VOITH

Bachelor thesis

Analyzing big data from a hydropower plant

Assignment 36 – Marius Lauvland, Børge Størvold and Anders Thon

All the data in this presentation are based on fabricated values.
Any similarity to real values or data is unintentional.

If you require access to the full report contact Nils Tesdal at
The Norwegian University of Science and Technology.

About Voith

- Multinational, family-owned German corporation founded in 1867
- Headquarters in Heidenheim, Germany
  - Norwegian departments in Oslo, Trondheim and Fredrikstad
- Separated into four divisions: Voith Hydro, Voith Paper, Voith Turbo and Voith Digital Solutions
- 4.3 billion Euro in revenue

Problem to be solved

The fourth Industrial Revolution is over us and the Internet of things takes over everyday life. The industry also has completely different requirements and expectations than just a few years ago, also in the power industry. How to address these problems requires you to investigate and try out the methods that exist to see what works.

The problem we are adressing is:
«How to apply analytics and machine learning to predict the behavior of chosen components in hydropower plants»

VOITH

Why we chose this assignment

- Challenging
- Combines engineering disciplines
- All new, modern technology
- Developed with open source technology
- Get first-hand experience with big data and machine learning
- Apply our knowledge to a real problem

Methods for analyzing data

- Plotting
- Physical model
- Heat mapping
- Correlation analysis
- Principal component analysis
- Partial least square regression

Correcting sampling frequency

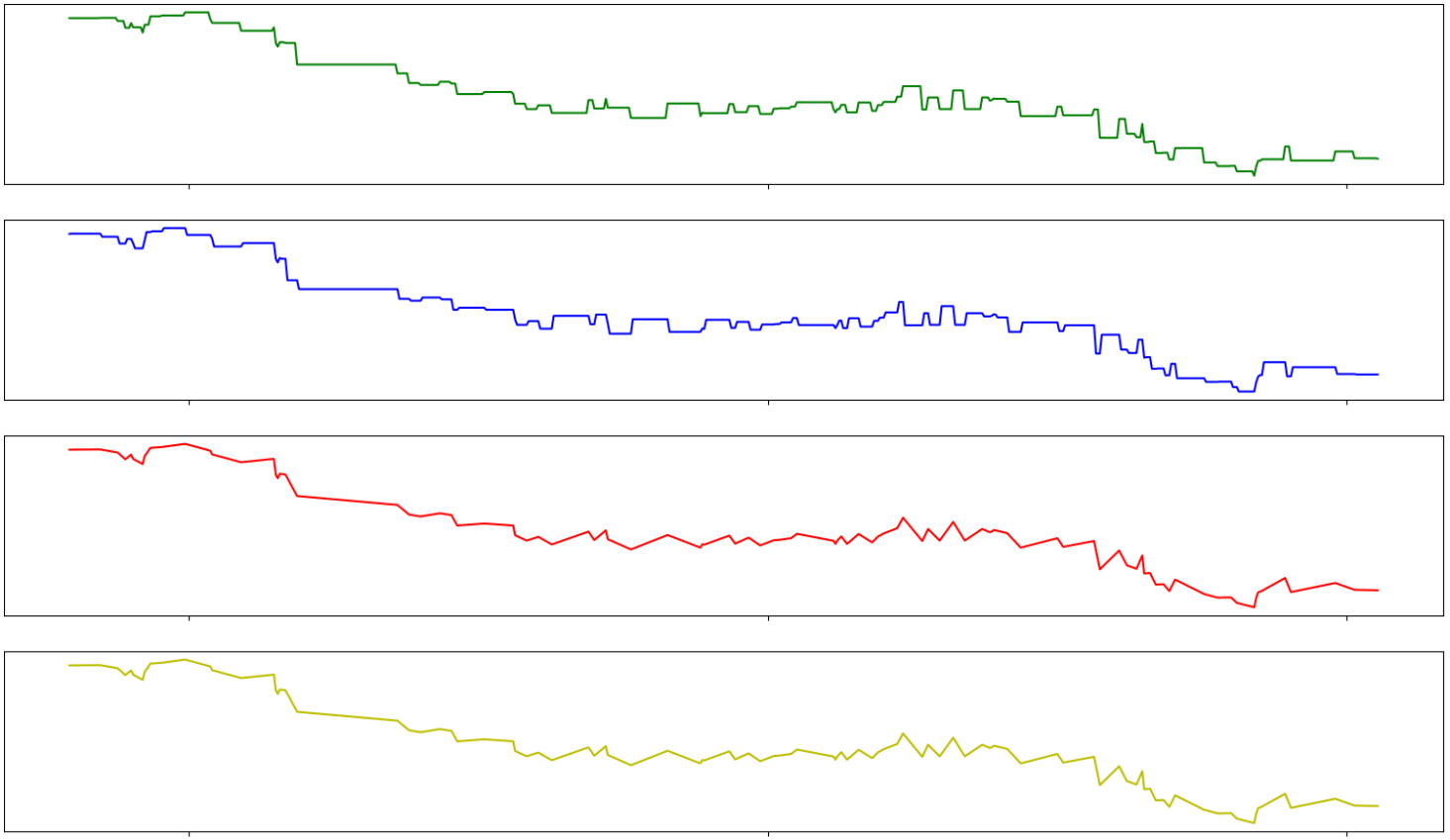
Time	value
23:59:57	690075
00:00:09	692969
00:00:35	695862
00:00:54	698756
00:01:13	701649
00:01:33	704543
00:02:01	707436
00:02:11	710330
00:02:37	713223

Time	value
23:59:57	690075
23:59:58	NaN
23:59:59	NaN
...	...
00:00:09	692969
00:00:10	NaN
..	..
00:00:35	695862
00:00:36	NaN

Time	value
23:59:57	690075
23:59:58	690075
23:59:59	690075
...	...
00:00:09	692969
00:00:10	692969
..	...
00:00:35	695862
00:00:36	695862

Correcting sampling frequency

- Different techniques for filling the gaps in the data
  - Forwardfilling, backwardfilling, linear interpolation and index interpolation



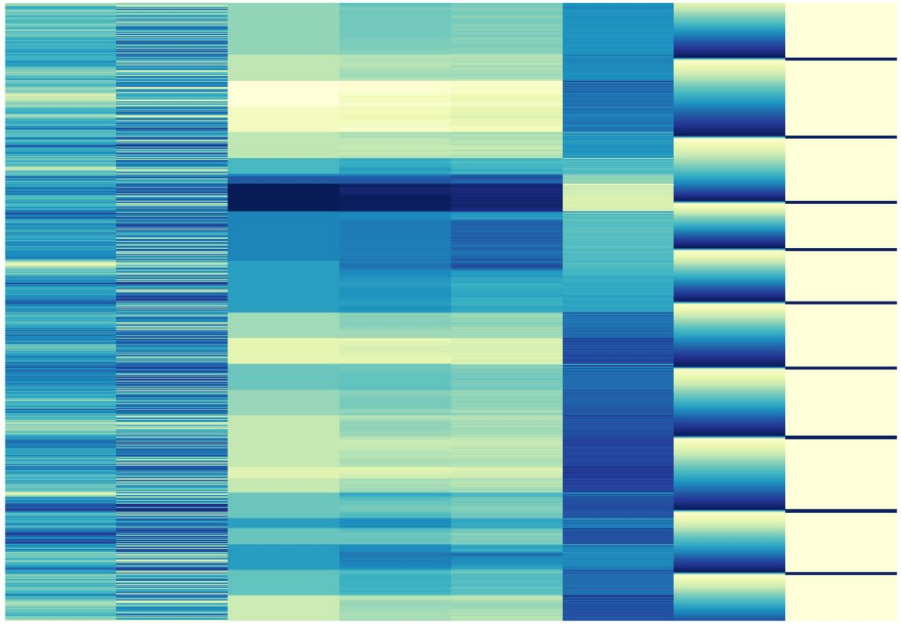
Physical model

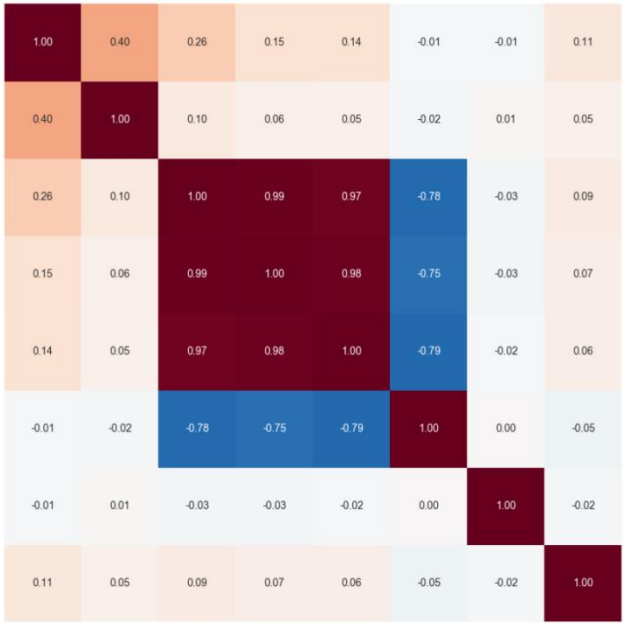
- Can be used as a soft sensor to:
  - Display information
  - Calculate new information
  - Derive signals
- Can be applied to other power plants
- Easy to expand and implement new functionality

Correlation analysis

Heatmapping and correlation matrix

- The heat map shows how the different signals change over time compared to each other
- Correlation matrix shows how they correlate

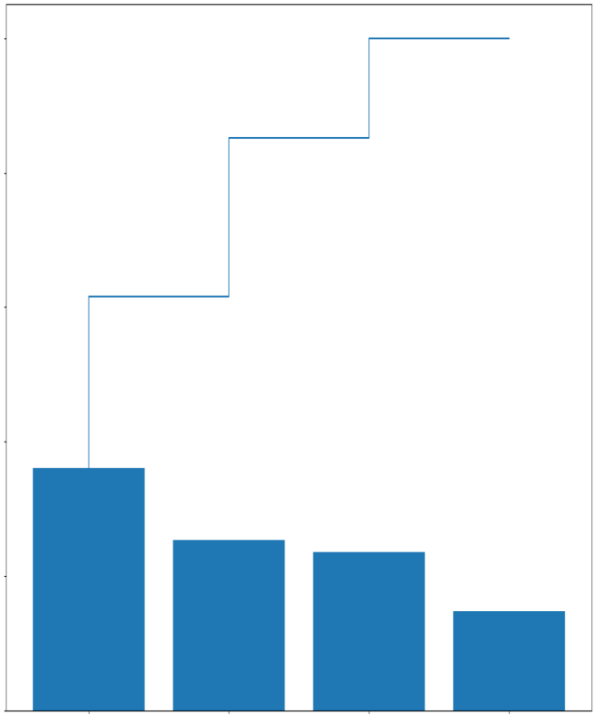


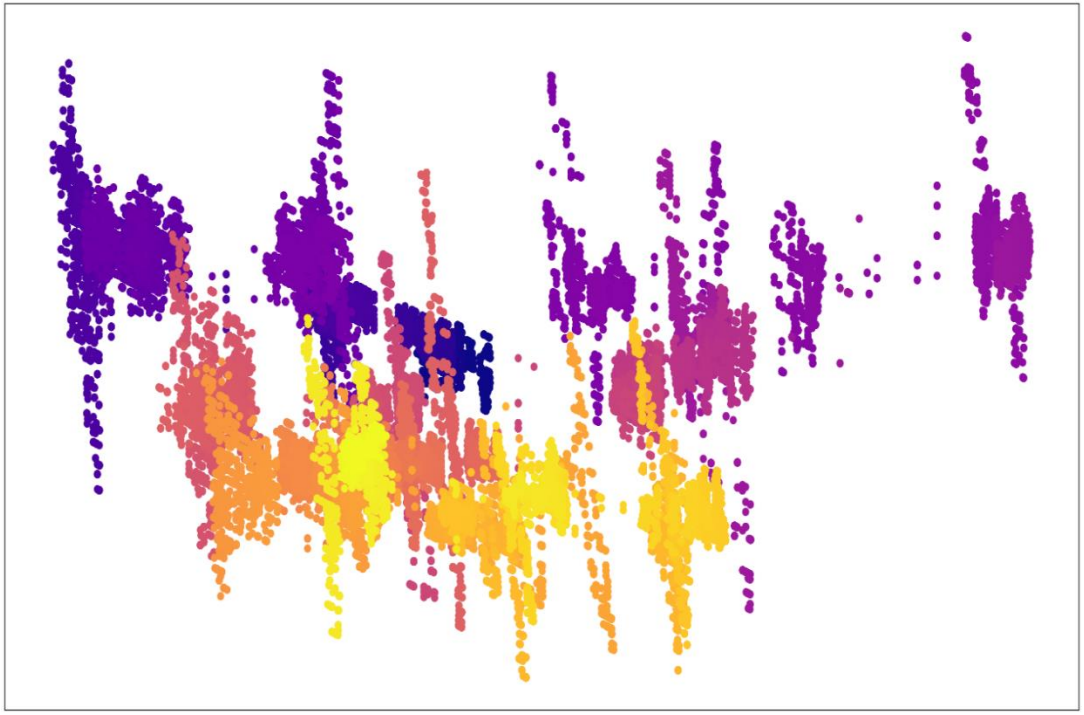


Principal component analysis

Explained variance and scatterplot

- Calculates the signals to a matrix and shows how much information is stored in each component





Partial least square regression(PLSR)

Data structure

Feature 1	Feature 2	Feature 3	Target
28	0,035	262	0,89
28	0,036	23	0,89
...	...	...	...

Training data

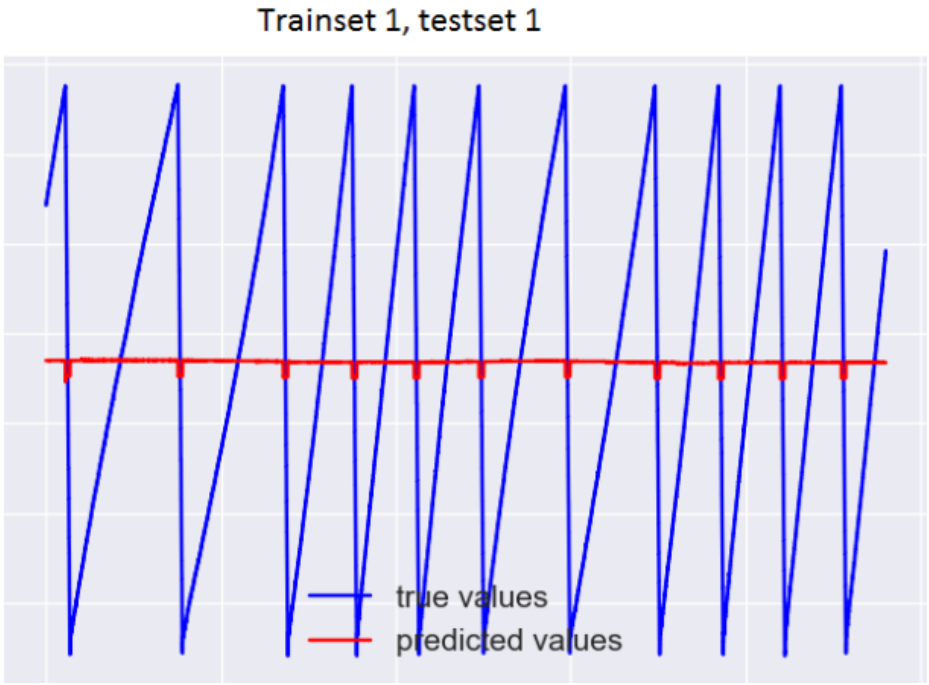
Feature 1	Feature 2	Feature 3
28	0,035	263
28	0,036	266
...	...	...

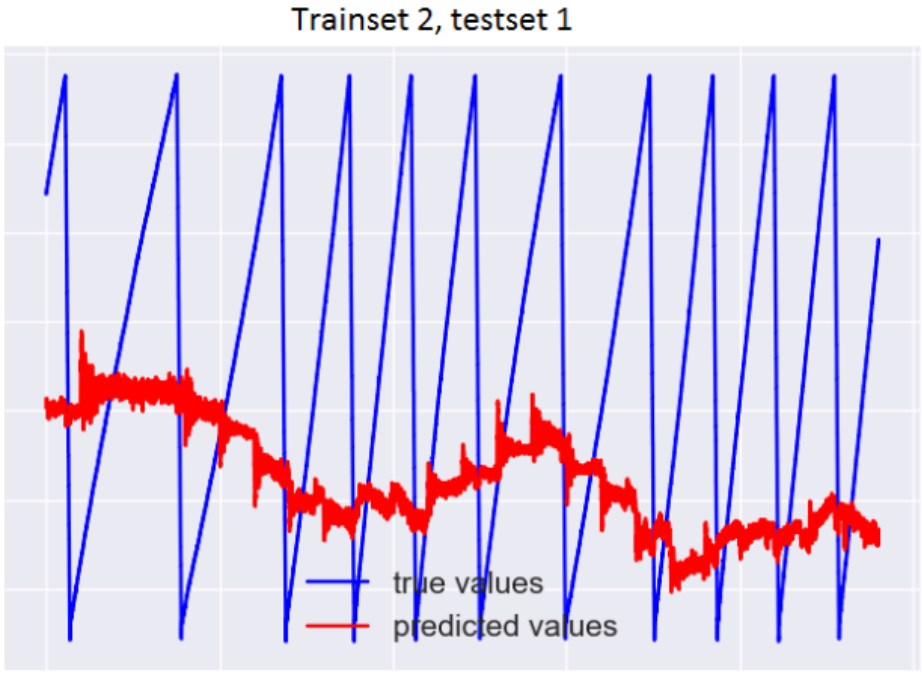
Test data

Target
0,83
0,91
...

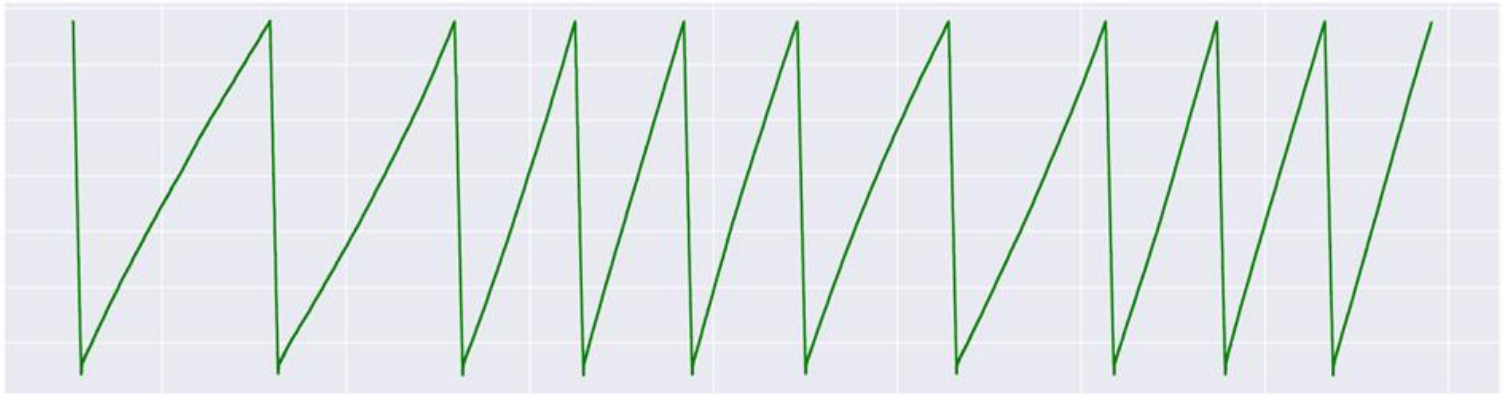
PLSR first try

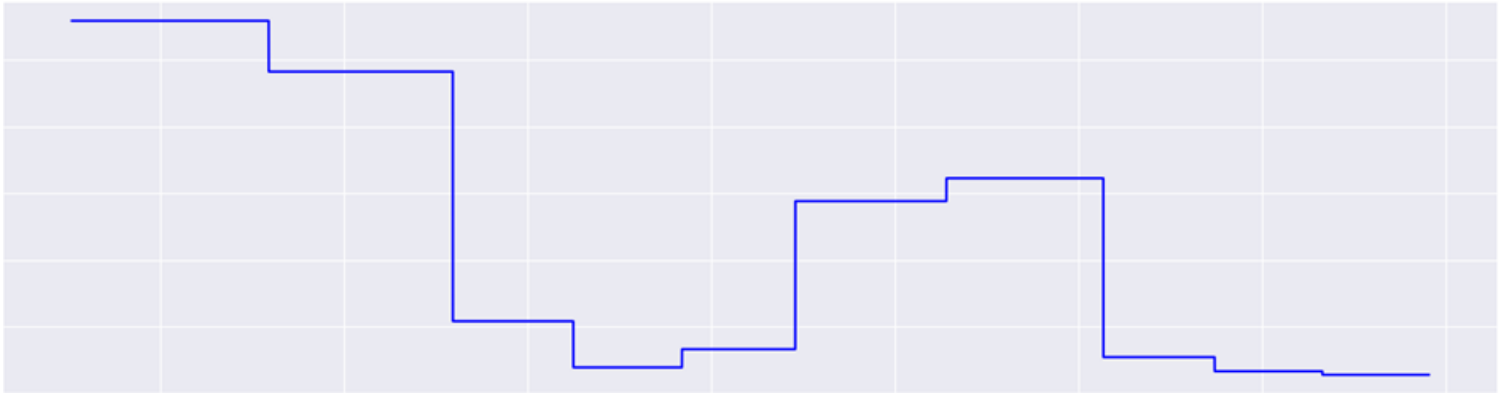
Comparing different datasets





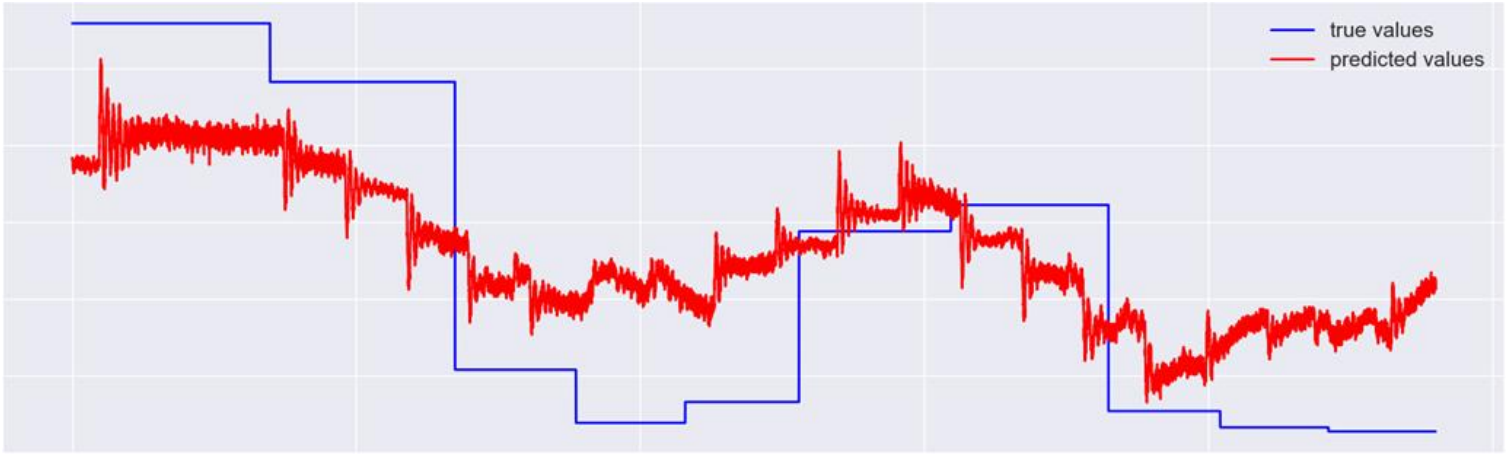
Derive a new signal to refine PLSR performance

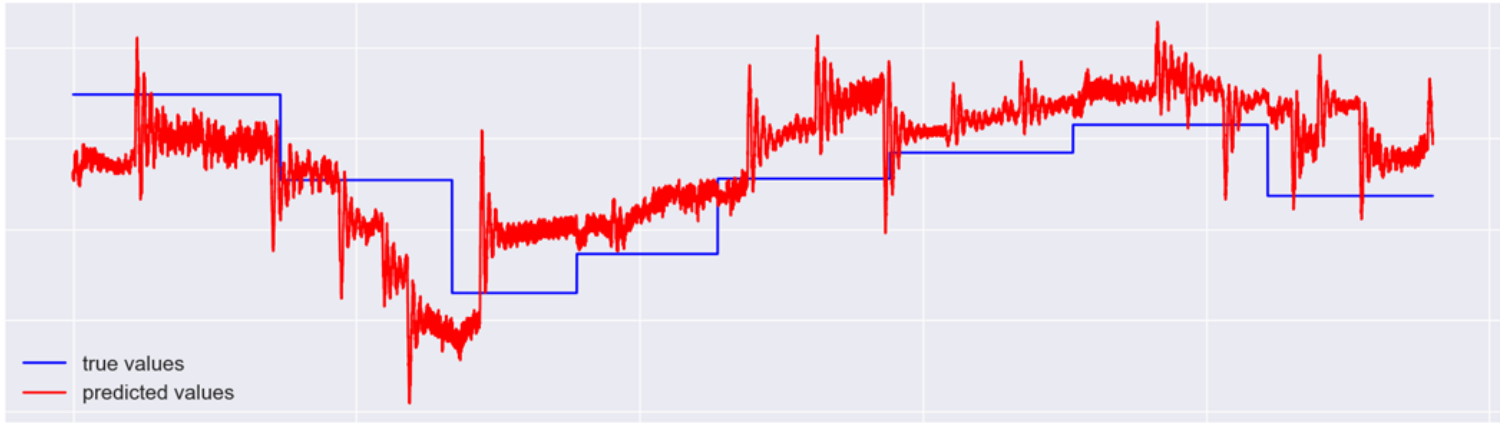




VOITH

PLSR performance on derived signal





Further work

- Try other analysis methods
- Improve results by refining the derived signal
- More functionalities for the physical model
- Refine algorithm
- Predict other signals or values
- Implement cloud processing
- Implement condition based maintenance using machine learning

References

- Voith. (out of date).*Voith i Norge.* Retrieved 26th of April 2017 from http://voith.com/nor-no/index.html