

IDENTIFIABILITY OF TIME VARYING PARAMETERS IN A GREY-BOX NEURAL MODEL: APPLICATION TO A BIOTECHNOLOGICAL PROCESS

Gonzalo Acuña
Francisco Cruz
Vicente Moreno

Departamento de Ingeniería Informática, Universidad de Santiago de Chile, USACH
Av. Ecuador 3659 – Santiago - Chile
{gacuna, fcruz, vmoreno}@usach.cl

KEYWORDS

Grey-Box, neural networks, identifiability, time-varying parameters, biotechnological processes.

ABSTRACT

Grey Box Neural Models (GBNM) constitute a real alternative for those processes for which the available *a priori* knowledge is incomplete. In this work an application to a biotechnological process has been performed. Good results of the GBNM acting as a software sensor for the non measured state variables has been shown. However even if the estimation performance is good, correct identification of the time varying parameters is not assured. Identifiability of these parameters has to be tested and some proposed techniques are used in this work showing that the specific growth kinetics and the specific production kinetics can be identified although the last one is more difficult because of its dependence on only one measured variable.

INTRODUCTION

In the development of dynamic system models it is better to take advantage of *a priori* knowledge of a process, generally expressed in terms of sets of ordinary differential equations which represent mass or energy balances. In complex biotechnological processes, the most difficult task is the modeling of time varying parameters, such as the specific kinetics. In order to address this problem, Psychogios and Ungar (1992) proposed to use grey-box models which combine *a priori* knowledge expressed in terms of a phenomenological, or white-box model, with a black-box model such as a neural network. These models have proved to be satisfactory for dynamic systems, they have better generalization characteristics, and they can be identified with a smaller amount of data (Psychogios and Ungar, 1992). Thompson and Kramer (1994) classified these grey-box models into two principal categories: those which deliver intermediate values (of parameters or variables) for use in phenomenological models (serial grey-box models), or those in parallel with the dynamic model, adjusted to compensate for modeling errors (parallel grey-box models). Van Can *et al.* (1996) showed that the series strategy resulted in grey-box models with superior results. More recently Thibault *et al.* (2000), and Acuña *et al.* (1999) have employed and analyzed this type of models demonstrating their performance and their use in complex processes.

Another problem that is often encountered concerns the identifiability of those time varying parameters. This could be a great problem because the aim of grey-box models is not only to minimize the difference between the model output variables and some targets but also to obtain a good model of the unknown time-varying parameters in terms of some relevant variables. So the question is how to know when, even if the outputs are correctly estimated, we can trust in the time-varying parameters values and hence in the determined model. So, the objective of this work is to develop a grey-box neural model for a biotechnological process taking care of the identifiability of the time-varying parameters by using some appropriate indices.

This article is organized as follow: first the biotechnological process is described, then a section is devoted to grey-box modeling. Identifiability methods are then presented followed by some results and conclusions.

BIOTECHNOLOGICAL PROCESS

The bioprocess considered is the production of giberelic acid (a vegetal growth hormone) by the filamentous fungi *Gibberella fujikuroi* growing in a solid state batch culture (SSC) at a laboratory level. A simplified model describing the evolution of the main variables is reported in (Gelmi *et al.*, 2002). This phenomenological model based on material balance laws considers 7 state variables: living Biomass (X), measured Biomass (X_m), urea (U), intermediate nitrogen (N_1), soluble starch (S), giberelic acid (GA_3), produced CO_2 and consumed O_2 . Only the last two variables can be directly measured on-line but for including parameter β into the identifiability analysis we will also consider GA_3 as an on-line measured variable. The model equations are the following:

$$\frac{dX_m}{dt} = \mu \cdot X \quad (1)$$

$$\frac{dX}{dt} = \mu \cdot X - k_d \cdot X \quad (2)$$

$$\frac{dU}{dt} = -k \quad (3)$$

$$\frac{dN_i}{dt} = \begin{cases} 0,47 \cdot k - \mu \cdot \left(\frac{X}{Y_{X/N_i}} \right), & si U \geq 0 \\ -\mu \cdot \left(\frac{X}{Y_{X/N_i}} \right), & U(t) = 0, si U < 0 \end{cases} \quad (4)$$

$$\frac{dS}{dt} = -\frac{\mu \cdot X}{Y_{X/S}} - m_s \cdot X \quad (5)$$

$$\frac{dGA_3}{dt} = \beta \cdot X - k_p \cdot GA_3 \quad (6)$$

$$\frac{dCO_2}{dt} = \mu \cdot \frac{X}{Y_{X/CO_2}} + m_{CO_2} \cdot X \quad (7)$$

$$\frac{dO_2}{dt} = \mu \cdot \frac{X}{Y_{X/O_2}} + m_{O_2} \cdot X \quad (8)$$

μ , corresponds to the specific growth rate and its intermediate nitrogen dependence is modeled by a Monod law. β corresponds to the specific production rate of giberelic acid.

$$\mu = \mu_m \cdot \frac{N_i}{k_n + N_i} \quad (9)$$

$$\beta = \frac{\beta_{etam}}{1 + K_i \cdot N_i} \quad (10)$$

The other model parameters were identified on the basis of some specific experiments and experimental considerations. Their values are included in Table 1 for controlled temperature and water activity conditions (T=25°C, Aw=0,992).

Table 1: Model parameters

Name/Description	Value
μ_m	Maximum specific growth rate 0.28 [1/h]
β_{etam}	Maximum specific production rate $6.5 \cdot 10^{-4}$ [g GA ₃ /g X h]
k	Urea degradation constant $1.33 \cdot 10^{-4}$ [g/h]
k_n	Constant $1.1 \cdot 10^{-3}$ [g N _i /g.i.s.]
k_d	Dead constant 0.031 [1/h]
k_p	Giberelic acid degradation rate $4.4 \cdot 10^{-4}$ [1/h]
k_i	Intermediate nitrogen production rate $7.86 \cdot 10^5$ [1/g N _i g.i.s.]
m_s	S maintenance coefficient 0.11 [g S / g X h]
m_{CO_2}	CO ₂ maintenance coefficient 0.11 [g CO ₂ / g X h]
m_{O_2}	O ₂ maintenance coefficient 0.06 [g O ₂ / g X h]
Y_{X/N_i}	Yield coefficient 20.8 [g X / g N _i]
$Y_{X/S}$	Yield coefficient 1.21 [g X / g S]
Y_{X/CO_2}	Yield coefficient 0.58 [g X / g CO ₂]
Y_{X/O_2}	Yield coefficient 2.11 [g X / g O ₂]

GREY BOX NEURAL MODEL

As previously mentioned GBNMs take advantage of the combination of *a priori* knowledge surrounding a given process expressed in terms of a set of differential equations that represent the first principles that govern that process with neural networks. The latter are responsible for the modeling of the interaction between variables that are relevant to the system, and certain parameters whose expressions are difficult to model. It is a well established fact that neural networks are capable of approximating non-linear functions. In particular, it has been demonstrated that perceptrons, with only one hidden layer and an adequate number of neurons in their internal layer, are universal approximators (Hornik *et al.*, 1989).

For the purposes of the present work it is important to distinguish between two training modes for neural networks inserted in GBNMs. The first type, also known as the direct learning mode (Acuña *et al.*, 1999), uses the error generated at the output of the neural network for the correct determination of its weights (Figure 1).

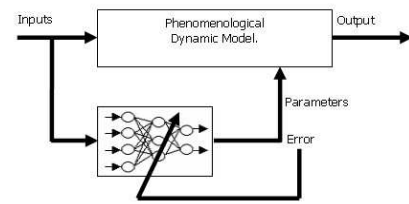


Figure 1: Grey-Box Neural Model in its direct learning mode.

The second type corresponds to an indirect mode by which the error generated at the output of the GBNM is used for the training of the neural network (Acuña *et al.*, 1999) (Figure 2).

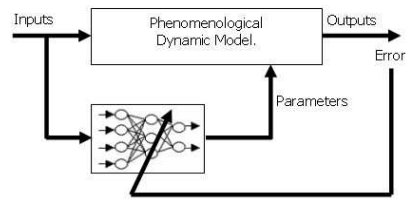


Figure 2: Grey-Box Neural Model in its indirect learning mode.

In the present work the indirect learning mode of the neural network is used. The neural networks used are multi-layered perceptrons with only one hidden layer. The training algorithm is error backpropagation combined with a Levenberg-Marquardt optimization.

The validation of the results obtained is carried out with tests that consist in evaluating the error produced when using the GBNM as a software sensor for the non-measured variables. The error index used is the Index of Agreement (IA), which is presented below :

$$IA = 1 - \frac{\sum_{i=1}^n (o_i - p_i)^2}{\sum_{i=1}^n (|o_i| + |p_i|)^2}$$

Where o_i and p_i are the observed and predicted values respectively, in time i , and N is the total number of data. $p_i' = p_i - o_m$ and $o_i' = o_i - o_m$, where o_m is the median value of the observations.

IDENTIFIABILITY METHODS

The identifiability of model parameters is determined using techniques proposed by (Brun *et al.*, 2002, Reichert and Vanrolleghem, 2001) based on sensitivity and uniqueness analyses. In fact model identifiability will be quantified by measuring the sensitivity of the model output variables to the time varying parameters determined by the neural network part of the grey-box model.

If we considered the measured output variables as $y = \varphi(x, t, \theta)$, where x corresponds to the state variables at time t and θ to the unknown parameter vector then the sensitivity matrix (dimensionless) is:

$$SS_{ij} = \left(\frac{\partial \hat{y}_i}{\partial \theta_j} \right) \theta_j w_i^{1/2} \quad (11)$$

$i = 1 \dots$ number of observations

$j = 1 \dots$ number of parameters

w_i correspond to a scale factor associated with the i^{th} output variable and is defined as the inverse of the measurement error variance.

A dimensionless index of sensitivity, defined by (Brun *et al.*, 2002) is:

$$\delta_j^{msqr} = \sqrt{\frac{1}{N} \sum_{i=1}^N S_{ij}^2} \quad (12)$$

Low sensitivity implies low influence of the parameter to the corresponding output hence low identifiability of this parameter.

Uniqueness of the parameters depends on their correlation (Zhang *et al.*, 2003) which is computed from the variance-covariance matrix of the estimated parameters.

$$Corr(\theta_u, \theta_v) = \frac{\text{cov}(\theta_u, \theta_v)}{\sqrt{\text{cov}(\theta_u, \theta_u) \cdot \text{cov}(\theta_v, \theta_v)}}$$

A high correlation between the parameters means that they cannot be uniquely identified from the available observations.

RESULTS

Grey-Box Neural Model

The detailed GBNN considering equations (1)-(8) and the fact that μ and β depend on N_i is shown in Figure 3. It is to notice that the discretized model of eqs. (1)-(8) is represented as a neural network with fixed weights. Only the black-box has variable weights which can be identified by backpropagation considering appropriate activation and transfer functions. 700 data points obtained from simulation of the complete model (eqs (1)-(10)) were used for training purposes while 300 data points were left for validation.

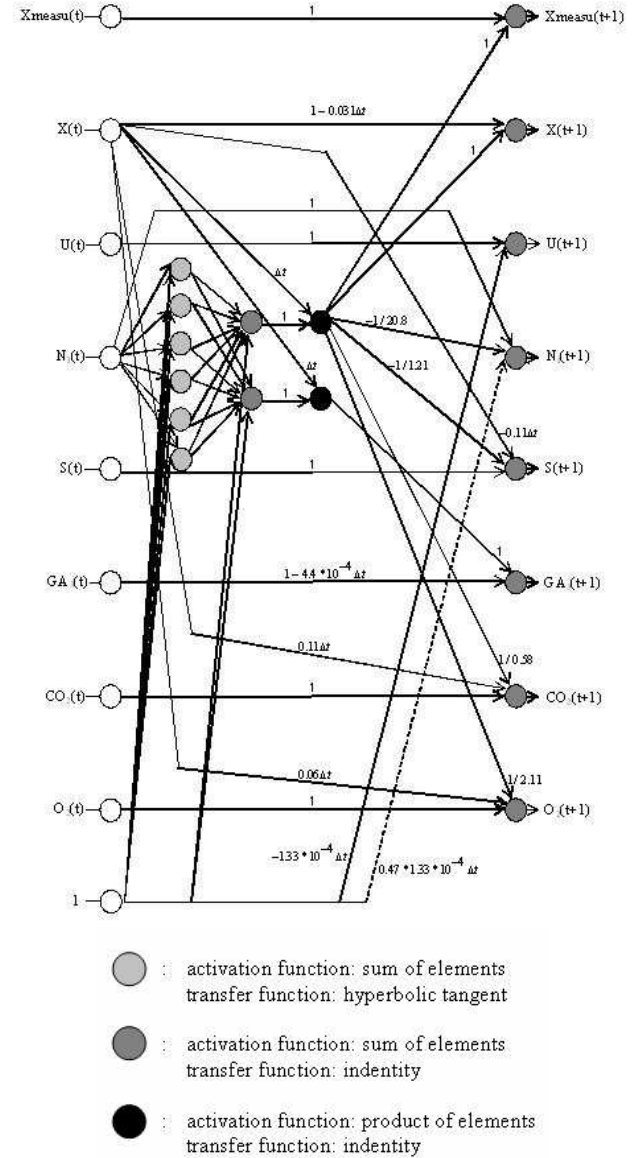


Figure 3: Grey-Box Neural Model for the SSC process.

Two different tests have been built to evaluate the performance and the robustness of the GBNN acting as a software sensor for the first five state variables (eqs (1)-(5)). That means that the GBNN is used in a Model Predictive Output (MPO) mode for those five variables (only from initial conditions) and in an One Step Ahead (OSA) mode

(Billings *et al.*, 1992) for the other three variables (eqs (6)-(8)):

1. Results under ideal conditions (0% error and no perturbations).
2. Results with an initial 250% error on the living biomass and affected by a 5% amplitude gaussian noise on all the state variables.

The initial value used for the living biomass (without noise and any perturbation) is 0.01 (gr/gr).

Test 1

For concision reasons only the results on the three most relevant variables or parameters for this analysis will be shown. A very good coherence can be seen between the simulated and estimated values of the living biomass (X) and the specific growth rate as it is shown in figures 4 and 5. Good results are also obtained for the estimation of the other state variables ($IA > 0.99$). A non as good estimation of the specific production rate can be observed. This is probably due to the exclusive dependence of this parameter to only one output variable (GA_3) as it is shown in Table II.

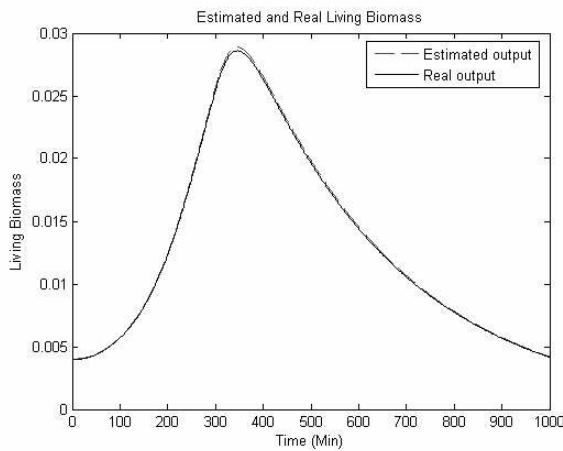


Figure 4: Simulated (Continuous lines) and estimated (dotted lines) living biomass

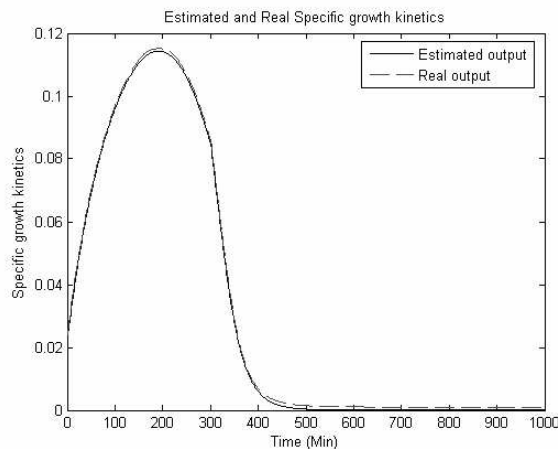


Figure 5: Simulated (Continuous lines) and estimated (dotted lines) specific growth rate.

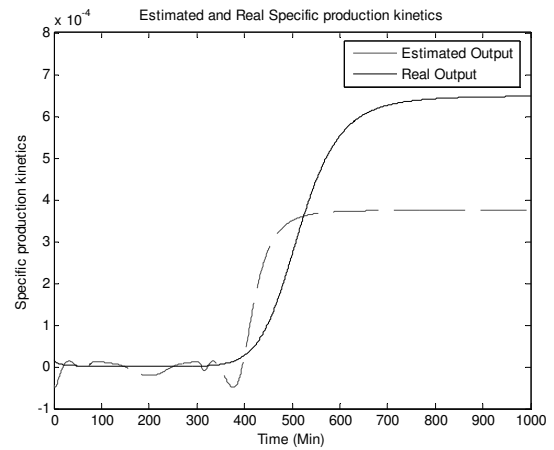


Figure 6: Simulated (Continuous lines) and estimated (dotted lines) specific production rate.

Test 2

A large initial perturbation on the state variable (living biomass) is quickly compensated by the GBNM acting as a software sensor and a 5% noise affecting the output is well filtered by the method. The error of 250% in the initial living biomass, although is not a real case, allows to clearly show the properties of convergence and stability of the method, supposing that in a practical application a great error in the initial conditions is committed, in the presence of noisy measurements. General results are shown in figure 7, 8 and 9. The same above mentioned remark concerning the specific production kinetics β has to be stated

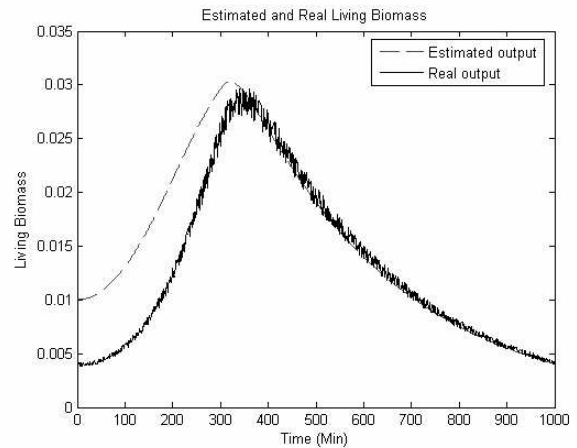


Figure 7: Simulated (Continuous lines) and estimated (dotted lines) living biomass

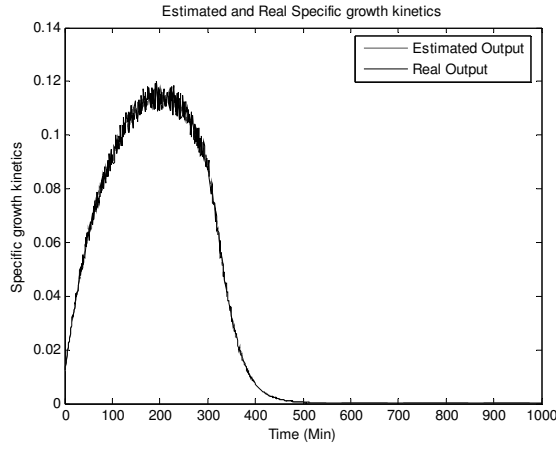


Figure 8: Simulated (Continuous lines) and estimated (dotted lines) specific growth rate.

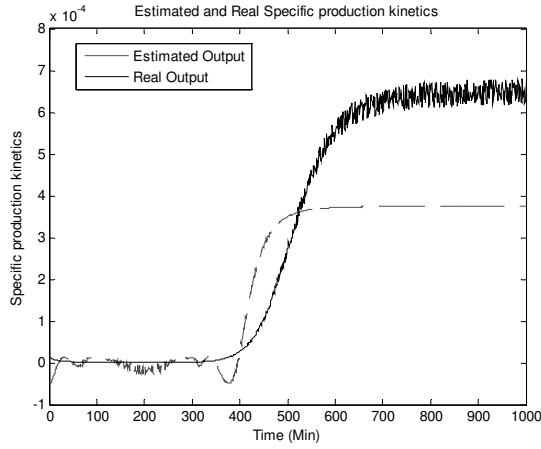


Figure 9: Simulated (Continuous lines) and estimated (dotted lines) specific production rate.

Identifiability Analysis

For computing the sensitivity matrix the derivative of the output to the corresponding parameters was determined from the following finite difference approximation:

$$\frac{\partial y_i}{\partial \theta_j} \approx \frac{y(x, t, \theta_j + \Delta \theta_j) - y(x, t, \theta_j)}{\Delta \theta_j} \quad (13)$$

A small enough parameter perturbation $\Delta \theta_j$ was used in order to assure a small truncation error in the finite difference approximation. The considered outputs and parameters were respectively CO_2, O_2, GA_3 and the specific kinetics μ and β . Hence the sensitivity matrix was constructed from the determination of:

$$\frac{\partial CO_2}{\partial \mu_i}, \frac{\partial CO_2}{\partial \beta_i}, \frac{\partial O_2}{\partial \mu_i}, \frac{\partial O_2}{\partial \beta_i}, \frac{\partial GA_3}{\partial \mu_i}, \frac{\partial GA_3}{\partial \beta_i} \quad (14)$$

In Table II results for the sensitivity index δ_j^{msqr} showing that β has no influence over CO_2 and O_2 are presented. It can also be noticed that μ_i has no influence over GA_3 .

Table 2: Sensitivity index δ_j^{msqr} for different output variables and parameters.

δ_j^{msqr}	μ_i	β_i
CO_2	0.0020	0
O_2	0.0012	0
GA_3	0	0.1311

Sensitivities of each parameter to the corresponding output variables are shown in Figures 10, 11 and 12. The greater influence of the parameters over the output variables ranges from 100 to 400 (min) which corresponds to the period of greater biomass growth.

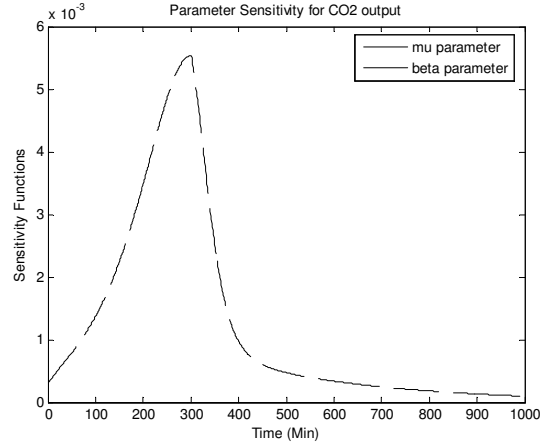


Figure 10: Relative sensitivity for CO2 output.

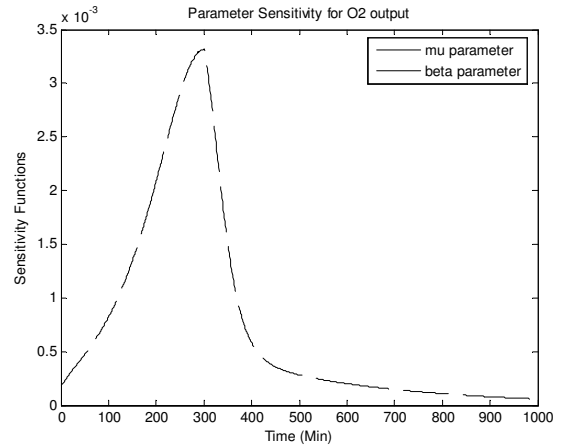
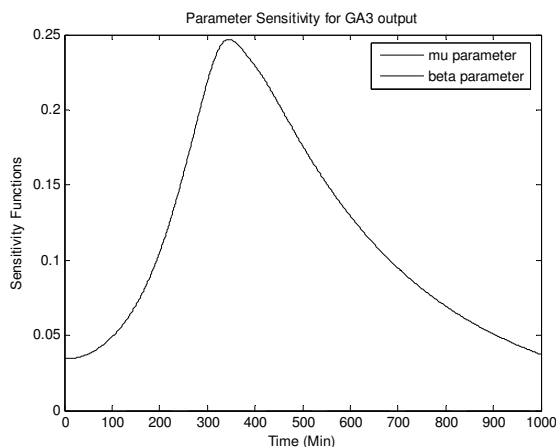


Figure 11: Relative sensitivity for O2 output.



Figures 12: Relative sensitivity for GA3 output.

Uniqueness of parameters was analyzed from the following correlation Table (III).

Table 3: Correlation coefficients between estimated parameters.

$Corr(\theta_u, \theta_v)$	μ	β
μ	1	-0.8554
β	-0.8554	1

It can be seen that the low absolute value of the correlation index between the parameters obtained (< 0.95) allows them to be uniquely determined from an adequately identification procedure.

CONCLUSIONS

GBNM constitute a real alternative for those real world processes for which the available *a priori* knowledge is incomplete, for example in a variety of industrial processes. As in GBNM only some of the physical and/or chemical laws that represent the model are known, and there are unknown parameters that must somehow be estimated, multi-layered perceptron neural networks have been employed for their notable capacity to approximate complex functions on the basis of observed data.

In this work an application to a biotechnological process has been performed. Good results of the GBNM acting as a software sensor for the non measured state variables has been shown. Convergence and noise rejection capacities were also some of the valuable features of this kind of software sensor. However even if the estimation performance is good, correct identification of the time varying parameters is not assured. Identifiability of these parameters has to be tested and some proposed techniques are used in this work showing that both time-varying parameters, the specific growth kinetics μ and the specific production kinetics β can be identified although the last one is more difficult to be correctly identified because of its dependence on only one measured variable (GA_3).

Acknowledgements: The authors would like to acknowledge grants from the Chilean Government (FONDECYT-1040208).

REFERENCES

- Acuña A., Cubillos F., Thibault J., Latrille E., 1999, Comparison of Methods for Training Grey-Box Neural Networks Models, Computers and Chemical Engineering Supplement, 23:561-564.
- Billings, S., Jamaluddin, B and Chen, S., 1992, "Properties of neural networks with applications to modeling non linear dynamical systems", International Journal of Control, 55(1):193-224.
- Brun, R.; Kühni, M.; Siegrist, H.R.; Gujer, W. and Reichert, P. 2002. "Practical identifiability of ASM2d parameters - systematic selection and tuning of parameter subsets". Water Research 36(16):4113-4127.
- Gelmi, C., Perez-Correa, R. and Agosin, E., 2002, "Modelling *Gibberella fujikuroi* growth and GA3 production on solid-state fermentation", Process Biochemistry, 37(9):1033-1040.
- Hornik, K., Stinchcombe, M. and White, H., 1989, "Multilayer feedforward networks are universal approximators", Neural Networks, 2:359-366.
- Psichogios D., Ungar L., 1992, A Hybrid representation approach for modeling complex dynamic bioprocesses. Bioprocess Engineering, 22: 547-556.
- Reichert P. and Vanrolleghem P. 2001. "Identifiability and uncertainty analysis of the river quality model no.1 (RQM1)". Water Sci Technol, 43(7):329-338.
- Thibault, J., Acuña, G., Pérez-Correa, R., Jorquera, H., Molin, P., Agosin, E., 2000, "A hybrid representation approach for modelling complex dynamic bioprocesses" Bioprocess Engineering, 22(6):547-556.
- Thompson M., Kramer M., 1994, Modeling Chemical Processes Using Prior Knowledge and Neural Networks, Computer & Chemical Engineering, 40:1328-1340.
- Van Can H, Braake H., Dubbelman S., Hellinga C., Luyben K., Heijnen J., 1996, Understanding and Applying the Extrapolation Properties of Serial Gray-Box Models, AIChE journal, 44: 1071-1089.
- Zhang, Z.F.; A.L. Ward. and G.W. Gee. 2003. "Estimating soil hydraulic parameters of a field drainage experiment using inverse techniques". Vadose Zone J. 2:201-211.