSCAN, hierarchical
- Validation
  - Cross-validation, scoring, ROC
- Deep Learning
  - Keras, DL4J
- External
  - R, Python, Weka, H2O, Keras

Some examples

- Preprocessing
- Data manipulation
- Cleaning
- Visualisation
- Statistic



One-way analysis of variance (ANOVA)

Descriptive Statistics

Confidence Interval (CI) Probability: 95.0%

	Group	N	Missing	Missing Group	Mean	Std. Deviation	Std. Error	CI (Lower Bound)	CI (Upper Bound)	Minimum	Maximum
Universe_0_0	Cluster_0	16	0	0	0,3555	0,0933	0,0233	0,3058	0,4052	0,1537	0,4756
Universe_0_0	Cluster_1	16	0	0	0,7063	0,0807	0,0202	0,6633	0,7493	0,5827	0,8436
Universe_0_0	Cluster_2	16	0	0	0,8518	0,0985	0,0246	0,7993	0,9043	0,6899	0,9793
Universe_0_0	Total	48	0	0	0,6379	0,2286	0,033	0,5715	0,7043	0,1537	0,9793

Levene Test

The Levene Test is used to test for the equality of variances.

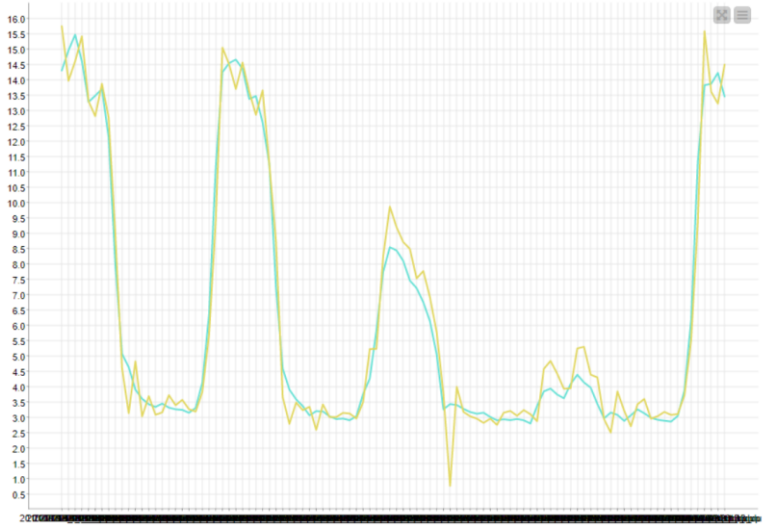
	F	df 1	df 2	p-Value
Universe_0_0	0,3366	2	45	0,7159

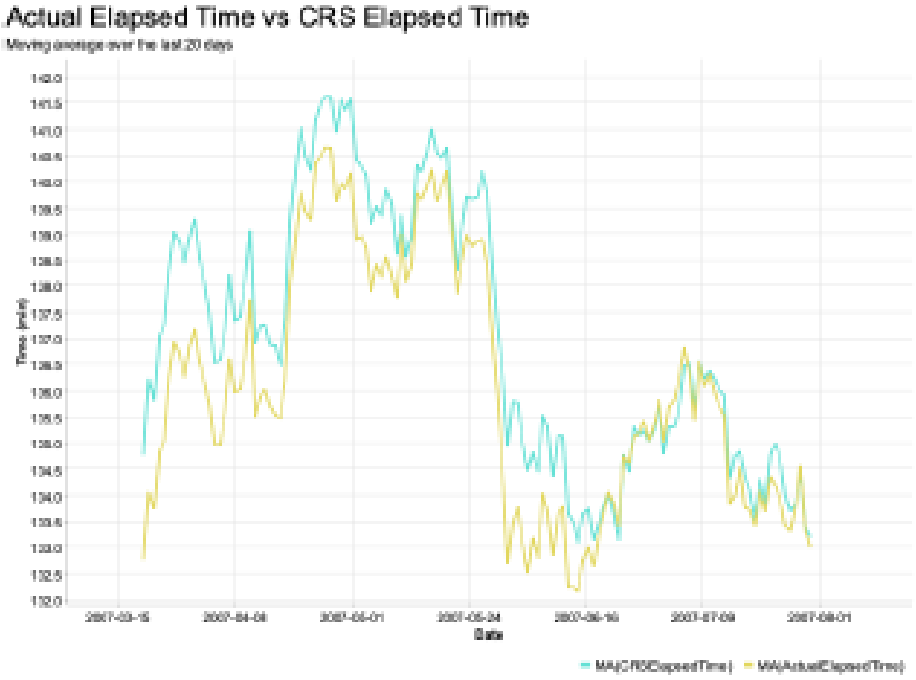
ANOVA

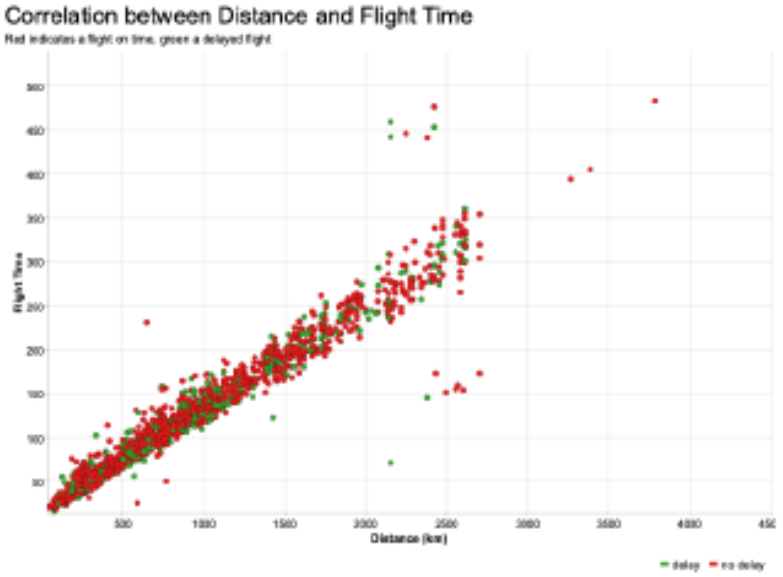
	Source	Sum of Squares	df	Mean Square	F	p-value
Universe_0_0	Between Groups	2,0826	2	1,0413	125,3747	0.0
Universe_0_0	Within Groups	0,3738	45	0,0083		
Universe_0_0	Total	2,4564	47			

Some examples

- Finding model
- Method
- Learning
- Showing results







Some examples

- First human brain is needed
- Create structure
- Computer AI/ML could be applied
- Using on the data

- Finding of the thinks people do not even dream
- Consolidate of the data
- Making traceability propagation to the ON-SITE measurements

The use of AI for improved moisture measurement

Solid biofuels, Liquid biofuels, Data mining and data-science techniques

Prometec

TÉCNICO LISBOA

DANISH TECHNOLOGICAL INSTITUTE

VERDO

TÜBİTAK UME

IMBIH

CETIAT
ensemble, innover et valider

PTB
Physikalisch-Technische Bundesanstalt
Braunschweig und Berlin

THE FRESNEL COMPANIES

ČESKÝ METROLOGICKÝ INSTITUT

biofMET

EMPIR

EURAMET

The EMPIR initiative is co-funded by the European Union's Horizon 2020 research and innovation programme and the EMPIR Participating States

Stakeholder workshop
2.-3. June 2022