I&EC research
Industrial & Engineering Chemistry Research

pubs.acs.org/IECR

Article

Toward Performance Improvement of a Baculovirus−Insect Cell System under Uncertain Environment: A Robust Multiobjective Dynamic Optimization Approach for Semibatch Suspension Culture

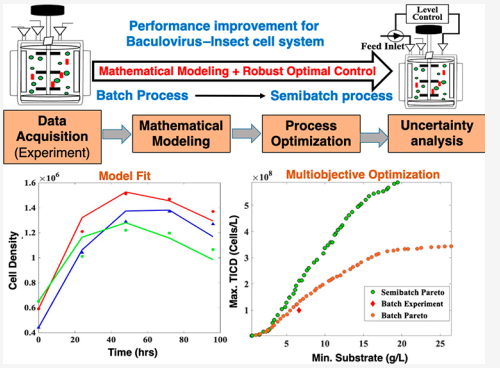
Surbhi Sharma, Pujari Nagasree Keerthi, Lopamudra Giri, and Kishalay Mitra*

Cite This: *Ind. Eng. Chem. Res.* 2023, 62, 111−125

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: The baculovirus expression vector system (BEVS) is one of the well-known versatile platforms for the recombinant protein/vaccine production. Mathematical modeling and optimization of a baculovirus−insect cell system can have significant industrial relevance as this reduces the number of expensive experiments and time involved in the experiment-based optimization. However, modeling and control of such a nonlinear system remains challenging due to the presence of uncertainties in the model. In this context, we propose a novel computational framework combining the principles of systems biology and dynamic optimization under uncertainty for optimizing a semibatch baculovirus−insect cell system. Toward this, first, a mathematical model replicating the dynamic experimental data on cell and virus growth was identified. Next, the proposed model was used for deterministic multiobjective dynamic optimization of the control variables, substrate, and multiplicity of infection (MOI) to achieve the conflicting objectives of productivity maximization and substrate minimization, simultaneously. Finally, based on the sensitivity analysis, six of the most influential parameters depicting model uncertainties have been considered for the robust multiobjective optimal control of the system. A comprehensive comparison displays up to 114% and 76% increases in the cell densities for the deterministic and stochastic semibatch processes, respectively, compared to the batch process. Semibatch operation also favors a minimum 40% reduction in MOI required to achieve the same level of infected cell density compared to the batch operation. This study provides a generic methodology for exhibiting a proof of concept that a semibatch suspension culture considering uncertainty in model parameters can give better productivity compared to a batch suspension culture for a BEVS.



1. INTRODUCTION

The baculovirus expression vector system (BEVS), which was first used in 1980s for the production of human β-interferons,[1] has now been successfully employed for the production of vaccines, antibodies, and many recombinant proteins.[2,3] Although, CHO (Chinese Hamster Ovary) cell-based manufacturing dominates over the other modes of recombinant protein-expression systems, insect cell-based BEVS assumes significance since the process is scalable and provides a cost-effective platform for protein synthesis.[4] In addition to protein/vaccine production, baculoviruses have also been used for biopesticide production because of their specificity to insects and environment friendliness.[5] However, as compared to chemical pesticide, the large-scale continuous production of baculovirus is costly. In principle, expression vectors on baculovirus are developed by expressing any gene of interest in BEVS. The host/insect cells are then infected with the baculovirus carrying the foreign gene, which further results in production of targeted products, e.g., virus-like particles (VLP), recombinant protein, or gene therapy vectors.[2] The FDA-approved human papilloma virus (HPV) vaccine is a VLP-based vaccine produced using BEVS. The BEVS has also been used for obtaining prophylactic vaccine candidates against several other diseases like influenza, Chikungunya, HIV 1, and severe acute respiratory syndrome (SARS).[3] Thus, BEVS remains crucial for vaccine/protein production, and hence, optimization of the baculovirus production process can have significant industrial impact for economic and large-scale production of vaccines, recombinant proteins, and biopesticides.[6]

Experimental process optimization for production of different vaccines and viruses has been reported in literature.[7−9] For optimum production, variables including concentration, pH, temperature, cell culture conditions, multiplicity of infection (MOI), etc. have been taken into account. To find the optimal value for each parameter, various sets of experiments were run

Received: September 16, 2022
Revised: December 3, 2022
Accepted: December 8, 2022
Published: December 23, 2022

https://doi.org/10.1021/acs.iecr.2c03355
Ind. Eng. Chem. Res. 2023, 62, 111−125

using different control values. Especially, experimental optimization and a large number of expensive fermenter runs make it more challenging in translating the biomolecule for industrial production and commercialization. This necessitates the development of a computational framework combining experimental techniques with modeling and optimal control strategies for process optimization in cases of virus/vaccine production.

Although computational model construction is crucial in this context, often, when the protein or antibody productions are performed in fed-batch and batch reactors, there is limited mechanistic information on viral protein production. Hence, identification of a detailed mechanistic model that can simultaneously emulate the experimental data on cell density, cell viability, cell death, and protein expression, remains challenging. To overcome this challenge, unstructured models for quantifying the growth kinetics of cells are often formulated using several empirical mathematical functions available in the literature.[10] The formulation of such predictive models is required for the efficient and extensive industrial production of vaccinations. The need for costly resources, such as sterilization, media, and cell maintenance schedules, as well as lengthy experimental times, further drives the development and identification of optimal unstructured models, which demand very little computational time for simulation and can be used for process optimization.[10]

For large-scale production via BEVS, batch and fed-batch are the two most preferred modes of operations due to their simplicity and flexibility. However, since all the nutrients and biomass are initially added in a batch culture, it results in an accumulation of metabolites as well as depletion of nutrients.[2] Hence, in a batch culture, only limited maximum cell concentrations can be achieved resulting in less productivity. The semibatch culture, where a continuous or intermittent feed addition is applied, can overcome these challenges and, thus, improve the productivity manifold as compared to the batch culture.[11−13] A large number of mathematical models[14] have been developed to simulate the protein expression in a baculovirus system in a batch culture. However, once an adequate kinetic model to mimic the process dynamics is developed, a crucial engineering problem to be solved concerns the development of an optimal operating recipe for transitioning a batch operation to a semibatch mode for optimal operation.[15] In this scenario, if the optimal control can be applied to the bioreactor, it can save costs of extensive experimentation to determine the optimal operating policies for product maximization. However, even if the mathematical model to replicate the bioprocess is available, solving this engineering problem is not an easy task due to the presence of multiple objectives, which are often conflicting in nature, apart from the technological constraints as well as uncertainties originating from multiple sources.[16,17] Since the problem is case dependent, the optimal operating recipe cannot be predetermined and must be derived from the dynamic optimization of the model, which can replicate the key process variable dynamics.

The dynamic optimization problem, also called an optimal control problem (OCP), can be solved by iterative dynamic programming, indirect and direct methods. Indirect approaches transform the original OCP into a Boundary Value Problem (BVP) using Pontryagin's Maximum Principle which can then be solved by gradient, collocation, or multiple shooting methods.[18] Direct methods transform the OCP into a nonlinear programming problem which can be solved using sequential (control vector parametrization, CVP) or simultaneous approaches. CVP has been commonly used for dynamic optimization of bioreactors. Case studies on dynamic optimization of semibatch reactors for ethanol, penicillin, and secreted protein have been reported, where many different approaches have been applied to solve the OCP.[8,19−22] However, in none of the prior studies has modeling and optimal control strategy been combined for determining the optimal operating recipe of a semibatch bioreactor for a baculovirus infection system.

Although such a modeling and optimization exercise can offer significant advantages for large-scale process development, in practice, this analysis can lead to suboptimal or even infeasible estimates without considering the uncertainties which are inherently present in the model.[23] These uncertainties can arise due to several factors such as parameter estimation using noisy data, external process disturbances, high degree of nonlinearity, limited experimental data causing model identifiability problems, model assumptions, etc.[23] Thus, incorporation of such uncertainties is important while performing the optimization in order to obtain robust control profiles (i.e., manipulated variables during the process), which guarantee better objective estimates without violation of any constraints. In the literature, a number of approaches for handling uncertainty while performing optimization have been reported. Among them, some of the more well-known ones include stochastic programming (SP), expected value model (EVM), chance constrained programming (CCP), fuzzy mathematical programming (FMP), and robust optimization (RO).[24] The first three techniques belong to the class of probability-based approaches, which assume that the statistical distribution of uncertain parameters can be made available through data.[25]

One such popular SP technique where decision variables can be used in two stages is two stage stochastic programming (TSSP).[26] Decision variables that are not linked to uncertain parameters are solved in the first step. In the second stage, realizations of uncertain parameters are revealed, and decision factors related to those realizations are also determined. The objective function is defined as the reduction of the total cost associated with the variables in the first stage and the expectation of the variables in the second stage with the related penalty terms for recourse. Assuming various combinations of scenarios for uncertain parameters, several realizations of uncertainties are simulated. The expansion in problem size caused by the increase in the number of unknown parameters and uncertain scenarios is one of the main issues with this technique. Also, decomposing the issue into two parts as indicated above may not always be simple.[27] Additionally, EVM employs constraints and expected values of objectives to counteract the effect of uncertainty, which could lead to conservative solutions.[28,29] Instead of requiring a prior knowledge of the distribution of uncertain parameters, Zimmerman's FMP represents uncertain parameters as fuzzy integers.[30] Despite the fact that it can handle a large number of uncertain parameters without the issue of an explosion in the problem size, FMP has the disadvantage of only partly utilizing the uncertain parameter space.[31,32] Another popular approach, CCP, handles the uncertainty by specifying a confidence level on constraint satisfaction.[33] Here, the concepts of probability are used to express the levels of constraint satisfaction. Thus, a higher probability of constraint violation results in a less reliable solution. However, the solutions obtained by the CCP may not be always robust and largely depend on the probability of constraint satisfaction. Hence, in this work, a Robust

https://doi.org/10.1021/acs.iecr.2c03355
Ind. Eng. Chem. Res. 2023, 62, 111−125

Table 1. Model Description in Terms of ODEs of Model Variables

Variables	ODEs	Description
	$\mu = \mu\max \frac{[S]}{Ks + [S]} \frac{[O2]}{K_{O_2} + [O_2]} \frac{K}{\exp\left(\frac{(CO_2)}{KCO_2}\right) + 1}$	Monod growth induced by O_2 and inhibited by CO_2
Substrate	$\frac{d[S]}{dt} = -\mu \frac{[U] + [I]}{[Y_S]}$	Substrate consumption by uninfected and infected cells
Oxygen	$\frac{d[O_2]}{dt} = k_L a(O^* - [O_2]) - \frac{[U] + [I]}{\left[Y_{O_2}\right]}$	1st term: oxygen transfer rate; 2nd term: oxygen uptake rate by uninfected and infected cells
Uninfected cells	$\frac{d[U]}{dt} = \mu[U] - a_1[V][U] - \frac{k_{d1}[U]}{\exp[k[S]]}$	1st term: uninfected cells generation; 2nd term: virus attachment to the uninfected cells; 3rd term: uninfected cells death rate dependent on substrate.
Infected cells	$\frac{d[I]}{dt} = a_1[V][U] - k_d[I]$	1st term: rate of infected cells generation; 2nd term: infected cells death
Dead cells	$\frac{d[D]}{dt} = \frac{k_{d1}[U]}{\exp(k[S])} + k_d[I]$	1st term: dead uninfected cells; 2nd term: dead infected cells
Virus	$\frac{d[V]}{dt} = k_a[I] - k_{dv}[V]$	1st term: virus production from infected cells; 2nd term: virus decay
Carbon dioxide	$\frac{d[CO_2]}{dt} = k_L a_{CO2}(CO_2^* - [CO_2]) + \frac{[U] + [I]}{\left[Y_{CO_2}\right]}$	1st term: carbon-dioxide transfer rate; 2nd term: carbon-dioxide produced from uninfected and infected cells

optimization framework has been applied for uncertainty handling, which ensures a feasible solution irrespective of the realization of the uncertainty in the parameters.[34] In the RO framework, the stochastic optimization problem is converted into an equivalent deterministic problem, known as the robust counterpart (RC), where the stochastic objective functions and constraints are computed for various realizations of uncertain parameters, and the problem can be solved using standard deterministic optimization techniques.

Optimization under uncertainty has also been applied for other biological processes.[23,34−39] Liu et al. applied the concepts of parametric uncertainty for ensemble modeling of a hybridoma cell culture for maximizing monoclonal antibody production.[36] Optimization under uncertainty has also been performed for various biochemical networks with single objectives as well as the conflicting objectives of minimizing the energy consumption while maximizing the production of a specific metabolite.[35,37] Logist et al. applied multiobjective dynamic optimization for the optimal design and operation of a jacketed tubular reactor and a fed-batch bioreactor with conflicting objectives.[38] However, to the best of our knowledge, this kind of framework has not been applied before for optimizing a baculovirus expression vector system. In practice, a robust multiobjective dynamic optimization problem, when combined with nonlinear dynamic models, can emerge as a very complex and nonconvex optimization task.

The goal of the proposed work is to devise a methodology, combining the experimental studies with modeling and a robust optimal control framework, to determine the optimal feeding recipe of the control variables, which can maximize the final cell density while limiting the consumption of the substrate in a baculovirus−insect cell system. To the best of our knowledge, this work represents the first instance where a multiobjective optimal control (MOOC) under an uncertainty study has been conducted for optimizing the semibatch productivity in context of BEVS. Here, first, the experimental study was conducted by infecting the insect cells with baculovirus in a shaker. Second, a mathematical model of the baculovirus infection process in the insect cells was proposed based on the experimental study conducted. Further, model validations with experimental data from the literature were also performed. Next, the proposed model was used for deterministic multiobjective dynamic optimization of the control variables to achieve the conflicting objectives of productivity maximization and substrate minimization. Finally, based on the parametric sensitivity analysis, six of the most influential parameters depicting uncertainties in the model and feed stream have been considered for robust multiobjective optimal control. A comprehensive comparison study has been conducted to illustrate (a) the effects of changing parametric uncertainty levels, (b) the effects of deterministic and stochastic optimizations on output, and (c) the performance of a semibatch operation as compared to the batch operation in terms of productivity and raw material consumption.

The rest of the paper is organized as follows: Section 2 describes the experimental data generation process followed by modeling, parameter estimation, and sensitivity analysis strategies. This section gives the details on deterministic multiobjective problem formulation and optimization under an uncertainty framework. Section 3 provides the detailed results and discussion followed by the concluding remarks presented in Section 4.

2. MATERIALS AND METHODS

2.1. Viral Strains and Culture Conditions. The *Spodoptera frugiperda* (Sf-21) cell line was grown at 27 °C in Sf-900II (Gibco) medium (without serum and antibiotics) in a 125 mL flask. AcMNPV E2[40] was propagated in Sf-21 cells in a 125 mL shaker flask. A recombinant baculovirus denoted as Ac-FPm was constructed to remove the 13 TTAA transposon target sites in the AcMNPV fp25k gene such that the amino acid sequence of the FP25K protein remained unchanged.[41] Sf-900II growth medium supplemented with 10% fetal bovine serum (FBS) (Gibco, Carlsbad, CA) was used for virus infections with AcMNPV and Ac-FPm during the passaging experiments.[42] The addition of FBS stabilizes the baculovirus stocks stored at 4 °C. Continuous passaging of the two viruses was simulated as described in ref 41.

2.2. Experimental Data. Insect cells (Sf-9) were cultured in nutrient media in 250 mL shaker flasks with working volumes of

https://doi.org/10.1021/acs.iecr.2c03355
Ind. Eng. Chem. Res. 2023, 62, 111−125

100 mL. The insect cell was infected with the wild type baculovirus (AcMNPV), in the batch suspension culture. The initial concentration of the virus used was ten times the concentration of the cells (MOI = 10). Observations of cell density and % viable cells were made every 24 h from $t = 0$ to $t = 96$ h. Thus, the dynamic data were recorded and analyzed at five time points between 0 to 96 h. Triplicate responses for viable cell density and percent cell viability were collected.

2.3. Model Formulation. Identification of a general mathematical structure and estimation of corresponding kinetic parameters are essential for predicting the growth and death profiles of any infection process as infection experiments are costly and time consuming. Hence, the specific challenges in analysis of baculovirus infection processes can be described as follows: (1) identification of a general mathematical model based on a system of nonlinear ODEs that will match with experiments performed in different days and (2) estimations of kinetic parameters that can be used for experiments performed on different days.

In this work, a mathematical model, as described in Table 1, has been proposed to capture the coupled interaction of the cell growth pattern, substrate consumption, oxygen consumption, carbon dioxide production, virus growth rate, and infected cell growth rate, when insect cells are infected with baculovirus. In accordance with the viral growth models, the wild type baculovirus (AcMNPV) first infects the insect cells (SF-9 cells) in the presence of the substrate and oxygen followed by carbon dioxide release. The infection leads to an increase in the number of infected cells, which subsequently produce more virus-like-particles. Specifically, substrate (S) depletion occurs due to consumption by uninfected and infected cells. Here, the yields (amount of cells produced per unit substrate consumed) of uninfected and infected cells have been considered the same. Similar assumptions have been made in the past for modeling a baculovirus−insect cell system.[43] As described by Power et al.,[43] this assumption is identical to assuming that a cell's metabolic activity remains same throughout the culture, whereas the end product changes from cell mass to recombinant products for infected cells. Next, the dissolved oxygen O_2 concentration has been measured as a function of the oxygen transfer and oxygen uptake rate by uninfected and infected cells. The first term in uninfected cell (U) kinetics shows the generation rate followed by the term for the virus attachment to the uninfected cells. The third term shows the substrate-dependent death of the cells, which implies a death rate increase if there is shortage of food/substrate for the cells.[44,45] Similar substrate-dependent death terms have also been used to model bacterial cultures in the past.[10,46] Next, infected cells (I) are formed due to infection of unifected cells by virus followed by the death term. The infected cells are known to halt the process of cell division due to alteration of the cell cycle checkpoints by the viral protein expression; hence, no multiplication term has been considered for cells infected by viruses.[43,47] The dead cell density (D) has been formulated by combining the death terms for the unifected and infected cells. The virus (V) kinetics has been formulated as a function of virus production from the infected cells as followed by decay. Finally, the first term for CO_2 shows a transfer rate similar to that of O_2, and the second term shows the production rate of CO_2 due to uninfected and infected cells.

The set of coupled ordinary differential equation initial value problems (ODE-IVPs), thus formulated (Table 1), was solved using fa ourth-order Runge−Kutta scheme using MATLAB ODE solver ODE45. For biological models, the estimation of kinetic parameters can often be formulated as an optimization problem. Here, the objective function is taken as the root-mean-square of differences (RMSE) between the simulated data generated using kinetic parameters from the model and the corresponding experimental measurements. The optimization algorithm then tries to optimize the kinetic parameters that generate outputs closest to the experimental measurements. All the simulations for model fitting and kinetic parameter estimation have been performed using the sequential quadratic programming algorithm in the MATLAB optimizer *fmincon*.

2.4. Defining the Deterministic Multiobjective Optimization Problem. Experimental studies have shown that a semibatch culture with controlled addition of feed material improves the productivity of the insect cell culture as compared to the batch culture.[2] Hence, the batch culture is not considered as an optimal bioreactor system to achieve high insect cell densities in suspension. During the batch operation, the substrate, virus, and cells are initially loaded; however, the semibatch operation is more complex, as the feeding recipe, whether continuous or intermittent, needs to be controlled throughout the process to ensure the optimal yields. The multiobjective optimal control (MOOC) problem formulated here is based on the semibatch process to show the effects of different addition amounts and patterns on the performance of the baculovirus infection process. Let us consider a general dynamical system of the form

$$X(t) := \left\{ x(t) \in \mathbb{R}^{n_x} u(t) \in \mathbb{R}^{n_u} \;\middle|\; \begin{array}{c} \exists\, x(\cdot): \\ \dot{x}(t) = f(t, x(t), u(t)) \\ x(0) \in X_0 \;\forall\; t \in [0, T] \end{array} \right\} \tag{1}$$

where t, $x(t)$ and $u(t)$ represent time, state vector, and control input, respectively. $X(t)$ represents the set of all the states that can be obtained by simulating the dynamic system in the given time horizon.[48] $X(t)$ can also be written as a formal differential equation of the form

$$X(t) = F(t, X(t), u(t)) \tag{2}$$

A general deterministic multiobjective dynamic optimization problem for such a system can be formulated as

$$\begin{aligned} &\inf_{u(\cdot), X(\cdot)} j_m(X, u) \subset \mathbb{R}^{n_m} \;\forall\; m = 1,\cdots, M \\ &\text{s.t.} \begin{cases} X(t) = F(t, X(t), u(t)) \\ X(0) = X_0 \\ 0 \geq h_i(t, X(t), u(t)) \subset \mathbb{R}^{n_i} \;\forall\; i = 1,\cdots, I \\ \forall\; t \in [0, T] \end{cases} \\ &u = \{u_k\} \subset \mathbb{R}^{n_k} | u_k^{LB} \leq u_k \leq u_k^{UB} \;\forall\; k = 1,\cdots, K \end{aligned} \tag{3}$$

Here, the goal is to devise a control strategy which minimizes the objectives $\{J_{1,\cdots}J_m\}$ subject to various constraints $h_i(x,u,t) \in \mathbb{R}^i$.

During batch operation, the substrate, virus, and cells are initially loaded; however the semibatch operation involves addition of control variables at different time points to achieve the desired objectives. Thus, the optimization problem for the semibatch operation of a baculovirus−insect cell system is formulated as described below.

https://doi.org/10.1021/acs.iecr.2c03355
Ind. Eng. Chem. Res. 2023, 62, 111−125

2.4.1. Selection of Objectives. For any baculovirus−insect cell system, the main goal is to maximize the production of the infected cells. This is because the final product of interest such as recombinant protein or virus-like particles are obtained from the infected cells. Hence, one of the objectives for the MOOC problem is considered as the maximization of the infected cell density. Despite years of insect cell culture, high culture media costs in batch fermentation continue to hamper the application of insect cells.[49] Thus, the second objective is formulated as minimizing the total substrate (amount of media in g/L) fed to the semibatch system. The production of infected cells and substrate is not dichotomous, and minimizing the substrate leads to a decrease in the cell density. Hence, an optimization problem considering the conflicting objectives of maximizing infected cell density and minimizing substrate has been formulated. The optimization exercise is expected to provide a set of decision variables which maximizes the product output while minimizing the substrate input, simultaneously.

2.4.2. Selection of Control Variables. Experimental studies have been performed to manipulate the nutrient addition for maximizing baculovirus production in a semibatch culture.[50,51] Experimental studies have also been performed to optimize the number of infected particles (MOI) for achieving high cell density culture.[9,52,53] These studies have proved that a semibatch addition where feeding is done at different time points during the operation substantially improves cell growth in a BEVS as compared to a batch addition. However, through experimental study, search of the optimal feeding strategy of control variables is not practically possible due to the huge cost and time involved in trial and error using experimentation. Hence, in this study, intermittent additions of the substrate and virus have been analyzed as a way to control the production in the BEVS system. Thus, substrate and virus have been added in a semibatch fashion at different time points. The decision/control variable space consists of discrete time values of substrate addition (described here as $U_1(t_j)$) and virus addition (described here as $U_2(t_j)$). Here, U_1 and U_2 are the two vectors of decision variables (substrate and virus), and t_j is the jth time instant at which semibatch addition is performed. Thus, $U_1(t_j)$) implies the amount of substrate added to the system at jth time instant.

The MOOP, thus formulated, has been described as follows:

Objectives:

$$\underset{U_1,U_2}{\text{Max}}\ \text{TICD} \tag{4}$$

$$\underset{U_1,U_2}{\text{Min}}\ \sum_{j=1}^{N} U_1(t_j) \tag{5}$$

The first objective represents maximizing total infected cell density (TICD), which is calculated by the area under the curve of the infected cells from t = 0 to 96 h. The second objective represents the minimization of total substrate fed to the system. Since the substrate is a control/decision variable defined by the vector U_1, the total amount of substrate is the sum of all the elements of that vector which represents the substrate addition to the system at different time points. These are subject to

Decision variable bounds:

$$U_i^{\min} \le U_i(t_j) \le U_i^{\max} \quad i = 1,2 \text{ and } j = 1,..,4 \tag{6}$$

N denotes the equally spaced time arcs obtained as $\Delta t = \frac{t_f}{N}$ where t_f = 96 h represents final time. In this study, for the sake of practicality in the considered biological system, the value of N is set to 4, which implies that the intermittent addition happens every 24 h in the bioreactor. Thus, both the profiles ($U_1(t_j)$) and $U_2(t_j)$) are discretized into four equally spaced points (one for each time instant t_j = 0, 24, 48, 72 h), thereby making the total number of decision variables as eight. Each of these variables is forced to lie between a lower bound ($U_i^{\min}$) and an upper ($U_i^{\max}$) bound. The upper bounds for decision variables were set same as the batch amounts for substrate and virus. The minimum bound was set to a very small amount, e.g., 0.01 mg/mL for the substrate and 100 mg/mL for the virus (which is equivalent to 0.001 MOI for this system).

Deterministic optimization provides a solution only for a particular instance and can be easily impacted by any possible variation in the uncertain set ξ.[25] In practice, this approach suffers from the drawback of under/overproduction, when the actual uncertain data are above or below the nominal condition. In general, mathematical models might have inherent uncertainties, which could be either due to intrinsic factors or external disturbances.[35] For the given problem, both the objectives (eqs 4 and 5) and decision variables are functions of certain model parameters as shown in Table 1. In this work, it is assumed that the presence of uncertainty is due to the model parameters. Such parametric uncertainty could be due to the sources of variations during experiments as well as the numerical uncertainty involved during regression.[35] To avoid the biased model prediction and control actions, such uncertainties are taken into account for robust optimal control of the process. Generally, these parameters are assumed to be known and fixed to their base (e.g., nominal) values when the optimization problem is solved in deterministic fashion. This, of course, makes the analysis of the problem easier. As these parameters are subject to uncertainty, a more realistic case is to assume them varying, and this way of handling uncertain parameters during the course of optimization is tackled under the paradigm of optimization under uncertainty.[33] In this study, a robust optimization (RO) framework has been adopted for modeling the parametric uncertainty in the baculovirus infection system as described in Section 2.4.

2.4.3. Sensitivity Analysis. Sensitivity analysis is an important tool that is used to identify the parameters affecting the response variable. We carried out a parametric sensitivity analysis of the model to determine the most influential parameters affecting the final infected cell density concentrations. The sensitivities were calculated as the effect of change in the input parameters on the model output over a time span of 96 h. The sensitivities of the system F (eq 7) were calculated by differentiating the system with respect to the kinetic parameter p, which gives a new ODE system to be simultaneously solved along with the original system.

$$F(t, y, y', p) = 0 \tag{7}$$

where the sensitivity with respect to a parameter p can be defined as

$$s_i = \frac{dy}{dp_i} \tag{8}$$

As proposed in ref 54, the sensitivity system may be approximated through a directional derivative finite difference

https://doi.org/10.1021/acs.iecr.2c03355
Ind. Eng. Chem. Res. 2023, 62, 111−125

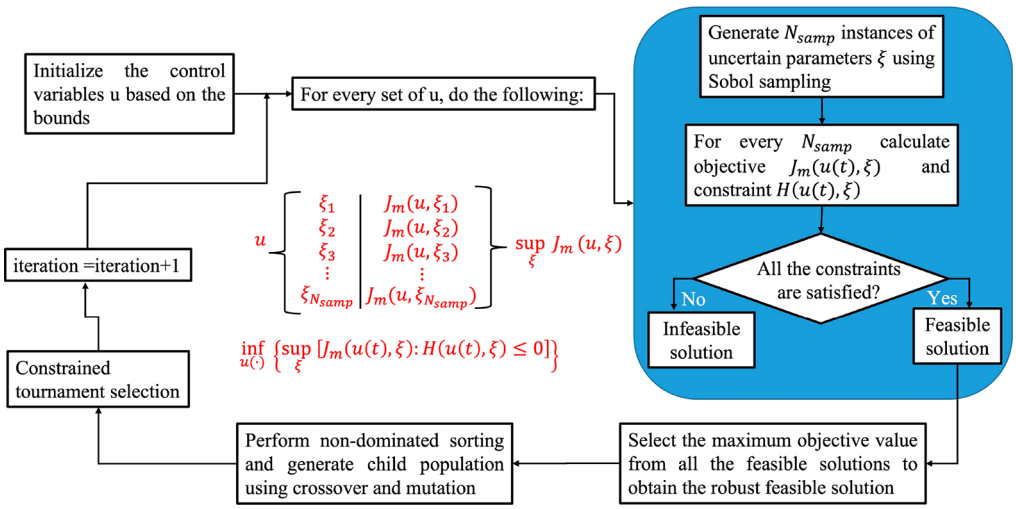


Figure 1. Step by step procedure for the worst-case Robust optimization (eq 11).

approximation. The normalized sensitivity coefficient for each parameter p_i was computed as follows:

$$s_i = \frac{\partial y(t)}{\partial p_i}\frac{p_i}{y(t)} \tag{9}$$

2.5. Multiobjective Dynamic Optimization under Uncertainty. Here, a robust optimization framework has been used to identify the multiobjective Pareto trade-off solutions between the total infected cell density and the substrate concentration under the influence of uncertain parameters selected based on sensitivity analysis. The targeted problem under parametric uncertainties has to be solved considering all uncertain realizations so that the decision variable vector can converge to the robust optimal solutions immune to all such uncertain scenarios. Let us consider a dynamic system as described in eq 1 with a general multi-objective optimization problem defined in eq 3 with M objectives and I constraints. Now, for the uncertain case, where both the objectives and constraints are functions of uncertain parameters ξ, this problem can be reformulated as follows:

Objectives:

$$\inf_{u(\cdot),X(\cdot)} j_m(X, u, \xi) \subset \mathbb{R}^{n_m} \ \forall\ m = 1,\cdots, M$$

$$\text{s.t.}\begin{cases} X(t) = F(t, u(t), X(t), \xi) \\ X(0) = X_0 \\ 0 \ge h_i(t, u(t), X(t), \xi) \subset \mathbb{R}^{n_i} \ \forall\ i = 1,\cdots, I \\ \forall\ t \in [0, T] \end{cases}$$

$$u = \{u_k\} \subset \mathbb{R}^{n_k} | u_k^{LB} \le u_k \le u_k^{UB} \ \forall\ k = 1,\cdots, K$$

$$\xi = \{\xi_d\} \subset \mathbb{R}^{n_d} | \xi_d^{LB} \le \xi_d \le \xi_d^{UB} \ \forall\ d = 1,\cdots, D \tag{10}$$

This standard OUU formulation can be presented in the worst-case robust optimization (Figure 1) framework as

$$\inf_{u(\cdot),X(\cdot)} \sup_{\xi} J_m(X, u, \xi) \subset \mathbb{R}^{n_m} \ \forall\ m = 1,\cdots, M$$

$$\text{s.t.}\begin{cases} X(t) = F(t, u(t), X(t), \xi) \\ X(0) = X_0 \\ 0 \ge H(t, u(t), X(t), \xi) \\ \forall\ t \in [0, T] \end{cases}$$

$$\text{Here we define: } H(t, u(t), X, \xi) = \sup_{\xi} h_i(t, u(t), X, \xi) \subset \mathbb{R}^{n_i} \ \forall\ i = 1,\cdots, I \tag{11}$$

First, the robust feasible solution set is obtained by selecting only those decision variable combinations for which the solution remains feasible across all possible uncertain realizations. Next, the supremum of the objective values (i.e., the worst case) obtained from the robust feasible solution set is minimized across decision variable space. This way of propagating uncertainties into the objective function is also called the worst-case formulation because the worst objective value with respect to the uncertain parameters (ξ) is minimized.

Similarly, a best-case formulation can also be presented (eq 12) by replacing the supremum operator in the inner optimization loop by the infimum operator. This case might be very idealistic to achieve practically under uncertain conditions.

$$\inf_{u(\cdot),X(\cdot)} \inf_{\xi} J_m(X, u, \xi) \subset \mathbb{R}^{n_m} \ \forall\ m = 1,\cdots, M$$

$$\text{s.t.}\begin{cases} X(t) = F(t, u(t), X(t), \xi) \\ X(0) = X_0 \\ 0 \ge H(t, u(t), X(t), \xi) \\ \forall\ t \in [0, T] \end{cases}$$

$$\text{Here we define: } H(t, u(t), X, \xi) = \inf_{\xi} h_i(t, u(t), X, \xi) \subset \mathbb{R}^{n_i} \ \forall\ i = 1,\cdots, I \tag{12}$$

The algorithm of the worst-case RO formulation is shown in Figure 1.

https://doi.org/10.1021/acs.iecr.2c03355
Ind. Eng. Chem. Res. 2023, 62, 111−125

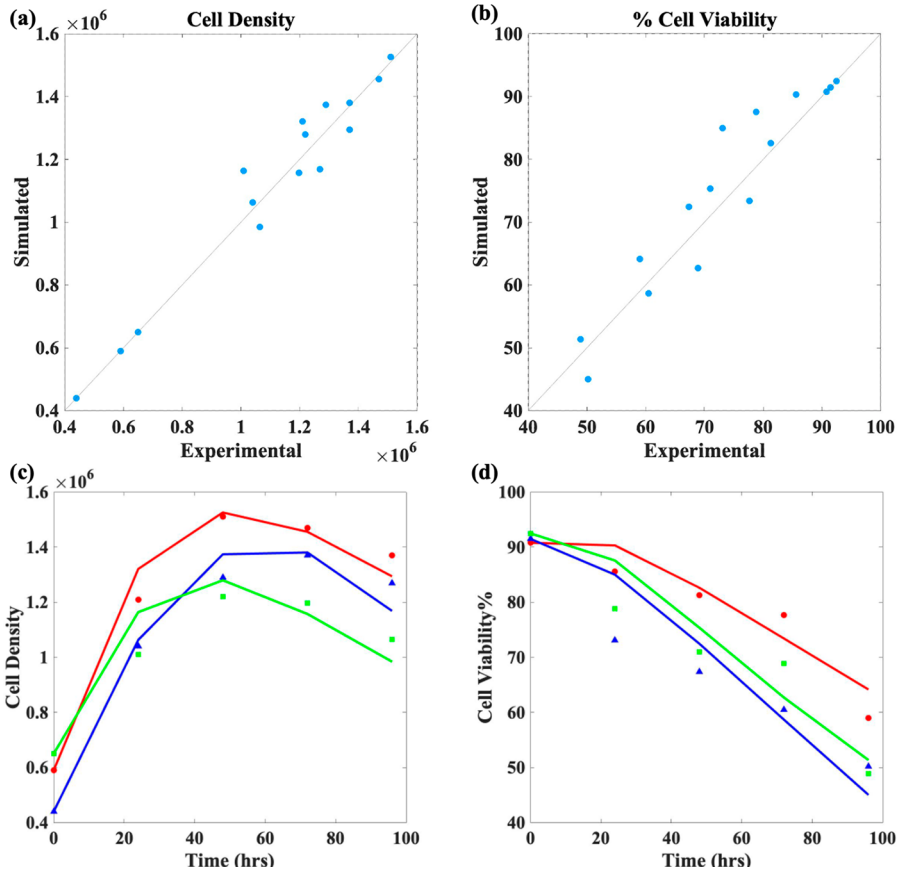


Figure 2. Comparison between experimental and simulated data shown by (a, b) the parity plot and (c, d) the dynamic profiles for cell density and % viability, respectively, obtained from the experimental study.

Here, first the population of decision variable values (population of chromosomes) have been initialized randomly within the given bounds under a multiobjective optimization algorithm framework (NSGA-II in this case). Next, for every chromosome in the population, several instances of uncertain parameters are generated in the given uncertain parameter bounds. Corresponding to every uncertain parameter realization, objective and constraint values are evaluated. A chromosome is declared as the robust feasible solution, if all the constraints are satisfied for all uncertain parameter realizations; otherwise, the chromosome is declared as infeasible. The maximum value of the objective function across several instances of uncertain realizations for a robust feasible solution (i.e., the worst case as we are targeting the minimization problem) is then reported as the fitness value for the chromosome. Once all population members (parent population) are categorized as robust feasible and infeasible, the entire population is sorted using the principles of constrain domination based nondominated sorting and nondominating fronts with different ranks are identified. Simulated binary crossover (SBX) and polynomial mutations are used to generate the child population. Elitist strategy as adopted in NSGA II is used to merge parent and child populations before selecting best solutions from the merged population (keeping the population size same as it was in the beginning). Constrained tournament selection is used for this purpose. This completes one generation of NSGA-II, which is continued in an iterative fashion until a user arrives at a predefined maximum number of generations. Like any other evolutionary multiobjective optimization (EMO) algorithm, NSGA II aims at finding a finite set of Pareto optimal solution sets out of an entire feasible search space, and such a Pareto optimal set provides useful insights about the trade-off solutions of a MOOP. Several different EMOs have been reported in the literature for solving a multiobjective optimization problems.[55] In this work, MOOCs with and without uncertainty have been solved using NSGA-II[56] under the Platemo platform,[57] where the choice of NSGA II is purely based on the past success of it on several challenging problems reported in the literature.

3. RESULTS AND DISCUSSION

Model Validation. The proposed model (Table 1) was first validated with the experimental data for cell density and cell viability obtained by infecting the Sf-9 insect cell with the wild type baculovirus (AcMNPV). Figure 2 shows the comparison of the experiment and the simulation results obtained by fitting the model to the experimental data. The result shows that the model was able to emulate the experimental data well, and an R^2 of 0.99 (RMSE 0.16) was obtained for the fit. Figure S1 also shows the dynamic profile of all the model variables obtained for the batch case. Corresponding parameter values of the fit obtained are presented in Table S1. Further, model validation was performed by fitting the model with other sources of data available in the literature. Figure S2 shows the model fit to experimental data on cell density, cell viability, oxygen, and substrate obtained from one of the studies.[58] Here, the model was able to fit the experimental data with an R^2 of 0.97 (RMSE 0.18). The dynamic data represent the growth of Sf-9 cells in a shaker flask and the corresponding consumption of substrate and oxygen. Further validation of dynamics was performed using another data set

https://doi.org/10.1021/acs.iecr.2c03355
Ind. Eng. Chem. Res. 2023, 62, 111−125

where Sf-9 cells were used to produce virus-like particles,[59] and an R^2 of 0.94 (RMSE 0.13) was obtained for the fit shown in Figure S3. Overall, the result indicates that the model was able to capture the dynamic trends of cell density, cell viability, substrate, and oxygen consumption as obtained from different studies. Unstructured and population based modeling studies showing the growth of baculovirus−insect cell cultures have been conducted in the past.[43,47,60] Licari and Bailey[61] proposed a model for simulating the baculovirus infection in a nonmotile insect cell with different MOIs in a confined environment without nutrient limitation in a batch culture. Power et al.[60] proposed a mathematical model to describe the growth and infection of insect cells by recombinant baculovirus in a batch suspension culture. Another model describing the dynamic process of insect cell infection with the baculovirus at low multiplicity of infection (MOI) in a batch culture was proposed by Zhang and Merchuk.[62] An updated model proposed by Power et al.[43] considers substrate depletion to account for a decrease in product yield. These models describe the dynamic interactions of cells and virus in batch cultures without considering the roles of substrate, oxygen, and carbon dioxide in the system dynamics. To the best of our knowledge, this work reports the first instance where the proposed unstructured model captures the dynamic interaction of cell and virus while also considering the effect of other important growth affecting molecules such as substrate, oxygen, and carbon dioxide.

Deterministic MOOC. The optimal control problem for the considered baculovirus−insect cell system has two conflicting objectives of maximizing total infected cell density and minimizing the total amount of substrate fed to the reactor with two control variables of virus and substrate addition at four different time intervals. The deterministic problem is solved first without considering any uncertainty involved in model parameters. Thus, all the parameters have been fixed to their nominal values as presented in Table S1, which has been estimated by minimizing the error between experimental and simulated data as described earlier. The deterministic MOOP has been solved using NSGA-II,[56] and the corresponding NSGA-II parameters are reported in Table 2. The Pareto optimal (PO) solutions obtained from NSGA-II are shown in Figure 3. To obtain the PO solutions, first, random values of eight decision variables (U_1 and U_2 vector of size four each) have been initiated within their given bounds. After 50 generations (maximum number of iterations), PO solutions obtained from the last generation are marked as semibatch Pareto solutions in Figure 3. Here, it can be seen that the emerged PO front is a nonconvex one, which could have posed difficulties to be obtained using classical derivative-based MOO techniques.[55] Moreover, NSGA-II has provided a wide variety of non-dominated solution alternatives to a decision maker due to its ability to generate well spread PO solutions in a single simulation run. Solutions in the PO set are conflicting in nature,

Table 2. List of Parameters for Solving MOOC Using NSGA-II

NSGA-II Parameters	Values
Number of generations	50
Population size	50
Number of real variables	8
Crossover Probability	0.9
Mutation Probability	0.01

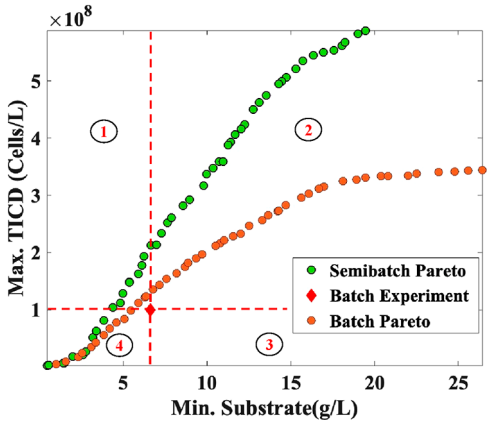


Figure 3. Comparison between deterministic semibatch Pareto optimal results obtained from the last generation simulation of NSGA-II with the batch simulated and experimental results. Red dashed lines represent the division of Pareto space taking batch point as reference point (origin). Thus, the first region is where all the Pareto optimal solutions are better than the batch case in terms of both the objectives. Second and fourth regions are better than the batch case in terms of only one objective, whereas points lying in the third region are worse than the batch case in terms of both the objectives.

where improvement in one objective can be achieved at the cost of the other objective. For the considered system with the objective of minimizing substrate amount and maximizing infected cell density, both the objectives cannot be achieved simultaneously. Thus, minimization of substrate is obtained at the cost of minimizing infected cell density, or in other words, objectives of minimizing substrate and maximizing TICD are conflicting (trade-off) in nature. This further means that none of the PO solutions obtained are better than the others in terms of both the objectives. Hence, it is not trivial to choose a single point from the PO set declaring it as the best (as all of them are nondominating in nature). Irrespective of objectives being proportional or inversely proportional based on inherent system behavior, the trade-offs between the objective functions in terms of the objectives set while formulating the optimization problem frame the ideal platform for multiobjective optimization.

Next, the results for optimized batch Pareto front are presented. This was obtained by setting the upper and lower bounds for all the decision variables (substrate and virus) as 0 for all the intervals except for $t = 0$. It is because for the batch case all the feed is added altogether at $t = 0$, and no intermittent addition is done. The optimizer was run to obtain the optimal decision variables for $t = 0$ for the batch operation. The obtained Pareto front for the batch case is shown in Figure 3 marked as batch Pareto. The result shows that a semibatch operation with an intermittent injection of substrate and virus at different time points provides a higher cell density as compared to a batch operation in all the cases. We further compared the semibatch process with the experimental batch point. Figure 3 shows the preferred Pareto space when the batch experimental case is taken as the reference point. The result shows that the Pareto solutions in region 1 are better than the batch point in terms of both the objectives. This implies that for the points in region 1, a semibatch operation provides a higher cell density for the lower amount of substrate as compared to the batch process. The points in regions 2 and 4 are better than the batch point only in terms of one of the objectives. Thus, the points in region 2 have higher cell density, but those are obtained at the cost of higher

https://doi.org/10.1021/acs.iecr.2c03355
Ind. Eng. Chem. Res. 2023, 62, 111−125

substrate consumption, whereas points in region 4 have lower cell density, but the amount of substrate consumption is also reduced. It can also be seen that the points in region 3 are the worst, as they are inferior to the batch point in terms of both the objectives. Some of the randomly generated points in the first generation of NSGA-II lie in the third region. However, none of the Pareto optimal points obtained from the last generation lie in the third region which proves the ability of NSGA-II in providing the optimized objective values.

Deterministic MOOC: Quantitative Comparison of Semibatch and Batch Case. A comparison between the infected cell density, MOI, and substrate consumed in batch and semibatch processes is presented in Table 3. From the final PO solution for the semibatch process, two solutions have been selected for comparison. For the first point, the amount of substrate consumed is similar order of magnitude to the batch case, whereas for the second point, the total infected cell density is similar in order of magnitude to the batch case. The results show that for the first case, where the amount of total substrate added for the semibatch is same as the batch (6.6 mg/mL), the total infected cell density obtained from the semibatch is more than 2-fold (114%) higher as compared to the batch process. Also, the amount of virus (MOI) required to achieve such a higher cell density in the semibatch is 45% less as compared to the batch case. For the second Pareto point, where the amount of total infected cells obtained from the semibatch process is similar to the batch case, the total substrate consumption was 34.8% less for the semibatch case as compared to the batch. For the same setting, the MOI required for semibatch was 51% lesser than the batch case. It was also observed that for all the final Pareto solutions obtained for the semibatch process, the MOI values were 45%−76% lesser compared to the batch process. This way of formulating the optimal control problem is advantageous since this enables handling either of a batch or semibatch operation. This is because in a semibatch formulation, all control variables added in the beginning could have symbolized a batch operation, and it is purely the choice of the optimizer to choose such a mode of operation, if found optimal. In this case, none of the PO solutions are found to symbolize batch operation. Hence, the results show that use of a suitably formulated feed medium and optimizing the feeding recipe in a semibatch process can substantially increase the product formation in a BEVS relative to a batch culture. A similar comparison between the batch and fed-batch operations for a baculovirus−insect cell system has been done in the past using the experimental methods.[63−66] Elias et al.[65] showed that by carefully manipulating the feeding strategy of the nutrients, a significantly higher cell density can be obtained using a fed-batch culture as compared to a batch culture. Chan et al.[63,64] also reported a 2- to 3-fold increase in the productivity by using a fed-batch operation for the baculovirus−insect cell culture instead of the batch operation. Meghrous et al.[66] reported a 2- to 3-fold enhancement in the production of recombinant protein for the production of influenza virus from the fed-batch process as compared to the batch process. Thus, the optimization exercise performed in this work shows an agreement with these findings. However, experimental optimization of the feeding strategy is a laborious and time-consuming task. Hence, this work can guide in process optimization, as the Pareto optimal solutions can be

Table 3. Comparing the Pareto Optimal Solutions Obtained from Deterministic and Nondeterministic MOOC with the Batch Case

Reactor type	Total infected cell density produced (mL)	Substrate consumed (mL)	MOI
Batch	9.95×10^7	6.61	10
	Deterministic		
Semibatch 1	2.13×10^8	6.64	5.49
Semibatch 2	1.04×10^8	4.37	4.94
	RO (5% uncertainty)		
Semibatch 1	1.7×10^8	6.4	6.11
Semibatch 2	9.7×10^7	4.66	5.78

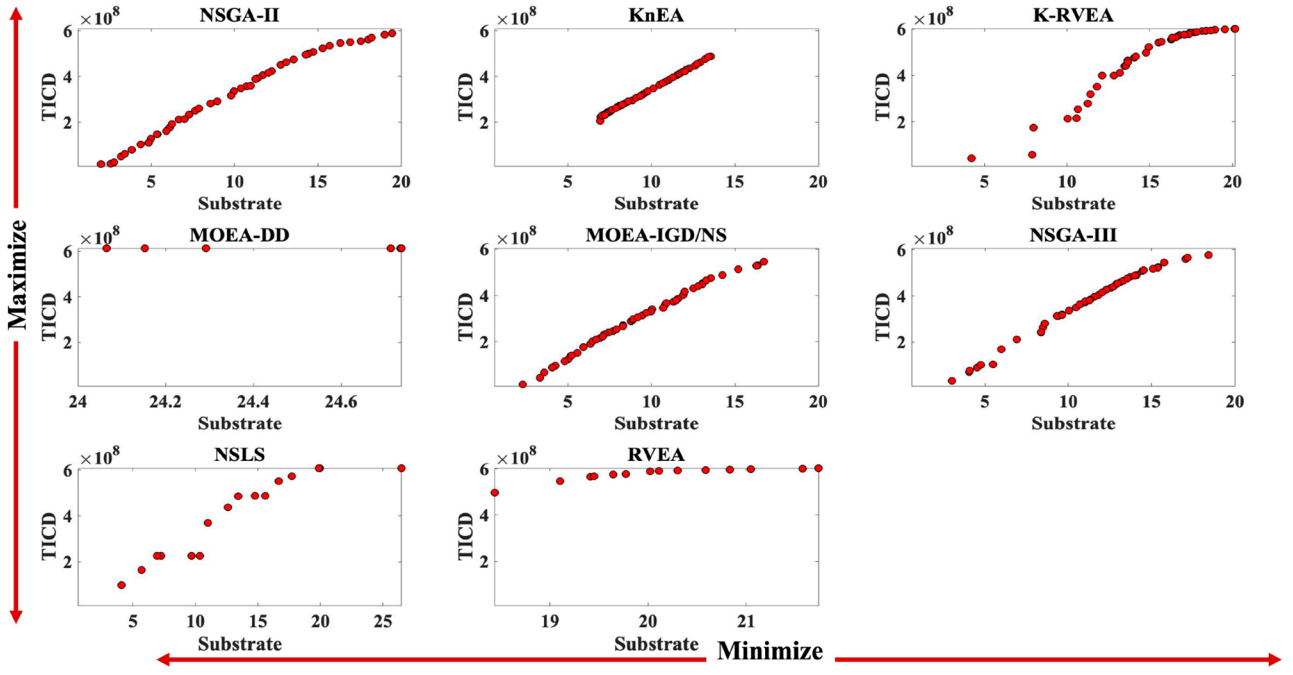


Figure 4. Pareto optimal solutions obtained for deterministic MOOC using different MOEAs.

https://doi.org/10.1021/acs.iecr.2c03355
Ind. Eng. Chem. Res. 2023, 62, 111−125

Table 4. Metrics Comparing the Quality of Pareto Fronts Obtained Using NSGA-II and Other Novel MOO Algorithms

Algorithms	DeltaP (min)	DM (max)	GD (min)	HV (max)	IGD (min)	PD (max)	Spread (min)
NSGA-II	**4344335.03**	**0.836**	65334.3698	0.586	**4344335.03**	**5182586255**	0.635
KnEA	56126584.7	0.379	81102.1752	**0.614**	56126584.7	3040832762	0.896
K-RVEA	12320490.1	0.515	**44465.5219**	0.504	12320490.1	3533782444	1.259
MOEA-DD	316001593	0.167	1221153.54	0	316001593	3799245.15	0.999
MOEA-IGD/NS	7444034.16	0.724	92403.3685	0.613	7444034.16	4601057942	**0.482**
NSGA-III	6838796.26	0.694	92664.7714	0.586	6838796.26	4266591888	0.814
NSLS	20333700.8	0.646	104834.924	0.534	20333700.8	3074028126	1.684
RVEA	213034460	0.191	155870.391	0.183	213034460	530436952	1.05

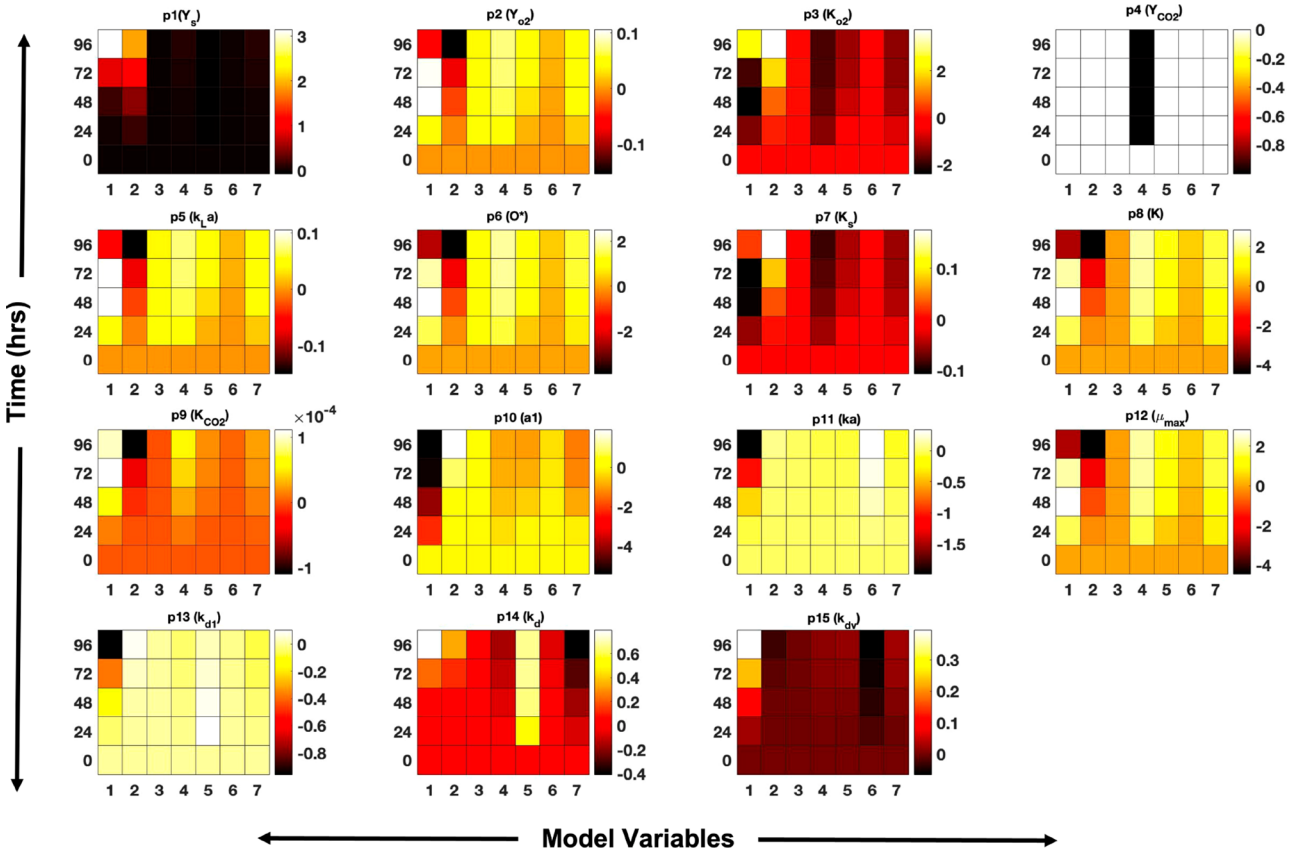


Figure 5. Dynamic sensitivities of all the model variables with respect to all the model parameters (p1−p15). Here *X*-axis represents the seven model variables in the order: uninfected cells, substrate, oxygen, carbon dioxide, dead cells, virus, and infected cells. *Y*-axis represents the time in hours. Color bar represents the sensitivity values.

experimentally validated saving the time and resources to run multiple experiments for determining the optimal feeding strategy to improve the productivity.

Deterministic MOOC: Comparison with Other Multiobjective Evolutionary Algorithms (MOEAs). To check the competence of NSGA-II with other novel MOO algorithms, the MOOC has been solved using seven other novel algorithms.[57] Figure 4 shows the results of the deterministic MOOC solved using several such algorithms, e.g., knee point-driven evolutionary algorithm (KnEA), kriging-assisted reference vector-guided evolutionary algorithm (K-RVEA), many-objective evolutionary algorithm based on dominance and decomposition (MOEA-DD), multiobjective evolutionary algorithm based on an enhanced inverted generational distance metric (MOEA/IGD-NS), NSGA-III, nondominated sorting and local search (NSLS), and reference vector-guided evolutionary algorithm (RVEA). The quality of the PO front obtained by each algorithm has been compared using seven different metrics:[57] averaged Hausdorff distance (DeltaP), diversity metric (DM), generation distance (GD), hyper-volume (HV), inverted GD (IGD), pure diversity metric(PD), and spread. The metric values corresponding to every algorithm are presented in Table 4. The results show that performance of NSGA-II is better/comparable to most of the novel MOO algorithms under consideration for the baculovirus system. Therefore, it has been decided to use NSGA-II for optimization related to other nondeterministic cases. The rationale behind trying several evolutionary algorithms is their stochastic nature. These algorithms have been run several times before reporting the PO solutions obtained from them. For reporting PO solutions from each algorithm, they are run several times, and the nondominated solutions obtained from solutions accumulated over several runs are reported.

Sensitivity Analysis. Next, in order to identify the most influential parameters, which affect the infected cell density and substrate consumption, the parametric sensitivity analysis was

https://doi.org/10.1021/acs.iecr.2c03355
Ind. Eng. Chem. Res. 2023, 62, 111−125

performed. Figure 5 shows the dynamic sensitivity of model variables with respect to all parameters as calculated by solving the system given in eqs 8 and 9 for the proposed model. The result shows that the highest sensitivity indices (>1) were observed with respect to the parameters representing the growth rate (μ_{max}), half velocity constant (K_{O_2}), equilibrium constant (K), yield with respect to substrate (Y_S), oxygen saturation constant (O*), and virus attachment rate (a_1). These highly sensitive parameters were finally considered for the optimization runs under model parameter uncertainty.

Optimization under Uncertainty (OUU). For solving the MOOC problem under uncertainty, a six-dimensional uncertain data set corresponding to six selected uncertain parameters has been generated. Different numbers of uncertain parameter samples were chosen as 50, 100, 200, and 500 to compare their effects on the objective functions. The uncertain parameters have been uniformly sampled using a Sobol sampling scheme (reference) within ±5% of the base parameter (nominal) values obtained from the parameter estimation exercise. Figure 6 shows

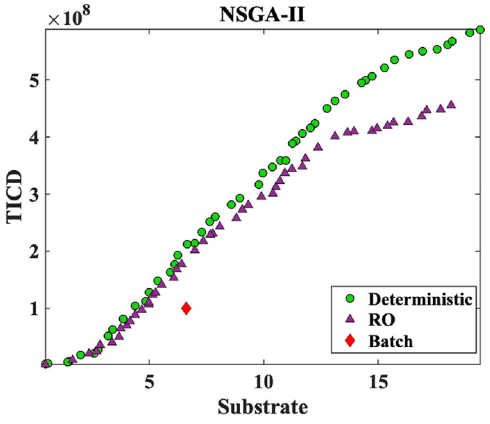


Figure 6. Comparison between deterministic and nondeterministic Pareto solutions obtained by solving MOOC problem.

the comparison between the Pareto sets obtained through deterministic and optimization under uncertainty formulations. The IGD value of 4.3×10^6 was obtained for the deterministic case, whereas an IGD value of 2.4×10^7 was obtained for the RO formulation with 50 sample points. The result shows that with RO formulation only a part of the deterministic Pareto optimal front overlaps with the robust front. This means there are many deterministic PO solutions, which seemed to be better and feasible for the deterministic case, but found infeasible when uncertainty in model parameters is considered. This is one of the serious drawbacks of considering deterministic formulations ignoring the sources of uncertainty, where the reality involves uncertainty in parameters. PO solutions corresponding to the RO formulation might appear conservative and inferior to the deterministic case; however, this is natural as these solutions are immune to various possible scenarios of uncertainty realizations and are obtained through the worst-case RO formulation.

OUU: Quantitative Comparison with Batch Case. As compared to the batch case, the robust solutions for the semibatch results in 70% more production of total infected cell density for a similar amount of substrate consumed (Table 3). The amount of virus particles (MOI) was also 40% less as compared to the batch case. Overall, the results show that the RO Pareto solutions provide better yield as compared to the batch case. Further, the effect of varying sample points on the uncertain Pareto front was also studied (Figure S4). The corresponding CPU time as well as IGD values for the uncertain parameter samples of 50, 100, 200, and 500 are also reported in Table S2. The result shows emergence of similar type of PO solution set for different types of sample sizes as mentioned above; however, a significant increase in the CPU time was observed (703 s for 50 sample size to 7280 s for 500 sample size). Hence, the sample size for uncertainty evaluation was chosen as 50 for rest of the analysis.

OUU: Comparison with Varying Degrees of Uncertainty. Next, a comparison has been made between the varying degrees of uncertainty in the parameters starting from 2.5% to 10% variation from the nominal value (Figure S5). As the degree of uncertainty increases, the Pareto front becomes more conservative, where a higher substrate amount is required to achieve a given cell density. A comparison between the control variable feeding recipe between the deterministic case as well as the stochastic case with 5% and 10% uncertainties is also shown in Figure 7. The results show that the deterministic case assumes a lower substrate consumption and comparatively lower MOI for achieving a higher cell density compared to the stochastic case. However, in the presence of parametric uncertainty, the deterministic solutions are practically meaningless, and considering the practical scenario of uncertainty, a higher level of uncertainty results in a more conservative solution. The CPU time and IGD values for each varying degree are shown in Table S3. As the uncertainty increases, the Pareto front becomes more conservative, and thus, a higher substrate consumption is required to achieve the same amount of infected cell density.

OUU: Considering Only Three Uncertain Parameters. Further uncertainty analysis was performed by considering only three uncertain parameter (K_{O_2}, O*, and μ_{max}) as shown in Figure S6. These three parameters were chosen based on the effects of uncertain parameters on both the infected cell density and substrate because infected cell density and substrate mainly determine the objective values. Since the parameter a_1 affected infected cell levels and did not affect substrate, it was not selected in the three-parameter group. Similarly, Y_S affected substrate levels, but infected cells did not show any sensitivity with respect to Y_S. This was also not considered. The remaining four parameters K, μ_{max}, K_{O_2}, and O* had influence on both infected cells and substrate. However, parameters K and μ_{max} had similar influences and reported the same sensitivity indices for both infected cells and substrate. Hence, only one of them is chosen to represent the effect in the reduced three parameter set. The IGD value for the three uncertain parameter case was 1.8×10^7, which is less compared to the six uncertain parameter case. The result shows that optimistic Pareto optimal solutions are obtained when a smaller number of parameters are considered uncertain. However, the lower part of the Pareto front almost coincides in both the cases.

Finally, it can be concluded that in all the cases of uncertainty considered for the semibatch process, a better performance in terms of decreased substrate consumption and increased total infected cell density as well as lower MOI can be obtained for the semibatch process as compared to the batch process.

4. CONCLUSION

This work focuses on the modeling, multiobjective optimization, and robust control of a baculovirus−insect cell system for obtaining maximum cell density and minimum substrate consumption, simultaneously. BEVS is crucial for the develop-

https://doi.org/10.1021/acs.iecr.2c03355
Ind. Eng. Chem. Res. 2023, 62, 111−125

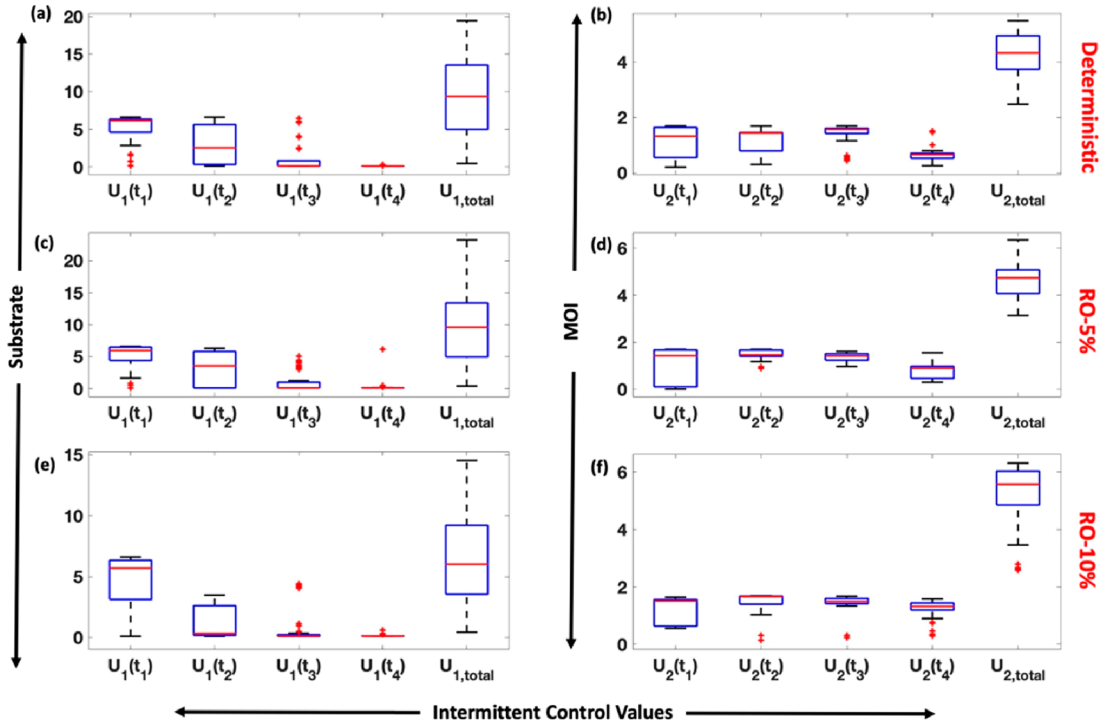


Figure 7. Box plots representing the variation of optimized control values of substrate and MOI at every interval for (a, b) deterministic case, (c, d) RO with 5% uncertainty in parameters, and (e, f) RO with 10% uncertainty in parameters.

ment of recombinant proteins, VLPs, and vaccines[2,3] apart from its extensive use as biopesticides.[5] Hence, modeling and optimization of the baculovirus infection system can greatly contribute to the large-scale and economic production of proteins/infected cells by limiting the use of expensive experiment-based routes. Several experimental and modeling studies have been conducted to study the growth of insect cells in a baculovirus−insect cell system.[6,13,14,59−62,67,68] However, none of the previous studies amalgamates the experimental findings with modeling and multiobjective dynamic robust control in a semibatch mode to maximize the productivity for a BEVS.

Here, we first conducted the experimental studies, where insect cells (Sf-9 cells) were infected with baculovirus (AcMNPV), and the dynamic evolutions of cell density and cell viability were observed. Next, a mathematical model with an objective of minimizing the error between experimental and simulated data was proposed with uninfected cell density, substrate, oxygen, carbon dioxide, dead cells, infected cells, and virus as the main model variables. Model validation with the obtained experimental data as well as other data obtained from the literature show that the model is capable of emulating the dynamic trends of cell density, cell viability, substrate, and oxygen.

We further showed that instead of adopting a batch approach, if a semibatch approach with intermittent addition of substrate and virus was adopted, a considerable increase in the infected cell density (more than 2 folds) could be obtained for a lesser amount of substrate consumption. Previous experimental studies also corroborate such facts of increase in productivity by adopting a semibatch operation compared to a batch operation.[2,49,64,69] Next, we performed a directional derivative-based parametric sensitivity analysis[54] to identify the most influential parameters for uncertainty analysis. In none of the previous studies for baculovirus−insect cell systems has such an analysis based on parameter uncertainty been performed to check the feasibility of the results. A comprehensive comparison between varying the sample size of uncertain parameters and uncertainty levels indicates that with 5% variation in uncertain parameters from their nominal values, a 70% higher cell density production and up to 40% less MOI are possible for a semibatch case compared to the batch process. It is also observed that higher level of uncertainty results in more conservative solutions to ascertain robustness.

The proposed model can also be updated by including aspects for protein expression which can be further used for improving performance of a BEVS. However, it remains challenging to obtain data on protein expression at a higher resolution as Western blot experiments are time intensive and tedious and require larger sample volumes.[70] Also, to build the dynamic model, continuous protein expression measurements at multiple time intervals are required, which is difficult to obtain. Further, other variables such as oxygen and cell density can also be manipulated using the proposed framework for process optimization at a higher resolution.

Hence, this work provides a computational framework and shows a proof of concept that an intermittent addition of reactants can improve the yield in a BEVS, instead of a batch operation. Alhough further experimental studies need to be conducted to validate the obtained results, this work can act as a starting point and significantly guide toward optimizing the process development for large-scale production in a BEVS.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.iecr.2c03355.

https://doi.org/10.1021/acs.iecr.2c03355
Ind. Eng. Chem. Res. 2023, 62, 111−125

Additional model validation results, Pareto comparison for varying sample sizes and uncertainty level, and model parameter values (PDF)

AUTHOR INFORMATION

Corresponding Author

Kishalay Mitra − *Department of Chemical Engineering, Indian Institute of Technology, Hyderabad, Telangana 502284, India;* orcid.org/0000-0001-5660-6878; Email: kishalay@iith.ac.in

Authors

Surbhi Sharma − *Department of Chemical Engineering, Indian Institute of Technology, Hyderabad, Telangana 502284, India*

Pujari Nagasree Keerthi − *Department of Chemical Engineering, Indian Institute of Technology, Hyderabad, Telangana 502284, India*

Lopamudra Giri − *Department of Chemical Engineering, Indian Institute of Technology, Hyderabad, Telangana 502284, India;* orcid.org/0000-0002-3099-9068

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.iecr.2c03355

Author Contributions

The manuscript was written through contributions of all the authors. All the authors have given approval to the final version of the manuscript.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

K. Mitra acknowledges the support provided by the project BT/PR34209/AI/133/19/2019 funded by the Department of Biotechnology (DBT), Government of India, for this work.

REFERENCES

(1) Smith, G. E.; Summers, M. D.; Fraser, M. J. Production of Human Beta Interferon in Insect Cells Infected with a Baculovirus Expression Vector. *Mol. Cell. Biol.* **1983**, *3* (12), 2156−2165.

(2) Contreras-Gómez, A.; Sánchez-Mirón, A.; García-Camacho, F.; Molina-Grima, E.; Chisti, Y. Protein Production Using the Baculovirus-Insect Cell Expression System. *Biotechnol. Prog.* **2014**, *30* (1), 1−18.

(3) Nooraei, S.; Bahrulolum, H.; Hoseini, Z. S.; Katalani, C.; Hajizade, A.; Easton, A. J.; Ahmadian, G. Virus-like Particles: Preparation, Immunogenicity and Their Roles as Nanovaccines and Drug Nanocarriers. *J. Nanobiotechnology* **2021**, *19* (1), 59.

(4) Tripathi, N. K.; Shrivastava, A. Recent Developments in Bioprocessing of Recombinant Proteins: Expression Hosts and Process Development. *Front. Bioeng. Biotechnol* **2019**, *7*, 420.

(5) Lengai, G. M. W.; Muthomi, J. W. Biopesticides and Their Role in Sustainable Agricultural Production. *J. Biosci. Med.* **2018**, *06* (06), 7−41.

(6) Demain, A. L.; Vaishnav, P. Production of Recombinant Proteins by Microbes and Higher Organisms. *Biotechnol. Adv.* **2009**, *27* (3), 297−306.

(7) Valkama, A. J.; Leinonen, H. M.; Lipponen, E. M.; Turkki, V.; Malinen, J.; Heikura, T.; Ylä-Herttuala, S.; Lesch, H. P. Optimization of Lentiviral Vector Production for Scale-up in Fixed-Bed Bioreactor. *Gene Ther.* **2018**, *25* (1), 39−46.

(8) Fricke, J.; Pohlmann, K.; Tatge, F.; Lang, R.; Faber, B.; Luttmann, R. A Multi-Bioreactor System for Optimal Production of Malaria Vaccines with Pichia Pastoris. *Biotechnol. J.* **2011**, *6* (4), 437−451.

(9) Aggarwal, K.; Jing, F.; Maranga, L.; Liu, J. Bioprocess Optimization for Cell Culture Based Influenza Vaccine Production. *Vaccine* **2011**, *29* (17), 3320−3328.

(10) Singh, R.; Sharma, S.; Kareenhalli, V. V.; Giri, L.; Mitra, K. Experimental Investigation into Indole Production Using Passaging of E. Coli and B. Subtilis along with Unstructured Modeling and Parameter Estimation Using Dynamic Optimization: An Integrated Framework. *Biochem. Eng. J.* **2020**, *163*, 107743.

(11) Joshi, P. R. H.; Cervera, L.; Ahmed, I.; Kondratov, O.; Zolotukhin, S.; Schrag, J.; Chahal, P. S.; Kamen, A. A. Achieving High-Yield Production of Functional AAV5 Gene Delivery Vectors via Fedbatch in an Insect Cell-One Baculovirus System. *Mol. Ther. - Methods Clin. Dev* **2019**, *13*, 279−289.

(12) Nguyen, B.; Jarnagin, K.; Williams, S.; Chan, H.; Barnett, J. Fed-Batch Culture of Insect Cells: A Method to Increase the Yield of Recombinant Human Nerve Growth Factor (RhNGF) in the Baculovirus Expression System. *J. Biotechnol.* **1993**, *31* (2), 205−217.

(13) Liu, Y.-K.; Yang, C.-J.; Liu, C.-L.; Shen, C.-R.; Shiau, L.-D. Using a Fed-Batch Culture Strategy to Enhance RAAV Production in the Baculovirus/Insect Cell System. *J. Biosci. Bioeng* **2010**, *110* (2), 187−193.

(14) Saxena, A.; Byram, P. K.; Singh, S. K.; Chakraborty, J.; Murhammer, D.; Giri, L. A Structured Review of Baculovirus Infection Process: Integration of Mathematical Models and Biomolecular Information on Cell−Virus Interaction. *J. Gen. Virol* **2018**, *99* (9), 1151−1171.

(15) Yuan, C.; Yang, H. Research on K-Value Selection Method of K-Means Clustering Algorithm. *J. Multidiscip. Res.* **2019**, *2* (2), 226−235.

(16) Maria, G.; Crişan, M. Operation of a Mechanically Agitated Semi-Continuous Multi-Enzymatic Reactor by Using the Pareto-Optimal Multiple Front Method. *J. Process Control* **2017**, *53*, 95−105.

(17) Smets, I. Y.; Claes, J. E.; November, E. J.; Bastin, G. P.; Van Impe, J. F. Optimal Adaptive Control of (Bio)Chemical Reactors: Past, Present and Future. *J. Process Control* **2004**, *14* (7), 795−805.

(18) Bock, H. G.; Plitt, K. J. A Multiple Shooting Algorithm for Direct Solution of Optimal Control Problems. *IFAC Proc. Vol* **1984**, *17* (2), 1603−1608.

(19) Logist, F.; Van Erdeghem, P. M. M.; Smets, I. Y.; Van Impe, J. F. Optimal Design of Dispersive Tubular Reactors at Steady-State Using Optimal Control Theory. *J. Process Control* **2009**, *19* (7), 1191−1198.

(20) Wang, L.; Liu, X.; Zhang, Z. A New Sensitivity-Based Adaptive Control Vector Parameterization Approach for Dynamic Optimization of Bioprocesses. *Bioprocess Biosyst. Eng.* **2017**, *40* (2), 181−189.

(21) Freitas, H.; Olivo, J.; Andrade, C. Optimization of Bioethanol In Silico Production Process in a Fed-Batch Bioreactor Using Non-Linear Model Predictive Control and Evolutionary Computation Techniques. *Energies* **2017**, *10* (11), 1763.

(22) Ochoa, S. A New Approach for Finding Smooth Optimal Feeding Profiles in Fed-Batch Fermentations. *Biochem. Eng. J.* **2016**, *105*, 177−188.

(23) Nimmegeers, P.; Vallerio, M.; Telen, D.; Impe, J.; Logist, F. Interactive Multi-objective Dynamic Optimization of Bioreactors under Parametric Uncertainty. *Chemie Ing. Tech* **2018**, *91* (3), 349−362.

(24) Virivinti, N.; Mitra, K. Fuzzy Robust Optimization for Handling Feed Stream and Model Parameter Uncertainties during Comminution Process. *J. Taiwan Inst. Chem. Eng.* **2017**, *70*, 411−425.

(25) Sahinidis, N. V. Optimization under Uncertainty: State-of-the-Art and Opportunities. *Comput. Chem. Eng.* **2004**, *28* (6−7), 971−983.

(26) Zakaria, A.; Ismail, F. B.; Lipu, M. S. H.; Hannan, M. A. Uncertainty Models for Stochastic Optimization in Renewable Energy Applications. *Renew. Energy* **2020**, *145*, 1543−1571.

(27) Liu, M. L.; Sahinidis, N. V. Optimization in Process Planning under Uncertainty. *Ind. Eng. Chem. Res.* **1996**, *35* (11), 4154−4165.

(28) Virivinti, N.; Mitra, K. Intuitionistic Fuzzy Chance Constrained Programming for Handling Parametric Uncertainty: An Industrial Grinding Case Study. *Ind. Eng. Chem. Res.* **2015**, *54* (24), 6291−6304.

(29) Souza, D. L.; Lobato, F. S.; Gedraite, R. Robust Multiobjective Optimization Applied to Optimal Control Problems Using Differential Evolution. *Chem. Eng. Technol.* **2015**, *38* (4), 721−726.

(30) Zimmermann, H.-J. Fuzzy Programming and Linear Programming with Several Objective Functions. *Fuzzy Sets Syst* **1978**, *1* (1), 45−55.

(31) Liu, M. L.; Sahinidis, N. V. Process Planning in a Fuzzy Environment. *Eur. J. Oper. Res.* **1997**, *100* (1), 142−169.
(32) Mitra, K.; Gudi, R. D.; Patwardhan, S. C.; Sardar, G. Resiliency Issues in Integration of Scheduling and Control. *Ind. Eng. Chem. Res.* **2010**, *49* (1), 222−235.
(33) Sharma, S.; Pantula, P. D.; Miriyala, S. S.; Mitra, K. A Novel Data-Driven Sampling Strategy for Optimizing Industrial Grinding Operation under Uncertainty Using Chance Constrained Programming. *Powder Technol.* **2021**, *377*, 913−923.
(34) Vallerio, M.; Telen, D.; Cabianca, L.; Manenti, F.; Van Impe, J. F. M.; Logist, F. Robust Multi-Objective Dynamic Optimization of Chemical Processes Using the Sigma Point Method. *Chem. Eng. Sci.* **2016**, *140*, 201−216.
(35) Nimmegeers, P.; Telen, D.; Logist, F.; Van Impe, J. F. M. Dynamic Optimization of Biological Networks under Parametric Uncertainty. *BMC Syst. Biol.* **2016**, *10* (1), 86.
(36) Liu, Y.; Gunawan, R. Bioprocess Optimization under Uncertainty Using Ensemble Modeling. *J. Biotechnol.* **2017**, *244*, 34−44.
(37) Schillings, C.; Sunnåker, M.; Stelling, J.; Schwab, C. Efficient Characterization of Parametric Uncertainty of Complex (Bio)Chemical Networks. *PLOS Comput. Biol.* **2015**, *11* (8), No. e1004457.
(38) Logist, F.; Houska, B.; Diehl, M.; Van Impe, J. F. Robust Multi-Objective Optimal Control of Uncertain (Bio)Chemical Processes. *Chem. Eng. Sci.* **2011**, *66* (20), 4670−4682.
(39) Achcar, F.; Kerkhoven, E. J.; Bakker, B. M.; Barrett, M. P.; Breitling, R. Dynamic Modelling under Uncertainty: The Case of Trypanosoma Brucei Energy Metabolism. *PLoS Comput. Biol.* **2012**, *8* (1), No. e1002352.
(40) Ayres, M. D.; Howard, S. C.; Kuzio, J.; Lopez-Ferber, M.; Possee, R. D. The Complete DNA Sequence of Autographa Californica Nuclear Polyhedrosis Virus. *Virology* **1994**, *202* (2), 586−605.
(41) Giri, L.; Li, H.; Sandgren, D.; Feiss, M. G.; Roller, R.; Bonning, B. C.; Murhammer, D. W. Removal of Transposon Target Sites from the Autographa Californica Multiple Nucleopolyhedrovirus Fp25k Gene Delays, but Does Not Prevent, Accumulation of the Few Polyhedra Phenotype. *J. Gen. Virol* **2010**, *91* (12), 3053−3064.
(42) Jarvis, D. L.; Garcia, A. Long-Term Stability of Baculoviruses Stored under Various Conditions. *Biotechniques* **1994**, *16* (3), 508−513.
(43) Power, J. F.; Reid, S.; Radford, K. M.; Greenfield, P. F.; Nielsen, L. K. Modeling and Optimization of the Baculovirus Expression Vector System in Batch Suspension Culture. *Biotechnol. Bioeng.* **1994**, *44* (6), 710−719.
(44) Käßer, L.; Harnischfeger, J.; Salzig, D.; Czermak, P. The Effect of Different Insect Cell Culture Media on the Efficiency of Protein Production by Spodoptera Frugiperda Cells. *Electron. J. Biotechnol* **2022**, *56*, 54−64.
(45) Meneses-Acosta, A.; Mendonça, R. Z.; Merchant, H.; Covarrubias, L.; Ramírez, O. T. Comparative Characterization of Cell Death between Sf9 Insect Cells and Hybridoma Cultures. *Biotechnol. Bioeng.* **2001**, *72* (4), 441−457.
(46) Singh, R.; Miriyala, S. S.; Giri, L.; Mitra, K.; Kareenhalli, V. V. Identification of Unstructured Model for Subtilin Production through Bacillus Subtilis Using Hybrid Genetic Algorithm. *Process Biochem* **2017**, *60*, 1−12.
(47) Licari, P.; Bailey, J. E. Modeling the Population Dynamics of Baculovirus-infected Insect Cells: Optimizing Infection Strategies for Enhanced Recombinant Protein Yields. *Biotechnol. Bioeng.* **1992**, *39* (4), 432−441.
(48) Houska, B.; Logist, F.; Van Impe, J.; Diehl, M. Robust Optimization of Nonlinear Dynamic Systems with Application to a Jacketed Tubular Reactor. *J. Process Control* **2012**, *22* (6), 1152−1160.
(49) Marteijn, R. C. L.; Jurrius, O.; Dhont, J.; de Gooijer, C. D.; Tramper, J.; Martens, D. E. Optimization of a Feed Medium for Fed-Batch Culture of Insect Cells Using a Genetic Algorithm. *Biotechnol. Bioeng.* **2003**, *81* (3), 269−278.
(50) Carinhas, N.; Bernal, V.; Monteiro, F.; Carrondo, M. J. T.; Oliveira, R.; Alves, P. M. Improving Baculovirus Production at High Cell Density through Manipulation of Energy Metabolism. *Metab. Eng.* **2010**, *12* (1), 39−52.
(51) Mendonça, R. Z.; Palomares, L. A.; Ramírez, O. T. An Insight into Insect Cell Metabolism through Selective Nutrient Manipulation. *J. Biotechnol.* **1999**, *72* (1−2), 61−75.
(52) Nie, J.; Sun, Y.; Peng, F.; Li, X.; Yang, Y.; Liu, X.; Li, Y.; Liu, C.; Bai, Z. Production Process Development of Pseudorabies Virus Vaccine by Using a Novel Scale-Down Model of a Fixed-Bed Bioreactor. *J. Pharm. Sci.* **2020**, *109* (2), 959−965.
(53) Tapia, F.; Vázquez-Ramírez, D.; Genzel, Y.; Reichl, U. Bioreactors for High Cell Density and Continuous Multi-Stage Cultivations: Options for Process Intensification in Cell Culture-Based Viral Vaccine Production. *Appl. Microbiol. Biotechnol.* **2016**, *100* (5), 2121−2132.
(54) Maly, T.; Petzold, L. R. Numerical Methods and Software for Sensitivity Analysis of Differential-Algebraic Systems. *Appl. Numer. Math* **1996**, *20* (1−2), 57−79.
(55) Deb, K.*Multi-Objective Optimization Using Evolutionary Algorithms*; John Wiley and Sons: New York, 2001.
(56) Deb, K.Multi-Objective Optimisation Using Evolutionary Algorithms: An Introduction. In *Multi-Objective Evolutionary Optimisation for Product Design and Manufacturing*; Springer London: London, 2011; pp 3−34. DOI: 10.1007/978-0-85729-652-8_1.
(57) Tian, Y.; Cheng, R.; Zhang, X.; Jin, Y. PlatEMO: A MATLAB Platform for Evolutionary Multi-Objective Optimization [Educational Forum]. *IEEE Comput. Intell. Mag* **2017**, *12* (4), 73−87.
(58) Monteil, D. T.; Shen, X.; Tontodonati, G.; Baldi, L.; Hacker, D. L.; Wurm, F. M. Disposable Orbitally Shaken TubeSpin Bioreactor 600 for Sf9 Cell Cultivation in Suspension. *Anal. Biochem.* **2016**, *505*, 26−28.
(59) Puente-Massaguer, E.; Gòdia, F.; Lecina, M. Development of a Non-Viral Platform for Rapid Virus-like Particle Production in Sf9 Cells. *J. Biotechnol.* **2020**, *322*, 43−53.
(60) Power, J.; Greenfield, P. F.; Nielsen, L.; Reid, S. Modelling the Growth and Protein Production by Insect Cells Following Infection by a Recombinant Baculovirus in Suspension Culture. *Cytotechnology* **1992**, *9* (1−3), 149−155.
(61) Licari, P.; Bailey, J. E. Modeling the Population Dynamics of Baculovirus-Infected Insect Cells: Optimizing Infection Strategies for Enhanced Recombinant Protein Yields. *Biotechnol. Bioeng.* **1992**, *39* (4), 432−441.
(62) Zhang, Y.-H.; Merchuk, J. C. A Mathematical Model of Baculovirus Infection on Insect Cells at Low Multiplicity of Infection. *Acta Biochim. Biophys. Sin. (Shanghai)* **2004**, *36* (11), 729−740.
(63) Chan, L. C. L.; Young, P. R.; Bletchly, C.; Reid, S. Production of the Baculovirus-Expressed Dengue Virus Glycoprotein NS1 Can Be Improved Dramatically with Optimised Regimes for Fed-Batch Cultures and the Addition of the Insect Moulting Hormone, 20-Hydroxyecdysone. *J. Virol. Methods* **2002**, *105* (1), 87−98.
(64) Chan, L. C. L.; Greenfield, P. F.; Reid, S. Optimising Fed-Batch Production of Recombinant Proteins Using the Baculovirus Expression Vector System. *Biotechnol. Bioeng.* **1998**, *59* (2), 178−188.
(65) Elias, C. B.; Zeiser, A.; Bédard, C.; Kamen, A. A. Enhanced Growth of Sf-9 Cells to a Maximum Density of 5.2 × 107 Cells per ML and Production of β-Galactosidase at High Cell Density by Fed Batch Culture. *Biotechnol. Bioeng.* **2000**, *68* (4), 381−388.
(66) Meghrous, J.; Mahmoud, W.; Jacob, D.; Chubet, R.; Cox, M.; Kamen, A. A. Development of a Simple and High-Yielding Fed-Batch Process for the Production of Influenza Vaccines. *Vaccine* **2009**, *28* (2), 309−316.
(67) Felberbaum, R. S. The Baculovirus Expression Vector System: A Commercial Manufacturing Platform for Viral Vaccines and Gene Therapy Vectors. *Biotechnol. J.* **2015**, *10* (5), 702−714.
(68) Roldão, A.; Vieira, H. L. A.; Charpilienne, A.; Poncet, D.; Roy, P.; Carrondo, M. J. T.; Alves, P. M.; Oliveira, R. Modeling Rotavirus-like Particles Production in a Baculovirus Expression Vector System: Infection Kinetics, Baculovirus DNA Replication, MRNA Synthesis and Protein Production. *J. Biotechnol.* **2007**, *128* (4), 875−894.

(69) Wang, M. Y.; Bentley, W. E. Continuous Insect Cell (Sf-9) Culture with Aeration through Sparging. *Appl. Microbiol. Biotechnol.* **1994**, *41* (3), 317−323.
(70) Mishra, M.; Tiwari, S.; Gomes, A. V. Protein Purification and Analysis: Next Generation Western Blotting Techniques. *Expert Rev. Proteomics* **2017**, *14* (11), 1037−1053.

https://doi.org/10.1021/acs.iecr.2c03355
Ind. Eng. Chem. Res. 2023, 62, 111−125