Appendix A   Environment Details

A.1   ReactorEnv

Recall that the reaction is of form A → B. Performing a component balance on reactant A, we obtain the following ordinary differential equation

$$\frac{dc_A}{dt} = \frac{q_{in}}{\pi r^2 h}(c_{Af} - c_A) - k_0 \exp(-\frac{E}{RT})c_A, \tag{1}$$

where c_A is the concentration of reactant A in $kmol/m^3$, t is the time in min, q_{in} is the volumetric flowrate of the inlet stream in m^3/min, r is the radius of the reactor in m, h is the level of reaction mixture in the reactor in m, c_{Af} is the concentration of reactant A in the feed stream in $kmol/m^3$, k_0 is the pre-exponential factor in min^{-1}, E/R is the ratio of reaction activation energy to the universal gas constant in K and T is the reaction mixture temperature in K.

Similarly, an energy balance can be conducted to obtain the following energy balance equation

$$\frac{dT}{dt} = \frac{q_{in}}{\pi r^2 h}(T_f - T) + \frac{-\Delta H}{\rho c_p} k_0 \exp(-\frac{E}{RT})c_A + \frac{2U}{r\rho c_p}(T_c - T), \tag{2}$$

where T_f is the temperature of the feed stream in K, ΔH is the heat of reaction in $kJ/kmol$, ρ is the density of the reaction mixture in kg/m^3, c_p is the specific heat capacity of the reaction mixture in $kJ/kg \cdot K$, U is the heat transfer coefficient in $kJ/min \cdot m^2 \cdot K$ and T_c is the coolant temperature.

Finally, deriving an overall material balance around the reactor leads to the following equation

$$\frac{dh}{dt} = \frac{q_{in} - q_{out}}{\pi r^2}, \tag{3}$$

where q_{out} is the volumetric flow rate of the contents out of the reactor in m^3/min.

A summary of the parameter values used in this project is presented in Table 3.

Table 3: Table of parameter values

Parameter	Unit	Value
q_{in}	m^3/min	0.1
r	m	0.219
c_{Af}	$kmol/m^3$	1.0
T_f	K	76.85
E/R	K	8750.0
k_0	min^{-1}	7.2×10^{10}
$-\Delta H$	J/mol	5.0×10^4
U	$kJ/min \cdot m^2 \cdot K$	5.0×10^4
c_p	$kJ/kg \cdot K$	0.239
ρ	kg/m^3	1000.0

In the CSTR process model described above, c_A, T and h are the state variables. The controlled variables are c_A and h while the manipulated variables are q_{out} and T_c.

A.2   AtropineEnv

A description of the streams in Figure 4 is summarized in Table 4.

The mixing in the mixers is assumed to occur instantaneously which implies zero dynamics. Thus, the mixer is modeled by the following set of algebraic equations

$$\dot{m}_{\text{out},i} = \sum_{k=1}^{n_s} \dot{m}_{\text{in},i,k} \tag{4}$$

where m denotes the mass flow rate in xx and n_s is the number of streams. In Equation (4), the subscripts i, k, out, in refer to species, stream number, reactor outlet and reactor inlet respectively.

15
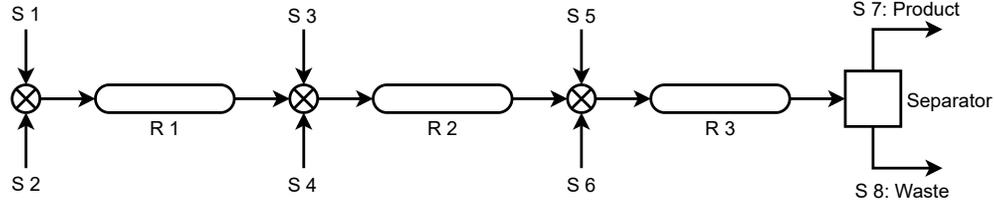
Figure 4: Process flow diagram of the continuous manufacturing process

Table 4: Description of streams in Figure 4

Stream	Description
S 1	Tropine in dimethylformamide (2 M)
S 2	Phenylacetylchloride (pure)
S 3	Formaldehyde (37 wt%)
S 4	Sodium hydroxide (4 M)
S 5	Buffer solution
S 6	Organic solvent (Toluene)
S 7	Product
S 8	Waste

Each reactor is described by the following partial differential equations obtained from their mass balance

$$\frac{\partial c_{i,z}}{\partial t} = -Q_{\text{tot}} \frac{\partial c}{\partial V}\bigg|_{i,z} + r_{i,z} \tag{5}$$

Equation 5 can be converted to ordinary differential equations using the method of lines (MOL) [54] to obtain

$$\frac{dc_{i,j}}{dt} = -Q_{\text{tot}} \frac{c_{i,j} - c_{i,j-1}}{\Delta V} + r_{i,j} \tag{6}$$

. In Equations 5 and 6, Q_{tot} is the total volumetric flow rate inside the reactor in mL/min, r is the rate of reaction and ΔV is the volume of a segment of the reactor. The subscripts i is as previously defined and j is the volume of a segment of the reactor in mL. The temperature dynamics in each reactor are assumed to be fast and therefore the energy balances are not required. The liquid-liquid separator is described by both ordinary differential equations and algebraic equations which results in a differential-algebraic equation (DAE) system of index 1. More details of the process model can be found in [23] and the references therein. A summary of the key process parameters is shown in Table 5.

Table 5: Key process parameters

Parameter	Description	Value [units]
V_1	Volume of Reactor 1	2 [mL]
V_2	Volume of Reactor 2	9.5 [mL]
V_3	Volume of reactor 3	9.5 [mL]
V_4	Volume of Liquid-liquid separator	110 [mL]
T_1	Temperature of Reactor 1	373.15 [K]
T_2	Temperature of Reactor 2	373.15 [K]
T_3	Temperature of Reactor 3	323.15 [K]
q_5	Volumetric flow rate of $S5$	0.2 [mL/min]
q_6	Volumetric flow rate of $S6$	0.5 [mL/min]
$\log(D_9)$	Separation coefficient of atropine	-2 [-]

In the continuous-flow manufacturing process, the volumetric flow rates of streams S1–S4 are manipulated to control the production of atropine while the volumetric flow rates of streams S5 and S6 are kept constant.

16

A.2.1 Process control with MPC

The entire system of DAEs described in the previous section can be written in the form

$$\dot{x}(t) = f(x(t), z(t), u(t))$$
$$0 = g(x(t), z(t), u(t))$$
$$y(t) = h(x(t), z(t), u(t)),$$

where $\dot{x}(t) \in \mathbb{R}^{1694}$ is the velocity of the state vector $x(t) \in \mathbb{R}^{1694}$ at time $t \in \mathbb{R}_+$, $u(t) \in \mathbb{R}^4$ is the vector of inputs, $z(t) \in \mathbb{R}^{30}$ is the vector of algebraic states and $y(t) \in \mathbb{R}$ is the output.

The control objective is to maximize atropine production while minimizing the waste produced. This metric is known as the environmental factor (E-factor) and is defined as

$$\text{E-factor} = \frac{\text{Mass of waste produced (excluding water)}}{\text{Mass of product obtained}}.$$

The above DAE system of equations, when used in a model-based controller such as MPC will result in a large-scale nonlinear and possibly non-convex optimization problem which is in general difficult to solve. Thus, to reduce the complexity of the controller, a simple linear model was identified from data and used to make predictions in the controller. A linear discrete-time subspace model relating the inputs to the output (E-factor) was obtained and used in the controller. Since the states of the linear subspace model have no physical meaning, a steady-state Kalman filter was designed to estimate the initial state from the inputs and outputs. A schematic diagram of the control system is shown in Figure 5. In the 5, $r(t)$ is the reference signal to be tracked (usually obtained from a higher decision making body such as Real-Time Optimizer (RTO)) and $\hat{x}(t)$ is the initial state estimate for the linear model in the controller. It is worth mentioning that the linear model in the controller may have to be re-identified if the new reference is far from the current reference point.
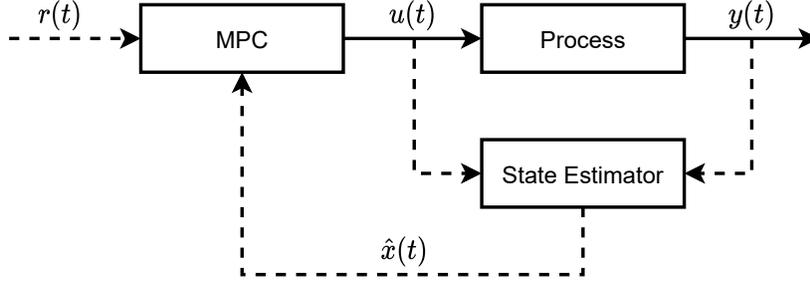
Figure 5: Schematic diagram of the control system

The parameters for the identified model are

$$x(k+1) = \begin{bmatrix} 0.8543 & -0.1164 \\ 0.0195 & 0.8576 \end{bmatrix} x(k) + \begin{bmatrix} -0.0382 & -0.0547 & 0.0103 & 0.1290 \\ -0.0051 & 0.0072 & 0.0020 & 0.0078 \end{bmatrix} u(k)$$
$$y(k) = \begin{bmatrix} -148.6124 & -46.8132 \end{bmatrix} x(k)$$

with the associated optimal steady-state Kalman filter gain being

$$K = \begin{bmatrix} -0.0093 \\ 0.0115 \end{bmatrix}.$$

A summary of the steady-state values as well as the system constraints are presented in Table 6.

A.3 mAbEnv

A.3.1 Mathematical model development

In this section, we present a physics-based mathematical model of the Monoclonal Antibody (mAb) production process. The mAb production process consists of two sub-processes referred to in this

17

Table 6: Summary of the input and output constraints, and their steady-state values

Input	Description	Steady state value [units]	Bounds
q_1	Volumetric flow rate of $S1$	0.4078 [mL]	[0, 5]
q_2	Volumetric flow rate of $S2$	0.1089 [mL]	[0, 5]
q_3	Volumetric flow rate of $S3$	0.3888 [mL]	[0, 5]
q_4	Volumetric flow rate of $S4$	0.2126 [mL]	[0, 5]
y	E-factor	13.057 [kg/kg]	unbounded

work as the upstream and downstream processes. The upstream model presented here is primarily based on the works by Kontoravdi et al. [55, 56] as well as other works in literature and the downstream model is mainly based on the works by Gomis-Fons et al. [57]. We begin the section by first describing the mAb production process. Subsequently, we present the mathematical models of the various units in the mAb production process.

A.3.1.1 Process description   As mentioned earlier, the mAb production process consists of the upstream and the downstream processes. In the upstream process, mAb is produced in a bioreactor which provides a conducive environment for mAb growth. The downstream process on the other hand recovers the mAb from the upstream process for storage. In the upstream process for mAb production, fresh media is fed into the bioreactor where a conducive environment is provided for the growth of mAb. A cooling jacket in which a coolant flows is used to control the temperature of the reaction mixture. The contents exiting the bioreactor are passed through a microfiltration unit which recovers part of the fresh media in the stream. The recovered fresh media is recycled back into the bioreactor while the stream with a high amount of mAb is sent to the downstream process for further processing. A schematic diagram of the upstream process is shown in Figure 6.
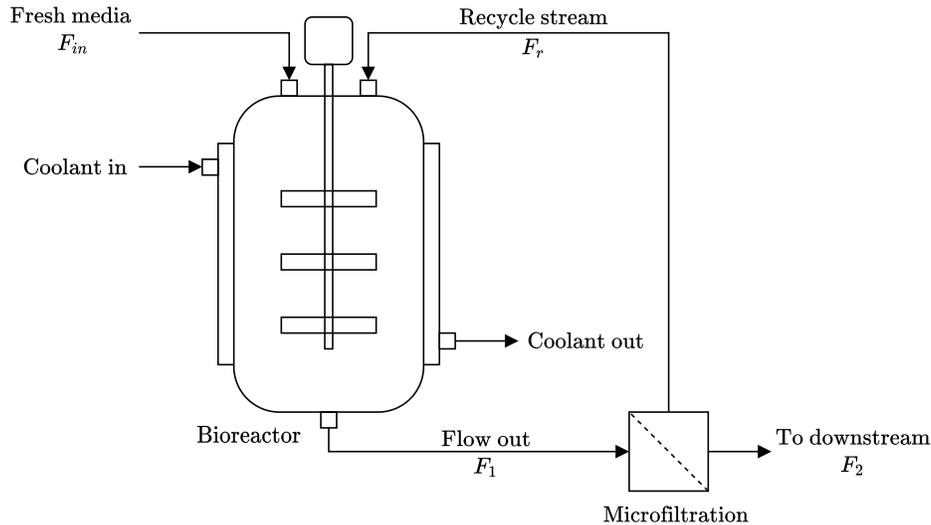
Figure 6: A schematic diagram of the upstream process for mAb production

The objective of the downstream process for mAb production is to purify the stream with a high concentration of mAb from the upstream and obtain the desired product. The configuration of the downstream is adopted from Gomis-Fons' work [57]. It is composed of a set of fractionating columns, for separating mAb from impurities, and holdup loops, for virus inactivation (VI) and pH conditioning. The schematic diagram of the downstream process is shown in Figure 7. Three main steps are considered in the scheme: capture, virus inactivation, and polish. It is worth mentioning that the ultrafiltration preparing the final product is not considered in this work, and hence is not included in the diagram. The capture step serves as the main component in the downstream and the majority of mAb is recovered in this step. Protein A chromatography columns are usually utilized to achieve this goal. The purpose of VI is to disable the virus and prevent further mAb degradation. At

18

last, the polish step further removes the undesired components caused by VI and cation-exchange chromatography (CEX) and anion-exchange chromatography (AEX) are generally used. In order to obtain a continuous manufacturing process, the perfusion cell culture, a continuous mAb culturing process is used in the upstream, however, the nature of chromatography is discontinuous. Therefore, a twin-column configuration is implemented in the capture step. According to the diagram, column A is connected to the stream from the upstream and loaded with the solutions. Simultaneously, column B is connected to the remaining components of the downstream and conducts further mAb purification. According to Gomis-Fons, et al. [57], the time needed for loading is designed as the same as the time required for the remaining purification steps. Hence, when column A is fully loaded, column B is empty and the resin inside is regenerated. Then, the roles of these two columns will be switched in the new configuration, meaning column B will be connected to the upstream and column A will be connected to the remaining components in downstream. In conclusion, a continuous scheme of downstream is achieved by implementing the twin-column configuration in the capture step.
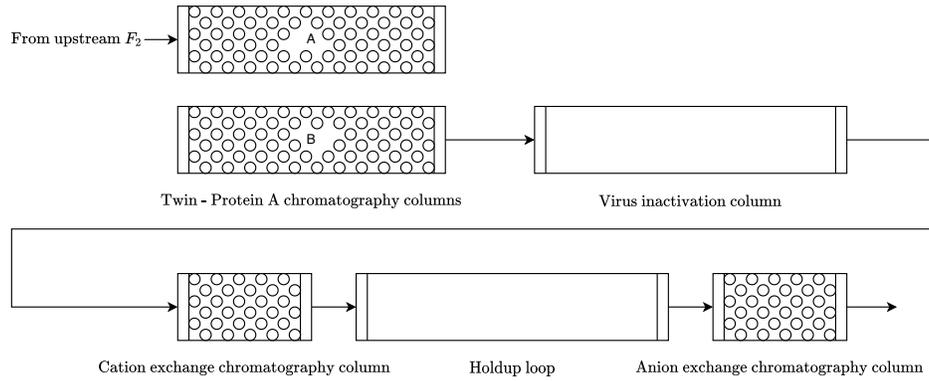
Figure 7: A schematic diagram of the downstream process for mAb production

A.3.1.2 Bioreactor modeling The mathematical model of the bioreactor can be divided into three parts, namely cell growth and death, cell metabolism, and mAb synthesis and production. Papathanasiou and coworkers described a simplified metabolic network of GN-NS0 cells using a Monod kinetic model [58]. In the study by Villiger et al. [59], while the specific productivity of mAb was observed to be constant with respect to viable cell density, it varied with respect to the extracellular pH. By considering these two models, we proposed one simplified model to describe the continuous upstream process. The following assumptions were used in developing the dynamic model of the bioreactor in the continuous production of the mAb process.

- The contents of the bioreactor are perfectly mixed

- The dilution effect is negligible

- The enthalpy change due to cell death is negligible

- There is no heat loss to the external environment

- The temperature of the recycle stream and the temperature of the reaction mixture are equal

A.3.1.2.1 Cell growth and death An overall material balance on the bioreactor yields the equation

$$\frac{dV_1}{dt} = F_{in} + F_r - F_{out}. \tag{7}$$

In Equation (7), V is the volume in L, and F_{in}, F_r, and F_{out} are the volumetric flow rate of the fresh media into the reactor, the volumetric flow rate of the recycle stream and the volumetric flow rate out of the bioreactor respectively in L/min. Throughout this report, the subscripts 1 and 2 represent the bioreactor and the microfiltration unit respectively.

19

The conversion of the viable and total cells within the culture can be described using a component balance on the viable and total number of cells as shown in Equations (8) and (9)

$$\frac{dX_{v1}}{dt} = \mu X_{v1} - \mu_d X_{v1} - \frac{F_{in}}{V_1} X_{v1} + \frac{F_r}{V_1}(X_{vr} - X_{v1}) \tag{8}$$

$$\frac{dX_{t1}}{dt} = \mu X_{v1} - \frac{F_{in}}{V_1} X_{t1} + \frac{F_r}{V_1}(X_{tr} - X_{t1}), \tag{9}$$

where X is the cell concentration in $cells/L$, μ is the specific growth rate in min^{-1} and μ_d is the specific death rate in min^{-1}. The subscripts v and t denote the viable and total cells respectively.

The specific cell growth rate is determined by the concentrations of the two key nutrients namely glucose and glutamine, the two main metabolites namely lactate and ammonia and temperature following the Monod kinetics

$$\mu = \mu_{max} f_{lim} f_{inh} \tag{10}$$

$$f_{lim} = (\frac{[GLC]_1}{K_{glc} + [GLC]_1})(\frac{[GLN]_1}{K_{gln} + [GLN]_1}) \tag{11}$$

$$f_{inh} = (\frac{KI_{lac}}{KI_{lac} + [LAC]_1})(\frac{KI_{amn}}{KI_{amn} + [AMN]_1}). \tag{12}$$

In Equation (10), μ_{max} is the maximum specific growth rate in min^{-1}, f_{lim} and f_{inh} are the nutrient limitation function and the product inhibition function which are described in Equations (11) and (12), respectively. In Equations (11) and (12), $[GLC]$, $[GLN]$, $[LAC]$ and $[AMM]$ stand for the concentrations of glucose, glutamine, lactate and ammonia in mM, and K_{glc}, K_{gln}, KI_{lac} and KI_{amm} represent the Monod constant for glucose, glutamine, lactate and ammonia respectively in mM.

The specific death rate is determined based on the assumption that cell death is only a function of the concentration of ammonia accumulating in the culture, and is shown as follows:

$$\mu_d = \frac{\mu_{d,max}}{1 + (\frac{K_{d,amm}}{[AMM]_1})^n}, \ n > 1. \tag{13}$$

In Equation (13), n is assumed to be greater than 1 to give a steeper increase of specific death as ammonia concentration increases.

Temperature is a key factor in the maintenance of cell viability and productivity in bioreactors. It is expected that the growth and death of the mAb-producing cells will be affected by temperature. The effect of temperature on the specific growth and death rates is achieved through the maximum specific growth and death rates. In this study, standard linear regression of data available in literature [60] was used to obtain a linear relationship between the temperature and the maximum cell growth rate μ_{max}.

$$\mu_{max} = 0.0016T - 0.0308. \tag{14}$$

Similarly, a linear relationship was obtained for the maximum cell death rate as shown in

$$\mu_{d,max} = -0.0045T + 0.1682. \tag{15}$$

In (14) and (15), T is the temperature of the bioreactor mixture in $°C$. The data was obtained for the maximum specific growth and death rates at 33 °C and 37 °C. Therefore, the Equations (14) and (15) are valid only within this temperature range. A heat balance on the bioreactor together with the following above assumptions leads to the following ordinary differential equation:

$$\frac{dT}{dt} = \frac{F_{in}}{V_1}(T_{in} - T) + \frac{-\Delta H}{\rho c_p}(\mu X_{v1}) + \frac{U}{V_1 \rho c_p}(T_c - T). \tag{16}$$

In Equation (16), T_{in} is the temperature of the fresh media in $°C$, ΔH is the heat of reaction due to cell growth in J/mol, ρ is the density of the reaction mixture in g/L, c_p is the specific heat capacity of the reaction in $J/(g \circ C)$, U is the overall heat transfer coefficient in $J/(hr°C))$, and T_c is the temperature of fluid in the jacket in $°C$.

The first term of Equation (16) represents the heat transfer due to the inflow of the feed and the second term represents the heat consumption due to the growth of the cells. The final term describes the external heat transfer to the bioreactor due to the cooling jacket.

20

A.3.1.2.2 Cell metabolism A mass balance on glucose, glutamine, lactate and ammonia around the bioreactor results in the following equations [58]:

$$\frac{d[GLC]_1}{dt} = -Q_{glc}X_{v1} + \frac{F_{in}}{V_1}([GLC]_{in} - [GLC]_1) + \frac{F_r}{V_1}([GLC]_r - [GLC]_1) \tag{17}$$

$$Q_{glc} = \frac{\mu}{Y_{X,glc}} + m_{glc} \tag{18}$$

$$\frac{d[GLN]_1}{dt} = -Q_{gln}X_{v1} - K_{d,gln}[GLN]_1 + \frac{F_{in}}{V_1}([GLN]_{in} - [GLN]_1) - \frac{F_r}{V_1}([GLC]_1 - [GLN]_1) \tag{19}$$

$$Q_{gln} = \frac{\mu}{Y_{X,gln}} + m_{gln} \tag{20}$$

$$m_{gln} = \frac{\alpha_1[GLN]_1}{\alpha_2 + [GLN]_1} \tag{21}$$

$$\frac{d[LAC]_1}{dt} = Q_{lac}X_{v1} - \frac{F_{in}}{V_1}[LAC]_1 + \frac{F_r}{V_1}([LAC]_r - [LAC]_1) \tag{22}$$

$$Q_{lac} = Y_{lac,glc}Q_{glc} \tag{23}$$

$$\frac{d[AMM]_1}{dt} = Q_{amm}X_{v1} + K_{d,gln}[GLN]_1 - \frac{F_{in}}{V_1}[AMM]_1 + \frac{F_r}{V_1}([AMM]_r - [AMM]_1) \tag{24}$$

$$Q_{amm} = Y_{amm,gln}Q_{gln}. \tag{25}$$

A.3.1.2.3 MAb production The rate of mAb production is described as

$$\frac{d[mAb]_1}{dt} = X_{v1}Q_{mAb} - \frac{F_{in}}{V_1}[mAb]_1 + \frac{F_r}{V_1}([mAb]_r - [mAb]_1) \tag{26}$$

$$Q_{mAb} = Q_{mAb}^{max}exp[-\frac{1}{2}(\frac{pH - pH_{opt}}{\omega_{mAb}})^2]. \tag{27}$$

In Equation (27), Q_{mAb}^{max} is the maximum specific productivity with unit $mg/cell/min$, and ω_{mAb} is the pH-dependent productivity constant. pH_{opt} is the optimal culture pH as shown in [59]. The pH value is assumed as a function of state and shown in Section A.3.1.3.2.

A.3.1.3 Mircofiltration

A.3.1.3.1 Cell separation In the cell separation process, a external hollow fiber (HF) filter is used as cell separation device. It is assumed that no reactions occur in the separation process. Hence, the concentration of each variable in recycle stream is shown as follows:

$$X_{vr} = \eta_{rec}X_{v1}\frac{F_1}{F_r} \tag{28}$$

$$X_{tr} = \eta_{rec}X_{t1}\frac{F_1}{F_r} \tag{29}$$

$$[GLC]_r = \eta_{ret}[GLC]_1\frac{F_1}{F_r} \tag{30}$$

$$[GLN]_r = \eta_{ret}[GLN]_1\frac{F_1}{F_r} \tag{31}$$

$$[LAC]_r = \eta_{ret}[LAC]_1\frac{F_1}{F_r} \tag{32}$$

$$[AMM]_r = \eta_{ret}[AMM]_1\frac{F_1}{F_r} \tag{33}$$

$$[mAb]_r = \eta_{ret}[mAb]_1\frac{F_1}{F_r}. \tag{34}$$

21

According to [61], the cell recycle rate (η_{rec}) is assumed to be 92% and the retention rates of glucose, glutamine, lactate, ammonia, and mAb (η_{ret}) are assumed to be 20%.

The material balance around the separation device is shown as:

$$\frac{dV_2}{dt} = F_1 - F_2 - F_r. \tag{35}$$

The mass balance for concentrations of glucose, glutamine, lactate, ammonia, and mAb can be described as:

$$\frac{dX_{v2}}{dt} = \frac{F_1}{V_2}(X_{v1} - X_{v2}) - \frac{F_r}{V_2}(X_{vr} - X_{v2}) \tag{36}$$

$$\frac{dX_{t2}}{dt} = \frac{F_1}{V_2}(X_{t1} - X_{t2}) - \frac{F_r}{V_2}(X_{tr} - X_{t2}) \tag{37}$$

$$\frac{d[GLC]_2}{dt} = \frac{F_1}{V_2}([GLC]_1 - [GLC]_2) - \frac{F_r}{V_2}([GLC]_r - [GLC]_2) \tag{38}$$

$$\frac{d[GLN]_2}{dt} = \frac{F_1}{V_2}([GLN]_1 - [GLN]_2) - \frac{F_r}{V_2}([GLN]_r - [GLN]_2) \tag{39}$$

$$\frac{d[LAC]_2}{dt} = \frac{F_1}{V_2}([LAC]_1 - [LAC]_2) - \frac{F_r}{V_2}([LAC]_r - [LAC]_2) \tag{40}$$

$$\frac{d[AMM]_2}{dt} = \frac{F_1}{V_2}([AMM]_1 - [AMM]_2) - \frac{F_r}{V_2}([AMM]_r - [AMM]_2) \tag{41}$$

$$\frac{d[mAb]_2}{dt} = \frac{F_1}{V_2}([mAb]_1 - [mAb]_2) - \frac{F_r}{V_2}([mAb]_r - [mAb]_2). \tag{42}$$

A.3.1.3.2 pH value pH is defined as the decimal logarithm of the reciprocal of the hydrogen ion activity in a solution. We assume our pH model as follows:

$$pH = \theta_1 - log_{10}(\theta_2[AMM] + \theta_3). \tag{43}$$

After applying nonlinear regression method, we fit the model as:

$$pH = 7.1697 - log_{10}(0.074028[AMM] + 0.968385). \tag{44}$$

A.3.1.4 Downstream modeling The mathematical model of the downstream is constructed based on each unit operation. Specifically, two different models are utilized to describe the loading mode and elution mode of the Protein A chromatography column separately. The models for CEX and AEX share the same mathematical structure with different parameters and the models for VI and holdup loop share the same structures and parameters. A detailed explanation of each model is shown in the following subsections.

A.3.1.4.1 Protein A chromatography column loading mode A schematic diagram [62] depicting a general chromatography column is shown in Figure 8. The column is packed with the porous media which have the binding sites with mAb. The porous media is defined as the stationary phase and the fluid which contains mAb and flows through the column is considered as the mobile phase. Three types of mass transfers are usually considered inside of the column. From the top of the figure, the convection caused by the bulk fluid movement is portrayed. Then by only considering a control volume of the column, which is illustrated in the second subfigure, the dispersion of mAb along the axial direction is shown. Within the beads, there is intra-particle diffusion and in the last subfigure, mAbs are adsorbed on the binding sites of beads.

The general rate model (GRM) simulates the mass transfer in a chromatography column, with the assumption that the transfer along the radial direction of the column is negligible and the transfer along the axial direction of the column and the radial direction in the beads is considered.

In this work, the GRM identified by Perez-Almodovar and Carta [63] is used to describe the loading mode of the Protein A chromatography column. The mass transfer along the axial coordinate is described below:

22
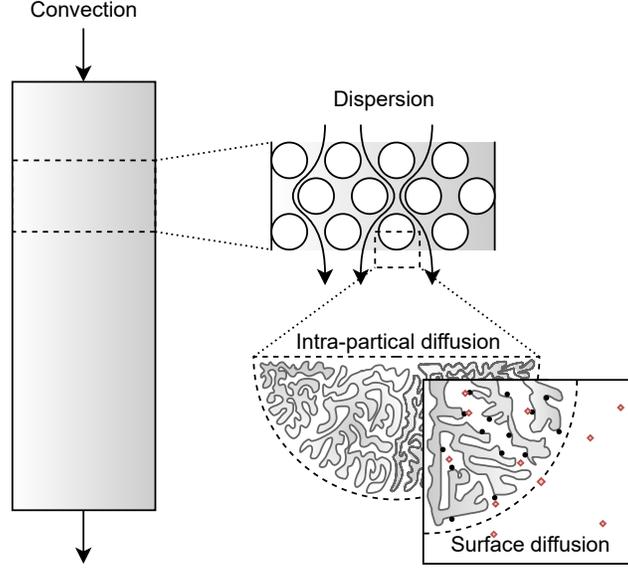
Figure 8: A schematic diagram of the chromatography column

$$\frac{\partial c}{\partial t} = D_{ax}\frac{\partial^2 c}{\partial z^2} - \frac{v}{\epsilon_c}\frac{\partial c}{\partial z} - \frac{1-\epsilon_c}{\epsilon_c}\frac{3}{r_p}k_f(c - c_p|_{r=r_p}), \tag{45}$$

where c denotes the mAb concentration in the mobile phase, changing with time (t) and along with the axial coordinates of columns (z). D_{ax} is the axial dispersion coefficient, v is the superficial fluid velocity, ϵ_c is the extra-particle column void, r_p is the radius of particles and k_f is the mass transfer coefficient.

On the right-hand side of Equation (45), there are three terms. The first term, $\frac{\partial^2 c}{\partial z^2}$, models the dispersion of mAb. In other words, it describes the movement of mAb caused by the concentration difference in the column. The second term $\frac{\partial c}{\partial z}$ denotes the change of concentration of mAb caused by the convection flow. The last term $k_f(c - c_p|_{r=r_p})$ describes the mass transfer between the mobile phase c and the surface of the beads $c_p|_{r=r_p}$.

The boundary conditions of Equation (45) are shown below:

$$\frac{\partial c}{\partial z} = \frac{v}{\epsilon_c D_{ax}}(c - c_F) \quad \text{at} \ \ z = 0 \tag{46a}$$

$$\frac{\partial c}{\partial z} = 0 \quad \text{at} \ \ z = L, \tag{46b}$$

where c_F stands for the harvest mAb concentration from the upstream process.

The concentration of mAb along radial coordinate in the beads (c_p) is the second component of GRM and the mass balance for protein diffusion inside the porous particles is shown in Equation (47) with boundary conditions in Equations (48a) and (48b)

$$\frac{\partial c_p}{\partial t} = D_{eff}\frac{1}{r^2}\frac{\partial}{\partial r}(r^2\frac{\partial c_p}{\partial r}) - \frac{1}{\epsilon_p}\frac{\partial(q_1 + q_2)}{\partial t} \tag{47}$$

$$\frac{\partial c_p}{\partial r} = 0 \quad \text{at} \ \ r = 0 \tag{48a}$$

$$\frac{\partial c_p}{\partial r} = \frac{k_f}{D_{eff}}(c - cp) \quad \text{at} \ \ r = r_p, \tag{48b}$$

23

where D_{eff} is the effective pore diffusivity, r is the distance from the current location to the center of the particle, and ϵ_p is the particle porosity.

At last, the description of adsorbed mAb concentration (q_1 and q_2) is shown as follows:

$$\frac{\partial q_i}{\partial t} = k_i[(q_{max,i} - q_i)c_p|_{r=r_p} - \frac{q_i}{K}] \quad \text{for} \ i = 1, 2, \tag{49}$$

where k_i is the adsorption kinetic constant, q_{max} is the column capacity, and K is the Langmuir equilibrium constant. The reason for having two $\frac{\partial q}{\partial t}$ is because there are two adsorption sites on the beads and one of them is a fast binding site and another one is the slow one.

A.3.1.4.2 Protein A chromatography column elution mode An adsorption kinetic model, convective-dispersive equation with adsorption, is used to describe the elution of the Protein A chromatography column. The setup of boundary conditions for this model can take Equations (46a) and (46b) as the reference, at the same time keeping the inlet and outlet conditions of elution mode in mind. The model is shown as follows:

$$\frac{\partial c}{\partial t} = D_{ax}\frac{\partial^2 c}{\partial z^2} - \frac{v}{\epsilon}\frac{\partial c}{\partial z} + \frac{1 - \epsilon_c}{\epsilon}\frac{\partial q}{\partial t} \tag{50}$$

$$\frac{\partial q}{\partial t} = k[H_0 c_s^{-\beta}(1 - \frac{q}{q_{max}})c - q] \tag{51}$$

$$\frac{\partial c_s}{\partial t} = D_{ax}\frac{\partial^2 c_s}{\partial z^2} - \frac{v}{\epsilon}\frac{\partial c_s}{\partial z}, \tag{52}$$

where c is the mAb concentration in the mobile phase, c_s stands for the modifier concentration, q is the adsorbed mAb concentration. k is the adsorption/desorption rate, H_0 is the Henry equilibrium constant, β is the equilibrium modifier-dependence parameter, and ϵ is the total column void.

On the right hand side of Equation (50), the first two terms are similar with those in Equation (45). The third term $\frac{\partial q}{\partial t}$ is detailed expressed in Equation (51), which is a Langmuir isotherm describing the adsorption and desorption of mAb on beads. This mass transfer is affected by the concentration of the modifier c_s whose dynamics are described in Equation (52).

A.3.1.4.3 CEX and AEX chromatography The adsorption kinetic model shown in Equations (50), (51) and Equation (52) can also be used to describe the CEX and AEX chromatography process. The same rule applies to the boundary conditions. Since the AEX column is in flow-through mode as described in [63], the product mAb is not adsorbed on the beads and the kinetic constant k is supposed to be zero.

A.3.1.4.4 Virus inactivation and holdup pool Equation (53) shows the model of loop for VI and holdup, which is modeled as a one-dimensional dispersive-convective transport, with boundary conditions in Equations (46a) and (46b). Since the loop is not packed, there is no intra-particle diffusion or mass transfer between mAb outside of particles and on the surface of the particles.

$$\frac{\partial c}{\partial t} = D_{ax}\frac{\partial^2 c}{\partial z^2} - v\frac{\partial c}{\partial z}. \tag{53}$$

A.3.2 Control problem formulation and controller design

In this chapter, we present preliminary results of implementing advanced process control (APC) techniques in the operation of the continuous mAb production process. Specifically, two variants of APC algorithms, namely model predictive control (MPC) and economic model predictive control (EMPC) were designed and tested on the mAb production process. We begin the chapter by presenting the control problem to be addressed. Subsequently, we present the various controller designs. Finally, we compare the results of MPC and EMPC.

A.3.2.1 Control problem formulation

24

A.3.2.1.1 Upstream process Before we begin this section, let us rewrite the model of the upstream mAb production process in the state space form

$$\dot{x}(t) = f(x(t), u(t)), \tag{54}$$

where $\dot{x}(t) \in \mathbb{R}^{15}$ is the velocity of the state vector $x \in \mathbb{R}^{15}$ at time t and $u(t) \in \mathbb{R}^{7}$ is the input vector. The variables in the input vector will be defined later in this section. For practical reasons, we assume that the state and input are constrained to be in the spaces $\mathbb{X}$ and $\mathbb{U}$ respectively.

The primary control objective in this work is to ensure that safety and environmental regulations are adhered to during the operation of the mAb production process. From an economic point of view, it is essential to maximize the production of mAb in the upstream process. Thus, two secondary economic objectives are considered. The first is the maximization of the mAb flow rate from the bioreactor while the second is the maximization of the mAb flow rate in the separator (microfiltration unit). These objectives are given as

$$\ell_{\text{bioreactor}} = \text{mAb concentration in bioreactor} \times \text{flow out of the bioreactor} \tag{55}$$

$$\ell_{\text{separator}} = \text{mAb concentration in separator} \times \text{flow out of the separator.} \tag{56}$$

Combining the two economic objectives, the following economic objective is obtained:

$$\ell_e(x, u) = \ell_{\text{bioreactor}} + \ell_{\text{separator}}. \tag{57}$$

To achieve these objectives, we manipulate (as input variables) the flow rates F_{in}, F_r, F_1 and F_2, the coolant temperature T_c together with the concentration of ammonia and glucose in the fresh media stream. Considering the objectives, advanced process control (APC) algorithms that consider the complex system interaction while ensuring constraint satisfaction must be used.

Let us define the steady-state economic optimization with respect to the economic objective ℓ_e as

$$(x_s, u_s) = \arg\min \ -\ell_e(x, u) \tag{58a}$$

$$\text{subject to} \quad 0 = f(x, u) \tag{58b}$$

$$x \in \mathbb{X} \tag{58c}$$

$$u \in \mathbb{U}, \tag{58d}$$

where Equation (58b) is the system model defined in Equation (54) with zero state velocity, and Equations (58c) and (58d) are the the constraints on the state and the input respectively. The negative economic cost function converts the maximization problem to a minimization problem. The optimal value function in (58) is used as the setpoint for MPC to track.

A.3.2.2 Controller design

A.3.2.2.1 Tracking Model Predictive Control (MPC) MPC is a multivariable advanced process control algorithm which has gained significant attention in the process control community. This is because of its ability to handle the complex system interactions and constraints in the controller design. At each sampling time t_k, the following dynamic optimization problem is solved:

$$\min_{\mathbf{u}} \ \int_{t_k}^{t_k + N\Delta} (x(t) - x_s)^T Q(x(t) - x_s) + (u(t) - u_s)^T R(u(t) - u_s) dt \tag{59a}$$

$$\text{subject to} \quad \dot{x}(t) = f(x(t), v(t)) \tag{59b}$$

$$x(t_k) = x(t_k) \tag{59c}$$

$$x(t) \in \mathbb{X} \tag{59d}$$

$$u(t) \in \mathbb{U}. \tag{59e}$$

In the optimization problem (59) above, Equation (59b) is the model constraint which is used to make predictions into the future, Equation (59c) is the initial state constraint, Δ is the sampling time, N is the prediction and control horizons, Equations (59d) and (59e) are the constraints on the state and input respectively, and Q and R are matrices of appropriate dimensions which represent the weights on the deviation of states and the inputs from the setpoint. The setpoint is obtained by solving the steady-state optimization problem in (60). The decision variable $\mathbf{u}$ in (59) is the optimal input sequence for the process. The first input $u(t_k)$ is applied to the system and the optimization problem is solved again after one sampling time.

25

A.3.2.2.2 Economic Model Predictive Control (EMPC) The MPC described in Section A.3.2.2.1 uses a quadratic cost in its formulation. However, in recent years MPC with a general objective is known as economic MPC (EMPC) has received significant attention. The objective function in an EMPC generally reflects some economic performance criterion such as profit maximization or heat minimization. This is in contrast with the tracking MPC described earlier where the objective is a positive definite quadratic function. The integration of process economics directly in the control layer makes EMPC of interest in many areas, especially in the process industry. There have been numerous applications of EMPC.

At each sampling time t_k, the following optimization problem is solved

$$\min_{\mathbf{u}} \quad \int_{t_k}^{t_k+N\Delta} -\ell_e(x(t), u(t))dt \tag{60a}$$

$$\text{subject to} \quad \dot{x}(t) = f(x(t), u(t)) \tag{60b}$$

$$x(t_k) = x(t_k) \tag{60c}$$

$$x(t) \in \mathbb{X} \tag{60d}$$

$$u(t) \in \mathbb{U}. \tag{60e}$$

In the optimization problem (60) above, the constraints are the same as the optimization problem in (59). However, a general cost function is used in place of the quadratic cost function. The benefits of EMPC over MPC will be demonstrated in the results section.

A.3.2.3 Simulation settings After conducting extensive open-loop tests, the control and prediction horizons N for both controllers were fixed at 100. This implies that at a sampling time of 1 hour, the controllers plan 100 hours into the future. The weights on the deviation of the states and input from the setpoint were identifying matrices. As mentioned earlier, the setpoint for the tracking MPC was determined by solving the optimization problem in (58). The optimization problems were implemented using the modeling environment casadi [64] in Python.

A.3.3 Model parameters

Table 7: Parameters for the upstream process model

Parameter	Unit	Value
$K_{d,amm}$	mM	1.76
$K_{d,gln}$	min^{-1}	0.00016
K_{glc}	mM	0.75
K_{gln}	mM	0.038
KI_{amm}	mM	28.48
KI_{lac}	mM	171.76
m_{glc}	$mmol/(cell \cdot min)$	8.2×10^{-16}
Q_{mAb}^{max}	$mg/(cell \cdot min)$	1.1×10^{-11}
$Y_{amm,gln}$	$mmol/mmol$	0.45
$Y_{lac,glc}$	$mmol/mmol$	2.0
$Y_{X,glc}$	$cell/mmol$	2.6×10^8
$Y_{X,gln}$	$cell/mmol$	8.0×10^8
α_1	$(mM \cdot L)/(cell \cdot min)$	5.7×10^{-15}
α_2	mM	4.0
$-\Delta H$	J/mol	5.0×10^5
rho	g/L	1560.0
c_p	$J/(g^\circ C)$	1.244
U	$J/(h^\circ C)$	4×10^2
T_{in}	$^\circ C$	37.0

The parameters of the downstream model are obtained from the work of Gomis-Fons et al. [57] and several parameters are modified because the process is upscaled from lab scale to industrial scale. They are summarized in Table 8.

26

A.3.4 Market value and importance

Drugs based on monoclonal antibodies (mAbs) play an indispensable role in the biopharmaceutical industry in aspects of therapeutic and market potential. In therapy and diagnosis applications, mAbs are widely used for the treatment of autoimmune diseases, cancer, etc. According to a recent publication, mAbs also show promising results in the treatment of COVID-19 [24]. Until September 22, 2020, 94 therapeutic mAbs have been approved by U.S. Food & Drug Administration (FDA) [65] and the number of mAbs approved within 2010-2020 is three times more than those approved before 2010 [66]. In terms of its market value, it is expected to reach a value of \$198.2 billion in 2023. Thus, with the fact that Canada is an active and competitive contributor to the development of high capacity mAb manufacturing processes [67], increasing the production capacity of mAb manufacturing processes is immediately necessary due to the explosive growth in the mAb market. Integrated continuous manufacturing of mAbs represents the state-of-the-art in mAb manufacturing and has attracted a lot of attention because of the steady-state operations, high volumetric productivity, and reduced equipment size and capital cost, etc. [68].

A.4 PenSimEnv

In this section, we briefly describe the mathematical model of the PenSimEnv as presented in Goldrick and coworkers [25]. The model of the fermenter consists of the growth, production, morphology and generation of the biomass. The following key equations describe the four regions in the fermenter, namely growing regions, non-growing regions, degenerated regions and autolysed regions.

$$\frac{dA_0}{dt} = r_b - r_{diff} - \frac{F_{in}A_0}{V} \tag{61}$$

$$\frac{dA_1}{dt} = r_e - r_b + r_{diff} - r_{deg}\frac{F_{in}A_1}{V} \tag{62}$$

$$\frac{dA_3}{dt} = r_{deg} - r_a - \frac{F_{in}A_3}{V} \tag{63}$$

$$\frac{dA_4}{dt} = r_a - \frac{F_{in}A_4}{V} \tag{64}$$

Equations (61) – (64) describe the four regions respectively. In these equations, r_b is the rate of branching, r_{diff} denotes the rate of differentiation, r_e is the rate of extension, r_{deg} is the rate of degeneration, r_a is the rate of autolysis, P is the rate of product formation, h is the rate of hydrolysis, r_m is the rate of maintenance, A_i where $i = 0, 1, 2, 3, 4$ refers to the actively growing regions, non-growing regions, degenerated regions formed through vocuolation and autolysed regions, t is the batch time, F_{in} refers to all the inputs to the process and V is the volume of the fermenter. The total biomass in the system is given as $\sum_{i=0}^{4} A_i$.

The product formation, substrate consumption, and the volume of the fermentation mixture is described in Equations 65 – 67:

$$\frac{dP}{dt} = r_p - r_h - \frac{F_{in}P}{V} \tag{65}$$

$$\frac{ds}{dt} = -Y_{s/X}r_e - Y_{s/X}r_b - m_s r_m - Y_{s/P}r_P + \frac{F_s c_s}{V} + \frac{F_{oil}c_{oil}}{V} \tag{66}$$

$$\frac{dV}{dt} = F_s + F_{oil} + F_{PAA} + F_a + F_b + F_w - F_{evp} - F_{dis} \tag{67}$$

where s is a combined oil and sugar as a single substrate, $Y_{s/X}$ and $Y_{s/P}$ denotes the substrate yield coefficients of biomass and penicillin respectively, m_s refers to the substrate maintenance coefficient, F_{oil} and c_{oil} denotes the feed flow rate and concentration of soya bean oil respectively, F_s and c_s denotes the feed flow rate and concentration of sugar respectively, F_{PAA} is refers to the flow rate of phenylacetic acid, F_a and F_b refers to the to flow rates of the acid and base respectively, F_w is the flow rate of injection water, F_{evap} is the rate of evaporation of the fermenter and F_{dis} is the rate of discharge from the fermenter during production.

Several other equations such as the component balance on the oxygen and nitrogen in the fermenter is also present in the model. A more detailed description can be found in [25].

27

A.5 BeerFMTEnv

The fermentation unit is a critical component in the beer manufacturing process. The dynamic model of the beer fermentation process, as presented in the work by Rodman et al. [69] and de Andres-Toro et al. [70] is described by 7 ordinary differential equations and several temperature dependent parameters. The equations are derived based on the component balances

$$\frac{d[X_A]}{dt} = \mu_x[X_A] - \mu_{DT}[X_A] + \mu_L[X_L] \tag{68}$$

$$\frac{d[X_L]}{dt} = -\mu_L[X_L] \tag{69}$$

$$\frac{d[X_D]}{dt} = \mu_{SD}[X_D] + \mu_{DT}[X_A] \tag{70}$$

$$\frac{d[S]}{dt} = \mu_S[X_A] \tag{71}$$

$$\frac{d[EtOH]}{dt} = f\mu_{EtOH}[X_A] \tag{72}$$

$$\frac{d[DY]}{dt} = \mu_{DY}[S][X_A] - \mu_{AB}[DY][EtOH] \tag{73}$$

$$\frac{d[EA]}{dt} = Y_{EA}\mu_X[X_A] \tag{74}$$

In Equations 68 – 74, the symbol $[\cdot]$ represents the concentration of a component, X_A denotes the active cells, X_L denotes the latent cells, X_D refers to the dead cells, S represents sugar, $EtOH$ denotes ethanol, DY denotes diacetyls and EA represents ethyl acetate. The parameter μ denotes the rates, f is the inhibition factor. More details about the parameters can be found in [69].

Appendix B   Compute

The compute is an RTX 3090 GPU, an RTX 2070 GPU and a GTX 1080 GPU with i9-12900k CPU for a total of 5380 GPU hours.

28

Table 8: Parameters of digital twin of downstream

Step	Parameter	Unit	Value
Capture	$q_{max,1}$	mg/mL	36.45
	k_1	$mL/(mg\ min)$	0.704
	$q_{max,2}$	mg/mL	77.85
	k_2	$mL/(mg\ min)$	$2.1 \cdot 10^{-2}$
	K	mL/mg	15.3
	D_{eff}	cm^2/min	$7.6 \cdot 10^{-5}$
	D_{ax}	cm^2/min	$5.5 \cdot 10^{-1}v$
	k_f	cm/min	$6.7 \cdot 10^{-2}v^{0.58}$
	r_p	cm	$4.25 \cdot 10^{-3}$
	L	cm	20
	V	mL	10^5
	ϵ_c	$-$	0.31
	ϵ_p	$-$	0.94
	$q_{max,elu}$	mg/mL	114.3
	k_{elu}	min^{-1}	0.64
	$H_{0,elu}$	M^β	$2.2 \cdot 10^{-2}$
	β_{elu}	$-$	0.2
Loop	D_{ax}	cm^2/min	$2.9 \cdot 10^2v$
	L	cm	600
	V	mL	$5 \cdot 10^5$
CEX	q_{max}	mg/mL	150.2
	k	min^{-1}	0.99
	H_0	M^β	$6.9 \cdot 10^{-4}$
	β	$-$	8.5
	D_{app}	cm^2/min	$1.1 \cdot 10^{-1}v$
	L	cm	10
	V	mL	$5 \cdot 10^4$
	ϵ_c	$-$	0.34
AEX	D_{app}	cm^2/min	$1.6 \cdot 10^{-1}v$
	k	min^{-1}	0
	L	cm	10
	V	mL	$5 \cdot 10^4$
	ϵ_c	$-$	0.34

29