

Optimal operation of parallel mini-bioreactors in bioprocess development using Multi-Stage MPC

Niels Krausch,^a Jong Woo Kim,^a Sergio Lucia,^b Sebastian Groß,^c Tilman Barz,^a
Peter Neubauer,^a Mariano N. Cruz Bournazou,^{a,d}

^a*Technische Universität Berlin, Bioprocess Engineering, Ackerstrasse 76, 13355 Berlin, Germany*

^b*Technische Universität Dortmund, Process Automation Systems, Emil-Figge-Str. 70, 44227 Dortmund, Germany*

^c*WEGA Informatik (Deutschland) GmbH, Obere Schanzstrasse 1/10, 79576 Weil am Rein, Germany*

^d*DataHow AG, Zürichstrasse 137, 8600 Dübendorf, Switzerland*
n.cruz@datahow.ch

Abstract

Bioprocess development is commonly characterized by long development times, especially in the early screening phase. After promising candidates have been pre-selected in screening campaigns, an optimal operating strategy has to be found and verified under conditions similar to production. Cultivating cells with pulse-based feeding and thus exposing them to oscillating feast and famine phases has shown to be a powerful approach to study microorganisms closer to industrial bioreactor conditions. In view of the large number of strains and the process conditions to be tested, high-throughput cultivation systems provide an essential tool to sample the large design space in short time. We have recently presented a comprehensive platform, consisting of two liquid handling stations coupled with a model based experimental design and operation framework to increase the efficiency in High Throughput bioprocess development. Using calibrated macro-kinetic growth models, the platform has been successfully used for the development of scale-down fed-batch cultivations in parallel mini-bioreactor systems. However, it has also been shown that parametric uncertainties in the models can significantly affect the prediction accuracy and thus the reliability of optimized cultivation strategies. To tackle this issue, we implemented a multi-stage Model Predictive Control (MPC) strategy to fulfill the experimental objectives under tight constraints despite the uncertainty in the parameters and the measurements. Dealing with uncertainties in the parameters is of major importance, since constraint violation would easily occur otherwise, which in turn could have adverse effects on the quality of the heterologous protein produced. Multi-stage approaches build up a scenario tree, considering the uncertainties and calculate optimal inputs considering the model uncertainties. Using the feedback information gained through the evolution along the tree, the control approach is significantly more robust than standard MPC approaches. We show in this study that the application of multi-stage MPC can increase the number of successful experiments, by applying this methodology to a mini-bioreactor cultivation operated in parallel.

Keywords: high-throughput, model-predictive control, scale-down, multi-stage

1. Introduction

1.1. High-throughput bioprocess development

The development of a process in biotechnology is enormously time-consuming. Not only does it take a long time to find a suitable strain from a huge collection of strains, but the identification of optimal process conditions for this strain also takes a lot of time and resources. Especially due to the increased demand for resource-saving and environmentally friendly bioproducts, the requirements for faster process development are also increasing. Miniaturization and parallelization are two recent technologies that can accelerate throughput in screening processes and process development (Hemmerich et al., 2018). In particular, the use of liquid handling stations and the application of computer-aided and model-based methods could make a decisive contribution to holistic and faster process development (Hans et al., 2020). In particular, these methods have already been used for optimal experimental re-design as well as process insights and faster strain phenotyping (Anane et al., 2019).

1.2. Advanced control methods of bioprocesses

Fed-batch is still the most widely used process strategy in bioprocessing. Here, the change of the feeding rate plays an important role. Classically, feeding sequences after the batch phase follow an exponential course, based on a previously defined growth rate. However, it can easily happen that the selected growth rate is either too low and thus the process takes a long time to reach a certain cell density or that the feeding rate was selected too high and oxygen limitation can occur. This is a problem for various processes, as it causes unwanted stress reactions and can significantly influence the yield (Baez & Shiloach, 2014). Model predictive control is a method that has been widely used in the chemical industry and has also shown good success in various simulations in Bioprocess engineering but has had little application in real experiments. This is mainly due to the fact that good estimates for the underlying parameters must be available for the model. Finding good values for this is often time consuming and the values are subject to large uncertainties. We show here our implementation of a combination of moving horizon estimation for parameter estimation and multi-stage model predictive control. This strategy can provide good parameter values for new strains or process conditions for which little or no information is available in a short time and calculate an optimal feeding rate that avoids oxygen limitation. Since the DOT is measured with a first order delay, it is not possible to control the process optimally by means of a PID controller, since the violation of the constraint always occurs after a glucose pulse has been added. In this respect, a model-based control with prediction of whether a selected pulse satisfies the constraint is necessary. This is another reason to take the uncertainties of the parameter values into account, as this approach has proven to be significantly more robust than previous approaches (Lucia et al., 2013). By using our fully automated Liquid Handling Station, we can quickly take many measurements to adapt our model well to the new strain and at the same time use the outputs of the MPC to optimally adjust the process online. The multi-stage approach in particular makes it possible to still ensure robust process control even with inaccurate estimated values for the parameters.

2. Material and Methods

Experiments were conducted on our high-throughput bioprocess development platform. The platform comprises two liquid handling stations (Freedom Evo 200, Tecan, Switzerland; Microlab Star, Hamilton, Switzerland), a mini bioreactor system (48

BioReactor, 2mag AG, Munich, Germany) and a Synergy MX microwell plate reader (BioTek Instruments GmbH, Bad Friedrichshall, Germany). With this setup, the platform is capable to perform up to 48 cultivations in parallel. DOT and pH are measured online, while the concentrations of the other state variables, i.e. glucose, acetate and biomass were measured at-line using enzymatic kits. The reader is referred to (Haby et al., 2019) for a detailed description of the facility and sampling procedure. The multi-rate sampling frequency of the different states poses another challenge for the optimization, as DOT is measured every 60 s, while the other state variables are only measured every 20 min. The MHE part for parameter and initial state estimation can be generally formulated according to (1), subject to (2).

$$\min_{\theta, x_{0,r}} \frac{1}{2} \|x_{0,r} - x_{0,r,old}\|_{W_x}^2 + \frac{1}{2} \|\theta - \theta_{old}\|_{W_p}^2 + \sum_{k=0}^{N_{MHE}} \frac{1}{2} \|h(x_r(t), u_r(t), \theta) - y_{meas}(t)\|_{W_y}^2 \quad (1)$$

$$\text{s.t.} \\ \dot{x}_r(t) = f(x_r(t), u_r(t), \theta) \\ \theta_{min} \leq \theta \leq \theta_{max} \quad (2)$$

The estimate for the states at the initial point of the window are then denoted by $\hat{x}_{0,r}$, the current parameter vector θ , the parameter vector from the previous horizon θ_{old} . The squared norm is applied to all calculations and the optimizer tries to minimize the deviations over the last time window N_{MHE} of the measurements y_{meas} and the predicted values $h(x_r(t), u_r(t), \theta)$, which is subject to the inputs $u_r(t)$ and current parameter vector. Each norm is further weighted by the weighting factors W_x, W_p, W_y .

The general MPC formulation is mathematically formulated as:

$$\min_{u_r} -W_M X_r(t + N_{MPC}\Delta t) - W_L \sum_{k=0}^{N_{MPC}-1} X_r(t + k\Delta t) \quad (3)$$

$$\text{s.t.} \\ \dot{x}_r(t) = f(x_r(t), u_r(t), \hat{\theta}) \\ x_r(t_0) = \hat{x}_{0,r} \quad (4)$$

$$\text{DOT}_r(t) \geq z_{lb}, \quad z_{lb} = 20\%, \quad u_r(t) \geq 3\mu\text{L}$$

Where W_V and W_L denote the weightings for the terminal- and the stage-cost, respectively. The goal was to maximize the biomass in the shortest time while at the same time avoiding oxygen limitation, i.e. having a level of at least 20% dissolved oxygen in the medium.

The system was solved using the do-mpc software which utilizes orthogonal collocation on finite elements to discretize the system, which can then be solved using NLP optimizers. One main challenge when operating this system is the non-continuous property of the inputs. Since glucose feeding is bolus-like, concentrations and volumes are changed at discrete time points. To deal with this, mass balances must be solved for the respective pulse additions, and it is mathematically more difficult to solve the system. The reader is referred to (Kim et al., 2021) for an in depth description of the underlying model and description of the optimization process. To consider the uncertainties in the parameters, a scenario tree is built. Here, each branch of the tree represents a possible combination of the uncertain parameter combinations. At each node, the optimization

problem is solved again considering the current parameter uncertainty and an optimal trajectory is identified. When the parameter uncertainty is large, this approach yields much better constraint satisfaction than nominal MPC approaches. In this contribution, we tested the multi-stage approach to consider the uncertainties for a prediction horizon of 180 min and a robust control horizon of 10 min. A schematic description of the scenario tree with arbitrary lengths is depicted in Figure 1.

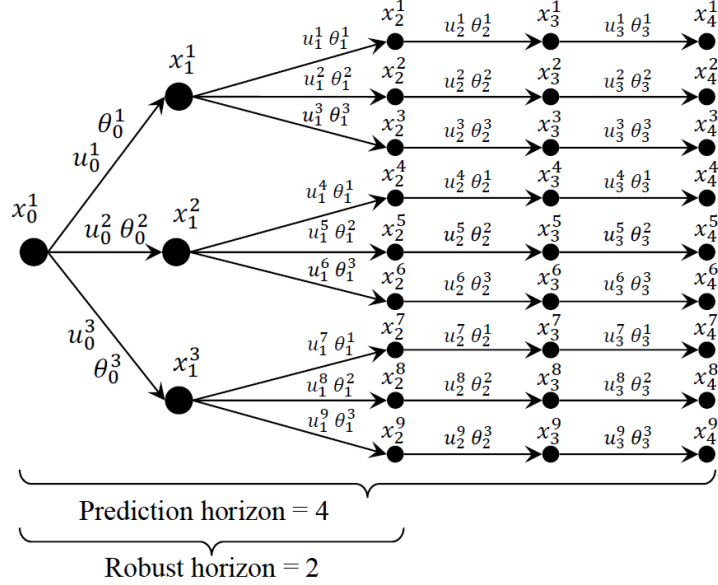


Figure 1: Scenario tree. Schematic overview about how the tree is built with an arbitrary prediction horizon of 4 and a robust horizon of 2. Uncertainties are considered until the second node.

3. Results and discussion

Based on our previous work, where we showed the advantage of using MPC to obtain higher biomass while avoiding oxygen limitation compared to a predefined feeding regime (Krausch et al., 2020), we extended our work to account for the uncertainties in the parameters. Compared to the nominal MPC, multi-stage MPC offers the opportunity to consider the inherent uncertainties of the model parameters to be more robust. This is especially important to ensure that the defined constraints are not violated. Especially for processes with a large uncertainty in the parameters or difficult to predict dynamics, this offers a much better possibility of process control. In this setup, the framework was used to consider the uncertainties of three important model parameters, namely the maximum substrate uptake rate $q_{S,max}$, the yield coefficient of biomass per substrate excluding maintenance $Y_{XS,em}$ and the volumetric oxygen transfer coefficient $k_L a$. All these parameters can affect the oxygen consumption and hence easily lead to constraint violation. Choosing the right parameters in the scenario tree could be supported by using a subset selection method with sensitivity analysis or further data-driven approaches like PCA (Thombre et al., 2019).

Many studies have shown the advantage of advanced control methods such as MPC, but have mainly been studying the nominal case for optimization (Chang et al., 2016; Ulonska et al., 2018). As can be seen in the *in silico* results in Figure 2, the nominal MPC can only

guarantee constraint satisfaction if the uncertainties are not considered, thus, small deviations in the parameter values can quickly lead to a violation of the constraint. The multi-stage approach, on the other hand, can identify an optimal feeding regime and account for the strength of uncertainties via a feedback loop. This is particularly advantageous for processes about which little information and measurement data are available like in early screening of new strains and therefore the parameter uncertainties are correspondingly large. Accordingly, taking these uncertainties into account leads to a lower feeding behavior in contrast to the nominal MPC, but to a much more robust process behavior. In particular, the system can react quickly in an unpredictable manner if the constraint is violated, since oxygen limitation also has a strong influence on the expression of numerous genes and thus could strongly change the dynamics of the system.

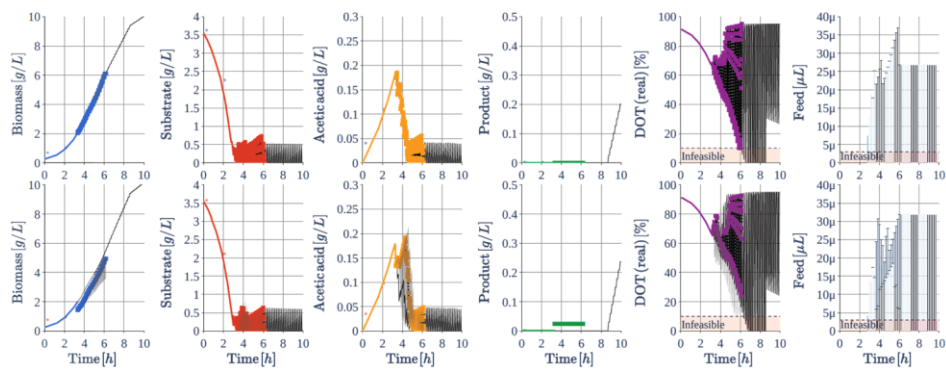


Figure 2: Comparison of nominal (upper) and multi-stage (lower) MPC at 3 h. The multi-stage approach guarantees constraint satisfaction even when the uncertainties are considered. Uncertainties for the DOT and feed are shown as shaded plots and error bars, respectively.

4. Conclusion and outlook

In this study, we have shown that a robust multi-stage MPC approach that takes into account the uncertainties in the parameters is even better than a nominal MPC. Violation of the constraint is a problem for many bioprocesses and there are various approaches to control bioprocesses. With the predictive approach, the constraints can be met well before they are violated, and the process can be guided to its optimum, so constraints are satisfied even in case of high uncertain parameters. However, the computational effort increases significantly with the number of parameters and the underlying uncertainties, so careful consideration must be given to which parameters the uncertainties should be considered for. Future work will need to deal with this problem by being more flexible in terms of what parameter uncertainties are considered and how large the prediction horizon should be, since the problem size is growing exponentially with a larger robust horizon or number of uncertain parameters to account for. Another aspect is to represent the product formation with the help of data-driven models, so that the process can also be optimized in this respect.

Acknowledgments

We thank Felix Fiedler for the development of the do-mpc software. We gratefully acknowledge the financial support of the German Federal Ministry of Education and

Research (BMBF) (project no. 01QE1957C – BioProBot and 01DD20002A – 568 KIWI Biolab).

References

- Anane, E., García, Á. C., Haby, B., Hans, S., Krausch, N., Krewinkel, M., Hauptmann, P., Neubauer, P [Peter], & Cruz Bournazou, M. N [Mariano Nicolas] (2019). A model-based framework for parallel scale-down fed-batch cultivations in mini-bioreactors for accelerated phenotyping. *Biotechnology and Bioengineering*, 116(11), 2906–2918. <https://doi.org/10.1002/bit.27116>
- Baez, A., & Shiloach, J. (2014). Effect of elevated oxygen concentration on bacteria, yeasts, and cells propagated for production of biological compounds. *Microbial Cell Factories*, 13, 181. <https://doi.org/10.1186/s12934-014-0181-5>
- Chang, L., Liu, X., & Henson, M. A. (2016). Nonlinear model predictive control of fed-batch fermentations using dynamic flux balance models. *Journal of Process Control*, 42, 137–149. <https://doi.org/10.1016/j.jprocont.2016.04.012>
- Haby, B., Hans, S., Anane, E., Sawatzki, A., Krausch, N., Neubauer, P [Peter], & Cruz Bournazou, M. N [Mariano Nicolas] (2019). Integrated Robotic Mini Bioreactor Platform for Automated, Parallel Microbial Cultivation With Online Data Handling and Process Control. *SLAS Technology*, 24(6), 569–582. <https://doi.org/10.1177/2472630319860775>
- Hans, S., Haby, B., Krausch, N., Barz, T [Tilman], Neubauer, P [Peter], & Cruz-Bournazou, M. N. (2020). Automated Conditional Screening of Multiple Escherichia coli Strains in Parallel Adaptive Fed-Batch Cultivations. *Bioengineering*, 7(4). <https://doi.org/10.3390/bioengineering7040145>
- Hemmerich, J., Noack, S., Wiechert, W., & Oldiges, M. (2018). Microbioreactor Systems for Accelerated Bioprocess Development. *Biotechnology Journal*, 13(4), e1700141. <https://doi.org/10.1002/biot.201700141>
- Kim, J. W., Krausch, N., Aizpuru, J., Barz, T [T.], Lucia, S., Cruz Bournazou, M. N [M. N.], & Neubauer, P [P.] (2021). Multi-stage model predictive control and moving horizon estimation for pulse-based parallel cultivation. *Preprint*.
- Krausch, N., Hans, S., Fiedler, F., Lucia, S., Neubauer, P [Peter], & Cruz Bournazou, M. N [Mariano N.]. (2020). From Screening to Production: a Holistic Approach of High-throughput Model-based Screening for Recombinant Protein Production. In *Computer Aided Chemical Engineering. 30th European Symposium on Computer Aided Process Engineering* (Vol. 48, pp. 1723–1728). Elsevier. <https://doi.org/10.1016/B978-0-12-823377-1.50288-3>
- Lucia, S., Finkler, T., & Engell, S. (2013). Multi-stage nonlinear model predictive control applied to a semi-batch polymerization reactor under uncertainty. *Journal of Process Control*, 23(9), 1306–1319. <https://doi.org/10.1016/j.jprocont.2013.08.008>
- Thombre, M., Krishnamoorthy, D., & Jäschke, J. (2019). Data-driven Online Adaptation of the Scenario-tree in Multistage Model Predictive Control. *IFAC-PapersOnLine*, 52(1), 461–467. <https://doi.org/10.1016/j.ifacol.2019.06.105>
- Ułonska, S., Waldschitz, D., Kager, J., & Herwig, C. (2018). Model predictive control in comparison to elemental balance control in an E. coli fed-batch. *Chemical Engineering Science*, 191, 459–467. <https://doi.org/10.1016/j.ces.2018.06.074>