

Towards “tighter” titer: model predictive control in fed-batch bioreactors

Anastasia Nikolakopoulou^{1,*}, Yu Luo², and Zach Hatzenbeller²

Abstract—Cell culture processes for the production of therapeutic monoclonal antibodies (mAbs) often suffer from high batch-to-batch variability in the final concentration of the mAb (titer) in the bioreactor. In this study, we investigated the adoption of model predictive control (MPC) to reduce this variability, even under process disturbances. Both Linear MPC (LMPC) and Nonlinear MPC (NMPC) were tested in a Design of Experiments (DoE) involving 24 bioreactors at a 250 mL scale using human-in-the-loop (HITL) tracking MPC. The LMPC used a linear state-space model, while the NMPC employed a polynomial-NARX model to describe the process dynamics. Different estimators were applied for each MPC scheme. The effect of two different disturbances was evaluated. Experimental results showed that MPC significantly reduced process variation, with titer tracking error decreasing from a range of 18 percentage points without MPC to approximately 5 percentage points for both LMPC and NMPC case studies. These findings suggest that MPC can effectively control titer in bioprocessing, offering potential for increased efficiency and productivity.

I. INTRODUCTION

Cell culture is a process where cells are cultivated under controlled conditions to produce valuable biological products. In biomanufacturing, therapeutic proteins such as monoclonal antibodies (mAbs) are manufactured by cell culture processes. One of the most widely used cell lines for mAb production is the Chinese hamster ovary (CHO) cell line. CHO cells are preferred for their ability to express soluble mammalian proteins in larger quantities than bacterial cell lines, and their resilience in continuous production, making them ideal for mAb manufacturing [1]. These cells are cultivated in a bioreactor, typically in fed-batch processes, where nutrients are supplied over time to sustain cell growth and mAb production [2].

Several key variables influence CHO cell culture performance, such as pH, dissolved oxygen (DO), viable cell concentration (VCC), metabolite levels, and nutrient concentrations. Controlled pH and DO ensure that the cells have the right environment to grow and produce the mAb of interest, while nutrient concentrations (e.g., glucose, glutamine) and waste metabolites (e.g., lactate, ammonia) are indicators of the cells’ metabolic state. The levels of these variables during each day of the cell culture can affect the final mAb concentration (titer) and its quality. The complex and complicated nature of mAb cell culture has resulted in a

wave of model-based strategies in an attempt to improve and control the process yield and product quality [3].

Bioreactor models can be used to reduce the number of experiments needed to enhance process understanding, decide the next experiment that maximizes information, establish the design space, and monitor or control the process in real time [4], [5]. Different models can be used – ranging from mechanistic models that describe the biological and chemical reactions in the system, to data-driven models, and hybrid models combining both mechanistic and data-driven components [6], [7]. Mechanistic models are constructed when extended understanding of the biological processes is already established, while empirical models provide a more flexible approach when data are available but their understanding is not very sophisticated. In the context of advanced process control, model-based control has emerged as a powerful approach to maintain optimal conditions throughout the cell culture process, ensuring consistency and robustness in mAb production.

Advanced process control in biomanufacturing industrial practices is not widely adopted yet but is slowly gaining traction thanks to its potential to increase process robustness by ensuring that process disturbances will not impact product quantity, quality, or safety. In a recent publication, mAb titer was maximized in a tracking MPC scheme using machine-learning models [8]. Advanced process control can also be used to address the practical issue of batch-to-batch titer variability and maintain consistent end-of-run (EoR) titer irrespective of process disturbances. In this work, we present the first industrial example of tracking MPC used to maintain titer to a setpoint while disturbances stretched to relevant parameter acceptable ranges were introduced into the process. In our MPC schemes, we used data-driven models and specifically a linear state-space model and a polynomial-NARX model [9], [10]. A polynomial-NARX model might have an advantage in terms of more accurately describing the process behavior, however, a linear model can have an implementation advantage in an industrial setting thanks to its simplicity. We experimentally implemented the MPC in 250 mL-scale bioreactors in a design-of-experiments (DoE) study with 24 batches which were statistically analyzed. While additional experimental work, not disclosed here, was successfully carried out to ensure that the models scale up in closed loop to industrial scales, MPC’s implementation in an industrial setting will face challenges. Data streaming and handling, model validation, incorporation in manufacturing ways-of-work, and approval by regulatory agencies are some of the most significant ones.

¹GSK, Process Engineering & Analytics, Collegeville, PA, USA, anastasianikolako@gmail.com

²GSK, Biopharm Drug Substance Development, Upper Merion, PA, USA, yu.8.luo@gsk.com, zach.a.hatzenbeller@gsk.com

This article is structured as follows: the models, controller design, and experimental setup are discussed in Section II. The results are presented and discussed in Section III and the conclusions are summarized in Section IV.

II. METHODOLOGY

A. Model formulation

1) *Linear state-space model*: State-space models are widely used in industrial control algorithms [11]. They provide a straightforward formulation for capturing the evolution of dynamical systems. Linear state-space models (LSS) are a very popular subclass of state-space models and have been used to model cell culture processes [12]. A LSS in continuous-time is written as

$$\frac{dx}{dt} = Ax(t) + Bu(t), \mathbf{y} = \mathbf{x}, \quad (1)$$

where $\mathbf{x}$ is the vector of system states, $\mathbf{u}$ is the vector of system inputs, $\mathbf{y}$ is the vector of system outputs, and A , and B are matrices of appropriate dimensions.

2) *Polynomial NARX model*: The non-linear autoRegressive moving average with exogenous inputs (NARMAX) model is a general input-output representation of nonlinear systems and contains, as special cases, models such as the NARX model [13]. The NARX model formulation in its multiple-input multiple-output (MIMO) form is written as

$$y(k) = g([\mathbf{u}(k-d-M+1)^T, \dots, \mathbf{u}(k-d)^T, \mathbf{y}(k-M_y)^T, \dots, \mathbf{y}(k-1)^T]^T), \quad (2)$$

where $g(\cdot)$ is a nonlinear map, $\mathbf{u}(\cdot)$ is the vector of system inputs, $\mathbf{y}(\cdot)$ is the vector of system outputs, d is the time delay, and M and M_y are positive integers indicating the number of past inputs and outputs included in the NARX model respectively. In this work, the focus is on models with $d = 1$, with g being a polynomial function of order r . During model training, the positive integers can take values $M = \{1, 2, 3\}$, $M_y = \{1, 2, 3\}$, and $r = \{1, 2, 3\}$. The resulting models (27 total) were evaluated against each other as described in Section II-B.2.

B. Model training

A total of 56 CHO cell culture batches for the production of a monoclonal antibody, each containing about 13 time-series data points, were available for training and testing. Daily measurements regarding cellular health (i.e., total cell concentration (TCC), viable cell concentration (VCC), viability), metabolism factors (i.e., glucose, lactate, ammonium concentration in the bioreactor), and operating conditions were collected for each batch. The operating conditions that varied among batches included pH, temperature, dissolved oxygen (DO), normalized nutrient feed (NNF=nutrient feed/cell culture volume), normalized glucose feed (NGF), and cell culture duration. A subset of the measurements related to cell health and metabolism were used either as system states in the LSS model or as system outputs in the polynomial NARX (p-NARX) model. A subset of the measurements related to the operating conditions were

used as system inputs. Different variables were included in each model formulation by design. The LSS model included more variables in its states x to allow for an accurate process description despite the limitations imposed by its formulation. Engineering judgment was used in both cases to determine which variables would be included in each model formulation. A full list of each model's variables can be found in Table I.

TABLE I
LSS AND p-NARX MODEL VARIABLES

Variable	LSS		p-NARX	
	u	x	u	y
Temperature, DO	✓		✓	
pH			✓	
pH setpoint	✓			
NNF	✓		✓	
NGF			✓	
VCC, lactate, titer		✓		✓
Viability, osmolality, pCO ₂		✓		

A p-NARX model was trained separately for each output variable (e.g., y_{VCC} , $y_{lactate}$, etc.) and a subset of the output variables can be used as part of $\mathbf{y}$ in the predictive model. For example, in the VCC predictive model

$$y_{VCC} = g([\mathbf{u}(k-M)^T, \dots, \mathbf{u}(k-1)^T, \mathbf{y}(k-M_y)^T, \dots, \mathbf{y}(k-1)^T]^T), \quad (3)$$

where $\mathbf{u} = [\text{pH, temperature, DO, NNF, NGF}]^T$ and $\mathbf{y} = [\text{VCC, lactate}]^T$, past values of lactate as well as VCC are included in the model, thus resulting in an autoregressive model with respect to VCC. A detailed breakdown of the model variables included in the $\mathbf{y}$ vector in each p-NARX model are given in Table II for completeness. Additionally, the day of cell culture is used as part of the system input vector for the titer predictive model.

TABLE II
OUTPUT VARIABLES INCLUDED IN EACH POLYNOMIAL NARX MODEL

Predictive model for ...	Variables included in $\mathbf{y}$		
	VCC	Lactate	Titer
... VCC	✓	✓	
... Lactate	✓	✓	
... Titer	✓	✓	✓

All data used for model training were collected from a CHO cell culture fed-batch process ran in 250 mL-scale bioreactors.

1) *Linear state-space model*: Linear regression was used to generate an initial estimate of the parameter matrices A and B . The parameters were refined by minimizing the weighted sum of squares between measurements and predictions. The data were split 80%–20% between the training and test sets.

2) *Polynomial NARX model*: Elastic net (EN), a regularization regression method, was used to train sparse models thus avoiding overparametrization and overfitting [14]. The model training was carried out using the *glmnet* package in MATLAB [15], [16]. Sixteen batches (29%) were held out as the test set and the remaining forty batches (71%) were used for training. A 3-fold time-series two-dimensional cross-validation was employed to identify the hyperparameters associated with EN [10]. In each fold, about 33% of the available time-series data was set aside for validation [10].

Upon model training completion, the final models for all values of M , M_y , and r (27 models in total) were evaluated against each other. The evaluation metrics were the testing data root mean square error (RMSE), the testing data residual distribution and autocorrelation, parity plots, and model sparsity. Models with lower RMSE, lower residual autocorrelation, with normally distributed residuals and therefore with low skewness in the parity plots, as well as high sparsity models were favored. The most parsimonious model was favored among models with similar evaluation metrics by prioritizing the selection of a model with slightly higher RMSE but significantly simpler/sparser structure. Training results are not presented here. It is worth noting that the trained model is as good at describing the process as the excitation of the available data used for training. To ensure that the process dynamics were sufficiently captured we assessed the test set residual autocorrelation. In practice, the NMPC would counter disturbances not available in the training set and some model-plant mismatch is expected. More details on p-NARX model evaluation and selection can be found in previously published work [10]. The structure-determining parameters for each of the final p-NARX models used in NMPC are given in Table III.

TABLE III
FINAL POLYNOMIAL NARX MODEL STRUCTURE-DETERMINING
PARAMETERS

Predictive model for ...	M	M_y	r
... VCC	2	2	2
... Lactate	1	1	3
... Titer	3	2	2

C. Controller design

The MPC algorithm determines the manipulated variables by solving an online optimization at every sampling instant. A widely used setpoint tracking MPC formulation is adopted

$$\begin{aligned} \min_{\Delta \mathbf{u}} \quad & W_y \sum_{i=1}^p (y_{sp}(k+i) - \hat{y}(k+i))^2 + W_u \sum_{i=0}^{c-1} \Delta u(k+i)^2 \\ \text{s.t.} \quad & \mathbf{u}_{\min} \leq \mathbf{u} \leq \mathbf{u}_{\max} \end{aligned} \quad (4)$$

where k is the current sampling instant, $y_{sp}(k)$ is the controlled variable setpoint at instant k , $\hat{y}(k)$ is the controlled variable estimate at instant k , $u(k)$ is the manipulated

variable at instant k , and $\Delta u(k) := u(k) - u(k-1)$ is the control move at instant k , p is the prediction horizon, c is the control horizon, W_y is a weighting factor on the controlled variable, W_u is a weighting factor on manipulated variables, $\mathbf{u}$ is a sequence of current and future manipulated variables, and $\Delta \mathbf{u}$ is a sequence of current and future control moves in the control horizon.

Titer was the controlled variable in both LMPC and NMPC and both controllers used NNF as the manipulated variable. Additionally to NNF, NGF was used as a manipulated variable in the NMPC ($\mathbf{u}_{\text{NMPC}} = [\text{NNF} \ \text{NGF}]^T$). The controller tuning parameters were identified such that the end-of-run titer deviation from the setpoint was minimized in simulations. Additional tuning was carried out during the experiment.

The NMPC optimization constraints were decided based on the system characteristics. NNF and NGF were constrained by their ranges in the model training data sets to ensure that the optimization was carried out within the model fidelity region. Additionally, a volume constraint $V_{\max}$ was implemented in NMPC to ensure that the cell culture volume does not exceed the feasible working volume. The initial guess for the optimization ($\mathbf{u}_{\text{start}}$) was set equal to the lower bound. The final controller tuning and optimization parameters are shown in Table IV.

A linear and a nonlinear model are used in linear and nonlinear MPC respectively to obtain the estimates $\hat{y}(k)$.

The estimate in the LMPC algorithm was obtained first by a moving horizon estimation step:

$$\begin{aligned} \hat{x}^*(k-2) &= \arg \min_{\hat{x}(k-2)} \sum_{s=k-2}^k w \|\hat{x}(s) - x(s)\|_W^2 \\ &\quad + (1-w) \|(\hat{x}(s) - \hat{x}_{\text{prev}}(s))\|_W^2 \\ \text{s.t.} \quad \hat{x}(k) &= A\hat{x}(k-1) + Bu(k-1) \\ \hat{x}(k-1) &= A\hat{x}(k-2) + Bu(k-2) \end{aligned}$$

where $\hat{x}^*(k-2)$ is the estimated state at the beginning of a 3-element moving horizon; $w = 0.9$ is the weight on the cost term based on measurement; $\hat{x}_{\text{prev}}$ is the previously estimated state; $W = \text{diag}(10^{-5}, 2, 10, 50, 0.007, 0.001)$ is weight matrix of the cost associated with comparing the model prediction and measurement (or previously estimated state) of titer, VCC, viability, lactate, OSMO, and pCO_2 .

The estimate is further modified by applying a multiplicative correction factor to the re-estimated state:

$$\begin{aligned} p^* &= \arg \min_p \sum_{s=k-2}^k \|\text{diag}(p)\hat{x}(s) - x(s)\|_W^2 \\ &\quad + \|p - 1\|^2 + \|p - p_{\text{prev}}^*\|^2 \\ \text{s.t.} \quad \hat{x}(k) &= A\hat{x}(k-1) + Bu(k-1) \\ \hat{x}(k-1) &= A\hat{x}^*(k-2) + Bu(k-2) \end{aligned}$$

where p^* is the updated correction factor, which is initialized as an array of ones; p_{prev}^* is the previously updated correction factor. This correction factor further reduces the difference between the re-estimated state and the measurement, and it

is applied to future predictions that are used in the controller objective function.

The estimate in the NMPC algorithm is obtained by assuming a constant unmeasured disturbance propagating across the prediction horizon

$$\begin{bmatrix} \hat{y}(k+1|k) \\ \vdots \\ \hat{y}(k+p|k) \end{bmatrix} = \begin{bmatrix} y(k+1|k) \\ \vdots \\ y(k+p|k) \end{bmatrix} + y_m(k) - \hat{y}(k). \quad (5)$$

The notation $y(i|k)$ stands for the predicted value of variable y at instant i given the information up to (and including) the instant k and $y_m(k)$ is the measured value at instant k . The unmeasured disturbance is equal to the difference of the current measurement and its corresponding estimated value obtained in the previous time instant [17], [18]. Thus by using eq. (5), an integrator is implemented in the NMPC. More details regarding the implementation of the NARX predictions and the estimator can be found in the literature [10].

It is noted that there are no available p-NARX model predictions during the first two days of the experiment. The structures of the final models suggest that two previous inputs (additionally to the current value obtained via optimization) and two previous outputs are required to obtain a prediction for titer. Therefore, p-NARX models with $M = M_y = 1$ are used for the prediction of VCC, lactate, and titer in NMPC during the first two days of the experiment.

TABLE IV
CONTROLLER & OPTIMIZATION PARAMETERS

Parameter	LMPC	NMPC
W_u	10^4	$\begin{bmatrix} 1 & 1 \end{bmatrix}$
W_y	10^{-6}	10^{-5}
p	until end of batch	until end of batch
c	3	$\min(p, 4)$
$\mathbf{u}_{\min}$	0.002	$\begin{bmatrix} 0.002 & 10^{-7} \end{bmatrix}$
$\mathbf{u}_{\max}$	0.05	$\begin{bmatrix} 0.024 & 0.014 \end{bmatrix}$
$V_{\max}$	–	230 mL
$\mathbf{u}_{\text{start}}$	previous MPC solution	$\mathbf{u}_{\min}$

The NMPC was solved in JuMP [19] while LMPC was implemented in Python. Both controllers resulted in extremely fast computational times allowing for real-time implementation.

D. Experimental design

A statistical design of experiments (DoE) with three factors—one categorical (“Controller”) and two numeric (“initial VCC (iVCC)” disturbance, “pH, T ” disturbance)—was conducted to assess if the controllers’ disturbance rejection was statistically significant. The numeric factors were decided based on a prior study’s results, which suggested that variations in the initial VCC during bioreactor inoculation (iVCC) had the highest impact on end-of-run (EoR) titer. The same study showed that changes in the bioreactor pH and temperature (T) had a relatively lesser impact on the

EoR titer compared to iVCC. Temperature and pH were lumped into a single factor to define a DoE structure with reasonable number of runs while still providing adequate statistical power. A singular value decomposition (SVD) analysis on the steady state gain matrix of the LSS model (B) showed that the largest impact on titer occurs when T and pH change in the same direction. During the experiment, temperature and pH in the bioreactor were maintained at their centerpoint values ($T = 34^\circ\text{C}$ and $\text{pH} = 7.2 \pm 0.1$). The pH and temperature disturbances were enforced between days 3 and 4 and days 9 and 10 of the experiment by changing the setpoint of these variables in the same direction. The disturbance levels for the two numeric factors are shown in Table V. Each of the factors had three levels in the DoE.

TABLE V
EXPERIMENTAL DESIGN

Batch No	Controller	iVCC (10^6 cells/mL)	pH, T
1	NMPC	18	7.1, 33°C
2	NMPC	12	7.3, 35°C
3	LMPC	12	7.2, 34°C
4	No MPC	18	7.3, 35°C
5	LMPC	18	7.3, 35°C
6	LMPC	12	7.3, 35°C
7	NMPC	18	7.3, 35°C
8	No MPC	15	7.2, 34°C
9	LMPC	18	7.1, 33°C
10	NMPC	18	7.2, 34°C
11	No MPC	18	7.1, 33°C
12	NMPC	15	7.1, 33°C
13	LMPC	15	7.1, 33°C
14	NMPC	15	7.3, 35°C
15	LMPC	18	7.2, 34°C
16	No MPC	12	7.1, 33°C
17	NMPC	12	7.1, 33°C
18	LMPC	15	7.3, 35°C
19	NMPC	12	7.2, 34°C
20	LMPC	12	7.1, 33°C
21	No MPC	12	7.3, 35°C
22	No MPC	15	7.2, 34°C
23	No MPC	15	7.2, 34°C
24	No MPC	15	7.2, 34°C

The DoE was carried out in 24 ambr250 bioreactors (250 mL scale, Sartorius Stedim, Göttingen, Germany). A CHO K1 glutamine synthetase (GS)–knockout cell line expressing an IgG1 monoclonal antibody was used. Proprietary and chemically defined seed expansion medium, production basal medium, and production feed medium were used. The initial working volume was 200 mL. The target seeding density was between 12 and 18 million cells/mL (see Table V). The reactors were agitated at 408 rpm. The dissolved oxygen was controlled at 50%. The feed rate of each reactor was determined based on MPC outputs or the cell-specific feed rate (CSFR) and the VCC measured every day for the “no MPC” reactors:

$$F_{\text{feed}} = \text{CSFR} \times x_v \times V$$

where F_{feed} (mL/h) is the volumetric flow rate of the nutrient feed, x_v (cells/mL) is the measured VCC, and V (L) is the

working volume of the bioreactor. A 400 g/L glucose solution was used to provide supplemental glucose as needed. All experiments were run for 12 days.

Once daily, the variables relevant to the cell culture (e.g., VCC, lactate, glucose, pH, temperature, titer etc.) were measured for each of the 24 batches in the DoE. The antibody titer was measured using a Cedex Bio HT analyzer (Roche, Mannheim, Germany). The cell densities and cell culture viability were measured using a Vi-Cell XR cell counter (Beckman, Indianapolis, IN). The metabolite concentrations were measured using a BioProfile FLEX2 analyzer (Nova Biomedical, Waltham, MA). The values were used in the relevant MPC code to calculate the NNF (and NGF for the NMPC batches) for the day. For the LMPC batches, glucose was added into the bioreactor based on the standard glucose feeding recipe for the process (proprietary).

This study was executed as human-in-the-loop (HITL) MPC due to lack of automation infrastructure in the ambr250 reactors. The daily inputs to the MPC code were the available experimental measurements and the daily outputs were NNF and NGF. These values were manually entered into the pump software and the pumps would deliver the right amounts to the cell culture. In subsequent scale-up studies, this limitation was removed and automation infrastructure was built to deploy automated MPC with minimal human intervention.

III. RESULTS & DISCUSSION

Each level of the “Controller” factor was assigned 8 reactors (Table V). Each of these reactors was inoculated with each of three target values for iVCC. The results are grouped based on “Controller” and “iVCC”. The daily titer setpoint was determined by the daily average experimental titer values of the nominal process such that the process controlled under disturbances approximates the process under nominal operating conditions.

The titer tracking error throughout the 12 days of the experiment are shown in Fig. 1. The final titer in bioreactors inoculated with low iVCC (12×10^6 cells/mL) was much lower than the titer in bioreactors inoculated with high iVCC (18×10^6 cells/mL) when no controller was acting on the process. In the former case, an average 8% titer loss compared to the nominal process (iVCC= 12×10^6 cells/mL) was incurred while the later case resulted in a 11% titer gain on average. Additionally, the process variation was lower under the presence of LMPC or NMPC; all bioreactors, irrespective of their iVCC, resulted in similar titer profiles. As a result, the tracking error towards the end of the run was very similar for bioreactors controlled by the same controller despite different iVCC at inoculation. On the other hand, the tracking error spanned 19 percentage points for the reactors without titer control. Overall, both controllers successfully controlled titer around the setpoint (blue dashed line) until day 9. However, LMPC undershot the setpoint from day 10 onwards which was attributed to suboptimal tuning, was investigated in subsequent studies, and successfully addressed. The results of the subsequent studies are not be disclosed at this time.

NMPC successfully tracked the setpoint until the end of the run.

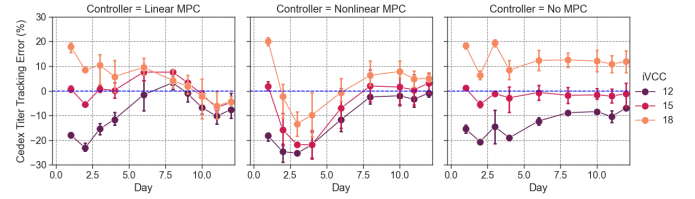


Fig. 1. Average titer tracking error per controller type and disturbance type. The blue dashed line corresponds to the MPC setpoint. The error bars correspond to the 95% confidence intervals.

The end-of-run tracking error signal was statistically analyzed in Design Expert. The results are shown in Fig. 2 and Table VI. A statistically significant model for tracking error was obtained. The p-values suggest that iVCC and the type of controller are the factors that most impact EoR tracking error and do so in a statistically significant way. The reduced slope of the lines corresponding to the controllers compared to the “no MPC” in Fig. 2 suggests that the controllers are able to reduce the impact of iVCC on EoR tracking error, thus successfully rejecting the disturbance.

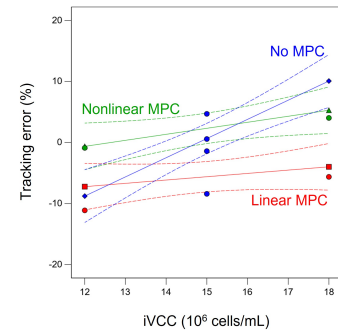


Fig. 2. End-of-run titer tracking error (%) for different levels of the “iVCC” and “Controller” factors.

TABLE VI
DOE STATISTICAL ANALYSIS RESULTS

Source	Sum of squares	Degrees of freedom	p-value
Model	787.09	6	< 0.0001
A: iVCC	265.88	1	0.0001
B: pH, T	83.24	1	0.0145
C: Controller	279.92	2	0.0005
AC	158.05	2	0.0059
Residual	190.79	17	
Lack of fit	100.85	14	0.9735
Pure error	89.94	3	

The cumulative nutrient feed amount (mL) and glucose feed (mL) are shown in Figs. 3 and 4 respectively. Both MPCs suggested higher nutrient feed amounts compared to the standard cell-specific feeding rate (CSFR) (confidential)

for the “no MPC” reactors. The CSFR suggests feeding proportionally to the VCC. Therefore, the reactors inoculated with lower iVCC will be supplemented with less nutrient, as their VCCs are lower. Higher iVCC implies that more nutrient is provided to the cells under CSFR which leads to more growth, higher VCCs (more healthy cells), and subsequently more nutrient being required. On the other hand, LMPC suggests on average higher nutrient feed amounts for the reactors inoculated with lower iVCCs compared to the reactors inoculated with higher iVCCs. This allows for bridging the VCC gap across reactors and consequently titer productivity. It is noted that there is a trade-off between cell growth, resulting in increasing VCC, and titer productivity by the healthy cells. Very high nutrient feeds can result in cell cultures with a high number of healthy cells (VCC) but low titer productivity. NMPC suggests feeding on average more nutrient to the low iVCC reactors. Interestingly, the average nutrient feed is almost identical across reactors controlled by NMPC past day 4. This is attributed to the fact that after day 4 the reactors seem to have very similar VCC (Fig. 5) and lactate (Fig. 6) profiles implying that the differences across reactors have been minimized and their dynamic behaviors are on average the same. The feeding patterns until day 4 in NMPC reactors might have resulted in this behavior.

The cumulative glucose feed provided in the cell cultures was very different across reactors controlled by different controllers (Fig. 4). LMPC and “no MPC” reactors were supplemented with glucose using the same rules, ensuring that the concentration of glucose in the reactor would not fall under a certain threshold. Thus, the cumulative glucose feed values were on average very similar between LMPC and “no MPC” reactors until day 7. Past day 7, LMPC reactors were supplemented with a lot more glucose due to lower glucose concentrations in the bioreactor (Fig. 7). The lower glucose concentration in the LMPC reactors was likely due to the higher VCC resulting in higher glucose consumption. NMPC optimization was responsible for determining the glucose feed amount for the NMPC reactors and resulted in much lower total glucose being provided to all NMPC reactors compared to all other reactors. These differences also led to different glucose concentration profiles (Fig. 7). Due to the complex nature of biological processes, further exploration is needed to enhance the understanding of factors contributing to these differences.

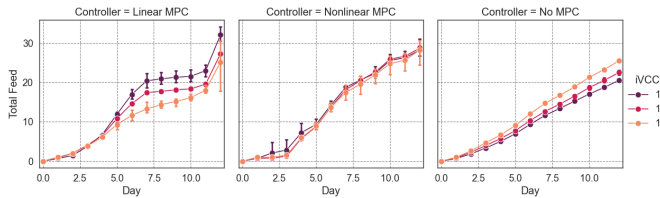


Fig. 3. Cumulative nutrient feed per controller type and disturbance type. The error bars correspond to the 95% confidence intervals.

The controllers also influenced the VCC profiles (Fig. 5). Both controllers resulted in similar VCC profiles across the

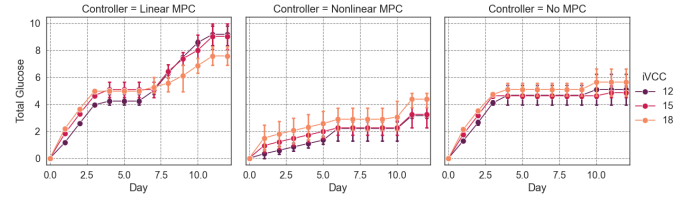


Fig. 4. Cumulative glucose feed per controller type and disturbance type. The error bars correspond to the 95% confidence intervals.

bioreactors they controlled past days 5 and 4 for LMPC and NMPC respectively. “Controlling” the VCC by controlling titer to a setpoint suggests that under the presence of advanced control, we can obtain a uniformity across reactors and therefore reduced process variation in process variables other than the ones we are actively trying to control. NMPC resulted in a prolonged peak in the VCC profiles (at about day 6) compared to LMPC and “no MPC” (at about days 3 and 4 respectively) suggesting prolonged cell health which is associated with higher productivity. Cell viability (Fig.

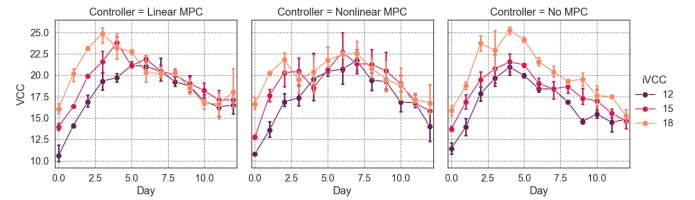


Fig. 5. Viable cell concentration (VCC) per controller type and disturbance type. The error bars correspond to the 95% confidence intervals.

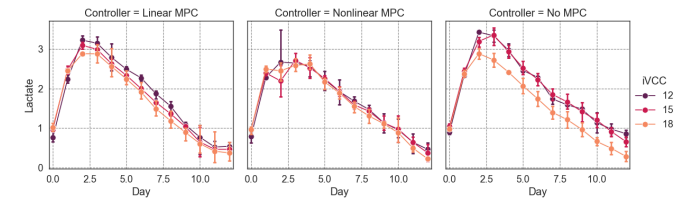


Fig. 6. Lactate concentration in the bioreactors per controller type and disturbance type. The error bars correspond to the 95% confidence intervals.

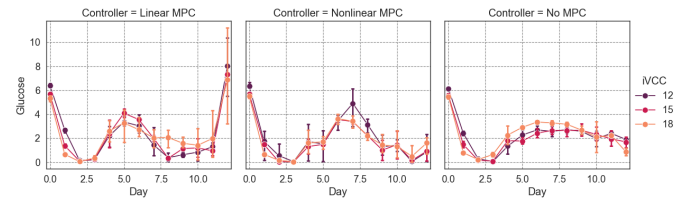


Fig. 7. Glucose concentration in the bioreactors per controller type and disturbance type. The error bars correspond to the 95% confidence intervals.

8) was higher for the “no MPC” reactors towards the end of the run but this did not negatively impact the productivity of the LMPC and NMPC reactors since the number of healthy

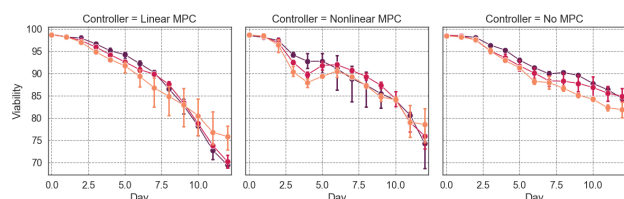


Fig. 8. Viability per controller type and disturbance type. The error bars correspond to the 95% confidence intervals.

cells (VCC) was still higher in these reactors. In subsequent versions of this work, viability models will be constructed and viability constraints will be incorporated in the MPCs to ensure no impact on relevant quality attributes. However, no issues were observed during this experiment despite the slightly lower viability values in the MPC reactors.

Since model–plant mismatch is expected, especially when using the linear model, the controller performance needs to be assessed within industrially relevant disturbance ranges and at larger scales. This effort is currently ongoing.

IV. CONCLUSIONS

In this study, we explored the potential of implementing model predictive control (MPC) in CHO cell culture fed-batch processes to control monoclonal antibody titer. Both linear MPC (LMPC) and nonlinear MPC (NMPC) were evaluated for their ability to maintain the titer setpoint in the presence of process disturbances in an experimental setting. A design of experiments (DoE) approach involving 24 bioreactors at 250 mL-scale was conducted using human-in-the-loop (HITL) MPC. Our findings demonstrated that “MPC” significantly reduced process variation in titer compared to reactors without MPC under the presence of three process disturbances, grouped into two groups (initial viable cell concentration at inoculation and lumped pH, temperature disturbance). Specifically, the end-of-run titer tracking error without MPC spanned 18 percentage points, while with MPC, it was reduced to approximately 5 percentage points. This rejection of the process disturbances was found to be statistically significant. Additionally, NMPC was observed to result in a delayed peak in the VCC profiles, associated with prolonged cell health, suggesting MPC’s potential for indirectly improving the bioreactor dynamic behavior. LMPC showed an undershooting of the titer setpoint during the final days of the batch, likely due to the linear model behavior and sub-optimal tuning. Future work will focus on testing these algorithms in larger bioreactor scales and developing the infrastructure required to transition from HITL to full automated MPC. Additionally, MPC has potential to improve titer productivity in economic MPC schemes. Drug quality constraints or objectives can also be included in MPC schemes as long as such models can be developed using available data.

ACKNOWLEDGMENT

We would like to thank the following people for their contributions to this work: Rachel Sharp, Kelsey Morgan,

Lingyi Huang, Alexis Parillo, Kona Haramoto, Ruya Ronish, Prairie Yang, Tia Lewis, Kevin Huff, Randal Allen, Ou Yang, Gabriele Bano, Sayantan Chattoraj, Brian Rhodes, Sameer Talwar, Diana Ritz, Brian Crump, Paloma Diaz-Fernandez, and Pavlos Kotidis.

REFERENCES

- [1] A. C. Szkodny and K. H. Lee, “Biopharmaceutical manufacturing: Historical perspectives and future directions,” *Annual Review of Chemical and Biomolecular Engineering*, vol. 13, no. Volume 13, 2022, pp. 141–165, 2022.
- [2] W.-P. Alejandra, J.-P. Miriam Irene, G.-S. Fabio Antonio, R.-G. R. Patricia, T.-A. Elizabeth, J. P. Aleman-Aguilar, and G.-V. Rebeca, “Production of monoclonal antibodies for therapeutic purposes: A review,” *International Immunopharmacology*, vol. 120, p. 110376, 2023.
- [3] O. Yang and M. Ierapetritou, “mab production modeling and design space evaluation including glycosylation process,” *Processes*, vol. 9, no. 2, 2021.
- [4] M. S. Hong, W. Sun, A. Lu, and R. Braatz, “Process analytical technology and digital biomanufacturing of monoclonal antibodies,” *Am. Pharm. Rev.*, vol. 23, no. 6, pp. 122–125, 2020.
- [5] A. Nikolakopoulou, O. Yang, and G. Bano, “Real-time statistical process monitoring,” in *The Digital Transformation of Product Formulation: Concepts, Challenges, and Applications for Accelerated Innovation* (A. Schmidt and K. Wallace, eds.), Boca Raton, FL: CRC Press, 1st ed., 2024.
- [6] *Basic Principles of GMP*. Geneva, Switzerland: World Health Organization, 2008.
- [7] Y. Chen, O. Yang, C. Sampat, P. Bhalode, R. Ramachandran, and M. Ierapetritou, “Digital twins in pharmaceutical and biopharmaceutical manufacturing: A literature review,” *Processes*, vol. 8, no. 9, 2020.
- [8] M. Rashedi, M. Rafiei, M. Demers, H. Khodabandehlou, T. Wang, A. Tulyan, C. Undey, and C. Garvin, “Machine learning-based model predictive controller design for cell culture processes,” *Biotechnology and Bioengineering*, vol. 120, no. 8, pp. 2144–2159, 2023.
- [9] A. Nikolakopoulou and R. D. Braatz, “Fast nonlinear model predictive control of distributed parameter systems,” in *2022 American Control Conference (ACC)*, pp. 994–999, 2022.
- [10] A. Nikolakopoulou and R. D. Braatz, “Polynomial NARX-based nonlinear model predictive control of modular chemical systems,” *Computers & Chemical Engineering*, vol. 177, p. 108272, 2023.
- [11] S. Qin and T. A. Badgwell, “A survey of industrial model predictive control technology,” *Control Engineering Practice*, vol. 11, no. 7, pp. 733–764, 2003.
- [12] Y. Luo, V. Kurian, L. Song, E. A. Wells, A. S. Robinson, and B. A. Ogunnaik, “Model-based control of titer and glycosylation in fed-batch mab production: Modeling and control system development,” *AIChE Journal*, vol. 69, no. 4, p. e18075, 2023.
- [13] S. Chen and S. A. Billings, “Representations of non-linear systems: the NARMAX model,” *International Journal of Control*, vol. 49, no. 3, pp. 1013–1032, 1989.
- [14] H. Zou and T. Hastie, “Regularization and variable selection via the elastic net,” *Journal of the Royal Statistical Society Series B: Statistical Methodology*, vol. 67, pp. 301–320, 03 2005.
- [15] J. H. Friedman, T. Hastie, and R. Tibshirani, “Regularization paths for generalized linear models via coordinate descent,” *Journal of Statistical Software*, vol. 33, no. 1, p. 1–22, 2010.
- [16] J. Qian, T. Hastie, J. Friedman, R. Tibshirani, and N. Simon, “Glmnet for Matlab.” http://hastie.su.domains/glmnet_matlab/, 2013.
- [17] C. R. Cutler and B. L. Ramaker, “Dynamic Matrix Control - a computer control algorithm,” in *AIChE National Meeting*, (Houston, Texas), 1979.
- [18] C. E. Garcia and A. M. Morshedi, “Quadratic programming solution of dynamic matrix control (QDMC),” *Chemical Engineering Communications*, vol. 46, no. 1-3, pp. 73–87, 1986.
- [19] M. Lubin, O. Dowson, J. Dias Garcia, J. Huchette, B. Legat, and J. P. Vielma, “JuMP 1.0: Recent improvements to a modeling language for mathematical optimization,” *Mathematical Programming Computation*, 2023.