Contribution ID: **93** Type: **not specified**

# Expert and data-driven approaches for big data analysis of industrial purification process

Nowadays, big data is generated real-time in the majority of industrial production processes. Happenstance data is characterized by high volume, variety, velocity and veracity (4v of big data).
In this study production data from industrial purification process is analyzed to assess process performance and its relations with product quality. For this purpose, a comprehensive data preprocessing strategy is developed first. It comprises of data driven methods such as parametric time warping for time profile synchronization, definition of process steps, outlier detection and missing values removal. Importantly, also expert-driven steps are applied to extract more than 200 relevant features for different process variables and different process steps.
Finally, different combinations of preprocessed process data from multiple runs, which are contributing to single batch quality data, are analyzed with univariate and multivariate statistical tools. This allows to explore and visualize relationships between process and product in efficient manner to ease process understanding and monitoring.

## Special/ Invited session

## Classification

Mainly application

## Keywords

industrial process data, process monitoring, process improvement, correlation, time-resolved data

**Primary authors:** Ms LEUSDEN, Irene (FrieslandCampina); Dr LI, WeiWei (FrieslandCampina); Mr VAN MIDDELKOOP, Bart (FrieslandCampina); SZYMANSKA, Ewa (FrieslandCampina)

**Presenter:** SZYMANSKA, Ewa (FrieslandCampina)

**Track Classification:** Statistical Process Monitoring