CHRISTOPH BERG[1],
NINA IHLING[1],
JOCHEN BÜCHS[1]

Aachener Verfahrenstechnik | RWTH AACHEN UNIVERSITY

COMBINATION OF MULTIVARIATE DATA ANALYSIS AND 2D-FLUORESCENCE SPECTROSCOPY IN HIGH-THROUGHPUT CULTIVATION EXPERIMENTS

MOTIVATION

- High dimensional data from fluorescence and scattered light measurements of bioprocesses demand advanced data processing techniques (PCA, PLS)
- Establishment of both data processing and data selection workflows for batch cultivations in high-throughput experiments needed
- Evaluation of workflow for cultivation of *H. polymorpha* under phosphate limitation

EXPERIMENTAL SETUP

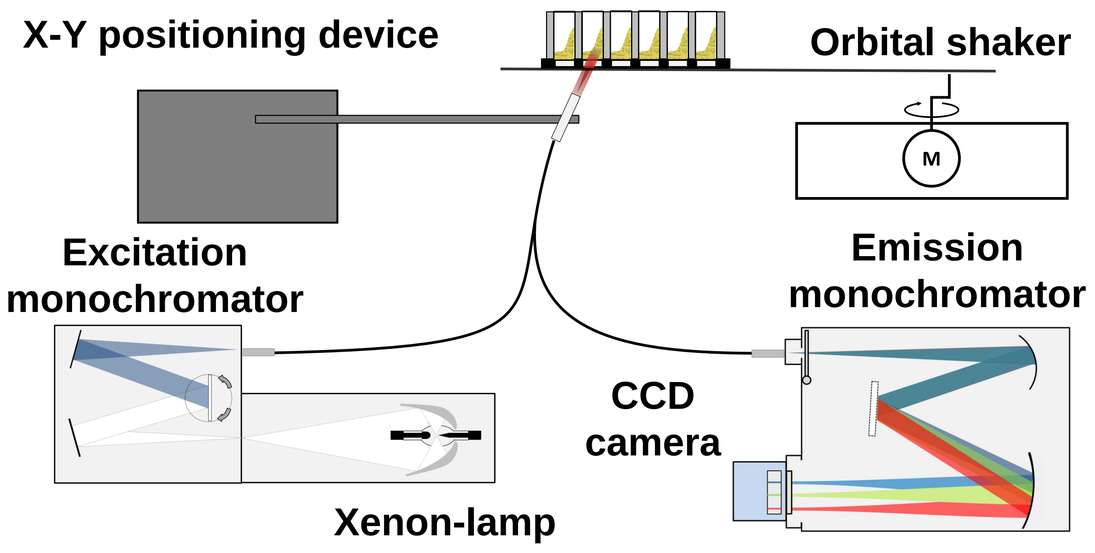


Figure 1: Schematic setup implementing 2D-spectroscopy in a 48-well microtiter plate[b]

RESULTING SPECTROSCOPICAL 2D-DATA

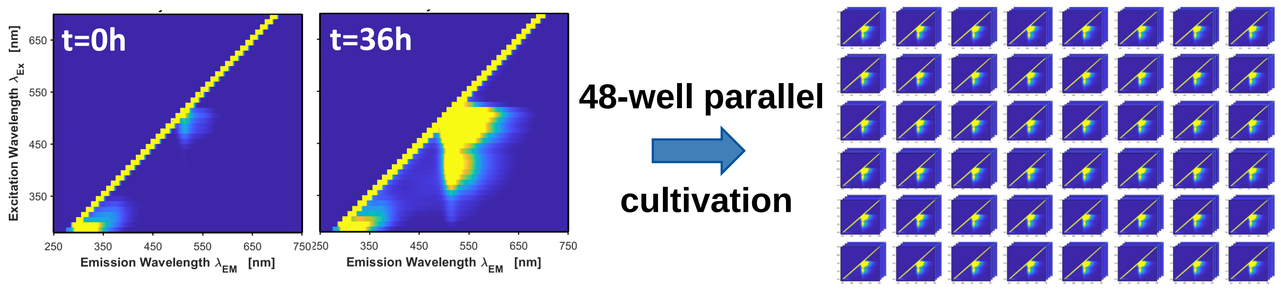


Figure 2: 2D-spectra from cultivation of *H. polymorpha* expressing GFP under phosphate-limitation[a]

Cultivation conditions: 48-well microtiter plate, V_L= 800 µl, T= 30 °C, n= 1000 rpm, d_0= 3 mm,

RESULTS AND DISCUSSION

Principal component analysis (unsupervised)

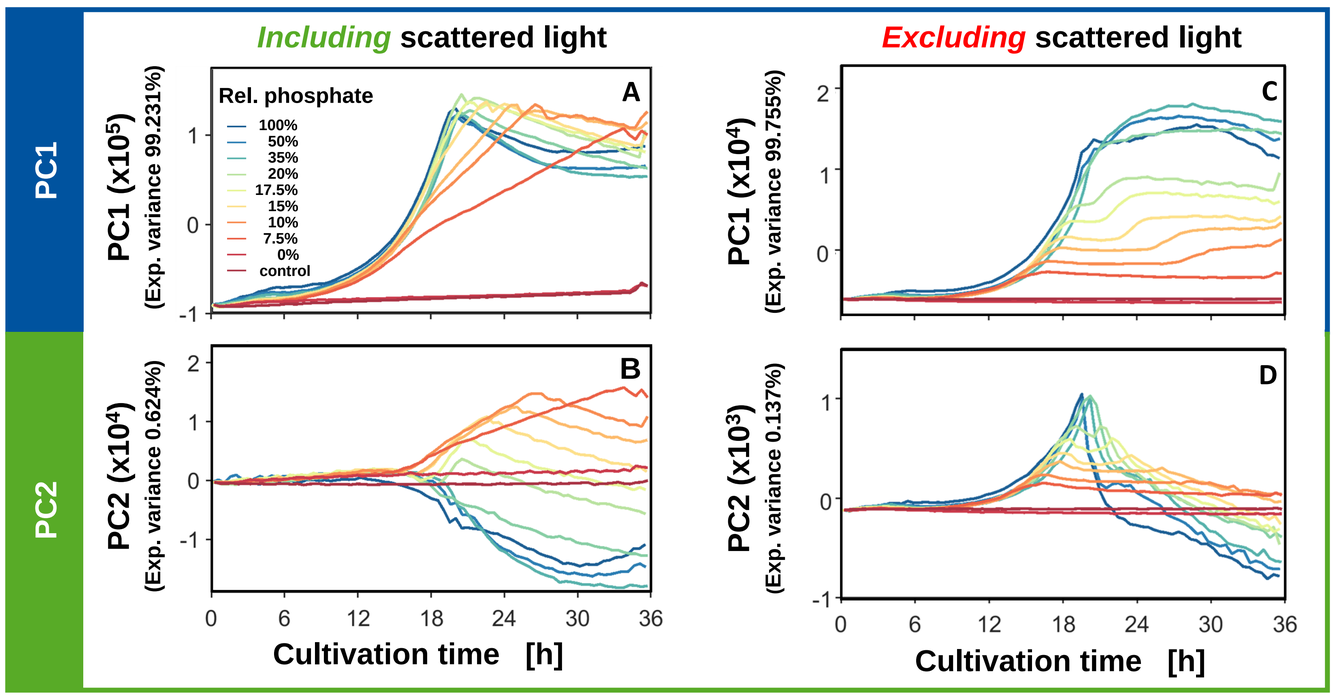
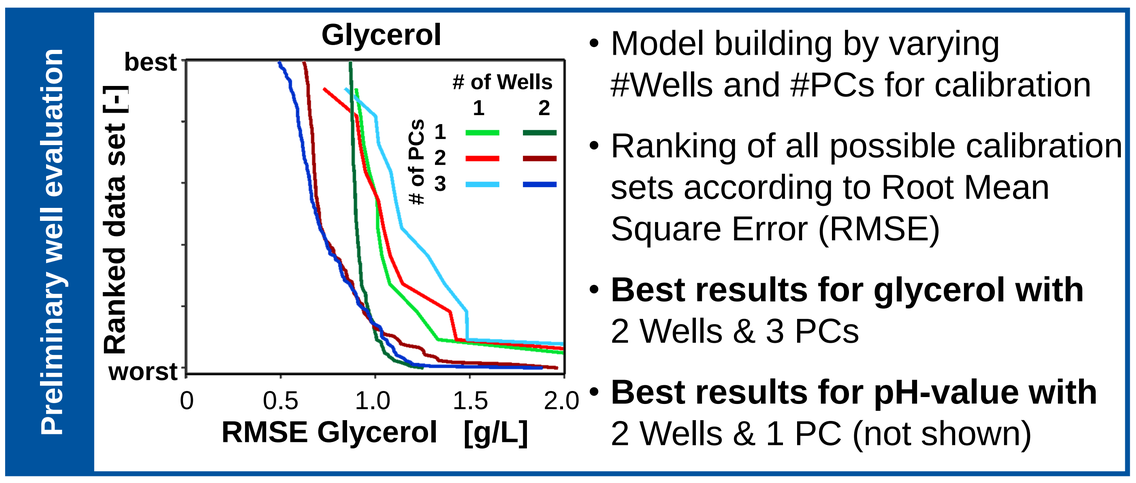


Figure 3: Scores of 2D Spectra A, B) including scattered light or C, D) excluding scattered light over cultivation time after applying principal component analysis on well-resolved data.

- PCA results vary depending on input data set:
- ***Results including* scattered light:** mostly affected by backscattered light
- ***Results excluding* scattered light:** only affected by fluorescing compounds
- Data de-correlation facilitates the interpretation of online signals

  → *i.e.* PC2 qualitatively helps identifying limitation conditions

Partial least square regression (supervised)



- Model building by varying #Wells and #PCs for calibration
- Ranking of all possible calibration sets according to Root Mean Square Error (RMSE)
- **Best results for glycerol with** 2 Wells & 3 PCs
- **Best results for pH-value with** 2 Wells & 1 PC (not shown)

Figure 5: Evaluation of calibration sets for glycerol prediction

PLS model building

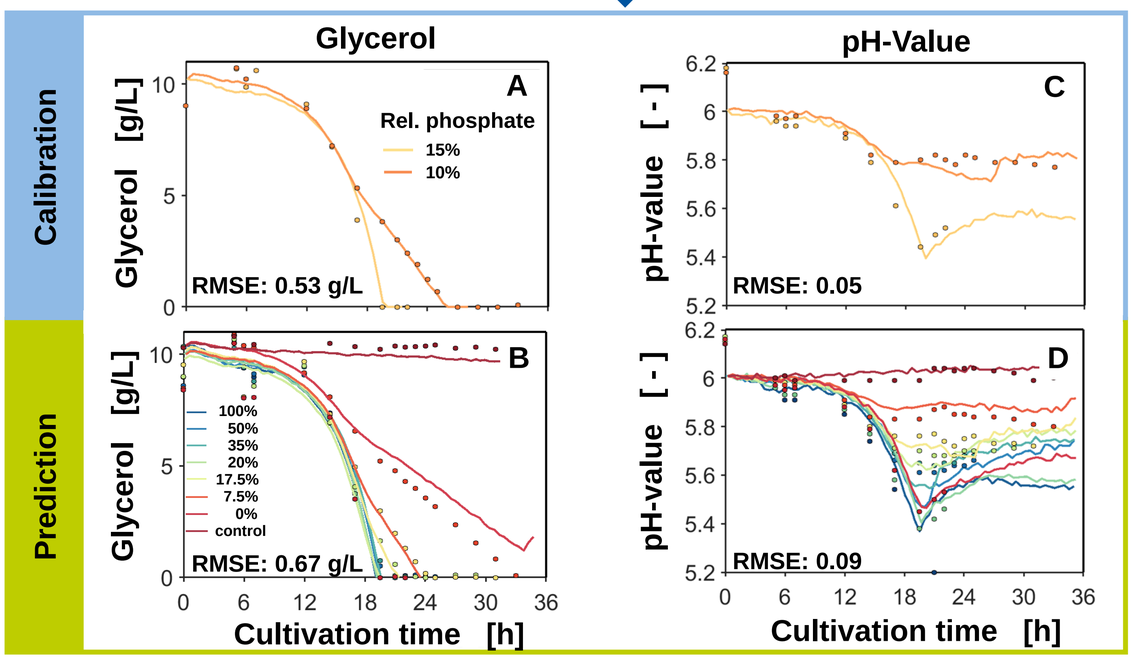


Figure 6: Calibration and prediction of A, B) Glycerol and C, D) pH-value using calibration well combinations according to preliminary evaluation method (Figure 5). ● Offline measurements, ─ Predicted values based on 2D-Spectra and PLS

CONCLUSION AND OUTLOOK

- Evaluation of bioprocess data quality using unsupervised learning is beneficial for process understanding and interpretation
- Application of supervised methods successful: Prediction of strongly correlated offline values *i.e.* glycerol and pH possible
- Outlook: Modelbuilding for non-biomass related signals such as non-fluorescent product formation

Correspondence:

[1]Aachener Verfahrenstechnik – Biochemical Engineering
RWTH Aachen University
Forckenbeckstraße 51
D-52074 Aachen

References:

a) **Kottmeier 2009, Biotechnol. Bioeng., 104: 554–561.**

b) **Ladner et al. 2016, Biotechnol. J., 11: 1605–1616.**

Acknowledgement:

www.iconsensus.eu

www.imi.europa.eu

We acknowledge the Innovative Medicines Initiative 2 Joint Undertaking [grant agreement No 777397]. This Joint Undertaking, project iConsensus, receives support from the European Union's Horizon 2020 research and innovation programme and EFPIA Partners Sanofi, GSK, Bayer, Rentschler Biopharma, UCB, Byondis and Pfizer.