

Improved calibration of a solid substrate fermentation model

Johannes Sacher¹ · Pedro Saa² · Martín Cárcamo² · Javiera López² · Claudio A. Gelmi² ✉ · Ricardo Pérez-Correa²

¹ Technische Universität Berlin, Fakultät III Prozesswissenschaften, Berlin, Germany

² Pontificia Universidad Católica de Chile, Escuela de Ingeniería, Departamento de Ingeniería Química y Bioprocesos, Macul, Santiago, Chile

✉ Corresponding author: cgelmi@ing.puc.cl
Received April 5, 2011 / Accepted June 29, 2011
Published online: September 15, 2011
© 2011 by Pontificia Universidad Católica de Valparaíso, Chile

Abstract

Background: Calibration of dynamic models in biotechnology is challenging. Kinetic models are usually complex and differential equations are highly coupled involving a large number of parameters. In addition, available measurements are scarce and infrequent, and some key variables are often non-measurable. Therefore, effective optimization and statistical analysis methods are crucial to achieve meaningful results. In this research, we apply a metaheuristic scatter search algorithm to calibrate a solid substrate cultivation model. **Results:** Even though scatter search has shown to be effective for calibrating difficult nonlinear models, we show here that a posteriori analysis can significantly improve the accuracy and reliability of the estimation. **Conclusions:** Sensibility and correlation analysis helped us detect reliability problems and provided suggestions to improve the design of future experiments.

Keywords: *Gibberella fujikuroi*, gibberellic acid, global optimization, metaheuristics, nonlinear dynamic models, parameter estimation, solid substrate cultivation

INTRODUCTION

Processes in biotechnology usually represent productions in relatively small quantities and are therefore mostly done in batches. These systems possess an inherently dynamic character making it difficult for scale-up and optimization. Therefore, accurate dynamic models are necessary to predict the behaviour of these processes, which is why models have gained great importance in recent years. However it is difficult to uniquely identify the parameters and state variables, due to the sheer number used.

The usual procedure to identify a certain model relies on the quest to identify a set of values for the unknown parameters, such that the integration of the model will represent the existing experimental data as precisely as possible. This procedure, referred to as regression, involves minimizing the differences between experimental data and the model predicted values-which is essentially an optimization problem.

Most of the problems in bioprocess engineering applications are highly constrained; they exhibit nonlinear dynamics and sometimes have a non-smooth nature. These properties often result in non-convex and multimodal problems, and therefore are not well predicted by gradient based local

methods. In order to provide an approximation of the global optimum, one must use global optimization (GO) methods (Banga et al. 2003). These methods are based on robust solvers that can locate the vicinity of the global solution in a reasonable number of iterations and handle noise and/or discontinuities. Stochastic algorithms, a particular classification of GO methods, treat the objective function as a black box; that is, as a simple relationship between inputs and outputs without further information about relationships among the decision variables (no derivative information). Metaheuristics are particularly useful stochastic global optimization methods which are often used in biotechnology. They consist of an iterative generation process that efficiently finds near-optimal solutions by intelligently combining different learning strategies to explore and exploit the search spaces (Jones et al. 2002). These methods are easy to implement, making them robust for a wide variety of problems. The metaheuristic applied in this work is a scatter search method (SS), which is a population-based method that has recently shown promising outcomes for solving combinatorial and nonlinear optimization problems (Egea et al. 2007).

Nonetheless, a good model fitting doesn't ensure that a suitable model can be found. Models with a great number of parameters generally fit the experimental data better, but they tend to estimate parameters with large confidence intervals. Therefore, the quality of the parameter estimation, in terms of accuracy, needs to be checked before reaching a meaningful interpretation of the results.

The procedures required to test the quality of the parameter estimation are called post-regression diagnostics, which encompass various methods (Jaqaman and Danuser, 2006):

Parameter sensitivity (how parameter variations affect response variables).

Parameter significance (determine parameter confidence intervals).

Parameter identifiability (detection of cross correlation between parameters).

Model validity, also known as structural identifiability, implies that the model's parameters can be uniquely identified. If the post-regression diagnostics are unsatisfactory, the model has to be modified and/or the experimental data has to be repeated or complemented with new experimental data. Furthermore, the diagnostic methods can give valuable recommendations for modifying experimental setups in order to increase the data's information content. It has to be noted that there are several methods that systematically optimize experiments so that their information content is maximized. The different criteria and problems encountered are described in other publications (Chu and Hahn, 2008; Franceschini and Macchietto, 2008), and will not be discussed here. Finally, post-regression diagnostics are similar in function to pre-regression diagnostics, since the process of linking data to models is inherently cyclical and iterative (Figure 1).

In this research, a published dynamic model (Gelmi et al. 2002) that consists of eight coupled ordinary differential equations and 14 unknown parameters was analyzed. The model presented in Gelmi et al. 2002 describes the growth and production of a secondary metabolite in solid substrate cultivation (SSC) (for more details see Kinetic model). The model was fitted to different experimental conditions using a modern version of the SS algorithm, SS_m (Rodriguez-Fernandez et al. 2006; Egea et al. 2007). Parameter sensitivity, significance and identifiability were applied in order to assess the model's validity. As a consequence of the post-regression analysis, a modified model with a reduced parameter set was postulated and fitted to the experimental data using SS_m. The fitted parameters were further compared to the ones presented in Araya et al. (2007), using an alternative orthogonal collocation on finite element (OCFE) optimization method.

METHODS

Kinetic model

The lumped parameter model (Gelmi et al. 2002) describes the changes in the biomass of *G. fujikuroi* growing in glass columns on an inert support (Amberlite IRA-900). The model does not incorporate the lag phase of the fungus growth and it is based on the following assumptions:

Oxygen transfer resistance is low.

Available nitrogen is the limiting substrate.

The carbon source is not limiting.

Temperature (T) and water activity (aw) remain constant throughout the cultivation.

The model parameters remain constant during the cultivation.

The total amount of measurable dry biomass (X_{measu}) considers active as well as inactive fungi and is expressed on a dry weight basis (same method used for the other seven variables as well as the other seven state variables):

$$\frac{dX_{measu}}{dt} = \mu X$$

[Equation 1]

Assuming a first order death rate, the active biomass is:

$$\frac{dX}{dt} = \mu X - k_d X$$

[Equation 2]

Here, μ and k_d represent the specific growth and death rate of the fungus, respectively.

Experimental data (not shown in this article, see Gelmi et al. 2000) shows that the rate of urea consumption follows a zero order kinetics and is expressed as:

$$\frac{dU}{dt} = -k$$

[Equation 3]

k is the rate of conversion of urea to available nitrogen, N_I , which can be directly metabolized into active biomass. The concentration of assimilable nitrogen is given by:

$$\frac{dN_I}{dt} = 0.47k - \mu \left(\frac{X}{Y_{X/N_I}} \right)$$

[Equation 4]

In Equation 4, 0.47 corresponds to the urea nitrogen content and $Y_{X/N}$ is the yield between biomass and assimilable nitrogen. Equation 4 can be used to describe observed biomass growth despite total urea depletion within the culture. The fungus seems to accumulate part of the nitrogen from the urea conversion and uses it for growth when the external source, urea, has been depleted.

The microorganism consumes starch for growth and maintenance,

$$\frac{dS}{dt} = -\mu \left(\frac{X}{Y_{X/S}} \right) - m_S X$$

[Equation 5]

Since GA_3 is a secondary metabolite, its production rate includes a growth-associated term with nitrogen inhibition, β , and a first order degradation rate, k_p .

$$\frac{dGA_3}{dt} = \beta X - k_p GA_3$$

[Equation 6]

The differential equations for CO_2 production and O_2 consumption rates include two terms: one associated with growth and the other with maintenance,

$$\frac{dCO_2}{dt} = \mu \left(\frac{X}{Y_{X/CO_2}} \right) + m_{CO_2} X$$

[Equation 7]

$$\frac{dO_2}{dt} = \mu \left(\frac{X}{Y_{X/O_2}} \right) + m_{O_2} X$$

[Equation 8]

The specific growth rate, μ , is modelled using Monod's expression with assimilable nitrogen as the limiting nutrient,

$$\mu = \frac{\mu_{\max} N_I}{N_I + k_n}$$

[Equation 9]

Here, μ_{\max} is the maximum specific growth rate and k_n is the substrate inhibition constant. A substrate inhibition expression describes the specific GA₃ production rate,

$$\beta = \frac{\beta_{\text{etam}}}{1 + K_I N_I}$$

[Equation 10]

where β_{etam} is the maximum specific GA₃ production rate and K_I is the associated substrate inhibition constant.

The above model was calibrated using experimental data of four cultures at different temperature and a_w conditions,

- 1) Temperature = 25°C, Water Activity = 0.992.
- 2) Temperature = 25°C, Water Activity = 0.999.
- 3) Temperature = 31°C, Water Activity = 0.985.
- 4) Temperature = 31°C, Water Activity = 0.992.

Further details regarding the experimental setup and the above model are available in Gelmi et al. (2000) and Gelmi et al. (2002).

Parameter fitting

The model parameters were estimated by minimizing the sum of squared residual errors between predicted and experimental data. Because not all variables were measured simultaneously and because variables have different scales, the sums of squared residuals were weighted by the squared number of measurements and by the squared mean experimental value. Thus, the following least square function was minimized:

$$\text{Min}_{\Theta} \sum_{i=1}^m \sum_{j=1}^{n_i} \left(\frac{X_{i,j}^{\text{model}} - X_{i,j}^{\text{exp}}}{n_i \cdot \bar{X}_i^{\text{exp}}} \right)^2$$

[Equation 11]

where Θ is the parameter space, X_{ij} is the measurement j of the variable i , m is the number of measured variables, n_i is the number of measurements and \bar{X}_i^{exp} is the mean of experimental measurements for the variable i .

The optimization procedure was solved numerically using SSm, which is a routine developed in MATLAB® (Rodriguez-Fernandez et al. 2006; Egea et al. 2007). SSm orients its explorations in an evolutionary procedure where, in each step, random linear combinations of a reference set of solutions are created. The reference set is then updated with some best new solutions found and also with some of the most diverse solutions found in order to avoid a local minimum. Before the first cycle of iteration, a large set of diverse trial solutions is created from which the first reference set is created. If at some point in the optimization, the solutions in the reference set impoverish in diversity, a diversification method will pass more diverse trial solutions to the reference set. If desired, a so-called improvement method can undertake local searches for the best solutions found. Several local solvers can be easily chosen in SSm options. We encountered the local solver *fminsearch* to be very effective. The iteration stops when the maximum number of objective function evaluations is reached and the best solution is drawn out. In our case, an evaluation number between 30,000 and 40,000 appeared to be sufficient.

Pre/post-regression diagnostics

The regression diagnostic tests were computed using the software package MATLAB® version 2009b running in a Windows PC 2 GB RAM.

Parameter sensitivity measures to what extent the model's outputs are influenced by variations in parameter values. Absolute sensitivity is defined as:

$$G_{ij}(t) = \frac{dX_i(t)}{d\theta_j}$$

[Equation 12]

It is particularly useful for complex systems that involve a large number of variables and parameters, in which it is crucial to identify either the most important ones or the non-sensitive parameters (Yue et al. 2008). Non-sensitive parameters are parameters that don't influence the state variables. In dynamic models, parameter sensitivity varies with time and thus parameters can be non-sensitive for certain time intervals and sensitive for others. Since parameters and response variables can differ extremely in scale, a normalized sensitivity is useful to facilitate comparison:

$$\bar{G}_{ij}(t) = \frac{\theta_j}{X_i} \frac{dX_i(t)}{d\theta_j}$$

[Equation 13]

The derivatives of Equation 13 were determined numerically by the finite difference method. A central difference approach was used with a 1% perturbation factor around the nominal parameter values.

Parameter determinability or identifiability calculates the correlation matrix of the model parameters in order to determine if the fitted parameters are locally identifiable for a given time interval (Jacquez and Greif, 1985). The matrix elements represent the parameter cross correlations, defined as:

$$K_{ij} = \text{correlation } (G)$$

[Equation 14]

G is the sensitivity matrix, where each column represents the sensitivity of a given parameter (j) computed at measured times (t_1, t_2, \dots, t_n) for all the states. In general, parameters that are *a priori identifiable* have correlations with all other parameters between -1 and $+1$. In practical terms, a high correlation coefficient value is defined as:

$$|K_{ij}| > 0.95$$

[Equation 15]

If two parameters are highly correlated, it is said that they are *a priori unidentifiable*. These parameters affect the measured variables in exactly the same way (or same opposite way).

Parameter significance and confidence intervals were calculated after the parameters were fitted to the experimental data. In order to establish the statistical significance of estimates, a Student t -value for each parameter was used (Franceschini and Macchietto, 2008):

$$t_i = \frac{\hat{\theta}_i}{\sqrt{\sigma_i^2}}$$

[Equation 16]

where σ_i is the standard deviation of the i^{th} parameter. The t -values may be tested by reference to a t -distribution with $(n-p)$ degrees of freedom; higher t -values than those in the distribution tend to indicate reliable estimates (t -value > 1.96 for large values of $n-p$). Very low t -values suggest that confidence intervals or regions involving that parameter may include zero and such a situation means that the parameter could be statistically dropped from the model.

The usual method to obtain the parameter variances relies on estimating them through the diagonal elements of the Fisher Information Matrix (FIM) (Landaw and DiStefano III, 1984; Petersen et al. 2001):

$$\sigma_i^2 \geq FIM_{ii}^{-1}$$

[Equation 17]

In the case of biologic and chemical dynamic processes, the FIM can be obtained by (Chu and Hahn, 2008):

$$FIM = \sum_1^N G^T Q^{-1} G$$

[Equation 18]

The matrix product is summed over all experimental times. In this equation, Q is the matrix of measurement error covariances and G is the sensitivity matrix of the measured states. Assuming that the measurement errors are independent, Q is a diagonal matrix with all covariance equal to zero. The only non zero entries are the diagonal elements, $\sigma_{meas,j}^2$, that are the measurement errors variances. These variances are usually approximated using the residuals of the state variables (Franceschini and Macchietto, 2008):

$$\sigma_{meas,j}^2 \approx \hat{\sigma}_{meas,j}^2 = \sum_k^N \left(y_{j,k}^{meas} - y_{j,k}^{model} \right)^2$$

[Equation 19]

We numerically estimated the derivatives of the sensitivity matrix using a central finite difference:

$$\left. \frac{\partial X_j}{\partial \theta_i} \right|_{\hat{\theta}} \approx \frac{X_j(\hat{\theta}_i + \Delta\theta_i) - X_j(\hat{\theta}_i - \Delta\theta_i)}{2\Delta\theta_i}$$

[Equation 20]

This task proved far from trivial. Usually, one would make the $\Delta\theta_i$ step just small enough to obtain accurate values. But it is important to be careful of numerical integration noise, which, in our case, appeared to be considerable. The presence of noise leads to a situation where small step size will produce false and large derivatives. These derivatives will propagate in large FIMs, producing small confidence intervals. Therefore, we always made the effort to check the precision of the estimated derivatives, carefully varying the step size if necessary.

Notice that in order to apply Equation 18, all experimental variables need to have measurements at the same experimental times. This is not our case. However, the frequency of the measurements was very similar and many of them were sampled at the same times. This made it reasonable to interpolate the different measurements using a uniform time interval vector. The interpolated values and times are shown in Figure 2, Figure 3, Figure 4 and Figure 5.

RESULTS AND DISCUSSION

In this section we will present and discuss the main results of the different diagnostic methods used in this paper (see Pre/post-regression diagnostics), along with the modifications performed based on our analysis. Later, we will show how the fitted parameters compare with those reported by Araya et al. (2007) using Orthogonal Collocation on Finite Elements (OCFE). Finally, we contrast the performance of the global optimization technique used here, with standard techniques used in previous publications.

Identifiability analysis

The identifiability analysis showed that several pairs of the 14 model parameters were highly correlated.

Primarily, parameter μ_{\max} was highly correlated with the inhibition constant k_N ($K_{ij} > 0.98$) in all conditions. The implication of this is that there are several combinations of μ_{\max} and k_N values that lead to identical biomass levels. The usual procedure to deal with this situation is to hold constant one of the correlated parameters. The parameter to fix should be the one that is well known in the literature so that a reliable value for the less known parameter is obtained. In our case we fixed μ_{\max} as we found reliable values in Gelmi et al. (2000).

Significant correlation was also found in parameters associated with starch, CO₂ and O₂ levels (Equation 5 and Equation 7) in several culture conditions ($K_{ij} > 0.95$). The highly correlated pairs were: $Y_{X/S}$ and m_S , Y_{X/CO_2} and m_{CO_2} , Y_{X/O_2} and m_{CO_2} . As before, the better known value should be fixed in order to make a reliable estimation for the other. However, no reliable values for those parameters exist in the literature, since to date few investigations in SSC have been done in *G. fujikuroi*. Consequently, other experiments have to be planned under each condition in order to identify one of the parameters independently. In this study we fixed $Y_{X/S}$, Y_{X/CO_2} and Y_{X/O_2} arbitrarily to the values obtained in Gelmi et al. (2002), as shown in Table 1. However, the oxygen yields are similar to (even smaller than) those previously reported in SSF-between 3.8 and 2.7 g/g- (Soccol et al. 1993; Rodríguez-León et al. 1999; Brand et al. 2001; Brand et al. 2002). Furthermore, a recent publication by Chavez-Parga et al. (2008) reported yields similar ($Y_{x/s} = 0.22-1.87$ g/g and $Y_{x/n} = 11-49.4$ g/g) to the ones reported by Gelmi et al. (2002). The confidence intervals shown in Figure 6 show that the estimated biomass/nitrogen yields are reliable.

Sensitivity analysis

Our sensitivity analysis showed that the degradation rate k_p present in the GA₃ equation shows that regardless of the culture conditions evaluated, state variables show no sensitivity on any state variable at any culture condition. Therefore, in order to facilitate the optimization effort, we set k_p to zero. All other parameters showed to affect at least one of the state variables of the model. The results are shown in Figure 7.

Parameter calibration

First, calibrations with the full set of 14 parameters were performed. The regression was unstable, in the sense that very different estimates produced equally good results. This is due to the effects of unidentifiable parameters described in Identifiability analysis. As a direct consequence, we reduced the parameter space to 9 parameters. After the reduction in dimensionality, the calibration results were stable.

Figure 2, Figure 3, Figure 4 and Figure 5 show our calibration curves in comparison with the curves resulting from OCFE. Under all culture conditions, good parameter fits were produced. In Table 6 the objective function values were compared to the ones found in Araya et al. (2007) and Gelmi et al. (2002).

Statistical significance

In this section we present the t -values and confidence intervals for the estimated set of 14 and 9 parameters, while also comparing our t -values to those reported by Araya et al. (2007).

To perform a fair comparison, we calculated the standard deviations for all three parameter sets using the same method described in Methods. For all three parameter sets, t -values are displayed in Table 2, Table 3, Table 4 and Table 5, and confidence intervals are shown in Figure 6 for all four conditions. In the case of the full set of 14 parameters, we calculated all 14 standard deviations leading to an (14 x 14)-FIM matrix, whereas only the statistical information of the 9 relevant parameters is compared. In the case of the 14 parameters found by Araya et al. (2007) the same 5 parameters were held fixed

leading to a (9 x 9)-FIM matrix for the remaining nine in order to make their estimates comparable to ours.

First of all, it can be seen that the significance generally increased by reducing the number of parameters (Table 2, Table 3, Table 4 and Table 5, t -value (14 par.) versus t -value (9 par.)). The reduced parameter set features greater t -values and smaller confidence intervals. However, the resulting confidence intervals are only trustworthy if the reliability of the fixed values is proven in the literature or by independent experiments. In this research, only μ_{max} was able to be assigned reliably. In general, the significance of the parameters found by Araya et al. (2007) was comparable to the one found in this research (Table 2, Table 3, Table 4 and Table 5, t -value (OCFE) versus t -value (14 par.)). Therefore, in this case the parameters will have similar confidence intervals.

Significance problems only occurred in the case of k_i , where its confidence intervals were extremely large compared to its estimated value. As a solution, the sensitivity analysis can help us to plan future experiments so that k_i can be estimated reliably by sampling more frequently in the zone where dGA_3/dk_i is most sensitive (i.e., $t > 50$ hrs, see Figure 1, column GA_3 , parameter k_i).

CONCLUDING REMARKS

The Scatter Search algorithm proved to be an effective and easy tool to handle a global minimization problem for a typical biologic dynamic model. Regression diagnostics are crucial tools to validate and modify the model, as they could eradicate important correlations between parameters and reduce the model's uncertainty. Furthermore, the diagnostic methods could give us important recommendations regarding sampling times for further experiments; this will help to obtain unique parameter estimates and will increase the statistical significance of the model.

REFERENCES

- ARAYA, M.M.; ARRIETA, J.J.; PÉREZ-CORREA, J.R.; BIEGLER, L.T. and JORQUERA, H. (2007). Fast and reliable calibration of solid substrate fermentation kinetic models using advanced non-linear programming techniques. *Electronic Journal of Biotechnology*, vol. 10, no. 1. [\[CrossRef\]](#)
- BANGA, J.; BALSACANTO, E.; MOLES, C.G. and ALONSO, A.A. (2003). Improving food processing using modern optimization methods. *Trends in Food Science & Technology*, vol. 14, no. 4, p. 131-144. [\[CrossRef\]](#)
- BRAND, D.; PANDEY, A.; RODRIGUEZ-LEON, J.A.; ROUSSOS, S.; BRAND, I. and SOCCOL, C.R. (2001). Packed bed column fermenter and kinetic modeling for upgrading the nutritional quality of coffee husk in solid-state fermentation. *Biotechnology Progress*, vol. 17, no. 6, p. 1065-1070.
- BRAND, D.; PANDEY, A.; RODRIGUEZ-LEON, J.A.; ROUSSOS, S.; BRAND, I. and SOCCOL, C.R. (2002). Relationship between coffee husk caffeine degradation and respiration of *Aspergillus sp.* LPBx in solid-state fermentation. *Applied Biochemistry and Biotechnology*, vol. 102, no. 1-6, p. 169-177. [\[CrossRef\]](#)
- CHAVEZ-PARGA, M.C.; GONZALEZ-ORTEGA, O.; NEGRETE-RODRÍGUEZ, M.L.; VALLARINO, I.G.; GONZÁLEZ-ALATORRE, G. and ESCAMILLA-SILVA, E.M. (2008). Kinetic of the gibberellic acid and bikaverin production in an airlift bioreactor. *Process Biochemistry*, vol. 43, no. 8, p. 855-860. [\[CrossRef\]](#)
- CHU, Y. and HAHN, J. (2008). Integrating parameter selection with experimental design under uncertainty for nonlinear dynamic systems. *AIChE Journal*, vol. 54, no. 9, p. 2310-2320. [\[CrossRef\]](#)
- EGEA, J.; RODRÍGUEZ-FERNÁNDEZ, M.; BANGA, J. and MARTÍ, R. (2007). Scatter search for chemical and bio-process optimization. *Journal of Global Optimization*, vol. 37, no.3, p. 481-503. [\[CrossRef\]](#)
- FRANCESCHINI, G. and MACCHIETTO, S. (2008). Model-based design of experiments for parameter precision: state of the art. *Chemical Engineering Science*, vol. 63, no. 19, p. 4846-4872. [\[CrossRef\]](#)
- GELMI, C.; PÉREZ-CORREA, R.; GONZÁLEZ, M. and AGOSIN, E. (2000). Solid substrate cultivation of *Gibberella fujikuroi* on inert support. *Process Biochemistry*, vol. 35, no. 10, p. 1227-1233. [\[CrossRef\]](#)
- GELMI, C.; PÉREZ-CORREA, R. and AGOSIN, E. (2002). Modelling *Gibberella fujikuroi* growth and GA_3 production in solid-state fermentation. *Process Biochemistry*, vol. 37, no. 9, p. 1033-1040. [\[CrossRef\]](#)
- JACQUEZ, J.A. and GREIF, P. (1985). Numerical parameter identifiability and estimability: Integrating identifiability, estimability, and optimal sampling design. *Mathematical Biosciences*, vol. 77, no. 1-2, p. 201-227.
- JAQAMAN, K. and DANUSER, G. (2006). Linking data to models: Data regression. *Nature Reviews Molecular Cell Biology*, vol. 7, no. 11, p. 813-819. [\[CrossRef\]](#)
- JONES, D.; MIRRAZAVI, S.K. and TAMIZ, M. (2002). Multi-objective meta-heuristics: an overview of the current state-of-the-art. *European Journal of Operational Research*, vol. 137, no. 1, p. 1-9. [\[CrossRef\]](#)
- LANDAW, E.W. and DISTEFANO III, J.J. (1984). Multiexponential, multicompartmental, and noncompartmental modeling. II. Data analysis and statistical considerations. *American Journal Physiology*, vol. 246, no. 5, p. R665-R677.

- PETERSEN, B.; GERNAEY, G. and VANROLLEGHEM, P.A. (2001). Practical identifiability of model parameters by combined respirometric-titrimetric measurements. *Water Science & Technology*, vol. 43, no. 7, p. 347-355.
- RODRIGUEZ-FERNANDEZ, M.; EGEA, J. and BANGA, J. (2006). Novel metaheuristic for parameter estimation in nonlinear dynamic biological systems. *BMC Bioinformatics*, vol. 7, no. 1, p. 483. [\[CrossRef\]](#)
- RODRÍGUEZ-LEÓN, J.A.; DOMENECH, F.; LEÓN, M.; MÉNDEZ, T.; RODRÍGUEZ, T.E. and PANDEY, A. (1999). Production of spores of *Trichoderma harzianum* on sugar cane molasses and bagasse pith in solid state fermentation for biocontrol. *Brazilian Archives of Biology and Technology*, vol. 42, no. 1.
- SOCOL, C.; LEON, J.R.; MARIN, B.; ROUSSOS, S. and RAIMBAULT, M. (1993). Growth kinetics of *Rhizopus arrhizus* in solid-state fermentation of treated cassava. *Biotechnology Techniques*, 1993, vol. 7, no. 8, p. 563-568. [\[CrossRef\]](#)
- YUE, H.; BROWN, M.; HE, F.; JIA, J. and KELL, D.B. (2008). Sensitivity analysis and robust experimental design of a signal transduction pathway system. *International Journal of Chemical Kinetics*, vol. 40, no. 11, p. 730-741. [\[CrossRef\]](#)

How to reference this article:

SACHER, J.; SAA, P.; CÁRCAMO, M.; LÓPEZ, J.; GELMI, C.A. and PÉREZ-CORREA, R. (2011). Improved calibration of a solid substrate fermentation model. *Electronic Journal of Biotechnology*, vol. 14, no. 5. <http://dx.doi.org/10.2225/vol14-issue5-fulltext-7>

APPENDIX FIGURES

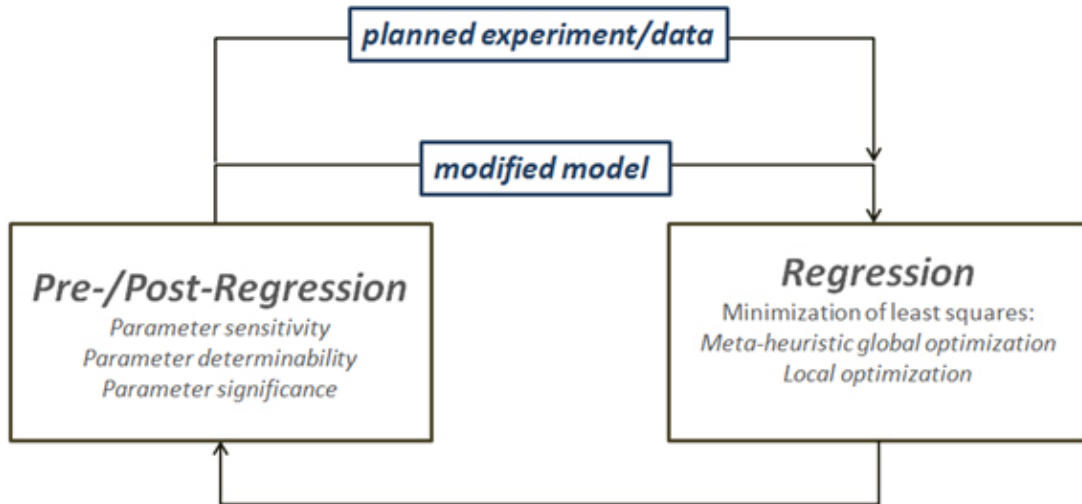


Fig. 1 Iterative cycle of linking data to models.

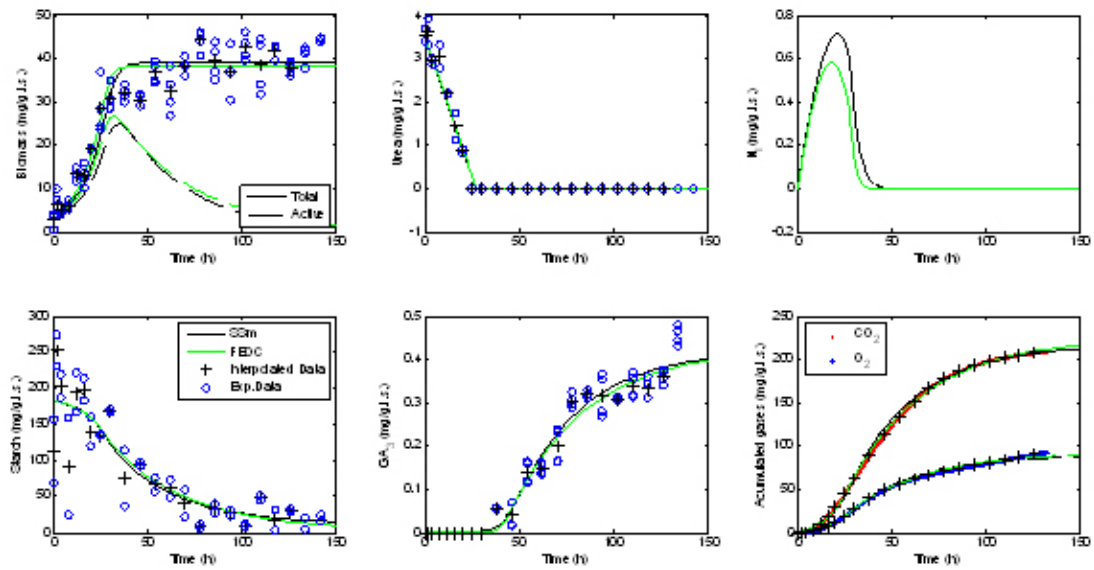


Fig. 2 Parameter calibration for culture condition 1.

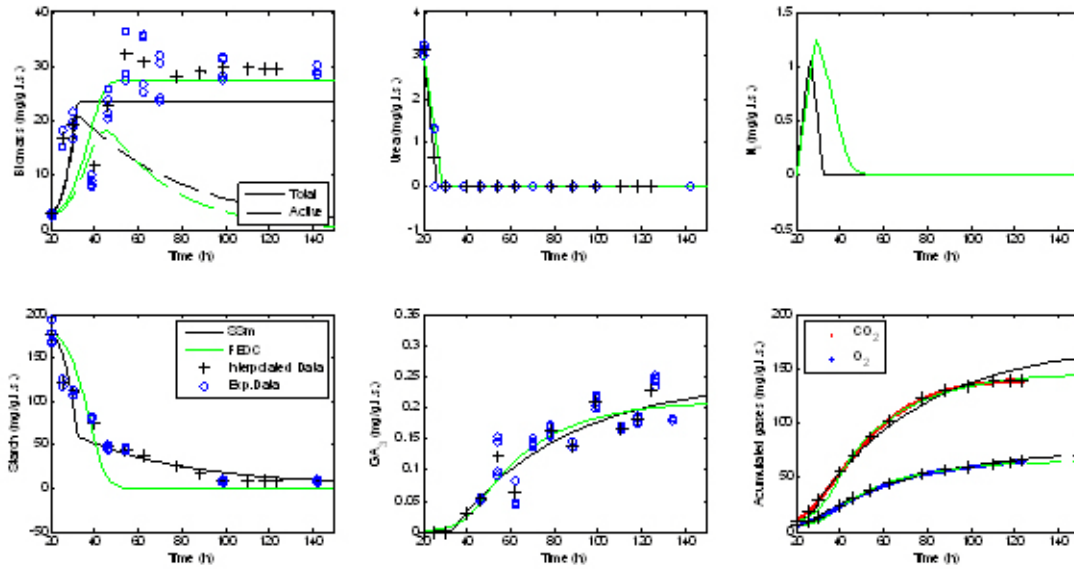


Fig. 3 Parameter calibration for culture condition 2.

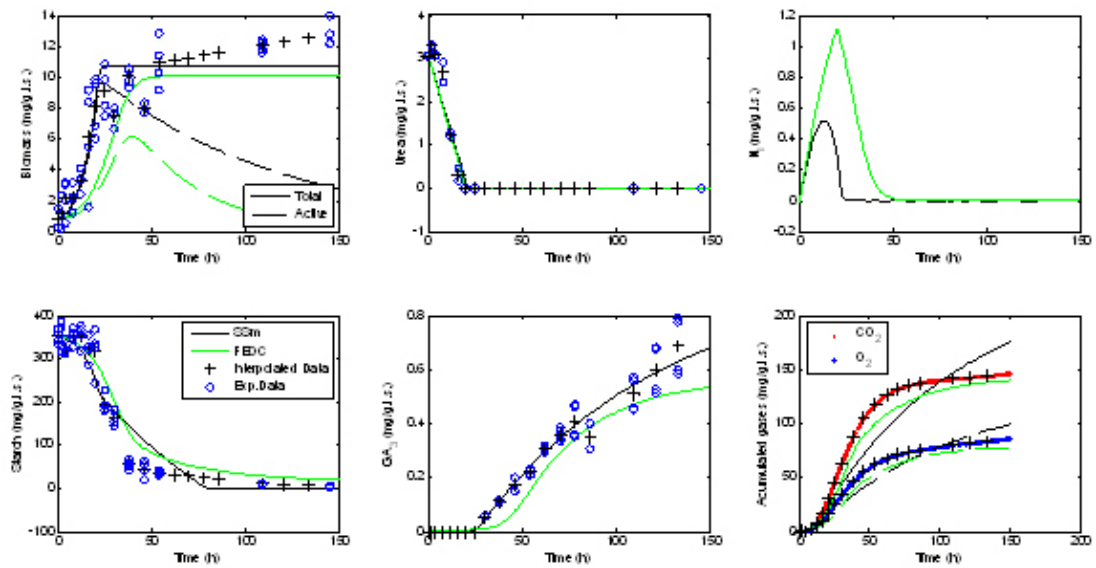


Fig. 4 Parameter calibration for culture condition 3.

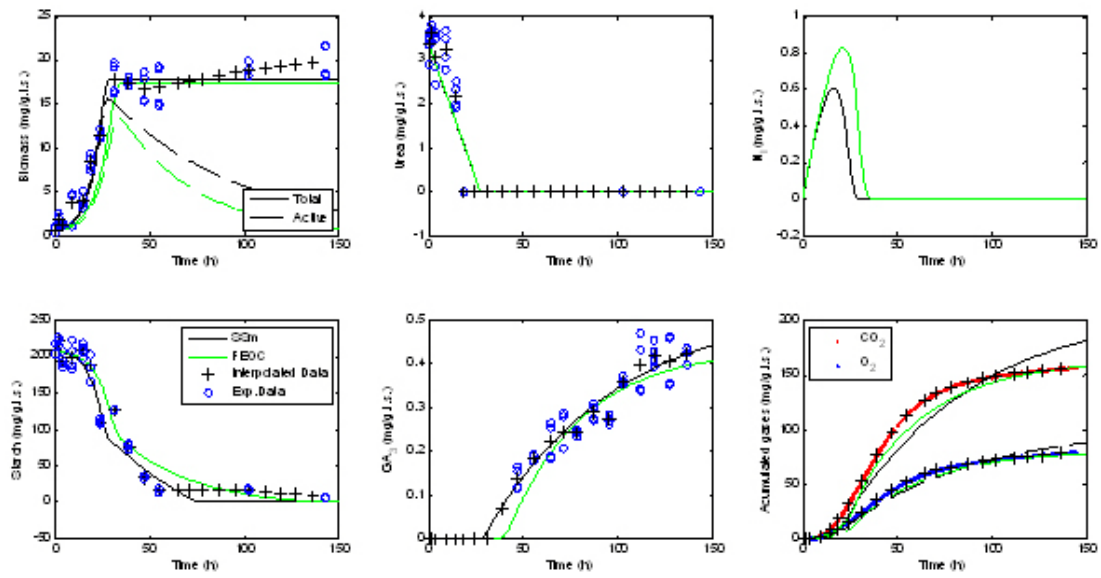


Fig. 5 Parameter calibration for culture condition 4.

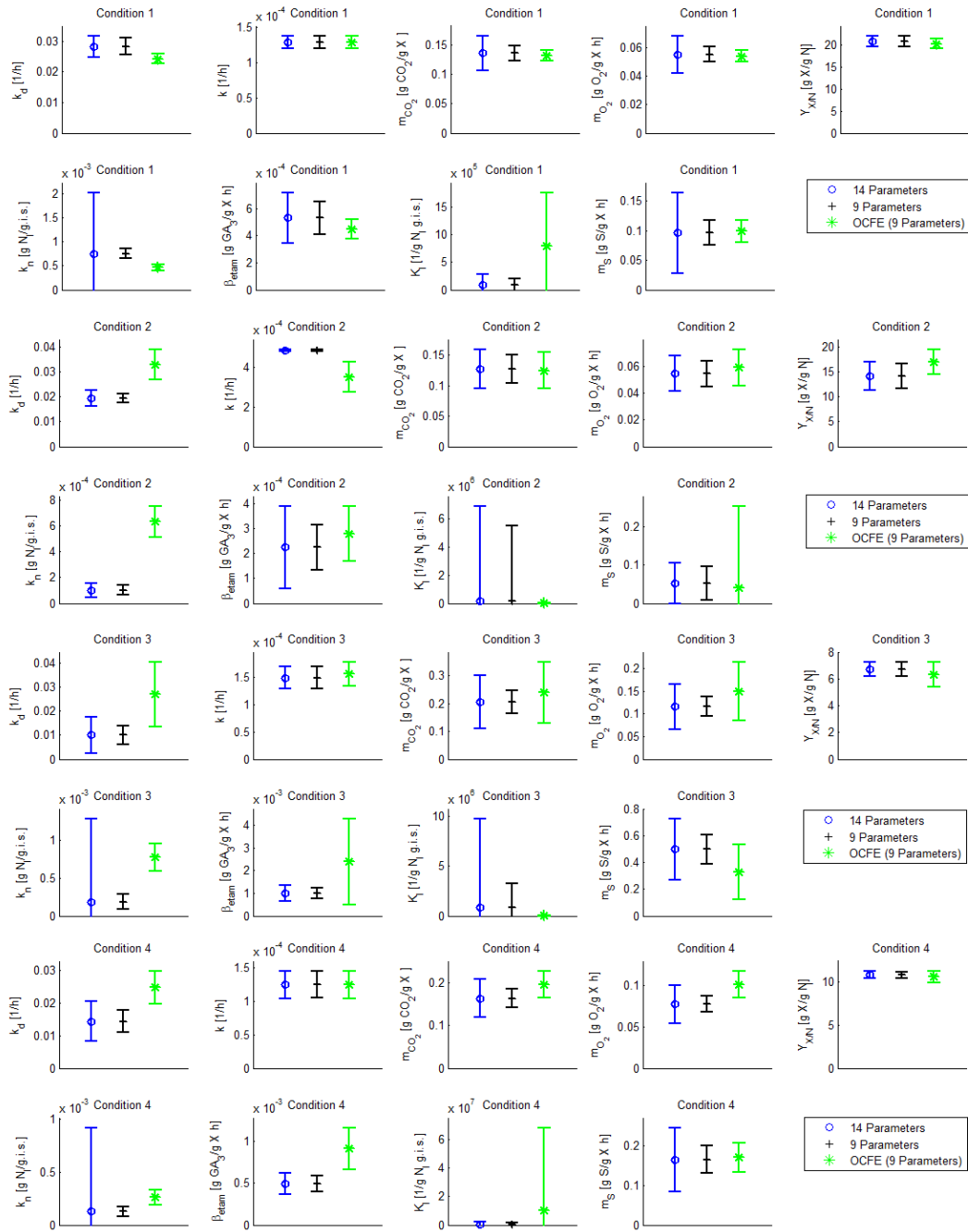


Fig. 6 Confidence intervals of the different parameter sets.

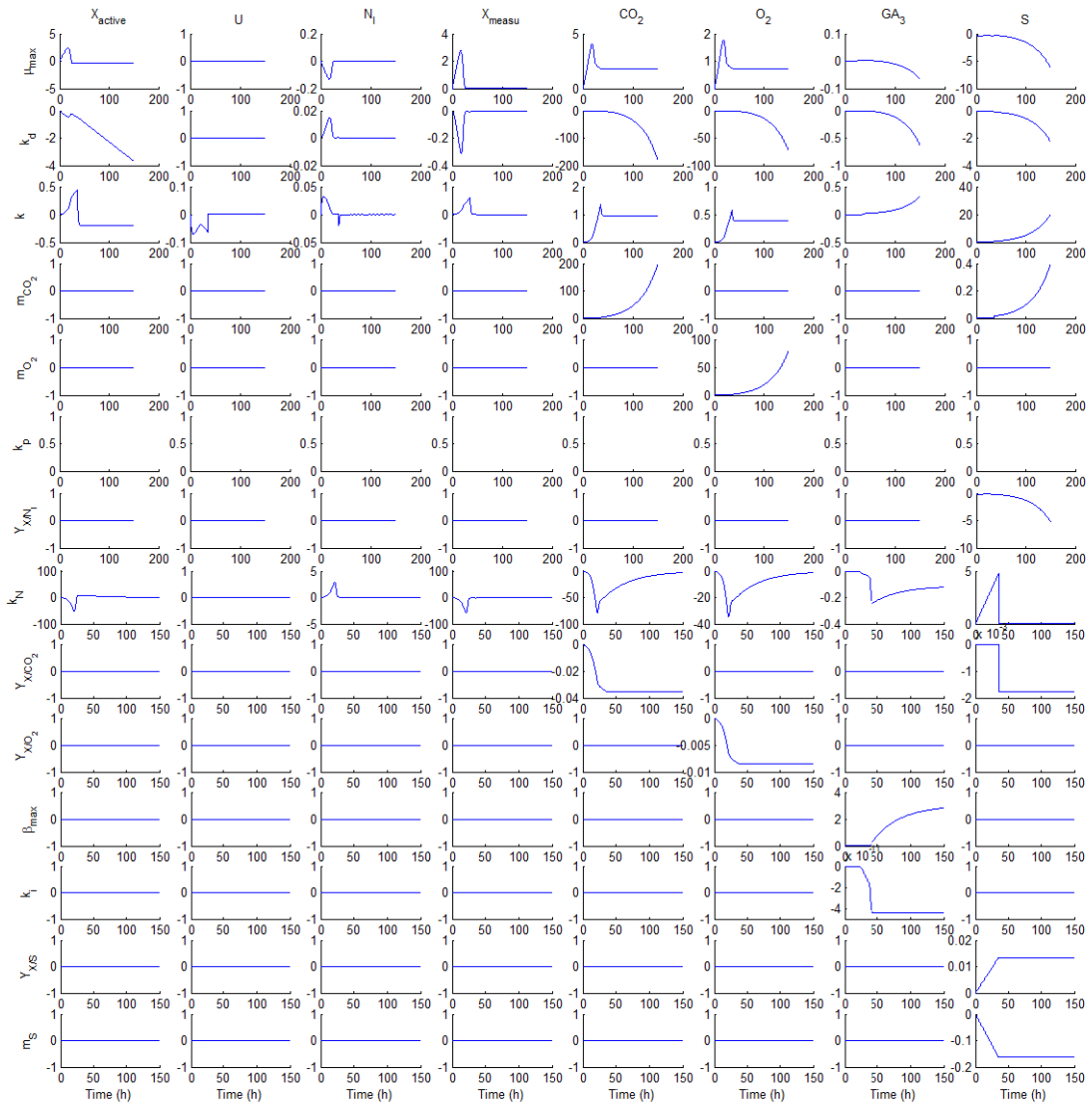


Fig. 7 Time-dependent sensitivities for culture condition 1.

APPENDIX TABLES

Table 1. Fixed parameters for each condition.

Parameter	Condition 1		Condition 2		Condition 3		Condition 4	
	SSm	OCFE	SSm	OCFE	SSm	OCFE	SSm	OCFE
μ_{max} (1/h)	0.21	0.1832	0.23	0.2253	0.19	0.1868	0.195	0.1968
Y_{X/CO_2} (g X/g CO ₂)	1.2	1.8	1.96	0.79	0.745	0.171	1.173	0.67
Y_{X/O_2} (g X/g O ₂)	2.6	3.7	3.8	2.4	1.22	0.38	2.6	2
k_p (1/h)	0	4E-009	0	2E-008	0	4E-010	0	1E-009
Y_{XS} (g X/g S)	0.9	1.4	0.183	0.151	0.101	0.043	0.18	0.18

Table 2. Estimates and statistical parameters for cultivation condition 1.

Parameter	Value	OCFE Value	t-val (OCFE)	t-val (14 par.)	t-val (9 par.)
k_d (1/h)	2.824E - 002	2.430E - 002	2.91E + 001	1.54E + 001	2.02E + 001
k (g/h)	1.290E - 004	1.287E - 004	2.77E + 001	2.77E + 001	2.79E + 001
m_{CO_2} (g CO ₂ /g X h)	1.366E - 001	1.320E - 001	2.74E + 001	9.00E + 000	2.04E + 001
m_{O_2} (g O ₂ /g X h)	5.527E - 002	5.400E - 002	2.67E + 001	8.26E + 000	1.98E + 001
Y_{X/N_I} (g X/g N _I g.i.s.)	2.086E + 001	2.030E + 001	3.48E + 001	3.14E + 001	3.22E + 001
k_n (g N _I / g.i.s.)	7.481E - 004	4.600E - 004	1.24E + 001	1.15E + 000	1.44E + 001
β_{etam} (g GA ₃ /g X h)	5.326E - 004	4.500E - 004	1.23E + 001	5.58E + 000	8.66E + 000
K_I (g N _I /g.i.s.)	8.898E + 004	8.000E + 005	1.63E + 000	9.20E - 001	1.59E + 000
m_S (g S/g X h)	9.613E - 002	9.900E - 002	1.06E + 001	2.75E + 000	9.06E + 000

g.i.s.: grams of inert support.

Table 3. Estimates and statistical parameters for cultivation condition 2.

Parameter	Value	OCFE Value	t-val (OCFE)	t-val (14 par.)	t-val (9 par.)
k_d (1/h)	1.942E - 002	3.300E - 002	1.04E + 001	1.17E + 001	1.99E + 001
k (g/h)	4.874E - 004	3.529E - 004	9.11E + 000	1.97E + 002	1.97E + 002
m_{CO_2} (g CO ₂ /g X h)	1.274E - 001	1.250E - 001	8.08E + 000	7.63E + 000	1.10E + 001
m_{O_2} (g O ₂ /g X h)	5.422E - 002	5.900E - 002	8.54E + 000	8.05E + 000	1.09E + 001
Y_{X/N_I} (g X/g N _I g.i.s.)	1.409E + 001	1.700E + 001	1.32E + 001	9.80E + 000	1.13E + 001
k_n (g N _I / g.i.s.)	1.019E - 004	6.300E - 004	1.00E + 001	3.65E + 000	5.54E + 000
β_{etam} (g GA ₃ /g X h)	2.247E - 004	2.800E - 004	4.99E + 000	2.65E + 000	4.85E + 000
K_I (g N _I /g.i.s.)	1.780E + 005	1.000E + 003	6.17E - 001	1.59E - 001	6.53E - 002
m_S (g S/g X h)	5.233E - 002	4.000E - 002	3.64E - 001	5.76E + 000	2.33E + 000

g.i.s.: grams of inert support.

Table 4. Estimates and statistical parameters for cultivation condition 3.

Parameter	Value	OCFE Value	t-val (OCFE)	t-val (14 par.)	t-val (9 par.)
k_d (1/h)	1.001E - 002	2.710E - 002	3.93E + 000	2.58E + 000	4.94E + 000
k (g/h)	1.485E - 004	1.553E - 004	1.35E + 001	1.42E + 001	1.42E + 001
m_{CO_2} (g CO ₂ /g X h)	2.047E - 001	2.400E - 001	4.28E + 000	4.18E + 000	9.93E + 000
m_{O_2} (g O ₂ /g X h)	1.154E - 001	1.500E - 001	4.56E + 000	4.55E + 000	1.07E + 001
Y_{X/N_I} (g X/g N _I g.i.s.)	6.731E + 000	6.300E + 000	1.31E + 001	2.41E + 001	2.57E + 001
k_n (g N _I / g.i.s.)	1.832E - 004	7.700E - 004	8.46E + 000	3.27E - 001	3.63E + 000
β_{etam} (g GA ₃ /g X h)	9.969E - 004	2.400E - 003	2.46E + 000	5.36E + 000	8.16E + 000
K_I (g N _I /g.i.s.)	8.872E + 005	1.900E + 004	9.14E - 001	1.95E - 001	7.28E - 001
m_S (g S/g X h)	5.000E - 001	3.300E - 001	3.11E + 000	4.22E + 000	8.99E + 000

g.i.s.: grams of inert support.

Table 5. Estimates and statistical parameters for cultivation condition 4.

Parameter	Value	OCFE Value	t-val (OCFE)	t-val (14 par.)	t-val (9 par.)
k_d (1/h)	1.431E - 002	2.470E - 002	9.45E + 000	6.27E + 000	8.28E + 000
k (g/h)	1.246E - 004	1.243E - 004	1.19E + 001	1.19E + 001	1.20E + 001
m_{CO_2} (g CO ₂ /g X h)	1.634E - 001	1.960E - 001	1.21E + 001	7.60E + 000	1.50E + 001
m_{O_2} (g O ₂ /g X h)	7.745E - 002	1.020E - 001	1.23E + 001	7.85E + 000	1.56E + 001
Y_{X/N_I} (g X/g N _I g.i.s.)	1.078E + 001	1.060E + 001	3.12E + 001	5.50E + 001	5.52E + 001
k_n (g N _I / g.i.s.)	1.269E - 004	2.600E - 004	7.56E + 000	2.97E - 001	5.24E + 000
β_{etam} (g GA ₃ /g X h)	4.901E - 004	9.100E - 004	7.19E + 000	6.35E + 000	9.82E + 000
K_I (g N _I /g.i.s.)	2.624E + 005	1.000E + 007	3.36E - 001	1.39E - 001	3.21E - 001
m_S (g S/g X h)	1.641E - 001	1.700E - 001	9.08E + 000	4.16E + 000	9.27E + 000

g.i.s.: grams of inert support.

Table 6. Objective function of the optimization for all cultivation conditions.

Condition	fobj _{SMM}	fobj _{OCFE}	fobj _{GELMI}
1	0.006381	0.006429	0.008763
2	0.013460	0.024046	0.013539
3	0.004888	0.007480	0.015695
4	0.004617	0.005526	0.006327