2015 IEEE Conference on Control Applications (CCA)
Part of 2015 IEEE Multi-Conference on Systems and Control
September 21-23, 2015. Sydney, Australia

Control Systems Technology in the Advanced Manufacturing of Biologic Drugs

Amos E. Lu[1†], Joel A. Paulson[1†], Nicholas J. Mozdzierz[1], Alan Stockdale[1],
Ashlee N. Ford Versypt[2], Kerry R. Love[1], J. Christopher Love[1], and Richard D. Braatz[1‡]

***Abstract*— The paper focuses on systems engineering challenges for biomanufacturing, as exemplified by the Integrated and Scalable Cyto-Technology (InSCyT) bio-manufacturing platform under the Defense Advanced Research Projects Agency (DARPA) Biologically-derived Medicines on Demand (Bio-MOD) program. One goal of the project is to apply modeling and simulation techniques to the design, control, and optimization of biomanufacturing operations. In this paper, we present models and control strategies for the unit operations within the InSCyT platform, including a perfusion bioreactor, several packed bed and membrane chromatography steps, and a conductivity- and pH-controlled buffer mixing unit.**

I. Introduction

A. Introduction to Biopharmaceuticals

Biopharmaceuticals are products of biological organisms used in the treatment of disease. There are a total of 246 biopharmaceutical products licensed within the United States and the European Union as of 2014 [1]. Total sales within the United States reached $63.6 billion in 2012 [2], an 18.2% increase over the previous year. The share of biologics are projected to increase from 18% in 2012 to 20% in 2017, relative to the overall share of pharmaceuticals [1]. It is clear that biologics will continue to be relevant into the foreseeable future, which strongly motivates the development of rapid (yet safe and robust) methods for their production.

B. Manufacturing Biologics on the InSCyT Platform

The Integrated and Scalable Cyto-Technology (InSCyT) platform is a prototype biomanufacturing platform developed under the Defense Advanced Research Projects Agency (DARPA) Biologically-derived Medicines on Demand (Bio-MOD) program, the goal of which is to develop a platform for rapid production and purification of protein biologics. An overview of the InSCyT platform is given in Figure 1. The InSCyT platform begins with genetically engineered *Pichia pastoris* inoculum and sterile media. The cells are inoculated into a perfusion bioreactor and outgrown to attain desired biomass. Following that, the media will be switched to induce biologic production and secretion into the cell culture fluid. Operating under perfusion, a continuous flow of media is introduced into the bioreactor while removing supernatant. This supernatant will undergo further purification, while the cells are physically retained within the reactor.

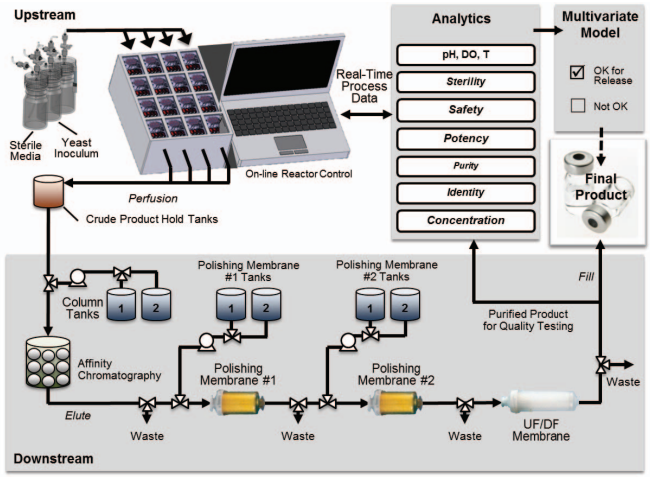

Fig. 1. InSCyT Platform Overview

In the first downstream unit operation, a peptide based affinity resin will be used to separate the protein of interest from the cell culture fluid using a bind and elute type chromatography process. A further two polishing steps will be used to remove host cell proteins (HCPs) and residual DNA from the protein of interest. Finally, a ultrafiltration and diafiltration membrane will be used to concentrate the protein of interest and effect a buffer exchange into the final formulation buffer.

The platform also includes a suite of analytics aimed at validating the product for real time release. This will include traditional UV/Vis spectrophotometer measurements for protein concentration, as well as novel Raman spectroscopy, carbon nanotube sensors, and nanochannel arrays for identity and purity determination.

The unit operations will be highly integrated, with automatic material handling and transfer between unit operations. In addition, the platform aims to produce single dose quantities of drug within a 24 hour time span. Lastly, the platform is developed for deployment in the field, potentially exposing the unit to a variety of environmental conditions that will induce disturbances. This motivates a quality by design (QbD) approach to ensure product quality.

Integrated strategies have been used for the biomanufacturing of proteins. Warikoo et al. [3] have demon-

Financial support is acknowledged from the Defense Advanced Research Project Agency (DARPA).

[1]Massachusetts Institute of Technology, Cambridge, MA 02139, USA {amoslu,jpaulson,mozdzn,astock,kerryluv, clove,braatz}@mit.edu

[2]Oklahoma State University, Stillwater, OK 74078, USA ashleefv@okstate.edu

[†]These authors contributed equally to this work.

[‡]To whom correspondence should be addressed.

strated integrated continuous production of a monoclonal antibody and a recombinant human enzyme in Chinese hamster ovary (CHO) cell culture using a perfusion bioreactor and a periodic counter-current chromatography purification setup.

C. Importance and Relevance of Work

Quality by design is a strategy adopted by the FDA to encourage a more scientific and proactive approach to pharmaceutical and biopharmaceutical development [4]. Under this framework, process development would involve defining a target product profile (TPP), identifying the critical quality attributes (CQAs) associated with the TPP, and understanding the impact of critical process parameters (CPPs) on the CQAs. This knowledge can then be used to design control systems to achieve quality in the final drug product.

The aim of modeling is to represent physical phenomena via mathematical means. This quantitative understanding of the process represents a high level of process understanding and is critical to the design of an effective control strategy for quality. Models can be broadly divided into two classes: data-based models or first principles models [5]. Data-based models aim to describe the behavior of the process through the use of data collected during model development [5]. Techniques such as multiple linear regression [6], principal component analysis (PCA) [7], partial least squares (PLS) [8], and artificial neural networks (ANN) [9] have been applied to the modeling of biomanufacturing processes.

The key benefits of data-based models are that they require less fundamental and mechanistic knowledge compared to first principles models, and thus, can be more straightforwardly developed [5]. This may explain their prevalence in the pharmaceutical industry [10]. However, the models perform poorly on extrapolation and do not yield mechanistic knowledge about the underlying processes, limiting its applicability to predictive design and optimization.

First principles models, on the other hand, aim to both explain and describe the behavior of the process [5]. This involves combining mathematical descriptions of the mass, energy, and momentum conservation laws, thermodynamics, chemical and biological kinetics, and the dynamics of transport phenomena into a model that describes the unit operation. The mechanistic considerations within the model allow them to be used for process synthesis, optimization, and scale up. Due to these advantages, this work will focus on the development and application of first principles models.

In the bioprocessing industry, most work has revolved around modeling unit operations [11]. Models for bioreactors [5], [12] and chromatography [13], [14] are mature and will be briefly reviewed in the subsequent subsections. However, the unit operations do not operate in isolation, and changing a unit's operating conditions will impact downstream processing [15], [16]. This motivates integrated plant models.

Some simple integrated models have been developed in literature. Groep et al. [17] have modeled the production of alcohol dehydrogenase under aerobic fermentation with *Saccharomyces cerevisiae*. Varga et al. [18] have modeled the production of a protein engineered form of alcohol dehydrogenase in *Saccharomyces cerevisiae*. Chhatre et al. [19] have modeled the production of CroFab, an FDA-approved IgG fragment biotheraputic, from ovine serum. Pinto et al. [20] have modeled a multi product protein production plant with *Saccharomyces cerevisiae*.

However, plant-wide predictive models are still lacking in literature [21]. To the author's knowledge, there have been no publications documenting an integrated model for both upstream and downstream biopharmaceutical using any of the following: (i) state-of-the-art models in chromatography, (ii) bioreactors with *Pichia pastoris*, or (iii) the use of models for integrated design and control.

The aims of this work (i) attempt to address the lack of work in the modeling and application of model-based design and control in biomanufacturing and (ii) provide a case study on the application of process systems engineering principles to a real example of a biomanufacturing process. With the regulatory agencies emphasis on QbD and the continued importance of biopharmaceuticals into the future, this work will be of great relevance to the biopharmaceutical industry.

D. Paper Overview

For this paper, we will focus on modeling and control of the individual unit operations of the InSCyT platform. Relevant modeling and control work is explored for the bioreactors, chromatography separation train, and conductivity- and pH-controlled buffer mixing system in Sections II, III, and IV, respectively. As future work, these models will be linked into an integrated plant model that can be used to perform optimal design, to analyze tradeoffs between the operation of the individual process units, and to develop plant-wide control strategies that meet specifications with reasonable performance and a desired level of robustness.

II. Bioreactor Modeling and Control

A. Dynamic model

Bioreactors are traditionally modeled in a similar way to that of traditional stirred tank chemical reactors [5], [12]. The following mass balance equation may be used:

$$\frac{d(Vy)}{dt} = F_i y_i + Q_i y_{g,i} - F_o \delta y - Q_o Y_{g,o} + V r_y \qquad (1)$$

where V is the media volume, y is the concentration of a component within the reactor, F is the media flow rate, Q is the gas flow rate, the subscripts i and o referring to in and out, the subscript g referring to the gas phase, δ is the separation factor at the outlet and r_y is the volumetric reaction rate of the component [22].

Biological reaction rate models can be classified as unstructured or structured, and as nonsegregated or segregated. In unstructured models, the reaction rate depends only on concentrations external to the cell, while in structured models, the reaction rate is additionally dependent on the concentrations in intercellular components. In nonsegregated models, all cells are taken to be identical, while in segregated

models, there may be multiple populations of cells with different characteristics.

Jahic et al. [23] have modeled the growth and energy metabolism of *Pichia pastoris* producing a fusion protein under glycerol and methanol feed using an unstructured, nonsegregated model. They modeled the kinetics with differential balances for oxygen (O), substrate (glycerol or methanol, S), and biomass (X). For the fed batch system,

$$\frac{dX}{dt} = \left(-\frac{F}{V} + \mu\right) X \tag{2}$$

$$\frac{dS}{dt} = \frac{F}{V}(S_i - S) - q_s X \tag{3}$$

$$\frac{dO}{dt} = q_O X V \tag{4}$$

where q_s is the specific substrate consumption rate, and q_O is the specific oxygen consumption rate.

Other types of models have also been discussed. Bibila et al. [24] used a structured dynamic model to identify rate limiting steps to antibody production in hybridoma cells. Craven et al. [25] used an unstructured, segregated model, based on Monod kinetics, to perform reactor scale up. Nolan et al. [26] used a structured dynamic model, containing 24 metabolites and 34 reactions, to predict the effects of temperature shift, seed density, and nutrient concentrations. Selvarasu et al. [27] used a combined metabolomics and modeling approach to gain insight into intercellular metabolism and identify key metabolic pathways limiting cell growth. Fernandes et al. [28] reviewed the application of population balance models (PBM) and computational fluid dynamics (CFD) to segregated models. Jimenez del Val et al. [29] presented a model for glycosylation within the Golgi apparatus, a CQA for therapeutic proteins.

B. Control of Dissolved Oxygen

The aforementioned bioreactor models, once accurately trained with process data, can be used in advanced control strategies e.g., model predictive control (MPC) to achieve high performance operation. First, a regulatory control layer should be designed to ensure that the underlying process is stabilized during development of higher-level (model-based) control strategies. In this section, we will describe an example of a regulatory controller for the dissolved oxygen content of a medium using the impeller stirrer speed as a manipulated variable for a fixed feed rate of oxygen.

Dissolved oxygen (DO) content (in percent) is defined as the relative amount of oxygen in a medium to that of an oxygen-saturated medium. During the outgrowth phase, *Pichia pastoris* will consume oxygen as it grows which, in turn, reduces the DO content in the biomass. However, it is important to maintain DO at a desired level to ensure the cell culture continues to thrive in the bioreactor environment. To avoid foaming, we would like to keep the feed rate of oxygenated gas constant. Since the speed of stirring effects the diffusion of gas into the culture, we can use this as our manipulated variable at a fixed feed of oxygenated gas.

We look to actively control the DO content using a proportional-integral-derivative (PID) controller, of the form

$$u(t) = K_p e(t) + K_i \int_0^t e(\tau)d\tau + K_d \frac{d}{dt} e(t) \tag{5}$$

where $u(t)$ is the stirrer speed, $e(t) = y(t) - y_{sp}(t)$ is the error defined as the difference between the DO measurement $y(t)$ and its setpoint $y_{sp}(t)$, and K_p, K_i, and K_d are the proportional, integral, and derivative gains, respectively. The gains K_p, K_i, and K_d are tuning parameters that should be chosen to achieve acceptable closed-loop performance. We look to use the internal model control (IMC) tuning strategy for the selection of these gains [30]. To do this, we must first obtain a process model of the system. By measuring how DO responds to a step change in stirrer speed, we can fit a first-order process model to this data (Fig. 2) i.e.,

$$G_p(s) = \frac{Ke^{-\theta s}}{\tau s + 1} \tag{6}$$

where $G_p(s) = Y(s)/U(s)$ denotes a transfer function model of the process, $Y(s)$ and $U(s)$ denote Laplace transforms of the DO content and stirrer speed, respectively, K denotes the process model gain, τ denotes the process model time constant, and θ denotes the dead time (i.e., time delay).

If the model of the process $G_p(s)$ is perfect, then we would like to choose a controller that is exactly the inverse of this process $G_p(s)^{-1}$. Tuning a controller using this principle is known as an IMC design strategy. In practice, however, there will be mismatch between the plant and model, the output will be affected by unknown disturbances, and the model may not be invertible. A feedback control strategy, based on practical IMC design, can be adopted to mitigate these affects. Approximating the time-delay term using a first-order Taylor series expansion i.e., $e^{-\theta s} \approx 1 - \theta s$, IMC tuning rules can be used to derive expressions for the PID gains [30]

$$K_p = \frac{\tau}{K(\theta + \tau_f)}, \quad K_i = \frac{1}{K(\theta + \tau_f)}, \quad K_d = 0 \tag{7}$$

where τ_f represents the time constant of a first-order low pass filter placed in series with the IMC controller, which is included to attenuate effects of plant-model mismatch that usually occur at high frequencies.

The measured DO content in the cell culture, controlled using this IMC-tuned PID, is shown in the top plot of Fig. 3 for a particular experimental run of the bioreactor. Note that an open-loop strategy is used for the first five hours after inoculation (at time zero). The DO setpoint of 20% is tracked fairly well over a fifty hour experiment. The minor oscillations that appear at certain points in the experiment are likely due to the evolution of the cell culture over time, which effectively change the process parameters K, τ, and θ. A gain scheduling method could be employed to achieve even better performance in these cases.

For comparison, we also show the measured DO content using a standard cascade control algorithm in the bottom plot of Fig. 3. Notice that significant oscillations (from nearly 0% to 100%) are present for almost the entire experiment.

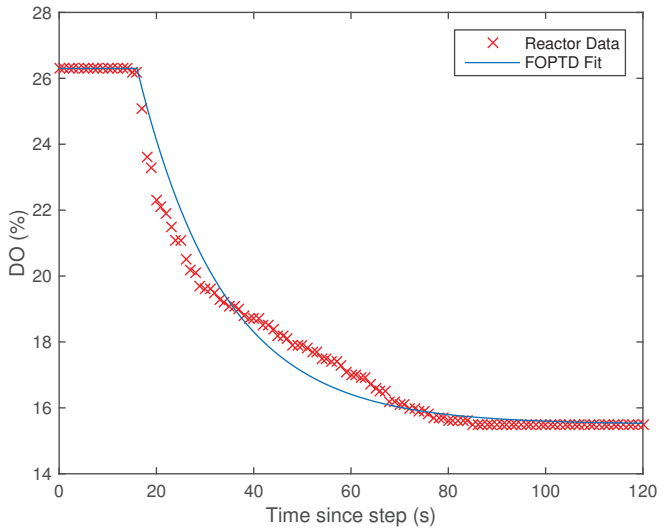

Fig. 2. Measured response of DO content (in the cell culture) to a step decrease in the stirrer speed (manipulated variable) for a particular bioreactor run. The fit of a first-order process with time delay (FOPTD) model to this measured DO data is shown in blue.

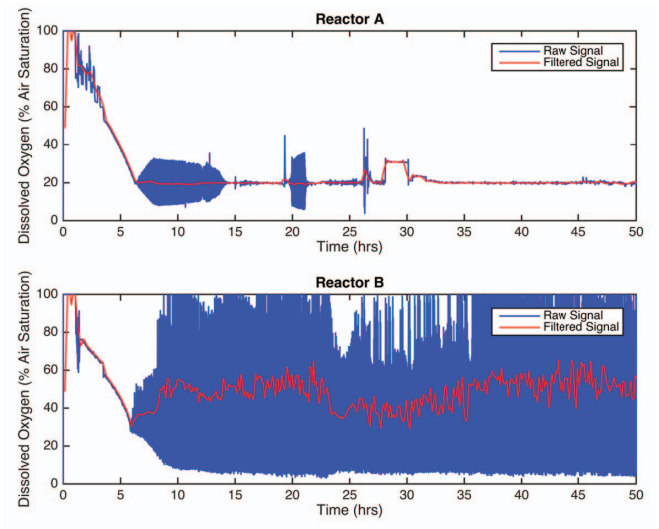

Fig. 3. (Reactor A, top) Measured DO using the proposed IMC-tuned PID controller for a fifty hour experiment. This does a good job of tracking the setpoint allowing the cell culture to thrive in the bioreactor. (Reactor B, bottom) Measured DO using a standard cascade control algorithm. The significant, persistent oscillations resulted in suboptimal performance and even killed a number of cultures due to lack of oxygen. In both plots, the same open-loop strategy is used for the first five hours of the experiment.

This control algorithm resulted in cell cultures dying during growth in a number of runs due to lack of oxygen. The IMC-tuned PID does a much better job of tracking the DO setpoint such that the cell culture is able to thrive in the bioreactor yielding desired amounts of supernatent.

III. Chromotography Separation Train

Chromatography is a separation technology that is widely used in research and industry as an analytical or preparative tool [31]. Separation can be achieved by differences in affinity, hydrophobicity, surface charge and molecular weight, and other factors [32].

A. Spatiotemporal model of mobile and stationary phases

The chromatography column for the separation of hGH from its dimer in a Capto Adhere multi-modal column is modeled based on the model equations and parameters from Sejergaard et al. [33]. They used the steric mass action (SMA) isotherm with a nonideal transport model

$$\frac{\partial C_i}{\partial \theta}\varepsilon_{tot} = -\frac{\partial C_i}{\partial z} + \frac{1}{\text{Pe}}\frac{\partial^2 C_i}{\partial z^2} - \frac{\partial q_i}{\partial \theta}\varepsilon_{pore} \tag{8}$$

where C_i is the concentration of species i in the liquid phase, θ is the dimensionless time, ε_{tot} is the total porosity, $z \in [0,1]$ is the dimensionless length, Pe is the Péclet number, q_i is the concentration of species i in the adsorbed phase, and ε_{pore} is the particle porosity.

The mass balance of salt in the liquid phase, whose concentration is denoted by C_{salt}, is described by the partial differential equation (PDE)

$$\frac{\partial C_{salt}}{\partial \theta}\varepsilon_{tot} = -\frac{\partial C_{salt}}{\partial z} + \frac{1}{\text{Pe}}\frac{\partial^2 C_{salt}}{\partial z^2} \tag{9}$$

The adsorbed phase was described by the SMA isotherm. The kinetics of the adsorption in the adsorbed phase are described by the PDE

$$\frac{\partial q_i}{\partial \theta} = k_{ads,i} C_i \bar{Q}^{\nu_i} - k_{des,i} q_i C_{salt}^{\nu_i} \tag{10}$$

where $k_{ads,i}$ is the rate constant of adsorption for component i, $\bar{Q}$ is the number of ligands available for adsorption and desorption, ν_i is the characteristic charge of component i, $k_{des,i}$ is the rate constant of desorption for component i. The rate constants of adsorption and desorption are further related by the equilibrium constant $K_i = k_{ads,i}/k_{des,i}$. Boundary conditions were the typical Danckwerks boundary conditons.

The equations were discretized using the finite volume method, to yield a system of ordinary differential equations (ODEs) i.e.,

$$\begin{aligned}\frac{dC_j}{d\theta}\varepsilon_{tot} = &-\frac{C_{j+0.5} - C_{j-0.5}}{\Delta z} - \frac{dq_j}{d\theta}\varepsilon_{pore} \\ &+ \frac{C_{j+1} - 2C_j + C_{j-1}}{\text{Pe}\Delta z^2},\ j = 0, \ldots, N\end{aligned} \tag{11}$$

where $N+1$ represents the total number of nodes, Δz represents the dimensionless length of each finite volume cell, j represents the cell index, and $j \pm 0.5$ represent the indices of the cell walls. To determine the concentration values at the wall, the Weighted Essentially Non Oscillatory (WENO) method was used due to its high accuracy and low computation requirements [34]. The WENO scheme relates the wall values to the node values by

$$\begin{aligned}C_{j+0.5} = &\frac{\alpha_{0,j}}{\alpha_{0,j} + \alpha_{1,j}}\left(\frac{1}{2}C_j + \frac{1}{2}C_{j+1}\right) \\ &+ \frac{\alpha_{1,j}}{\alpha_{0,j} + \alpha_{1,j}}\left(\frac{3}{2}C_j - \frac{1}{2}C_{j-1}\right)\end{aligned} \tag{12}$$

TABLE I
OPERATING VARIABLES FOR CHROMATOGRAPHY SIMULATION

#	Stage	Time	Protein Conc.	Salt Conc.
1	Load	t_l	$c^{\#}_{p,i}$	$c^{\#}_{s,l}$
2	Wash	t_w	0	$c_{s,w}$
3	Purge	t_p	0	$c_{s,e}$
4	Elute	t_e	0	$c_{s,e}$

$^{\#}$Variables specified by upstream units.

where $\alpha_{0,j} = \frac{2}{3(c_{j+1}-c_j+\delta)^4}$ and $\alpha_{1,j} = \frac{1}{3(c_j-c_{j-1}+\delta)^4}$ are weighting parameters and δ is a small number e.g., 10^{-10}.

The set of ODEs were solved by ODE15s [35] within MATLAB. The Jacobian matrix was computed by automatic differentiation using the Adigator [36] package to speed up the solution process.

The column was assumed to operate in a 4-stage bind and elute mode. A total of 6 operating variables were relevant to the problem, listed in Table I.

B. Open-loop Optimal Control for Chromatography

The inlet protein and salt concentration, entering the chromatography columns, are set from the bioreactor upstream to the separation unit. In a chromatography setup, sensors can only be placed at the entrance and exit of the column–not within its interior. This limits the applicability of classical feedback control. For example, if a load solution has an unusually high protein concentration, this will only be detected by premature breakthrough during the loading phase. Even if loading was terminated upon the detection of breakthrough, additional protein would be lost during the wash phase, compromising recovery.

To mitigate these effects, open loop optimal control can be used as a control strategy. In this case, measures of the load solution concentration, obtained by UV280 absorbance or Raman spectroscopy, can be used as parameters for the chromatography model presented earlier. Optimization can then be used to select the operating conditions to be used to achieve optimal process performance.

The column simulation yields four competing objectives: purity, recovery, concentration, and productivity,

$$\text{Purity} = \frac{\int_0^{t_e} c_{pdt,e} dt}{\int_0^{t_e} c_{pdt,e} dt + \int_0^{t_e} c_{imp,e} dt} \tag{13}$$

$$\text{Recovery} = \frac{\int_0^{t_e} c_{pdt,e} dt}{t_l c_{pdt,l}} \tag{14}$$

$$\text{Concentration} = \frac{\int_0^{t_e} c_{pdt,e} dt}{t_e} \tag{15}$$

$$\text{Productivity} = \frac{Q \int_0^{t_e} c_{pdt,e} dt}{t_l + t_w + t_p + t_e + t_r} \tag{16}$$

where t_i is the volume of the i^{th} stage of a bind and elute cycle, $c_{j,i}$ is the concentration of the j^{th} component in the i^{th} phase, Q is the volumetric flow rate for column operation, the subscripts l, w, p, e and r refer to the load, wash, purge, elute and regeneration phase, respectively, and the subscripts pdt and imp refer to the product of interest and the impurities, respectively.

In order to enable a tractable solution to the multi-objective optimization problem, the number of competing objectives must be reduced. This can be done by constraining some of the objective values. In this case, the purity and recovery were constrained to be at least 99% and 90%, respectively, and the concentration and productivity tradeoffs considered. The constraints were implemented via a penalty function method, with a linear penalty and a step change at the boundary.

$$x_1 = -\text{Concentration} + \omega_1 + \omega_2 \tag{17}$$

$$x_2 = -\text{Productivity} + \omega_1 + \omega_2 \tag{18}$$

where x_1 and x_2 are the two resultant objectives to be simultaneously minimized, and ω_1 and ω_2 are penalty terms

$$\omega_1 = \begin{cases} 0 \;\mid\; \text{if Purity} > \text{Purity}_{con} \\ \alpha_1 + \alpha_2(\text{Purity}_{con} - \text{Purity}) \;\mid\; \text{otherwise} \end{cases}$$

$$\omega_2 = \begin{cases} 0 \;\mid\; \text{if Recovery} > \text{Recovery}_{con} \\ \alpha_3 + \alpha_4(\text{Recovery}_{con} - \text{Recovery}) \;\mid\; \text{otherwise} \end{cases}$$

where α_1, α_2, α_3, and α_4 are positive coefficients that govern the relative weights of the step and linear penalty terms.

The multi-objective optimization problem was solved by using a non-linear multi-objective optimizer, nondominated sorting genetic algorithm II (NSGA-II) [37] for multiple concentrations of the feed solution, corresponding to a disturbance in the feed solution. The Pareto curve is shown in Fig. 4, with the decision variables plotted against concentration in Fig. 5. These curves show the tradeoff between concentration and productivity for feed solutions of various concentrations, and also show the optimal value of each decision variable for a feed and outlet concentration. These results can be put into a lookup table, and optimal operating parameters referenced during online operation, to achieve open loop optimal control of the chromatography column under feed solution disturbance.

Faster simulations, based on the method of characteristics solution of the ideal-nonlinear model, can be explored to allow the optimization to be performed online. In this case, gradient-based optimization methods may provide guaranteed and/or faster convergence properties.

IV. BUFFER MIXING UNIT: pH AND CONDUCTIVITY

During the operation of the integrated platform, buffers of various compositions, conductivities and pH are required. To avoid carrying all possible buffers directly on the platform, which is impractical due to space limitations, the buffer make-up is done by adding concentrated stock solutions to deionized water within a batch mixing tank. Closed-loop control is required to ensure that the buffers achieve the specified composition, conductivity and pH.

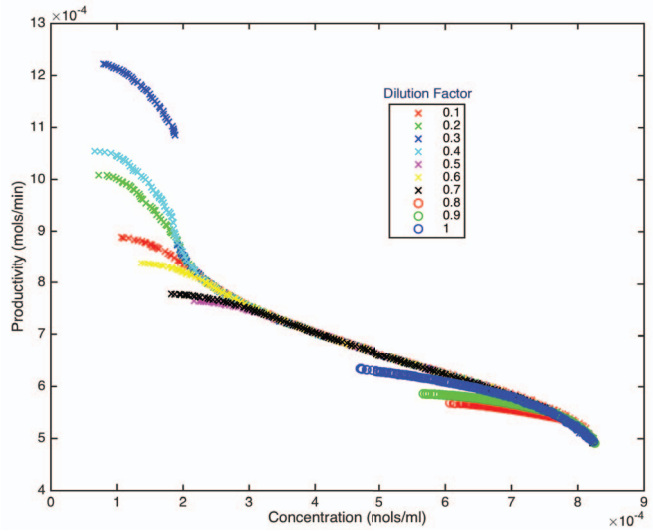

Fig. 4. Multi-objective Optimization: Concentration and Productivity

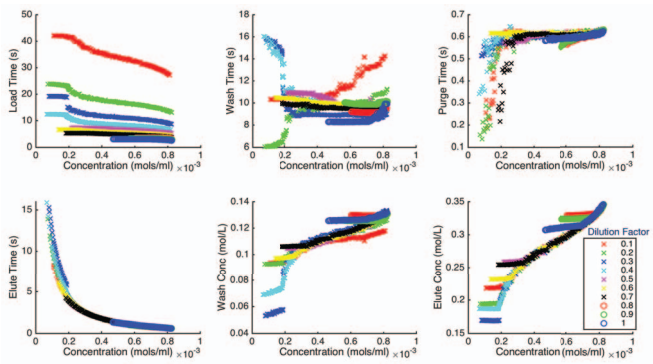

Fig. 5. Multi-objective Optimization: Decision Variables and Concentration

The stock solution addition is necessarily one way. For example, conductivity can be raised by adding salt solution, but there is no way to lower the conductivity without dilution. If a classical PID controller was used, a conservative controller tuning would be required to prevent overshoot. The small length scales and cleaning requirements of the mixing tank limit the speed at which mixing can occur. The slow mixing, coupled with the conservative controller action, will result in slow buffer make-up.

Model-based control is well suited for overcoming these limitations. In this case, a model of ionic dissociation, pH and conductivity can be used to predict the impact of addition of various buffer solutions to the mixture. These predictions can be used, together with pH and conductivity sensor data, to provide rapid control to the process.

In this section, we first discuss models for the speciation of ions in solution. These models provide a way to estimate the concentration of all dissociated species and ions within a given mixture. Next, we show how this speciation analysis can be used to compute the solution pH (using reaction invariants) and conductivity (using models of ionic conductivity). The model for conductivity, however, involves a number of parameters that should be estimated from data collected from the specific system of interest. Lastly, we illustrate the usefulness of these models in (i) conductivity design of experiment for parameter estimation and (ii) design and evaluation of pH control strategies.

A. Specitation Analysis

Accurate modeling of pH and conductivity for the electrolytes in the buffer make up system would rely on the accurate speciation of ions in solution. The speciation of aqueous electrolyte systems have been well described in literature [38], [39], [40], [41], [42].

Speciation can be written in the form of equilibrium equations:

$$K = \frac{\prod_{\mathrm{P}}(\gamma_i m_i)^{\nu_i}}{\prod_{\mathrm{R}}(\gamma_i m_i)^{\nu_i}} \tag{19}$$

where the products over P refer to the reaction products and the products over R refer to the reactants, γ_i refers to the activity coefficient, m_i refers to the molality, and ν_i refers to the stoichiometric coefficient.

The simplest model, valid at vanishingly small concentrations, is the ideal solution model, $\gamma_i = 1$. For dilute solutions of 0.001 molal or less, the Debye-Hückel limiting law is applicable:

$$\log_e \gamma = -A\,|z_+ z_-|\,\sqrt{I} \tag{20}$$

where z_+ and z_- are the ionic charges of the ionic species, I is the ionic strength, and A is,

$$A = \frac{1}{2.303}\left(\frac{e}{\sqrt{DkT}}\right)^3 \sqrt{\frac{2\pi d_0 N_A}{1000}} \tag{21}$$

where e is the electronic charge, D is the dielectric constant, k is the Boltzmann's constant, T is the absolute temperature, d_0 is the solvent density, and N_A is Avogadro's constant.

Further refinements have been made to this model. Debye and Hückel proposed a correction term to the limiting law:

$$\log_e \gamma = -\frac{A\,|z_+ z_-|\,\sqrt{I}}{1 + \beta a \sqrt{I}} \tag{22}$$

where a is a parameter describing the distance of closest approach, and β is

$$\beta = \sqrt{\frac{8\pi e^2 N_A d_0}{1000 DkT}} \tag{23}$$

Other models, including further refinements to the Debye Hückel model, are proposed in literature. These include Guggenheim's equation, Bromley's Equation, Pitzer's method, and Chen's Method [38].

In particular, Chen's method decomposes excess Gibbs free energy into the sum of a long range and short range interactions. The short range interactions are modeled by a nonrandom, two liquid type model, while the long range interactions are similar to that of the Pitzer model. The equations for Chen's model are too bulky to be represented here, and are well described by Zemaitis et al. [38] and in Chen et al.'s original paper [43].

B. Reaction Invariants for pH Modeling

Gustafsson and Waller [44] have applied a reaction invariant approach to the dynamic modeling and control of pH in solution. This approach identifies quantities that are invariant under the dissociation and speciation reactions associated with the addition of acids and bases (under the assumption of fast reactions that quickly go to equilibrium). This allows for the nonlinear pH dynamics in a continuously-stirred tank or semi-batch reactor to be converted to linear dynamics in the reaction invariants with a static nonlinearity. The static nonlinearity is written in terms of the reaction invariants using the equilibrium relationships between the species in solution. A general procedure for determining the reaction invariants for a general acid-base system is discussed in [44], and invariants for an example system are constructed below.

C. Conductivity Modeling

From the speciation models discussed earlier, volumes of solutions added to a mixture can be converted into concentrations of all dissociated species. Conductivity models would then use the concentrations of all ions in solution to determine a mixture conductivity. Robinson and Stokes present a fundamental treatment of the modeling of ionic conductivity in solution [45], and some of the key models are described here. The simplest of which is the limiting conductivity, which describes the conductivity at very low concentrations:

$$\lambda = \lambda^0 - \sqrt{c} \tag{24}$$

where λ is the equivalent conductivity, λ_0 is the limiting equivalent conductivity, A is a coefficient, and c is the ionic concentration

There are more sophisticated theory available in literature that describe conductivity at high solution concentrations, such as the Fuoss-Onsager theory. Recently the mean spherical approximation (MSA) has emerged as an accurate description of conductivity of an ion pair at moderate ionic concentration [46]. The theory describes the equivalent conductivity λ_i as

$$\lambda_i = \lambda_i^0 \left(1 + \frac{v_i^{el}}{v_i^0}\right)\left(1 + \frac{\delta X}{X}\right) \tag{25}$$

where λ_i^0 is the limiting equivalent conductivity of ion i at infinite dilution, $\frac{v_i^{el}}{v_i^0}$ represents the electrophoretic correction, and $\frac{\delta X}{X}$ represents the relaxation contribution. Detailed forms of these terms are described by Bernard et al. [46]

Anderko and Lencka extended the MSA theory to a multicomponent systems over a wider concentration range [47]. They proposed several mixing rules to compute multicomponent conductivity from considering the conductivity of all cation and anion pairs. This allows the single component conductivities obtained from the MSA theory to be combined to calculate multicomponent conductivity.

D. Example: pH Control

The InSCyT platform will have a feature to that enables mixing buffers of various compositions, conductivities, and pH. This unit operation is important for ensuring the chromatography separation train works properly as the buffer properties are key in its efficacy, and change from product to product. We can use a model of the process to design and evaluate various control strategies. In this section, we use the method of reaction invariants to develop a preliminary model of pH in the buffer mixing unit.

Here, we look at an example system composed of water, phosphate, base (NaOH), and arginine. Three possible reaction invariants for this system are

$$w_p = [\mathrm{H_3PO_4}] + [\mathrm{H_2PO_4^-}] + [\mathrm{HPO_4^{2-}}] + [\mathrm{PO_4^{3-}}] \tag{26}$$

$$w_a = [\mathrm{H_3Arg^{2+}}] + [\mathrm{H_2Arg^{+}}] + [\mathrm{HArg}] + [\mathrm{Arg^{-}}] \tag{27}$$

$$w_b = [\mathrm{Na^{+}}] \tag{28}$$

We define the vector of invariant concentrations and moles to be $\mathbf{w} := [w_p,\ w_a,\ w_b]^\top$ and $\mathbf{N} := \mathbf{w}V$, respectively, where V is the total volume of solution. The system is assumed to have three inlet flows (pure phosphoric acid, arginine, and NaOH) of fixed concentration that can be manipulated to control the pH. The time evolution of these quantities can then be computed as

$$V(k+1) = V(k) + [1,\ 1,\ 1]\mathbf{u}(k) \tag{29}$$

$$\mathbf{N}(k+1) = \mathbf{N}(k) + \mathbf{C}_{in}\mathbf{u}(k) \tag{30}$$

$$\mathbf{w}(k) = \mathbf{N}(k)/V(k) \tag{31}$$

where $k \in \{0, 1, \ldots\}$ denotes the discrete-time index, $C_{p,in}$, $C_{a,in}$, $C_{b,in}$ denote the fixed concentration of the inlet phosphoric acid, arginine, and NaOH inlet streams, respectively, $\mathbf{C}_{in} = diag(C_{p,in}, C_{a,in}, C_{b,in})$, and $\mathbf{u}(k) := [V_{p,in}(k),\ V_{a,in}(k),\ V_{b,in}(k)]^\top$ denotes the control input vector, with elements representing the volume of each inlet stream added to the system at time k.

The reaction invariants $\mathbf{w}$ can be related to the pH of the solution via the equilibrium conditions for the reactions in the system (see e.g., [44] for details). This results in a static nonlinear equation $f(\mathbf{w}, \mathrm{pH}) = 0$ for pH as a function of the reaction invariants $\mathbf{w}$. The time evolution of pH can be determined by solving this equation at each time instant k after computing the updated invariant values from (29)–(31).

We assume the mixing dynamics can be represented as a first-order filter with mixing time constant τ_m. Discretizing the first-order ODEs yields a difference equation

$$\mathbf{N}_{mix}(k) = \lambda \mathbf{N}_{mix}(k) + (1-\lambda)\mathbf{N}(k) \tag{32}$$

where $\lambda = \tau_m/(\tau_m + \Delta t)$, Δt is the sampling time of the system, and $\mathbf{N}_{mix}(k)$ is the state measured by the sensor, which will be delayed due to the time it takes for the system to become well-mixed.

To show how this model can be used to gain understanding that is useful in the development of control strategies for the buffer mixing unit, let us look at an example buffer make-up via the titration of phosphoric acid with base. A titration curve for this system is shown in Fig. 6. Notice that system exhibits sharp changes in pH about halfway between the pKa values of the phosphate ions (these are known equivalence

points). For pH setpoints in the vicinity of these equivalence points, one would like to design a controller that is not overly sensitive to measurement errors or disturbances. However, this controller would likely perform poorly in other pH regions that exhibit a much slower system response.

This analysis indicates that this application is well-suited to a gain scheduled PID controller, where a different PID controller is applied to the system for different regions of the measured state of the system. To design this type of controller, we can use the titration curve to look for regions where the system behaves approximately linearly. For this example, the locally similar regions of pH are approximately $[1, 3]$, $[3, 6]$, $[6, 8]$, $[8, 11]$, and $[11, 14]$. Next, we design local PID controllers that perform well in each of these regions. A particular controller is applied at a given time instant whenever the measured pH falls within that region. The system response with the the gain scheduled PID controller, for a pH setpoint of 4.5, is shown in Fig. 7. In this simulation, we assume a mixing time τ_m of one minute and a sampling time Δt of 0.1 min. If the pH goes above the setpoint, some phosphoric acid is added based on a proportional control law.

As seen in Fig. 7, the pH response exhibits some slight overshoot due to the difference between the true and measured state. Compare this, however, to exclusively applying the controller tuned for the low pH region (Fig. 8). This controller is too aggressive near the setpoint, which causes significant overshoot and oscillations to occur. Therefore, it takes slightly longer to reach the setpoint. More importantly, it requires an extra liter of phosphoric acid to be added to the system compared to the gain scheduled controller (Fig. 7). This is important because (i) we only want to mix (as close as possible) to the amount of buffer needed during a run to avoid waste and (ii) the integrated platform is meant for deployment in the field such that we want to design control strategies that minimize volume requirements.

E. Example: Conductivity Design of Experiments

Design of experiments techniques can be used to determine experimental conditions to obtain the most accurate estimates of model parameters. Each conductivity experiment involves adding specified volumes of several solutions to a mixer and measuring the conductivity. We explore various solutions form a five-component system, consisting of phosphate buffer, base, arginine, salt, and water. A Bayesian d-optimal experimental design procedure [48] was used to determine the volumes of the solutions to be added to the mixer for each experiment.

The experimental design procedure begins by computing sensitivities

$$\mathbf{F}_j = \left.\frac{\partial \tilde{y}_j(\mathbf{u}_j, \mathbf{Q})}{\partial \mathbf{Q}}\right|_{\mathbf{Q}^*} \tag{33}$$

where $\mathbf{F}_j$ is the sensitivity vector for the j^{th} experiment, $\tilde{y}_j$ is the model predicted conductivity, $\mathbf{u}_j$ is the vector of j^{th} experimental conditions, $\mathbf{Q}$ is the vector of model parameters, and $\mathbf{Q}^*$ is the best estimate of the parameters obtained from literature.

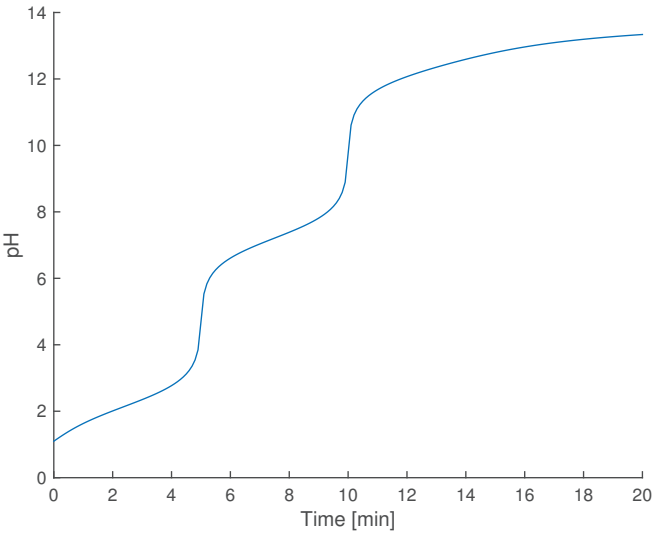

Fig. 6. Titration curve for the addition of base (at 0.2 mol/min) to a 1 M solution of phosphoric acid. Note that the pKa values for phosphate are 2.16, 7.12, and 12.32.

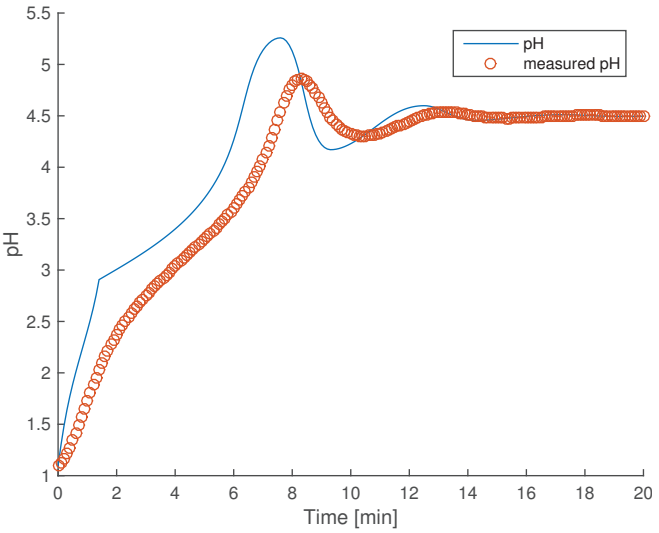

Fig. 7. Time evolution of pH in buffer mixing tank using a gain scheduled controller for a setpoint of 4.5 to determine inlet flows to add to the tank.

The maximum *a posteriori* (MAP) estimate on the covariance of the model parameters is

$$\mathbf{P}_{MAP} = \left[\sum_{j=1}^{N_e} \mathbf{F}_j^T \mathbf{V}^{-1} \mathbf{F}_j + \mathbf{P}_{prior}^{-1}\right]^{-1} \tag{34}$$

where N_e is the total number of experiments, $\mathbf{V}$ is the covariance matrix of the measurement, and $\mathbf{P}_{prior}$ is the covariance matrix of the prior distribution.

The optimal set of experiments minimizes the determinant of $\mathbf{P}_{MAP}$. Alternatively, the problem can be formulated in a

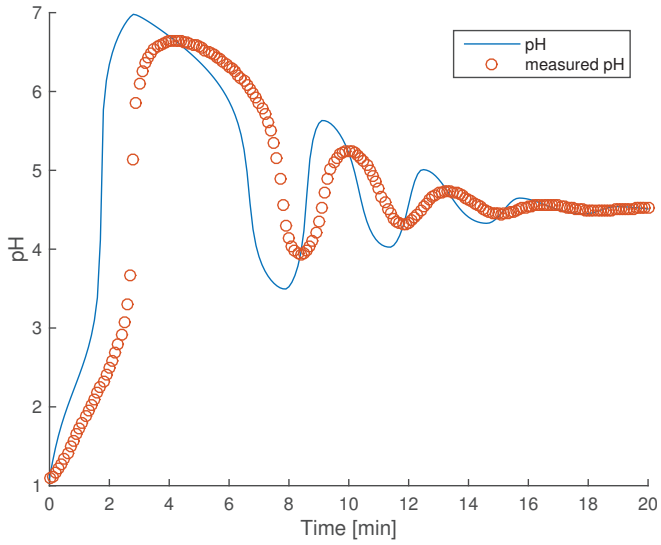

Fig. 8. Time evolution of pH in buffer mixing tank for a setpoint of 4.5, exclusively using the PID controller tuned for the low pH region.

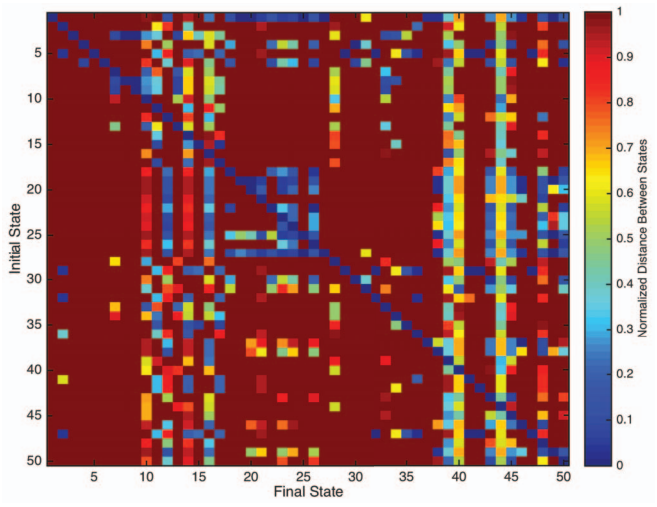

Fig. 9. Distance between any two experiments. Blue points represent easy to reach experiments while red points represent hard to reach experiments.

way that avoids taking matrix inverses:

$$\max_{\mathbf{u}} \left| \sum_{j=1}^{N_e} \mathbf{F}_j^T \mathbf{V}^{-1} \mathbf{F}_j + \mathbf{P}_{prior}^{-1} \right| \tag{35}$$

This optimization involves $n_u \times n_e$ decision variables, where n_u is the number of conditions for each experiment. To reduce the scale of the optimization, a local optimum was sought by using a greedy algorithm. In such an algorithm, the first experiment is designed alone, creating a n_u size optimization problem. The next experiment is designed with the first experiment fixed, and so on, resulting in n_e optimization problems of size n_u. This was used to find a list of experiments to be conducted.

Each experiment involves adding specific volumes of solutions to a tank, mixing the tank contents, measuring the conductivity, and cleaning the tank. Optimal sequencing of the experiments can reduce the time and feedstock volume required by allowing multiple mixture conductivities to be measured before the tank is cleaned.

It is desired to sequence the experiments to allow for the most data to be collected before emptying the tank and cleaning. The number of solutions that can be measured within each cycle is constrained by the minimum and maximum measurable volumes within the tank. The additional volume necessary to move from one composition to another is

$$\frac{V_{final}}{V_{initial}} = \max_i \frac{x_{i,initial}}{x_{i,final}} \tag{36}$$

where V_{final} and $V_{initial}$ are the initial and final volumes, respectively, and $x_{i,initial}$ and $x_{i,final}$ are the mole fractions of component i for the initial and final mixtures, respectively.

The volume ratio calculated is a multiplicative factor and was linearized by taking the natural logarithm,

$$d = \ln\left(\frac{V_{final}}{V_{initial}}\right) \tag{37}$$

where d represents the additive distance between two points.

The distance between two successive points can thus be represented as a digraph. The matrix representation of the digraph is shown in Fig. 9, with blue points being easier to reach and the red points being harder to reach.

A greedy algorithm was used to obtain approximate solutions to the optimal sequencing problem. In this algorithm, a run is defined as a set of experiments, performed by sequentially adding solutions to a tank. A run begins at a minimum volume, corresponding to a minimum volume capable of being measured by the sensors, and a maximum volume, corresponding to the size of the mixing vessel. The algorithm begins at the best unvisited experimental point, and adds to the run by choosing the point with the shortest distance. This continues until the tank is full, and the algorithm beings again with the next best unvisited experimental point.

The algorithm was run to sequence the list of experiments generated via the earlier algorithm, and the results are tabulated in Table II. This algorithm allows for the optimal design and sequencing of conductivity and pH experiments for any future set of ions proposed for the InSCyT platform.

V. Conclusions and Future Work

This paper focuses on the various systems engineering challenges for biomanufacturing, as exemplified by the Integrated and Scalable Cyto-Technology (InSCyT) bio-manufacturing platform under the Defense Advanced Research Projects Agency (DARPA) Biologically-derived Medicines on Demand (Bio-MOD) program. Models and control strategies for the major unit operations within the InSCyT platform are discussed, which include a perfusion-operated bioreactor, a packed bed chromatography separation

TABLE II

APPROXIMATE SOLUTION TO THE OPTIMAL SEQUENCING PROBLEM FOR THE MIXING MODULE

Expt	Volume of Conc. Solution (L)				Expt	Volume of Conc. Solution (L)				Volume of Solution Added (L)
	Buffer	Base	Arg	Salt		Buffer	Base	Arg	Salt	
Run 1					Run 3					
1	0.00	0.50	0.00	0.00	3	0.00	0.00	0.50	0.00	0.00
13	0.00	0.50	0.00	0.00	24	0.00	0.00	0.49	0.00	0.05
25	0.14	0.32	0.00	0.00	47	0.07	0.02	0.30	0.08	0.10
5	0.16	0.26	0.02	0.05	49	0.02	0.01	0.14	0.31	0.15
45	0.18	0.11	0.05	0.07	Run 4					0.20
30	0.22	0.11	0.05	0.07	4	0.05	0.00	0.45	0.00	0.25
31	0.16	0.13	0.10	0.07	41	0.05	0.00	0.45	0.00	0.30
44	0.09	0.06	0.11	0.06	16	0.05	0.00	0.45	0.00	0.35
Run 2					22	0.02	0.00	0.43	0.00	0.40
2	0.00	0.00	0.00	0.50	40	0.01	0.00	0.22	0.15	0.45
33	0.00	0.00	0.00	0.48	32	0.01	0.06	0.11	0.30	0.50
7	0.00	0.01	0.00	0.49	Run 5					
8	0.03	0.01	0.01	0.45	6	0.15	0.35	0.00	0.00	
36	0.06	0.01	0.01	0.41	28	0.14	0.31	0.00	0.00	
9	0.05	0.01	0.02	0.40	43	0.08	0.28	0.02	0.10	
39	0.04	0.03	0.02	0.35	Run 6					
46	0.03	0.04	0.05	0.33	10	0.01	0.02	0.09	0.37	
11	0.02	0.02	0.29	0.16	37	0.01	0.06	0.01	0.30	

train, and a conductivity- and pH-controlled buffer mixing unit. Experimental and simulation-based case studies are presented for each unit operation.

As future work, the models and control strategies will continue to be improved as relevant data is collected or theory is developed. Design of experiment procedures, as discussed in the context of conductivity models, will be applied to the bioreactor and chromatography models to reduce residual uncertainty in parameter estimates. Models and control strategies will also be explored for the membrane ultrafiltration module, which is used to effect a buffer exchange at the end of the process.

A primary future goal of this work will be to link these models, of the individual process units, into a integrated plant model that represents the entire InSCyT platform. Both single and multi-objective optimization algorithms will be used to perform optimal design of the integrated plant, and the tradeoffs between the operation of the individual process units understood. Plant-wide control strategies will also be developed and implemented on models and on a prototype of the InSCyT platform. A challenge will be developing simulations that are fast enough for real time control and iterative optimization. To address these issues, tools, such as fast model predictive control [49], will be applied to integrated plant when possible.

REFERENCES

[1] G. Walsh, "Biopharmaceutical benchmarks 2014," *Nature Biotechnology*, vol. 32, no. 10, pp. 992–1000, 2014.

[2] S. R. Aggarwal, "What's fueling the biotech engine - 2012 to 2013," *Nature*, vol. 201, p. 4, 2014.

[3] V. Warikoo, R. Godawat, K. Brower, S. Jain, D. Cummings, E. Simons, T. Johnson, J. Walther, M. Yu, B. Wright, *et al.*, "Integrated continuous production of recombinant therapeutic proteins," *Biotechnology and Bioengineering*, vol. 109, no. 12, pp. 3018–3029, 2012.

[4] A. S. Rathore and H. Winkle, "Quality by design for biopharmaceuticals," *Nature Biotechnology*, vol. 27, no. 1, p. 26, 2009.

[5] A. Cinar, S. J. Parulekar, C. Undey, and G. Birol, *Batch Fermentation: Modeling, Monitoring, and Control*. CRC Press, 2003.

[6] S. H. Gordon, S. C. Whatel, B. C. Wheeler, and C. James, "Multivariate FTIR analysis of substrates for protein, polysaccharide, lipid and microbe content: potential for solid-state fermentations," *Biotechnology Advances*, vol. 11, no. 3, pp. 665–675, 1993.

[7] J. C. Gunther, J. S. Conner, and D. E. Seborg, "Fault detection and diagnosis in an industrial fed-batch cell culture process," *Biotechnology Progress*, vol. 23, no. 4, pp. 851–857, 2007.

[8] H. Zhang and B. Lennox, "Integrated condition monitoring and control of fed-batch fermentation processes," *Journal of Process Control*, vol. 14, no. 1, pp. 41–50, 2004.

[9] S. M. Rosa, M. A. Soria, C. G. Vélez, and M. A. Galvagno, "Improvement of a two-stage fermentation process for docosahexaenoic acid production by Aurantiochytrium limacinum SR21 applying statistical experimental designs and data analysis," *Bioresource Technology*, vol. 101, no. 7, pp. 2367–2374, 2010.

[10] G. M. Troup and C. Georgakis, "Process systems engineering tools in the pharmaceutical industry," *Computers & Chemical Engineering*, vol. 51, pp. 157–171, 2013.

[11] S. Chhatre, "Modelling approaches for bio-manufacturing operations," in *Measurement, Monitoring, Modelling and Control of Bioprocesses*, pp. 85–107, Springer, 2013.

[12] I. J. Dunn, E. Heinzle, J. Ingham, and J. E. Prenosil, *Biological Reaction Engineering*. John Wiley & Sons, 2003.

[13] G. Guiochon and B. Lin, *Modeling for Preparative Chromatography*. Academic Press, 2003.

[14] G. Guiochon, A. Felinger, and D. G. Shirazi, *Fundamentals of Preparative and Nonlinear Chromatography*. Academic Press, 2006.

[15] C. E. Hogwood, A. S. Tait, N. Koloteva-Levine, D. G. Bracewell, and C. M. Smales, "The dynamics of the CHO host cell protein profile during clarification and protein A capture in a platform antibody purification process," *Biotechnology and Bioengineering*, vol. 110, no. 1, pp. 240–251, 2013.

[16] A. S. Tait, R. D. Tarrant, M. L. Velez-Suberbie, D. I. Spencer, and D. G. Bracewell, "Differential response in downstream processing of CHO cells grown under mild hypothermic conditions," *Biotechnology Progress*, vol. 29, no. 3, pp. 688–696, 2013.

[17] M. Groep, M. Gregory, L. Kershenbaum, and I. Bogle, "Performance modeling and simulation of biochemical process sequences with interacting unit operations," *Biotechnology and Bioengineering*, vol. 67, no. 3, pp. 300–311, 2000.

[18] E. Varga, N. Titchener-Hooker, and P. Dunnill, "Prediction of the pilot-scale recovery of a recombinant yeast enzyme using integrated models," *Biotechnology and Bioengineering*, vol. 74, no. 2, pp. 96–107, 2001.

[19] S. Chhatre, C. Jones, R. Francis, K. O'Donovan, N. Titchener-Hooker, A. Newcombe, and E. Keshavarz-Moore, "The integrated simulation and assessment of the impacts of process change in biotherapeutic antibody production," *Biotechnology Progress*, vol. 22, no. 6, pp. 1612–1620, 2006.

[20] J. M. Pinto, J. M. Montagna, A. R. Vecchietti, O. A. Iribarren, and J. A. Asenjo, "Process performance models in the optimization of multiproduct protein production plants," *Biotechnology and Bioengineering*, vol. 74, no. 6, pp. 451–465, 2001.

[21] C. Kontoravdi, N. J. Samsatli, and N. Shah, "Development and design of bio-pharmaceutical processes," *Current Opinion in Chemical Engineering*, vol. 2, no. 4, pp. 435–441, 2013.

[22] S.-O. Enfors, *Fermentation Process Engineering*. Royal Institute of Technology, Stockholm, 2011.

[23] M. Jahic, J. Rotticci-Mulder, M. Martinelle, K. Hult, and S.-O. Enfors, "Modeling of growth and energy metabolism of pichia pastoris producing a fusion protein," *Bioprocess and Biosystems Engineering*, vol. 24, no. 6, pp. 385–393, 2002.

[24] T. A. Bibila and M. C. Flickinger, "Use of a structured kinetic model of antibody synthesis and secretion for optimization of antibody production systems: II. Transient analysis," *Biotechnology and Bioengineering*, vol. 39, no. 3, pp. 262–272, 1992.

[25] S. Craven, N. Shirsat, J. Whelan, and B. Glennon, "Process model comparison and transferability across bioreactor scales and modes of operation for a mammalian cell bioprocess," *Biotechnology Progress*, vol. 29, no. 1, pp. 186–196, 2013.

[26] R. P. Nolan and K. Lee, "Dynamic model of CHO cell metabolism," *Metabolic Engineering*, vol. 13, no. 1, pp. 108–124, 2011.

[27] S. Selvarasu, Y. S. Ho, W. P. Chong, N. S. Wong, F. N. Yusufi, Y. Y. Lee, M. G. Yap, and D.-Y. Lee, "Combined in silico modeling and metabolomics analysis to characterize fed-batch CHO cell culture,"

Biotechnology and Bioengineering, vol. 109, no. 6, pp. 1415–1429, 2012.

[28] R. Lencastre Fernandes, M. Nierychlo, L. Lundin, A. E. Pedersen, P. Puentes Tellez, A. Dutta, M. Carlquist, A. Bolic, D. Schäpper, A. C. Brunetti, *et al.*, "Experimental methods and modeling techniques for description of cell population heterogeneity," *Biotechnology Advances*, vol. 29, no. 6, pp. 575–599, 2011.

[29] I. Jimenez del Val, J. M. Nagy, and C. Kontoravdi, "A dynamic mathematical model for monoclonal antibody N-linked glycosylation and nucleotide sugar donor transport within a maturing Golgi apparatus," *Biotechnology Progress*, vol. 27, no. 6, pp. 1730–1743, 2011.

[30] D. Seborg, T. F. Edgar, and D. Mellichamp, *Process dynamics & control*. John Wiley & Sons, 2006.

[31] L. R. Snyder, J. J. Kirkland, and J. W. Dolan, *Introduction to Modern Liquid Chromatography*. John Wiley & Sons, 2011.

[32] E. Heftman, *Chromatography - Fundamentals and Applications of Chromatography and Related Differential Migration Methods*. Elsevier, Amsterdam, 2004.

[33] L. Sejergaard, H. S. Karkov, J. K. Krarup, A. B. B. Hagel, and S. M. Cramer, "Model-based process development for the purification of a modified human growth hormone using multimodal chromatography," *Biotechnology Progress*, 2014.

[34] E. von Lieres and J. Andersson, "A fast and accurate solver for the general rate model of column liquid chromatography," *Computers & Chemical Engineering*, vol. 34, no. 8, pp. 1180–1191, 2010.

[35] L. F. Shampine and M. W. Reichelt, "The MATLAB ODE suite," *SIAM Journal on Scientific Computing*, vol. 18, no. 1, pp. 1–22, 1997.

[36] M. A. Patterson, M. Weinstein, and A. V. Rao, "An efficient overloaded method for computing derivatives of mathematical functions in MATLAB," *ACM Transactions on Mathematical Software*, vol. 39, no. 3, p. 17, 2013.

[37] K. Deb, A. Pratap, S. Agarwal, and T. Meyarivan, "A fast and elitist multiobjective genetic algorithm: NSGA-II," *IEEE Transactions on Evolutionary Computation*, vol. 6, no. 2, pp. 182–197, 2002.

[38] J. F. Zemaitis Jr, D. M. Clark, M. Rafal, and N. C. Scrivner, *Handbook of Aqueous Electrolyte Thermodynamics: Theory & Application*. John Wiley & Sons, 2010.

[39] K. S. Pitzer, *Activity Coefficients in Electrolyte Solutions*. CRC press, 1991.

[40] S. I. Sandler, *Models for Thermodynamic and Phase Equilibria Calculations*. Dekker, 1994.

[41] J. R. Loehe and M. D. Donohue, "Recent advances in modeling thermodynamic properties of aqueous strong electrolyte systems," *AIChE Journal*, vol. 43, no. 1, pp. 180–195, 1997.

[42] A. Anderko, P. Wang, and M. Rafal, "Electrolyte solutions: from thermodynamic and transport property models to the simulation of industrial processes," *Fluid Phase Equilibria*, vol. 194, pp. 123–142, 2002.

[43] C.-C. Chen, H. I. Britt, J. F. Boston, and L. B. Evans, "Extension and application of the pitzer equation for vapor-liquid equilibrium of aqueous electrolyte systems with molecular solutes," *AIChE Journal*, vol. 25, no. 5, pp. 820–831, 1979.

[44] T. K. Gustafsson and K. V. Waller, "Dynamic modeling and reaction invariant control of pH," *Chemical Engineering Science*, vol. 38, no. 3, pp. 389–398, 1983.

[45] R. A. Robinson and R. H. Stokes, *Electrolyte Solutions*. Courier Dover Publications, 2002.

[46] O. Bernard, W. Kunz, P. Turq, and L. Blum, "Conductance in electrolyte solutions using the mean spherical approximation," *The Journal of Physical Chemistry*, vol. 96, no. 9, pp. 3833–3840, 1992.

[47] A. Anderko and M. M. Lencka, "Computation of electrical conductivity of multicomponent aqueous systems in wide concentration and temperature ranges," *Industrial & Engineering Chemistry Research*, vol. 36, no. 5, pp. 1932–1943, 1997.

[48] J. V. Beck and K. J. Arnold, *Parameter Estimation in Engineering and Science*. John Wiley & Sons, 1977.

[49] J. A. Paulson, A. Mesbah, S. Streif, R. Findeisen, and R. D. Braatz, "Fast stochastic model predictive control of high-dimensional systems," in *Proceedings of the IEEE Conference on Decision and Control, page In Press. Los Angeles*, 2014.