A chemometric approach to the optimization of bio-industrial processes

Anna Klimkiewicz[1,2,*] and Frans W.J. van den Berg[2]

1) Novozymes A/S, Kalundborg, Denmark

2) University of Copenhagen, Denmark

*) akcz@novozymes.com

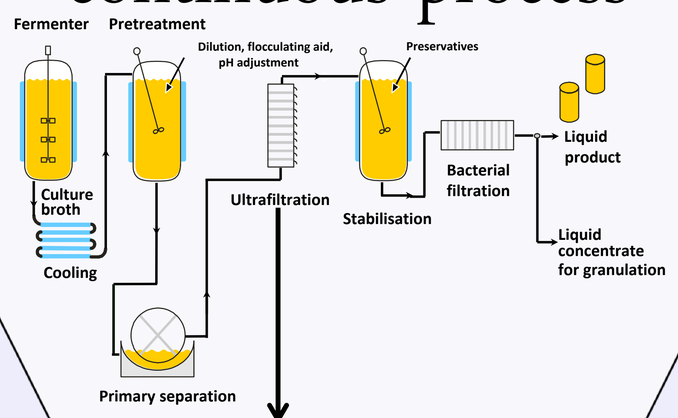
Introduction

Massive amounts of *process tags* are collected with high frequency in biomanufacturing for specific and dedicated univariate monitoring tasks and closed-loop control applications. This produces large amounts of data which are seldom used outside their direct scope. It can however be of great benefit to combine all available data, structure it and extract relevant information from the full set via process chemometrics tools. In this work we highlight the value of historical production data and multivariate statistical data analysis as a starting point of variance reduction and capacity optimization in an industrial-scale continuous enzyme purification process.

Conclusions

The problem of flux decline in ultrafiltration is very complex and multifaceted, even more so in case of a full scale bio-industrial recovery process with numerous sources of variation. The method of connecting historical data from continuous processes to the flux quality parameters is a subject of ongoing research. Process chemometric tools applied on these large amounts of production data offer a good overview during exploratory problem solving, for instance, help optimization engineers to focus attention on suitable target areas or to asses the effect when manipulating process parameters upstream so that no adverse effect downstream is observed.

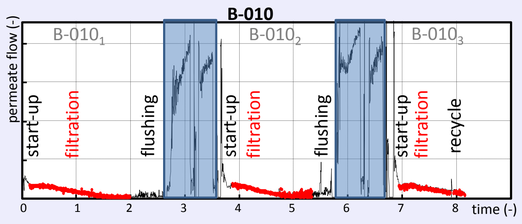
Enzyme recovery, an (almost) continuous process



What's a batch?

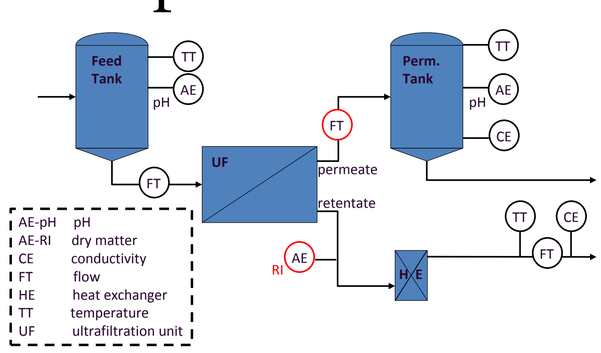
UF *batches* vary considerably in length and volume

- **Operation mode:** discontinuous feed supplied to the membrane plant



- **Unit specific nuances** - i.e. start-up phase finishes when dry matter set-point is reached in the retentate stream leaving the UF
- **Filtration sequences**: separation based on Cleaning-In-Place demands → batch

UF capacity optimization



- **Problem**: *UF flux* (permeate volume per unit area of the membrane per unit of time) varies for both the same product and between different products.
- **Aim:** understand *batch-to-batch* variations in the performance of ultrafiltration.

Query the historian

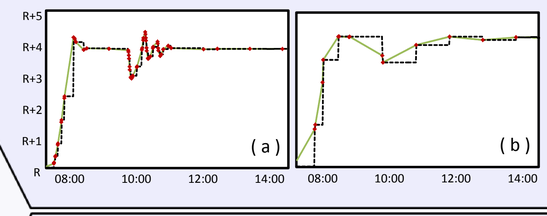
Data can be compressed at any or all of the following steps:

Sensor → Process Control System → Interface Node → Historian

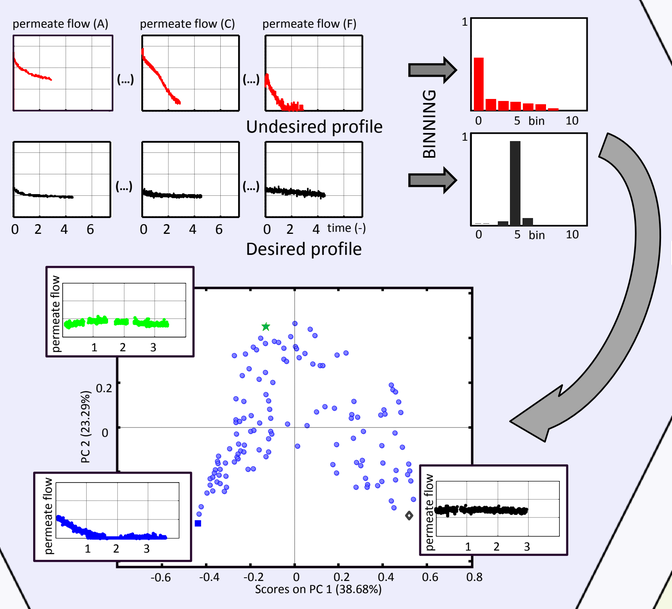
Three most relevant functions for data retrieval:

- **Sampled** (–) interpolated between the two last logged points (equidistant)
- **Archived** (–) return the previous logged value (equidistant)
- **Compressed** (♦) only the true logged points (non-equidistant)

Example - two on-line dry matter measurements and the effect of the dead band settings, 0.2 R (a) and 1.0 R (b)



How to qualify the flux?

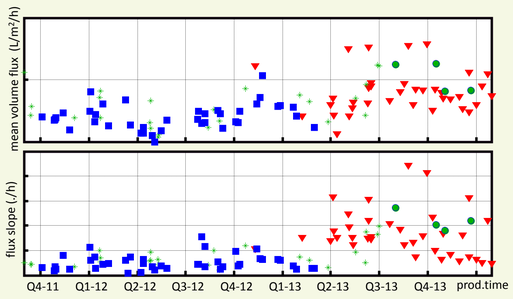


Comparing batches

- Mathematical basis set

Flux as a function of batch production (calendar) time

→Higher concentrated group

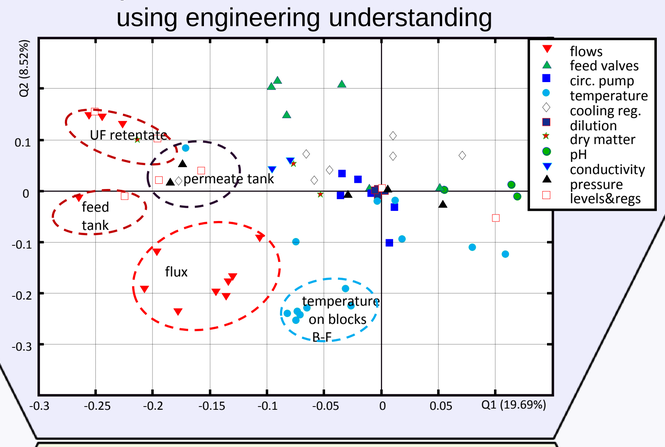


= Introduction of changes in fermentation and recovery SOPs

How to select sensors?

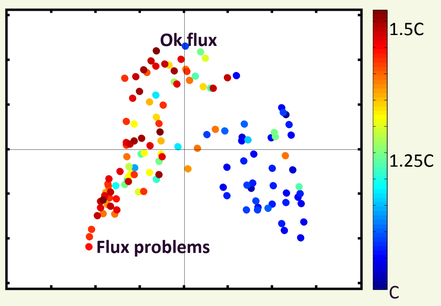
Unsupervised methods for batches varying in length:

- Multilevel Simultaneous Component Analysis (MLSCA)
- Common Components and Specific Weights Analysis (CCSWA)
  - → plot similarities between process tags
  - → analysis for clusters and outliers using engineering understanding

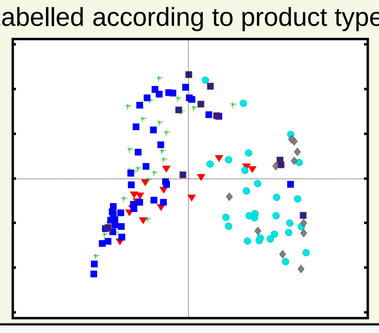


• Flux quality map

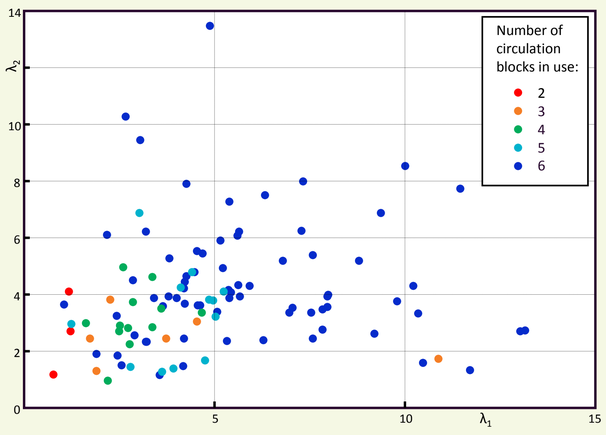
→ PC1 is related to concentration factor



→ Labelled according to product type

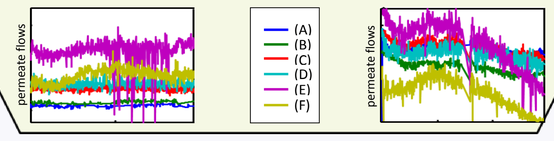


• CCSWA



Saliences = How representative is a batch for the *average batch*

↓ λ ↑ λ



novozymes®

Rethink Tomorrow