Identification of key bioprocess variables using global sensitivity analysis

J.M.P. King, N.J. Titchener-Hooker and Y. Zhou

The Advanced Centre for Biochemical Engineering, Department of Biochemical Engineering, University College London, London, UK

Abstract

The selection of appropriate operating conditions for bioprocessing is complex due to the large number of interacting stages and variables. Models for each processing stage require significant numbers of variables and thus the whole process model will consist of a large number of variables overall. Interactions typically exist within a bioprocess meaning that the operation of one unit or the value of a variable may adversely effect the operation of subsequent units. It is therefore necessary to consider the process as a whole – selecting the optimal conditions for individual units will not yield optimal performance for the process. In addition to these complexities, bioprocesses also operate under tight regulation and it is necessary to demonstrate that performance is satisfactory over the likely operating range. Therefore, tools to analyse the sensitivities of the variables and to identify the key variables will assist bioprocess design and be of significant utility.

Conventional approaches for the analysis of variable sensitivities are inadequate. Since they only consider one variable at a time and are unable to consider interactions between variables. We propose the use of global sensitivity analysis to determine the level of importance of each variable and their interactions. Global sensitivity analyses enable the effects of variable and parameter changes to be determined over the range of likely operation, but more importantly it determines the impact of all variables simultaneously rather than individually, which is necessary where interactions between variables are expected. Quantitative sensitivity indices for each variable and their interactions can be quantified which are used to determine the relative importance of each variable. Once key variables have been determined, the designer may focus on the most significant variable subset and investigate the effects of any process interactions.

In this paper a case study is presented which illustrates the utility of the application of global sensitivity analysis to bioprocess operation. The case study investigates a two-stage sequence of a fermentation and subsequent centrifugal harvest and studies how the variable sensitivities change as the fermentation yield increases. The impact of such changes on the operation of the process is then considered. It is found as the cell density increases the importance of the fermentation stage increases. Thus at higher cell densities control of the fermentation will have a larger impact on the overall performance than the subsequent harvest centrifuge.

Keywords: Global sensitivity analysis, fermentation, centrifugation, bioprocessing, simulation

Introduction

The selection of appropriate operating conditions for bioprocesses is a complex task due to significant interactions which occur between processing stages (Siddiqi et al, 1995, and Clarkson et al, 1996). In addition, models to describe the unit operations which comprise the process, typically have many variables each of which may interact. Thus the operation of unit operations must be considered within the context of the whole process model – selection of unit operating conditions considered in isolation will not result in optimal operation for the process as a whole, for example the operation of upstream units may negatively impact upon units downstream. As well as these operational considerations, it is important to note that biopharmaceuticals are produced within a tight regulatory framework, which requires the demonstration of consistent robust operation of the process within defined operating limits. This is known as process validation. Therefore, the ability to identify key variables within a bioprocess will assist significantly with both the determination of satisfactory operation and with the demonstration of robust operation, since knowledge about which variables contribute most to the model output will enable engineers to focus efforts on these areas.

Traditional approaches to variable sensitivity analysis are inadequate as they investigate the impact of variables one at a time which by its very nature precludes the investigation of the impact of variable interactions. Global sensitivity analysis however investigates the impact of a number variables over a defined range and therefore this methodology is suitable for the analysis of a complex systems such as bioprocesses. The ability to carry out the analysis over a set range is particularly relevant since bioprocesses will be validated over defined ranges and knowledge from models about which variables are most significant will assist with this task by reducing its complexity. Previous work has demonstrated the applicability of global sensitivity analysis for determining further data and information on the significance of bioprocess variables (King et al, 2005).

In this paper a case study is presented which illustrates the application of global sensitivity analysis to bioprocess operation. The case study investigates a two-stage sequence of a fermentation and subsequent centrifugal harvest, and studies how the variable sensitivities change as the fermentation biomass yield increases. The impact of such changes on process operation is then considered.

Global Sensitivity Analysis

Sensitivity analysis techniques quantify how the output from a model depends upon each of its input variables. Global sensitivity, in particular, attempts to determine the relative effect of variables on the model outputs considering all variables simultaneously. This is in contrast to local sensitivity analysis, which typically determines the rate of change of model output with respect to individual model variables. Global sensitivity analysis therefore determines sensitivities of a multidimensional system as opposed to local sensitivity analysis which finds a

gradient at the operating point with respect to a single variable. In addition, global sensitivity techniques may be used to determine the strength of interactions between model variables.

There are a number of global sensitivity analysis techniques including FAST (Saltelli and Bolado, 1998) and Sobol's method; in this paper Sobol's method is used due to the ease of formulation for the complex equations involved in bioprocesses. Sobol's method is based on the ANOVA (Analysis of Variances) representation of functions. This representation decomposes the model function under study into a summation, assuming that the model may be represented by a function of n variables, e.g $f(x_1, ..., x_n)$:

$$f(x_1,...,x_n) = f_0 + \sum_{s=1}^{n} \sum_{i_1<...<i}^{n} f_{i_1..i_s}(x_{i_1}...x_{i_s}) \qquad \text{(equation 1)}$$

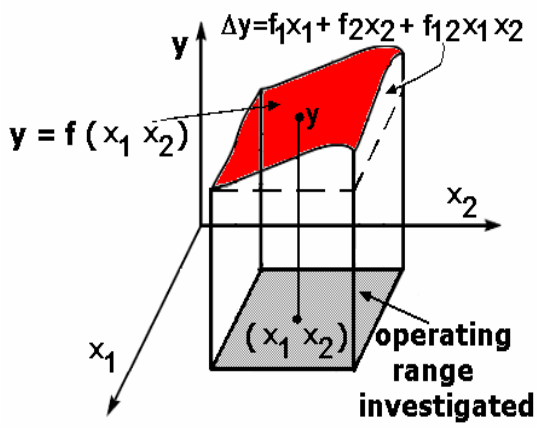

Figure 1: This diagram shows the concepts involved in global sensitivity analysis of a system with two variables, x_1 and x_2, which control a function f. The surface at the top of the figure represents the response of the function $f(x_1\ x_2)$ over the operating range. The operating range is represented by the rectangular region at the base of the graph. The volume above the dashed line represents the total change in performance over the operating range due to changes in all variables and this change in performance may be represented by the equation shown for Δy which is based on the ANOVA concept (see equation 1). Global sensitivity analysis determines the contribution from each term in this equation and hence from each variable or combination of variables.

Figure 1 shows a two-dimensional example of the concepts involved in Sobol's global sensitivity analysis. In particular, the figure indicates the operating range investigated, the resulting range of model response and the significance of the ANOVA representation, i.e. Δy represents the change in model output over the response surface shown, with a contribution due to each variable and combination of variables. Global sensitivity analysis determines the significance of the contribution of each variable or combination, using the ANOVA representation and advanced mathematics techniques (Sobol, 2001). Previous work details our application of Sobol's method to bioprocess models and in this work we use the same implementation (King et al, 2005).

The following section describes a case study which investigates changes in variable sensitivities which occur with modifications in process operating conditions, demonstrating their impact over a sequence of two units. This enables the engineer to determine appropriate operating strategies and equipment.

Case Study

Centrifuge Models

In this paper, a previously developed centrifuge model (Clarkson et al, 1996) has been used to determine the impact of changes in operating conditions and changes in physical properties on model sensitivities. The model is based upon standard centrifugation theory, namely a modified Stokes' Law (Richardson and Zaki, 1954), which takes into account hindered settling. Hindered settling is a phenomenon which occurs when at higher solids concentrations (>2%) where particles interact, slowing the rate of sedimentation. With hindered settling the critical diameter is defined as (Richardson and Zaki, 1954):

$$d_c = \sqrt{\frac{18Q\mu}{\Delta\rho\Sigma(1-C_v)^{\sigma}\,\mathrm{g}}} \qquad \text{(equation 2)}$$

where Q is the volumetric throughput, μ is the dynamic viscosity of the fluid, $\Delta\rho$ is the difference in the densities between the liquid and solids phases, g is the acceleration due to gravity, Σ is the equivalent settling area of the centrifuge, C_v is the volume concentration of solids and σ is a geometric factor dependent upon solid particle shape (4.6 – for spherical particles). The calculation of Σ is dependent upon the type of centrifuge, but is dependent upon factors such as bowl dimensions and spin speed. Application of the grade efficiency curve, $T(x)$, accounts for any non-idealities of fluid flow which may exist in industrial machines (Mannweiler, 1990):

$$T(x) = 1 - \exp[-k(x/d_c)^n] \qquad \text{(equation 3)}$$

The experimentally-determined parameters k and n describe the grade efficiency curve, while x and d_c are the particle and critical particle diameters respectively.

Application of a mass balance over the centrifuge stage, using the grade efficiency curve, and assuming a Normal feed particle size distribution with a mean m and a standard deviation σ, enables the clarification efficiency, $c.e.$, to be calculated:

$$c.e.(x) = \int_0^{\infty} p(x).T(x)dx$$

$$= \int_0^{\infty} \frac{1}{s\sqrt{2\pi}} e^{-\left(\frac{(x-m)^2}{2s^2}\right)} \left(1\text{-}e^{-k\left(\frac{x}{d_c}\right)^n}\right) dx \qquad \text{(equation 4)}$$

The function $p(x)$ represents the Normal particle size distribution of the feed material. Clarification efficiency, as defined by equation 4, will be the performance criterion used in the following case studies.

Case Study

In this case study the variation of sensitivities with changes in operating conditions are investigated to determine their impact on the sensitivity of model variables. Operational conditions may change during the lifetime of a plant, for example if strain and culture improvements occur enabling fermentations to be carried out at higher cell densities. This will require that existing equipment is operated differently in order to maintain the desired product targets throughout the process. For example, the flow rate would have to be adjusted to suit a changed solids volume concentration as the critical diameter, d_c, is directly affected by the volume concentration of solids (see equation 2) and also by any change in the viscosity arising from an increased solids concentration. Therefore, the flow rate must be reduced in such a case to ensure that the desired level of clarification was maintained. The study determined the effects of a change in solids volume concentration, from 5% to 30% v/v, on the predicted variable sensitivities.

Two studies were carried out, the first using particle size, flow rate, viscosity and density difference and in the second solids volume concentration replaced density difference. In both studies the flow rate was adjusted so that at each operating point the clarification level was maintained at 98-99% and these changes are shown on Figure 3. The range of each variable was +/- 5% of the operating point. Viscosity values were linked to the solids concentration, using experimental data (Salte et al, 2005).

Results

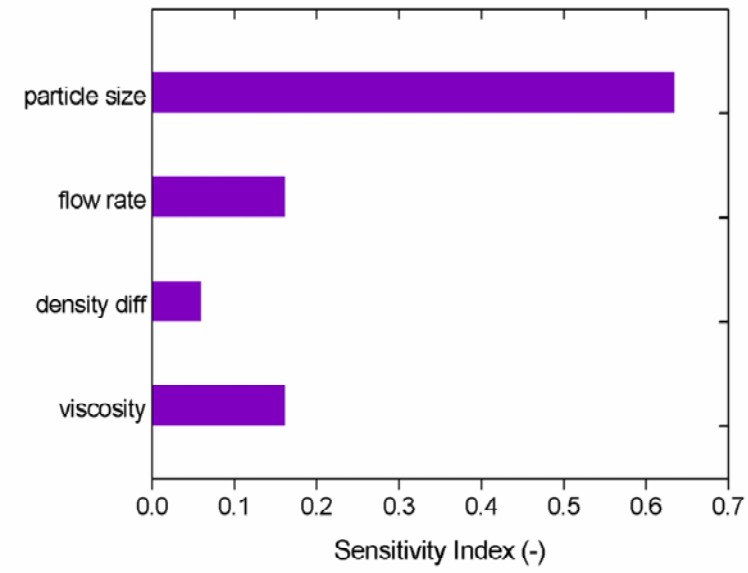

Figure 2: Sensitivities for the first case the study, where changes in sensitivity of particle size, flow rate, density difference and viscosity were analysed with respect to changes in solids fraction. These variable sensitivities were not found to change significantly over the range studied (solids fraction was varied from 5-30%).

For the first study, where changes in the sensitivity of particle size, flow rate, density difference and viscosity were analysed, these variable sensitivities did not change significantly from the values shown in Figure 2, over the range of solids fractions studied (5-30%). This lack of change can be understood by inspecting equation 2. Since the target clarification level remains fixed at 98-99% and although the viscosity increases at higher solids densities, changes in flow rate compensate for this resulting in a similar critical diameter and thus similar clarification, over the ranges used in the global sensitivity analysis. The outcome of this is that the sensitivities remain unchanged.

Figure 3: The change in variable sensitivities as a function of solids fraction changes – flow rate ; solids fraction ; particle size ; viscosity . Flow rate was adjusted to ensure clarification was between 98 and 99% in all cases.

The second case study produces more illuminating results, which are shown in Figure 3. Note that the lines for flow rate and viscosity sensitivity indices are co-incident. The results shown in Figure 3 can be used to illustrate how the level of solids fraction produced in the fermentation broth impacts on the possible complexity of the process validation. When operating at the higher solids fractions it can be seen that there are two variables which significantly influence the performance, namely solids fraction and particle size but in contrast at lower solid fractions only particle size is a significant factor. Thus, the complexity of the validation task will be greater and more time consuming for the process operated at a higher solids fraction as the robust limits of operation for two variables rather than one will need to be investigated in detail. In addition, information gain from sensitivity analyses, such as this, will also impact on process operation and monitoring. For example, above a solids concentration of approximately 0.175 (v/v), then the sensitivities which have the greatest effect on clarification are all associated with the physical properties of the fermentation broth, rather than the operation of the centrifuge. Thus changes in the operation of the fermentation give

the biggest opportunity to influence the clarification efficiency. Therefore, as the solids concentration increases, the importance of monitoring the progress of the fermentation increases, since any change in the physical properties has the potential to have a big impact upon the output of the process, namely the clarification efficiency.

Conclusions

The influence of changes in operating conditions on variable sensitivities have been determined and this has been shown to give valuable insight when considering process operating strategies. The information has also been shown to give insight into the potential complexity of process validation, part of the regulatory approval process for therapeutics.

Furthermore, knowledge of the significance of each variable may be used to make an assessment of the potential robustness of a process. For example, where a process is found to be highly sensitive to a certain variable, known to be subject to a significant level of uncertainty during operation, then this may serve as a warning that other process options should be considered, such as alternative ranges of operation or equipment.

Acknowledgements

The support of the Engineering and Physical Sciences Research Council Life Science Interface and Innovative Manufacturing Research Centre initiatives for the IMRC in Bioprocessing is gratefully acknowledged. The IMRC is part of The Advanced Centre for Biochemical Engineering, UCL with collaboration from a range of academic partners and biopharmaceutical and biotechnology companies.

References

1. Clarkson A I, Lefevre P and Titchener-Hooker NJ, “A study of process interactions between cell disruption and debris clarification stages in the recovery of yeast intracellular products”, Biotechnology Progress 3: 462-467 (1993).

2. Clarkson AI, Bulmer M and Titchener-Hooker NJ, “Pilot-scale verification of a computer-based simulation for the recovery of biological particles”. Bioprocess Engineering 14:81–89 (1996).

3. King J.M.P, Zhou Y., and Titchener-Hooker N.J., “Ranking bioprocess variables using global sensitivity analysis: A case study in centrifugation.”, 7th World Chemical Engineering Congress, Glasgow, Scotland (2005).

4. Mannweiler K., “The recovery of biological particles in high-speed continuous centrifuges with special reference to feed zone break-up effects”, PhD Thesis, University of London, (1990).

5. Richardson J.F. and Zaki W.N., “Sedimentation and Fluidisation: Part1”, Transactions of the Institution of Chemical Engineers, 32, 35-53.

6. Salte H., King J.M.P., and Titchener-Hooker N.J., “Centrifuge selection for recovery of high solids density cell broths and visualization of separation performance via Windows of Operation”, The 229th ACS National Meeting, in San Diego, CA, March 13-17, 2005

7. Saltelli A and Sobol IM, ”About the use of rank transformation in sensitivity analysis of model output”, Reliability Engineering and Systems Safety, 50: 225-239 (1995).

8. Saltelli A and Bolado R, “An alternative way to compute Fourier amplitude sensitivity test FAST”, Computational Statistics and Data Analysis, 26: 445-460 (1998).

9. Siddiqi SF, Bulmer M, Ayazi Shamlou P and Titchener-Hooker NJ, “The effect of fermentation conditions on yeast cell debris particle size distribution during high pressure homogenisation”, Bioprocess Engineering 14: 1 – 8 (1995).

10. Sobol IM, “Sensitivity estimates for non-linear mathematical models”, Mathematical Modelling and Computational Experiments, 1:407-414 (1993).

11. Sobol IM, “Global sensitivity indices for non-linear mathematical models and their Monte Carlo estimates”, Mathematics and Computers in Simulation, 5: 271-280 (2001).