Chemical Product and Process Modeling

Volume 4, *Issue* 2 2009 *Article* 6

MODELING AND CONTROL

Iterative Design of Dynamic Experiments in Modeling for Optimization of Innovative Bioprocesses

Mariano Cristaldi* Ricardo Grau†

Ernesto Martinez‡

*INTEC (UNL-CONICET), mcristaldi@santafe-conicet.gov.ar

†INTEC (UNL-CONICET), cqfina@santafe-conicet.gov.ar

‡INGAR (CONICET-UTN), ecmarti@santafe-conicet.gob.ar

Iterative Design of Dynamic Experiments in Modeling for Optimization of Innovative Bioprocesses*

Mariano Cristaldi, Ricardo Grau, and Ernesto Martinez

Abstract

Finding optimal operating conditions fast with a scarce budget of experimental runs is a key problem to speed up the development and scaling up of innovative bioprocesses. In this paper, a novel iterative methodology for the model-based design of dynamic experiments in modeling for optimization is developed and successfully applied to the optimization of a fed-batch bioreactor related to the production of r-interleukin-11 (rIL-11) whose DNA sequence has been cloned in an *Escherichia coli* strain. At each iteration, the proposed methodology resorts to a library of tendency models to increasingly bias bioreactor operating conditions towards an optimum. By selecting the 'most informative' tendency model in the sequel, the next dynamic experiment is defined by re-optimizing the input policy and calculating optimal sampling times. Model selection is based on minimizing an error measure which distinguishes between parametric and structural uncertainty to selectively bias data gathering towards improved operating conditions. The parametric uncertainty of tendency models is iteratively reduced using Global Sensitivity Analysis (GSA) to pinpoint which parameters are keys for estimating the objective function. Results obtained after just a few iterations are very promising.

KEYWORDS: modeling, optimization, biotechnology, experimental design, dynamic experiments

*Please send correspondence to ecmarti@santafe-conicet.gob.ar.

1. Introduction

Findings in the 1950s that DNA is the molecule that encodes proteins, which in turn controls all cellular processes including metabolic pathways, have provided the impetus for the biotechnology era (Crommelin and Sidelar, 2002; Walsh, 2007). Since the first gene cloning in the beginning of the 1970s, it is now possible to modify a microorganism to produce a desired substance, often a protein. This has led to the advent of recombinant DNA technology which marks the birth of modern biotechnology. Recombinant DNA techniques permit the creation of exactly known changes in the DNA sequence. The production of biopharmaceuticals via recombinant technologies has led to a plethora of new, innovative bioproducts, as well as significant improvements in quality and yield of existing biotech processes. In the pharmaceutical industry many drugs are now being produced using genetically modified microorganisms. DNA modification is also used to manipulate protein sequences to give them better properties which has opened new possibilities to improve the productivity of industrial strains and to produce new high valued-added products such as insulin, humanized antibodies, interferons or interleukins along with biofuels (Fischer *et al*., 2008) and new materials (e.g. biodegradable polymers).

The recombinant microorganism is typically grown in a fed-batch bioreactor to high cell concentration and then expression of an heterologous protein is triggered so as to obtain considerable quantities of the target product (Cooney, 1983). For this purpose, the cell environment inside the bioreactor should allow optimal growth and product synthesis. One of the obstacles in attaining high product yields and high productivity is the accumulation of the metabolic by-products such acetate, which inhibits growth (Luli and Strohl, 1990) as well as production of the desired recombinant protein (Akesson et al., 2001). Formation of acetate in *Escherichia coli* cultures occurs under anaerobic conditions but also under fully aerobic conditions in situations with excess carbon source. In a fed-batch culture, the feed rate of the carbon source, usually glucose, must be manipulated in order to restrict overflow metabolism and glucose repression. To this aim, model-based optimization of a bioreactor operating condition seems to be the safe and economic approach to resort with. However, considering the large uncertainty and poor reproducibility in novel bioprocesses along with the issue of metabolic regulation (Ramkrishna, 2003), the development of an accurate mathematical representation of bioreactor dynamic behavior is a costly and very difficult undertaking.

Dynamic optimization of fed-batch bioprocesses has been an active area of research assuming that an accurate model of the bioreactor is available. Application of Pontryagin's maximum principle for fed-batch bioreactor optimization has been studied by several researchers (Lim et al., 1986; San and

Stephanopoulos, 1989; Rahman and Palanki, 1996; Mahadevan and Doyle III, 2003). Typically, fed-batch bioreactor optimization problems are linear in the manipulated variables and thus admit solutions where the manipulated variables are either on a singular arc or on the constraints. When discontinuities are present or multiple objective functions are considered, then direct approaches may converge to a local minimum (Luus, 1993). Iterative dynamic programming (Luus, 1990) or stochastic optimization algorithms (Banga et al., 1997) can be applied to such problems to obtain the global solution. Several of the optimization approaches discussed above have been applied to the optimization of recombinant protein in a fed-batch bioreactor (Park and Ramirez, 1988; Lee and Ramirez, 1994; Tholudur and Ramirez, 1996; Balsa-Canto et al., 2001). Also, lacking of relevant on-line measurements has prompted the use of estimation algorithms for bioprocesses (Bastin and Dochain, 1986; Bastin and Van Impe, 1995; Gudi et al., 1997). However, most of these optimization methodologies for bioprocess scaling up and productivity inprovement have not been widely accepted for industrial use since the perfect model assumption is far from real and bioreactor behavior is quite often deviant from model predictions. As a result, several modeling strategies have been proposed which have different characteristics ranging from purely statistical modeling methods going through unstructured models to more structured ones (Bailey, 1998).

Bioprocess modeling and optimization is a challenging task due to the complexity of both metabolic switches and organism's response to changes in its environment as consequence of the implementation of a given operating policy. Even though taking into account metabolism in bioreactor modeling is highly desirable, the issue of what can be reliably measured is a major obstacle to the level of detail that can be incorporated to improve model structure. For example, compartimental models (see Tang et al., 2007) require to sample not only extracellular substrate and metabolites but also inside the cell substrate and building blocks such ribosomes, mRNA and tRNA. Despite these advances, parameterizable structures of bioreactor models are still quite shallow which renders the issue of significant modeling errors to be addressed in model-based optimization by combining exploration with exploitation in model selection (Martínez, 2000). Keeping above considerations in mind, a new point of view called *modeling for optimization*, which has been developed for batch processes in general (Bonvin, 1998; Martínez and Wilson, 2003), is here applied to bioreactor operation for fast experimental optimization in bioprocess development.

One central concern in modeling of innovative bioprocessess is how to design optimally informative experiments taking into account poor knowledge about metabolic regulation, sparse and biased measurements of key intracellular pools of species and uncontrollable variations in the bioreactor dynamics from batch to batch. As a result, a model for an innovative bioprocess cannot be

DOI: 10.2202/1934-2659.1298

entirely knowledge-driven or data-driven alone. Modeling for optimization is a systematic way of combining scarce data with simple tendency models (unstructured models) aimed at reducing extrapolation errors in experimental optimization. In order to achieve the goal of optimal operation of innovative processes in the face of gross modeling errors, a number of requirements are imposed on modeling for optimization to make an impact on bioprocess development. A crucial issue is how to design a rather short sequence of dynamic experiments that are most informative in order to reduce the uncertainty about the parameters of the optimal operating policy. Despite the importance of this problem there is no previous work on the development of experimental design techniques addressing the more specific objective of modeling for optimization of bioreactors during the scaling up of innovative products. The problem addressed in this work is formulated as: "How does one adjust the time-varying controls, sampling strategy, initial conditions and length of each dynamic experiment to generate the information needed for the purpose of significantly reducing the uncertainty regarding the location of the process optimum operating condition in a bioreactor?" The notion of dynamic experiments (Asprey and Machietto, 2002) highlighted the fact that some control variables are time-varying during the experiment which rules out using standard experimental design techniques.

To address the above issues it is proposed here a policy iteration strategy which combines policy evaluation in designed dynamic experiments with increasingly bias data gathering towards improved operating conditions. Model selection is done by computing an error measure which distinguishes between parametric and structural uncertainty. Even though initial parametric uncertainty of tendency models is significantly high it is iteratively reduced using Global Sensitivity Analysis (GSA) to pinpoint which parameters are key for better estimating the objective function. Even though convergence towards an optimum cannot be guaranteed due to modeling errors, significant improvement in process performance is achieved after few iterations. Results obtained for the case study demostrate that despite modeling errors the proposed approach for model-based policy iteration does indeed converge to the optimal policy producing the maximum of protein IL-11 corresponding to the *in silico* bioreactor.

2. Modeling for optimization

In an attempt to compensante for bioprocess-model mismatch, optimal operation under uncertainty requires using measurements from carefully designed experiments to improve on a run-to-run basis from an initial input policy. The standard procedure consists of iteratively using new measurements to increasingly bias model parameter estimation and later resorting to the updated model for policy improvement. The underlying idea of *modeling for optimization* is to select

from a library of tendency models one which allows computing inputs that increasingly improve the operating policy and bias data gathering accordingly. Since the utility of model for a system or process must be assessed with regards to a purpose, in modeling for optimization the library of tendency models (Visser *et al*, 2000) is understood as a means to find a near-optimal operation policy despite incomplete understanding of process dynamics and uncontrollable disturbances affecting state evolution of a batch. Thus, tendency model development is not an end in itself as it is in kinetic studies where optimal design of experiments is aimed specifically at reducing model parametric uncertainty.

In modeling for optimization it will be assumed hereafter that initially the predictive capability is constrained to a library of tendency models which are valid qualitatively but quantitatively highly uncertain due to model parameterization errors and data bias. Model discrimination to handle uncertainty regarding model structure in modeling for optimization is addressed by model selection using a total error measure which accounts for parametric uncertainty and modeling errors. After each dynamic experiment, some parameters of all tendency models are selectively re-estimated using data obtained at properly chosen sampling times. Modeling for optimization thus revolves around iteratively improving the input policy based on the proper design of dynamic experiments that provide meaningful model parameter updates and model selection upon which process performance is incrementally improved on a run-to-run basis.

2.1 Model-based policy iteration

In what follows let's assume that the dynamic behavior of the bioreactor under study may be modeled by the set of ODEs comprising a given tendency model

$$\frac{dx}{dt} = f(x(t), \wp(w,t), \theta, t) \;\; 0 \le t \le t_f, \quad x(0): \text{given}, \tag{1}$$

and the optimization objective to be maximized is

$$J(w) = h(x(t_f)) + \int_0^{t_f} g(x(t), \wp(w,t), \theta, t)dt \tag{2}$$

where $x(t)$ is an n_s-dimensional vector of time dependent state variables, w is an m-dimensional vector of parameters for the input policy $\wp$, $\theta \in \Theta$ is a p-dimensional vector of model parameters and t_f is the final time of a batch run which in turn may also be optimized. The function g is the instantaneous reward

DOI: 10.2202/1934-2659.1298

function along the state trajectory defined by a given policy parameterization whereas the function *h* is the specific reward for the final state of the batch run when using the input policy $\wp(w,t)$. It is worth noting that Eq. (2) defines the *value* (to be maximized) of a policy in the dynamic programming jargon (Powell, 2007). Accordingly, a policy defined by the set of parameters w_2 with value J_2 is better than (or preferred to) a policy defined by w_1 with value J_1 if and only if $J_2 > J_1$. The sensitivity of process perfomance to policy parameterization is a central issue for designing optimally informative dynamic experiments to bias data gathering in modeling for optimization and when deciding which subset of model parameters should be re-estimated using data gathered in the current iteration.

For a given model parameterization $\hat{\theta} \in \Theta$, the optimal policy $\wp^*$ for the deterministic continuous-time optimal control problem defined by Eq. (1) and Eq. (2) above should satisfy the well-known *Hamilton-Jacobi-Bellman* (*HJB*) optimality condition:

$$\frac{\partial J(w^*)}{\partial t} = \max_{u \in U} \{ g(x, \wp^*, \hat{\theta}, t) + \left[\frac{\partial J(w^*)}{\partial x} \right]^T f(x, \wp^*, \hat{\theta}, t) \}, \text{ for all } t,x, \tag{3}$$

with the boundary condition $J(w^*) = h(x(t_f))$.

The *HJB* optimality conditions is a partial differential equation which must be satisfied for all time-state pairs (*t,x*) by the reward-to-go function $J(w^*)$. The optimal policy $\wp^*(w^*,t)$ can be found by iteratively maximizing the right-handside of Eq. (3) using a typical policy iteration approach of Dynamic Programming (Powell, 2007) . Each iteration starts with an initial policy which is evaluated in a specifically designed dynamic experiment (See Section 3.2. below). Data gathered in the evaluation experiment is used to change selectively model parameters so that a new policy is obtained by solving the dynamic optimization problem defined by Eqs. (1) and (2). As long as the policy evaluation experiment produces meaningful information to further improve policy performance the identification-optimization cycle continues. Assuming the performance prediction mismatch can be driven to zero, policy iteration will converge to the optimal policy $\wp^*$ which gives rise to optimal performance *J**. Ideally, if model parameters θ be perfectly known *a priori,* model-based policy iteration is able to provide the optimal solution once the first iteration has been completed. As model parametric uncertainty is significantly high for bioprocesses an optimally informative dynamic experiment must be designed in each policy iteration step.

To start with model-based policy iteration, model parameters are initially set to $\hat{\theta}$ using mid-values of their current feasible interval or better, if available, expert judgment. This rough model parameterization allows obtaining an initial

policy $\wp_1$ which will be progressively improved using a sequence of designed experiments. Alternative, policy iteration may start with a known operating policy based on good practices and heuristics for the bioprocess to be optimized. Despite the initial policy is often very conservative and sub-optimal, it is able to provide a sensible bias to guide data sampling in the first iteration. Data gathered in policy evaluation experiments is increasingly used to reduce selectively parametric uncertainty in each tendency model and through dynamic optimization an improved policy $\wp_2$ is obtained and so on and so forth. To deal with the issue of persistent excitation in the generated data with the current policy (re)estimation is only done for the subset θ_ℓ of model parameters which are the most relevant for reducing the performance prediction mismatch. To determine which parameters are in θ_ℓ, GSA techniques (Saltelli et al., 2004, 2006; Sobol, 1993) is used. For this subset of parameters new estimations and confidence intervals are obtained as follows.

For the current iteration, once the most relevant parameters comprising θ_ℓ has been identified using GSA, this sub-set of model parameters is re-estimated by minimizing model response errors while maintaining other model parameters in their values from the previous iteration.

$$\hat{J} = \min_{\theta_\ell \in \Theta_\ell} \left\{ tr \left\| \left(Y\left(\theta_\ell^*\right) - Y_{\exp} \right)^T * \left(Y\left(\theta_\ell^*\right) - Y_{\exp} \right) \right\| \right\} \tag{4}$$

where $\hat{J}$ is the least-square error function for parameter estimation, $Y(\theta_\ell^*)$ is the vector of state variables which are predicted by varying $\hat{\theta}_\ell^*$, whereas $Y_{\exp}$ correspond to experimental data gathered in a given experiment. To compute the (1-α) confidence intervals for each parameter whose value has been re-estimated using data from a designed experiment the following approach based on Monte Carlo simulations is used. Let's denote by $\hat{J}_{\min}$ the minimum value for $\hat{J}$ obtained from solving the minimization problem in (4); choosing a given value for α, a maximun admissible error $\hat{J}_{\max}$ is defined (typically the confidence level is chosen so that $\hat{J}^*$ cannot be more than 5% greater than $\hat{J}_{\min}$). In each simulation, the values of parameters in the subset θ_ℓ are allowed to vary randomly in the subspace defined for the previously known uncertainty space Θ_ℓ and the corresponding $\hat{J}^*$ value is computed. If $\hat{J}^* \in [\hat{J}_{\min}, \hat{J}_{\max}]$ then the corresponding vector of parameters θ_ℓ is selected, otherwise is excluded from the confidence interval set. After a sufficiently high number of realizations of the vector θ_ℓ have been tried the confidence intervals are re-calculated for each parameter in θ_ℓ with a confidence level of (1-α) %.

To end this sub-section some comments regarding *identifiabiliy* and *estimability* of tendency models used in modeling for optimization of

DOI: 10.2202/1934-2659.1298

bioprocesses are required. Parameter identifiability is concerned with establishing whether the values of measured state variables (model outputs) correspond to an unique realization of parameter values (Walter, 1987). That is, can unique parameter estimates be recovered, in theory, from noise-free output data? If a model is unidentifiable, then this means that several parameter vectors may exist that correspond to exactly the same dynamic behavior of the model and it is impossible to distinguish any of them using (even noise-free) data alone. Parameter estimability has a broader definition and involves an analysis concerned with actual samples or sampling errors and includes, for example, estimation of variances and the impact of cross-correlation of parameters (Sidoli, et al., 2005). For tendency models with few parameters and equations, techniques exist that make them amenable to guarantee practical identifiability and estimatibility (Walter, 1987; Walter and Pronzato, 1996), assumming data have been pre-treated to reduce significantly the signal-to-noise ratio and outliers are not present (Cinar, et al., 2003). Finally, it is worth noting that in modeling for optimization the objective is to increasingly improve the operating policy using models as *guidelines* for designing optimally informative experiments and not to provide very accurate predictions which renders the issues of identifiability and estimability as much less critical for success.

2.2 *Global Sensitivity Analysis*

Global Sensitivity Analysis (GSA) (Saltelli et al., 2004, 2006) takes into account the fact that parametric uncertainty in complex models can propagate, compensate or suffer many kinds of interactions which may affect the *output* of interest (e.g., the performance index *J*) in different ways. GSA is a variance-based technique that decomposes model *output*s variability as a combination of uncertainty from each *i-th* independent *input* factor and its interactions with other factors. This decomposition attempts to rank the importance of uncertainty sources by mean of sensitivity indices. Briefly, let's suppose that the value of a model *output* of interest y is estimated with an uncertainty due to a set of k independent parameters x_i i=1,2,…,k. Furthermore, those parameters can interact among them and as a whole influence on model output y . Unconditional variance $V(y)$ of y is decomposed as follow:

$$V(y) = \sum_i V_i + \sum_i \sum_{j>i} V_{ij} + \ldots + V_{ij\ldots k} \tag{5}$$

$$V_i = V_{x_i}(E_{x_{-i}}(y|x_i))$$
$$V_{ij} = V_{x_i x_j}(E_{x_{ij}}(y|x_i x_j)) - V_{x_i}(E_{x_{-i}}(y|x_i)) - V_{x_j}(E_{x_{-j}}(y|x_j)) \tag{6}$$

where V_i is the amount of the total variance in the model response which can be explained based on the values of *i-th* parameter and it is known as main effect term for the *i-th* parameter; V_{ij} is the amount of variability generated due to the interaction between the *i-th* and *j-th* parameters. Note that in computing V_i, it is necessary to compute and integrate over x_{-i} (all factors except x_i) and then a new integral over the marginal distribution of x_i to finally know the conditional variance V_i. The objective of applying GSA is to rank factors so as to know how *V(y)* would be reduced if some of those factors were fixed in their true value. Accordingly, a first measure of the fraction of *V(y)* which accounts for the uncertainty of x_i is the so-called first order sensitivity index Si_i defined as:

$$Si_i = \frac{V_i}{V} \tag{7}$$

Estimators for Si_i's can be obtained following different approaches and here they have been computed by the Sobol's method (Sobol, 1993) which has been recently improved by Saltelli *et al.* (2004, 2006). In this method a quasi-random sampling in the multi-dimensional space spanned by the parameter set X is used in order to find the sensitivity indices which ensures exploration over the whole range of variation for all *input* factors. To facilitate assessing the significance of sensitivity indices computed using Monte Carlo simulations it can be useful to normalize each *Si* as follows:

$$Si_i^n = \frac{Si_i - \min(Si_j)}{\sum_j Si_j}; \quad j = 1,...,k \tag{8}$$

2.3 *Optimal sampling*

Due to the *a priori* significant uncertainty about model parameters their values should be estimated selectively using data gathered in policy evaluation experiments. For a given policy $\wp_n$ in the *n-th* policy iteration of Fig. 1 optimal sampling times ψ^{opt} along a batch run must be calculated so as to bring new information to selectively reduce parametric uncertainty which affect the most the value estimation of the performance index *J*. Assuming model parameters are set to $\hat{\theta}$ and the current policy is $\wp(w_n, t)$, the issue of optimal sampling is related to calculating at which times $\psi^{opt} \in \Psi$ in a dynamic experiment the values of measured process variables are most informative in modeling for optimization assuming that the policy evaluation step should narrow down the uncertainty

DOI: 10.2202/1934-2659.1298

about the optimal input. To this end, the following optimization problem is solved:

$$\psi^{opt} = \max_{\psi \in \Psi} \det \left| M(\hat{\theta}, \wp(w_n, t), \psi) \right| , \; M = Q^T Q$$
$$Q = \begin{pmatrix} Si_{11} & \cdots & Si_{1n} \\ \vdots & \ddots & \vdots \\ Si_{m1} & \cdots & Si_{mn} \end{pmatrix} \tag{9}$$

where each entry of the matrix Q , Si_{ij}, measures the sensitivity of the performance index $J(w_n)$ at the *i-th* sampling time with respect to *j-th* parameter of the operating policy. The number of samplings along each run will be defined in accordance to the budget for processing samples and bearing in mind that this number should be the at least the number parameters defining the input policy.

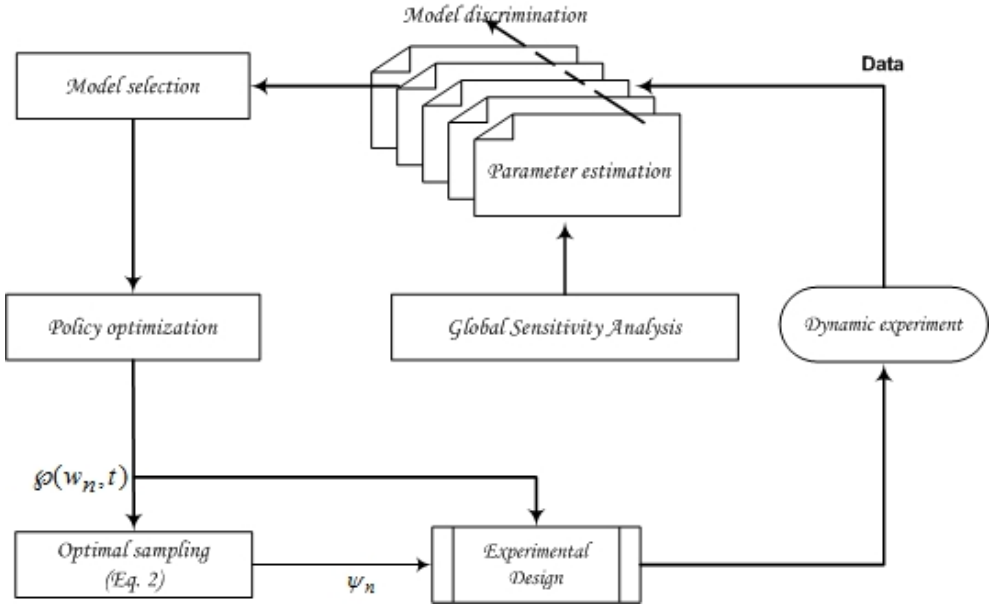

Figure 1. Modeling for optimization using designed dynamic experiments.

Fig. 1 provides a summary of the proposed methodology for experimental design in modeling for optimization. At each iteration, a dynamic experiment is designed around the current policy $\wp_n$, and optimal sampling times ψ^{opt} are calculated by solving (9). The experiment is carried out and new data is collected. Based on these data the sub-set $\hat{\theta}_\ell$ of model parameters for each tendency model

are re-estimated which reduces parametric uncertainty. Based on total modeling errors a tendency model is selected for policy re-optimization. With the new input policy $\wp_{n+1}$ and a new iteration begins. The identification-optimization cycle is continued until no performance improvement is obtained and the parameters *w* for the calculated input policy converge.

It worth mentioning that the very purpose of the above procedure is to increasingly bias model structure/parameters using data from optimally designed experiments in the search for an improved operating policy. Model structure and parameter setting is thus biased to find better operating policies. As soon as no significant improvement in the performance index *J* is achieved from one iteration to the next, there is no point in doing more experiments.

2.4 *Model selection*

In this work, model selection is based on distinguishing between parametric uncertainty and structural errors in performance prediction using tendency models (see Fig. 2 for details). For the *rth* realization of model parameters, the corresponding simulated trajectory of process performance is $\hat{J}_r$. At each sampling point, a sample average $\left\langle \hat{J} \right\rangle$ of different model parameterizations can be used to characterize parametric uncertainty for each tendency model as follows (Asprey 2000; Asprey and Machietto, 2002; Chen and Asprey, 2003):

$$\varepsilon_{\hat{J}_i \to \langle \hat{J}_i \rangle} = \frac{1}{n_{sp} n} tr\left[\left(\hat{J}_{ir} - \left\langle \hat{J}_i \right\rangle \right) W_{ir} \left(\hat{J}_{ir} - \left\langle \hat{J}_i \right\rangle \right)^T \right]; \quad i = 1,2,...,n_{sp} \quad r = 1,2,...,n \tag{10}$$

where W_{ir} is a weighting matrix. As parametric uncertainty is iteratively reduced the importance of structural errors in each model regarding performance predictions are more evident when the operating condition is changed in the search for policy improvement. As a measure of structural uncertainty, the average performance trajectory $\left\langle \hat{J}_i \right\rangle$ is compared at *ith* sampling point to the actual trajectory J_i to define the structural error:

$$\varepsilon_{\langle \hat{J} \rangle \to J} = \frac{1}{n_{sp}} \left[\left(\left\langle \hat{J}_i \right\rangle - J_i \right)^T W_{ii} \left(\left\langle \hat{J}_i \right\rangle - J_i \right) \right]; \quad i = 1,2,...,n_{sp}; \quad W_{ii} : weighting \;\; matrix \tag{11}$$

DOI: 10.2202/1934-2659.1298

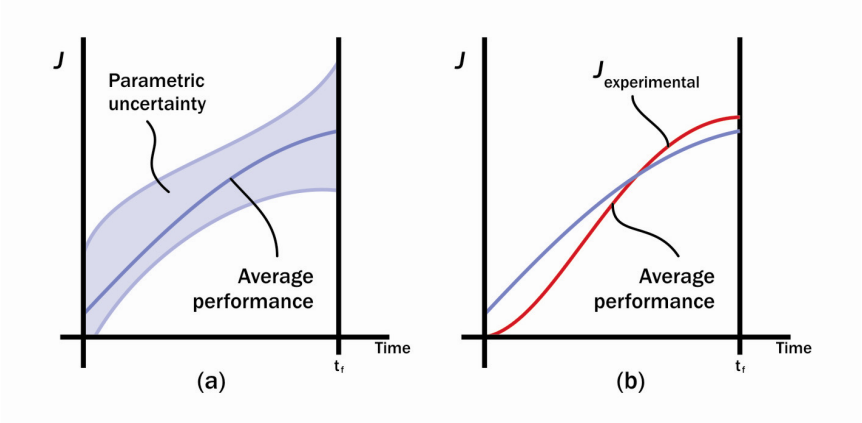

Figure 2. Monte Carlo estimation of parametric uncertainty and structural errors (a) Parametric uncertainty; (b) structural mistmach.

The total error of the ℓth model can be expressed as a weighted sum of errors in (10) and (11). Model selection may simply be done so that the tendency model whose total error is the lowest is chosen for policy optimization. More elaborated strategies for model selection can also be developed. For example, initially model selection may emphasize reducing parametric uncertainty and as more data are gathered model selection is more based on structural errors.

3. Case study

To illustrate the proposed methodology results obtained in the optimization of fed-batch fermentation process for the recombinant protein rIL-11 using a genetically modified *E. coli* strain are presented. Production of recombinant proteins in *E. coli* has been widely applied in both laboratory research and bioproduct manufacturing since this microorganism is considered a reliable source of proteins. This method may achieve profitable mass productivity due to high density cell growth and fast product formation. A structured kinetic model proposed by Tang et al. (2007) which describes state variables trajectories such as: biomass (X), substrate (S), intracellular recombinant protein concentration (P) will be used as an *in silico* bioreactor to generate the required data in the modeling for optimization approach. Four unstructured (tendency) models which differ in their biomass growth kinetics are used as guidelines for policy optimization so that the mismatch between the "real" bioprocess and alternative models of the fed-batch bioreactor is accounted for by increasingly biasing data gathering. Also, the operation policy has been defined based on the substrate

feeding rate and induction time as the main components subject to optimization, including the initial culture condition. The performance index $J(t)$ is related to the amount of recombinant protein obtained at the final time. Tendency model equations and their alternative biomass growth kinetics are:

$$\frac{dX}{dt} = \mu X;\ \frac{dS}{dt} = -\frac{\mu}{Y_{xs}} X - f(X,t);\ \frac{dP}{dt} = r_P - \mu P \tag{12}$$

$$r_P = \begin{cases} 0, & t \geq t_{ind} \\ K_P^{\max}\left(\frac{S}{K_s + S}\right)\left[\frac{1}{1+\left(\frac{P}{KI_P}\right)^5}\right], & t \geq t_{ind} \end{cases}$$

$$\textit{First order}: \quad \mu = \mu_{\max} S; \qquad f(X,t) = 0$$

$$\textit{Monod}: \quad \mu = \mu_{\max} \frac{S}{K_S + S}; \qquad f(X,t) = 0$$

$$\textit{Contois}: \quad \mu = \mu_{\max} \frac{S}{K_x X + S}; \qquad f(X,t) = 0$$

$$\textit{Maintenance}: \mu = \mu_{\max} \frac{S}{K_S + S}; \qquad f(X,t) = mX$$

Based on experimental data provided by Tang et al. (2007), a rather rough estimation of each tendency model parameters was made and referred to as "initial values" in Table 1. Due to the significant level of parametric uncertainty a ±50% confidence interval around these initial values for each parameter is assumed in the first policy optimization iteration. Moreover, a uniform distribution over its confidence interval is assumed for each parameter. This level of uncertainty is important to provide ample room for exploration in model selection, mainly in the initial steps of modeling for optimization.

At any time t, the input policy $\wp(w,t)$ is defined by a vector w of parameters corresponding to two different degrees of freedom for process optimization. A subset of the policy parameters (vector w) corresponds to inputs that can be modified from run-to-run but are time-invariant in a given run such as the substrate feeding concentration, run duration or induction time. The remaining entries are parameters which are used here for describing the profile of time-

DOI: 10.2202/1934-2659.1298

varying control variables such as the feeding rate. In the latter case, a key issue is the mathematical description to be used so as to provide ample room for different variability patterns within economic and safety constraints with a minimum number of independent parameters. Without any loss of generality the following quadratic inverse polynomial is used hereafter:

$$F_{in} = \begin{Bmatrix} 0 & t < t_0 \\ \dfrac{At}{1 + Bt + Ct^2} & t \geq t_0 \end{Bmatrix} \tag{13}$$

Table 1. Initial parameterization for each tendency model based on experimental data

Parameter	Unit	Model			
		1st Order	Monod	Contois	Maintenance
μ_{max}	h^{-1}	0.2000	0.6301	0.5607	0.5261
K_s	$g\ L^{-1}$	2.0184	1.4956	-	0.7190
Y_{xs}	$g_{biomass}\ g_{substrate}^{-1}$	0.3982	0.4506	0.4826	0.4464
K_P^{max}	$g\ L^{-1}$	0.0759	0.0629	0.0557	0.0536
KI_p	$g\ (L\ h)^{-1}$	0.0877	0.0609	0.0627	0.0600
K_x	$g_{substrate}\ g_{biomass}^{-1}$	-	-	1.7291	-
m	h^{-1}	-	-	-	0.0100

In the past there have been various approaches to implement bioreactor feeding policies which can be defined as constant, piecewise constant, piecewise continuous or fully continuous functions of time. In this work, the feeding rate profile is described using inverse polynomials of low order with respect to time. Inverse polynomials resort to a small number of parameters to define time trajectories which are quite flexible for modeling a rich variety of continuous feeding patterns for bioreactor optimization. It is worth noting that the methodology proposed in Section 2 is by no means limited by the family of mathematical functions used to describe time-varying input controls. However, bioreactor dynamics slowly unfolds cell responses to environmental changes which require using smooth profiles for time-varying control inputs such as the one in Eq. (13).

In the first experiment a constant feeding rate of 1 liter/hour is used whereas the final time t_f is set to 12 hours (see Table 2 for other policy parameters in the first experiment). Sampling times are arbitrarily chosen as evenly space over the fed-batch run:

$$\psi_0[h] = [1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12]$$

Table 2. Optimum operation policy for *E. coli* culture for rIL-11

Parameter	Units	Initial Condition	1st iter	2nd iter	3rd iter	4th iter	5th iter	6th iter
Z	$L\ h^{-1}$	1	-	-	-	-	-	-
A	$L\ h^{-2}$	0	0.0544	0.0121	0.0911	0.1431	0.2385	0.2389
B	h^{-1}	-	$3\ 10^{-4}$	$3\ 10^{-4}$	$3\ 10^{-4}$	$3\ 10^{-4}$	$3\ 10^{-4}$	$3\ 10^{-4}$
C	h^{-2}	-	$3\ 10^{-4}$	$4\ 10^{-4}$	$8\ 10^{-4}$	0.0156	0.0222	0.0223
S_f	$g\ L^{-1}$	10	30	30	30	30	30	30
t_{feed}	h	6	0	0	3.19	4.05	5	5
t_{ind}	h	4	4	4	4	4	4	4
t_f	h	12	16	16	16	16	16	16
V_0	L	6	5.31	10	5	6.61	5	5
X_0	$g\ L^{-1}$	0.05	0.1	0.1	0.1	0.1	0.1	0.1
S_0	$g\ L^{-1}$	6	3.93	7	7	7	7	7
$J \cdot V_f$	g	**1.60**	**6.06**	**3.75**	**7.15**	**6.40**	**7.22**	**7.22**

The final amount of protein obtained from this initial experiment is 1.60 g. Data obtained from this experiment (see Fig. 3) are going to be used to re-estimate those parameters in each tendency model which are most influential in determining the final amount of protein. In Table 3, GSA indices based on the current policy are provided whereas Table 4 the new values and confidence intervals following selective parameter re-estimation are shown.

To define the policy for the next experiment a tendency model must be selected. To this end parametric uncertainty and structural errors are calculated. From data shown in Table 5, the Contois model seems to be the one having the lowest total error.

Table 3. Normalized Global Sensitivity Indices for models parameters

	S_i^n (initial)			
	Model			
Parameter	1st Order	Monod	Contois	Maintenance
μ_{max}	0.0103	0.5318	0.5501	0.5679
K_s	0	0	-	0.0074
Y_{xs}	0.6276	0.2600	0.1747	0.2220
K_P^{max}	0.1898	0.0541	0.1921	0.0664
KI_p	0.1723	0.1541	0.0831	0.1363
K_x	-	-	0	-
m	-	-	-	0

DOI: 10.2202/1934-2659.1298

Table 4. Parameters re-estimation after 1st experiment

Model	$\mu_{\max}$	Y_{xs}
1st Order	-	0.687±0.012
Monod	0.4882±0.0015	-
Contois	0.485±0.011	-
Maintenance	0.4452±0.0014	-

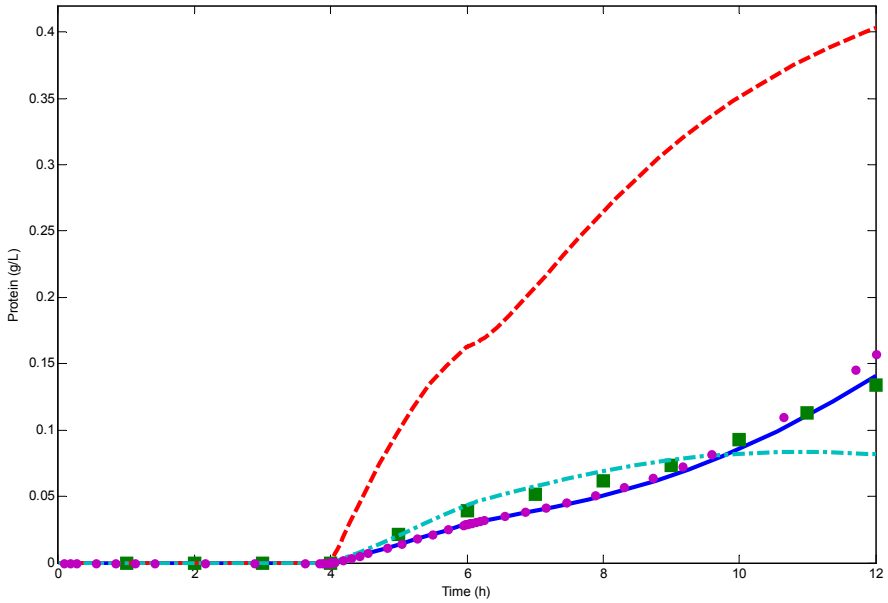

Figure 3. Predicted trajectories for 1st order (- -),Monod (-), Contois (.-) and Manteinance (..) models estimations for state variables and "experimental data" from *in silico* bioreactor run (squares) for the 1st experimental design (iteration # 0).

Table 5: Errors for different models (initial)

Model	$E_{j_i \to \langle j \rangle}$	$E_{\langle j \rangle \to J}$	E_{total}
1st Order	0.0044	0.0209	0.0254
Monod	$7.13\ 10^{-12}$	$6.23\ 10^{-5}$	$6.23\ 10^{-5}$
Contois	$1\ 10^{-4}$	$5.78\ 10^{-5}$	$1.58\ 10^{-4}$
Manteinance	1.1403	$6.35\ 10^{-5}$	1.1404

Using the Contois model with the parameter μ_{max} being updated as shown in Table 4, an improved policy identified as 1st iteration in Table 2 is obtained. It is worth noting that a significant increase in the amount of protein obtained is

achieved (6.06 g). Optimal sampling times for this experiment are then calculated using Eq. (9) as:

1^{st} iteration:

Sampling times:

$$\psi_{1st}[h] = [0.5, 1.0, 1.5, 2, 3.83, 5.0, 6.33, 6.83, 12.17, 12.67, 13.17, 13.67]$$

$$M = 4.37 \cdot 10^{-26}$$

Data obtained from this second experiment (see Fig. 4) is going to be used to re-estimate those parameters in each tendency model which are most influential in determining the final amount of protein. In Table 6, GSA results around the current policy are provided whereas Table 7 the new values and confidence intervals following selective parameter re-estimation are shown.

Table 6. Normalized Global Sensitivity Indices for models parameters

	S_i^n (1st iter)			
	Model			
Parameter	1st Order	Monod	Contois	Maintenance
μ_{max}	0.1053	0	0	0
K_s	0.0745	0.0178	-	0.0033
Y_{xs}	0	0.5153	0.5101	0.5567
K_P^{max}	0.2978	0.0850	0.0284	0.0609
KI_p	0.5224	0.3818	0.2576	0.3779
K_x	-	-	0.2039	-
m	-	-	-	0.0012

Table 7. Parameters re-estimation using 1st iter data

Model	KI_p	Y_{xs}
1st Order	0.0707±0.0014	-
Monod	-	0.487±0.003
Contois	-	0.780±0.009
Maintenance	-	0.438±0.002

In Fig. 4, data from the 1st iteration and tendency model predictions are shown wheras total errors for each tendency model are indicated in Table 8. From these estimated errors the Contois model is again selected for the 2nd iteration.

DOI: 10.2202/1934-2659.1298

Table 8. Errors for different models (1st iter)

Model	$E_{j_i \to \langle j \rangle}$	$E_{\langle j \rangle \to J}$	E_{total}
1st Order	0.0049	0.0485	0.0534
Monod	0.0034	$3.96\ 10^{-4}$	0.0038
Contois	0.0013	0.0023	0.0033
Manteinance	0.0060	$2.79\ 10^{-4}$	0.0062

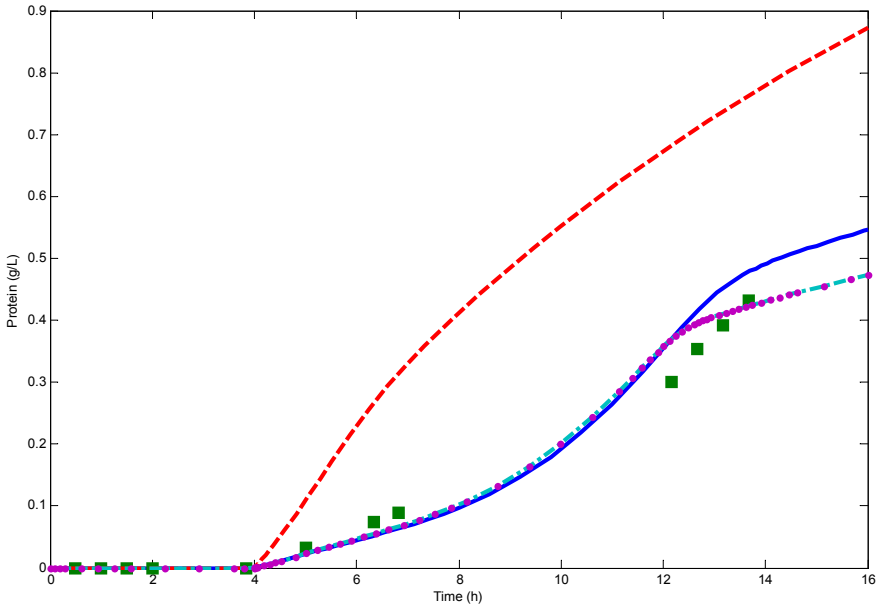

Figure 4. Predicted trajectories for 1st order (- -),Monod (-), Contois (.-) and Manteinance (..) models estimations for state variables and "experimental data" from *in silico* bioreactor run (squares) applying the optimal policy in the 1st iteration (Contois model).

In the second iteration, parametric uncertainty and structural errors bias data gathering in such a way that protein production is eventually lower. However, the resulting sequence of dynamic experiments is informative enough so as to converge to the optimal operating condition of the *in silico* reactor after a few iterations. A summary of results obtained until convergence are given as follows. As can be seen, there exists a remarkable shrinking of parametric uncertainty in all models included in the library for optimization when the proposed methodology is applied. Moreover, the relative weight of structural uncertainty is increased when *in silico* bioreactor is operated far away of initial

condition despite some compensations done by re-parametrization of models used for optimization.

2nd iteration:

Sampling times:

$\psi_{2nd}[h] = [1.17, 1.67, 2.33, 2.83, 3.33, 3.83, 5.17, 6.17, 10.67, 12.0, 12.5, 15.83]$

$M = 1.87 \cdot 10^{-26}$

Table 9. Normalized Global Sensitivity Indices for models parameters

	S_i^n (2nd iter)			
	Model			
Parameter	1st Order	Monod	Contois	Maintenance
$\mu_{\max}$	0.4718	0.0012	0.0050	0.0003
K_s	0.1894	0	-	0
Y_{xs}	0	0.0140	0	0.0010
$K_P^{\max}$	0.3372	0.1695	0.0289	0.1902
KI_p	0.0016	0.8153	0.5957	0.8046
K_x	-	-	0.3704	-
m	-	-	-	0.0039

Table 10. Parameters re-estimation for 2nd iter

Model	$\mu_{\max}$	KI_p
1st Order	0.373±0.007	-
Monod	-	0.0701±0.0013
Contois	-	0.0777±0.0016
Maintenance	-	0.080±0.002

Table 11. Errors for different models (2nd iter)

Model	$E_{\hat{J}_i \to \langle \hat{J} \rangle}$	$E_{\langle \hat{J} \rangle \to J}$	E_{total}
1st Order	0.0023	0.0064	0.0087
Monod	$4.32\ 10^{-4}$	$5.61\ 10^{-5}$	$4.88\ 10^{-4}$
Contois	$2.65\ 10^{-4}$	0.0011	0.0014
Maintenance	$5.50\ 10^{-4}$	$1.01\ 10^{-4}$	$6.51\ 10^{-4}$

DOI: 10.2202/1934-2659.1298

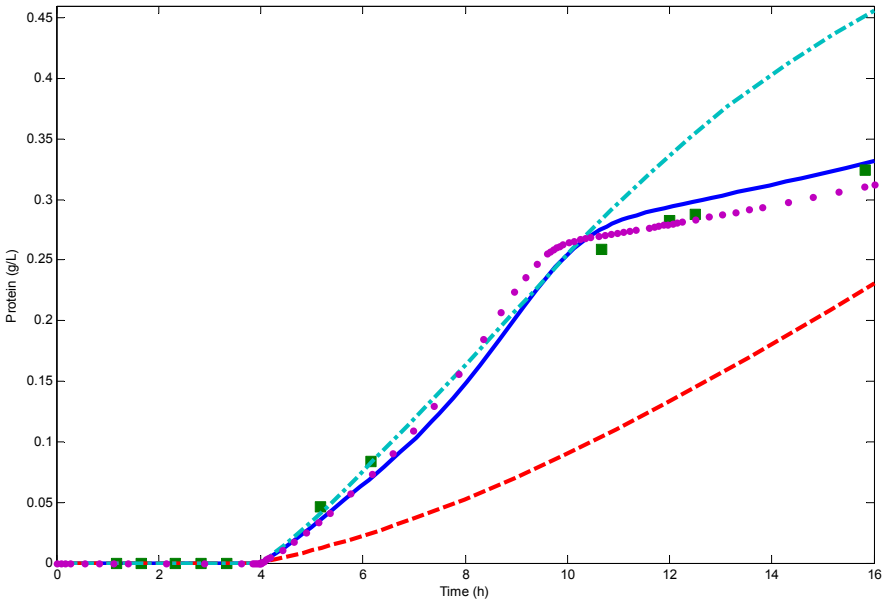

Figure 5. Predicted trajectories for 1st order (- -),Monod (-), Contois (.-) and Manteinance (..) models estimations for state variables and "experimental data" from *in silico* bioreactor run (squares) applying the optimal policy in 2nd iteration.

3rd iteration:

Sampling times:

$\psi_{3rd}[h] = [2.33, 4.33, 7.17, 7.67, 8.17, 8.67, 9.17, 9.67, 13.17, 13.67, 14.33, 14.83]$

$M = 1.74 \cdot 10^{-34}$

Table 12. Normalized Global Sensitivity Indices for models parameters

	S_i^n (3rd iter)			
	Model			
Parameter	1st Order	Monod	Contois	Maintenance
$\mu_{\max}$	0.0003	0.0053	0.0008	0.0157
K_s	0.3514	0	-	0.0132
Y_{xs}	0	0.0721	0	0.0174
$K_P^{\max}$	0.6465	0.9131	0.0242	0.9455
KI_p	0.0017	0.0095	0.0001	0
K_x	-	-	0.9750	-
m	-	-	-	0.0083

Table 13. Parameters re-estimation for the 3rd iter

Model	$K_P^{\max}$	K_x
1st Order	0.104±0.002	-
Monod	0.0528±0.0011	-
Contois	-	1.131±0.008
Maintenance	0.0583±0.0011	-

Table 14. Errors for different models (3rd iter)

Model	$E_{\hat{J}_i \to \langle \hat{J} \rangle}$	$E_{\langle \hat{J} \rangle \to J}$	E_{total}
1st Order	$4.87\ 10^{-4}$	0.0465	0.0470
Monod	$5.24\ 10^{-4}$	0.0029	0.0034
Contois	$1.15\ 10^{-4}$	0.0020	0.0021
Maintenance	$5.24\ 10^{-4}$	0.0117	0.0119

4^{th} iteration:

Sampling times:

$\psi_{4th}[h] = [0.5, 1.17, 1.67, 2.17, 3.17, 4.17, 4.67, 5.17, 8.33, 10.83, 11.33, 15.17]$

$M = 9.31 \cdot 10^{-33}$

Table 15. Normalized Global Sensitivity Indices for models parameters

	S_i^n (4th iter)			
	Model			
Parameter	1st Order	Monod	Contois	Maintenance
$\mu_{\max}$	0	0.2405	0.0367	0.0137
K_s	0.9950	0.3614	-	0
Y_{xs}	0.0010	0.3157	0.0185	0.1217
$K_P^{\max}$	0.0033	0.0824	0.9263	0.1336
KI_p	0.0006	0	0	0.3953
K_x	-	-	0.0186	-
m	-	-	-	0.3357

DOI: 10.2202/1934-2659.1298

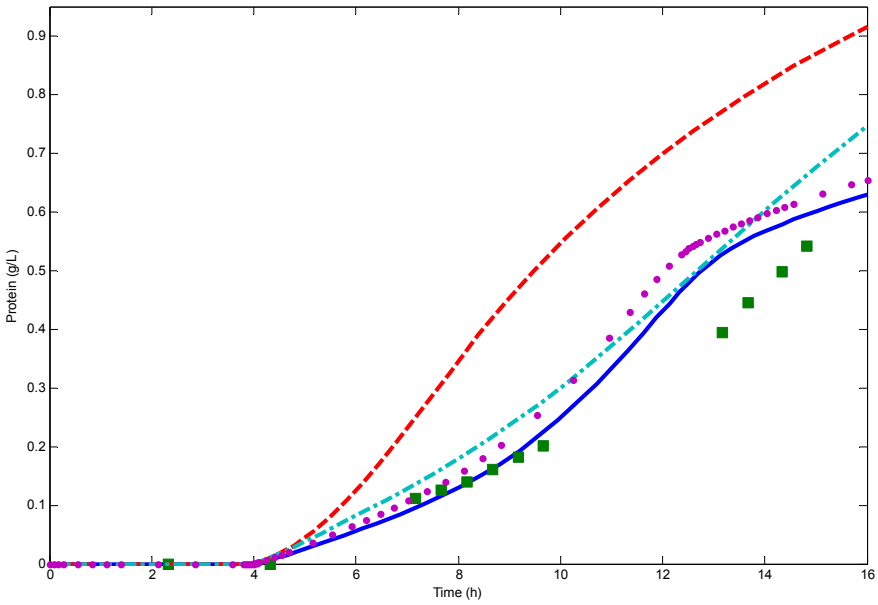

Figure 6. Predicted trajectories for 1st order (- -),Monod (-), Contois (.-) and Manteinance (..) models estimations for state variables and "experimental data" from *in silico* bioreactor run (squares) applying the third optimal policy.

Table 16. Parameters re-estimation for the 4th iter

Model	K_s	Y_{xs}	$K_P^{\max}$	KI_p	m
1st Order	2.95±0.06	-	-	-	-
Monod	1.380±0.008	0.478±0.003	-	-	-
Contois	-	-	0.0413±0.0008	-	-
Maintenance	-	-	-	0.094±0.002	0.00581±0.00012

Table 17. Errors for tendency models after the 4th iter

Model	$E_{\hat{j}_i \to \langle \hat{j} \rangle}$	$E_{\langle \hat{j} \rangle \to J}$	E_{total}
1st Order	$6.41\ 10^{-7}$	0.0279	0.0279
Monod	$6.92\ 10^{-6}$	0.0011	0.0011
Contois	$1.65\ 10^{-5}$	$9.85\ 10^{-4}$	0.0010
Maintenance	$2.01\ 10^{-4}$	0.0045	0.0047

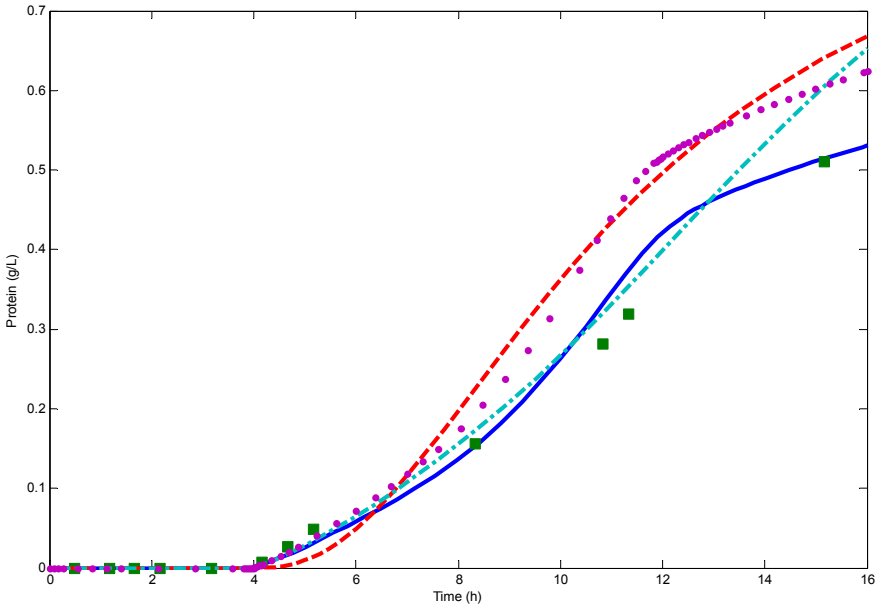

Figure 7. Predicted trajectories for 1st order (- -),Monod (-), Contois (.-) and Manteinance (..) models estimations for state variables and "experimental data" from *in silico* bioreactor run (squares) applying the optimal policy in the 4th iteration.

5th iteration:

Sampling times

$\psi_{5th}[h]=[1.67,3.17,3.67,5.17,9.5,10.0,10.5,11.0,11.5,15.0,15.5,16]$

$M=1.20\cdot10^{-37}$

Table 18. Normalized Global Sensitivity Indices for models parameters

	S_i^n (5th iter)			
	Model			
Parameter	1st Order	Monod	Contois	Maintenance
$\mu_{\max}$	0.1578	0.1793	0.5578	0.0297
K_s	0.3202	0.1429	-	0
Y_{xs}	0.1375	0	0.1518	0.2011
$K_P^{\max}$	0.3845	0.2092	0.1046	0.2353
KI_p	0	0.4686	0	0.4781
K_x	-	-	0.1857	-
m	-	-	-	0.0558

DOI: 10.2202/1934-2659.1298

Table 19. Parameter re-estimation in the 5th iter

Model	$K_P^{\max}$	$\mu_{\max}$	K_s	KI_p
1st Order	0.046±0.003	-	3.70±0.07	-
Monod	-	-	-	0.0752±0.0015
Contois	-	0.472±0.002	-	-
Maintenance	-	-	-	0.119±0.002

Table 20. Errors for different tendency models (5th iter)

Model	$E_{\hat{j}_i \to \langle \hat{j} \rangle}$	$E_{\langle \hat{j} \rangle \to J}$	E_{total}
1st Order	$1.02\ 10^{-4}$	0.0189	0.0190
Monod	$3.59\ 10^{-5}$	0.0064	0.0065
Contois	$1.18\ 10^{-5}$	0.0069	0.0070
Manteinance	$9.38\ 10^{-4}$	0.0218	0.0227

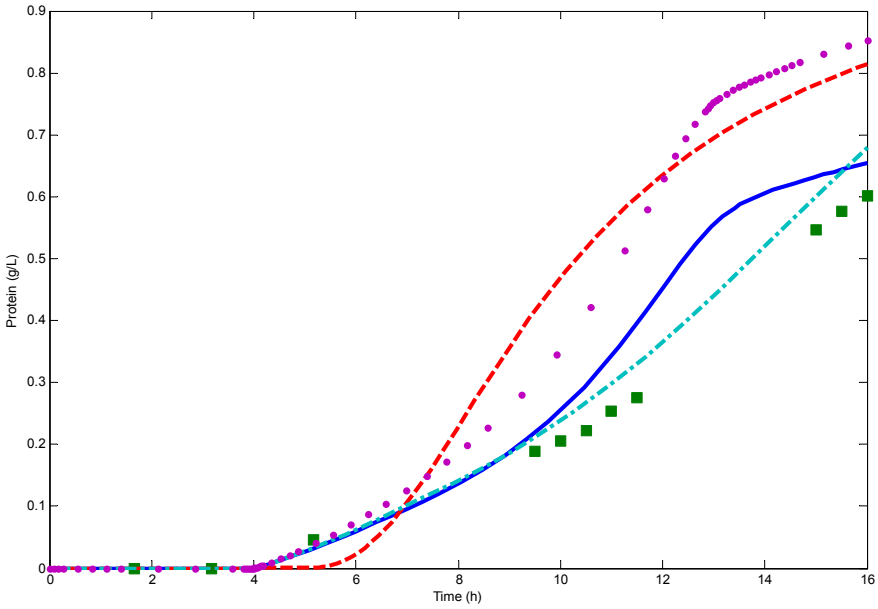

Figure 8. Predicted trajectories for 1st order (- -),Monod (-), Contois (.-) and Manteinance (..) models estimations for state variables and "experimental data" from *in silico* bioreactor run (squares) applying the fifth optimal policy.

4. Final remarks

This paper has presented a systematic procedure for sequential design of dynamic experiments in modeling for optimization using a library of tendency models which allows safe exploration of alternative parameterizations of the input policy to increasingly improve operating conditions. At each iteration, model selection is based on the total error which accounts separately for parametric uncertainty and structural errors. Since tendency models are initially plagued with uncertainty model selection using poorly estimated errors does provide a certain amount of safe exploration which is instrumental for model-based optimization with imperfect models. Global sensitivity analysis has been used to formulate the optimal sampling in each dynamic experiment as an optimization problem whose solution provides the optimal sampling times at which the performance objective is most sensitive to the uncertainty in the selected model parameters. Once new data are available, global sensitivity analysis is used to determine the subset of model parameters that should be re-estimated for each tendency model. Following the update of all tendency models, the input policy is recalculated by dynamic optimization and a new iteration begins. It is worth noting from the results obtained for the case study that after a small number of designed dynamic experiments the input policy has converged to the optimal one corresponding to the *in silico* bioreactor model used for data generation in each experiment. Results obtained in the production of r-interleukin-11 (r-IL-11) case study are very promising and industrial applications in the development of novel specialty-chemicals and bioprocesses are currently underway.

REFERENCES

Asprey, S. M.; Machietto, S. (2000). Statistical tools for optimal dynamic model building. *Computers and Chemical Engineering,* 24, 1261-1267.

Asprey, S.; Machietto, S. (2002). Designing robust optimal dynamic experiments. *J. of Process Control,* 12, 545-556.

Akesson, M.; Hagander, P.; Axelsson J.P. (2001). Avoiding acetate accumulation in *Escherichia coli* cultures using feedback control of glucose feeding, *Biotechnol. Bioeng*. 73 223–230.

Bailey, J. E. (1998). Mathematical Modeling and Analysis in Biochemical Engineering: Past Accomplishments and Future Opportunities. *Biotechnol. Prog*. **14**, 8-20.

Balsa-Canto, E.; Banga, J. R.; Alonso, A. A.; Vassiliadis, V. S. (2001). Dynamic optimization of chemical and biochemical processes using restricted second-order information. *Comput. Chem. Eng*., 25, 539-546.

DOI: 10.2202/1934-2659.1298

Banga, J. R.; Alonso, A. A.; Singh, R. P. (1997). Stochastic dynamic optimization of batch and semicontinuous bioprocesses. *Biotechnol. Prog.,* 13(3), 326-335.

Bastin, G.; Dochain, D. (1986). On-line estimation of microbial specific growth rates. *Automatica*, 22, 705-709.

Bastin, G.; Van Impe, J. F. (1995). Nonlinear and adaptive control in biotechnology: A tutorial. *Eur. J. Control,* 1, 37-53.

Bonvin, D. (1998). Optimal operation of batch reactors: a personal viewpoint. *J. Proc. Control*, 355-368.

Chen, B. H., Asprey. S. P. (2003). On the design of optimally informative dynamic experiments for model discrimination in multiresponse nonlinear situations." *Industrial Engineering Chemical Research*, 42, 1379-1390.

Cinar, A.; Parulekar, S. J.; Ündey, C.; Birol, G. (2003). Batch fermentation: Modeling, Monitoring and Control. Marcel Dekker, New York.

Cooney, C. L. (1983). Bioreactors: design and operation; *Science* **19**, 728-740,.

Crommelin, D.J.A.; Sidelar, R.D. (2002). *Pharmaceutical biotechnology: an introduction for pharmacist and pharmaceutical scientist*; John Wiley & Sons: New York.

Fischer, C. R.;. Klein-Marcuschamer, D; Stephanopoulos, G. (2008). Selection and optimization of microbial hosts for biofuels production. *Metabolic Engineering*, (in press).

Gudi, R. D.; Shah, S. L.; Gray, M. R.; Yegneswaran, P. K. (1997). Adaptive multirate estimation and control of nutrient levels in a fed-batch fermentation using off-line and on-line measurements. *Can. J. Chem. Eng.*, 75, 562-573.

Lim, H. C.; Tayeb, Y. J.; Modak, J. M.; Bonte, P. (1986). Computational algorithms for optimal feed rates for a class of fed-batch fermentation: Numerical results for penicillin and cell mass production. *Biotech. Bioeng*., 28, 1408-1420.

Lee, J.; Ramirez, W. F. (1994). Optimal fed-batch control of induced foreign protein production by recombinant bacteria. *AIChE J*., 40(5), 899-907.

Luli, G.W.; Strohl, W.R. (1990). Comparison of growth, acetate production, and acetate inhibition of *Escherichia coli* strains in batch and fed-batch fermentations, *Appl. Environ. Microbiol*. 56 1004–1011.

Luus, R. (1990). Application of dynamic programming to high-dimensional nonlinear optimal control problems. *Int. J. Control*, 52(1), 239-250.

Luus, R. (1993). Optimization of fed-batch fermentors by iterative dynamic programming. *Biotech. Bioeng*., 41, 599-602.

Mahadevan, R.; Doyle III, F. J. (2003). On-Line optimization of recombinant product in a fed-batch bioreactor. *Biotechnol. Prog.*, 19, 639-646.

Martinez, E. C. (2000). Batch process modeling for optimization using reinforcement learning, *Comput. Chem. Eng.* 24, 1187-1193.

Martínez, E. C.; Wilson, J. A. (2003). Evolutionary optimization of batch process systems using imperfect models, *Proc. Indian Natn. Sci. Acad*., 403-428.

Park, S.; Ramirez, W. F. (1988). Optimal production of secreted protein in fed-batch bioreactors. *AIChE J*., 34, 1550-1558.

Powell, W. B. (2007). *Approximate dynamic programming*. John Wiley & Sons: New Jersey.

Rahman, S.; Palanki, S. (1996). On-line optimization of batch processes in the presence of measurable disturbances. *AIChE J*., 42(10), 2869-2882.

Ramkrishna, D. (2003). On bioreactor modeling for control. *J. Process Control* **13**, 581-589.

Saltelli, A.;Tarantola, S.; Campolongo, F.; Ratto, M. (2004). *Sensitivity Analysis in Practice: A Guide to Assessing Scientific Models*. John Wiley & Sons Ltd, Chichester, England.

Saltelli A., Ratto M.; Tarantola S.; Campolongo F. (2006). Sensitivity Analysis Practices. Strategies for Model-Based Inference, *Reliability Engineering & System Safety*, 1109-1125.

San, K.; Stephanopoulos, G. (1989). Optimization of fed-batch pencillin fermentation: A case of singular optimal control with state constraints. *Biotech. Bioeng*., 34, 72-78.

Sidoli, F. R.; Mantalaris, A.; Asprey, S. P. (2005). Toward global parametric estimability of a large-scale kinetic single-cell model for mammalian cell cultures. Ind. Eng. Chem. Res., 44, 868-878.

Sobol', I. M. (1993). Sensitivity analysis for non-linear mathematical models. *Mathematical Modelling & Computational Experiment*, 407-414.

Tang, S.; Chen, J.; Zhang, Z. (2007). Structured models for recombinant human interleukin-11 fermentation. *Biochemical Engineering J.*, 35, 210-217.

Tholudur, A.; Ramirez, W. F. (1996). Optimization of fed-batch bioreactors using neural network parameter function models. *Biotechnol. Prog*., 12, 302-309.

Visser, D.; van der Heijden, R.; Mauch, K.; Reuss, M.; Heijnen, S. (2000). Tendency Modeling: A New Approach to Obtain Simplified Kinetic Models of Metabolism Applied to Saccharomyces cerevisiae. *Metabolic Engineering*, 252-275.

Walsh, G. (2007). *Pharmaceutical biotechnology: concepts and applications*; John Wiley & Sons Ltd, Chichester, England.

Walter, E., Ed. (1987). *Identifiability of Parametric Models*. Pergamon Press: Oxford, U.K.

Walter, E.; Pronzato, L. (1996). On the identifiability and distinguishability of nonlinear parametric models. *Math. Comput. Simul*., 42, 125-134.

DOI: 10.2202/1934-2659.1298