J. Basic. Appl. Sci. Res., 2(8)7772-7780, 2012

ISSN 2090-4304
Journal of Basic and Applied Scientific Research
www.textroad.com

Application of a Continuous-Discrete Recursive Prediction Error (RPE) Algorithm for Toxicity Detection

Mochamad Bagus Hermanto[1,*], J.D. Stigter[2]

[1]Department of Agricultural Engineering, University of Brawijaya, Malang (INDONESIA)
[2]Biometris Group, University of Wageningen, Wageningen (THE NETHERLANDS)

ABSTRACT

A recursive prediction error algorithm (RPE) has been used as a state/parameter reconstruction tool and also used in combination with a Model Predictive Control (MPC) to solve a problem of optimal input design for a fed-batch bioreactor. In case of environmental engineering, a declining biomass concentration can be used as an indication for toxicity in a system. But the biomass concentration is difficult to be measured. Therefore, it is a challenge to apply the method for toxicity detection. The objective of this work was to apply a so-called continuous-discrete RPE algorithm as a state and parameter reconstruction tool for systems with continuous dynamics and a discrete-time sampling strategy as a toxicity detection mechanism. The RPE algorithm was applied in a non-linear type of bioreactor which are a fed-batch bioreactor and continuous stirred-tank reactor (CSTR). By using only dissolved oxygen as the measured data, state reconstruction by a continuous-discrete RPE algorithm can obtain information about the system, including biomass concentration and substrate concentration. Toxicity can be well recognized when the RPE algorithm is tested in a system which have less hazardous toxic and which deal with highly noisy measurement data. Toxicity can be detected not only by monitoring the biomass concentration and the gain matrix for biomass concentration, but also indicated by the local parametric sensitivity.

Keywords: recursive prediction error, bioreactor, toxicity detection.

INTRODUCTION

A mathematical model is a representation of a real system which is usually focused on a set of selected properties and features of the latter. Models are the essential components for modern process systems engineering methods (i.e. simulation, optimisation and control), and they are usually classified into three categories [1]:

- First-principles (or white-box) models, which are derived from well known physical and chemical relationships, reflecting the underlying principles that govern the process behaviour.
- Data-driven (or black-box) models, which are of empirical nature (e.g. artificial neural networks, time series).
- Hybrid (grey-box) models: a combination of the above.

When complex design problems are introduced, then mathematical models are useful in explaining the underlying mechanisms of the observed system's behaviour.

The model is needed in order to support decisions that have to be taken online during the operation. The model is then updated at each time instant when some new data become available. The update of the state variables and parameters in the system model is performed by recursive algorithm.

In environmental and biochemical case studies, it is difficult to detect toxicity in a system. In the experimental practice, a declining biomass concentration can be used as an indication for toxicity in a system. But the biomass concentration is difficult to be measured.

From a study to determine the bio kinetic parameters μ_{max} and K_s, an identification experiment with a so-called Rapid Oxygen Demand TOXicity (RODTOX) device can be conducted and respirometric data can be obtained [2]. A respirometer is usually installed at the influent of a treatment plant. Toxicity is evaluated by comparing the respirometric data for the influent with the respirometric data obtained during a calibration cycle with a reference substrate [3]. Generally, a respirogram characterizes the healthy state of biomass concentration and may be used, for instance, to identify a toxic alarm, meaning that biomass does not perform healthily because of the presence of a toxic substance in the feeding substrate [2]. Therefore, an output, which is respirometric data, can be observed to detect toxicity in a system.

Another important issue is a method which not only has a capability as a toxic alarm but also reduce the effect of noise at the measured output. The noise may contribute to the bias and inaccuracy of the estimated parameters in a mathematical model. The parameter in a model contributes to the quality of the model itself to be used in the simulation.

In Stigter and Beck [2], a recursive prediction error algorithm was used as a state/parameter reconstruction tool for systems with continuous dynamics and a discrete-time sampling strategy to estimate parameters. A

***Corresponding Author:** Mochamad Bagus Hermanto, Department of Agricultural Engineering, University of Brawijaya, Malang, Indonesia. Email: mbhermanto@ub.ac.id

recursive prediction error algorithm was also used in combination with a Model Predictive Control (MPC) algorithm to solve a problem of optimal input design for a fed-batch bioreactor case study [2]. Hence, we aim to develop the potential for this method for a reliable toxicity detection mechanism that helps the environmental engineer to decide whether a toxic event has occurred yes or no.

The research aims to apply a so-called continuous-discrete recursive prediction error (RPE) algorithm as a state and parameter reconstruction tool for systems with continuous dynamics and a discrete-time sampling strategy as a toxicity detection mechanism.

MATERIALS AND METHODS

Stigter and Beck [2] have described the RPE algorithm. In the model it is assumed that the system model is in a continuous-discrete innovations format:

$$\frac{dx(t)}{dt} = f(x(t),u(t),\theta) + K_\infty \eta(t) \quad \text{..... Equation 1}$$

$$y(t_k) = h(x(t_k),u(t_k),\theta) + \eta(t_k) \quad \text{..... Equation 2}$$

Where

$x(t)$ is the n-dimensional state vector

$y(t_k)$ is m-dimensional observation vector containing the measurements of m-instruments at time constant t_k

$u(t)$ is r-dimensional input vector

θ is p-dimensional parameter vector (assumed constant)

K_∞ is assumed to be an extra set of parameters to be estimated by proposed recursive algorithm which has dimension n x m

$\eta(t_k)$ is white Gaussian noise with spectral density and variance $R_c(t)$ and $R_d(t_k)$ respectively which has dimension m x m

The aim of the RPE is to design an algorithm that is primarily developed for parameter estimation instead of state estimation as in the extended Kalman filter [4]. In this paper, the algorithm is developed for the toxic detection through estimation of the gain matrix K_∞. For that purpose, the gain matrix K_∞ is parameterized into θ.

The RPE algorithm is used to estimate the parameters in the single Monod model which is a non-linear type of bioreactor model as described by Dochain et al. [5] and Stigter et al., [2]. In combination with oxygen uptake rate (OUR) measurements, this model is identifiable [5, 6]**:**

$$\dot{S} = -\frac{\mu_{\max} x}{Y}\frac{S}{S+K_S} \quad \text{..... Equation 3}$$

$$OUR = -(1-Y).\dot{S} \quad \text{..... Equation 4}$$

A fed-batch type of biokinetic model was also proposed for the Rapid Oxygen Demand TOXicity or RODTOX [2] and substrate (including dissolved oxygen) is fed into the fed-batch bioreactor. In this paper, one initial impulse substrate concentration is given at the very beginning of the experiment. We assume that the dissolved oxygen concentration in the feeding substrate can be affected by the solute or the presence of toxic. The consumption of substrate by biomass reduces in time and directly affects the dissolved oxygen concentration in the vessel. Some additional state variables are given to describe the dynamic model of the process:

$$\dot{x} = (\bar{x} - x) - k_{tox}.u_1 \quad \text{..... Equation 5}$$

$$\dot{DO} = k_{La}.(C_{sat} - DO) - OUR \quad \text{..... Equation 6}$$

where,

C_{sat} = saturation concentration of dissolved oxygen (DO) [mg.l^{-1}]
DO = dissolved oxygen concentration [mg.l^{-1}]
k_{tox} = toxic concentration constant [mg.l^{-1}l]
k_{La} = reaeration constant [l.min^{-1}]
OUR = oxygen uptake rate [mg.l^{-1}.min^{-1}]
x = biomass concentration [mg.l^{-1}]
u_1 = toxic introduction [-]

For a continuously stirred-tank reactor (CSTR) (Figure 1), the substrate is fed into the bioreactor with a feed rate $F_{in}(t)$ and washed out with a feed rate $F_{out}(t)$.

$$\dot{S} = \frac{F}{V}(S_{in} - S) - \frac{\mu_{\max} x}{Y}\frac{S}{S + K_S} \quad \text{..... Equation 7}$$

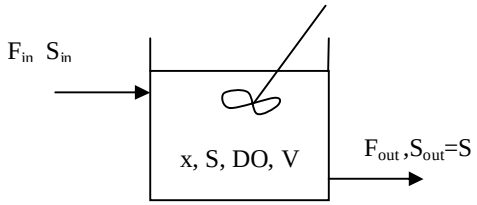


Figure 1 Continuous stirred-tank reactor (CSTR).

So-called Dirac delta function is used in order to make a substrate feeding of multiple cycles. The Dirac delta function is a generalized function representing an infinitely sharp peak bounding unit area, where a function has the value zero everywhere except for several points.

The value μ_{max} and K_s which are obtained from the literature in Table 1 and the initial value of the biomass concentration is 4000 g/l according to [2]. The oxygen concentration is always maintained above 3 mg/l so that the oxygen is not limiting.

The experiments are performed in such a way that [6]:

- The change in biomass concentration can be assumed negligible due to small value of μ_{max} and short period of experimental time.
- The oxygen uptake rate (OUR) data are due to exogenous respiration, only when substrate is induced.

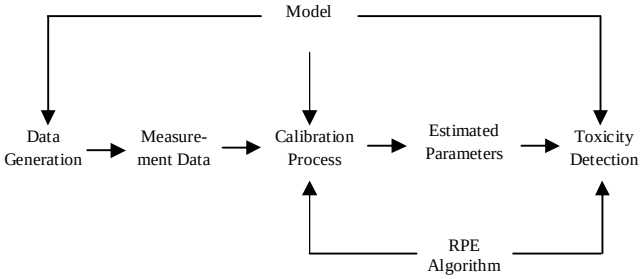
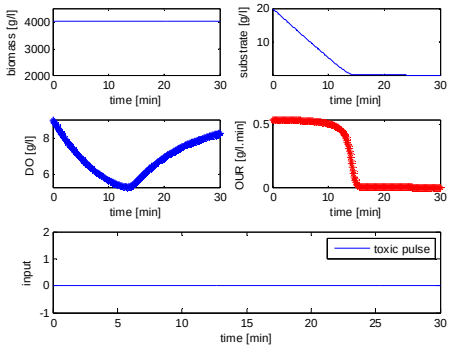


Figure 2 Flowchart of experiment.

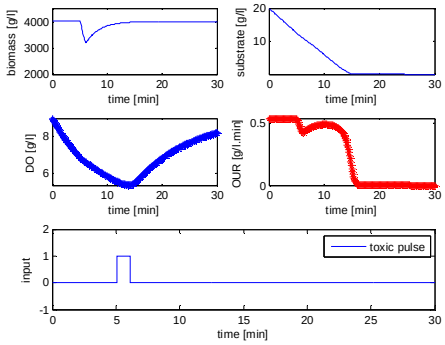
RESULTS AND DISCUSSION

A. Data Generation

The measurement dissolved oxygen data are generated with the model described in Figure 2. A continuous-discrete measurement dataset is obtained. The biomass concentration decreases when toxic substances come into the system. The toxic is given by making u_1 =1 in Equation 5 in a short time period. Then, datasets are generated with 1% of white measurement noise (Figure 3).



(a) without toxic substance



(b) With toxic substance

Figure 3 Single cycle experimental data.

B. Calibration Process

RPE algorithm is used to estimate the parameters known as calibration process. It is known that in the system, RPE algorithm needs initialisation at the initial stage for the initial state ($\hat{x}_0$), initial parameters ($\hat{\theta}(0)$) which consists of 2 bio kinetic parameters, and 3 gain matrix elements (K_∞), variance-covariance matrix ($P(0)$), Gaussian noise variance (R_d) with the following values:

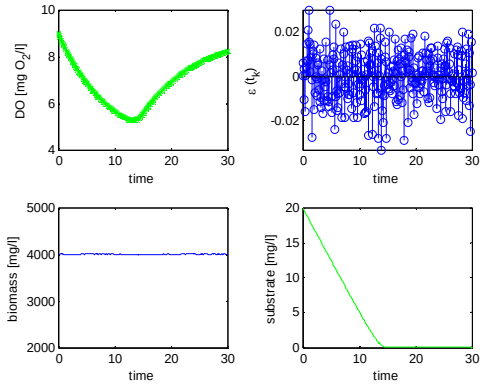
$$\hat{x}(0) = \begin{bmatrix} 4000 \\ 20 \\ 9 \end{bmatrix}; \hat{\theta}(0) = \begin{bmatrix} 1E-4 \\ 0.3 \\ 0 \\ 0 \\ 0 \end{bmatrix}; \; P(0) = \begin{bmatrix} 5E-8 & 0 & 0 & 0 & 0 \\ 0 & 5E-3 & 0 & 0 & 0 \\ 0 & 0 & 5E-1 & 0 & 0 \\ 0 & 0 & 0 & 5E-3 & 0 \\ 0 & 0 & 0 & 0 & 10 \end{bmatrix},$$

the gain matrix K_∞ is estimated recursively together with the parameter θ. So, P has 2 bio kinetics parameter and 3 extra parameter K_∞ diagonally; $R_d = [1E-2]$.

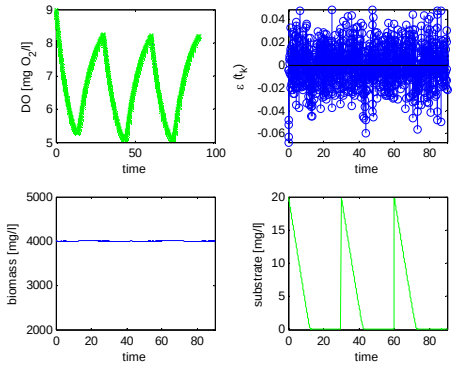
The initial values of the variance-covariance matrix P elements for the parameter θ (including the gain value K_∞) gives the opportunity to modify the search direction range for error minimization system inside the algorithm based on the parametric sensitivity value calculation. A big value of P causes a sharp movement from recent value to the next prediction and a small value of P causes a slow movement. In order to obtain a good calibration, the calculated error $\varepsilon(t_k)$ has to be zero in average (over time) due to the properties of the measurement. It can be observed in Figure 4. The accurate initialisation of the variance-covariance matrix P gives a possibility to make the error distribution become independent.

It is further observed that the RPE allows inclusion of the instrumental error variance. The Gaussian noise variance R_d is obtained from the instrument tolerance or sensor tolerance in measurement. Therefore, this value is equal to the instrument tolerance. A lower and higher value of R_d than instrument tolerance cause unreliable result. The overall calibration process can be observed in Table 1.

By using a continuous-discrete RPE algorithm, more data are needed in order to reduce bias and for a better accuracy of the result. Therefore, multiple cycles are recommended for the calibration process. Furthermore, the estimated parameters are used in the initialisation for the next step for detecting a toxic substance in a system.



(a) Single cycle



(b) Multiple cycles

Figure 4 State reconstructions for parameter estimation in fed-batch bioreactor.

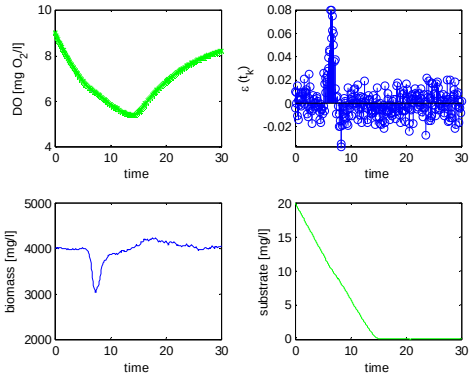
TABLE I **BIO KINETIC PARAMETER VALUES**

Parameter	Reference values	Initial values	Estimated values		
	Vanrolleghem et al., [6]		Fed-batch Single Cycle	Fed-batch Multiple Cycle	CSTR Multiple Cycle
μ_{max} **[/min]**	2.457 x10 $^{-4}$	1.000 x10 $^{-4}$	2.4720x10^{-4}	2.4576x10^{-4}	2.4826x10^{-4}
K_s **[mg/l]**	**0.456**	**0.300**	**0.4645**	**0.4456**	**0.4774**

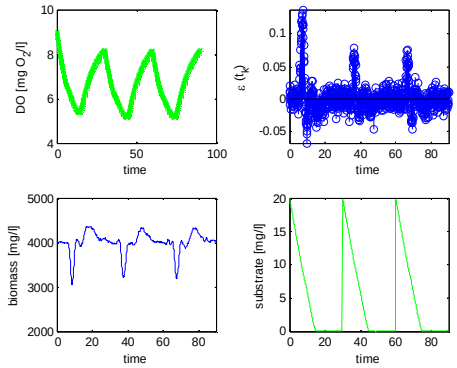
C. Toxicity Detection

The toxicity detection is conducted through keeping constant the bio kinetic parameters while estimating gain matrix elements. The state variables, gain elements, and other variables of recursive algorithm were observed. The experiment was conducted for single and multiple cycles of fed-batch bioreactor and continuous stirred tank reactor (CSTR) experiment.

By using only dissolved oxygen as the measured data, state reconstruction by a continuous-discrete RPE algorithm can obtain information about the system, including biomass concentration and substrate concentration (Figure 5).



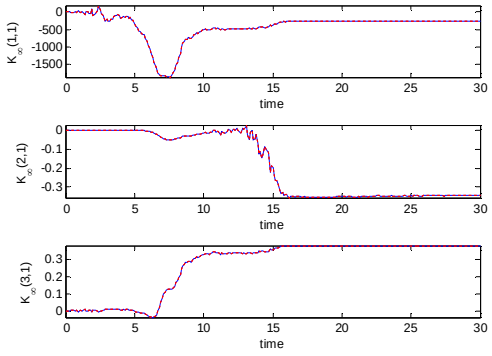
(a) Single cycle



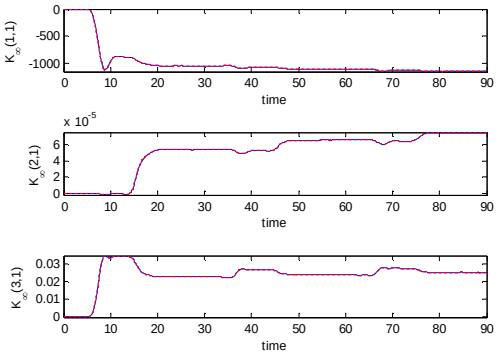
(b) Multiple cycles

Figure 5 State reconstructions for toxicity detection in fed-batch bioreactor.

The most important task is to detect the toxicity by observing the change in biomass concentration. For that purpose, it is necessary to give more priority to the gain matrix element for biomass concentration K_∞ (1,1) by making its value bigger that the other gain matrix elements so that K_∞ (1,1) is the strongest indicator for the toxicity compared to the other gain elements (Figure 6).



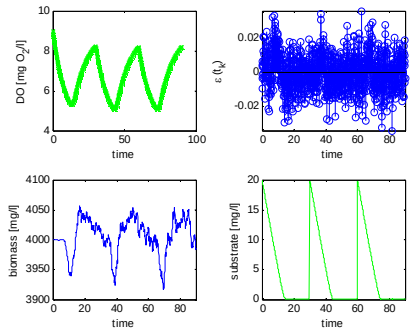
(a) Single cycle



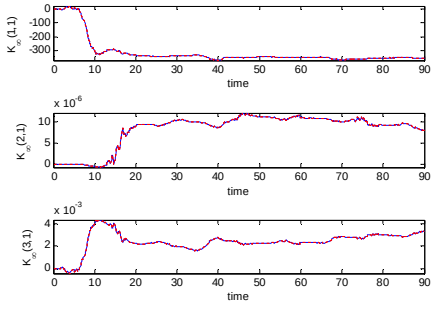
(b) Multiple cycles

Figure 6 Gain matrix elements K_∞ indicator in fed-batch bioreactor.

The robustness of the method is tested by making 2 examinations which are (i) a system which have less hazardous toxic and (ii) deal with highly noisy measurement data. From the experiments, a decrease of 2% of biomass concentration from its steady state can be well recognized by the algorithm as well as 20%, where less decrease means less hazardous toxic (Figure 7). In addition, toxicity detection can be conducted from the measurement data with 10% level of measurement noise as well as 1% level of measurement noise (Figure 8). The ability of toxic detection also depends on the determination of P value to detect the biomass concentration in a system.

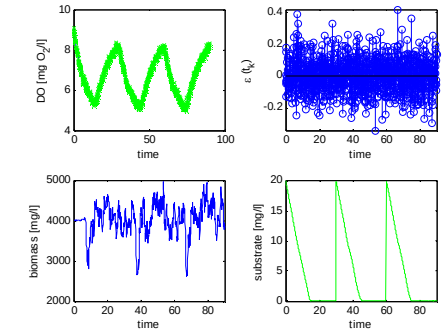


(a) State reconstruction

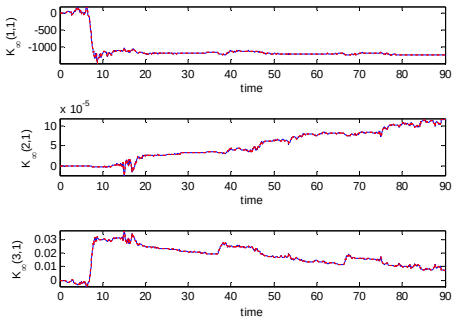


(b) Gain elements

Figure 7 Multiple cycles experiment for less hazardous toxic detection.



(a) State reconstruction

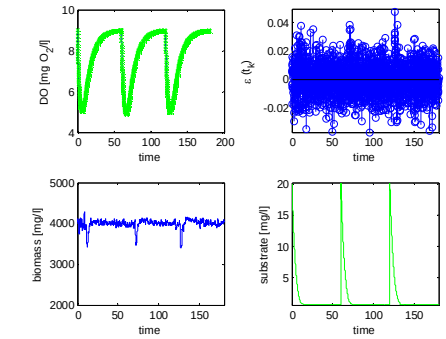


(b) Gain elements

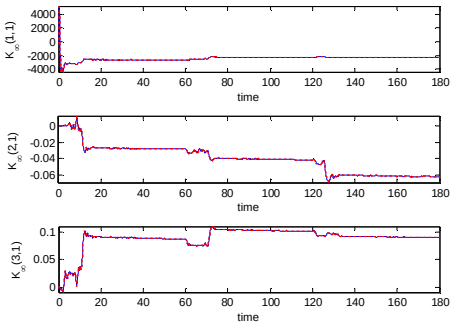
Figure 8 Multiple cycles experiment of toxic detection for more noisy measurement data.

The case in CSTR is almost similar to fed-batch bioreactor with continuous substrate inflow and outflow. The toxic substrate enters the system during the substrate feeding process. The detection is better conducted after the dissolved oxygen (DO) has reached the saturation concentration. Therefore, several fluctuations of the observed variables trajectory have to be distinguished. It is observed that every turning point in the DO data has an effect on the gain matrix for biomass concentration $K_\infty(1,1)$. However, the changing value of the gain matrix element $K_\infty(1,1)$ graph can be used to distinguish between toxicity or otherwise. A valley in gain matrix trajectory indicates the toxicity in the system.

From Figure 9, the toxic can be detected from biomass graph, gain matrix trajectory for biomass concentration $K_\infty(1,1)$, gain matrix trajectory for substrate $K_\infty(2,1)$, and gain matrix trajectory for dissolved oxygen $K_\infty(3,1)$ which are at time =10 min, time =70 min, and time =125 min.



(a) state reconstruction



(b) Gain elements

Figure 9 Multiple cycles of CSTR for 180 minutes experiment.

The indication of toxicity also can be observed through the local parametric sensitivity $W(t,\theta)$. This local parametric sensitivity determines the prediction error gradient ψ for the gain matrix for the function acceleration. The toxicity effect on the decrease of biomass concentration, influence the local parametric sensitivity to produce a strong signal in order to eliminate error between measurement and model.

CONCLUSION

The continuous-discrete RPE algorithm has successfully robust not only for the bio kinetic parameters estimation in calibration process but also for state variables reconstruction including the biomass concentration and gain elements. The multiple cycles approach produced a better estimated of bio kinetic parameters in the calibration process. By using only discrete samples of dissolved oxygen as the measured data, state reconstruction by a continuous-discrete RPE algorithm can obtain information about the system, including biomass concentration and gain elements as the toxicity indicators which make a continuous-discrete recursive prediction error is a potential method to be used for toxicity detection.

REFERENCES

[1] Banga JR, Balsa-Canto E, Moles CG, Alonso AA. Improving food processing using modern optimization methods. Trends in Food Science & Technology 2003;14:131.

[2] Stigter JD, Vries D, Keesman KJ. On adaptive optimal input design: A bioreactor case study. AIChE Journal 2006;52:3290.

[3] Peterson B, Gernaey K, Ottoy J-P, Vanrolleghem P. Application of biosensor in wastewater treatment. 22e Symposium sur les eaux usées et 11e atelier sur l'eau potable. Montréal, Canada, 1999. p.9.

[4] Stigter JD, Beck MB. On the development and application of a continuous-discrete recursive prediction error algorithm. Mathematical Biosciences 2004;191:143.

[5] Vanrolleghem PA, Daele MV, Dochain D. Practical identifiability of a biokinetic model of activated sludge respiration. Water Research 1995;29:2561.

[6] Dochain D, Vanrolleghem PA, Van Daele M. Structural identifiability of biokinetic models of activated sludge respiration. Water Research 1995;29:2571.