Received: 1 December 2022 | Revised: 2 February 2023 | Accepted: 12 February 2023

DOI: 10.1002/aic.18075

AIChE JOURNAL

RESEARCH ARTICLE

Biomolecular Engineering, Bioengineering, Biochemicals, Biofuels, and Food

Model-based control of titer and glycosylation in fed-batch mAb production: Modeling and control system development

Yu Luo[1] | Varghese Kurian[1] | Liqing Song[2] | Evan A. Wells[2] | Anne Skaja Robinson[2] | Babatunde A. Ogunnaike[1]

[1]Department of Chemical and Biomolecular Engineering, University of Delaware, Newark, Delaware, USA

[2]Department of Chemical Engineering, Carnegie Mellon University, Pittsburgh, Pennsylvania, USA

Correspondence
Anne Skaja Robinson, Department of Chemical Engineering, Carnegie Mellon University, Pittsburgh, PA, USA.
Email: asrobins@andrew.cmu.edu

Funding information
National Institute for Innovation in Manufacturing Biopharmaceuticals (NIIMBL), Grant/Award Number: QSP 1.23

Abstract

Therapeutic monoclonal antibodies (mAbs) are typically manufactured using mammalian cell cultures in fed-batch bioreactors, with increasing emphasis on meeting productivity *and* product quality attribute targets that depend strongly on such process variables as nutrient feed rates and bioreactor operating conditions. In this article, we identify, categorize, and address the challenges of achieving both productivity and product quality goals simultaneously, by developing a multivariable, model-based control system that can satisfy multiple production objectives in a fed-batch cell culture process. Here, we discuss model development and present theoretical concepts of observability and controllability that are essential to understanding and handling effectively these intrinsic challenges. Subsequently, we evaluate via simulation the performance of the outer-loop model predictive control and demonstrate the overall capability to satisfy complex production objectives in a laboratory scale bioreactor, as a first step toward the ultimate goal of creating an advanced control system for fed-batch mAb manufacturing processes.

KEYWORDS
bioprocess engineering, control, multiscale modeling, optimization, process control

1 | INTRODUCTION

Monoclonal antibody (mAb) drugs enjoy a high share of the global biopharmaceutical products market, valued at about US$115 billion in 2018 and predicted to grow at a rate of over 14% in the next few years.[1] Achieving desired productivity *and* product quality targets in a consistent manner remains one of the major challenges of biopharmaceutical manufacturing.[2] For mAb manufacturing, protein titer and biomass are the most common productivity attributes; extent of glycosylation, glycan distribution, charge variants, and protein aggregates are the most common critical quality attributes (CQAs).[3] Because glycosylation—an enzymatic, post-translational process by which sugar molecules (i.e., glycans) are added to or removed from mAbs—is non-templated, the resulting distribution of glycans is often highly variable and difficult to control. However, the distribution of glycans is an important CQA of mAb therapeutics because glycans can alter the *in vivo* function of antibodies, either triggering or inhibiting desirable therapeutic effects, depending on the drug's target and mechanism of action.[4–6] In addition, glycan distribution affects the serum half-life and immunogenicity of antibodies, which can factor into dosage frequency and patient safety. Consequently, an effective and robust process control strategy that can maintain high protein titer *and* ensure that all CQAs (including glycan distribution) are within allowable ranges is essential to efficient mAb manufacturing.

Significant progress has been made in regulatory control of nutrient and metabolite levels in Chinese hamster ovary (CHO) cell cultures using online sensing technologies (e.g., Raman spectroscopy).[7,8] However, online control strategies at the product-attribute level (for controlling productivity and quality attributes) currently do not exist in commercial mAb manufacturing for a variety of reasons, including process complexity, lack of reliable online sensors, and other technical factors reviewed subsequently in this article and in reference 9. Instead, various attempts at using predetermined (open loop) and heuristic-based process operating protocols to meet the production targets and improve yield and quality have been attempted, but with limited effectiveness. Some notable examples that have been studied and tested experimentally are: (1) *Feeding Strategies*, such as volume-based feeding (which maintains a constant ratio between the amount of feed medium added into the culture and the working volume of the bioreactor throughout a run); cell count-based feeding[10]; glucose concentration-based feeding[11,12]; pH-based, turbidity or oxygen uptake rate-based feeding,[13] all primarily developed by trial-and-error and from heuristics. (2) *Predetermined Temperature Profiles*,[14] (3) pH,[15] and (4) *Nutrient Supplementation*.[16] While some of these techniques have improved the productivity and product quality of some specific CHO cell culture processes, they are all strictly open-loop strategies, with no feedback control to correct for the effect of unavoidable disturbances; they are therefore unable to achieve specific production targets precisely, and updating them for new processes requires re-running an inordinately large number of experiments on the new process. Most importantly, however, except for a few studies,[16] many approaches are capable of improving only one production attribute at a time, not all simultaneously.

To summarize: the goal of modern biomanufacturing is relatively easy to state—it is to achieve desired specification targets for the final mAb titer *and* CQA profile, simultaneously, consistently, and repeatably—but achieving this goal has been difficult. Our objective with this work is to present a rational framework for achieving this goal consistently. Our approach is based on using appropriate dynamic process models for two central tasks: (1) to estimate critical productivity and product quality attributes frequently throughout a fed-batch reaction run, and (2) to determine how best to adjust appropriate process inputs, judiciously and continuously, in order to control the final mAb titer *and* glycosylation profile simultaneously. This leads us to identify four broad technical challenges to be addressed for successful development and deployment of such a framework for effective advanced control of mAb production processes: modeling, estimation, controller design, and controller implementation. It is instructive therefore to present first an overview of the challenges to be addressed before presenting the details of how we propose to handle the challenges in developing the proposed framework. As such, the rest of the article is organized as follows: first, in Section 2, we present an Overview of these four challenges in general, with special emphasis on the specific problems posed by bioprocesses, before discussing our proposed approaches in the Methods section (Section 3). In Results and Discussion (Section 4), we present the first set of results obtained from an experimental laboratory bioreactor and evaluate the proposed control system's performance first in simulation, as a precursor to a detailed discussion of the full experimental implementation discussed in a follow-up paper. We draw insight from, and discuss the implications of, the results (experimental and simulation), and end the article with the Summary and Conclusions section to put this article in context and to preview the implementation of the control system experimentally on a laboratory scale bioreactor. A follow-up paper is devoted to a detailed discussion of the control system implementation on a laboratory scale bioreactor and of the experimental results.

2 | OVERVIEW

2.1 | Modeling

Regardless of the specific process in question, a validated process model is useful for process understanding, process state estimation, process output prediction, and process control.[17] However, developing useful models for mAb manufacturing processes is particularly challenging because of the complexity of bioprocess operations over widely disparate length- and time-scales. For instance, the dynamics of cell culture components (medium composition, cell population) evolve slowly in time, and at a macroscopic length scale comparable to the bioreactor size; the glycosylation of antibodies occurs at a much faster rate at the molecular scale in an organelle within the cell. The challenges associated with modeling these sub-processes appropriately at each scale are also different.

At the macro scale in the bioreactor, cell culture dynamics involve cell metabolism, cell growth, and protein synthesis, with the concentrations of cells, metabolites, nutrients, and antibodies interconnected through a vast metabolic network. Modeling the entire network and including all the species involved would lead to a nonlinear high-dimensional model with a large number of ordinary differential equations (ODEs) to solve and many (kinetic and stoichiometric) parameters to estimate. The values of most of those parameters are not available in the literature and cannot be estimated easily. Further, different process configurations (cell lines, media, operating conditions, etc.) affect kinetic parameters and render parameters estimated from one process configuration invalid for predicting process behavior under new conditions. Determining such parameters empirically each time is too time-consuming and too costly to be practical. In contrast, semi-mechanistic cell culture models can be developed (e.g., reference 18) by merging multiple metabolic pathways and lumping kinetic equations to obtain a system of ODEs of manageable size. However, many semi-mechanistic model parameters no longer carry physical meaning, and their values can be estimated only from experimental data. While less cumbersome in size, these semi-mechanistic models may still not be able to predict process dynamics accurately enough under new process conditions because, having been developed with limited mechanistic information specifically for certain predetermined baseline process configurations, extrapolation to other conditions is neither advisable nor effective.

At the micro scale, the molecular level glycosylation process occurs as a sequence of enzymatic reactions that modify antibodies by adding or removing glycans. Developing appropriate models for controlling glycosylation is challenging for the following reasons: (i) the complexity of the underlying reaction network, which contains more than 7000 unique glycans and over 20,000 reactions[19]; (ii) the stochasticity of the process[20]; and (iii) the non-availability of precise means of quantifying how process inputs affect glycosylation.

Furthermore, cell culture and glycosylation dynamics evolve over widely differing time- and length-scales: macro-scale cell culture dynamics evolve on the order of hours and days, while the micro-scale glycosylation dynamics evolve on the order of minutes. Developing a high-fidelity multiscale model that combines dynamics of disparate time- and length-scales presents the first major challenge to developing an effective model-based control system.

2.2 | Estimation

A defining characteristic of mAb manufacturing processes is that often, glycan distribution and other CQAs can be assessed only off-line and at the end of a bioreactor run, rendering such assessments ineffective for direct use in a feedback control scheme that otherwise requires more frequent sampling. When a product attribute is off-target, end-of-run information is useless for taking any corrective control action, for the obvious reason that such information would have arrived too late. Off-line measurements, by nature, are infrequent, and take a considerable amount of time to complete. Consequently, during the long intervals between samples, process operators are blind to the true state of the process; and by the time the analyses are completed, the information produced would be obsolete. Ensuring that disturbances affecting the process can be compensated for expeditiously requires reliable, on-time, and sufficiently frequent information about the true state of the process variables to be controlled. Recently, mass spectrometry (MS) and chromatography-based technologies have been deployed to provide at-line glycan measurements, but the associated cost and labor requirements prevent these measurements from being practical for real-time process control.[21]

These limitations argue for a "soft sensing" approach where the entire collection of measurable information, such as online bioreactor sensor data, at-line cell culture measurements, and spectroscopic data, can be combined with appropriate process models to provide, as frequently as necessary, reliable estimates of the process states, glycan distribution, and other CQAs that are not directly measurable by "hard" sensors. In addition to the practical matter of how such "state estimators" can be designed and implemented, is the more fundamental question: *Is it always possible to estimate all the required process states and product characteristics from the available measurements and an appropriate model*? This is the issue of observability, an important concept in control theory. A system is observable if all its states can be estimated from the available measurements; otherwise, it is unobservable. The design and implementation of a technique to generate estimates of desired but unmeasured process and product variables, at the appropriate frequency, using available measurements from a bioprocess that may be "unobservable," is the second major challenge to address.

2.3 | Controller design

In the current context, controller design is the process by which one obtains a scheme for adjusting appropriate process manipulated variables automatically, online, to meet the productivity and quality attribute objectives. To be effective, the control scheme must have access to desired process and product information and must be able to determine what variables to manipulate when and by how much. Solving the problem of "estimation" as summarized above can provide the required information; but the complexity of the bioprocess argues for a model-based approach to the design of the controller. In this latter sense, the inherent nonlinearity, high dimensionality, and multiscale nature of the underlying dynamics of mAb metabolism and glycosylation raises the issue of controllability, another concept in control theory that is complementary to observability noted earlier.

Informally, controllability is a measure of whether a system's available manipulated variables are capable of steering the process from any initial state to any other arbitrarily specified desired state in the entire allowed operating space. Even without formal analysis, it is evident, by mere observation, that a mAb manufacturing process has far more glycan species to be controlled than available manipulated variables, in which case controlling the entire glycan state is impossible; there simply are not enough degrees of freedom. However, while protein glycosylation is uncontrollable in this sense, a legitimate question to ask is: what extent of control can be achieved? A systematic answer to this important question can be found in reference 22 where St. Amand et al. quantified the controllability of glycosylation in a latent space, showing which output modes (linear combinations of the controlled variables) can be controlled using which input modes (linear combinations of the manipulated variables in the original space) and the strength or degree to which an input mode affects its corresponding output mode.

2.4 | Controller implementation

The development of an advanced model-based controller for mAb manufacturing processes, while necessary, is by no means sufficient. Achieving the ultimate objective of real-time implementation in actual practice on a real process requires that we contend with several non-trivial practical issues such as: real-time data acquisition; data transmission from the process to the computer and back, over a reliable interface between the computer and the process; and control action computation in real time. For model-based controllers, one must also contend with inevitable model imperfections and their effect on control system performance.

This article is devoted to a detailed presentation of our approach to tackling these challenges and a discussion of our results.

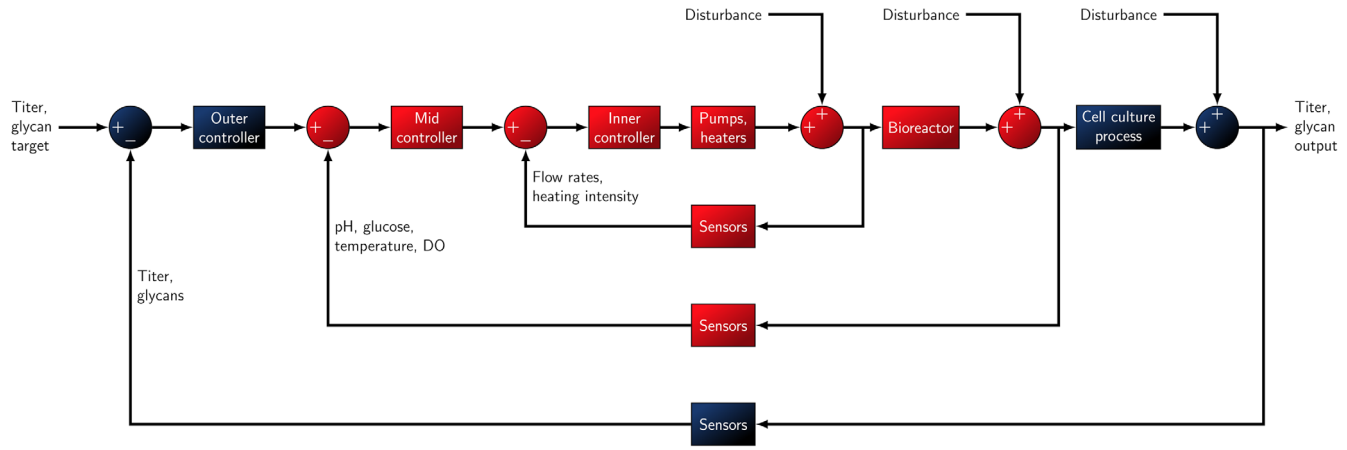

FIGURE 1 Block diagram of the proposed multivariable cascade control system.

Specifically, first, we introduce the rationale behind—and the methodology for—the design of the proposed control system and develop an appropriate fed-batch CHO cell culture process model for the specific lab scale bioreactor used in this study. Next, we show how the model's parameters are estimated from experimental data; present our approach to state estimation and controllability analysis; and then assess the performance of the control system via simulation. A detailed discussion of the experimental implementation of the control system on a laboratory scale bioreactor is deferred to the follow-up paper.

3 | METHODS

3.1 | Proposed control system

The natural structure of the mAb manufacturing bioprocess (especially the intrinsic separation of time- and length-scales in the components of the process) motivates us to propose the multivariable cascade control strategy shown in Figure 1. The overall objectives are: (1) at the base levels, to maintain the bioreactor at the desired operating conditions; and (2) at the production level, to meet the final product titer and product quality targets simultaneously.

As a result of the differing frequencies at which various measurements are available, and the differences in the natural response times of various components of the process, the cascade system involves different controller types each with different control objectives. The inner loop and the middle loop, shown in red, constitute two classic—albeit multivariable—cascade regulatory controllers designed to maintain the bioreactor environment variables (pH, glucose concentration, temperature, and dissolved oxygen (DO)) at the desired operating conditions, using multiple classical PID (proportional-integral-derivative) controllers, with the middle-loop controllers establishing the setpoints for the inner-loop controllers. In the outer loop shown in blue, productivity and product quality attributes are controlled using a multivariable model predictive controller (MPC). The MPC determines the setpoints for the middle-loop controllers. While using a standard (two-level) cascade control structure for single-input-single-output (SISO) control of process variables is common in the chemical industry, achieving effective control of product attributes (not just process variables) *and* productivity requires the introduction of the additional third level here. To the best of our knowledge, such a comprehensive three-level cascade control system does not exist in biomanufacturing.

The inner control loop consists of several regulatory PID controllers responsible for regulating the primary bioreactor input variables, such as the mass flow rates of feed material, metered through different pumps, the power of a heating unit, and the motor power for the impeller. These input variables affect such bioreactor operating conditions as the concentrations of medium components, pH, DO, temperature, and agitation speed directly. Often taken for granted, these base-level regulatory controllers are essential to effective process control because the effectiveness of higher-level, more sophisticated controllers depends significantly on the performance of these regulatory controllers. For example, a poorly tuned and slow-responding alkali flow controller that causes the alkali pump to respond too slowly to setpoint changes, will result in excessive amounts of alkali added to the reactor, with the undesirable consequence that the pH will overshoot its setpoint, regardless of how well the middle-loop controllers and the more sophisticated outer-loop MPC perform.

The middle control loop also consists of several PID controllers designed to maintain the bioreactor operating conditions, pH, DO, temperature, and stirring speed, at desired setpoints by appropriate manipulation of the bioreactor input variables (feed flow rate; heating power; motor power) that are regulated by the inner-loop controllers. Traditionally, the desired setpoints for these operating conditions are predetermined prior to the beginning of a fed-batch run and remain constant over the course of the run, unless, in an attempt to improve productivity, special protocols, such as a shifting-temperature strategy, are implemented.[23] In the proposed cascade control system, these setpoints are determined instead by the outer-loop MPC, and can change as needed.

The outer MPC uses a process model to determine explicitly the middle-loop controller setpoints required to meet production objectives, by solving an optimization problem iteratively. Because the MPC framework was developed to solve control problems with complex and competing objectives, subject to constraints, it is ideal for the highest-level control of typical bioprocesses whose objectives include simultaneously achieving a high protein titer while confining the glycan distribution within an acceptable range. At every sample time, the inputs to the MPC are the current conditions within the bioreactor, and the desired product attribute objectives; the outputs are the optimum desired values of feed rates and a set of bioreactor operating conditions, which are attained and maintained by the inner and middle control loops.

While all three levels of control are indispensable components of the overall control system, we focus primarily on the outer MPC in this article because the design and implementation of the inner-loop and the middle-loop controllers depend on mature techniques widely available in biomanufacturing. As the theory for PID controllers used in the inner and middle loops is mature, we assume that they are tuned properly and can deliver effective regulatory control.

3.2 | Process modeling

Good mathematical models are critical to the success of model-based control. In the context of biomanufacturing, models are used to predict dynamic cell behavior based on such process inputs as bioreactor operating conditions, feeding schedules, and medium compositions, all of which have been shown to affect glycosylation and cell metabolism,[24–26] and hence are potential manipulated variables for controlling productivity and product quality attributes.

Although models that can be used to predict antibody formation and/or glycosylation dynamics under predetermined process configurations[19,20,27,28] currently exist, adapting these models to new processes remains a challenging task that often involves modifying the model's structure, which can be time consuming. In a prior work,[29] we developed a modular approach to adapting an existing first-principles, mechanistic, process model, f_0, (which consists of systems of Monod and Michaelis–Menten equations), to new process conditions, without having to modify the model's structure. Instead, a supplemental input effect model, Δ, was trained using input response data from designed experiments, to "patch" f_0 such that the augmented model, f, can describe the dynamics under new process conditions more accurately. The resulting hybrid model is represented by Equation (1):

$$\frac{dx}{dt} = f(x,u_0,u;\theta,\beta) = f_0(x;\theta) + \Delta(x,u-u_0;\beta), \tag{1}$$

where x and u are, respectively, vectors of the state and input variables, and $u = u_0$ represents the process inputs at the baseline levels; θ and β are, respectively, vectors of the base model parameters and the supplemental input effect model parameters.

We adopt the same approach in developing a process model to use in the outer MPC loop to control the productivity and product quality attributes of interest. The vector of input variables, u, in Equation (1) contains agitation speed, DO, temperature, and the volume of feed medium added to the reactor during each day of a run. By contrast to prior work (e.g., in reference 29), the concentrations of a few medium supplements were modeled as input variables. While product attributes can be improved by optimizing the medium recipe, medium composition cannot be altered easily (if at all) during production. Therefore, we eliminated medium recipe from consideration as a viable manipulated variable. Bioreactor operating conditions (specifically, agitation speed, DO, and temperature) along with feed amounts, on the other hand, can be—and sometimes are—adjusted in real time during a fed-batch run. For instance, temperature shifts (i.e., changing the reactor temperature once during a fed-batch run) have been shown to increase the final titer[23,30]; an appropriate amount of DO is required to maintain culture growth and metabolism[31]; agitation rate is known to affect glucose consumption rates and viable cell density (VCD)[32]; feed rates determine the availability of nutrients, which affect cell metabolism and proliferation.

For two related reasons, the process model used to design and implement the outer-loop MPC was chosen to be a linearized approximation of f. First, the approximate linear model runs faster on a computer than the original nonlinear model, without a significant loss in prediction accuracy. Second, one of the characteristic benefits of feedback control is an intrinsic ability to compensate robustly for prediction errors arising from process/model mismatch. In addition, because MPC is a discrete-time strategy, we employ a discrete-time version of the model for the design and implementation of the outer-loop controller.

For our specific production process, the model components are as follows: the state variables—protein titer (C_{mAb}), cell densities (VCD and total cell density (TCD)), metabolite concentrations (C_{Lac}, C_{Amm}, and C_{Glc}), and glycan distributions; the input variables—feed rates and operating conditions. We did not include additional state variables—for example, concentrations of amino acids and other metabolites—in the process model as the objective was not to elucidate the underlying biological mechanisms or create a digital twin of the process. Instead, a model that contained commonly measured process variables was sufficient for illustrating MPC in this article. The linearized discrete-time cell-culture model, f_c, is shown in Equation (2):

$$\begin{aligned} x_c(k+1) &= f_c(k;x_c,u,V,V_f) \\ &= \frac{V(k)}{V(k)+V_f(k)}(A_c x_c(k) + B_c u(k)) + \frac{V_f(k)}{V(k)+V_f(k)} x_f, \end{aligned} \tag{2}$$

where the vector $x_c(k) \in \mathbb{R}^6_{\geq 0} = [C_{mAb}, VCD, TCD, C_{Lac}, C_{Amm}, C_{Glc}]^\top$ denotes the cell culture state variables at time point k (mAb, lactate, ammonium, and glucose concentrations are in g/L; VCD and TCD are in 10^6 cells/mL). One distinct feature of a fed-batch process is the periodic and abrupt cell culture volume and concentration changes due to feeding and, to a much lesser degree, sampling activities. The

vector $x_f \in \mathbb{R}^6_{\geq 0}$ denotes the concentrations of the same six cell culture state variables in the feed medium, which contains no glycans, and $V_f(k)$ denotes the volume of feed added into the culture at time point k. The cell-culture model has two parameter matrices. The matrix $A_c \in \mathbb{R}^{6\times 6}$ is an autoregressive coefficient, which relates the state variables of a previous time step to the state variables of the next time step. The matrix $B_c \in \mathbb{R}^{6\times 4}$ is an exogenous coefficient, which represents the effect of input variables on the state variables. The coefficients A_c and B_c were estimated by minimizing the model prediction error, as explained in a later section.

The linearized discrete-time glycosylation model, f_G, is shown in Equation (3):

$$\begin{aligned} x_{g,0}(k) &= A_g x_c(k) + B_g u(k), \\ x_g(k) &= \frac{x_{g,0}(k)}{1^\top x_{g,0}(k)} \times 100\%, \\ x_G(k) = f_G(1,\ldots,k;x_c,x_g,V) &= \frac{\sum_{s=1}^{k} x_g(s)(C_{\mathrm{mAb}}(s) - C_{\mathrm{mAb}}(s-1))V(s)}{C_{\mathrm{mAb}}(k)V(k)}, \end{aligned} \tag{3}$$

where the vector $x_{g,0} \in \mathbb{R}^4_{\geq 0} = [C_{\mathrm{G0F}}, C_{\mathrm{G1F}}, C_{\mathrm{G2F}}, C_{\mathrm{AFuc}}]^\top$ denotes the glycan concentrations (G0F, G1F, and G2F are three fucosylated glycan species measured in the experiments; AFuc represents all afucosylated glycan species), the vector $x_g(k) \in \mathbb{R}^4_{\geq 0} = [\mathrm{G0F}\%, \mathrm{G1F}\%, \mathrm{G2F}\%, \mathrm{AFuc}\%]^\top$ denotes the instantaneous glycan percentages of the antibodies produced between time points $k-1$ and k, and the vector $x_G(k) \in \mathbb{R}^4_{\geq 0}$ denotes the cumulative glycan percentages of all the antibodies in the reactor at time point k, calculated based on the antibody titer, C_{mAb}, and the working volume, V, values up to the time point k. The glycosylation model also has two parameter matrices. The matrix $A_g \in \mathbb{R}^{4\times 6}$ relates the cell-culture state variables to the glycosylation state variables. The matrix $B_g \in \mathbb{R}^{4\times 4}$ represents the effect of input variables on the glycosylation state variables.

The overall linearized discrete-time model, f, is shown in Equation (4):

$$\begin{aligned} x(k+1) &= f(x(k),u(k);A_c,B_c,A_g,B_g,\ldots), \\ x(k) = y(k) &= \begin{bmatrix} x_c(k) \\ x_G(k) \end{bmatrix}, \end{aligned} \tag{4}$$

where the vector y that combines x_c and x_G denotes the process output variables that can be measured (because every state variable is measurable, the dimensions of the output vector, y, and the state vector, x, are identical).

We wrote all model simulation and MPC codes from scratch in MATLAB 2019a (MathWorks, MA). No third-party MATLAB tools were used.

3.3 | Bioreactor experiments

For the process in question (described shortly), we designed two orthogonal experiments to generate data to train the process model. Initially, we designed a 2^{3-1}_{III} fractional factorial experiment where agitation, DO, and temperature were varied between low (-1) and high ($+1$) levels shown in Table 1. Subsequently, we designed a 2^2 factorial experiment to study two new factors—feed-to-volume ratio and feeding frequency. See Table 1 for a full description of the components of the two experimental designs and Figure 2 for the measurement data of titer, viability, VCD, glucose, lactate, and ammonium from the experiments.

The process on which the experiments were preformed used an IgG1-producing CHOZN cell line (provided by AMBIC). Before each bioreactor run, frozen cell stocks were thawed and added to vented-cap Erlenmeyer shake flasks at a working volume of 30 mL. Cell cultures were passaged every 3–4 days at a density of 3×10^5 cells/mL in a proprietary AMBIC 1.1 media formulation,[33] and were grown in a shaking incubator, which was maintained at 5% CO_2, 37°C, and 130 RPM. Subsequently, cells were inoculated at an initial VCD of approximately 5×10^5 cells/mL into four 1-L bench-top Applikon bioreactors (Delft, Netherlands) at a working volume of 300 mL. The setpoints of temperature, DO, and agitation rate were determined according to the experimental designs in Table 1. pH was controlled at 7.0 throughout the runs by the bioreactor's own inner- and middle-loop controllers by manipulating the flow rates of CO_2 gas and a 1 M sodium bicarbonate solution (pH 8.5). Three peristaltic feeding pumps were connected to one bottle of "AMBIC feed medium A," one bottle of "AMBIC feed medium B," and one bottle of 200 g/L glucose solution.[33] The bioreactors were sampled daily or every 2 days. VCD and percentage viability were measured using a trypan blue exclusion assay on an automated cell counter (Nexcelom Biosciences, MA). Metabolite (glucose, lactate, galactose, and ammonia) concentrations were measured using a CEDEX bioanalyzer instrument (Roche, Germany). Glucose concentration was brought back to 5 g/L every day by adding the requisite amount of 200 g/L glucose feed calculated according to the measured glucose concentration before feeding. Antibodies were purified from the supernatant of samples via Protein A, and the antibody titer was determined by a BioLogic Duoflow medium pressure chromatography instrument (Bio-Rad, CA) using a Protein-A column (Bio-Rad, CA).

4 | RESULTS AND DISCUSSION

4.1 | Parameter estimation

We estimated the parameters of the model in Equation (4) by minimizing the sum of squared differences between the measurements taken over T culture days and the corresponding model predictions using the same objective function that was used for parameter estimation in reference 29 and in the equation below. We trained the model on all but the last two data points in each fed-batch run, which were reserved for validating the model. The training was terminated when the validation error reached its lowest value, indicating that a further reduction in the training error would increase the risk of overfitting.

TABLE 1 Experimental design used to generate data for model development.

Label	Agitation (RPM)	DO (%)	Temperature (°C)	Feed ratio (%)	Feeding frequency
AF01–04	Study of agitation, DO, and temperature effects on process dynamics				
AF01	100	40	37	3	Odd-day
AF02	110	60	34	3	Odd-day
AF03	120	40	34	3	Odd-day
AF04	120	60	37	3	Odd-day
FE01–04	Study of feeding effects on process dynamics				
FE01	120	40	34	3	Odd-day
FE02	120	40	34	5	Odd-day
FE03	120	40	34	3	Daily
FE04	120	40	34	5	Daily

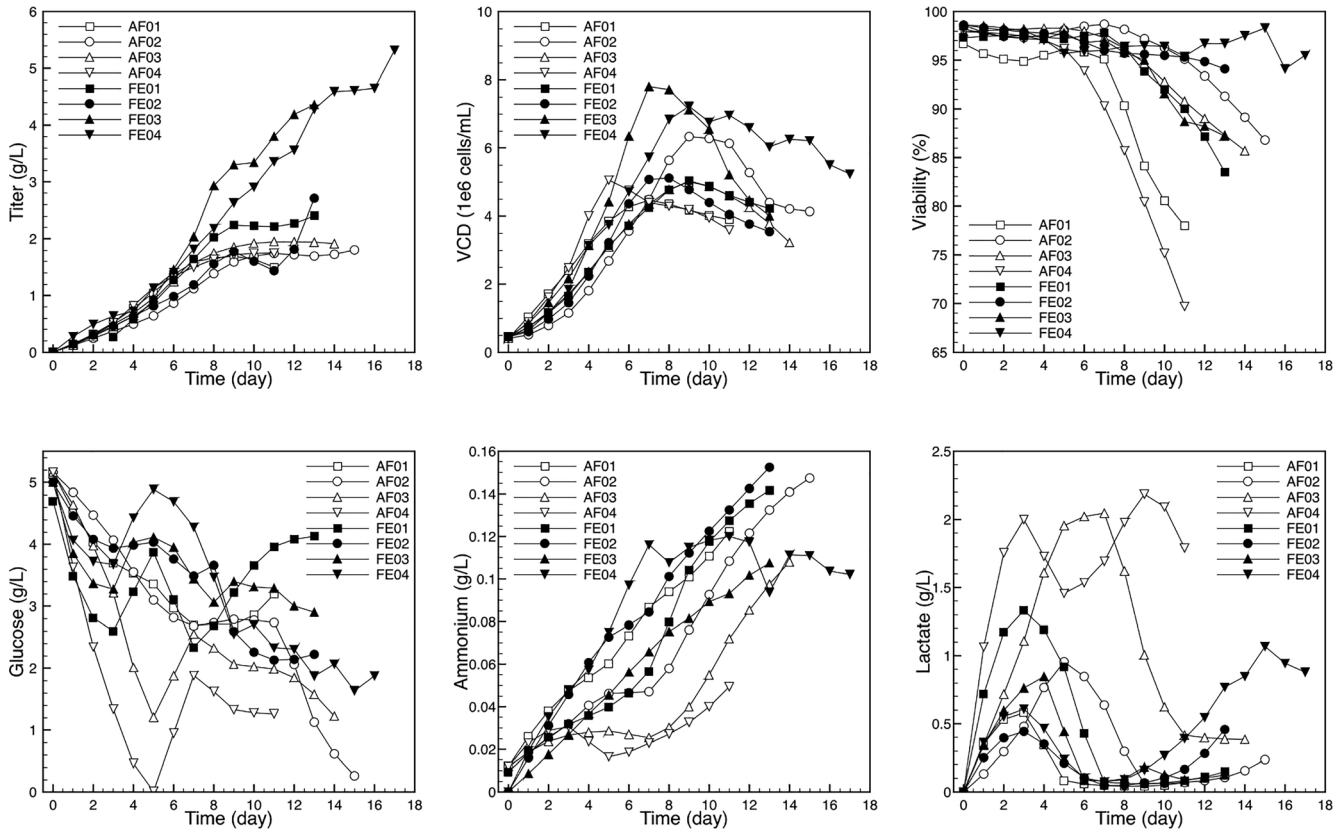

FIGURE 2 Measurement data from the experiments.

$$S = \sum_{i=1}^{n} \sum_{k=1}^{T} w_i^2 (y_i(k) - \hat{y}_i(k))^2 \quad (5)$$

where S is the objective function to be minimized; $w_1,...,w_n$ are the weights chosen to scale the n variables and to assign higher priority to variables of greater interest (e.g., titer); $y_i(k)$ is the ith variable measured at the kth time point; $\hat{y}_i(k)$ is the corresponding model prediction; T is the number of time points.

The estimated values for the model parameters—A_c,B_c,A_g, and B_g—are listed in Table S4 in Supporting Information. For the specific process in question, the inputs are agitation, DO, temperature, and feed-volume-to-working-volume ratio; the cell culture state (and output) variables are antibody concentration, VCD, TCD, lactate concentration, ammonium concentration, and glucose concentration (viability is an extra output variable calculated simply by dividing VCD by TCD); the glycosylation state (and output) variables are the percentages of antibodies glycosylated with G0F, G1F, G2F, and other afucosylated glycans (AFuc). Figure 3 shows a comparison between the model prediction and the measurement data for selected process and product variables. The relatively low prediction accuracy under certain

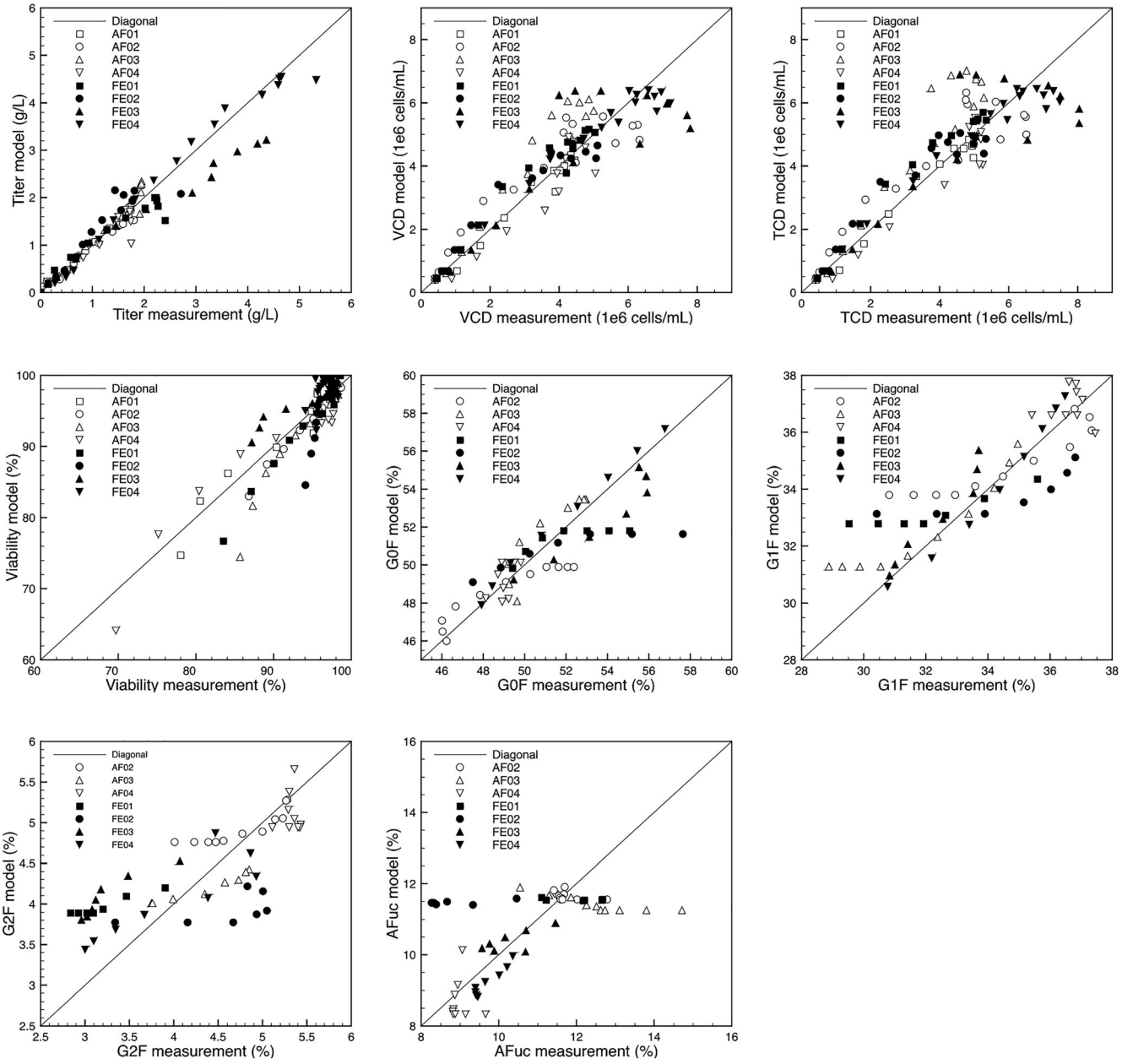

FIGURE 3 Regression plots of model prediction (y axis) versus measurement (x axis). The closer the data points align with the diagonal lines, the more accurate the model predictions are. The root-mean-square errors are 0.30 g/L (titer), 0.78 $\times$ 10^6 cells/mL (VCD), 0.88 $\times$ 10^6 cells/mL (TCD), 2.7% (viability), 1.5% (G0F), 1.2% (G1F), 0.56% (G2F), and 1.3% (AFuc).

experimental conditions can be attributed partly to the intrinsic batch-to-batch variability commonly observed in cell culture processes. For instance, even though reaction runs labeled AF03 and FE01 (see Table 1 for the corresponding experimental design conditions) were based on the same conditions, the resulting measurements differed somewhat. Such biological variability is not unusual and cannot be captured adequately by any mechanistic model. Within the context of process control, these are rightly considered as unmodelled disturbances, which, as discussed previously, are routinely corrected for by feedback control.

4.2 | Observability and controllability discussion

The model in Equation (4) is a linear time-invariant state-space model, of which the standard observability matrix and controllability matrix can be calculated easily.[34] However, the practical challenges associated with state estimation and control of bioprocesses go beyond merely calculating the observability and controllability matrices. In this section, we discuss these challenges in general, and their implications for designing effective bioprocess sensors and control systems beyond the specific process at hand.

4.2.1 | Observability

Observability, an intrinsic characteristic of a system, indicates whether the complete set of the system's internal state can be estimated from the available measured outputs. In general, a system is observable if all its state variables can be determined based on the measurements of the system output and the control action taken. The ability to determine accurately the entire system states allows us to predict the complete future trajectory of the process, which in turn can be used to calculate the optimal control action required to steer the process states to their desired target values. For bioprocesses, because process attributes such as CQAs are not measurable directly during production, state estimation is particularly important for online tracking of process attributes, to ensure that cells grow normally, and protein synthesis follows desired specifications.

For the specific process discussed in this article, the state variables are concentrations of mAb, cells, ammonia, lactate, and glucose, along with antibody galactosylation and fucosylation levels. The cell culture and glycosylation states are denoted by vectors x_c and x_G, respectively, in Equation (4). Similarly, y_c and y_G are the measured external outputs. Equation (4) is an approximate linear model developed particularly to describe the dynamics of the measured state variables only; therefore, by design, $y_c = x_c$, $y_G = x_G$, and $y = x$. In a typical fed-batch run, the cell culture state, x_c, is fully available through at-line measurements, from samples that are analyzed within minutes to hours of drawing. However, as noted shortly, these measurements are subject to noise, sometimes of magnitude too significant to ignore. The glycans, on the other hand, are typically not measured during a run but instead analyzed off-line after a run completes. Therefore, these states cannot be estimated from such end-of-run measurements during the run when the estimates are needed most. However, because of the connection to the measurable culture states, it may be possible to estimate some, or all of them, depending on the nature of the glycosylation process in question. Consequently, state estimation for the specific bioprocesses considered here involves (1) filtering process and measurement noise, and (2) estimating the full glycan state when only off-line or partial measurements are available.

For all practical processes, both model predictions and measurements are subject to noise—the former due to unmeasured disturbances and unmodelled dynamics; the latter from natural variability associated with all sensors. It is customary for modern implementations of MPC to employ a Kalman filter (Equation 6) to provide estimates of the system states, using such process models as the one in Equation (4). Equation (6) gives the general representation of a Kalman filter for a linear system with $\hat{x}$ and y being the (estimated) state and measurement vectors, respectively, f being the model, C being the matrix that relates the state vector to the measurement vector, and K being the Kalman gain.

$$\hat{x}(k+1) = f(\hat{x}(k), u(k), \ldots) + K(y(k) - Cf(\hat{x}(k), u(k), \ldots)). \quad (6)$$

From stochastic estimation theory,[34] optimal state estimates (in the least-squares sense) result from using the optimal Kalman gain, K, an optimal combination of the covariance matrices of the measurement noise and of the process (model) noise. In general, it is not always easy to determine these covariance matrices objectively. Under such circumstances, the elements of the Kalman gain matrix are selected by the user as tuning parameters, with specific chosen values influencing the performance of the Kalman filter. If neither covariance matrix is available or easily determined, and all state variables are measured (i.e., $C = 1$ by design), as is the case with our specific example, the Kalman gain matrix becomes a square matrix. Furthermore, in the absence of any objective reason to justify off-diagonal terms, the Kalman gain matrix reduces to a diagonal matrix, with individual scalar gains $0 \le \kappa_i \le 1$, one for each state variable, x_i. Selecting values for each scalar may be guided by the following considerations. When $\kappa_i = 0$, the current measurements do not influence the model prediction and hence, the estimated state will be identical to the uncorrected model prediction of the state (i.e., the filter operates in "open loop," with no feedback correction from measurements). When $\kappa_i = 1$, the estimated state is overridden entirely by the measurement. If additional information on the process noise and the observation noise is available, one can adjust each κ_i accordingly. In this work, we selected $\kappa_i = \kappa = 0.5$ as a reasonable compromise that is impartial toward either the model prediction or the measurement. In this special case, the expression in Equation (6) is simplified to the one given in Equation (7) with κ—a scalar—being the Kalman gain.

$$\hat{x}(k+1) = f(\hat{x}(k), u(k), \ldots) + \kappa(x(k) - f(\hat{x}(k), u(k), \ldots)). \quad (7)$$

When reliable online glycan measurements are unavailable during a fed-batch run, the glycan state can only be *predicted* using either the approximate linear model in Equation (4) or a full mechanistic/semi-mechanistic model, such as the one in reference 27, without the benefit of actual measurements to use in updating the estimates. However, glycosylation occurs *after* antibodies are synthesized, at a much faster rate than the rates of cell growth and metabolism. As a result, the glycosylation state depends more heavily on the cell culture state (e.g., cell density and titer determine the total amount of glycans) than on the past glycosylation states alone. Consequently, the development and implementation of useful "soft sensors" for glycans primarily requires a reliable glycosylation model that can predict the glycan distribution accurately from cell culture conditions.

4.2.2 | Controllability

Prior to designing the controller, we analyzed the controllability of our process using the method developed by reference 22 to identify which process outputs could be controlled using available manipulated variables. For most, if not all, bioprocesses, the number of outputs to be controlled will almost always be greater than the number of manipulated variables available for controlling them (particularly if the objective involves controlling the distribution glycans of which there are many species). Therefore, it is impossible to control *all* the

Max titer	0.1605	−0.1326		0.1679	0.9506
Max VCD					1.084
TCD					1.078
Lactate	0.8319	0.7639	0.6781	0.3902	
Ammonium		−0.6282			
Glucose		−0.6738			
G0F%		−0.5519		0.4155	
G1F%		0.5356	0.7882		
G2F%			0.783		
AFuc%			−0.7813	−0.938	
	Agitation	DO_2	Temperature	F1 added	F1 frequency

TABLE 2 Process gain matrix estimated from the changes in process inputs and process outputs (see reference 22 for methodology).

Note: Here the rows correspond to the state variables, and the columns correspond to the inputs. The state variables are scaled by their respective standard deviations, and the inputs are given the values −1 (low) or 1 (high).

outputs because the system lacks sufficient degrees of freedom. Specific to our process, we aim to control glycan groups that represent fucosylation, galactosylation, high mannose, or other structures instead of controlling individual glycan species independently. Consequently, we carried out a singular value decomposition (SVD) of the process gain matrix to determine which part of the system could best be controlled with the limited available degrees of freedom, and how to implement control.

The process gain matrix, K_p, connects the process outputs to the inputs, according to Equation (8), specifically indicating "by how much" the vector of process outputs will change at the end of a run (Δy), in response to a step change of Δu in the process inputs. The values of the process gain matrix elements were determined from data generated by using the experimental designs in Table 1 to change the indicated inputs, while measuring the response in titer, cell density, and antibody glycosylation (galactosylation, fucosylation) as the process outputs of interest. In interpreting the results of the data analysis using ANOVA, to balance the risk of missing truly significant gain values because of random noise against that of mistaking noise for significance, we chose to err on the side of lowering the former risk by selecting the threshold of statistical significance as $\alpha = 0.3$, as opposed to the traditional (but equally arbitrary) $\alpha = 0.05$.

$$\Delta y = K_p \Delta u. \tag{8}$$

The resulting estimate of the 10-by-5 gain matrix, K_p is shown in Table 2. An SVD of this matrix produced a rank-ordered list of output modes (linear combinations of outputs) naturally connected to input modes (linear combinations of inputs) with the associated singular values as the gain between each input–output mode pair. Rank-ordering the singular values in descending order from the largest to the smallest allows a re-arrangement and sorting of the original process variables into modes, from the most controllable mode (associated with the largest singular value) to the least controllable (associated with the smallest singular value). The results, presented in Figure 4, show the most controllable modes based on the largest 9 singular values of the process gain matrix (Figure 5), and the corresponding input modes (Figure 6).

It is instructive to examine and interpret these results and the insight they provide into the nature of the process in question and how it might best be controlled.

1. The most controllable mode, η_1, features primarily galactosylation, fucosylation, and lactate concentration, with very little else (Figure 4). This mode is connected by the highest singular value (1.89) to input mode 1, μ_1, consisting primarily of temperature, feed rate, DO, and agitation. The implication is that galactosylation and fucosylation are the most controllable product characteristics; that they can be increased (or decreased) simultaneously, without affecting titer and other product attributes much; and that such control is best carried out by manipulating temperature, feed rate, DO, and agitation.
2. The second most controllable mode, η_2, primarily involves titer, cell densities (VCD, TCD), with very little else; it is almost perfectly complementary to the first mode, η_1. This second mode is connected by the second highest singular value (1.80) to input mode 2, μ_2, which is dominated by feed frequency and not much else.
3. Such a decomposition into two virtually decoupled modes might appear fortuitous; in fact, it is in perfect keeping with the intrinsic structure of the process. Observe that the first mode is entirely about product quality; the second is about productivity. Furthermore, the primary driver for controlling the second mode (feed frequency) is also complementary to the primary drivers for the control of the first mode.

The most significant practical implication of the results of this theoretical modal analysis is that simultaneous control of the productivity attributes (titer and cell densities) and the product quality attributes (glycan distributions) is not only feasible, but the means of achieving such control have also been revealed.

The 3rd, 4th, and 5th largest singular values (1.18, 0.76, and 0.49)—nowhere as large as the first two—indicate that output mode

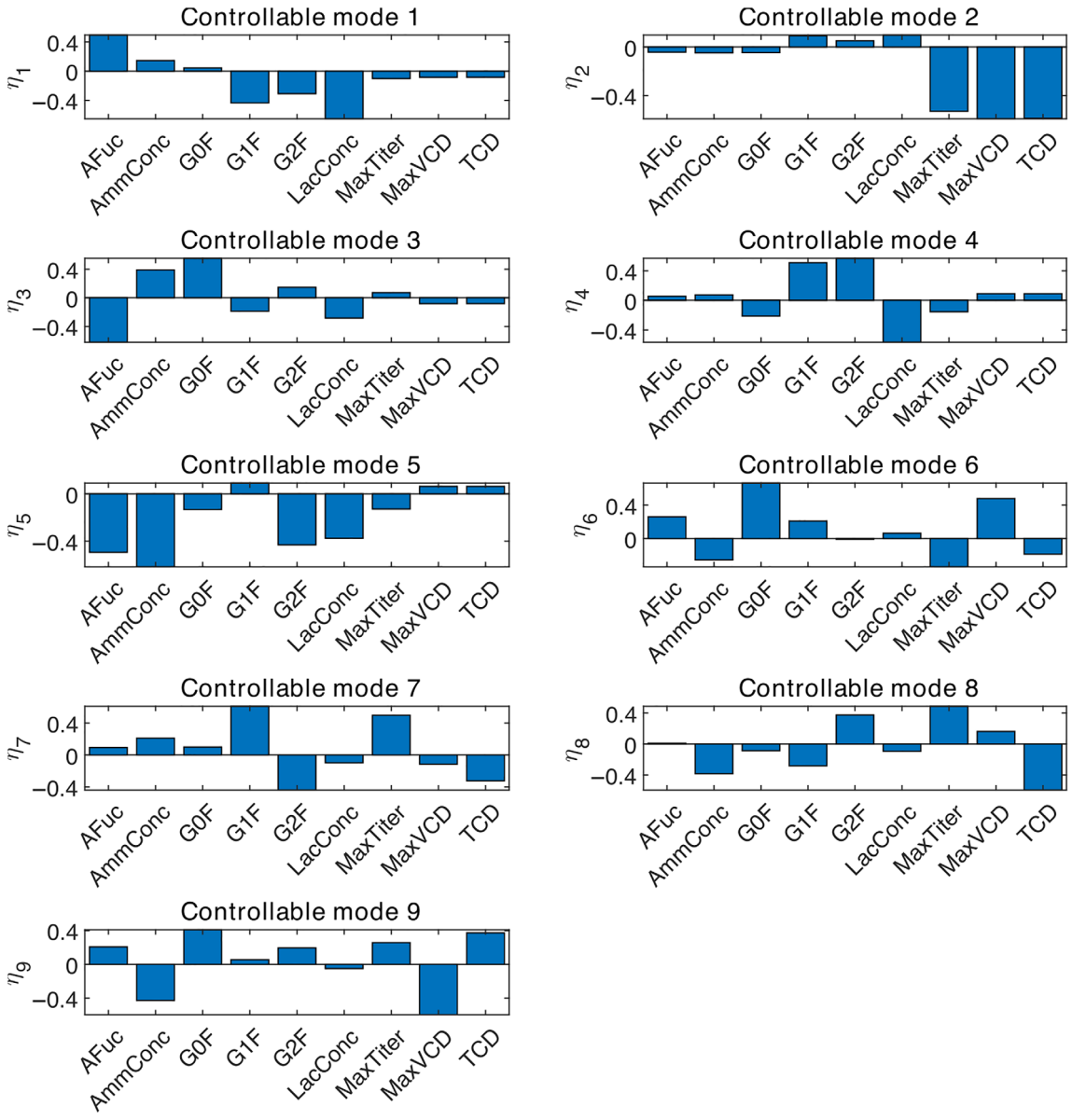

FIGURE 4 Loading plots of the most controllable modes of the system.

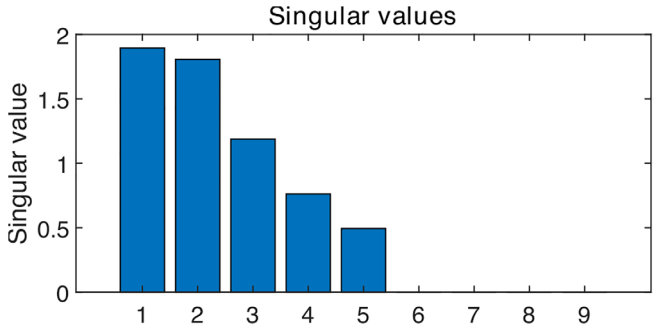

FIGURE 5 Singular values corresponding to the output modes.

3 (dominated by AFuc and G0F), output mode 4 (dominated by G1F and G2F) and to an even lesser extent, output mode 5 (dominated by ammonia, lactate, AFuc and G2F), are weakly controllable by the corresponding input modes. All other modes from the 6th to the 9th are practically uncontrollable, because of the negligible associated singular values.

4.3 | Control system performance evaluation in simulation

Prior to implementing a control system experimentally, it is customary (and wise) to evaluate the control system performance in simulation first. In this case, recall that the complete control system consists of three nested control loops—the inner loop, for base-level regulatory control; the middle loop for control of process variables such as agitation, DO, temperature, and pH; and the outer-loop

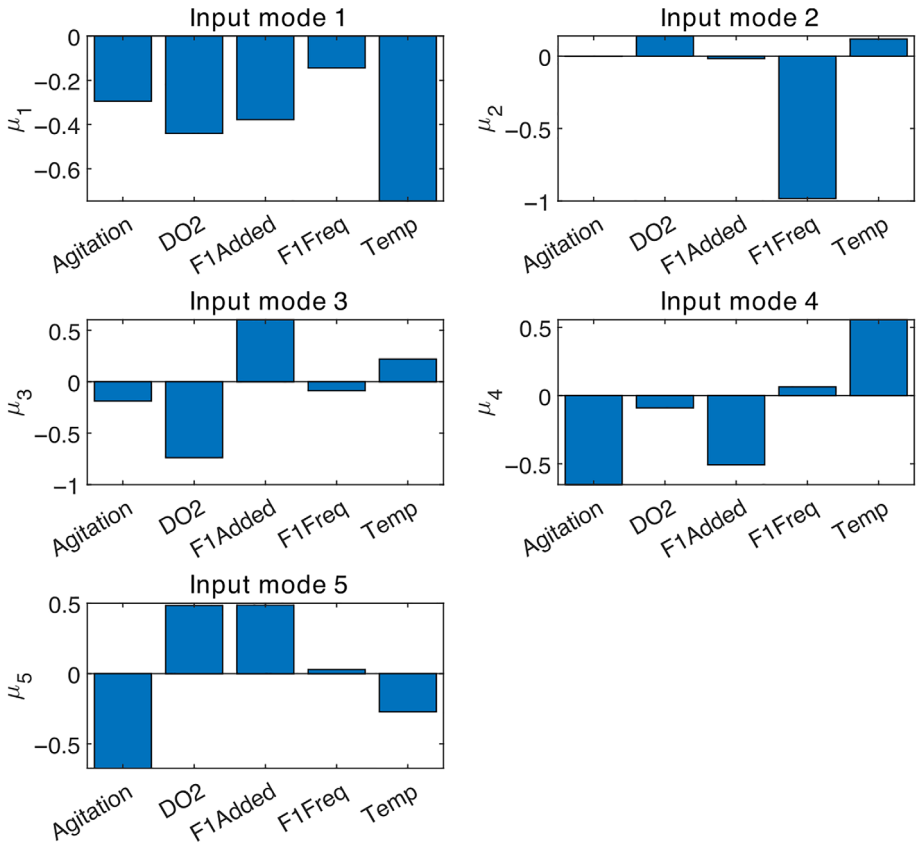

FIGURE 6 Input modes corresponding to the output modes.

MPC for higher-level control of the productivity and product quality attributes. Practically, however, for these bioprocesses we needed only to simulate the outer-loop MPC system for the following primary reasons: (i) the natural response times of the lower-level components of the bioprocess in question are of the order of seconds and minutes; on the other hand, (ii) the natural response time of the biological processes involved in product formation are much longer—on the order of several hours to days. Consequently, (iii) the MPC will not need to update its control action (the setpoints to the middle-loop controllers) any more frequently than once a day, and the lower-level processes will have reached steady state well before any new setpoint changes from the MPC. For all intents and purposes, therefore, the process to be simulated is a discrete dynamic system with a sample period of 1 day, where the lower-level processes behave like pure gain systems that achieve the desired setpoints in between samples with no noticeable dynamics. (In practice, it will be important to tune the lower-level regulatory controllers appropriately but the techniques for such tuning are widely available and familiar to most practitioners).

As an optimization-based control technique, MPC is perfectly suited for use in satisfying complex control objectives and for handling constraints explicitly. Unlike PID control, MPC requires a process model for state estimation and for prediction of future process behavior, and the required control action is determined by solving optimization problems recursively at each sample point, as the process evolves.[35] The MPC objective function, decision variables, and constraints are defined as follows. The objective function contains two cost terms: (i) the cost of the state variables deviating from the setpoint and (ii) the cost of the exerted control action. At each sample instance, the future state variables are simulated using the control action and the process model developed previously in the article. Minimizing the former ensures that the production targets are met, and minimizing the latter avoids taking overly aggressive control action in meeting the targets. The sequence of changes in the manipulated variable to be implemented as control actions constitute the decision variables. For our process, these decision variables are bounded by the lower and upper levels shown in Table 1 because the model was trained using data from experiments confined to such an input space. The optimization problem involved minimizing the objective function, J, as defined by Equation (9):

$$J = \sum_{k=1}^{T} \left\| y(k) - y_{sp}(k) \right\|_{Q}^{2} + \sum_{k=1}^{T-1} \left\| u(k) - u(k-1) \right\|_{S}^{2}, \tag{9}$$

where $y_{sp}(k)$ represents the vector of reference setpoint outputs, that is, the desired vector of outputs for the kth day; Q, and S represent the weighing matrices that penalize deviations from the reference outputs, and changes in control action, respectively. The implementation of an MPC in a batch or fed-batch process is different from that of a continuous process. For example, Nagy and Braatz[36] and Bradford and Imsland[37] used a shrinking horizon MPC to control batch processes.

To evaluate the controller performance under realistic conditions, we introduced two features:

1. *Plant/Model Mismatch*: We used the model presented earlier as the "plant model" (to represent the "true process") but used a different variation, having the same form but with different parameter values, to determine control action. This second model (or the "controller model") was obtained by replacing some of the parameters of the "plant model" with different values, which are listed in Table S5 in Supporting Information. We generated the "controller model" by relaxing the termination criteria of the optimization solver used in parameter estimation. As a result, the goodness of fit of the "controller model" is slightly worse than that of the "plant model."
2. *Measurement Noise*: Two zero-mean random variables with a small standard deviation $\left(0.05 \times 10^6\right.$ cells/mL) were added to the two cell density outputs to simulate the effect of measurement noise of the ViCell instrument. This choice is for the practical reason that cell density measurements using trypan blue are subject to higher levels of uncertainties than other at-line measurements such as titer.

We tested the MPC under two different scenarios where in Scenario 1, only titer and viability were controlled, and in Scenario 2, in addition to titer and viability, the level of fucosylation (or AFuc) was also controlled. In each scenario, the "plant model" was used to simulate the controlled process, and the MPC determined the optimal control actions for each day by solving an optimization problem using the "controller model" and adjusting its estimated process state based on the measurement feedback from the simulated process. In Table 3, the first row represents the final titer and AFuc levels that were calculated from simulating the FE02 condition (Table 1) without the MPC being active. The second and the third row represent the final titer and AFuc setpoints given to the MPC in Scenario 1 and Scenario 2.

TABLE 3 Summary of the final titer and AFuc levels under different conditions.

	Final titer (g/L)	Final afucosylation
FE02 (simulation)	2.1	11.5%
Scenario 1 (setpoint)	2.5	–
Scenario 2 (setpoint)	2.5	12%

4.3.1 | Scenario 1

In Scenario 1, we controlled the final titer on Day 13 at 2.5 g/L and maintained viability above 75% throughout the process. The titer setpoint trajectory was defined as a sigmoid that increases gradually from 50% (Day 0) to nearly 100% (end of run) of the desired final titer. To maintain viability above 75%, we specified a target for viability at

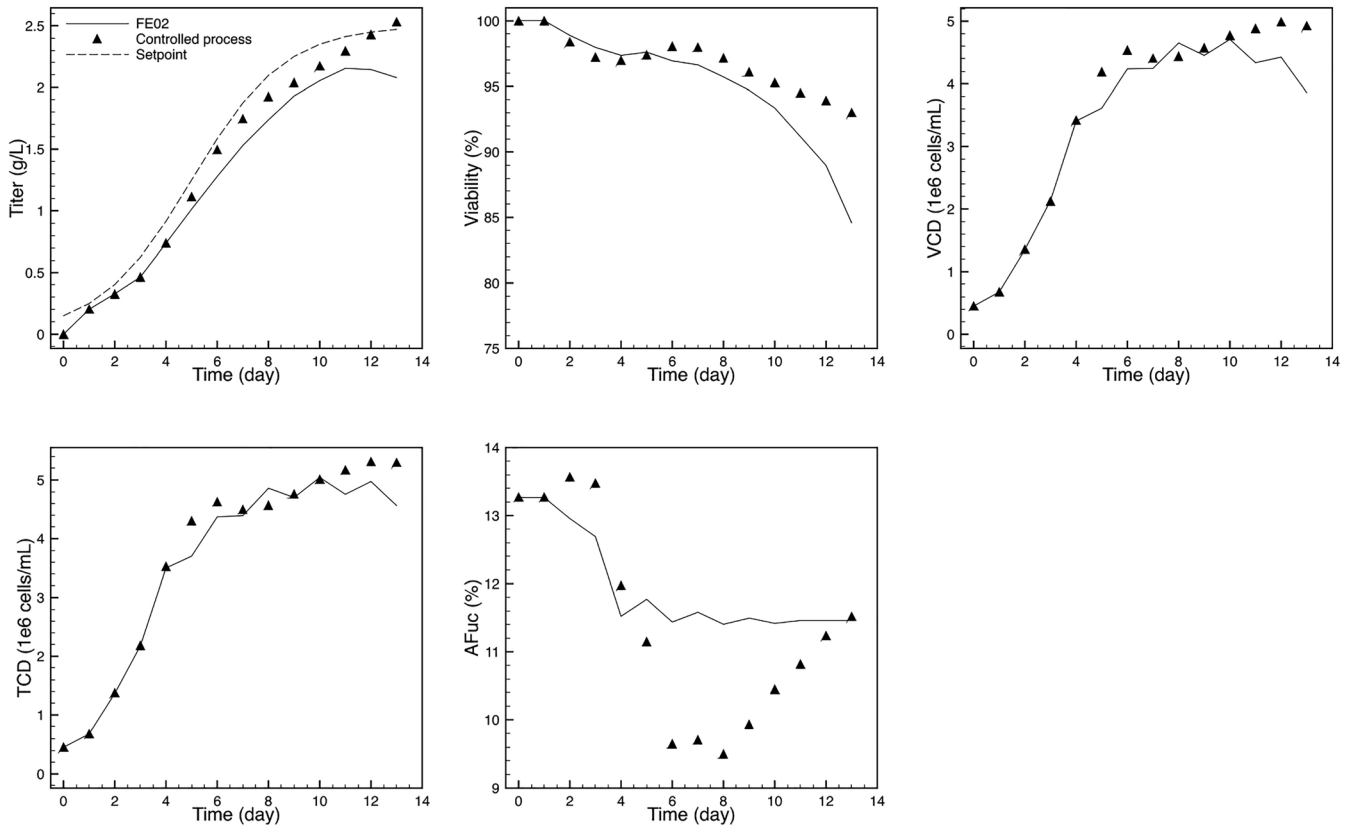

FIGURE 7 Time series of selected process and product variables under the FE02 condition (solid lines) and in Scenario 1 controlled by the MPC (triangles), and sigmoidal titer setpoint trajectory (dashed line).

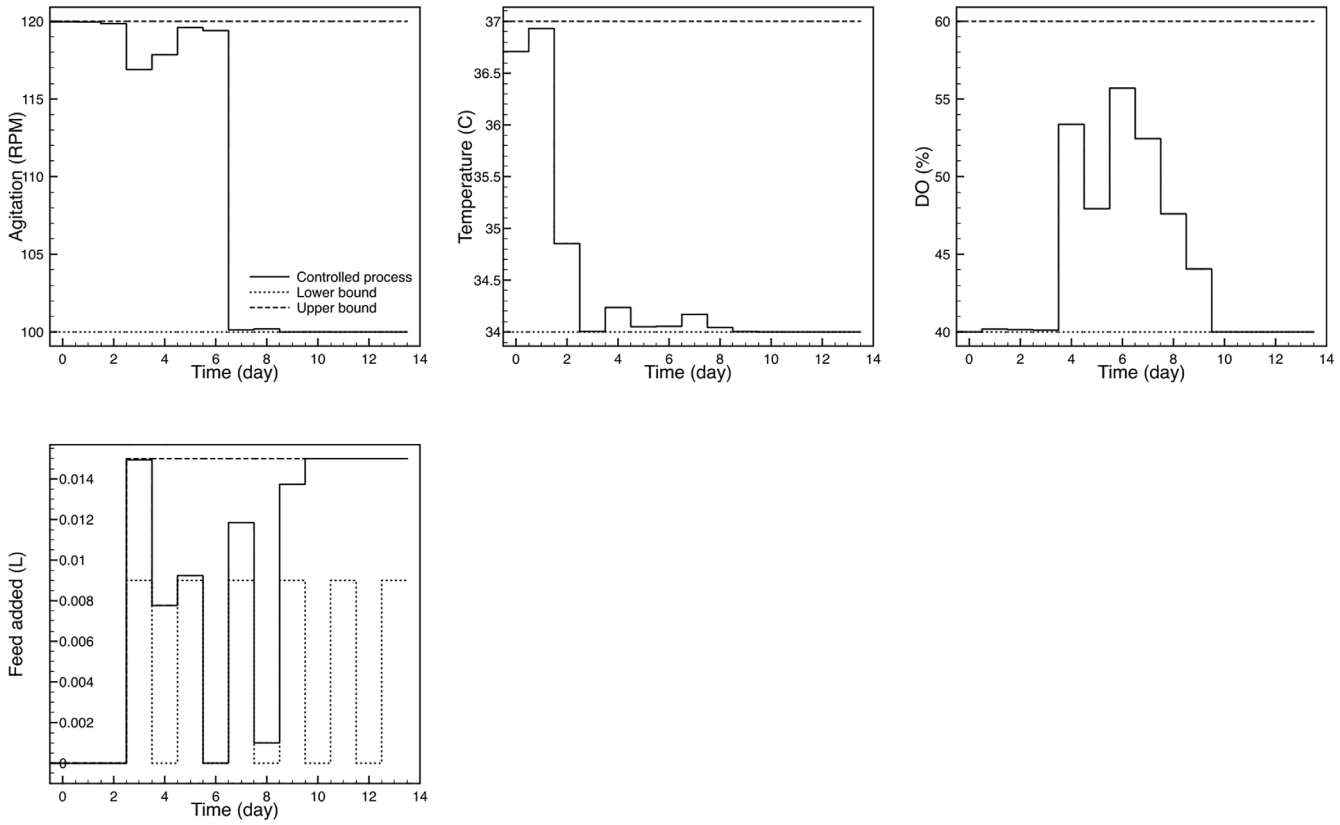

FIGURE 8 Time series of the operating conditions and the volume of feed added in Scenario 1 controlled by the MPC (solid lines), and the lower and upper bounds (dashed lines).

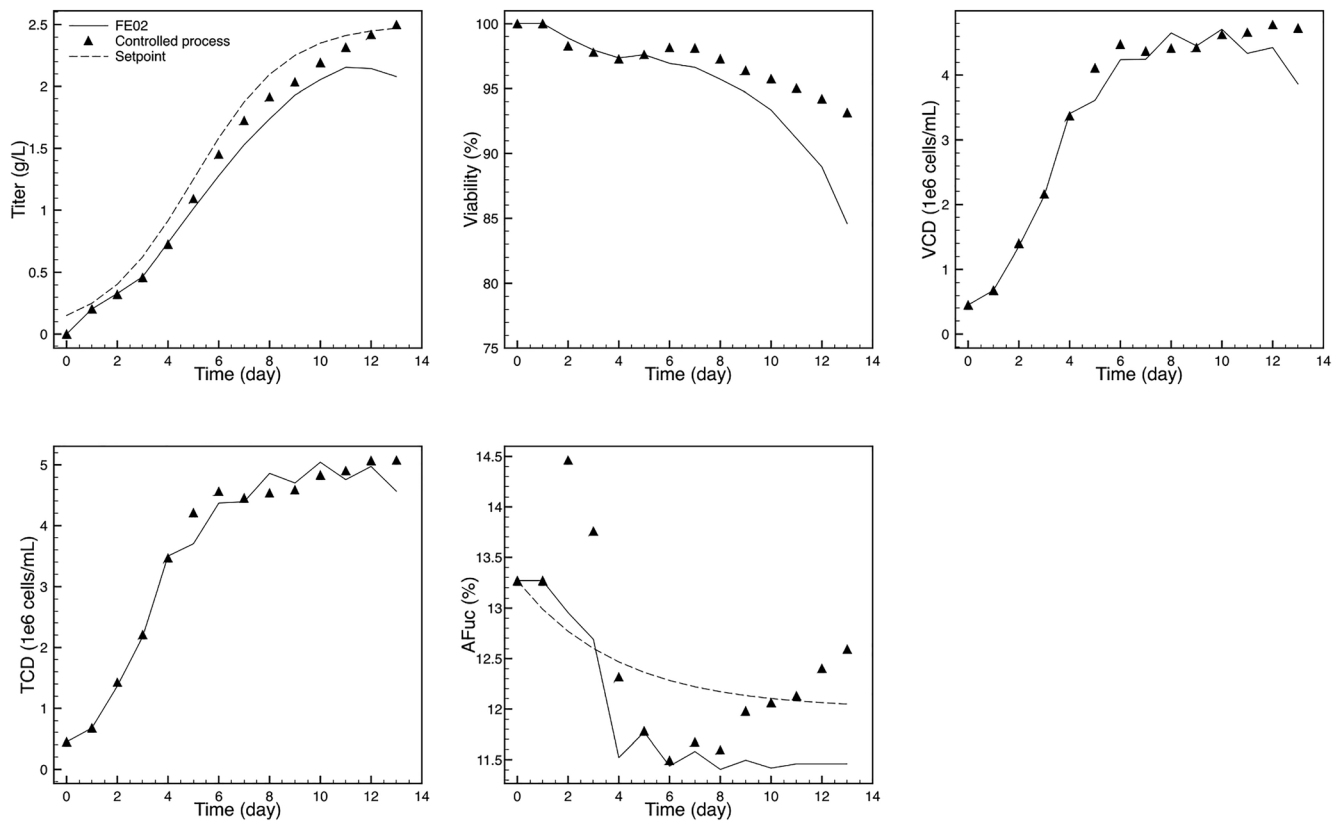

FIGURE 9 Time series of selected process and product variables under the FE02 condition (solid lines) and in Scenario 2 controlled by the MPC (triangles), sigmoidal titer setpoint trajectory (dashed line), and sigmoidal AFuc setpoint trajectory (dashed line).

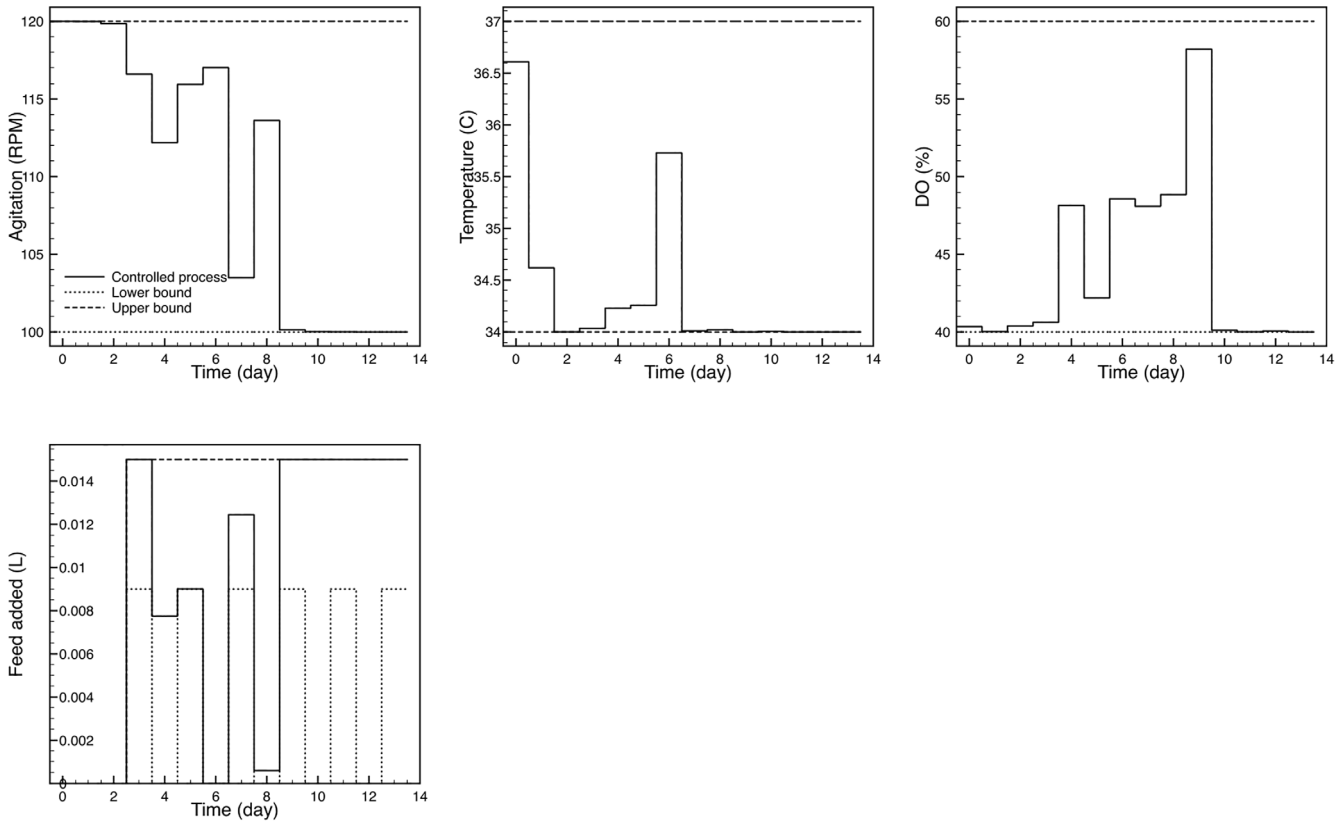

FIGURE 10 Time series of the operating conditions and the volume of feed added in Scenario 2 controlled by the MPC (solid lines), and the lower and upper bounds (dashed lines).

75%, and the first cost term in Equation (9) associated with viability becomes positive only when viability drops below 75%. In other words, the objective function does not penalize viability between 75% and 100%.

Figure 7 shows the time series of titer, viability, VCD, TCD, and AFuc levels in Scenario 1 and, for comparison, the process simulation based on the FE02 condition, without MPC. Figure 8 shows the times series of agitation, temperature, DO, and the volume of feed added. Although both VCD and TCD climbed steadily between Day 0 and Day 6, there was a drop in viability in the initial days, attributed potentially to the absence of feeding prior to Day 3. Between Day 1 and Day 4, titer increased at a slower rate than that of VCD. Between Day 4 and Day 13, when cells shifted metabolically from proliferation to production, the climb of VCD slowed down while the climb of titer accelerated. At the end of the run, the titer was close to the setpoint of 2.5 g/L. Note that the setpoint trajectory (dashed line) for the MPC is a sigmoid curve created artificially that reaches the 2.5 g/L final titer target. It is expected that the controlled process does not track the setpoint trajectory perfectly. Compared with the FE02 process, the controlled process used a slightly larger volume of the nutrient feed, and the feeding frequency was higher (which is consistent with the controllability analysis results in Figure 6 where feeding frequency is a prominent component of the input mode associated mostly with controlling product titer). Compared with the operating conditions of the FE02 process, agitation was lower between Day 7 and Day 13, temperature was higher between Day 0 and Day 2, and DO was higher between Day 4 and Day 9.

4.3.2 | Scenario 2

The fucosylation level of the antibodies is critical in improving the antibody dependent cell cytotoxicity (ADCC).[21] In Scenario 2, we controlled the final titer on Day 13 at 2.5 g/L, maintained viability above 75%, and controlled the final afucosylation level on Day 13 at 12%. The afucosylation setpoint trajectory was defined as a sigmoid that begins at the average afucosylation among all initial measurements in the reference experiment and ends at the desired final target.

Figure 9 shows the time series of titer, viability, VCD, TCD, and AFuc levels in Scenario 2, and Figure 10 shows the times series of agitation, temperature, DO, and the volume of feed added. There was a steady rise in titer after Day 4 and the titer reached its final target of 2.5 g/L on Day 13. The AFuc level surpassed its target of 12% and overshot to 12.5% at the end of the run. While the feeding profiles appeared similar in both scenarios, temperature and DO profiles were substantially different. This finding is consistent with the controllability analysis results in Figure 6 where temperature and DO are prominent components of the input mode associated mostly with controlling product glycosylation attributes. We demonstrated in the results in Scenario 1 and Scenario 2 that it is possible to control the

productivity and glycosylation attributes (independently) using the available input variables to achieve the desired production targets for this process. Note that the demonstration was performed under a limited amount of data and two specific control scenarios. Given the inherent variability of bioprocesses, how robust this control system is against disturbances and severe process/model mismatch is worth further investigation.

5 | SUMMARY AND CONCLUSIONS

In this article, we reviewed the challenges associated with controlling fed-batch mAb manufacturing processes and presented a case for model-based control of bioprocess productivity and product quality attributes, with a particular emphasis on controlling mAb product titer, viability, and glycosylation levels. We discussed the defining characteristics of these bioprocesses, the challenges they pose to effective bioprocess control, and then present a control scheme whose structure is dictated by the defining characteristics of these bioprocesses, and whose components are designed to address the noted challenges. The proposed control scheme was evaluated in MATLAB simulations, by design, reserving a detailed discussion of our implementation on a laboratory-scale reactor in the follow-up paper.

The results in this article show how controllability analysis provides unparalleled insight into mAb manufacturing processes, indicating what aspects of the process can be controlled, how well one can expect to be able to control them, and most importantly, how best to achieve such control along with which manipulated variables will be most effective.

Our simulation results indicate that with MPC, even in the presence of some modest process/model mismatch, the control system can meet productivity and product quality targets *simultaneously*, compared with the baseline process (FE02) performance without control. The ultimate test of the proposed model-based control scheme is how it performs experimentally on an actual process. We present in the follow-up paper the key challenges to practical implementation using a lab-scale bioprocess and discuss the results of the control system performance.

The lack of inline or online sensing will continue to constitute impediments to the realization of the full benefits of advanced control systems for bioprocesses that will be on par with what one routinely obtains for chemical processes. Some of the disadvantages of off-line measurements acquired via periodic manual sampling include the amount of time and labor they require, high operating cost, and loss of materials due to sampling (from both the sample volume and the dead volume). Even though addressing these challenges lies outside the intended scope of this article, we note in conclusion that the proposed scheme is applicable to and will work just as well (perhaps even better) with advanced inline or online sensing technologies, such as Raman spectroscopy, when they become routine components of bioprocesses.

AUTHOR CONTRIBUTIONS

Yu Luo: Formal analysis (equal); methodology (equal); software (equal); validation (equal); writing – original draft (lead). **Varghese Kurian:** Formal analysis (equal); investigation (equal); methodology (equal); validation (equal); writing – original draft (supporting); writing – review and editing (supporting). **Liqing Song:** Data curation (supporting); formal analysis (equal); methodology (supporting); writing – review and editing (supporting). **Evan A. Wells:** Data curation (supporting); formal analysis (supporting); investigation (supporting); validation (supporting); writing – review and editing (supporting). **Anne Skaja Robinson:** Conceptualization (lead); funding acquisition (lead); project administration (lead); supervision (lead); writing – review and editing (supporting). **Babatunde A. Ogunnaike:** Conceptualization (lead); formal analysis (lead); funding acquisition (supporting); supervision (lead); writing – review and editing (lead).

ACKNOWLEDGMENTS

This work was performed under a Project Award Agreement from the National Institute for Innovation in Manufacturing Biopharmaceuticals (NIIMBL). The authors are also grateful to AMBIC for providing the CHOZN cell line and the Ambic 1.1 media used in the study.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding author upon reasonable request.

ORCID

Yu Luo https://orcid.org/0000-0002-1379-6373
Varghese Kurian https://orcid.org/0000-0001-5881-5841
Evan A. Wells https://orcid.org/0000-0002-4039-2921
Anne Skaja Robinson https://orcid.org/0000-0001-7235-1481
Babatunde A. Ogunnaike https://orcid.org/0000-0002-8246-070X

REFERENCES

1. Lu R-M, Hwang YC, Liu IJ, et al. Development of therapeutic antibodies for the treatment of diseases. *J Biomed Sci*. 2020;27(1):1-30.
2. Xu J, Xu X, Huang C, et al. Biomanufacturing evolution from conventional to intensified processes for productivity improvement: a case study. *MAbs*. Vol 12. Taylor & Francis; 2020.
3. Torkashvand F, Vaziri B. Main quality attributes of monoclonal antibodies and effect of cell culture components. *Iran Biomed J*. 2017; 21(3):131-141.
4. Jefferis R. Glycosylation of recombinant antibody therapeutics. *Biotechnol Prog*. 2005;21(1):11-16.
5. Beck A, Wagner-Rousset E, Bussat MC, et al. Trends in glycosylation, glycoanalysis and glycoengineering of therapeutic antibodies and Fc-fusion proteins. *Curr Pharm Biotechnol*. 2008;9(6):482-501.
6. Reusch D, Tejada ML. Fc glycans of therapeutic antibodies as critical quality attributes. *Glycobiology*. 2015;25(12):1325-1334.
7. Webster TA, Hadley BC, Dickson M, Busa JK, Jaques C, Mason C. Feedback control of two supplemental feeds during fed-batch culture on a platform process using inline Raman models for glucose and phenylalanine concentration. *Bioprocess Biosyst Eng*. 2021;44(1):127-140.
8. W Eyster T et al. Tuning monoclonal antibody galactosylation using Raman spectroscopy-controlled lactic acid feeding. *Biotechnol Prog*. 2021;37(1):e3085.
9. Luo Y, Kurian V, Ogunnaike BA. Bioprocess systems analysis, modeling, estimation, and control. *Curr Opin Chem Eng*. 2021;33:100705.
10. Xie L, Wang DI. Fed-batch cultivation of animal cells using different medium design concepts and feeding strategies. *Biotechnol Bioeng*. 1994;43(11):1175-1189.
11. Lu F, Toh PC, Burnett I, et al. Automated dynamic fed-batch process and media optimization for high productivity cell culture process development. *Biotechnol Bioeng*. 2013;110(1):191-205.

12. Mellahi K, Brochu D, Gilbert M, et al. Assessment of fed-batch cultivation strategies for an inducible CHO cell line. *J Biotechnol*. 2019; 298:45-56.
13. Wlaschin KF, Hu W-S. Fedbatch culture and dynamic nutrient feeding. *Adv Biochem Eng Biotechnol*. 2006;101:43-74.
14. Sou SN, Sellick C, Lee K, et al. How does mild hypothermia affect monoclonal antibody glycosylation? *Biotechnol Bioeng*. 2015;112(6): 1165-1176.
15. Hogiri T, Tamashima H, Nishizawa A, Okamoto M. Optimization of a pH-shift control strategy for producing monoclonal antibodies in Chinese hamster ovary cell cultures using a pH-dependent dynamic model. *J Biosci Bioeng*. 2018;125(2):245-250.
16. Kotidis P, Jedrzejewski P, Sou SN, et al. Model-based optimization of antibody galactosylation in CHO cell culture. *Biotechnol Bioeng*. 2019; 116(7):1612-1626.
17. Ogunnaike BA. On-line modelling and predictive control of an industrial terpolymerization reactor. *Int J Control*. 1994;59(3):711-729.
18. Baughman AC, Huang X, Sharfstein ST, Martin LL. On the dynamic modeling of mammalian cell metabolism and mAb production. *Comput Chem Eng*. 2010;34(2):210-222.
19. Krambeck FJ, Betenbaugh MJ. A mathematical model of N-linked glycosylation. *Biotechnol Bioeng*. 2005;92(6):711-728.
20. Spahn PN, Hansen AH, Hansen HG, Arnsdorf J, Kildegaard HF, Lewis NE. A Markov chain model for N-linked protein glycosylation—towards a low-parameter tool for model-driven glycoengineering. *Metab Eng*. 2016;33:52-66.
21. Wells E, Song L, Greer M, et al. Media supplementation for targeted manipulation of monoclonal antibody galactosylation and fucosylation. *Biotechnol Bioeng*. 2020;117(11):3310-3321.
22. St. Amand MM et al. Controllability analysis of protein glycosylation in CHO cells. *PLoS One*. 2014;9(2):e87973.
23. Xu J et al. Systematic development of temperature shift strategies for Chinese hamster ovary cells. *MAbs*. Taylor & Francis; 2019.
24. Crowell CK, Grampp GE, Rogers GN, Miller J, Scheinman RI. Amino acid and manganese supplementation modulates the glycosylation state of erythropoietin in a CHO culture system. *Biotechnol Bioeng*. 2007;96(3):538-549.
25. Hossler P. Protein glycosylation control in mammalian cell culture: past precedents and contemporary prospects. In: Hu W, Zeng AP, eds. *Genomics and Systems Biology of Mammalian Cell Culture. Advances in Biochemical Engineering Biotechnology*. vol 127. Springer. 2011.
26. St. Amand MM et al. Identification of manipulated variables for a glycosylation control strategy. *Biotechnol Bioeng*. 2014;111(10):1957-1970.
27. Jimenez del Val I, Nagy JM, Kontoravdi C. A dynamic mathematical model for monoclonal antibody N-linked glycosylation and nucleotide sugar donor transport within a maturing Golgi apparatus. *Biotechnol Prog*. 2011;27(6):1730-1743.
28. Kontoravdi C, Asprey SP, Pistikopoulos EN, Mantalaris A. Development of a dynamic model of monoclonal antibody production and glycosylation for product quality monitoring. *Comput Chem Eng*. 2007; 31(5–6):392-400.
29. Luo Y, Lovelett RJ, Price JV, et al. Modeling the effect of amino acids and copper on monoclonal antibody productivity and glycosylation: a modular approach. *Biotechnol J*. 2021;16(2):2000261.
30. Abu-Absi SF, Yang LY, Thompson P, et al. Defining process design space for monoclonal antibody cell culture. *Biotechnol Bioeng*. 2010; 106(6):894-905.
31. Li F, Vijayasankaran N, Shen A(Y), Kiss R, Amanullah A. Cell culture processes for monoclonal antibody production. *MAbs*. Vol 2. Taylor & Francis; 2010:466-479.
32. Erklavec Zajec V, Novak U, Kastelic M, et al. Dynamic multiscale metabolic network modeling of Chinese hamster ovary cell metabolism integrating N-linked glycosylation in industrial biopharmaceutical manufacturing. *Biotechnol Bioeng*. 2021;118(1):397-411.
33. Cordova LT et al. Generation of reference cell lines, media, and a process platform for CHO cell biomanufacturing. *Biotechnol Bioeng*. 2022; 120:715-725.
34. Ray WH. *Advanced Process Control (Chemical Engineering)*. McGraw-Hill; 1980.
35. Rawlings JB, Risbeck MJ. Model predictive control with discrete actuators: theory and application. *Automatica*. 2017;78:258-265.
36. Nagy ZK, Braatz RD. Robust nonlinear model predictive control of batch processes. *AIChE J*. 2003;49(7):1776-1786.
37. Bradford E, Imsland L. Output feedback stochastic nonlinear model predictive control for batch processes. *Comput Chem Eng*. 2019;126: 434-450.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Luo Y, Kurian V, Song L, Wells EA, Robinson AS, Ogunnaike BA. Model-based control of titer and glycosylation in fed-batch mAb production: Modeling and control system development. *AIChE J*. 2023;69(4):e18075. doi:10.1002/aic.18075