From screening to production: a holistic approach of high-throughput model-based screening for recombinant protein production

Niels Krausch,[a] Jong Woo Kim[b], Sebastian Hans,[a,b] Felix Fiedler,[c] Sergio Lucia,[b,c] Peter Neubauer,[a,b] Mariano N. Cruz Bournazou[b,d]

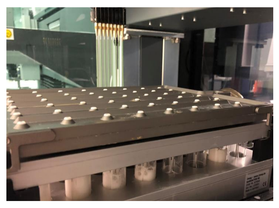
[a]Technische Universität Berlin, Bioprocess Engineering, Ackerstrasse 76, 13355 Berlin, Germany
[b]Technische Universität Berlin, KIWI Biolab, Ackerstrasse 76, 13355 Berlin, Germany
[c]Technische Universität Berlin, Internet of Things for Smart Buildings, Einstein Center Digital Future, Einsteinufer 17, 10587 Berlin, Germany
[d]DataHow AG, Vladimir-Prelog-Weg 1, 8093 Zurich, Switzerland

Abstract

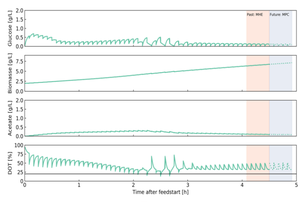
Efficient and robust screening of production strains in early bioprocess development is usually very time-consuming and laborious. Even though the introduction of high-throughput liquid handling stations allow a larger number of strains to be tested in parallel, it is not possible to gain insight into the dynamical phenotype of the strains [1]. Following upon our previous work of establishing a robotic high-throughput cultivation platform [2], we introduced a model predictive control feeding regime to the cultivations which exposes the cells to stress conditions similar to those present in large-scale bioreactors with the goal to obtain highest biomass at the end of the process and tested this approach with *in-silico* data. This approach helps in identifying the best strain for usage in industrial-scale bioreactors.

Concept and results

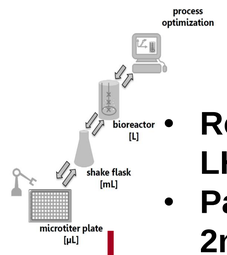
Parallel cultivation



MHE/MPC

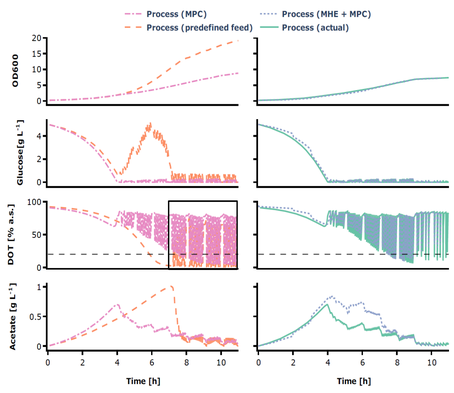


Scale down



- **Robotic HT facility consisting of 2 LHS**
- **Parallel *E. coli* cultivations in 2mag Minibioreactors**
- **Dynamical parameter estimation**
- **Optimal feeding regime to screen for best growing strain**

Strain discrimination



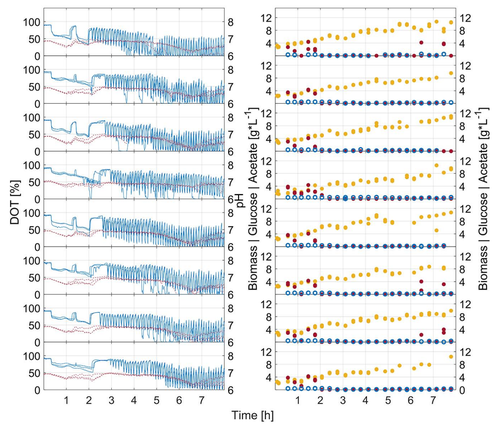
Monitoring via:

- Online and at-line Measurements of DOT, pH, OD_{600} Glucose and Acetate
- Data storage in central database

Process control via:

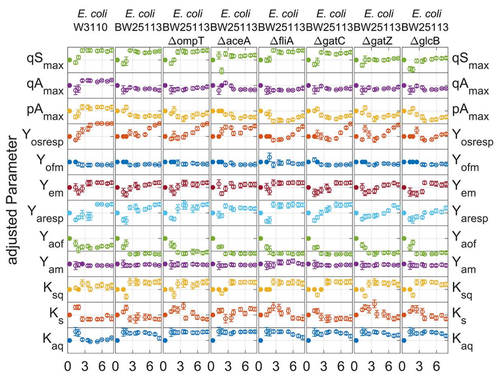
- Setpoints for feed, temperature, aeration, pH and stirring speed are stored in the database
- New setpoints for the feed can be based on future simulations, considering recent measurements

Process measurements



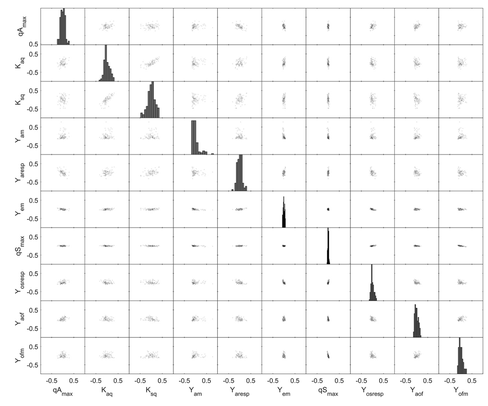
Experimental data of the different cultivations running in parallel. Left: DOT and pH measurements (online). Right: Measurements of Biomass, Glucose and acetate (at-line).

Parameter estimation



Results from the parameter estimation over time after feed start for the different strains. Values are normalized to their initial values. Only a subset of the complete parameter set was estimated for the sake of identifiability.

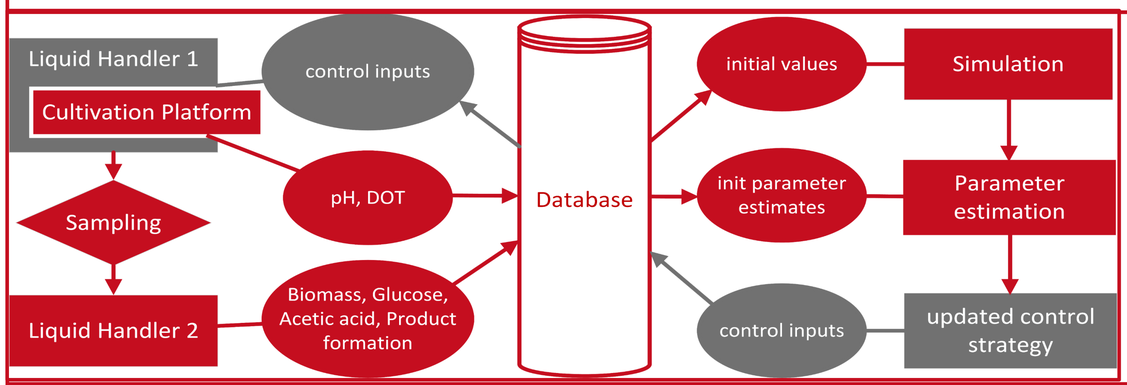
Parameter uncertainties



Uncertainty distribution and correlation analysis of the different parameters using 500 Monte Carlo simulations, which were carried out with σ = 0.15 for biomass, glucose, acetate and σ = 0.05 for DOT.

Nonlinear Model Predictive Control approach

Eight *E. coli* K-12 strains were cultivated in 3 parallels (=24 minibioreactors) with an industrial process-relevant feeding design (batch and subsequent exponential fed-batch phases). The mathematical model describing the cultivations is shown below. The dynamic states are denoted by $x_r \in \mathbb{R}^{n_x}$ and include biomass, substrate, dissolved oxygen tension (DOT) as well as acetate. The control inputs for each mini bioreactor are $u_r \in \mathbb{R}^{n_u}$ and $\theta \in \mathbb{R}^{n_\theta}$ denote the unknown parameters of the strains and the respective bioreactor. Measurements for all the states were obtained for each sampling at the following time intervals: 30 sec for DOT, and 20 min for the other states. At each iteration step, the inputs of the system, i.e. optimal feeding rates for each bioreactor, were computed with an MPC controller, using the dynamical model with the current estimation of the uncertain parameters $\hat{\theta}$. The MPC controller ensured that all strains were cultivated based on their maximum capabilities to assure that the best performing strain is selected for further process development. Since estimation of the parameters plays an important role in the computation of the optimal inputs, and as they are known to vary during cultivation, the parameters were iteratively estimated using a moving horizon approach, which uses a sliding time window and only considers the N_{mhe} last measurements.



Bioreactor dynamics:

$$\dot{x_r}(t) = f(x_r(t), u_r(t), \theta) \quad (1)$$
$$x_r(t_0) = x_{0,r}$$

MHE equations:

$$\min_{\theta} \left\| \hat{\theta}_{old} - \hat{\theta} \right\|^2_{W_p} + \sum_{k=0}^{N_{MHE}} \left\| h\big(x(k)\big) - y(k) \right\|^2_{W_y} \quad (2)$$

$$\dot{x_r}(t) = f\big(x_r(t), u_r(t), \hat{\theta}\big) \forall r \in \Re$$

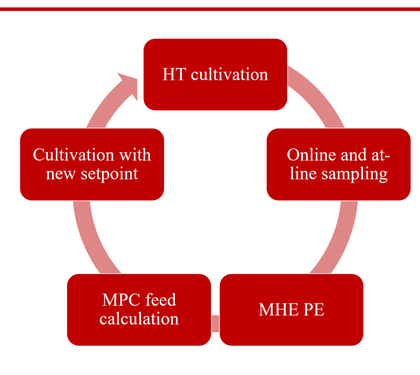
s.t. $$x_r(t_0) = x_{0,r} \quad (3)$$
$$\theta_{min} \leq \theta \leq \theta_{max}$$

MPC equations:

$$\min_{u_r} \sum_{r=1}^{N_{MPC}} -X_r(t_{end}) \quad (4)$$

$$\dot{x_r}(t) = f\big(x_r(t), u_r(t), \hat{\theta}\big) \forall r \in \Re$$

s.t. $$x_r(t_0) = x_{0,r} \quad (5)$$
$$\mathrm{DOT}_r(t) \geq 20\%$$



Conclusion

New strains could successfully be screened at their maximum growth rates, due to the application of a model predictive control approach. It can be seen, that the adaptive MPC feeding regime combined with an MHE approach to estimate the parameters of the different strains is superior to a classical fixed feeding regime, preventing overfeeding and limiting exposure of the cells to anoxic conditions. This framework will soon be further evaluated experimentally and extended to optimize product formation.

- MPC approach successful in minimizing overfeeding
- Dynamical estimation of model parameters enables adaptive control regime
- Further research necessary to consider product formation

Acknowledgements / References

[1] Haby, B., Hans, S., Anane, E., Sawatzki, A., Krausch, N., Neubauer, P., & Bournazou, M. N. C. (2019) Integrated Robotic Mini Bioreactor Platform for Automated, Parallel Microbial Cultivation With Online Data Handling and Process Control. *SLAS TECHNOLOGY: Translating Life Sciences Innovation*,

[2] Anane, E., Neubauer, P., & Bournazou, M. N. C. (2017). Modelling overflow metabolism in Escherichia coli by acetate cycling. *Biochemical engineering journal*, *125*, 23-30.

We gratefully acknowledge the financial support of the German Federal Ministry of Education and Research (BMBF) (project no. 031L0018A - LEANPROT and 01QE1957C - BioProBot).

Institute of Biotechnology
Technische Universität Berlin
Bioprocess Engineering