INTERNSHIPS AT ANALYTICAL CHEMISTRY & CHEMOMETRICS

Currently, we are looking for students for the following projects, the length and objectives of this internship can be adapted to bachelor and master internships in consultation with the supervisor at the start of the internship:

1. Anomaly detections in HIS using data inner polygon
2. Predicting properties of aramid yarn using near-infrared spectra
3. Data Analysis for optimization of Redox Flow Battery Systems

Anomaly detection in HIS using data inner polygon

Supervisor: dr. Mahdiyeh Ghaffari

This internship will help to increase your:

- Ability to image pre-processing/processing
- Programming experience (MATLAB)
- Ability to critically review scientific literature
- Ability in big data analysis/visualization

- HyperSpectral Images (HSI)

Hyperspectral imaging measures the spatial and spectral features of any object at wavelengths ranging from ultraviolet through long infrared, including the visible spectrum. A Hyperspectral image (HSI) can include dozens or hundreds of channels unlike RGB images, which use only three types of sensors sensitive to the red, green, and blue portions of the visible spectrum. Hyperspectral images have more information than RGB images. Two identical objects in an RGB image can differentiate in HSI, due to different characteristic properties.

- Anomaly detection

Anomaly detection is an important branch in chemometrics/machine learning and has important applications in computer vision and data mining. The purpose of anomaly detection is to identify a part of the data that is not normal or is significantly different from the rest. Anomaly detection in HSI refers to the identification of pixels whose spectral characteristics in an image are significantly different from global background pixels (it can be a pixel or multiple pixels). In fruit quality control using HSI, anomaly detection can be used to distinguish fresh fruit from damaged ones. Figure 1 presents two strawberries that are fresh (a) and fungi infected (b). This can be an example to confirm/reject the healthiness of and strawberry. In the current example, RGB is enough to distinguish. However, in the case of a minor number of pixels, it will be difficult and HSI is needed. In addition, HSI will help to characterize the type of impurity as well.

Figure 1. a and b present fresh and fungi-infected strawberries.

In this internship, you will have a look at the literature on anomaly detection methods first. Later you will get familiar with Principal Component Analysis (PCA) and data inner polygon in PCA space. Next, you will work on the application of data inner polygon in anomaly detection in HSIs. Finally, you will use the modified/generated algorithm in real cases/HSIs. All these analyses will be automated in MATLAB.

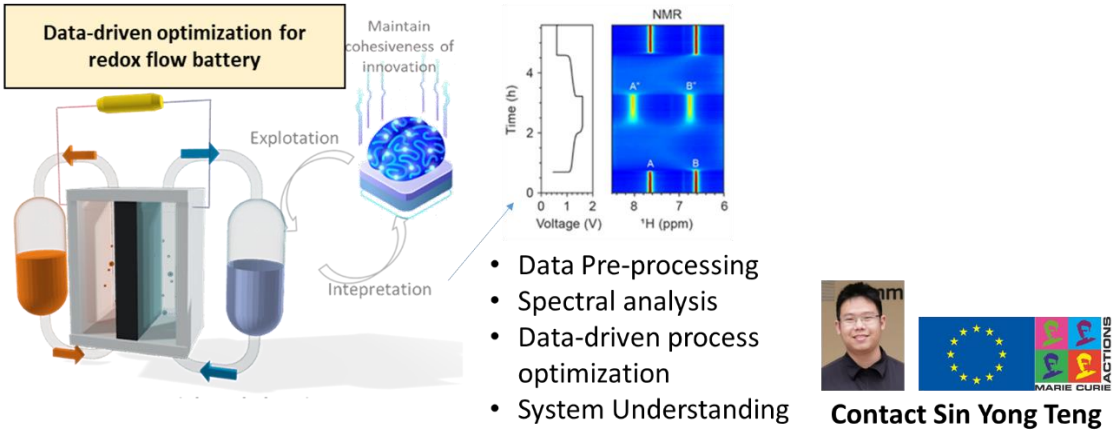
Predicting properties of aramid yarn using near-infrared spectra

Supervisor: dr. Mahdiyeh Ghaffari

This internship is performed in collaboration with Teijin Aramid BV (https://www.teijinaramid.com/), a company with headquarters in Arnhem that produces aramid fibers for high performance applications (bullet-proof vests, reinforcement for composites in tires, aerospace, construction, etc). The objective of this study is to predict the mechanical properties of Twaron yarn from Near InfraRed (NIR) spectra measured at the plant using various types of chemometric models. These models will also have to account for the variation of the NIR spectra with time, for which a few sets of repeated measures were recorded. The project might also include the estimation of the implications of including the resulting NIR-based models in quality control.

Figure 1. Data measurement over Twaron yarn production.

Data analysis for optimization of Redox Flow Battery Systems



Supervisor: dr. Sin Yong Teng

battery systems. Redox flow batteries represent a promising technology for energy storage due to their scalability and flexibility. However, their efficiency and performance can be further enhanced through systematic data-driven optimization. This internship includes skills such as data analysis, including pre-processing, spectral analysis, data-driven process optimization, and system understanding.

The first phase involves data pre-processing, where raw data from redox flow battery experiments are cleaned, normalized, and organized for further analysis. This step ensures that the data are reliable and suitable for subsequent analytical processes. Following pre-processing, spectral analysis techniques are applied to extract meaningful insights from the data. This includes identifying key spectral features related to battery performance and understanding their significance in the context of system optimization.

The internship also focuses on data-driven process optimization, where advanced statistical and machine learning methods are employed to develop predictive models for optimizing redox flow battery operation. These models utilize historical/NMR data to identify optimal operating conditions, maximizing efficiency and performance while minimizing costs and resource consumption.

If you are interested, please contact: sinyong.teng@ru.nl

Investigating the Impact of Uncertainty in Multivariate Analysis of Operational Data

In collaboration with Royal Cosun

Supervisor: dr. Mahsa Akbari Lakeh

Objective:

Investigate the application of process variable uncertainties in multivariate process control and quality predictions.

Introduction:

In process monitoring, traditional statistical methods like Principal Component Analysis (PCA) are widely used. However, these methods often assume that noise in process variables is equal and independent, which isn't always true for complex processes. This can lead to inefficient and unreliable monitoring.

To tackle this, Maximum Likelihood Principal Component Analysis (MLPCA) has been developed. MLPCA incorporates noise structure information, making it more robust for data with unequal and correlated noise. Additionally, combining MLPCA with Multivariate Curve Resolution by Alternating Least Squares (MCR-ALS) allows for better handling of complex noise structures and more accurate process monitoring.

However, a challenge remains: MLPCA requires a known error covariance matrix, which is often not available. Your task will be to explore methods for estimating uncertainty in operational data to improve the application of MLPCA in real-world scenarios.

What You'll Do:

- Investigate existing methods for estimating uncertainty in operational data.
- Apply MLPCA and MLPCA-MCR-ALS to process data, incorporating these uncertainties.
- Analyze the impact of considering uncertainties on process monitoring and quality predictions.

What We're Looking For:

- Background in statistical methods and process control.
- Analytical skills and familiarity with multivariate analysis techniques.
- Programming skills in Matlab or Python.
- Ability to work independently and collaboratively.

What You'll Gain:

- hands-on experience with advanced statistical methods.
- Insight into real-world applications of process control.
- Collaboration with industry experts at Royal Cosun.