Good Modeling Practice for PAT Applications: Propagation of Input Uncertainty and Sensitivity Analysis

Gürkan Sin and Krist V. Gernaey
Dept. of Chemical and Biochemical Engineering, Technical University of Denmark, DK-2800 Kgs. Lyngby, Denmark

Anna Eliasson Lantz
Center for Microbial Biotechnology, Dept. of Systems Biology, Technical University of Denmark, DK-2800 Kgs. Lyngby, Denmark

DOI 10.1021/bp.166
Published online June 30, 2009 in Wiley InterScience (www.interscience.wiley.com).

The uncertainty and sensitivity analysis are evaluated for their usefulness as part of the model-building within Process Analytical Technology applications. A mechanistic model describing a batch cultivation of Streptomyces coelicolor *for antibiotic production was used as case study. The input uncertainty resulting from assumptions of the model was propagated using the Monte Carlo procedure to estimate the output uncertainty. The results showed that significant uncertainty exists in the model outputs. Moreover the uncertainty in the biomass, glucose, ammonium and base-consumption were found low compared to the large uncertainty observed in the antibiotic and off-gas CO_2 predictions. The output uncertainty was observed to be lower during the exponential growth phase, while higher in the stationary and death phases - meaning the model describes some periods better than others. To understand which input parameters are responsible for the output uncertainty, three sensitivity methods (Standardized Regression Coefficients, Morris and differential analysis) were evaluated and compared. The results from these methods were* mostly *in agreement with each other and revealed that only few parameters (about 10) out of a total 56 were mainly responsible for the output uncertainty. Among these significant parameters, one finds parameters related to fermentation characteristics such as biomass metabolism, chemical equilibria and mass-transfer. Overall the uncertainty and sensitivity analysis are found promising for helping to build reliable mechanistic models and to interpret the model outputs properly. These tools make part of good modeling practice, which can contribute to successful PAT applications for increased process understanding, operation and control purposes.* © 2009 American Institute of Chemical Engineers *Biotechnol. Prog.,* 25: 1043–1053, 2009

Keywords: uncertainty, sensitivity, first principles modeling, PAT, antibiotic production, Morris screening, standardized regression coefficients

Introduction

Model-building is typically preceded by performing process characterization and validation experiments whose results are interpreted during construction of the process model. The end result is a mechanistic model describing behavior of the process in question. This activity is frequently used in life sciences for developing an improved understanding of microbiological processes.[1–4] The resulting mechanistic model can be considered as a summary of available process knowledge and understanding of a system. Process understanding is precisely what is asked for in the recent Process Analytical Technology (PAT) initiative introduced in the pharmaceutical industry.[5] Moreover, the PAT initiative also seeks active exploitation of that process knowledge to reach an optimised process operation. Modeling can thus be perceived as a means for reaching an end destination such as the creation of improved process understanding in combination with optimised and stable process operation.

However, considering the complexity of a typical mechanistic model of a microbiological process, good modeling practice (GMoP) is a necessity to minimize the risk of errors in the model-building process or in the subsequent application and interpretation of the model. GMoP can be understood broadly as a set of mathematical and statistical tools that allow improvement of the usage and the reliability of the model. This article is about mechanistic modeling, but contrary to traditional modeling articles that describe the development of a specific model or class of models, this article has focused on some of the GMoP tools. More specifically, the article will evaluate the usefulness of applying a number of uncertainty and sensitivity analysis techniques as part of the model-building process within PAT applications. These uncertainty and sensitivity analysis techniques are generally applicable in a broad range of engineering and scientific disciplines.[6–8]

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

Correspondence concerning this article should be addressed to G. Sin at gsi@kt.dtu.dk.

Table 1. Expert Review of Uncertainty of Input Parameters

Uncertainty Class	Variation Factor(%)*	Parameters
Low	5	Y_{SX}, MW_X, MW_S, γ_S, MW_{act}, γ_{ACT}, i_{NRED}, MW_{red}, γ_{RED}, T, pH, pK_{NH}, pK_{HCO_3} $pK_{H_2PO_4}$, pKw, $k_{f,NH}$, k_{f,CO_2}, k_{f,H_2PO_4}, $k_{f,W}$, R, and V_L
Medium	25	f_{XI}, i_{NX}, i_{PX}, γ_X, μ_{max}, m_S, k_d, t_{lag}, $K_La_{O_2}$, P_{in}, P_{O_2}, P_{CO_2}, P_{out}, D_{O_2}, D_{CO_2}, D_{N_2}, K_{HO_2}, K_{HCO_2}, K_{HN_2},V_G, and Q_{Gin}
High	50	K_S, K_O, K_N, K_P, Y_{SACT}, β_{ACT}, α_{ACT}, S_{ACT}^{max}, Y_{SRED}, β_{RED}, α_{RED}, S_{RED}^{max}, S_{NH3}^{*}, and $K_La_{MH_3}$

* Stands for the maximum percent of deviation from the mean value of a parameter that is allowed when sampling the parameters in Step 2, where the deviation is allowed to be positive or negative.

Uncertainty associated with the predictions of simulation models is generally classified as: (a) *stochastic uncertainty* that arises from stochastic components of a simulation model required to describe a stochastic system (e.g., random failure event of the aeration system in a fermenter); (b) *subjective (or input) uncertainty* that represents incomplete knowledge about the fixed values used as input to the model (e.g., process kinetics and stoichiometry related parameters, physical–chemical constants, operational parameters); and (c) *structural uncertainty* that relates to the mathematical formulation or the model structure, because most model structures are a form of approximation of a physical system rather than being its carbon copy.[6,7] In this study, the focus was specifically placed on the input uncertainty (b), although the discussion of the results is extended to also reflect on the impact of other classes of uncertainty (a and c) when evaluating the prediction uncertainty of fermentation models.

Several uncertainty analysis methods are available. For example, Bayesian analysis in combination with evolutionary optimization algorithms is emerging as a feasible method for performing uncertainty analysis in complex numerical models.[9] In this study, however, the engineering standard Monte Carlo procedure is chosen to propagate and analyse the input uncertainty, because the method is generally accepted as computationally-effective and reliable.[7]

Sensitivity analysis is complimentary to uncertainty analysis and can be viewed as an analysis of variance, in which the aim is to decompose the output variance with respect to input parameters.[7,8] Such information is valuable as it reveals which input parameters explain most of the variance in model predictions, thereby providing a parameter significance ranking among others. Sensitivity analysis-based parameter ranking is very attractive and highly relevant for process characterization and validation purposes—an important step in PAT applications indeed. This is because such a ranking can be exploited to direct experimental work towards reducing the uncertainty in the most influential parameters, instead of wasting experimental efforts on parameters that have little or no influence on the model output. In this study, three commonly used sensitivity analysis techniques are evaluated for sensitivity analysis of the input parameters, (i) differential analysis,[7,10] (ii) Standardised Regression Coefficients (SRC)[7,8] and (iii) Morris screening.[11]

The novelty in this article is the fact that uncertainty analysis is applied in the model-building phase. Indeed, the above mentioned GMoP tools for uncertainty and sensitivity analysis were evaluated on a complex dynamic model (56 parameters) that describes biological, physical, and chemical processes occurring in batch cultivations of *S. coelicolor* for antibiotic production.[4] The article is organised as follows. First the model, the uncertainty and sensitivity analysis techniques are introduced. Then, the uncertainty analysis results are presented followed by the sensitivity analysis results. The results of the uncertainty and sensitivity analysis techniques are critically discussed and compared.

Methods

Dynamic model of S. coelicolor *cultivation*

The model was developed in a previous study by Sin et al.[4] within a PAT application to a fermentation process. In this study, we focused exclusively on process characterization and identification for predicting dynamically all monitored variables including glucose, oxygen, biomass, actinorhodin, base-addition, and off-gas CO_2 during a batch cultivation lasting 150 h.[4] The model contains 56 parameters relating to biological, chemical equilibrium, and gas–liquid exchange processes. Among these parameters, 12 were readily available from routine experimental measurements, 10 parameters were determined by applying model identification (least-squares method),[4] and the remaining ones were adopted from relevant literature sources.[4] The complete list of the parameters together with their description and nominal values are also provided in the supporting material (**S1**).

Uncertainty analysis

Monte Carlo Procedure. For notational convenience, the dynamic model structure is represented by f, the output vector by y, the state vector by x, the input parameter vector by θ and time is represented by t:

$$\frac{dx}{dt} = \mathbf{f}(\mathbf{x}(t), t, \boldsymbol{\theta})$$
$$\mathbf{y}(t) = \mathbf{g}(\mathbf{x}(t)) \quad (1)$$

Monte Carlo analysis of uncertainty involves three steps: (1) specifying input uncertainty (2) sampling input uncertainty, and (3) propagating the sampled input uncertainty through **f** to obtain prediction uncertainty for **y**.

Step 1. In this study, the subjective input uncertainty was defined after an expert review process, which is typically performed in this field.[7,10] Expert review involves asking the opinion of process experts (and/or consulting the relevant literature resources) about the uncertainty of the parameters (in this case related to the cultivation of *S. coelicolor*), e.g., what is the upper and lower bound of a kinetic parameter. To structure the expert review process, all the model parameters were assumed to have a uniform probability distribution. Moreover, three classes of uncertainty, namely low, medium and high, were defined that correspond to 5, 25, and 50% of variability around the mean values (see Table 1). This way of input uncertainty characterization is similar to the approach of Brun et al.[10] and is typically performed in

this field. In this way, the minimum and maximum values of the uniform distribution can be calculated as follows:

$$\theta_{\min} = (1 - \%\text{Variation}) * \theta_{mean}$$
$$\theta_{\max} = (1 + \%\text{Variation}) * \theta_{mean} \quad (2)$$

For example, the chemical composition of substrate (i.e., glucose) is known with great accuracy, and hence the molecular weight and degree of reduction parameters were classified as low uncertainty (see Table 1). Likewise from microbiological studies the biomass growth yield, $Y_{SX,}$ of aerobic cultures of bacteria is relatively well known[1]; hence, it was also classified as low uncertainty. To continue the example, given a mean value for the yield (0.46 c-mol/c-mol),[4] the corresponding minimum and maximum values of Y_{SX} were determined as follows:

$$Y_{SX,\min} = (1 - \%5) * 0.46 = 0.44$$
$$Y_{SX,\max} = (1 + \%5) * 0.46 = 0.48$$

On the other hand, the kinetic parameters of antibiotic production are known with less accuracy,[4,12] and hence these parameters were classified in the high uncertainty group. By applying this line of reasoning to the remaining model parameters, the characterization of input uncertainty was completed (Table 1).

Step 2. The well-known Latin Hypercube Sampling method[13] was used for probabilistic sampling of the input space. In total, 500 samples were selected from the input parameter space, where each sample, $\boldsymbol{\theta}_i$, contains one value for each input parameter:

$$\boldsymbol{\theta}_i = [\theta_{1i}, \theta_{2i}, \theta_{3i} \ldots\ldots \theta_{Mi}], \quad \text{for } i = 1, 2 \ldots\ldots N \quad (3)$$

Where M stands for the total number of model parameters (56), N is the total number of Latin-Hypercube samples (e.g., 500), and $\theta_{1,i}$ represents the samples for the first input parameter. The correlation between the input parameters was induced using the rank-based method of Iman and Conover.[14] The correlation matrix between the input parameters was available for 10 parameters of the model (provided as supporting material, **S2**), whereas for the remaining parameters, no correlation was assumed as no previous information was available. The sampling results for some input parameters are also provided as supporting material (**S2**).

Step 3. The sampled input matrix, $\boldsymbol{\theta}_{N\times M}$, was propagated through **f** (the dynamic model) by performing N dynamic simulations, i.e., one simulation for each input parameter sample resulting from Step 2. These simulations resulted in a three dimensional matrix $\mathbf{y}_{T\times K\times N}$ that contained T time instants (15 minutes interval between 0 and 150 h) predictions of K output variables (the model contains 18 variables) and for N Latin Hypercube Samples (500 samples). The complete Monte Carlo results provided a cumulative distribution function for each output variable at each time instant. The uncertainty of the model outputs was then represented using simple statistics such as mean and percentile calculations.

Sensitivity analysis techniques

Differential Analysis. In this method, the first-order derivative of the model outputs was taken with respect to model inputs to assess the effect of each input parameter on the output[10]:

$$s_{jk}(t) = \left.\frac{\partial y_k(t)}{\partial \theta_j}\right|_{\boldsymbol{\theta}_0} \text{ and } ns_{jk}(t) = \frac{sc_j}{sc_k} s_{jk}(t) \quad (4)$$

Where $s_{jk}(t)$ is the dynamic sensitivity function of input parameter, θ_j, on the output y_k. $\frac{\partial y_k}{\partial \theta_j}$ is the first-order derivative of the output y_k with respect to the input parameter. θ_j and $\boldsymbol{\theta_0}$ represents the nominal values of the input parameters at which the model, **f**, was evaluated. Further, $ns_{jk}(t)$ is the non-dimensional sensitivity function obtained by scaling the dynamic sensitivity function with appropriate scaling factors for the input and the output, sc_j and sc_k, respectively. In this study, we used the standard deviations of the inputs and the outputs as sc_j and sc_k. The significance of the contribution of the input parameter to the output variance can then be assessed by aggregating the dynamic sensitivity functions, e.g., as mean-squared, $\boldsymbol{\delta}_{jk}^{\text{msqr}}$:

$$\delta_{jk}^{\text{msqr}} = \sqrt{\frac{1}{T}\sum_{t=1}^{T} ns_{jk}^2(t)} \quad (5)$$

Standardised Regression Coefficients. The standardised regression coefficients were obtained by constructing linear regression models on the outputs obtained from the Monte Carlo procedure (see above). Because this technique requires scalar output, a certain aggregate property of dynamic model outputs were used, e.g., mean. The scalar model output matrix was denoted as sy and has the dimensions of KxN. With that output matrix, one regression model was built for each model output k:

$$sy_{ik} = b_{0k} + \sum_{j=1}^{M} b_{jk}\theta_{ij} + \varepsilon_{ik} \quad \text{for } i = 1, 2, \ldots, N \quad \text{and for } k = 1, 2, \ldots., K \quad (6)$$

sy_{ik} is the scalar value for the kth output, b_{jk} is the coefficient of the jth input parameter, θ_j, for the kth output, θ_{ij} is the value of the jth parameter and ε_{ik} is the error of the regression model. The above equation can also be written in a dimensionless form by scaling the outputs and the parameters using their corresponding mean and standard deviations[15]:

$$\frac{sy_{ik} - \mu_{sy_k}}{\sigma_{sy_k}} = \sum_{j=1}^{M} \beta_{jk} \cdot \frac{\theta_{ij} - \mu_{\theta j}}{\sigma_{\theta j}} + \varepsilon_{ik} \quad (7)$$

β_{jk} is called the standardized regression coefficient (SRC) of the jth input parameter, θ_j, for the kth model output, y_k, and its magnitude relates to how strong the input parameter contributes to the output.

Morris Screening. The method, named after its developer Morris,[11] relies on estimating the distribution of the elementary effects (EE) of each input parameter on the kth model output called EE_{jk}. This distribution function is denoted as F_{jk}, which stands for the distribution of the effects of the jth input parameter on the kth output. Because this method also requires scalar values, the scalar model output matrix was used as data for the Morris screening, i.e., sy. The EE_{jk} attributable to each input parameter was obtained from the following differentiation of model output, sy_k, with respect to the input, θ_j:

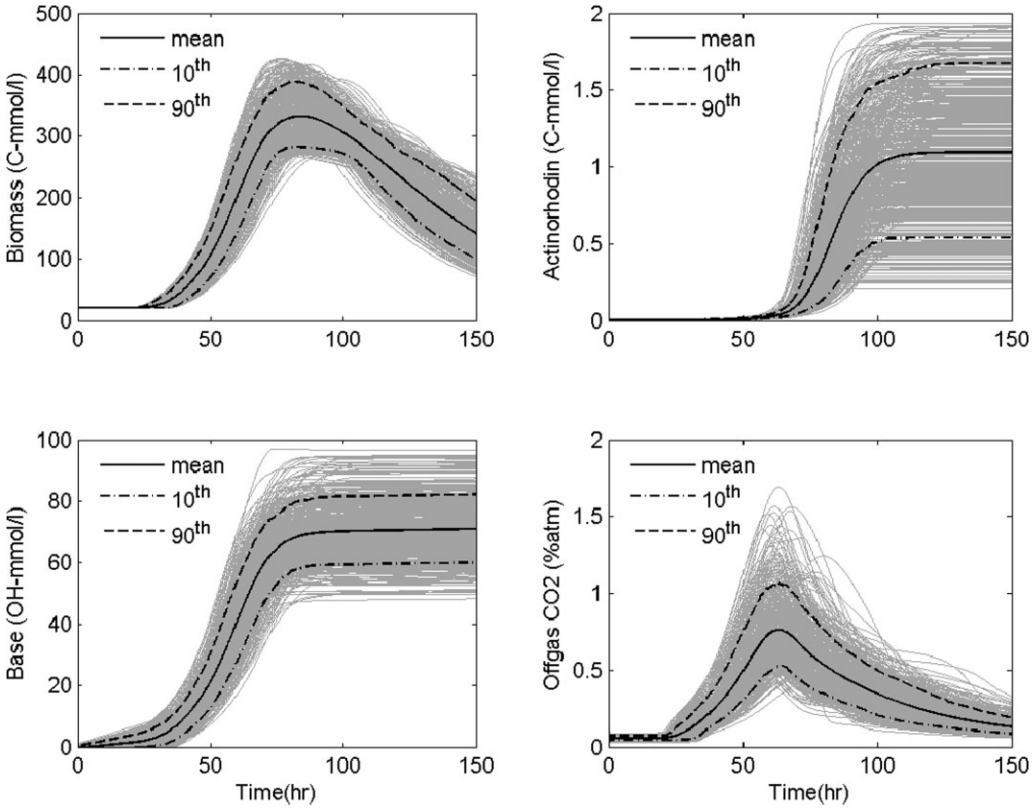

Figure 1. Representation of uncertainty in the model predictions for biomass, actinorhodin, base addition, and off-gas CO_2: Monte Carlo simulations (grey), mean and 10th and 90th percentile of the predictions.

$$\mathrm{EE}_{jk} = \frac{\partial sy_k}{\partial \theta_j} = \frac{sy_k(\theta_1, \theta_2, \theta_j + \Delta, \ldots, \theta_M) - sy_k(\theta_1, \theta_2, \theta_j, \ldots, \theta_M)}{\Delta} \quad (8)$$

Where Δ is a predetermined perturbation factor of θ_j, $sy_k(\theta_1, \theta_2, \theta_3, \theta_j, \ldots, \theta_M)$ is the scalar model output evaluated at input parameters $(\theta_1, \theta_2, \theta_3, \theta_j, \ldots, \theta_M)$, whereas sy_k $(\theta_1, \theta_2, \theta_3, \theta_j+\Delta, \ldots, \theta_M)$ is the scalar model output corresponding to a Δ change in θ_j. Each input parameter, θ_j, can only take values corresponding to (a predefined set of) p levels from its range (imagine a grid in which the range of each parameter is subdivided into p levels). The value for p could be 4, 6, and 8 which corresponds to the 25th, 17th, and 12.5th percentile of the uniform distribution of the inputs.[11,16] These elementary effects could be (i) negligible or zero, (ii) a constant function of x_i, (*iii*) a *non-constant* function of x_i or (iv) a *non-constant* function of more than one factor. These effects can be identified by analyzing the mean and standard deviation of the distribution function, F_{jk}.

The F_{jk} was estimated by performing calculations of the elementary effects, EE_{jk}, at randomly sampled points in the input space, and this procedure was repeated a number of times, r. To do that, an effective one-factor-at-a-time (OAT) design has been proposed by Morris.[11] In this design, calculation of one elementary effect for each input requires (M + 1) model simulations. Because a number of repetitions, r, is needed (typically 10–50), the total number of model simulations needed becomes r * (M + 1). In summary, the Morris method has three degrees of freedom that need to be specified, Δ, p, and r, which for this study were defined as 2/3, 4, and 35, respectively. The results of the Morris sampling of the input space are provided and discussed in the supporting material section, (**S3**).

Modeling and simulation platform

The dynamic fermentation model was implemented and simulated in Matlab (Mathworks, R14). All the above mentioned methods for performing the uncertainty and sensitivity analysis were also implemented in Matlab as m script files and can be provided upon request.

Results and Discussion

Uncertainty in the model outputs: Monte Carlo simulations

The Monte Carlo simulations obtained by simulating the 500 LHS samples using the dynamic fermentation model resulted in 500 time-series (dynamic profiles) for each model output. Although the model has in total 19 output variables, in this study we focused on the typically monitored/measured variables in fermentation technology which are glucose, oxygen, ammonium, phosphate, biomass, antibiotic products (actinorhodin), base addition, and off-gas CO_2. The raw data obtained from the Monte Carlo simulations were plotted in Figures 1 and 2. In general, the simulation results indicate typical batch fermentation curves consisting of a lag, exponential, stationary, and death phase. This general profile can be particularly discerned for the biomass profile in Figure 1. In particular, each simulated fermentation profile is different as different parameter values were used in each simulation. Overall, one observes that each model output has a time

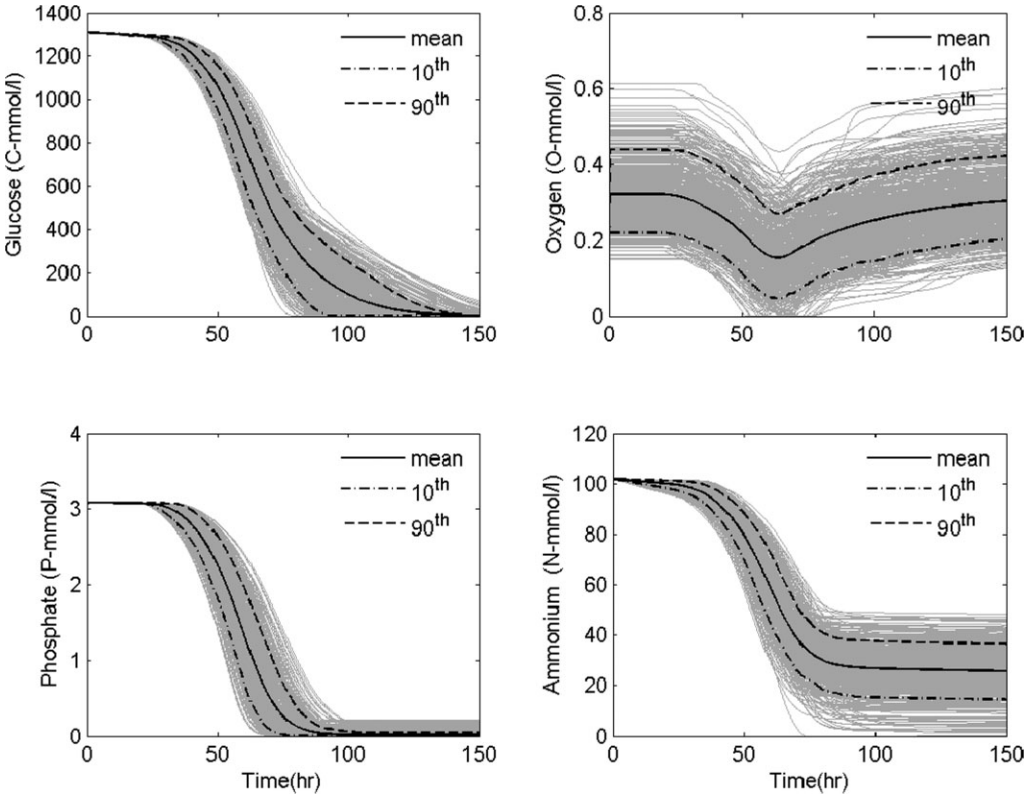

Figure 2. Representation of uncertainty in the model predictions for glucose, oxygen, ammonium, and phosphate: Monte Carlo simulations (grey), mean and 10th and 90th percentile of the predictions.

varying band profile (visually speaking). Statistically speaking the bands observed on these spaghetti plots correspond to the distribution of the model outputs at each time instant.

The uncertainty in the model outputs can also be represented using mean, 10th, and 90th percentile of the distribution of each model output at each time instant (see bold dashed and full lines in Figures 1 and 2). In general, the larger the spread of the distribution of an output (i.e. the width of the output range at each time instant), the higher the uncertainty is. The output uncertainty shown in Figures 1 and 2 reflects the uncertainty in the inputs, which were characterised as uniform distributions with different degree of uncertainty (see Table 1). This is often called mapping input uncertainty onto output uncertainty.

Overall, one observes that uncertainty exists in model outputs. The extent of uncertainty was different on different outputs, e.g., the uncertainty on biomass, glucose, phosphate, ammonium, and base addition predictions was relatively smaller compared with the uncertainty on the predictions of oxygen, off-gas CO_2, and actinorhodin production. Furthermore, for some of the outputs, the uncertainty was observed to be changing over time during the batch operation. This was particularly clear for the ammonium concentration and base addition predictions, where the spread of the uncertainty band increased (almost doubled) after about 75 h of operation (see Figure 1 bottom-left and Figure 2 bottom-right). This point of time in the batch operation roughly corresponded to the end of the exponential growth phase and the start of the stationary/decay phases. In short, this indicated that the model was able to provide rather reliable predictions of the batch system for some variables during the lag and exponential phases, but its certainty decreased when predicting the transition period from the exponential phase to the stationary/decay phases.

To understand what underlies this output uncertainty and to be able to address the above-mentioned observation that some outputs have more uncertainty than the others, one needs to perform a sensitivity analysis—which basically decomposes the output uncertainty with respect to the input characteristics of the fermentation process.

Sensitivity analysis

Standardised Regression Coefficients. The scalar outputs required for the calculation of SRCs were obtained by using the mean of the time series profiles for each model output. This choice was made on purpose for two particular reasons: (a) to check the sensitivity of the process characteristics on the average behavior of the process and, (b) to keep the study focused on comparing three sensitivity analysis methods using one type of scalar product (since otherwise one may easily get side-tracked by the technical details of each method). For each of these outputs, a linear regression model was constructed using Eq. 7 and the corresponding regression coefficients – SRCs – were obtained from linear least-squares. The SRCs were then ranked for each output, and a summary of the ranking is given in Table 2.

The degree of linearization indicated by the coefficient of model determination, R^2, was found to be high for all outputs except for actinorhodin (see the first row in Table 2), which was below the recommended value 0.7.[15] This indicated that the linearized model was able to explain most of the variance in five of the outputs in Table 2, and hence, the corresponding coefficients can reliably be used to assess the

Table 2. Standardized Regression Coefficients (SRCs): The 10 Highest Ranking Parameters on Six Model Outputs

	Glucose		Oxygen		Biomass		Actinorhodin		Base		Off-gas CO_2	
R^2	0.83		0.98		0.88		0.63		0.86		0.90	
Rank	θ	SRC*	θ	SRC*	θ	SRC*	θ	SRC*	θ	SRC*	θ	SRC*
1	K_P	0.74	P_{in}	0.60	i_{PX}	−0.75	K_P	−0.87	i_{PX}	−0.90	P_{in}	−0.55
2	μ_{max}	−0.65	P_{O_2}	0.58	K_P	−0.59	μ_{max}	0.79	i_{NX}	0.70	P_{out}	0.53
3	i_{PX}	0.61	K_{HO_2}	0.58	K_S	−0.50	S_{ACT}^{max}	0.56	K_P	−0.65	Q_{Gin}	−0.47
4	K_S	0.50	K_P	0.12	μ_{max}	0.42	K_S	−0.33	μ_{max}	0.54	K_P	−0.40
5	m_S	−0.47	$K_La_{O_2}$	0.09	$pK_{H_2PO_4}$	0.24	i_{PX}	0.32	pK_{NH_4}	−0.44	μ_{max}	0.33
6	t_{lag}	0.18	μ_{max}	−0.09	k_d	−0.19	t_{lag}	−0.26	K_S	−0.37	pK_{HCO_3}	0.23
7	$pK_{H_2PO_4}$	−0.16	i_{PX}	0.06	m_S	0.15	β_{ACT}	0.26	m_S	0.26	i_{PX}	−0.21
8	K_{HO_2}	0.06	Y_{SX}	0.04	P_{O_2}	−0.10	$pK_{H_2PO_4}$	0.17	$K_La_{NH_3}$	0.16	P_{O_2}	−0.18
9	P_{O_2}	0.06	γ_X	0.03	P_{in}	−0.10	Y_{SX}	0.09	$pK_{H_2PO_4}$	0.15	$pK_{H_2PO_4}$	0.14
10	$K_La_{O_2}$	0.06	$pK_{H_2PO_4}$	−0.03	$K_La_{O_2}$	−0.08	K_{HCO_2}	0.07	t_{lag}	−0.10	m_S	0.13

*The coefficients may take values between 0 and ±1. The larger the magnitude of a coefficient, the more significant the parameter. The signs indicate a negative or a positive impact of the parameter on the outputs respectively.

importance of the input parameters on the outputs. However, the coefficients obtained for actinorhodin should be considered with caution as the unexplained fraction of the output variance was significantly high. Thus, for this output the Morris screening results will be used (see below).

Looking closer at the classes of the 10 highest ranking parameters for each output, one observes that there were some inputs that are influential on all outputs. Among these common parameters, the phosphate dynamics related parameters (phosphorus content of biomass, i_{PX}, the equilibrium constant of phosphate buffer, $pK_{H_2PO_4}$ and the Monod affinity constant for phosphate, K_P) appeared to have consistently large influence on all of the outputs. This is fully understandable because the batch fermentations were simulated to operate under phosphate limited conditions, and therefore all the model predictions were very sensitive on assumptions concerning phosphate uptake kinetics. Further, the maximum specific growth rate of biomass, μ_{max}, also appeared to affect all the outputs, indicating the importance of the rate of biomass metabolism in the fermenter. Other parameters such as the maintenance coefficient, m_S and substrate half saturation coefficient, K_S affected most—but not all—of the outputs, indicating that substrate uptake related parameters also had a large impact on the output predictions. The lowest ranking parameters among the 10 parameters presented in Table 2 relate to the glucose uptake dynamics that in the conditions set for this batch simulation first became limited after the phosphate had been completely consumed by the biomass.

On the other hand, some of the input parameters were observed to have individual effects on the outputs. For instance, the equilibrium constant for ammonium dissolution, pK_{NH_4}, and ammonia stripping rate, $K_La_{NH_3}$, affected only the base addition. The airflow rate (Q_{Gin}) and the equilibrium constant of CO_2 dissolution (pK_{HCO_3}) were influential only on the off-gas CO_2 content. Moreover, the gas–liquid mass transfer related parameters, including the composition of the airflow to the reactor (P_{in}, P_{O_2}), were influential both on the dissolved oxygen and off-gas CO_2 content. The latter indicates that the mass-transfer kinetics was rather important in view of the prediction of gas–liquid phase related outputs in this particular operation.

Overall, the results showed that the input parameters both of biological or physical–chemical origin simultaneously affected all of the outputs. As an example, it is not only the biomass physiology related parameters that influence the prediction of a biological output such as biomass but also the parameters related to physical–chemical processes in the system. These results reinforce, therefore, the importance of using an integrated approach (i.e., include biological and physical–chemical processes in one model) when applying mechanistic modeling studies to understand complex interactions in systems such as fermentation.

Morris Screening In this study, the Morris sampling was designed to calculate 35 elementary effects for each input. The Morris sampling performed using r equal to 35 resulted in a sampling matrix, $\boldsymbol{\theta}_{Morris}$, having the dimensions 1995 × 56, in which 56 corresponds to the number of input parameters and 1995 is in fact 35 * (56+1) (see section 2.3 and supporting information S3). Performing model simulations using the sampling matrix, $\boldsymbol{\theta}_{Morris}$, resulted in 1995 time-series profile for each model output. Similar to the SRCs, these dynamic outputs were aggregated by using the mean of time series profiles. Finally, the elementary effects were calculated using Eq. 8, which provides random observations of the distribution function, F_{jk}. For illustration purposes, F_{jk} values were plotted as histograms in Figure 3 for a number of selected parameters. Important to note is that the elementary effects were scaled (using the standard deviation for the inputs and outputs respectively) to allow the comparison of the effects of parameters on different outputs.

The Morris results are evaluated by comparing the mean and the standard deviations of the distribution function, F_{jk}, of each input (Figure 4). In these plots, each circle corresponds to one input of the model (inputs are given without explicit labeling because labeling 56 data points with the parameter names makes the figure too overloaded to read). In addition, there are two lines provided in each plot that correspond to $\text{Mean}_i = \pm 2 \cdot \text{sem}_i$.[11] In this equation, sem_i is the standard error of the mean and estimated as $\text{sem}_i = \text{standard deviation}_i/\sqrt{r}$.[11] The wedge formed by these two lines in each subplot helps interpret the type of effects the input parameters exert on the outputs. Accordingly, if an input parameter (i.e., a circle) lies inside the wedge, then it indicates that its effect on the output is negligible and can be deemed insignificant. On the contrary, if the input parameter lies outside the wedge, then it is said to have a significant effect on the output. We also observe that these significant parameters are involved in non-linear interactions as their standard deviations (see the y-axis in the subplots) are non-zero.[11] Moreover, not a single parameter was found to have a linear effect on the outputs (i.e., no parameter has a zero standard deviation with non-zero mean). This is not surprising as the fermentation models are typically non-linear.

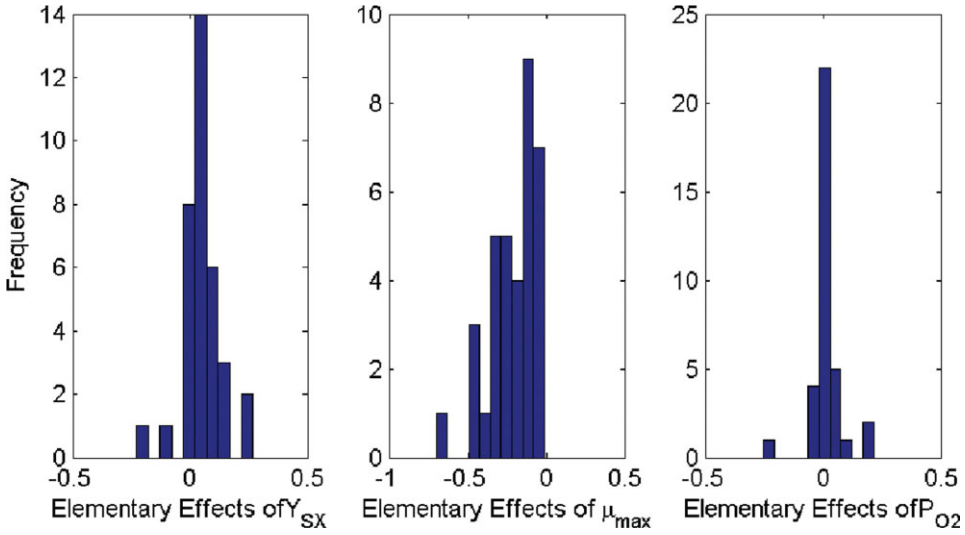

Figure 3. Histogram of the distribution function, F_{jk}: the elementary effects of Y_{SX}, μ_{max}, and P_{O_2} on the glucose output are shown.

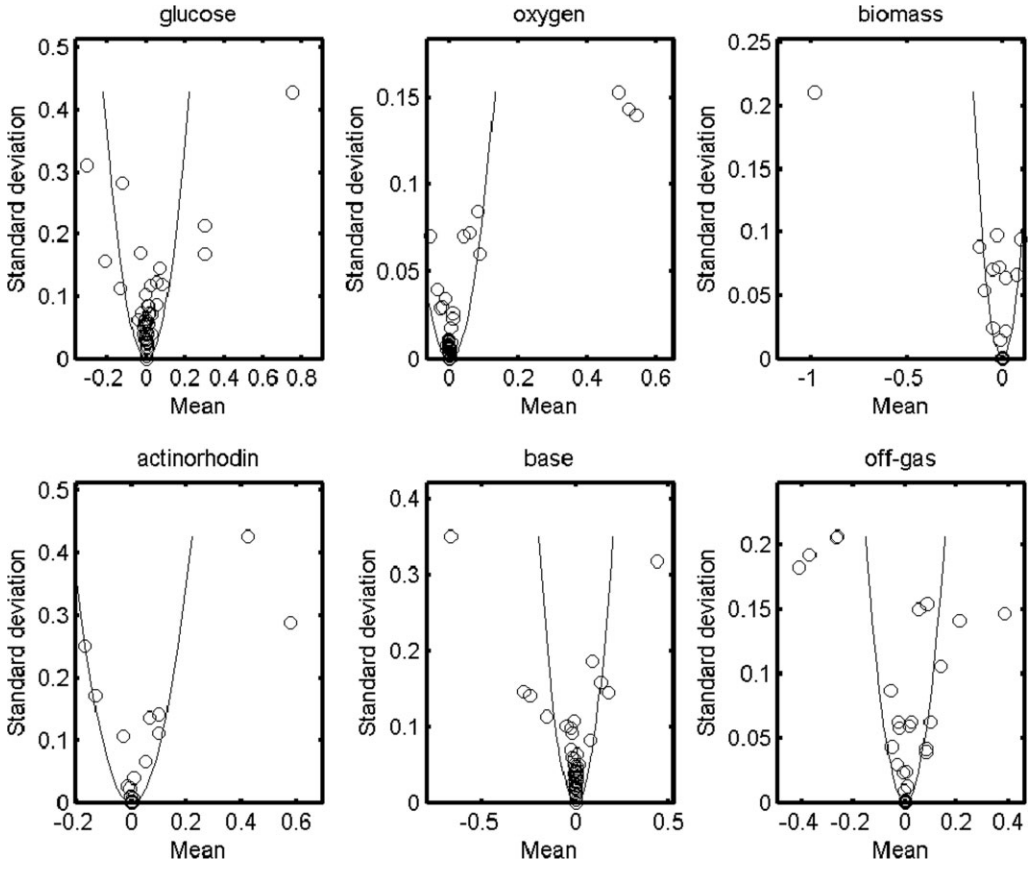

Figure 4. Estimated mean and standard deviation of the distribution of elementary effects of the inputs (in total 56 parameters denoted as anonymous circles) on the model outputs glucose, oxygen, biomass, actinorhodin, base, and off-gas CO_2. The two lines drawn in each subplot correspond to $Mean_i = \pm 2sem_i$ (see text).

Having provided a picture of all the results of Morris screening in Figure 4, the focus will now be on detailed investigation of the first 10 highest ranking parameters for each output (Table 3). The parameter significance ranking was performed according to the absolute mean of the distribution function, F_{jk} – the higher the mean value, the more significant the effect is. The results of the Morris screening of input parameters were found to be in large agreement with the ranking obtained by the SRCs above—albeit that the exact order of importance may be slightly different for both methods. Particularly, all model outputs were sensitive to certain common parameters such as biomass physiology related parameters (i_{PX}, K_P, K_S, μ_{max}). In addition to that, the off-gas predictions (dissolved oxygen and off-gas CO_2) were most sensitive to the parameters related to mass transfer dynamics including airflow rate and composition (Q_{Gin}, P_{in}, P_{O_2}, and $K_La_{O_2}$)—which confirms the findings of the SRCs (see above, explanation with Table 2).

Table 3. Estimated Means of the Distribution of Scaled Elementary Effects of the Inputs—The Effect of the First 10 Highest Ranking Parameters on Six Model Outputs are Shown

	Glucose		Oxygen		Biomass		Actinorhodin		Base		Off-gas CO_2	
Rank	θ	μ^*	θ	μ^*	θ	μ^*	θ	μ^*	θ	μ^*	θ	μ^*
1	i_{PX}	0.75	Pin	0.54	i_{PX}	−0.98	S_{ACT}^{max}	0.58	i_{PX}	−0.67	Q_{Gin}	−0.41
2	m_S	−0.30	K_{HO_2}	0.52	K_P	−0.12	i_{PX}	0.42	i_{NX}	0.44	P_{out}	0.39
3	K_P	0.30	P_{O_2}	0.49	μ_{max}	0.10	K_P	−0.17	pK_{NH_4}	−0.28	P_{in}	−0.37
4	K_S	0.30	$K_La_{O_2}$	0.09	K_S	−0.09	m_S	−0.13	K_P	−0.24	i_{PX}	−0.27
5	μ_{max}	−0.21	K_P	0.08	$pK_{H_2PO_4}$	0.08	β_{ACT}	0.10	μ_{max}	0.18	K_P	−0.26
6	$pK_{H_2PO_4}$	−0.13	i_{PX}	0.06	m_S	−0.05	μ_{max}	0.10	K_S	−0.15	μ_{max}	0.21
7	i_{NX}	−0.12	μ_{max}	−0.06	k_d	−0.05	$pK_{H_2PO_4}$	0.07	$K_La_{NH_3}$	0.14	$pK_{H_2PO_4}$	0.14
8	pK_{NH}	0.08	γ_X	0.04	i_{NX}	−0.03	Y_{SX}	0.05	m_S	0.09	m_S	0.10
9	t_{lag}	0.07	$pK_{H_2PO_4}$	−0.03	$K_La_{NH_3}$	−0.02	i_{NX}	−0.03	$pK_{H_2PO_4}$	0.08	$K_La_{O_2}$	0.08
10	γ_X	0.05	m_S	−0.03	pK_{NH}	0.02	t_{lag}	−0.01	γ_X	−0.05	V_L	0.08

* The estimated means of distribution of scaled elementary effects of the inputs on the outputs. A positive sign indicates that the input has a positive effect on the output, and vice versa the negative sign indicates that the input has a negative effect on the output.

However, one important difference between the Morris screening and the SRCs was related to the effects of nitrogen dynamics-related parameters (i_{NX} and pK_{NH_4}). Although the Morris method detected these parameters to have significant impact on glucose, oxygen, biomass, actinorhodin, and base addition, the SRCs could only detect their impact on the base addition predictions. This issue is also discussed in more detail below.

Concerning the actinorhodin, overall the parameter rankings by both Morris and SRCs allow to conclude that biomass metabolic parameters are influential on the antibiotic production, whereas physical–chemical related parameters are less influential. However, the two methods differ on the evaluation of the impact of K_{HCO_2}, the Henry's constant for CO_2 dissolution. This parameter was found insignificant by the Morris method, whereas SRCs suggested that it was important. From a process/physiological knowledge point of view,[17] this parameter is only remotely related to the actinorhodin dynamics. Hence, it is not expected to be significant as suggested by the Morris method. On the basis of that process knowledge and on the result of the Morris method, one may suspect this parameter to be misclassified (Type I error, which is the error committed when the statistical method identifies an input parameter as significant when in fact it is not)[16] by SRCs. This is likely to be caused by the low coefficient of model determination (R^2 equal to 0.63) obtained for the regression model built on the actinorhodin when calculating the SRCs.

Differential Sensitivity Analysis. The outcome from the differential analysis is a time series profile of the sensitivity function of the model outputs with respect to the input parameters. An example of the nondimensional sensitivity function for oxygen and biomass as model outputs with respect to the input parameters μ_{max} and Y_{SX} is shown in Figure 5. Typically these sensitivity functions are used to examine the correlation between parameters and time-dependent effects of parameters on the outputs.[18,19] For instance, one can observe that the sensitivity functions of oxygen with respect to both parameters were closely and negatively correlated throughout the fermentation curve (in fact the correlation coefficient is about −0.80), whereas the sensitivity functions of the parameters on the biomass are only correlated during the stationary and decay phase ($t > 100$ h) of the fermentation (Figure 5-middle). Moreover, one also observes that both the magnitude and the sign of the effect change during the progress of the batch (see e.g., oxygen sensitivity functions).

The sensitivity measure, δ^{msqr}, is a mean squared summary of these time series profiles of sensitivity functions whose magnitude indicates the extent of the effect of a parameter on the outputs. Similarly to SRCs and the Morris screening method, this measure allows ranking of the input parameters. The results of this ranking are shown in Table 4 for the first 10 highest ranking parameters. A caveat must be made concerning these rankings: not all of the ranked parameters may be significant. One needs to consider the magnitude of the δ^{msqr} values relative to each other. For instance, the value suddenly drops from 0.12 to 0.01 in the glucose column for the eight and ninth ranked parameters, which raises doubts whether the ninth ranked parameter is significant at all. An appropriate clustering algorithm applied to the sensitivity measure values would generate this separation line (distinguishing significant from insignificant parameters) automatically, but clustering is not within the scope of this contribution.

These results have a lot in common with the parameter significance ranking obtained by Morris screening and SRCs, especially for those parameters that were found to have significantly high effects on the outputs. This agreement between the three methods on the most sensitive parameters is positive as it reveals that the parameter significance ranking results are independent of the methods used. This issue is further discussed below.

However, there is an important difference between the parameter significance ranking based on the δ^{msqr} measure, which is the emergence of t_{lag} as a prominent parameter affecting all the model outputs significantly (its order in the significance ranking is much higher when compared with the order found by the SRC and the Morris method) (Table 4). This means that at the given initial conditions (nominal input parameter values where the differential analysis is carried out), the model outputs appear to be quite sensitive to the onset of the exponential phase parameter, t_{lag}. Two possible reasons can be offered to explain this difference in the ranking of t_{lag}: (a) the difference in the way model outputs were averaged: the SRC and Morris screening were performed on the averaged outputs, whereas differential analysis was performed on dynamic (time-series) outputs and later an average was taken as the mean-squared error, δ^{msqr} (refer to Eqs. 4 and 5 to see how the δ^{msqr} measure is calculated), and (b) the differential analysis is a local technique, whereas the SRC and Morris screening are global methods. Hence, the resulting rank order of t_{lag} may be relative to the initial conditions at which the analysis is performed.

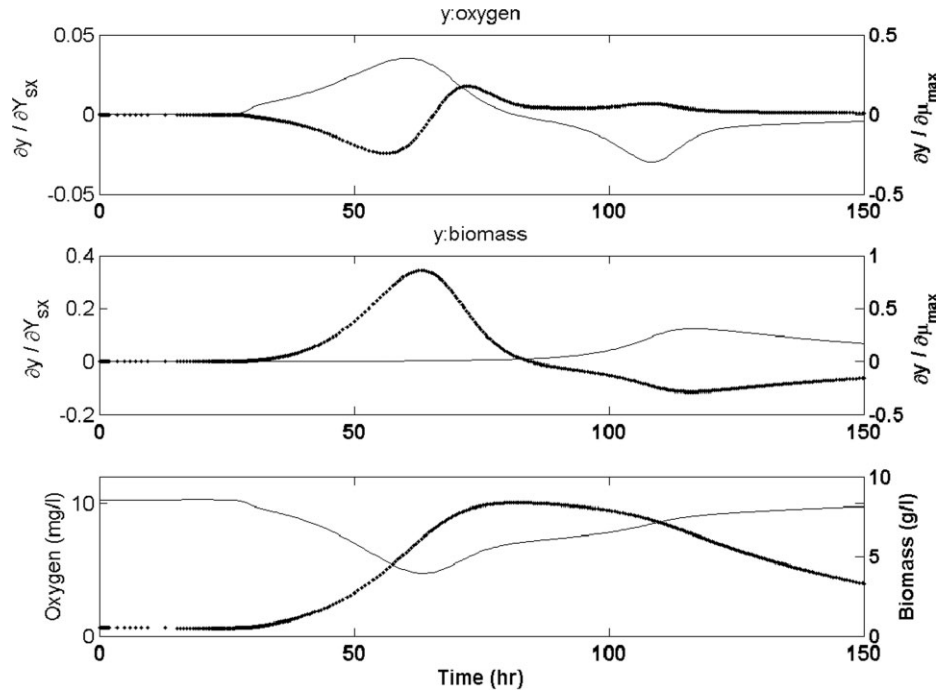

Figure 5. Non-dimensional sensitivity functions of oxygen (top) and biomass (middle) with respect to parameters μ_{max} and Y_{SX}. The oxygen and biomass profiles during the batch are also shown for comparison purposes (bottom). Grey lines corresponds to the y-axis on the left-hand side, whereas the bold lines corresponds to the y-axis on the right-hand side.

Table 4. Parameter Significance Ranking based on the δ^{msqr} Sensitivity Measure: The First 10 Highest Ranking Parameters on Six Model Outputs are Shown

	Glucose		Oxygen		Biomass		Actinorhodin		Base		Off-gas CO_2	
Rank	θ	δ^{msqr}	θ	δ^{msqr}	θ	δ^{msqr}	θ	δ^{msqr}	θ	δ^{msqr}	θ	δ^{msqr}
1	i_{PX}	0.76	K_{HO_2}	0.80	i_{PX}	1.27	S^{max}_{ACT}	2.19	i_{NX}	1.53	P_{out}	0.83
2	t_{lag}	0.71	P_{O_2}	0.80	μ_{max}	1.09	K_P	1.59	i_{PX}	1.46	i_{PX}	0.80
3	μ_{max}	0.70	P_{in}	0.80	t_{lag}	1.07	μ_{max}	1.39	μ_{max}	0.67	μ_{max}	0.78
4	K_P	0.63	i_{PX}	0.18	K_P	0.97	$pK_{H_2PO_4}$	0.86	t_{lag}	0.64	t_{lag}	0.77
5	$pK_{H_2PO_4}$	0.42	μ_{max}	0.18	m_S	0.97	t_{lag}	0.80	K_P	0.60	P_{in}	0.77
6	m_S	0.35	t_{lag}	0.17	$pK_{H_2PO_4}$	0.64	i_{PX}	0.47	pK_{NH_4}	0.58	Q_{Gin}	0.75
7	Y_{SX}	0.13	$K_La_{a_2}$	0.16	k_d	0.46	β_{ACT}	0.44	$pK_{H_2PO_4}$	0.39	pK_{HCO_3}	0.74
8	K_S	0.12	K_P	0.16	Y_{SX}	0.23	K_S	0.21	$K_La_{NH_3}$	0.25	K_P	0.69
9	*k_d*	*0.01*	m_S	0.11	K_S	0.18	m_S	0.18	*K_S*	*0.06*	m_S	0.48
10	Y_{SACT}	*0.01*	$pK_{H_2PO_4}$	0.10	Y_{SACT}	*0.02*	Y_{SX}	*0.05*	pK_{HCO_3}	*0.04*	$pK_{H_2PO_4}$	0.45

*A caveat: those parameter values *italicized* are not necessarily considered to be significant (see the text)

Comparison of the sensitivity analysis methods

Each of the three sensitivity analysis methods evaluated above was found valuable on its own for achieving certain objectives, while at the same time each method was demonstrated to have some limitations. The differential analysis provides an in-depth insight into how the effect of model parameters on key process variables changes as function of time. The parameter significance ranking obtained by the differential analysis appeared to agree *mostly* with the ranking obtained by the two other methods. A difference, however, was found concerning the ranking of one input parameter (see above), which cautions that the ranking based on differential analysis may not be safely extrapolated to other conditions. This is in fact not surprising and somewhat expected of local methods. However, we feel it is important to emphasize that one should not dismiss the ranking obtained from local methods to be totally irrelevant on global scale—at least in our study we have seen the ranking to be quite consistent with other methods.

The standardized regression coefficients (SRCs) provide a good approximation to the global sensitivity measure of the parameters at an affordable computational cost. It is important to mention that the computational cost associated with SRC is independent of the number of model parameters. This is not the case with neither Morris screening nor the differential analysis methods—both heavily depend on the number of parameters. However, the SRC method is only valid to the extent that the model outputs obtained from the Monte Carlo simulations can be linearized (i.e., the criterion to be satisfied is $R^2 \geq 0.7$). Finally, the Morris method provides a global measure of the sensitivity measure of a parameter, but may be computationally too demanding depending on the number of model parameters that are considered for the analysis. On the basis of our experiences, we suggest that (i) the differential analysis be used for detailed investigation of parameter sensitivity functions, (ii) the intuitively simple SRCs method be used for approximating the global sensitivity measure of a parameter effect on the

output. On the basis of the heavy computational cost associated with the Morris method, it is deemed useful only to verify the SRCs if the R^2 criterion doesn't hold.

Apart from the three methods evaluated above, statistical analysis has some more methods such as Sobol's and FAST sensitivity indices, which allows decomposition of output variance with respect to inputs.[8] These methods, although providing a truly global sensitivity measure, however, still suffer from a heavy computational demand, which is in the order of 10,000 model evaluations.[8] This issue makes them infeasible at this point of time to apply for pragmatic ends such as sensitivity analysis of complex fermentation models, but should be considered in the future as computational tools render them feasible.

Impact of stochastic and structural uncertainty on simulating fermentation systems

Overall, the uncertainty analysis helped to quantify uncertainty on the model outputs. Furthermore, one also observes that a significant uncertainty exists in the model outputs, which should not be ignored from a good modeling practice point of view. That being said, the output uncertainty as obtained in this study concerns the input uncertainty alone and excluded other sources of uncertainty namely stochastic and structural uncertainty (as mentioned in the introduction). As demonstrated in this study, mature statistical analysis methods exist to deal with input uncertainty. In what follows, we discuss other sources of uncertainty relevant for fermentation technology, and highlight their possible implications and future research needs.

The stochastic uncertainty relates to random events in the batch cultivations such as failure of mixing or aeration equipment, which may strongly affect the process performance. Uncertainty that may arise from such random events can be addressed by incorporating a stochastic process into the model structure.[7] However, incorporating such uncertainty is only useful if the predictions (of fermenter models) are intended to be used for specific purposes such as controller robustness evaluations (e.g., how does a controller react when the aeration equipment pump fails?) or risk analysis (what is the risk of failure of a fermentation process?). When models are used in a research setting, the predictions are usually used to interpret data collected from successful fermentation runs; hence, it remains more important to focus on input uncertainty.

The structural uncertainty concerns the adequacy of the mathematical model to represent the fermenter under study including biological and physical–chemical processes. For the biological part, one observes that many alternative model structures exist to simulate biological processes varying from as simple unstructured[1,20] to structured models including metabolic flux analysis and dynamic metabolic networks.[2] As for the physical processes, the mixing in the fermenter is modeled using a range of models, ranging from ideal completely mixed reactors to Computational Fluid Dynamics (CFD) methods.[21] Obviously, each choice of a model structure will imply some degree of approximation of reality and hence some degree of uncertainty on the predictions. Although CFD can help in reducing uncertainty on the mixing part (but at the expense of a high computational cost), the mathematical description of highly complex biological phenomena in the fermenters is expected to remain challenging. Although preliminary research exists to quantify structural uncertainty, it has to be emphasized, however, that this research field is still in its infancy and requires further research.[6] At this point of time, it is suggested to use the valuable process knowledge about the fermentation systems to reflect on this source of uncertainty when evaluating the predictions of the model.

Perspectives on the use of uncertainty and sensitivity analysis as part of GMoP in PAT

If the PAT initiative can be seen as promoting a continuous quest to increase process knowledge and understanding as a necessity for improving the quality and efficiency of the manufacturing process, then mechanistic modeling becomes a valuable tool for serving this purpose (of course in addition to multivariate chemometric tools).[22,23] In this respect, we firmly believe that the uncertainty and sensitivity analysis methods form an essential part of good modeling practice (GMoP), which is a necessity for building reliable models and correctly interpreting their results, as demonstrated here for a complex fermentation system example, an example that is highly relevant for industrial biotechnology. Overall, these methods help to quantify what the quality of the model predictions are (uncertainty) and help finding out how the quality of the predictions can be further improved (sensitivity analysis).

One additional advantage of using the uncertainty analysis method is that results of uncertainty analysis can be presented as cumulative probability distribution functions. Having such information available for decision makers (researchers, engineers, and managers alike) sets the stage for risk-based decision making within PAT applications, which may potentially allow more efficient/optimal process developments.

These methods have been successfully used in a wide range of scientific and engineering disciplines for improving the model based decision making, for process understanding, for design, optimization and control.[7,8,24] These methods are also expected to be equally useful for tackling this rather challenging engineering problem of improving manufacturing practice in the pharmaceutical industry[5] and in bioprocess development—e.g., process characterization and validation—in industrial biotechnology.

Conclusions

The uncertainty and sensitivity analysis of a mechanistic model describing a batch cultivation of *S. coelicolor* for antibiotic production was studied. The input uncertainty resulting from assumptions made during model-building was effectively propagated using the well-known Monte Carlo method. The output uncertainty was found lower for some predicted variables including biomass, glucose, and base addition, whereas it was large for off-gas and antibiotic production.

The sensitivity analysis performed using Morris screening, SRC, and differential methods identified that only few (around 10) of a total of 56 model parameters were mainly responsible for the variance in the outputs. The significant parameters are related to biomass metabolism but also chemical equilibrium and mass transfer related process parameters. This information helps understand what causes the uncertainty and hence how it can be remedied.

Although three of the sensitivity analysis techniques were found useful on their own, it is suggested to use (i) the differential analysis for detailed investigation of the sensitivity functions, (ii) the SRC method for approximating the global measure of the parameter effect, and (iii) because of its computational demand, the Morris method is only advised to be performed when necessary, e.g., when the SRCs results are not reliable because of low degree of linearization of model outputs.

Overall the uncertainty and sensitivity analysis methods presented and evaluated above are part of good modeling practice (GMoP) that seeks to ensure building and applying reliable mechanistic simulation models. Hence, these methods are expected to be useful in tackling the rather challenging problem of improving manufacturing practice in the pharmaceutical industry as encouraged within PAT applications.

Acknowledgments

The research work of Dr. Gürkan Sin was financed by a H.C. Ørsted postdoctoral fellowship of the Technical University of Denmark.

Literature Cited

1. Esener AA, Roels J, Kossen NWF. Theory and applications of unstructured growth models: kinetic and energetic aspects. *Biotechnol Bioeng*. 1983;25:2803–2841.
2. Teusink B, Smid EJ. Modeling strategies for the industrial exploitation of lactic acid bacteria. *Nat Rev Microbiol*. 2006;4:46–56.
3. Thilakavathi M, Basak T, Panda T. Modeling of enzyme production kinetics. *Appl Microbiol Biotechnol*. 2007;73:991–1007.
4. Sin G, Ödman P, Petersen N, Lantz AE, Gernaey KV. Matrix notation for efficient development of first-principles models within a PAT framework: integrated modeling of antibiotic production with *Streptomyces coelicolor*. *Biotechnol Bioeng*. 2008;101:153–171.
5. FDA. PAT—A framework for innovative pharmaceutical development, manufacturing and quality assurance. Guidance for Industry. 2004; Available at: (<http://www.fda.gov/cder/guidance/index.htm>)
6. McKay MD, Morrison JD, Upton SC. Evaluating prediction uncertainty in simulation models. *Comput Phys Commun*. 1999;117:44–51.
7. Helton JC, Davis FJ. Latin hypercube sampling and the propagation of uncertainty in analyses of complex systems. *Reliab Eng Syst Saf*. 2003;81:23–69.
8. Saltelli A, Ratto M, Tarantola S, Campolongo F. Sensitivity analysis practices: strategies for model-based inference. *Reliab Engng Syst Saf*. 2006;91:1109–1125
9. Gelman A, Carlin JB, Stern HS, Rubin DB. *Bayesian Data Analysis*. London, U.K.: Chapman & Hall/CRC; 2003.
10. Brun R, Kühni M, Siegrist H, Gujer W, Reichert P. Practical identifiability of ASM2d parameters-systematic selection and tuning of parameter subsets. *Wat Res*. 2002;36:4113–4127.
11. Morris MD. Factorial sampling plans for preliminary computational experiments. *Technometrics* 1991;33:161–174.
12. Kiviharju K, Salonen K, Leisola M, Eerikainen T. Modeling and simulation of *Streptomyces peucetius* var. caesius N47 cultivation and erhodomycinone production with kinetic equations and neural networks. *J Biotechnol*. 2006;126;365–373.
13. McKay MD, Beckman RJ, Conover WJ. A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. *Technometrics* 1979;21:239–245.
14. Iman LR, Conover WJ. A distribution-free approach to inducing rank correlation among input variables. *Commun Statist-Simula Computa*. 1982;11:311–334.
15. Cariboni J, Gatelli D, Liska R, Saltelli A. The role of sensitivity analysis in ecological modeling. *Ecol Modell*. 2007;203:167–182.
16. Campolongo F, Cariboni J, Saltelli A. An effective screening design for sensitivity analysis of large models. *Environ Model Soft*. 2007;22:1509–1518.
17. Doull JL, Vining LC. Nutritional control of actinorhodin production by *Streptomyces coelicolor* A3(2): suppressive effect of nitrogen and phosphate. *Appl Microbiol Biotechnol* 1990;32:449–454.
18. Holmberg A. On the practical identifiability of microbial growth models incorporating Michealis-Menten type nonlinearities. *Math Biosci*. 1982;62:23–43.
19. Baltes M, Schneider R, Sturm C, Reuss M. Optimal experimental design for parameter estimation in unstructured growth models. *Biotechnol Prog*. 1994;10:480–488.
20. Dassau E, Zadok I, Lewin DR. Combining six-sigma with integrated design and control for yield enhancement in bioprocessing. *Ind Eng Chem Res*. 2006;45:8299–8309.
21. Enfors SO, Jahic M, Rozkov A, Xu B, Hecker M, Jürgen B, Krüger E, Schweder T, Hamer G, O'Beirne D, Noisommit-Rizzi N, Reuss M, Boone L, Hewitt C, McFarlane C, Nienow A, Kovacs T, Trägårdh C, Fuchs L, Revstedt J, Friberg PC, Hjertager B, Blomsten G, Skogman H, Hjort S, Hoeks F, Lin HY, Neubauer P, van der Lans R, Luyben K, Vrabel P, Manelius A. Physiological responses to mixing in large scale bioreactors. *J Biotechnol*. 2001;85:175–185.
22. Junker BH, Wang HY. Bioprocess monitoring and computer control: key roots of the current PAT initiative. *Biotechnol Bioeng*. 2006;95:226–261.
23. Maes I, Van Liedekerke B. The need for a broader perspective if process analytical technology implementation is to be successful in the pharmaceutical sector. *J Pharm Innov*. 2006;1:19–21.
24. Refsgaard JC, van der Sluijs JP, Højberg AL, Vanrolleghem PA. Uncertainty in the environmental modeling process e A framework and guidance. *Environ Modell Soft*. 2007;22:1543–1556.

Manuscript received July 28, 2008, and revision received Sept. 2, 2008.