

Process Model Comparison and Transferability Across Bioreactor Scales and Modes of Operation for a Mammalian Cell Bioprocess

Stephen Craven, Nishikant Shirsat, Jessica Whelan, and Brian Glennon

School of Chemical and Bioprocess Engineering, University College Dublin, Belfield, Dublin 4, Ireland

DOI 10.1002/btpr.1664

Published online December 20, 2012 in Wiley Online Library (wileyonlinelibrary.com).

A Monod kinetic model, logistic equation model, and statistical regression model were developed for a Chinese hamster ovary cell bioprocess operated under three different modes of operation (batch, bolus fed-batch, and continuous fed-batch) and grown on two different bioreactor scales (3 L bench-top and 15 L pilot-scale). The Monod kinetic model was developed for all modes of operation under study and predicted cell density, glucose, glutamine, lactate, and ammonia concentrations well for the bioprocess. However, it was computationally demanding due to the large number of parameters necessary to produce a good model fit. The transferability of the Monod kinetic model structure and parameter set across bioreactor scales and modes of operation was investigated and a parameter sensitivity analysis performed. The experimentally determined parameters had the greatest influence on model performance. They changed with scale and mode of operation, but were easily calculated. The remaining parameters, which were fitted using a differential evolutionary algorithm, were not as crucial. Logistic equation and statistical regression models were investigated as alternatives to the Monod kinetic model. They were less computationally intensive to develop due to the absence of a large parameter set. However, modeling of the nutrient and metabolite concentrations proved to be troublesome due to the logistic equation model structure and the inability of both models to incorporate a feed. The complexity, computational load, and effort required for model development has to be balanced with the necessary level of model sophistication when choosing which model type to develop for a particular application. © 2012 American Institute of Chemical Engineers *Biotechnol. Prog.*, 29: 186–196, 2013

Keywords: mathematical models, statistical models, mode of operation and scale-up, model transferability, parameter sensitivity

Introduction

The biopharmaceutical industry is rapidly expanding and to date, a large proportion of the products are produced in mammalian cell cultures, which demand a high level of process management. The number of biopharmaceuticals currently on the market is just in excess of 200 and in 2009, they generated \$99 billion in sales.¹ In 2004, the FDA replaced the “Quality-by-Inspection” methodology by “Quality by Design” (QbD). Since then QbD has been playing a crucial role in assisting both the industry and the FDA to move toward a more scientific and proactive approach to the development of the biopharmaceutical industry. The application of QbD in biopharmaceuticals has received a lot of interest in the literature recently.^{2–5} QbD demands a complete understanding of the process and the identification of the critical process parameters, which affect the critical quality attributes (CQAs) so that the process can be operated in

an intelligent way compensating for the inherent variability in order to assure consistent product quality.

Process models form a core part of the QbD framework, playing a role in process understanding, process development, process and automation design, and online diagnostics. Many types of models exist. Choice of model type depends on the intended application and also on the quantity, quality, and nature of experimental data available. Process models can be classified as qualitative, mathematical, and statistical models (Figure 1a). Mathematical models can be further subdivided into mechanistic and empirical (black-box) models or categorized by the level of structure and segregation considered (Figure 1b).

One of the most common mathematical models used for bioprocesses is the mechanistic Monod kinetics model. It is an unstructured, nonsegregated, mechanistic model. There are a number of examples of this type of model presented in the literature.^{6–12} The specific growth and death rates are generally modeled using Monod-type functions based on nutrient and metabolite concentrations.

Unstructured models do not account for the intracellular processes and acknowledge only implicitly the change of cellular physiology with the environment.¹³ They are based

Additional Supporting Information may be found in the online version of this article.

Correspondence concerning this article should be addressed to B. Glennon at brian.glennon@ucd.ie.

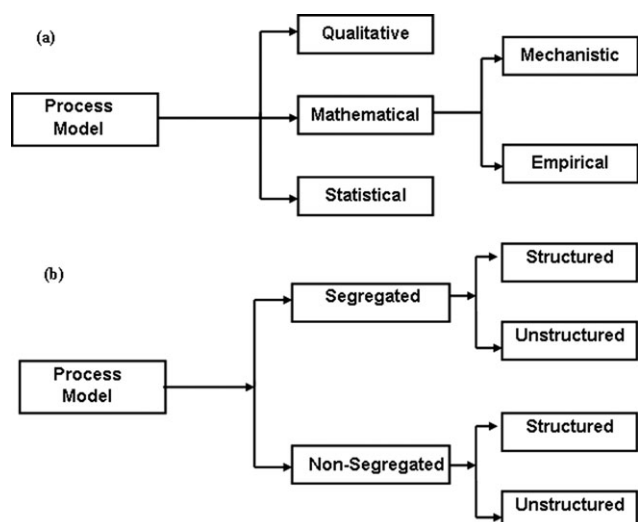


Figure 1. (a) Model classification for cell culture systems and (b) model classification based on the level of structure and segregation.

on the fundamental assumption of balanced growth. This common assumption states that during cell growth, the internal metabolites are at a quasi-steady state, and their net conversion rate equals zero.¹⁴ Such models treat the cell as a single homogeneous unit, hence their biological significance is limited, and the mathematical equations involved are only phenomenological descriptions of the actual system. Validity holds only for batch growth during the exponential growth stage and also for continuous culture during steady-state operation. Portner and Schafer¹⁵ conducted a survey of unstructured models for hybridoma cell growth and metabolite production, which revealed that even though unstructured models contain limited process mechanistic information compared with structured models, they can still serve as an effective tool for process design.

Structured models, on the other hand, are more complex than unstructured models because they attempt to incorporate biological knowledge by separating the biomaterial into compartments that are chemically and physically distinct. The compartments' interactions with each other and with the environment are described by stoichiometric equations that account for various metabolic pathways and/or kinetic rate expressions. For example, Batt and Kompala¹⁶ developed a structured kinetic model, which distinguishes four compartments inside the cell and considers the major substrates (glucose and amino acids) and metabolites (lactate and ammonia). The development of such models is time consuming and requires a large knowledge base.

Not all cells in a mammalian cell culture are alike; they are a heterogeneous population. Process models may or may not take this heterogeneity into account. A nonsegregated model views the population as consisting of identical "average cells" and uses a lumped variable such as total biomass per unit volume, to describe the entire population. Segregated models, in contrast, may take the cell cycle into consideration or distinguish different cell states.^{17,18} Differentiating groups of cells within the population with similar characteristics is more representative of the true physiological state. Therefore, segregated models can account for the degree of heterogeneity of the cell culture with regards to cell age, size, growth rate, and metabolic state. However,

segregated models are more computationally difficult to handle.¹⁹

The aforementioned QbD approach demands a high-level of process understanding for ensuring control of CQAs. Levels of process understanding and knowledge contained within the process models can be categorized from lowest to highest as unstructured and nonsegregated to structured and segregated (Figure 1b). While structured, segregated process models provide the greatest level of understanding and prediction capability, the large effort required for their development and use within biopharmaceutical industry may be too expensive due to the complex nature of most bioprocesses.

The accurate estimation of the model parameters is essential for successful model development. Although mechanistic Monod kinetic models can adequately describe experimental data, they are computationally intensive to fit due to the presence of a large number of kinetic parameters. Because of the challenges associated with developing the classical Monod kinetic model, alternatives have been developed such as semi-empirical logistic equation and empirical statistical regression models, which have been successfully used in a variety of applications to describe population dynamics.^{20,21}

The first logistic equation was developed by Malthus²² to describe population growth, which was subsequently modified by Verhulst²³ to account for finite resources for growth. Verhulst²³ proposed the concept of a stable maximum population, K , called the overall saturation constant (cell density/unit volume), which forms an upper limit for the size of the population. He introduced a logistic growth model, a modification of the exponential model, to incorporate the overall saturation constant K . Furthermore, the intrinsic growth rate constant, r acquires the identity of the intrinsic or relative growth rate $[(1/X_V)dX_V/dt]$, where, X_V , is the viable cell density during the bioprocess and t is time. Thus, both K and r are experimentally determinable parameters. Logistic equation models have been successfully applied to describe population dynamics in a variety of applications. Most reported applications involve bacterial growth curves characterized by lag, log, and stationary phases.^{20,24} Mammalian cell bioprocesses operated in batch and fed-batch exhibit a sharp decline in cell density, a behavior that cannot be described by the standard logistic growth equation. Therefore, alternative logistic equations have been developed to account for both cell growth and death.²⁵ The main difference between the logistic equation model and the Monod kinetic model is that the logistic equation model aims to describe the kinetics of the bioprocess with as few parameters as possible, while still being able to accurately describe the distinct stages of growth. However, the logistic equation model has no extrapolation capabilities.

The complexities inherent in biological systems can make them difficult to model satisfactorily with logistic equations and Monod kinetic models, whereas, statistical methods can perform well where the mechanism of the underlying biochemical reactions is too complex to resolve. Regression analysis is frequently used for determining functional relationships between discrete observations of categorical and response variables.²⁶ Ergun and Mutlu²⁷ applied statistical modeling techniques to describe the rate of production of ethanol from sugar beet molasses by *Saccharomyces cerevisiae*.

Bioreactor mode of operation is an important consideration when modeling bioprocesses. Mammalian cell

bioprocesses are usually operated in fed-batch mode in industrial applications due to the increase in product titer that results from the control it allows over nutrient concentrations. Feed delivery in a fed-batch bioprocess may be manual via the bolus addition of a feed medium or be automated based on an open or closed loop control strategy.

In this research, both automated continuous and manual bolus-feeding strategies were studied. Monod kinetic models were built to model all modes of operation (batch, bolus feeding, and continuous feeding fed-batch) and the transferability of the model structures and kinetic parameters across bioreactor scales and modes of operation was investigated. The applicability of both a semi-empirical logistic equation model and an empirical statistical regression model was also considered as modeling alternatives to the mechanistic Monod kinetic model due to the potentially lower effort required for development.

Materials and Methods

Cell culture

The cell line used in this study was a Chinese Hamster Ovary (CHO) cell line, CHO 320, which produced interferon- γ . The cells were grown in glucose-free Ex-cell serum-free medium (Sigma-Aldrich) which was supplemented with glutamine and glucose to 4 and 20 mM, respectively. Pluronic F68 (0.1% v/v), 1 μ M methotrexate, and antifoam C (5 ppm) were also added. The feed consisted of glucose (653 mM), glutamine (58.8 mM), and soy hydrolysate powder (58.8 g/L) (Irvine Scientific) dissolved in glucose-free Ex-cell serum-free medium.

Stirred tank bioreactor cultures

Experiments were conducted in both 3 L bench-top and 15 L pilot scale bioreactors (Applikon Biotechnology, Netherlands). The pH, dissolved oxygen, temperature, and agitation were controlled at constant values of 7.2, 50% of air saturation, 37°C, and 120 rpm, respectively. The substrates (glucose and glutamine) and by-products (lactate and ammonia) were measured online with Raman spectroscopy using a Kaiser Rxn2 system (Kaiser Optical Systems). Samples were taken daily at a minimum and the cell density and viability were measured using the trypan blue dye exclusion method.

Feeding regime

Two fed-batch feeding regimes were modeled in this study. The first was a fed-batch culture manually fed with bolus additions, at 24-h intervals, the volume of which was proportional to off-line integral viable cell density (IVC) measurements for the previous 24-h interval, calculated using Eq. 1.

$$IVC = (X_{V_{i+1}} + X_{V_i})(t_{i+1} - t_i) \quad (1)$$

The second regime was a continuous feed, the rate of which was determined and adjusted automatically using a model predictive controller and Raman-determined glucose values to keep the glucose concentration in the bioreactor at a set-point of 11 mM throughout the culture.

Process models

First Principle Engineering Model (Monod-Type Kinetics). The first principle engineering model was a

mechanistic mathematical model, which described the cell growth and the cell metabolism in an unstructured, nonsegregated manner. It was based on certain standard assumptions, such as a well-mixed bioreactor and perfect control of culture pH, temperature, and dissolved oxygen concentration. The balanced growth assumption was also adhered to, where the internal metabolites, c , are considered to be at quasi-steady state during cell growth (Eq. 2).

$$\frac{dc}{dt} \cong 0 \quad (2)$$

The model consisted of a number of first-order ordinary differential equations representing the rate of change of state variables of the process and was solved by numerical integration using an ordinary differential equation solver (ODE23s) in MATLAB (Version R2010b). The state variables included the total cell density (X_T), viable cell density (X_V), dead cell density (X_D), glucose concentration (G), glutamine concentration (Q), lactate concentration (L), and ammonia concentration (A). The model detailed below was developed for the batch mode of operation and then extended for both fed-batch scenarios (manual bolus feeds and closed loop continuous feeding). The model parameters are described in the nomenclature.

The rates of change of the total (X_T), viable (X_V), and dead (X_D) cell density are represented in Eqs. 3, 4, and 5, respectively.

$$\frac{dX_T}{dt} = \mu X_V - K_{LYSIS} X_D \quad (3)$$

$$\frac{dX_V}{dt} = (\mu - k_d) X_V \quad (4)$$

$$\frac{dX_D}{dt} = k_d X_V - K_{LYSIS} X_D \quad (5)$$

The specific growth rate, μ , depends on the concentration of the substrate, S . Many cell culture processes exhibit saturation type kinetics, i.e., the rate of the growth μ is limited by a certain factor when its concentration, S , is low, but the limiting effect disappears and the growth rate reaches a maximum value, μ_{max} , as the concentration increases. This behavior can be described by a Monod-type equation (Eq. 6), which is the most commonly used expression that relates the specific growth rate of the cell to the substrate concentration.²⁸

$$\mu = \mu_{max} \frac{S}{K_S + S} \quad (6)$$

It is plausible that instead of one metabolite being the controlling factor, there may be a group of metabolites that control the metabolic processes within the culture as in the present case. In such a situation, Zeng et al.¹² has shown that Eq. 6 can be modified to incorporate additional terms for substrates (glucose and glutamine) and inhibitory byproducts (lactate and ammonia) as shown in Eq. 7. However, this form of the Monod equation can lead to errors as shown by Bader.²⁹ However, with only four terms this multiplicative model was seen to be acceptable.

$$\mu = \mu_{max} \frac{G}{K_G + G} \frac{Q}{K_Q + Q} \frac{K_L}{K_L + L} \frac{K_A}{K_A + A} \quad (7)$$

The cell death rate, k_d , was modeled using the kinetic expression in Eq. 8,^{11,12} where the parameter $k_{d,max}$ is the maximum death rate, and k_u is the intrinsic death rate.

$$k_d = k_{d,max} \left(\frac{k_u}{\mu + k_u} \right) \quad (8)$$

The rate of consumption of glucose was represented by Eq. 9. The consumption of glucose for the maintenance of viable cells was accounted for by the maintenance coefficient m_G .

$$\frac{dG}{dt} = \left(-\frac{\mu}{Y_{X,G}} - m_G \right) X_V \quad (9)$$

The rate of glutamine consumption is modeled in a similar fashion with the exception of the inclusion of a first-order degradation coefficient, $k_{d,Q}$, which represents the spontaneous degradation of glutamine under reactor conditions.

$$\frac{dQ}{dt} = \left(-\frac{\mu}{Y_{X,Q}} - m_Q \right) X_V - k_{d,Q} Q \quad (10)$$

Equation 11 represents lactate production, which is assumed to be due to the consumption of glucose.

$$\frac{dL}{dt} = -Y_{L,G} \left(-\frac{\mu}{Y_{X,G}} - m_G \right) X_V \quad (11)$$

and finally, Eq. 12 shows the rate of production of ammonia due to the consumption and degradation of glutamine in the reactor.

$$\frac{dA}{dt} = -Y_{A,Q} \left(-\frac{\mu}{Y_{X,Q}} - m_Q \right) X_V + k_{d,Q} Q \quad (12)$$

The model parameters were determined either experimentally or fitted using the differential evolution (DE) technique with constraints set based on a literature survey. The parameters determined experimentally included; $Y_{A,Q}$, $Y_{L,G}$, $Y_{X,G}$, $Y_{X,Q}$, $k_{d,max}$, and μ_{max} . Based on the experimental data, the maximum growth rate, μ_{max} , and the maximum death rate, $k_{d,max}$, were calculated using differential analysis methods,³⁰ and the yield ratios ($Y_{A,Q}$, $Y_{L,G}$, $Y_{X,G}$, and $Y_{X,Q}$) were determined by performing a linear regression on the plots of A against Q , L against G , X_V against G , and X_V against Q , respectively. The fitted parameters were determined using the DE algorithm. DE is a very simple but effective population-based stochastic function minimizer based on minimizing the sum of the squares of the error (SSE) between the experimental data and model predictions.³¹ In the present work, the DE MATLAB source code by Buehren³² was used.

Logistic Model for Growth. A semi-empirical logistic equation model was used to predict viable cell density over the time course of the batch and was estimated with a statistical method. Statistical methods are used to correlate random variables with the response variables. The random variables are normally considered as noninterconnected, discrete observations of a scalar nature. However, in the dynamics of cell proliferation, the random observations are vectors because they are all biologically interconnected, qualitatively and quantitatively. Hence, each observation of viable cell number reflects the growth status of the cells in the previous observation and together, they represent the dynamics of the cell growth as expressed by the least square

equation, called the parametric equation, of the trend line. This conversion of discrete observations into state space implies that the state variables are continuous, at least in the region of interest, in the present case, the exponential phase. The cells are known for their predictable behavior in the exponential phase, thus increasing the probability of the reproducibility of observations. Therefore, the variations in the cell density with time are estimated from the parametric equation of the trend line and the data was used for modeling the cell growth density. Thus, the estimation of stochastic models by time course data makes it possible to describe biological dynamics in a quantitative framework.²¹

The logistic model used is of the form:

$$\frac{dX_V}{dt} = rX_V \left(1 - \frac{X_{V0}}{K} \right) \quad (13)$$

where, r (h^{-1}) is the intrinsic growth rate constant when the effects of the limiting factors are negligible, X_{V0} (cells L^{-1}), viable cell density at $t = 0$ and K (cells L^{-1}), overall growth saturation constant. Thus, r can be thought of as the rate constant for potential exponential growth of the population. The logistic equation (Eq. 13) introduces the multiplicative factor, $1 - (X_{V0}/K)$ to account for a saturation level characteristic (K) of the environment that will give a stable population level. The multiplicative factor slows growth as K is approached. When the population size equals K , then from Eq. 13, the growth rate becomes zero and the population remains at that level.

Solving Eq. 13 by separable variables yields

$$X_V(t) = \frac{K}{1 + \left(\frac{K - X_{V0}}{X_{V0}} \right) e^{-rt}} \quad (14)$$

Equation 14 is called the logistic equation, describing the exponential growth rate with undefined limitations.³³ However, Eq. 14 describes the exponential growth phase only. The cell growth and decline phases indicate symmetry at the point of maximum growth. Therefore, the rate of decline in the vicinity of the peak of the cell density is the same as the rate of growth. Hence, the decline can be described by replacing the intrinsic growth rate constant from $+r$ to $-r$. Thus, Eq. 14 becomes

$$X_V(t) = \frac{K}{1 + \left(\frac{K - X_{Vmax}}{X_{Vmax}} \right) e^{r(t-t_m)}} \quad (15)$$

X_{Vmax} is the model calculated maximum viable cell density occurring at time t_m , which can be determined from the experimental data. The remaining model parameters, K and r , were determined by plotting the intrinsic growth rate, $(1/Ln(X_V))(Ln(X_V)/dt)$ vs. the natural log of the viable cell density, $Ln(X_V)$. A linear trend line was fitted to the data in Excel with K (cells L^{-1}), the intercept at the x -axis and r (h^{-1}), the intercept on the y -axis. Therefore, all three model parameters (K , r , and t_m) are experimentally determinable.

Statistical Regression Model. A multivariate regression method was used to develop generalized linear models for predicting viable cell density from the concentrations of glucose, glutamine, lactate, and ammonia in both batch and fed-batch cultures. Regression analysis provides a systematic technique for estimating, with confidence limits, the unspecified regression coefficients from a new set of data, or for

Table 1. Monod-Type Model Parameters and Values in Brackets Show the Error at the 95% Confidence Interval for the 3 L Batch Mode-Fitted Parameters

Parameters	Unit	3 L Batch	3 L Bolus Fed-batch	3 L Continuous Fed-batch	15 L Continuous Fed-batch
$k_{d,Q}^{\dagger}$	h^{-1}	0.001 (9.6×10^{-5})	0.001	0.001	0.001
m_G^{\dagger}	mmol cell $^{-1}h^{-1}$	8.0×10^{-13} (1.1×10^{-14})	9.5×10^{-11}	1.1×10^{-10}	1.4×10^{-10}
$Y_{A,Q}^{\ddagger}$	—	0.68	0.80	0.90	0.90
$Y_{L,G}^{\ddagger}$	—	1.6	2.0	2.0	2.0
$Y_{X,G}^{\ddagger}$	cells mmol $^{-1}$	9.23×10^7	1.60×10^8	2.20×10^8	1.70×10^8
$Y_{X,Q}^{\ddagger}$	cells mmol $^{-1}$	8.8×10^8	1.5×10^9	1.5×10^9	1.3×10^9
K_L^{\dagger}	mM	150 (1.16)	150	150	150
K_A^{\dagger}	mM	40 (1.38)	40	40	40
$k_{d,max}^{\ddagger}$	h^{-1}	0.01	0.01	0.01	0.01
μ_{max}^{\ddagger}	h^{-1}	0.035	0.044	0.044	0.048
K_G^{\dagger}	mM	1.0 (0.018)	1.0	1.0	1.0
K_Q^{\dagger}	mM	0.22 (0.003)	0.22	0.22	0.22
m_Q^{\dagger}	mmol cell $^{-1}h^{-1}$	3.0×10^{-12} (1.4×10^{-13})	5.0×10^{-12}	0	7.0×10^{-12}
k_{μ}^{\dagger}	h^{-1}	0.01 (1.0×10^{-4})	0.01	0.01	0.01
K_{LYSIS}^{\dagger}	h^{-1}	4.0×10^{-2} (0.004)	0	2.0×10^{-2}	0
$C_i^{*\dagger}$	mM	—	100	100	100
q_i^{\dagger}	h^{-1}	—	3.5×10^{-10}	2.5×10^{-10}	3.5×10^{-10}

† Fitted model parameter.

‡ Experimentally determined model parameter.

testing whether the new data are consistent with the hypothesis.

The generalized linear models are of the form:

$$y = b_0 + b_1x_{i1} + \dots + b_nx_{ik} + \varepsilon_i \quad (16)$$

where b_i is the partial regression coefficients of each independent variable and ε_i is the residual difference between the observed and estimated y . When $k = 1$, i.e., only one independent variable, then the model is called a simple regression model. When $k > 1$, the model is described as a multiple regression model.

The method of least squares is used to estimate the partial regression coefficients, b_i . This is done by minimizing the squared value of the residual differences i.e. ε_i , using the regression function within the data analysis tool in Excel.

Results and Discussion

A first principle engineering model, based on Monod type kinetics, was developed and its transferability to three modes of operation (batch, bolus fed-batch, and continuous fed-batch), and two scales (3 L bench-top and 15 L pilot scale) was investigated. The performance, transferability, and development effort of the mechanistic Monod-type kinetic model was compared with a semi-empirical logistic equation modeling and empirical statistical regression modeling approach. The experimental data used was an average of two bioreactor runs.

Batch Model Development (Monod-Type Kinetics). The model parameters were determined either experimentally or fitted using the DE technique with constraints set based on a literature survey. The parameters determined experimentally were $Y_{A,Q}$, $Y_{L,G}$, $Y_{X,G}$, $Y_{X,Q}$, $k_{d,max}$, and μ_{max} , and the remaining parameters were determined by using the DE technique where the lowest sum of the squares was sought between the model predicted values and the experimental data. A 95 % confidence interval was used for these fitted model parameters. The values of the parameters were initially optimized by fitting the model to the experimental results (average of two batches) obtained from the bench-top 3 L batch mode bioreactor runs. Kontoravdi et al.³⁴ similarly used batch experiments for providing initial model parameter estimates. Table 1 details the optimized model parameter sets for all modes of operation and scales considered. A satisfactory agreement was

observed between the 3 L batch model predicted values and experimental data (Figures 2a,e,i).

Figure 2i shows the depletion of glucose at 120 h, which resulted in an increase in the rate of cell death. The glutamine consumption rate and thus the ammonia production rate also respond to this change in growth phase and the associated levels of metabolism by decreasing after 120 h (Figure 2e). The total cell density (X_T) begins to decline at this time because of the onset of cell lysis in the 3 L bioreactor (Figure 2a). This behavior has been incorporated into the unstructured model where the rate of cell lysis is represented by K_{LYSIS} (Eq. 2).

A parameter sensitivity analysis was conducted on the fifteen experimental and fitted parameters of the 3 L batch process model to determine which parameters are most important for model optimization. Each parameter was varied by $\pm 25\%$ of its absolute value individually, and the effect on viable cell density was quantified by calculating the percentage difference in the SSE between the experimental data and the model predicted values. The viable cell density was chosen as the basis for evaluation because it is the key variable, affecting all other variables (Figure 3).

The viable cell density was most sensitive to the maximum specific growth rate, μ_{max} , and the yield of biomass from glucose, $Y_{X,G}$. Interestingly, an antisymmetry was noticed between the nutrient and metabolite Monod saturation constants due to the nature of the metabolism. When the metabolite saturation constants (K_L , K_A) had a positive percentage change in the SSE, the nutrient saturation constants (K_G , K_Q) had a negative percentage change. However, the viable cell density was relatively insensitive to these Monod saturation constants. Overall, it is found that the viable cell density is very sensitive to the majority of the experimentally determined parameters (yields and rates) and not as much for the nonexperimental, fitted parameters. This information is very useful when fitting the model to the experimental data and its applicability is seen in the parameter transferability analysis between bioreactor scales and between modes of operation.

Transferability of the 3 L batch model to 3 L fed-batch operation

Bolus Fed-Batch. Initially, during the fed-batch process development, a manual bolus-feeding regime at 24-h

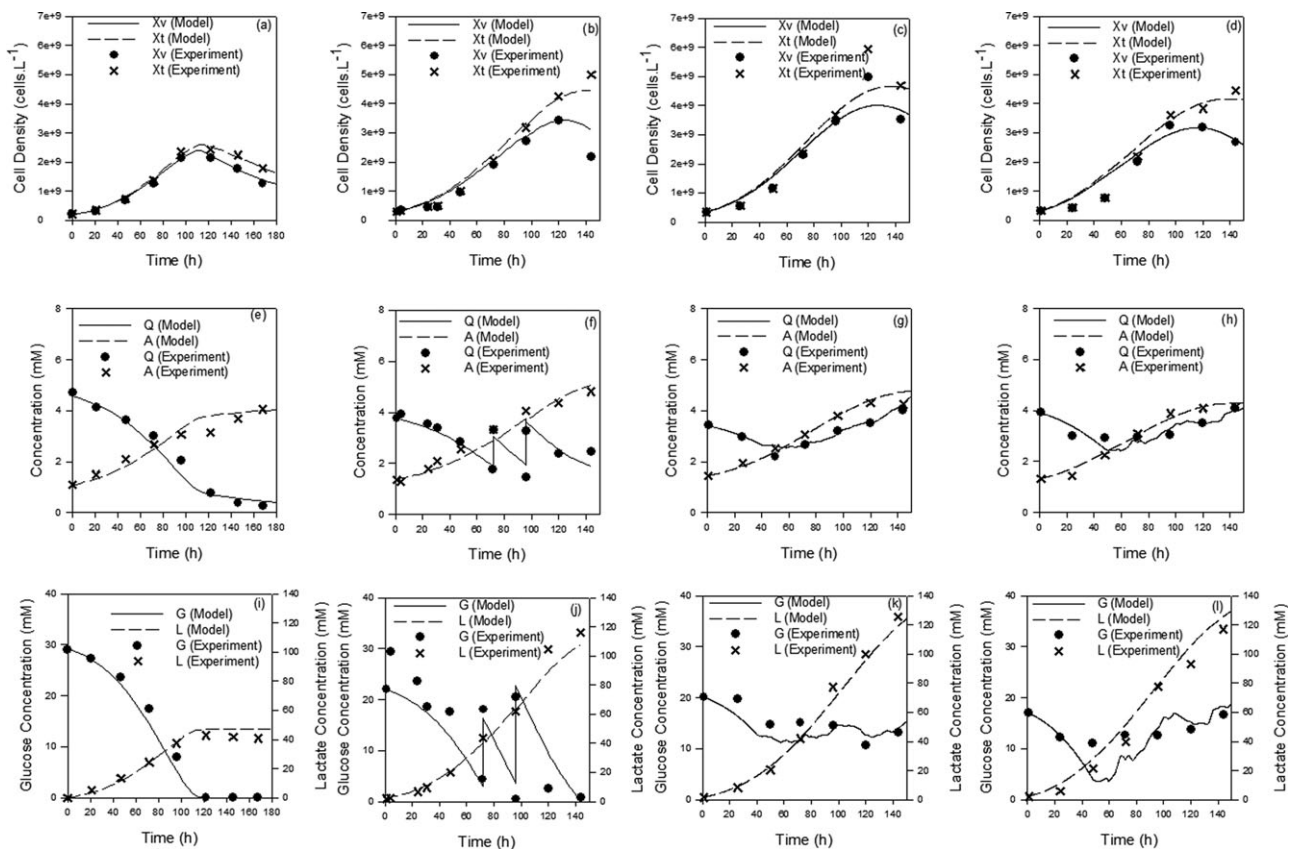


Figure 2. Monod-type kinetic model predictions for (a, b, c, and d) cell density, (e, f, g, and h) glutamine and ammonia, and (i, j, k, and l) glucose and lactate concentrations in a 3 L batch, bolus, and continuous fed-batch bioreactor and a 15 L continuous fed-batch bioreactor, respectively.

intervals was used. Variable volumes of feed determined by evaluating the IVC were added manually once daily.

The 3 L batch model structure was adapted to incorporate the bolus-feeding regime. The 3 L bolus fed-batch experimental data indicated that an additional inhibitory term in Eq. 5 was needed to account for the cessation of growth in the 3 L fed-batch process. This alteration to the 3 L batch is usual. A number of studies^{35–37} has shown that an inhibitory factor other than ammonia and lactate was responsible for the limitation of growth and for the transition to death in both hybridoma and CHO cultures. The inhibitory species has been shown to be dialyzable and therefore, of low molecular weight, less than 5000 Daltons. This inhibitory factor was successfully incorporated into the mathematical 3 L fed-batch model (Supporting Information appendix 1). The experimentally determinable parameters for the 3 L bolus fed-batch model were calculated and the transferability of the remaining fitted parameters from batch mode to bolus fed-batch mode was examined (Table 1). It is evident from Table 1 that the experimentally determined yields and rates differ from those calculated for the batch mode of operation. This is due to the increased availability of nutrients for growth in fed-batch mode. The majority of the fitted parameters remained unchanged from those for the batch mode of operation with the exception of the maintenance coefficients for both glucose and glutamine and the rate of cell lysis. Cell lysis did not occur to any great extent in this fed-batch process. However, all the Monod saturation constants fitted for the 3 L batch bioreactor were found to be applicable when the mode of operation was changed to bolus feed additions.

The results predicted by the adapted model for the 3 L bolus feed fed-batch process are presented in Figures 2b,f,j. The bolus feeding strategy yielded higher cell densities. These higher cell densities can be attributed to increased nutrient availability due to feeding. However, the bolus-feeding regime led to increased levels of lactate and ammonia when compared with the 3 L batch culture as is expected due to the greater total consumption of glucose and glutamine.

Continuous Closed Loop Fed-Batch. The bioreactor was also operated in continuous closed loop fed-batch mode and the resulting experimental continuous feed profile was inputted into the model. Real-time nutrient and metabolite concentration measurements from in-situ Raman spectroscopy was used to close the loop.

The resulting model structure developed for this mode of operation is shown in Supporting Information Appendix 1 and does not differ in structure from that developed for the bolus feed additions. Apart from the maintenance coefficients, rate of the cell lysis, and the specific production rate of the inhibitor, the fitted parameters (Table 1) did not alter, indicating good transferability between the three modes of operation.

The zero glutamine maintenance coefficient, m_Q , allows the model to account for the surplus in glutamine concentration (Figure 2g) as a result of a reduced glutamine consumption rate toward the end of the fed batch process. This behavior was also witnessed by Wong et al.³⁸ A higher glucose maintenance coefficient, m_G , was used within the model to account for the greater glucose consumption rates. The reduction in the effect of the inhibitor concentration on cell growth allows the model to predict the higher cell densities experienced due

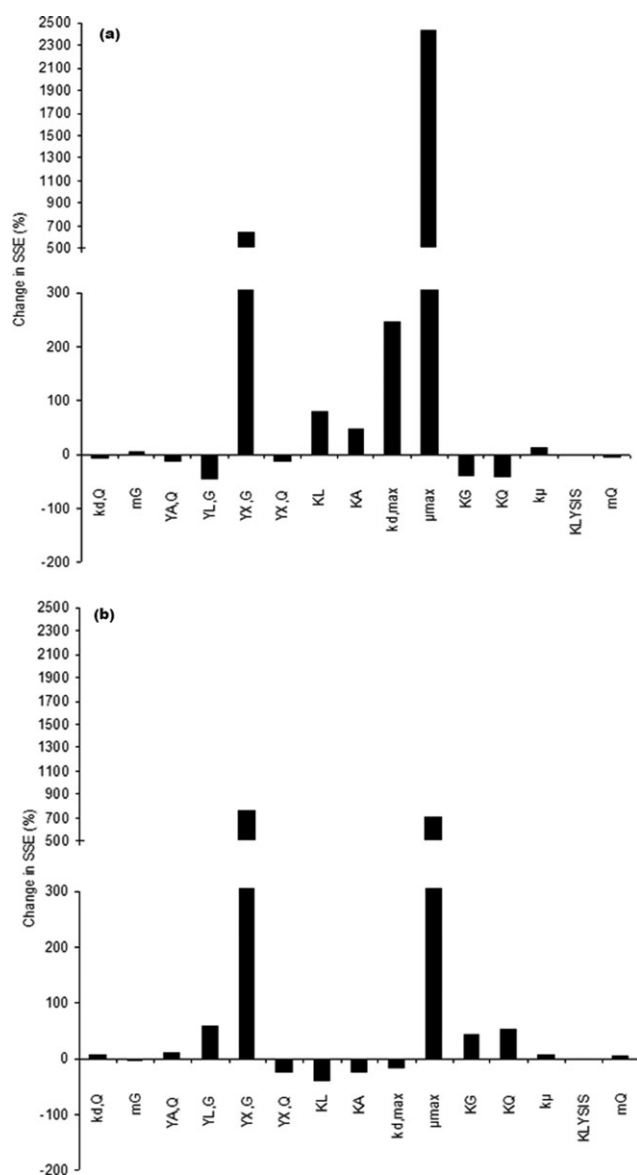


Figure 3. Parameter sensitivity analysis for a 25% parameter (a) increase and (b) decrease.

to closed loop continuous feeding (Figure 2c). Such improvements can be attributed to the more constant concentrations of substrates in the reactor, which prevented the short periods of nutrient depletion that were seen on occasion just prior to the bolus feed additions. Lactate and ammonia production rates change when the glucose concentration was limited to 11 mM over the course of the continuous feed fed-batch process. However, Wong et al.³⁸ showed that the specific ammonia and lactate production can be reduced by controlling to lower glucose/glutamine levels (~0.3 mM).

Overall, it has been demonstrated that once the experimentally determinable parameters are calculated for each mode of operation, the parameter estimation can be reduced to a small subset of model parameters as the majority of the fitted parameters showed good transferability across the different modes of bioreactor operation. It should also be noted that the majority of the experimentally determined parameters showed little or no variation between the bolus and continuous fed-batch feeding strategies. Portner and Schafer¹⁵ simi-

Table 2. Logistic Equation Model Parameters

	Unit	3 L Batch	3 L Bolus Fed-Batch	3 L Continuous Fed-Batch
K	cells L ⁻¹	21.60	22.31	22.71
r	h ⁻¹	0.0184	0.0152	0.0148
t_m	h	122	119	119

larly saw that an unstructured model developed from batch culture data can also describe fed-batch cultures with minor model modifications.

Transferability of model across bioreactor scales

15 L Pilot Scale Continuous Closed Loop Fed-Batch. The application of the model to process scale up was tested by scaling the bioprocess up from a 3 L bench-top scale bioreactor to a 15 L pilot scale bioreactor. The model structure developed for the 3 L continuous closed loop fed-batch was applicable to the 15 L bioprocess (Supporting Information Appendix 1) and the majority of the model parameters remained identical (Table 1). The only parameters that changed between scales were the experimentally determinable yields and the maximum growth rate and the fitted rate of cell lysis, the maintenance coefficients for both glucose and glutamine and the specific production rate of the inhibitor. Thus, the model structure and parameter set transferred very well between scales in this study. This finding provides a large operational and hence, commercial benefit when scaling up bioprocesses due to the avoidance of the large burdens associated in redefining process models and the nontrivial task of parameter estimation.

Figures 2d,h,l show the performance of the model against the experimental results. It is evident that with a larger scale (15 L), the total cell density does not decline due to cell lysis as found in the case of a smaller scale (3 L), justifying the absence of cell lysis in the developed model ($K_{LYSIS} = 0$). The increase in the specific production rate of the inhibitor was necessary to model the lower cell densities (Figure 2d) experienced when the bioprocess was scaled up. Glucose and glutamine consumption and lactate and ammonia production showed similar trends to that of the cells grown in the continuous feed fed-batch 3 L bench-top bioreactor.

Comparison with other modeling approaches

Both semi-empirical logistic models and empirical statistical regression models were considered as alternatives to the Monod kinetic model due to the potentially lower effort required for development.

Logistic Equation Modeling. A modified logistic equation model was developed for both the exponential and decline phases of batch and fed-batch cultures to predict viable cell density (Eq. 14). The logistic equation depends on the estimation of only three experimentally determinable model parameters: intrinsic growth rate constant (r), the overall saturation constant (K) and the time at which the maximum viable cell density is reached (t_m), whereas, up to 17 parameters were required for the more comprehensive unstructured kinetic model. The estimation of such a large number of kinetic parameters is not a trivial task.

The logistic equation model parameters are presented in Table 2. The values of K increased as the mode of operation changed from batch to continuous feed fed-batch with a simultaneous decrease observed in the intrinsic growth

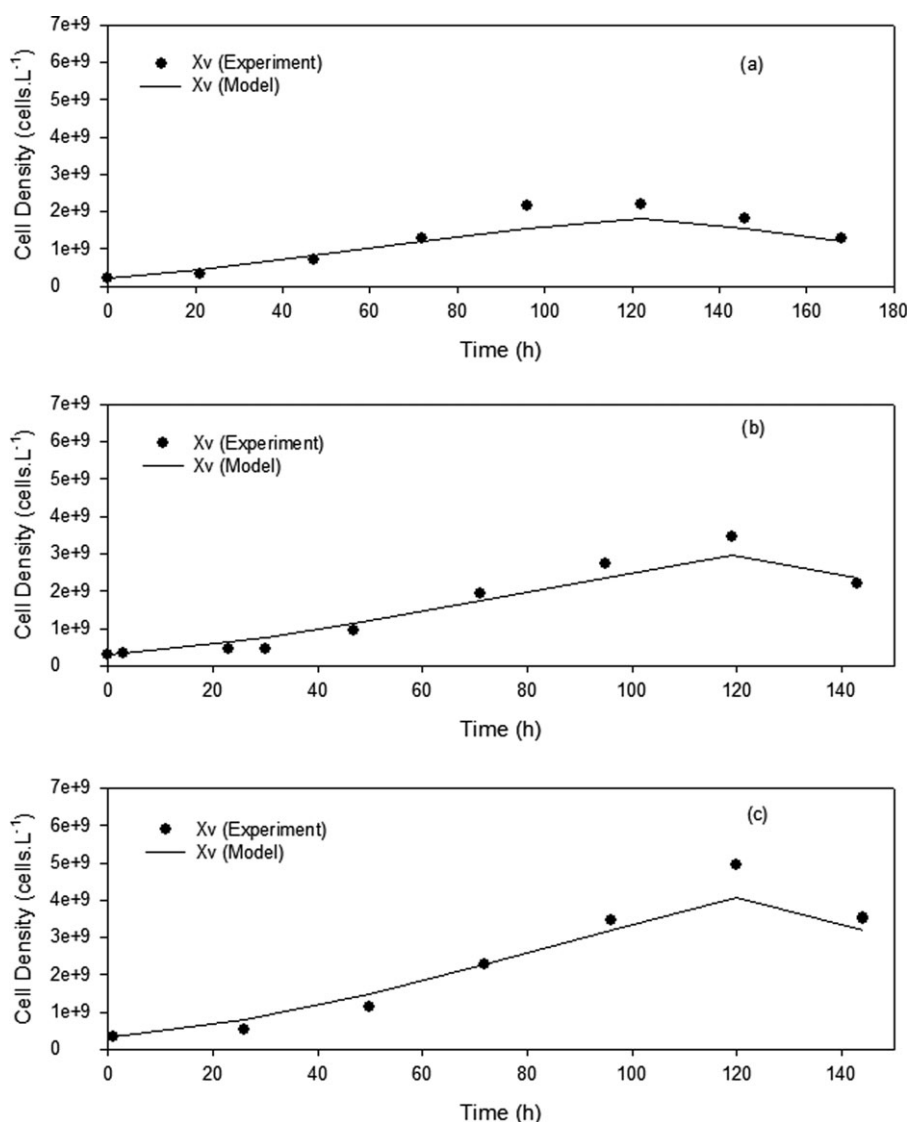


Figure 4. Logistic model predictions for viable cell density in a 3 L (a) batch, (b) bolus, and (c) continuous fed-batch bioreactor.

rate, r . Furthermore, the lower the intrinsic growth rate, r the longer the exponential phase. This explains why the continuous feed fed-batch bioprocess is the most efficient of the three modes of operation studied by reaching higher viable cell densities.

From a practical standpoint, the use of a single standalone equation to describe the cell growth adds to the simplicity of the logistic modeling method. The logistic equation model was fitted to the 3 L batch, bolus feed and continuous feed fed-batch experimental viable cell density data. In the Monod-type kinetic model there are several saturation constants pertaining to each of the substrates and metabolites, whereas in the Verhulst logistic equation model, K , represents the overall saturation constant of the substrates and the end metabolites. Therefore, model fits could not be produced for individual nutrient and metabolite concentrations of the batch and the fed-batch data via Verhulst logistic equation modeling. Logistic equation models for the nutrient and metabolite concentrations can also be affected by the distortion caused by the feeding of substrates in fed-batch mode.³⁹ Figure 4 illustrates the observed prediction ability of the logistic equation model to the viable cell density experimental data for all three modes of operation examined. Overall,

the model produced satisfactory fits. However, prediction improvements were seen with the more complex Monod type kinetic model predictions (Figures 2a–c) resulting in smoother optimal prediction profiles. This was due to the consideration of independent terms for growth-related and maintenance-related metabolism within the Monod kinetic type model framework.³⁹

Statistical Regression Models. Conceptually, logistic modeling and the statistical regression modeling approaches are similar. The number of parameters involved in statistical regression depends on the order of the regression models used. For instance, a 1st order regression model involves two parameters; a 2nd order regression model has three parameters and so on. The viable cell density logistic model has three parameters. However, it is difficult to give a physical interpretation of the model parameters, which are statistically estimated from the experimental results.

The nonlinear experimental data was linearized by a natural log function so as to derive a general linear equation. Expressing the data in this fashion resulted in a better measurement of “fit,” i.e. R^2 value.⁴⁰ It should also be noted that the size of the dataset available for the regression analysis is significant statistically. A stronger model could be built

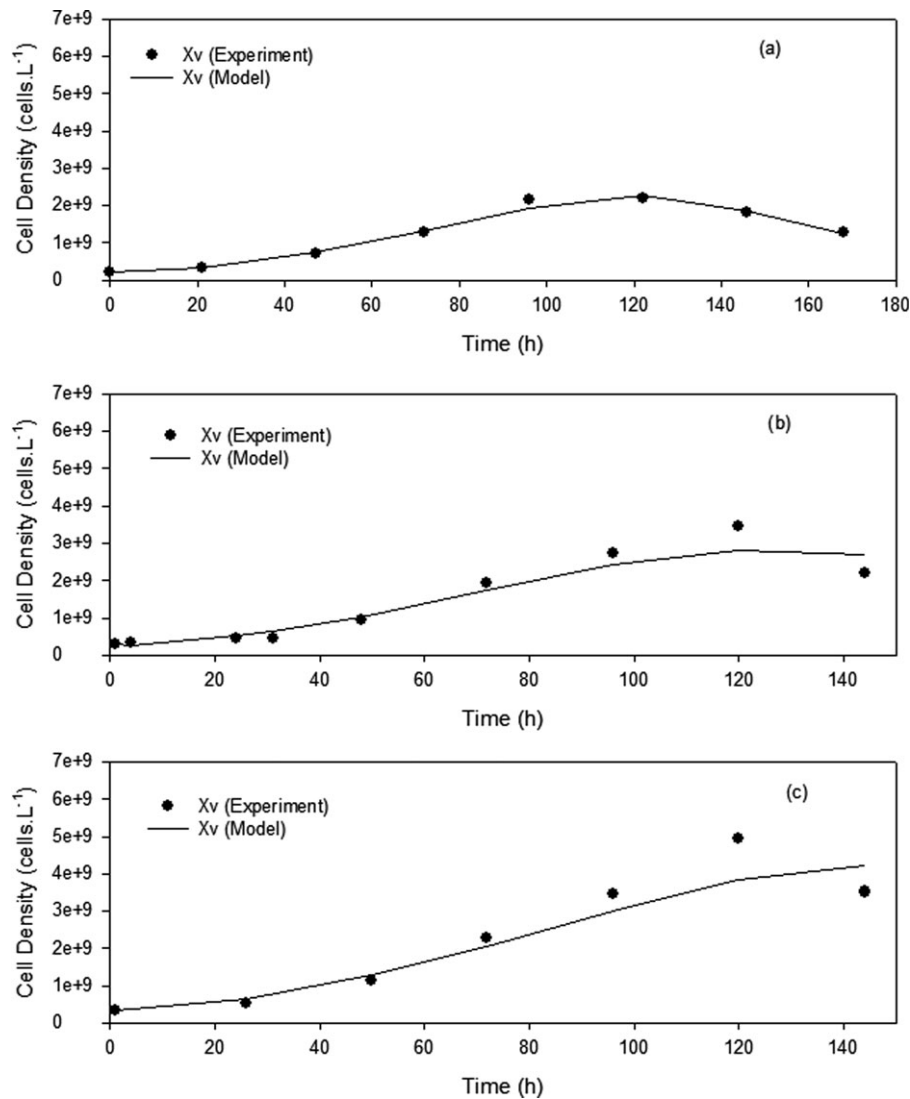


Figure 5. Statistical model predictions for viable cell density in a 3 L (a) batch, (b) bolus, and (c) continuous fed-batch bioreactor.

with a larger experimental dataset,⁴¹ whereas, the mechanistic Monod kinetic model and semi-empirical logistic equation model developed do not require large datasets for model development.

Similar to the logistic equation models, the regression modeling technique cannot accurately model the substrate concentrations during the fed-batch modes of operation due to the distortions caused by the feeding. Therefore, only polynomial approximations of the by-products, lactate and ammonia, were utilized in the empirical regression models for the viable cell density of the two fed-batch operations under study. The generalized linear regression models for the three modes of operation under study are shown in Eqs. 17, 18, and 19, where, $\ln(X_v) = f(G, Q, L, A)$. The categorical variables are G = glucose, Q = glutamine, L = lactate, A = ammonia and the response variable $\ln(X_v)$ = natural log of the viable cell density.

Batch

$$\ln(X_v) = 16.652 - 0.015G + 0.726Q + 0.014L - 0.452A \quad (17)$$

Bolus fed-batch

$$\ln(X_v) = 16.908 - 0.042L + 1.847A \quad (18)$$

Continuous fed-batch

$$\ln(X_v) = 18.112 - 0.007L + 1.120A \quad (19)$$

The performance of the statistical regression models for the 3 L bioreactor runs under the three modes of operation are shown in Figure 5.

Overall, this empirical statistical modeling approach is seen to produce satisfactory viable cell density predictions for the batch mode of operation whilst requiring no more than five model parameters compared up to 17 parameters necessary for the mechanistic kinetic model. However, the applicability of the statistical regression models suffered for the fed-batch modes of operation due to the inability to incorporate the main substrates into the viable cell density statistical regression models.

In summary, it was observed that even though both the semi-empirical logistic equation and empirical statistical regression models were less labor-intensive to develop, their lack of model complexity hindered their prediction capabilities.

Conclusions

In this work, the application of three different model types to a CHO 320 bioprocess operated under three different

modes of operation and grown on two different bioreactor scales has been illustrated.

The logistic equation model is essentially a semi-empirical mathematical model, which simply describes a set of data using a convenient mathematical relationship with little consideration of underlying phenomena. Similarly, statistical regression models cannot capture the full system dynamics. On the contrary, mechanistic Monod type kinetic models allow for the response to be considered in terms of known physical, chemical, and biological parameters. Interpretation of model parameters is of paramount importance for understanding the bioprocess.

The semi-empirical logistic equation and empirical statistical regression models consisted of a small parameter set, and as a result, they can be very useful where the complexity of system makes more detailed mechanistic models impractical to develop. However, the main downfalls of these models are their inability to model the nutrient and metabolite concentrations due to distortions caused by the feeding of nutrients and model structure limitations.

Sensitivity analysis is vital for screening the parameters for their relative influence on the bioprocess and for the evaluation of transferability across the modes of operation, feeding regime, and scale of operation. Even though initially the Monod kinetic model has a greater computational burden in optimizing a large number of parameters, the present work has demonstrated that many of the Monod kinetic model parameters do not have a significant effect on the model. Such a feature is important for the commercial implementation of such models, particularly when moving toward different modes of operation and across bioreactor scales. Overall, the results from this study show that the Monod kinetic model could be of more use within the QbD framework for control strategy design, feed profile optimization, online diagnostics, and greater bioprocess understanding than both the semi-empirical logistic equation and empirical statistical regression models.

Acknowledgments

This work was carried out as part of the “Bio Application of PAT (apPAT)” project, an Enterprise Ireland funded Industry Led Research Project.

Nomenclatures

A	= ammonia concentration (mM)
b_i	= partial regression coefficients (–)
C_i	= inhibitor concentration (mM)
C_{i*}	= inhibitor saturation concentration (mM)
F	= feed-rate ($L\ h^{-1}$)
G	= glucose concentration (mM)
$k_{d,Q}$	= degree of degradation of glutamine (h^{-1})
K	= overall growth saturation constant ($cells\ L^{-1}$)
K_L	= lactate saturation constant (mM)
K_A	= ammonia saturation constant (mM)
K_G	= glucose saturation constant (mM)
K_Q	= glutamine saturation constant (mM)
k_d	= death rate (h^{-1})
$k_{d,max}$	= maximum death rate (h^{-1})
k_μ	= intrinsic death rate (h^{-1})
K_{LYSIS}	= rate of cell lysis (h^{-1})
L	= lactate concentration (mM)
m_G	= glucose maintenance coefficient ($mmol\ cell^{-1}h^{-1}$)
m_Q	= glutamine maintenance coefficient ($mmol\ cell^{-1}h^{-1}$)

Q	= glutamine concentration (mM)
q_i	= specific inhibitor production rate (h^{-1})
r	= intrinsic growth rate constant (h^{-1})
S_G	= glucose concentration in the feed (mM)
S_Q	= glutamine concentration in the feed (mM)
t_m	= time at which X_{Vmax} occurs (h)
V	= volume (L)
X_T	= total cell density ($cells\ L^{-1}$)
X_D	= dead cell density ($cells\ L^{-1}$)
X_V	= viable cell density ($cells\ L^{-1}$)
X_{V0}	= initial viable cell density ($cells\ L^{-1}$)
X_{Vmax}	= maximum viable cell density ($cells\ L^{-1}$)
$Y_{A,Q}$	= yield of ammonia from glutamine (–)
$Y_{L,G}$	= yield of lactate from glucose (–)
$Y_{X,G}$	= yield of cells from glucose ($cells\ mmol^{-1}$)
$Y_{X,Q}$	= yield of cells from glutamine ($cells\ mmol^{-1}$)

Greek Symbols

μ	= growth rate (h^{-1})
μ_{max}	= maximum growth rate (h^{-1})

Literature Cited

- Walsh G. Biopharmaceutical benchmarks 2010. *Nat Biotech.* 2010;28:917–924.
- Rathore AS, Branning R, Cecchini D. *BioPharm Int.* 2007;36–40.
- Cook S, Patton KA, Bazemore LR. *BioPharm Int.* 2007;28–35.
- Read EK, Shah RB, Riley BS, Park JT, Brorson KA, Rathore AS. Process analytical technology (PAT) for biopharmaceutical products: Part II. Concepts and applications. *Biotechnol Bioeng.* 2010;105:285–295.
- Rathore AS, Winkle H. Quality by design for biopharmaceuticals. *Nat Biotech.* 2009;27:26–34.
- Glacken MW, Huang C, Sinskey AJ. Mathematical descriptions of hybridoma culture kinetics III. Simulation of fed-batch bioreactors. *J Biotechnol.* 1989;10:39–66.
- de Tremblay M, Perrier M, Chavarie C, Archambault. Optimization of fed-batch culture of hybridoma cells using dynamic programming: single and multi-feed cases. *J Bioprocess Eng.* 1992;7:229–234.
- Dalili M, Sayles GD, Ollis DF. Glutamine-limited batch hybridoma growth and antibody production: experiment and model. *Biotechnol Bioeng.* 1990;36:74–82.
- Bree MA, Dhurjati P, Georghegan RF, Robnett. Kinetic modeling of hybridoma cell growth and immunoglobulin production in a large-scale suspension culture. *Biotechnol Bioeng.* 1988;32:1067–1072.
- Zhang L, Hirai M, Shoda M. Removal characteristics of dimethyl sulfide by a mixture of *Hyphomicrobium* sp. I55 and *Pseudomonas acidovorans* DMR-11. *J Ferment Bioeng.* 1992;74:174–178.
- Liu YH, Bi JX, Zeng AP, Yuan JQ. A simple kinetic model for myeloma cell culture with consideration of lysine limitation. *Bioprocess Biosyst Eng.* 2008;31:569–577.
- Zeng AP, Hu WS, Deckwer WD. Variation of stoichiometric ratios and their correlation for monitoring and control of animal cell cultures. *Biotechnol Prog.* 1998;14:434–441.
- Fredrickson A. Formulation of structured growth models. *Bio-tech Bioeng.* 1976;18:1481–1486.
- Provost A, Bastin G. Dynamic metabolic modeling under the balanced growth condition. *J Process Contr.* 2004; 14:717–728.
- Portner R, Schafer. Modeling hybridoma cell growth and metabolism—a comparison of selected models and data. *J Biotechnol.* 1996;49:119–135.
- Batt BC, Kompala DS. A structured kinetic modeling framework for the dynamics of hybridoma growth and monoclonal antibody production in continuous suspension cultures. *Biotechnol Bioeng.* 1989;34:515–531.
- de Andrés-Toro B, Girón-Sierra JM, López-Orozco JA, Fernández-Conde C, Peinado JM, García-Ochoa F. A kinetic model for beer production under industrial operational conditions. *Math Comput Simul.* 1998;48:65–74.

18. Uchiyama K, Shioya S. Modeling and optimization of α -amylase production in a recombinant yeast fed-batch culture taking account of the cell cycle population distribution. *J Biotechnol.* 1999;71:133–141.
19. Bailey JE, Ollis DF. *Biochemical Engineering Fundamentals*. New York: McGraw-Hill; 1986.
20. Gibson A, Bratchell N, Roberts T. The effect of sodium chloride and temperature on the rate and extent of growth of *Clostridium botulinum* type A in pasteurised pork slurry. *Appl Bacteriol.* 1987;62:479–490.
21. Wilkinson DJ. Stochastic modeling for quantitative description of heterogeneous biological systems. *Nat Rev Genet.* 2009;10:122–133.
22. Malthus T. "An essay on the principle of population", (1798 1st edition) with "A summary view", and introduction by Professor Antony Flew: *Penguin Classics*; 1830.
23. Verhulst PF. Notice sur la loi que la population poursuit dans son accroissement. *Quetelet's Corresp.* 1838;10:113–121.
24. Tsoularis A, Wallace J. Analysis of logistic growth models. *Math Biosci.* 2002;179:21–55.
25. Jolicoeur P, Pontier J. Population growth and decline: a four-parameter generalization of the logistic curve. *Theor Biol.* 1989;141:563–571.
26. Amemiya T. Multivariate regression and simultaneous equation models when the dependent variables are truncated normal. *Econometrica.* 1974;42:999–1012.
27. Ergun M, Mutlu SF. Application of a statistical technique to the production of ethanol from sugar beet molasses by *Saccharomyces cerevisiae*. *Bioresour Technol.* 2000;73:251–255.
28. Shuler ML, Kargi F. *Bioprocess Engineering: Basic Concepts*. New Jersey: Simon & Schuster Company; 1992.
29. Bader FG, Kinetics of double-substrate limited growth, in Bazin MJ (ed.), *Microbial Population Dynamics*, CRC Press, Boca Raton, FL, 1982, pp. 132.
30. Levenspiel O. *Chemical Reaction Engineering*, 2nd ed. New York: Wiley; 1972.
31. Karaboga D, Okdem S. A simple and global optimization algorithm for engineering problems: differential evolution algorithm. *Turk J Elec Eng.* 2004;12:53–60.
32. Buehren M. Matlab Central. File Exchange. February 3, 2008. Available at <http://www.mathworks.com/matlabcentral/fileexchange/18593-differential-evolution>.
33. Brown D, Rothery P. *Models in Biology: Mathematics, Statistics and Computing*. Chichester: Wiley; 1993.
34. Kontoravdi C, Pistikopoulos EN, Mantalaris A. Systematic development of predictive mathematical models for animal cell cultures. *Comput Chem Eng.* 2010; 34:1192–1198.
35. Dodge TC, Ji GY, Hu WS. Loss of viability in hybridoma cell culture—a kinetic study. *Enzyme Microb Technol.* 1987;9:607–611.
36. Lee YK, Yap PK, Teoh AP. Correlation between steady-state cell concentration and cell death of hybridoma cultures in chemostat. *Biotechnol Bioeng.* 1995;45:18–26.
37. Buntmeyer H, Siwiora S, Lehmann J. Inhibitors of cell growth: accumulation and concentration. In: Carrondo MJT, Griffiths B, Moreira JL, editors. *Animal Cell Technology: From Vaccines to Genetic Medicine*. Dordrecht, The Netherlands: Kluwer Academic Publishers; 1997; 651–655.
38. Wong DCF, Wong KTK, Goh LT, Heng CK, Yap MGS. Impact of dynamic online fed batch strategies on metabolism, productivity and N-glycosylation quality in CHO cell cultures. *Biotech Bioeng.* 2004;89:164–177.
39. Goudar CT, Joeris K, Konstantinov KB, Piret JM. Logistic equations effectively model mammalian cell batch and fed-batch kinetics by logically constraining the fit. *Biotech Bioeng.* 2005;21:1109–1118.
40. Motulsky HJ, Christopoulos A. *Fitting models to biological data using linear and nonlinear regression. A practical guide to curve fitting*. Graphpad software Inc.: San Diego, CA; 2003.
41. Gold HJ. *Mathematical Modelling of Biological systems, An Introductory Guide*, A Wiley-International Publication; Wiley, New York; 1987.

Manuscript received Jun. 22, 2012, and revision received Sept. 28, 2012.