EXTENDED KALMAN FILTER FOR INVESTIGATION OF INHOMOGENEOUS DYNAMICS IN INDUSTRIAL BIOREACTORS

— research paper —

Silviya POPOVA*[1], Maya IGNATOVA*, Velislava LYUBENOVA*, Stefan JUNNE, Peter NEUBAUER****

Institute of System Engineering and Robotics, Bulgarian Academy of Sciences, Acad. G. Bontchev str. Bl. 2, PO. Box 79, Sofia, Bulgaria (tel: +359-2979-2052; e-mail: popova@icsr.bas.bg)

***Chair of Bioprocess Engineering, Department of Biotechnology, Technische Universität Berlin, Ackerstrasse 71-76 ACK24, D-13355 Berlin, Germany (e-mail: stefan.junne@tu-berlin.de)*

Abstract: A stepwise software sensor (SS) based on extended Kalman filter (EKF) is proposed. For monitoring scheme derivation, two methodologies concerning SS with a cascade structure and a novel formalization of kinetics are used. The Kalman's filtering approach is applied for state and parameter estimation under assumption that the dry cell weight concentration is measured on-line only. The estimation algorithm is demonstrated using experimental data of fed-batch cultivation of a non-sporulating *B. subtilis* mutant realized in two compartment reactor as scale-down process simulator. The measurements obtained by new EKF-SS could improve the process monitoring and studying process dynamics.

Keywords: Extended Kalman Filter, cascade monitoring scheme, industrial scale fed-batch processes, scale down reactor

INTRODUCTION

Monitoring and control of industrial scale bioprocesses is still a major challenge. In difference to laboratory scale cultures, large-scale bioprocesses are characterized by inhomogeneities. Such inhomogeneous processes show a high batch to batch variability compared to well mixed processes (i) due to the continuous change of the composition of the liquid phase and the

[1] Corresponding author. Mailing address: Institute of System Engineering and Robotics, Bulgarian Academy of Sciences, Acad. G. Bontchev str. Bl. 2, PO. Box 79, Sofia, Bulgaria (tel: +359-2979-2052. E-mail: popova@icsr.bas.bg)

hydrodynamic conditions (volume, viscosity, density) during a cultivation, and (ii) the strong dependence of the biological system from the local environment and its fast reaction and adaptation which is in the order or milliseconds for metabolic fluxes and in the order of seconds to minutes for adaptation at the level of gene expression. The challenge is the multitude of parameters which influence the biological system on one side and the limited number of available *on line* sensors on the other side. Therefore it is widely accepted that the only way to proper monitoring and control of a large scale bioprocess is to include feed-back control strategies, which allow guiding the process with a steady adaptation within a narrow band of low variation.
In TU-Berlin, a laboratory system for process simulation of industrial installations is developed. It consists of interrelated stirred tank bioreactor (STR) and a plug flow reactor (PFR) which tries to mimic the conditions present in industrial scale fed-batch cultivations (Figure 1).

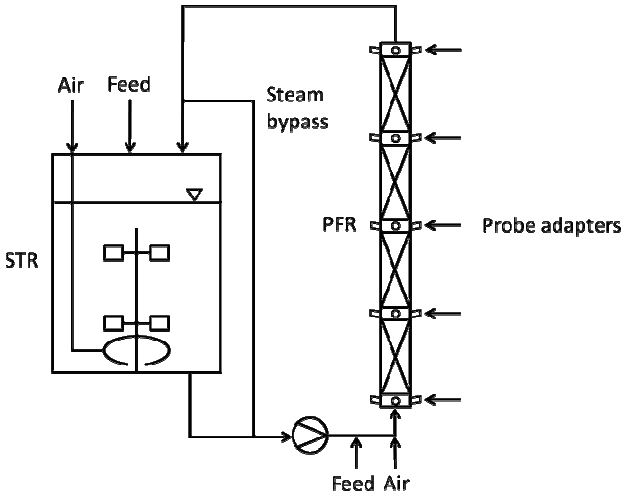

Figure 1. Scheme of laboratory simulator

The effectiveness of investigations in the laboratory simulator would be increased if some parameters, characterizing micro-organisms growth and their ability to produce different metabolites, could be observed on-line at oscillating conditions. A multi disciplinary team including biotechnologies from TU-Berlin and bioengineers from ISER-BAS was formed to solve this problem. The specialists – bioengineers are working on *on-line* monitoring of inhomogeneous dynamics in industrial bioreactors applying different model-based software sensor (SS). A model-based SS is an algorithm built from a dynamical model of a process which uses certain on line

measurements in order to estimate unmeasured variables or to estimate unknown parameters. The process model used for this purpose has to be (i) as accurate as possible to mimic the main characteristics and dynamics of the process, and (ii) simple enough for monitoring and control design. This is not a trivial task due to processes nonlinearities, instationaryties and the lack of reproducibility of experiments. The general dynamic models proposed by Bastin and Dochain in (Bastin and Dochain, 1990), are considered as classical operational models in different approaches for SS design as extended Kalman and Luenberger filters (Soons et al., 2006) (Veloso et al., 2009), the high-gain approach (Selişteanu et al., 2012) and others. Recently, a new approach for monitoring of very complex biotechnological processes has been proposed by Ignatova and Lyubenova in (Ignatova and Lyubenova, 2011). This approach is based on reduced operational models which present process kinetics with single terms including time-varying yield coefficients that comprise unmodeled dynamics.

In this paper, a stepwise SS based on extended Kalman filter for kinetic monitoring of aerobic fed-batch bioprocesses at industrial-scale bioreactors is proposed. For monitoring scheme derivation, two methodologies concerning SS with a cascade structure and a novel formalization of kinetics (Ignatova and Lyubenova, 2011) are applied. The operational model is derived considering main process variables and key process parameters only. The Kalman's filtering approach (Anderson and Moore, 1979) (Bastin and Dochain, 1990) is used for the state and parameter estimation under the assumption that the dry cell weight concentration is measured on-line only.

MONITORING SCHEME

In Figure 2, the monitoring scheme is shown. As can be seen, the scheme is realized in two steps. The first one includes Extended Kalman Filter 1 (EKF1) design with input on-line measurements of cell dry weight (C_{Xm}) only. The measurements of its specific growth rate are used as reference data to demonstrate the observation quality. EKF1 output is estimated values of cell dry weight growth rate, $\hat{R}_X$. These measurements are used as input information of EKF2. This second software sensor (SS) estimates simultaneously the yield coefficient $\hat{Y}_{XS}$, glucose concentration $\hat{C}_S$, and cell dry weight growth rate $\hat{R}_{X2}$.

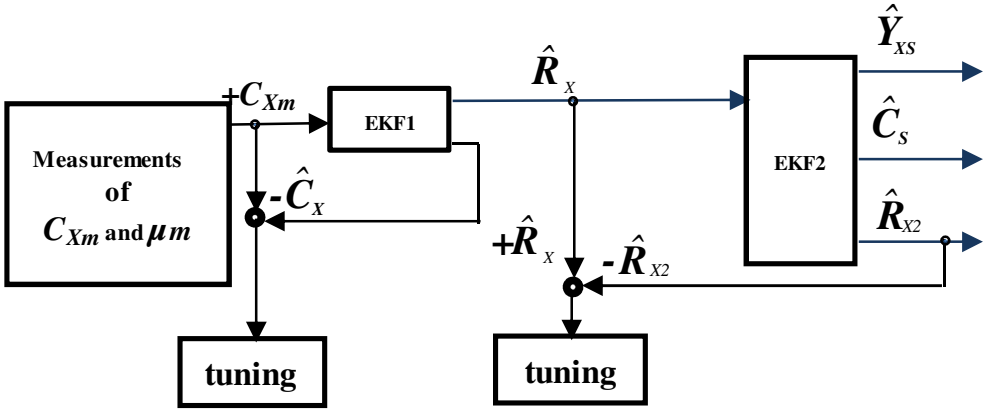

Figure 2. EKF step-wise monitoring scheme

The proposed estimation scheme is investigated by simulations in MATLAB environment on the basis of the data of one experiment of fed-batch cultivation of a non-sporulating *B. subtilis* mutant realized in two compartment reactor as scale-down process simulator at the Laboratory of Bioprocess Engineering at TU Berlin (Neubauer et al., 1995 (Neubauer and Junne, 2010) (Junne et al., 2011).

PROCESS MODEL

Usually, bioprocesses are simulated based on the biochemical network of reactions considering the corresponding stoichiometry. For process model derivation, a choice of key variables has to be made from the expert's point of view. The model including the dynamics of these variables is assumed to be appropriate for the derivation of monitoring and control algorithms. For the investigated process measurable *on line* variable is dry cell weight concentration. The glucose is not measured *on line,* but its dynamics is also included in the model as main carbon source. The following reduced model is considered:

$$\frac{dC_X}{dt} = R_X(t) - \frac{F}{V}C_X \qquad (1)$$

$$\frac{dC_S}{dt} = -R_S(t) + \frac{27.01}{V} - \frac{F}{V}C_S \qquad (2)$$

where C_X and C_S, are dry cell weight and glucose concentrations; R_X - cell dry weight growth rate, R_S –glucose consumption rate, F – substrate feed rate, V- culture volume.

According the approach proposed in (Ignatova and Lyubenova, 2011), process kinetics is presented with single term including time-varying yield coefficient Y_{XS}, that comprises unmodeled dynamics:

$$R_X(t) = R_S(t) Y_{XS}(t) \tag{3}$$

EKF STEP WISE ALGORITHM DERIVATION

The main monitoring task is to receive on-line information for glucose concentration (C_S), dry cell weight growth rate (R_X), and time-varying yield coefficient (Y_{XS}) using on-line measurements of dry cell weight concentration (C_X) only.

The two step estimation scheme (shown in Figure 2) includes as software sensors two Extended Kalman algorithms named EKF1 and EKF2. One of advantages of these algorithms is that decreasing of the estimation error is guaranteed by solving the system of Riccati equations including in the EKF algorithm.

In the first step, dry cell weight growth rate is estimated on the basis of on-line measurement of dry cell weight concentration.

The algorithm of EKF1 is described with following system:

$$\frac{d\hat{C}_X}{dt} = \hat{R}_X - \frac{F}{V}\hat{C}_X + w_{11}(C_X - \hat{C}_X) \tag{4a}$$

$$\frac{dV}{dt} = F \tag{4b}$$

$$\frac{d\hat{R}_X}{dt} = w_{31}(C_X - \hat{C}_X) \tag{4c}$$

where w_{11} and w_{31} are EKF1 tuning parameters; $\hat{C}_X$, $\hat{R}_X$ are estimates of dry cell weight concentration and its growth rate respectively.

Following the EKF theory, tuning parameters w_{11} and w_{31} are calculated applying Riccati equations (Anderson and Moore, 1979) considering the following initial values of elements of Riccati matrix $\mathcal{R}$:

$$\mathcal{R}_o = \begin{bmatrix} r_1 & r_2 & r_3 \\ r_4 & r_5 & r_6 \\ r_3 & r_6 & r_7 \end{bmatrix} = \begin{bmatrix} 0.001 & 1 & 6 \\ 6000 & 0 & 0 \\ 6 & 0 & 0.00001 \end{bmatrix} \tag{5}$$

As results, the estimates of state, $\hat{C}_X$ and specific growth rate $\hat{\mu} = \hat{R}_X / \hat{C}_X$ are obtained and shown in Figure 3a and 3b respectively. In Figure 3c, the estimates of dry cell weight growth rate $\hat{R}_X$ are presented.

On the second step, the estimation of glucose concentration (C_S), dry cell weight growth rate (R_X), and time-varying yield coefficient (Y_{XS}) is realized using on-line R_X estimates received as output of EKF1.

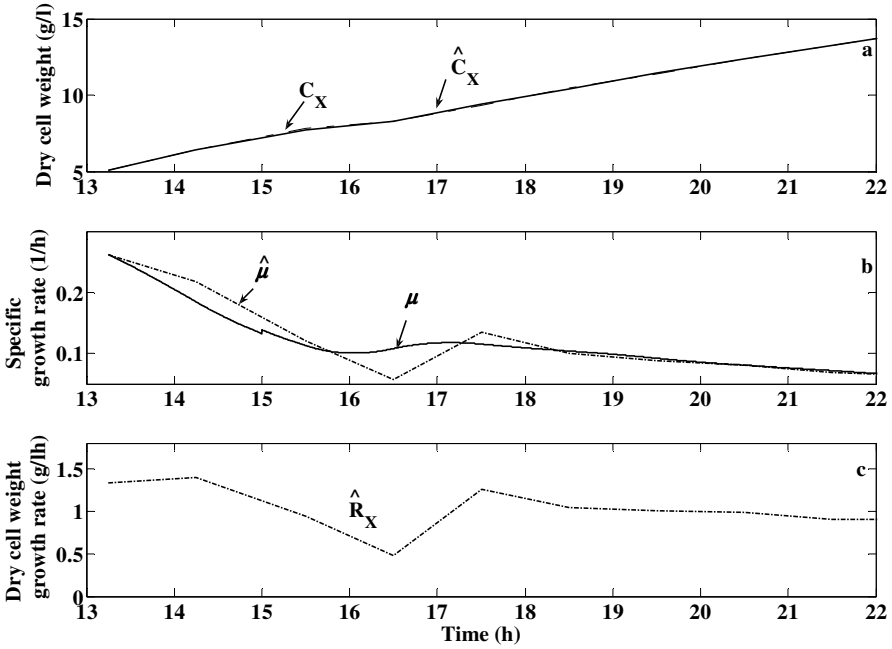

Figure 3. EKF1 results

The algorithm of EKF2 is described with following system:

$$\frac{d\hat{C}_S}{dt} = -\frac{\hat{R}_X}{\hat{Y}_{XS}} - \frac{F}{V}\hat{C}_S + w_{13}(\hat{R}_X - \hat{R}_{X2}) \quad (6a)$$

$$\frac{dV}{dt} = F \quad (6b)$$

$$\frac{d\hat{R}_{X2}}{dt} = w_{33}(\hat{R}_X - \hat{R}_{X2}) \quad (6c)$$

$$\frac{d\hat{Y}_{XS}}{dt} = w_{43}(\hat{R}_X - \hat{R}_{X2}) \quad (6d)$$

where $w_{13,}$ w_{33} and w_{43} are EKF2 tuning parameters; $\hat{C}_S$, $\hat{R}_{X2}$ and $\hat{Y}_{XS}$ are estimates of glucose concentration, cell dry weight growth rate and time-varying yield coefficient respectively.

Similar to EKF1, the tuning parameters $w_{13,}$ w_{33} and w_{43} are calculated applying Riccati equations using the following initial values of elements of Riccati matrix R:

$$\mathcal{R}_o = \begin{bmatrix} r_1 & r_2 & r_3 & r_4 \\ r_5 & r_6 & r_7 & r_8 \\ r_3 & r_7 & r_9 & r_{10} \\ r_4 & r_8 & r_{11} & r_{12} \end{bmatrix} = \begin{bmatrix} 10 & 0.1 & 3.15 & 0 \\ 1 & 0 & 0.1 & 0.1 \\ 3.15 & 0.1 & 2 & 0 \\ 0 & 0.1 & 0.05 & 10 \end{bmatrix} \quad (7)$$

The outputs of EKF2, $\hat{C}_S$, $\hat{R}_{X2}$ and $\hat{Y}_{XS}$, are shown in Figure 4a, 4b and 4c, respectively.

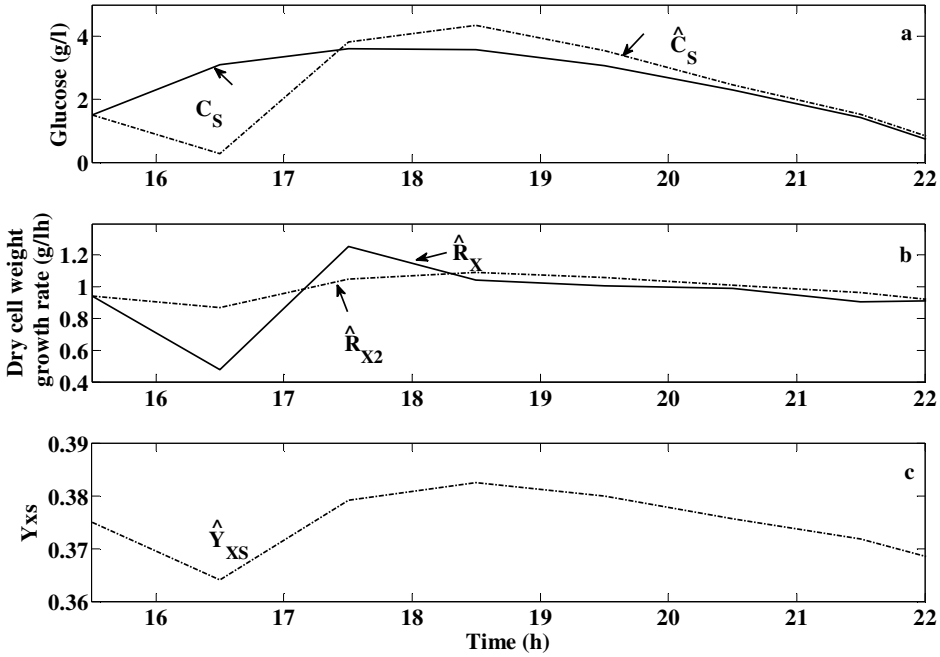

Figure 4. EKF2 results

In the following section, the results, shown in the Figures 3 and 4, will be disccused.

RESULTS AND DISCUSSION

The experiment realized in the Laboratory of Bioprocess Engineering at TU Berlin is carried out at a constant feed rate F=0.06 l/h. In this way, different substrate per cell density ratios (that is amount of substrate provided per cell) are examined. For computer investigations, a program package for system

dynamics simulations was prepared in the MATLAB environment. The results from the first EKF1 are shown in Figure 3. A coincidence between the measurements (C_X solid line) and estimates ($\hat{C}_X$ - dashed line) of dry cell weight concentration is observed in subfigure 3a. In subfigure 3b, a comparison between specific growth rate experimental values (μ - solid line) of and its estimates ($\hat{\mu}$ - dashed line) are presented. As can be seen, the estimate follows the dynamics of experimental data and after 17 hour of cultivation a coincidence between them may be remarked. The good results shown in subfigures 3a and 3b guarantee that estimates of dry cell weight rate, $\hat{R}_X$ are accurate enough to be used as input information for EKF2.

The results from the second EKF2 are shown in Figure 4. In subfigure 4b, a comparison between input dry cell weigh growth rate values ($\hat{R}_X$ - solid line) and its output estimates ($\hat{R}_{X2}$ - dashed line) are presented. The deviation at the beginning is overcome after 18 hour and output estimates converge to input ones. In subfigure 4a, off-line measurements of glucose concentration, (C_S - solid line) are compared with the estimates of the same variable ($\hat{C}_S$ - dashed line). After 17 hour, the estimate follows glucose concentration dynamics with decreasing error. On the basis of the results from subfigures 4a and 4b we have good reason to consider that the estimates $\hat{Y}_{XS}$ converge to true values after 18 hours. The observed inaccuracies can be explained with i) existence of partially observability of EKF2 (it uses one on-line measurement to estimate one variable and one parameter) ii) accumulation of estimation error coming from EKF1 and iii) sensibility of estimation error with respect to initial values of elements of Riccati matrix $\mathbf{R}$.

CONCLUSION

The derived step-wise EKF scheme is verified using the experimental data. The results from simulation investigations are promising for future investigations and could be improved by 1) using additional measurements and 2) applying optimization tuning procedure concerning initial values of elements of Riccati matrix $\mathbf{R}$.

Acknowledgments

This work was supported by Bulgarian National Fund "Scientific Researches" under contract No DTK – 02/27/17.12.2009 and by Joint projects titled: i)*"Contemporary methods for control and characterisation of EPS and PHB produced by cultivation*

of Haloferax mediterran" financed by Bilateral agreements between ISER-BAS and "Lucian Blaga" University of Sibiu – Romania and ii) "*Advanced Intelligent Control in Chemical and Bio-chemical Industries*", financed by Bilateral agreements between ISER-BAS and Petroleum-Gas University, Ploiesti, Romania.

REFERENCES

1. Anderson, B.D.O., J.L. Moore, Optimal Filtering, Prentice Hall New Jersey, 1979
2. Bastin G, Dochain D., On-line estimation and adaptive control of bioreactors. Amsterdam, Oxford, New York, Tokyo: Elsevier. 1990, 378 p.
3. Ignatova MN, Lyubenova VN., Control of biotechnological processes - new formalization of kinetics: Theoretical aspects and applications, Saarbrücken: LAP LAMBERT Academic Publishing, GmbH & Co. 2011, 120 p.
4. Junne S, Klingner A, Kabisch J, Schweder T, Neubauer P., A two-compartment bioreactor system made of commercial parts for bioprocess scale-down studies: Impact of oscillations on Bacillus subtilis fed-batch cultivations, Biotech J., 2011, 6 (8), 1009-1017
5. Neubauer P, Haggstrom L, Enfors S.O., Influence of substrate oscillations on acetate formation and growth-yield in *Escherichia coli* glucose-limited fed-batch bultivations, Biotech & Bioeng., 1995, 47, 139-146
6. Neubauer P, Junne S., Scale-down simulators for metabolic analysis of large-scale bioprocesses, Curr Opin Biotech., 2010, 21, 114-121
7. Selişteanu D, Petre E, Roman M, Şendrescu D., Estimation of kinetic rates in a baker's yeast fed-batch bioprocess by using non-linear observers, Control Theory & Applications, IET, 2012, 6 (2), 243 – 253
8. Soons, Z. J.A. Voogt, G. van Straten, A.J.B. van Boxtel, Constant specific growth rate in fed-batch cultivation of *Bordetella pertussis* using adaptive control, Journal of Biotechnology, 2006, 125 (2), 252-268
9. Veloso A, Rocha I, Ferreira EC., Monitoring of fed-batch E. coli fermentations with software sensors, Bioprocess Biosys Eng., 2009, 32 (3), 381-388