Chemical Engineering Science 177 (2018) 234–244

Contents lists available at ScienceDirect

Chemical Engineering Science

journal homepage: www.elsevier.com/locate/ces

State estimation for a penicillin fed-batch process combining particle filtering methods with online and time delayed offline measurements

Julian Kager [a], Christoph Herwig [a,b], Ines Viktoria Stelzer [c,d,*]

[a] Institute of Chemical, Environmental and Biological Engineering, TU Wien, Gumpendorfer Straße 1a, 1060 Vienna, Austria
[b] Christian Doppler Laboratory for Mechanistic and Physiological Methods for Improved Bioprocesses, TU Wien, Gumpendorfer Straße 1a, 1060 Vienna, Austria
[c] Department of Health Economics, Center for Public Health, Medical University of Vienna, Kinderspitalgasse 15, 1090 Vienna, Austria
[d] Ludwig Boltzmann Institute Applied Diagnostics, General Hospital of Vienna, Währinger Gürtel 18-20, 1090 Vienna, Austria

HIGHLIGHTS

- Algorithm for monitoring of difficult-to-measure physiological characteristics.
- Framework combines PF, parameter estimation, online and offline measurements.
- Importance of parameter estimation and recalculation step due to offline data.
- Application on real experimental data shows convincing performance.
- Provides a tool for process understanding and control.

ARTICLE INFO

Article history:
Received 10 October 2017
Accepted 30 November 2017
Available online 2 December 2017

Keywords:
Particle filter algorithm
Online measurements
Time delayed offline measurements
State estimation
Parameter estimation
Monitoring of fed-batch processes

ABSTRACT

Real time monitoring of physiological characteristics during a cultivation process is of great importance in the pharmaceutical industry. Measuring biomass, product, substrate and precursor concentrations continuously however is limited due to time-consuming laboratory analysis or expensive and hard-to-handle devices. In this work, a particle filter algorithm for estimating these difficult-to-measure process states in a *Penicillium chrysogenum* fed-batch cultivation is presented. The implemented particle filter represents a new algorithmic framework, combining several already existing methods and techniques for state estimation. It is based on nonlinear process and measurement models and takes into account both online measurements for state estimation and time delayed offline measurements, ensuring the observability of the considered system and being essential for the adaptation of dynamic model parameters. The application on real experimental data showed the convincing performance of the algorithm, estimating biomass, precursor and product concentration, as well as the specific growth rate, requiring standard measurements only. Furthermore, the positive effect of parameter estimation with respect to estimation quality was analyzed and the effect of the time delay was highlighted exemplarily. Despite of being computationally expensive due to time delayed data, the algorithm can be considered as an alternative monitoring strategy for industrial applications. Thus, it can be used further for process understanding and control.

1. Introduction

The importance of pharmaceuticals in today's health care is indisputable. In order to accelerate their development, to improve their manufacturing and to assure their quality in pharmaceutical industry the Process Analytical Technology (PAT) and Quality by Design (QbD) approach, proposed by the United States Food and Drug Administration (FDA), recommends to understand and control manufacturing processes throughout the product life cycle (USDHHS, 2004, 2009). Considering cultivation processes, this

Abbreviations: PAT, Process Analytical Technology; QbD, Quality by Design; FDA, Food and Drug Administration; pdf, probability density function; CER, carbon evolution rate; OUR, oxygen uptake rate; EKF, extended Kalman filter; PF, particle filter; PEN, penicillin V; POX, phenoxyacetate; Exp. A, experiment A; Exp. B, experiment B; RMSE, root mean square error.

* Corresponding author at: Department of Health Economics, Center for Public Health, Medical University of Vienna, Kinderspitalgasse 15, 1090 Vienna, Austria.
E-mail addresses: julian.kager@tuwien.ac.at (J. Kager), christoph.herwig@tuwien.ac.at (C. Herwig), ines.stelzer@lbiad.lbg.ac.at (I.V. Stelzer).

https://doi.org/10.1016/j.ces.2017.11.049
0009-2509/

implies a real time monitoring and control strategy for critical process parameters, product quality attributes and physiological characteristics of the underlying organism such as biomass, substrate, product or by-product concentrations (Rathore and Winkle, 2009), yields as well as product formation and growth rates (Montague et al., 1989). Thus, process deviations can be detected in time and counteracted (Glassey et al., 2011), leading to a robust production with constant performance (Mondal et al., 2014). Unfortunately, these variables are often difficult to measure in real time, resulting from the complex nature of biological systems and a lack of available measurement devices (Gonzalez et al., 2001; Farza et al., 1998). In order to circumvent this difficulty, state observers can be used. They estimate process states such as certain concentrations from partial and possibly noisy input and output measurements based on a proper process model with inexactly known initial conditions (Krener, 2009). The Bayesian approach consists in approximating the conditional probability density function (pdf) of the states recursively on condition that available measurements arrive (Simon, 2006; Ali et al., 2015). Thus, process states can be estimated in real time as online measurements arrive (Patwardhan et al., 2012).

In the recently published work by Golabgir and Herwig (2016), standard available online off-gas measurements were combined with Raman spectroscopy in order to estimate real time product and precursor concentrations in a penicillin production process. Using near-infrared spectroscopy and atline turbidity measurements Krämer and King (2016) demonstrated the online measurement of biomass and substrates in yeast cultivations. In both cases the spectroscopic measurement methods could improve estimation accuracy considerably using Bayesian state observers. Although these online measurement techniques are very promising (Vojinovic et al., 2006; Kroll et al., 2014), their use is limited due to the high cost of devices (Hulhoven et al., 2006), harsh environmental conditions and sometimes poor sensitivity and selectivity (Vojinovic et al., 2006). In addition to that, information from the complex spectra has to be extracted using multivariate data analysis (Krämer and King, 2016), which requires calibrated multivariate models relying on historical data and experts' experience (Luoma et al., 2017; Koch et al., 2014).

As it is well-known that valuable information regarding product quality and quantity can be obtained by standard offline analysis of a sample (Goffaux et al., 2009), including these offline measurements for process monitoring is common and relevant to various industrial processes (Patwardhan et al., 2012; Gopalakrishnan et al., 2011). However, using offline measurement data for real time monitoring and control is challenging due to time-consuming laboratory analysis or long sampling intervals (Vojinovic et al., 2006). In the scientific community different methodologies (Gopalakrishnan et al., 2011; Guo and Huang, 2015) and several implementations (Gudi et al., 1994; Soons et al., 2007) have emerged over the years aiming at estimating desired process states by Bayesian state observers efficiently, taking into account different measurement rates and time delayed measurements (Patwardhan et al., 2012; Oreshkin et al., 2011; Larsen et al., 1998). In most multi-rate measurement scenarios, data can be divided into measurements being available on a frequent basis, such as online measurements and on an infrequent basis, which are often time delayed, such as offline data. For efficient incorporation of various measurement rates special measurement-dependent state observers have to be designed. These algorithms estimate process states both as frequent measurements arrive being called minor update steps and as infrequent measurements are available being called major update steps (Gopalakrishnan et al., 2011; Guo and Huang, 2015). In the case of time delayed data historical adaptations need to be made by the state observer as soon as the measurements are available and retransferred into a real time context. Therefore historical information has to be stored temporarily and estimation time courses have to be recalculated, which is always associated with increased demands regarding data storage and processing time (Oreshkin et al., 2011; Larsen et al., 1998).

The estimation accuracy of any state observer strongly depends on the quality of the underlying models. In case of systems following nonlinear dynamics, to which biological systems often belong, nonlinear state observers as described for example in Krener (2009) and Simon (2006) have to be applied for state estimation. As model parameters do often change during a running process and are not static as assumed, their estimation in time can increase the quality of the observer and keep it non-degenerative. Successful implementation of algorithms, including offline measurements and online parameter estimation in the field of biotechnology are presented in Gudi et al. (1994) and Soons et al. (2007), where the biomass concentration was included as time delayed infrequent measurement. Gudi et al. (1994) reported to improve system observability by extending the frequently measured carbon evolution rate (CER) with the measured biomass concentration. Soons et al. (2007) reestimated the reactor specific volumetric oxygen transfer coefficient, which is changing over time, in order to improve the calculation of the oxygen uptake rate (OUR), which was used as frequently obtained measurement input for the state observer. In both cases the extended Kalman filter (EKF) was chosen as nonlinear Bayesian state observer, linearizing the underlying models and applying the well-known Kalman filter (Kalman, 1960). As estimation obtained by an EKF can be unsatisfactory for complex models with severe nonlinearities (Simon, 2006), a particle filter (PF) algorithm has been used in this work as state observer, as in Oreshkin et al. (2011), Golabgir and Herwig (2016) and Goffaux and Wouwer (2005). Particle filters are sequential Monte Carlo methods (Doucet et al., 2001) which approximate the pdf of estimated variables or process states by drawing a sample of N so-called particles from a predefined distribution that converges to the real pdf as N tends to infinity under certain conditions (Crisan and Doucet, 2002). The advantage of these algorithms lies in their independence of the model structure and the distributions of possibly added process and measurement noise, giving better results for nonlinear models or non-Gaussian noise (Gordon et al., 1993). Despite being computationally more expensive they have been used increasingly throughout the last years due to increasing computational capacities (Doucet et al., 2001; Gustafsson, 2010) being also relevant for industry.

The implemented PF possesses real time capability and estimates the concentration of biomass, the product penicillin V (PEN), the precursor phenoxyacetate (POX) as well as the specific growth rate in a *Penicillium chrysogenum* fed-batch process. It incorporates the standard online measurements CER and OUR, giving information about the organism's growth behavior and the time delayed offline measurements PEN, necessary for product information, and POX as it can be determined with the same HPLC method as PEN providing an additional information source. The underlying process model is based on those developed by Paul and Thomas (1996) and Paul et al. (1998), taking into account the heterogeneous structure of fungi filaments and dividing them into different compartments with growing tips and hyphal bodies where production occurs. As soon as the online measurements arrive, the algorithm estimates the desired variables according to these minor update steps. Since the observability of this system according to Hermann and Krener (1977) cannot be guaranteed using online measurements only in contrast to Golabgir and Herwig (2016), the offline measurements PEN and POX are incorporated, leading to an observable system. According to the major update steps, as soon as offline measurements are available, the PF is set back to the point in time when offline data were

measured, corrects the estimation at this point in time by using online and offline measurements and recalculates process states using historical online data until new offline data are available again. Despite of the expensive recalculation steps the additional information given by offline data improves estimation accuracy and enables parameter estimation which is performed additionally when offline data are available. More precisely, the pseudostoichiometric coefficient for the conversion of POX to PEN and the specific PEN production rate are estimated by the PF algorithm simultaneously as these two parameters have been shown to have the highest influence on POX and PEN. As parameter estimation method the straightforward approach of Kitagawa (1998) has been used. This Bayesian estimation technique extends the process model by random walk models for the parameters to be estimated instead of using maximum likelihood methods. The advantage of this method is that no time-consuming optimization problem needs to be solved. Being computationally efficient this methodology requires a significant amount of tuning since the dynamic trajectory of the parameters is not known. However, after satisfactory tuning it was shown to perform well in practice (Liu and West, 2001). An overview on state observer and PF methods for parameter estimation can be found in Tulsyan et al. (2013) and Kantas et al. (2015). Thus, the newly implemented algorithmic framework represents a state observer for the real time estimation of physiological characteristics in cultivation processes. PF methods used in Golabgir and Herwig (2016), Goffaux et al. (2009) and Oreshkin et al. (2011) are combined with online measurements, time delayed offline data considered in Gopalakrishnan et al. (2011), Guo and Huang (2015) and Gudi et al. (1994) and parameter estimation as performed in Soons et al. (2007). In order to assess the performance of the algorithm, quasi real time tests have been carried out using real experimental data generated by two fed-batch cultivation processes, showing satisfactory estimation results. Furthermore, the positive effect of parameter estimation regarding PEN and POX estimation is analyzed and the effect of time delay and recalculation is discussed. Thus, despite of being more expensive with respect to memory capacity and computing time, the presented algorithm could be a possible alternative for process monitoring in industry using standard online and offline measurements and operating without expensive, complicated measurement techniques as used in Golabgir and Herwig (2016) or Krämer and King (2016).

The paper is organized as follows. In Section 2 the experimental set-up and the measurements being taken into account are summarized. Describing the mathematical models and explaining the observer design in Section 3, results regarding the performance of the algorithm are discussed in Section 4. Finally, conclusions are presented in Section 5.

2. Material and methods

2.1. Experimental set-up

Using the spore suspension of an industrial *P. chrysogenum* strain, two cultivations, denoted by experiment A (Exp. A) and experiment B (Exp. B), were performed in a 30L and 15L stirred tank bioreactor (Infors AG, Switzerland) respectively. Batch and fed-batch processes were carried out in the same reactors. At the end of the batch processes, indicated by an increase in pH, cell broth was harvested, the reactors were filled up with sterilized, defined fed-batch media (11L and 5L) and inoculated with 160 mL/L cultivation broth from the batch processes. Both fed-batch processes were run until a biomass concentration of approximately 30 g/L was reached. Details on batch and fed-batch medium composition can be found in Posch and Herwig (2014). Dissolved oxygen was controlled above 40% by stirrer speed (350–800 rpm), while the reactor was kept at 1 bar overpressure and was aerated with 1 vvm pressurized air. The temperature was kept at 25 °C and the pH was maintained at 6.5 by addition of KOH or H_2SO_4. Glucose (500 g/L), POX (160 g/L) and ammonia (100 g/L) were supplied as feeds. In the two experiments different linear increasing glucose feeding profiles were used, leading to different cultivation conditions. Furthermore, POX was fed linearly. As glucose should be the limiting component only, ammonia was kept at nonlimiting concentrations (higher than 0.85 g/L) by increasing the feed rate manually.

2.2. Measurements

For online analytics, CO_2 and O_2 in the off-gas were quantified by a gas analyzer (Servomex, Marino Müller AG, Switzerland), using infrared and paramagnetic principle, respectively. Air inflow was controlled by a mass flow controller (Vogtlin GmbH, Switzerland). The conversion of O_2 to CO_2 was calculated by the difference between mass inflow and outflow, assuming an equilibrium between liquid and gaseous phase, giving the CER and the OUR in mol/h (Aehle et al., 2011) at nonequidistant points in time.

Offline samples were taken every 12 h and analyzed in triplicates. Analysis of PEN and POX concentration in the culture media was performed by HPLC using a ZORBAX C-18 Agilent column and 28% acetonitrile, 6 mM H_3PO_4 and 5 mM KH_2PO_4 as elution buffer. Solids were removed by centrifugation (16,000rpm for 10 min) before injection of an appropriate dilution of the supernatant. As PEN and POX eluate after 2.5 and 9.5 min respectively these offline results are available with a time delay of approximately 60 min after sampling. As reference measurement biomass was determined offline gravimetrically by separating the cells from 5 mL culture broth via centrifugation at 4800rpm for 10 min at 4 °C. The cell pellet was dried at 105 °C after a washing step with 5 mL of deionized water in weighted glass tubes and the weight of the dried pellet was determined.

Glucose and ammonia concentration were quantified from the supernatant by enzymatic photometric principle in a robotic system (Cedex Bio HT, Roche GmbH, Switzerland) with a detection limit of 0.02 g/L for glucose and 0.00048 g/L for ammonia. Gluconate was analyzed by HPLC (Agilent Technologies, USA) with a Supelco gel C-610 H ion exchange column (Sigma-Aldrich, USA) and a refractive index detector (Agilent Technologies, USA). The mobile phase was 0.1% H_3PO_4 with a constant flow rate of 0.5 mL/min, and the system was run isocratically at 4 °C.

3. Models and observer design

3.1. Derivation of process and measurement models

For the sake of completeness we sketch the derivation of the underlying models used by the state observer. The process model is mainly based on the models developed by Paul and Thomas (1996) and Paul et al. (1998) describing the heterogeneous structure of *P. chrysogenum*. According to structure and activity filamentous hyphae can be divided into actively growing parts A_0, nongrowing PEN producing parts A_1, degenerated parts and vacuoles (Megee et al., 1970). Corresponding to the experimental set-up described in Section 2, the rate of change of the volume V in the reactor is described by

$$\frac{dV}{dt} = u_{\mathrm{Glc}}(t) + u_{\mathrm{POX}}(t) + F_{\mathrm{in,ammonia}}(t) + F_{\mathrm{in,titration}}(t) - F_{\mathrm{out}}(t) = F_{\mathrm{in}}(t) - F_{\mathrm{out}}(t), \quad (1)$$

where u_{Glc}, u_{POX}, $F_{\mathrm{in,ammonia}}$ and $F_{\mathrm{in,titration}}$ are functions of time t, denoting the added glucose feed, the POX feed, the ammonia feed

and the sum of acid and base addition for pH regulation, respectively. The function F_{out} is a nonnegative function equal to zero except for points in time when a sample is taken for offline analysis. As this paper focusses on the growth and penicillin production process, biomass has been divided into A_0 and A_1, neglecting all other morphological compartments. As described in Paul and Thomas (1996) and Paul et al. (1998) it is assumed that A_0 is formed from A_1 due to branching and differentiates to A_1 depending on the glucose concentration. This can be formulated by the branching and differentiation rates

$$r_{b,0} = \frac{\mu_0 \cdot c_{A_1} \cdot c_{Glc}}{K_0 + c_{Glc}}, \quad r_{d,1} = \frac{\gamma_1 \cdot c_{A_0}}{K_1 + c_{Glc} + c_{Gln}},$$

where c_{A_0}, c_{A_1}, c_{Glc} denotes the concentration of A_0, A_1 and glucose respectively. The parameters γ_1, μ_0, K_0 and K_1 describe the maximum rate of glucose and gluconate utilisation for differentiation, the specific rate constant for branching in response to glucose, the saturation constant for branching in response to glucose and the glucose and gluconate inhibition constant for differentiation respectively. The concentration of gluconate c_{Gln} has been added to $r_{d,1}$ as it is formed extracellularly by glucose oxidase, if glucose is available in excess (Nielsen, 1997; Schmitz et al., 2013), being used as effective chelator and extracellular carbohydrate storage by fungi (Ruijter et al., 2002). Similarly to $r_{b,0}$, branch formation of A_0 by gluconate, whose utilization is inhibited by glucose (Schmitz et al., 2013) can be described by

$$r'_{b,0} = \frac{\mu'_0 \cdot c_{A_1} \cdot c_{Gln}}{(K'_0 + c_{Gln}) \cdot \left(1 + \frac{c_{Glc}}{K''}\right)},$$

where μ'_0 denotes the specific rate constant for branching in response to gluconate, K'_0 denotes the saturation constant for branching in response to gluconate and K' denotes the inhibition constant for branching on gluconate in response to glucose. Thus, for the state of growing parts A_0, we obtain

$$\frac{dc_{A_0}}{dt} = r_{b,0} + r'_{b,0} - r_{d,1} - \frac{F_{in} \cdot c_{A_0}}{V}, \tag{2}$$

where $F_{in} \cdot c_{A_0}/V$ denotes the corresponding dilution term. As the formation of A_0 causes a simultaneous extension of nongowing parts A_1 by

$$r_{e,1} = \frac{\mu_e \cdot c_{A_0} \cdot c_{Glc}}{K_e + c_{Glc}},$$

where μ_e and K_e are the corresponding specific rate and saturation constants in response to glucose, the rate of change of c_{A_1} can be described by

$$\frac{dc_{A_1}}{dt} = r_{e,1} + r'_{e,1} + r_{d,1} - r_{b,0} - r'_{b,0} - \frac{F_{in} \cdot c_{A_1}}{V}. \tag{3}$$

In Eq. (3) the effect of gluconate has been taken into account again by

$$r'_{e,1} = \frac{\mu'_e \cdot c_{A_0} \cdot c_{Gln}}{(K'_e + c_{Gln}) \cdot \left(1 + \frac{c_{Glc}}{K''}\right)},$$

where μ'_e, K'_e and K'' denotes the corresponding specific rate, saturation and inhibition constant, respectively. The change of the glucose concentration is given by

$$\frac{dc_{Glc}}{dt} = -\alpha_0 \cdot r_{b,0} - \alpha_e \cdot r_{e,1} - \alpha_p \cdot r_p - \frac{\mu_{GOD} \cdot c_{Glc}^2}{K_{GOD} + c_{Glc}} + \frac{u_{Glc} \cdot c_{f,Glc}}{V} - \frac{F_{in} \cdot c_{Glc}}{V}, \tag{4}$$

where α_0, α_e and α_p denotes the pseudostoichiometric coefficient for the conversion of glucose to A_0, to A_1 and to PEN respectively. The parameters μ_{GOD} and K_{GOD} are the specific rate constant and the saturation constant regarding the extracellular glucose oxidase enzyme (Leiter et al., 2004) and $c_{f,Glc} = 500$g/L denotes the glucose feed concentration. The rate of PEN production

$$r_p = \frac{\mu_p \cdot c_{A_1} \cdot (c_{Glc} + c_{Gln})}{K_p + (c_{Glc} + c_{Gln}) \cdot \left(1 + \frac{c_{Glc} + c_{Gln}}{K_I}\right)} \cdot \frac{c_{POX}}{(K_{POX} + c_{POX})}$$

is described similarly to Paul and Thomas (1996) and Paul et al. (1998), taking into account gluconate as additional C-source and POX as precursor, whose conversion to PEN is described by Monod kinetics. The specific PEN production rate is denoted by μ_p, depending on c_{A_1} as production occurs in the nongrowing parts. K_p and K_I denotes the saturation and inhibition constant for PEN production in response to substrate respectively, K_{POX} denotes the saturation constant of the conversion of POX to PEN and c_{POX} the POX concentration. For gluconate, being added as additional state and produced by glucose oxidase, we obtain

$$\frac{dc_{Gln}}{dt} = -\alpha_{Gln} \cdot (r'_{b,0} + r'_{e,1}) + \frac{\mu_{GOD} \cdot c_{Glc}^2}{K_{GOD} + c_{Glc}} - \frac{F_{in} \cdot c_{Gln}}{V}, \tag{5}$$

where the parameter α_{Gln} denotes the pseudostoichiometric coefficient for the conversion of gluconate to A_0 and A_1. The rate of change of PEN concentration c_{PEN} can be described by

$$\frac{dc_{PEN}}{dt} = r_p - \mu_h \cdot c_{PEN} - \frac{F_{in} \cdot c_{PEN}}{V}, \tag{6}$$

assuming that PEN is produced according to r_p and hydrolysed with the first order PEN hydrolysis constant μ_h. We refer to Revilla et al. (1984) and Menezes et al. (1994) for further details. Instead of lactose considered in Paul et al. (1998) POX has been added to the system as second substrate by the equation

$$\frac{dc_{POX}}{dt} = -r_p \cdot \alpha_{POX/PEN} + \frac{u_{POX} \cdot c_{f,POX}}{V} - \frac{F_{in} \cdot c_{POX}}{V}, \tag{7}$$

where $\alpha_{POX/PEN}$ denotes the pseudostoichiometric coefficient for the conversion of POX to PEN and $c_{f,POX} = 160$g/L denotes the POX feed concentration.

Altogether, the process model consists of 7 states and 24 parameters. The parameters and their values, taken from literature, are listed in Table 1. The values of K', K'' and K_{POX} have been determined experimentally using batch experiments. In order to identify the most sensitive parameters, local sensitivity and identifiability analysis according to Brun et al. (2002) has been carried out. It shows that the parameters μ_p and $\alpha_{POX/PEN}$ are the most important parameters to be estimated with respect to their effect on the states of interest PEN and POX. As described in Section 1 these parameters can be estimated simultaneously by the PF, according to the method in Kitagawa (1998) by including them as the additional states

$$\frac{d\alpha_{POX/PEN}}{dt} = 0, \tag{8}$$

$$\frac{d\mu_p}{dt} = 0. \tag{9}$$

These states are assumed to be static, which means that their rate of change is assumed to be zero.

As measurement model for the online measurements CER and OUR, the equations

$$y_{CER} = \left(\frac{dc_{Glc}}{dt} + \frac{dc_{Gln}}{dt} + \frac{dc_{POX}}{dt} - \frac{dc_{A_0}}{dt} - \frac{dc_{A_1}}{dt} - \frac{dc_{PEN}}{dt}\right) \cdot V \tag{10}$$

and

$$y_{OUR} = \left(-\frac{dc_{Glc}}{dt} - \frac{dc_{Gln}}{dt} - \frac{dc_{POX}}{dt} + \frac{dc_{A_0}}{dt} + \frac{dc_{A_1}}{dt} + \frac{dc_{PEN}}{dt}\right) \cdot V \tag{11}$$

Table 1
List of model parameters and their values.

Parameter	Value	Unit	Description	Source
α_0	1.850	g Glc $g^{-1}A_0^{-1}$	Pseudostoichiometric coefficient for the conversion of glucose to A_0	Paul et al. (1998)
α_e	1.830	g Glc $g^{-1}A_1^{-1}$	Pseudostoichiometric coefficient for the conversion of glucose to A_1	Paul et al. (1998)
α_{Gln}	2.100	g Gln g^{-1} $(A_0+A_1)^{-1}$	Pseudostoichiometric coefficient for the conversion of gluconate to A_0 and A_1	Nielsen (1997)
α_p	1.000	g Glc g^{-1} PEN^{-1}	Pseudostoichiometric coefficient for the conversion of glucose to PEN	Paul and Thomas (1996) and Menezes et al. (1994)
$\alpha_{POX/PEN}$	–	g POX g^{-1} PEN^{-1}	Pseudostoichiometric coefficient for the conversion of POX to PEN	estimated
γ_1	0.082	g (Glc + Gln) L^{-1} h^{-1}	Maximum rate of glucose and gluconate utilisation for differentiation from A_0 to A_1	Paul and Thomas (1996) and Paul et al. (1998)
μ_0	0.016	h^{-1}	Specific rate constant for branching in response to glucose	Paul and Thomas (1996) and Paul et al. (1998)
μ_0'	0.016	h^{-1}	Specific rate constant for branching in response to gluconate	Paul and Thomas (1996) and Paul et al. (1998)
μ_e	0.409	h^{-1}	Specific rate constant for extension in response to glucose	Paul et al. (1998)
μ_e'	0.409	h^{-1}	Specific rate constant for extension in response to gluconate	Paul et al. (1998)
μ_{GOD}	0.500	h^{-1}	Specific rate constant for glucose oxidation	Nielsen (1997) and Schmitz et al. (2013)
μ_h	0.002	h^{-1}	First order PEN hydrolysis constant	Paul and Thomas (1996) and Paul et al. (1998)
μ_p	–	h^{-1}	Specific rate constant for PEN production	Estimated
K'	0.100	g Glc L^{-1}	Inhibition constant for branching on gluconate in response to glucose	–
K''	0.100	g Glc L^{-1}	Inhibition constant for extension on gluconate in response to glucose	–
K_0	0.040	g Glc L^{-1}	Saturation constant for branching in response to glucose	Paul and Thomas (1996) and Paul et al. (1998)
K_0'	0.100	g Gln L^{-1}	Saturation constant for branching in response to gluconate	Nielsen (1997) and Schmitz et al. (2013)
K_1	0.091	g (Glc + Gln) L^{-1}	Inhibition constant for differentiation in response to glucose and gluconate	Paul et al. (1998)
K_e	0.079	g Glc L^{-1}	Saturation constant for extension in response to glucose	Paul et al. (1998)
K_e'	0.079	g Gln L^{-1}	Saturation constant for extension in response to gluconate	Paul et al. (1998)
K_{GOD}	0.010	g Glc L^{-1}	Saturation constant for glucose oxidation	Nielsen (1997) and Schmitz et al. (2013)
K_I	0.261	g (Glc + Gln) L^{-1}	Saturation constant for PEN production in response to glucose and gluconate	Paul et al. (1998)
K_p	0.041	g (Glc + Gln) L^{-1}	Inhibition constant for PEN production in response to glucose and gluconate	Paul et al. (1998)
K_{POX}	0.300	g POX L^{-1}	Saturation constant of the conversion of POX to PEN	–

are considered, using a carbon balance for CER and an electron balance for OUR. According to local observability for nonlinear models basing upon Lie derivatives (Hermann and Krener, 1977), system (1)–(7) and (10) and (11) is not fully observable. Thus, the process states

$$x = (V, c_{A_0}, c_{A_1}, c_{Glc}, c_{Gln}, c_{PEN}, c_{POX})^T,$$

where the superscript T denotes the transpose of a vector, given by the process model (1)–(7), cannot be estimated appropriately by the measurements CER and OUR. Extending the measurement model (10) and (11) by the offline measurements for PEN and POX given by

$$y_{PEN} = c_{PEN} \tag{12}$$

and

$$y_{POX} = c_{POX}, \tag{13}$$

however, observability can be guaranteed which has also been observed in Golabgir et al. (2015), where the original model of Paul et al. (1998) has been analyzed.

3.2. Observer design

For real time application of the state observer an OPC server has been developed that is based on the Python based Django framework (Django Software Foundation). The server guarantees the bidirectional communication between devices, sensors and algorithms implemented in Matlab R2015b (Mathworks, USA). All needed raw signals are loaded to the server every 30 s and are transferred to Matlab, where the calculations are performed iteratively. Incoming data are pretreated in Matlab using outlier detection and Savitzky-Golay smoothing algorithms. CER and OUR measurements are calculated as described in Section 2.2, feed rates are obtained by the first-order Savitzky-Golay derivative basing on the corresponding weights. Furthermore, signals arriving at different points in time are interpolated according to the largest time interval. These pretreated data are used as input by the PF for state estimation. The PF version implemented in this work is the so-called SIR or bootstrap filter according to Gordon et al. (1993). A detailed derivation of this algorithm can be found in Simon (2006) or Doucet et al. (2001), where different resampling techniques are compared theoretically and empirically in Hol et al. (2006). Alternative particle filter methods are described for example in Doucet et al. (2001), Arulampalam et al. (2002), Pitt and Shephard (1999) and Johansen and Doucet (2008). Roughly speaking, for the states x described by the process model, such as (1)–(7), the SIR algorithm draws a sample of N particles,

$$x_0^i, \quad i = 1, \ldots, N$$

by the known pdf $p(x_0)$ of the initial state x_0 as initialisation step, where N is a fixed predefined natural number. Iteratively, for time steps $k \geqslant 1$, N a priori particles $\tilde{x}_k^i$, $i = 1, \ldots, N$, are drawn according to the conditional pdf of x_k given x_{k-1}^i, $p(x_k|x_{k-1}^i)$, using the process model and possibly added process noise. Basing upon these particles, importance weights

$$\tilde{w}_k^i = p(y_k|\tilde{x}_k^i), \quad i = 1, \ldots, N,$$

defined as the conditional pdf of the measurement y at time step k given $\tilde{x}_k^i$, are calculated using the measurement model, possibly added measurement noise and real measurements and normalized by

$$w_k^i = \frac{\tilde{w}_k^i}{\sum_{j=1}^N \tilde{w}_k^j}, \quad i = 1, \ldots, N.$$

Applying multinomial resampling N a posteriori particles x_k^i, $i = 1, \ldots, N$, are resampled from $\tilde{x}_k^i$ by the normalized

importance weights. The distribution of these particles is an approximation of the desired probability density function, i.e.

$$x_k^i \sim p(x_k|y_k), \quad i = 1, \ldots, N.$$

Thus, the bootstrap filter estimates the states V, c_{A_0}, c_{A_1}, c_{Glc}, c_{Gln}, c_{PEN} and c_{POX} iteratively according to the minor update steps by the process model (1)–(7) and the measurement model (10)–(11), including added process and measurement noise, as soon as CER and OUR measurements arrive. In order to reduce the computational effort this updating step is performed every 30 min being sufficient regarding the slower dynamics of the considered process. When offline data for PEN and POX are available, the particle filter algorithm is reset, according to the major update step, taking into account the time delay of the signals and estimates all states and the parameters $\alpha_{POX/PEN}$ and μ_p by the extended process model (1)–(9) and the extended measurement model (10)–(13), including added process and measurement noise. The model (1)–(7) is updated by these estimated parameters and the particle population for the states obtained by the extended models is taken as starting point for reestimating the states based on the reduced models and historical online signals until current CER and OUR measurements can be incorporated. The state estimation using the current online measurements guarantees the real time application of the state observer. As soon as offline measurements are available this procedure is repeated. The principle of the calculation and recalculation step of the implemented algorithm taking into account online and offline measurements is illustrated in Fig. 1.

4. Results and discussion

In order to test the implemented state observer in quasi real time, the experimental data generated by the two *P. chrysogenum* fed-batch experiments A and B have been used and simulated as to arrive in real time iteratively. The initial state

$$x_0 = (V(0), c_{A_0}(0), c_{A_1}(0), c_{Glc}(0), c_{Gln}(0), c_{PEN}(0), c_{POX}(0))^T$$

of the process model (1)–(7) has been assumed to be normally distributed with known mean

$$\mu_{x0,A} = (13, 0.5, 4.5, 0, 0.6, 0, 5.8)^T \quad \text{and}$$
$$\mu_{x0,B} = (6.4, 0.5, 4.5, 0, 0.1, 0, 6.3)^T$$

for Exp. A and Exp. B and covariance matrix

$$\Sigma_{x0} = \mathrm{diag}\left(10^{-8}, 4\cdot10^{-4}, 4\cdot10^{-4}, 1\cdot10^{-4}, 1\cdot10^{-2}, 1\cdot10^{-2}, 2.5\cdot10^{-2}\right).$$

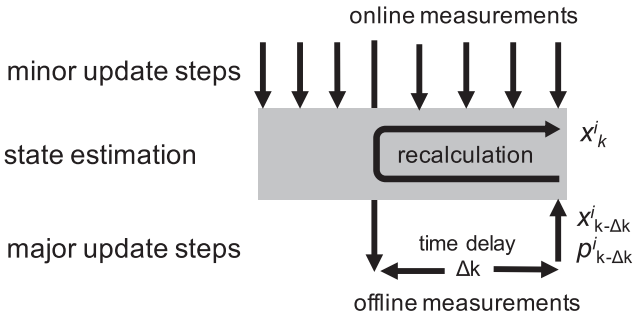

Fig. 1. Principle of PF algorithm taking into account online and time delayed offline measurements. According to minor update steps of online measurements the state population $x_k^i, i = 1, \ldots, N$ is calculated by the algorithm. As soon as offline measurements, according to major update steps with a time delay of Δk, are available, the state and parameter population $x_{k-\Delta k}^i$ and $p_{k-\Delta k}^i, i = 1, \ldots, N$ is estimated and a recalculation is performed.

The values of the vectors $\mu_{x0,A}$ and $\mu_{x0,B}$ have been chosen according to experience. As they describe the start conditions in a sufficient way the diagonal entries of Σ_{x0} have been assumed to be small. Furthermore, no correlation between the states has been assumed. Similarly, for the extended model (1)–(9) a multivariate normal distribution has been assumed for the initial parameter states

$$\begin{pmatrix} \alpha_{POX/PEN}(0) \\ \mu_p(0) \end{pmatrix} \sim \mathcal{N}\left(\begin{pmatrix} 0.431 \\ 0.023 \end{pmatrix}, \begin{pmatrix} 1\cdot10^{-8} & 0 \\ 0 & 1\cdot10^{-8} \end{pmatrix}\right),$$

where the mean of the initial parameter value $\alpha_{POX/PEN}(0)$ has been chosen as the theoretical yield of the corresponding reaction and the mean of $\mu_p(0)$ has been set to the parameter value proposed in Paul and Thomas (1996). For time $k \geqslant 0$, additive independent normally distributed white process noise $w_{k,minor}$ and $w_{k,major}$ has been added to the process model (1)–(7) and to the extended model (1)–(9) with zero mean and covariance matrices

$$\Sigma_{wk,minor} = \mathrm{diag}(1\cdot10^{-8}, 4\cdot10^{-4}, 4\cdot10^{-4}, 10^{-4}, 10^{-2}, 10^{-2}, 2.5\cdot10^{-2})$$

and

$$\Sigma_{wk,major} = \mathrm{diag}(1\cdot10^{-8}, 9\cdot10^{-2}, 0.36, 10^{-4}, 10^{-2}, 0.1, 4\cdot10^{-2}, 8.6\cdot10^{-2}, 1.3\cdot10^{-2})$$

respectively. According to the two different update steps two different covariance matrices have been considered for the process noise, where the process noise regarding the major update step has been assumed to be greater than or equal to those for the minor update step due to the longer sampling intervals. For $k > 0$ additive independent white noise $v_{k,minor}$ and $v_{k,major}$ has been added to the measurement model (10) and (11) and (10)–(13) distributed by

$$v_{k,minor} \sim \mathcal{N}\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 4\cdot10^{-2} & 0 \\ 0 & 3\cdot10^{-2} \end{pmatrix}\right)$$

and

$$v_{k,major} \sim \mathcal{N}\left(\begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 4\cdot10^{-2} & 0 & 0 & 0 \\ 0 & 3\cdot10^{-2} & 0 & 0 \\ 0 & 0 & 1\cdot10^{-1} & 0 \\ 0 & 0 & 0 & 2\cdot10^{-2} \end{pmatrix}\right)$$

respectively according to the real measurement errors. Thus, both the process and the measurement noise has been assumed to be zero on average with small deviations from that. Correlation between the noise on the various states has not been taken into account. Judging from the algorithm's performance, these assumptions seem to be sufficient. For the simulations the particle size N has been chosen to be 2 500 as it has shown to give the best accuracy with reasonable calculation time. State estimates have been represented by the median of the corresponding particle population.

4.1. Estimation results

In Fig. 2 estimation results obtained by the PF algorithm, using data given by Exp. A, are summarized. Giving important information during a cultivation process, the estimated concentration of biomass, PEN and POX versus time is presented, where the beginning of the fed-batch is denoted by time zero. The total biomass concentration is obtained by summing up growing and nongrowing parts according to $c_X = c_{A_0} + c_{A_1}$. Glucose and gluconate concentrations are not shown here as they are at very low ranges and below detection limits during the C-source limited fed-batch

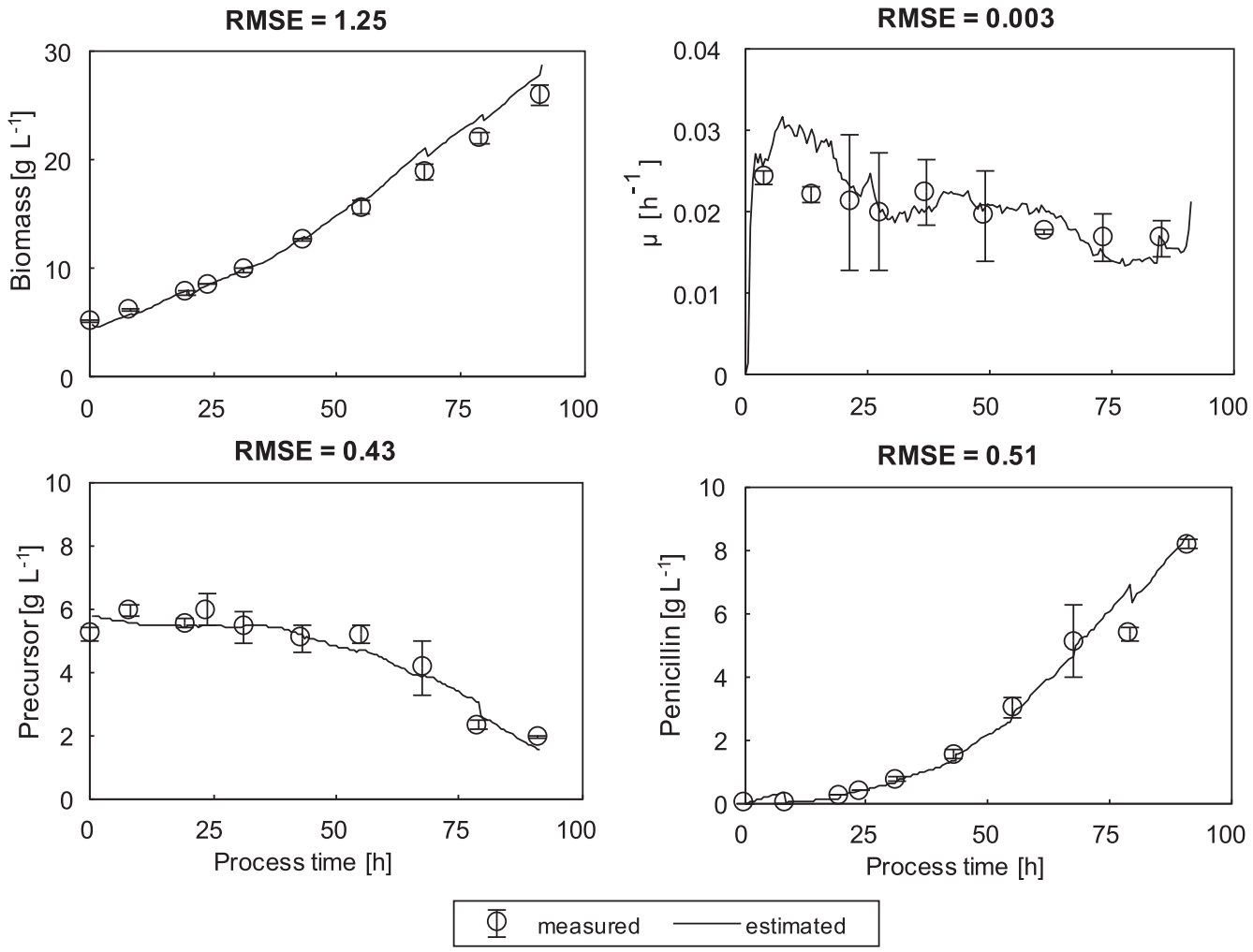

Fig. 2. Estimation of biomass, PEN and POX concentration and specific growth rate versus time for Exp. A.

processes. Furthermore, the specific growth rate, denoted by μ, being calculated as the relative rate of change of total biomass, is considered as additional model output. The estimated variables have been compared to the corresponding offline measurements, where the mean and the standard deviation of the measured triplicates have been used. The reference value for the specific growth rate has been calculated from offline biomass measurements. Root mean square errors (RMSEs) between the estimated values and the mean of the measured values have been calculated for both cultivations as performance indicators. The RMSEs for Exp. A can be seen in Fig. 2 showing a good performance of the estimation algorithm. Although it is well-known that there are more robust particle filter versions (Li et al., 2014), the behavior of the implemented SIR filter towards outliers can be seen in Fig. 2 after 79 h for PEN concentration. In this case the estimation is only slightly affected by the low PEN measurement, indicating that the algorithm is able to handle single outliers. For Exp. B similar results have been obtained with an RMSE of 2.07 g/L for biomass concentration, 0.22 g/L for PEN concentration, 0.40 g/L for POX concentration, and 0.007 1/h for the specific growth rate μ. Thus, for Exp. B the estimation accuracy for product and precursor concentration is even better than for Exp. A, where the estimation of the biomass and the resulting specific growth rate shows higher deviations from the reference measurements, which could result from higher dynamic growth behavior. Considering the average RMSEs of both experiments for PEN and POX concentration, the average deviation between estimation and measurement with 0.37 g/L for PEN and 0.42 g/L for POX, is not considerably higher than the standard deviations of the corresponding measurements (0.27 g/L for PEN and 0.33 g/L for POX). Comparing the average RMSE of biomass with 1.66 g/L to the standard deviation with 0.50 g/L, a higher difference can be observed. However, both values are in a similar range. Thus, the real time estimation of PEN, POX and biomass concentration seems to be reliable. Very promising results have been obtained further for the estimation of the not directly measurable specific growth rate μ although giving least satisfactory estimation results regarding all four states. Due to error propagation the calculated reference measurement of this rate shows an average standard deviation for Exp. A and Exp. B of approximately 17% which is rather high. The estimated specific growth rate with an average RMSE of 15%, thus represents a considerable alternative for the monitoring of this rate and could be used for feedback control and control tasks beyond that. Altogether, the estimation results obtained by the state observer are satisfactory for both experiments.

4.2. Parameter estimation

As parameter estimation has been included in the state estimation process as soon as offline data are available, its effect on estimation quality is analyzed in the following. In Table 2 the evolution of the estimated parameters $\alpha_{POX/PEN}$ and μ_p over time

Table 2
Evolution of parameters over time estimated by the state observer and Exp. A.

Process time [h]	0	8	19	24	31	43	55	68	79	91
$\alpha_{POX/PEN}$ [g POX g^{-1} PEN^{-1}]	0.434	0.348	0.328	0.334	0.350	0.384	0.398	0.417	0.488	0.494
μ_p [h^{-1}]	0.023	0.006	0.010	0.007	0.101	0.014	0.018	0.020	0.017	0.014

is given for Exp. A and illustrated in Fig. 3 (bottom left and bottom right plot) for Exp. B, where their values have been obtained by calculating the median of the particle population obtained by the state observer. Furthermore, the confidence interval with a confidence level of 70% has been calculated from this population in Fig. 3. The means of the initial parameter states have been used as static parameter values.

It can be seen that the pseudostoichiometric coefficient for the conversion of POX to PEN $\alpha_{POX/PEN}$ has been only adapted slightly by the state observer, indicating a less dynamic behavior of the parameter. In contrast, the specific PEN production rate shows a strong change in time and reaches different levels during the process. After dropping sharply at the beginning of the process, the productivity remains stable on a lower level as initially assumed, before it decreases again at the end of the process. Thus, setting this parameter constant seems to be an oversimplification. The behavior indicates that the description of product formation is complex and may be influenced by several factors, not only glucose and POX concentration as suggested by the model. This might be of interest and could be investigated further in order to increase process understanding.

The influence of the parameters for state estimation is reflected further in Fig. 3 for the most affected states PEN (top right) and POX (top left) for Exp. B. The black and the gray line represent state estimation results with and without the estimation of both parameters respectively. At points in time when offline measurements are available both estimation results coincide due to the full information obtained. However, neglecting parameter estimation, the

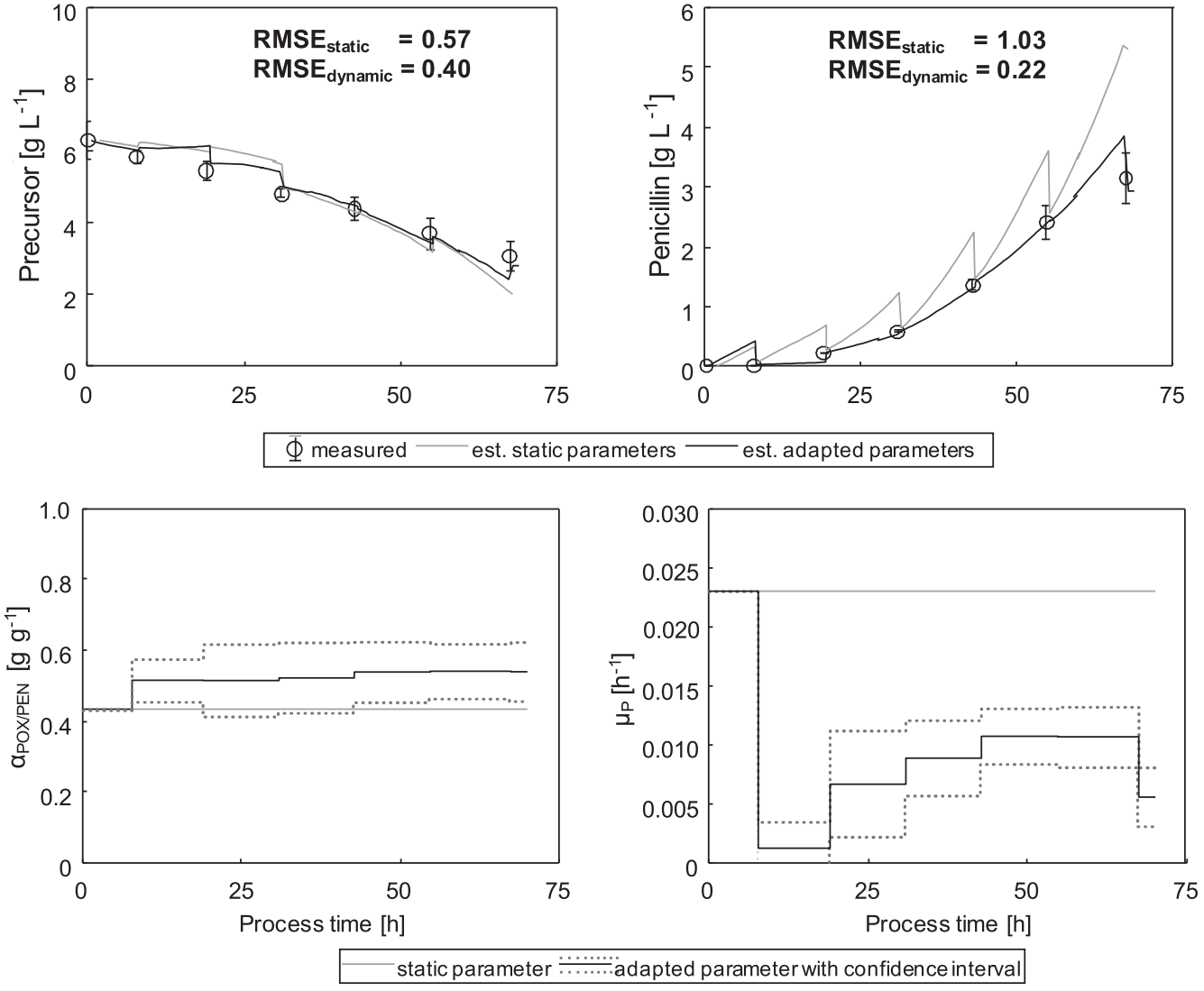

Fig. 3. Comparison of POX (top left) and PEN (top right) estimation with and without parameter adaptation and evolution of estimated state parameters $\alpha_{POX/PEN}$ (bottom left) and μ_p (bottom right) versus static parameters over time for Exp. B.

state observer shows higher deviations from measured concentrations for points in time when only online measurements are available. These deviations propagate with time until new offline data can be incorporated resulting from the model not being updated with respect to the parameters. Since the pseudostoichiometric coefficient for the conversion of POX to PEN is only adapted slightly during the estimation process, differences between POX concentration determined with and without parameter estimation are present but not so pronounced which can be seen in the corresponding RMSEs in Fig. 3. Considering PEN estimation, however, parameter adaptation strongly influences estimation accuracy, leading to a reduction of the RMSE from approximately 1.03 g/L to 0.22 g/L which is almost 79% lower. This highlights the importance of updating a model with respect to dynamic parameters.

4.3. Time delay

Apart from parameter estimation, the incorporation of time delayed data and the subsequent recalculation step, as shown in Fig. 1, is a key element of the state observer. In Fig. 4 the effect of time delay and recalculation for state estimation of Exp. B is highlighted by simulating PEN and POX measurements to arrive with a time delay of 6 h, although being available after 60 min. The black line corresponds to state estimation as it appears in real time taking into account measurements that arrive with a delay of 6 h. The dotted line shows the estimation with the recalculation step of the filter being performed between the points of time of sampling and the points of time when the offline measurement is available.

Obviously, the estimation of the considered states biomass, PEN and POX is improved by the recalculation step, although this is not reflected in a lower RMSE value for biomass as well as for the specific growth rate. Considering PEN, the concentration is strongly overestimated at the beginning of the process without the recalculation step due to the high initial specific PEN production rate. This results in an RMSE increase from 0.22 g/L to 0.33 g/L of 50%. For POX this increase is even higher with 60%. Nevertheless, considering the entire process, both estimations coincide as soon as offline data are available as the amount of information is all the same. Thus, the incorporation of time delayed offline data to the observer is relevant for improving estimation quality even though no recalculation step can be performed. However, it has to be taken into account that in that case real time estimation gets worse with an increased time delay. This can be significant in industrial applications. Thus, despite of being time-consuming and expensive with respect to data storage of historical information, which can be

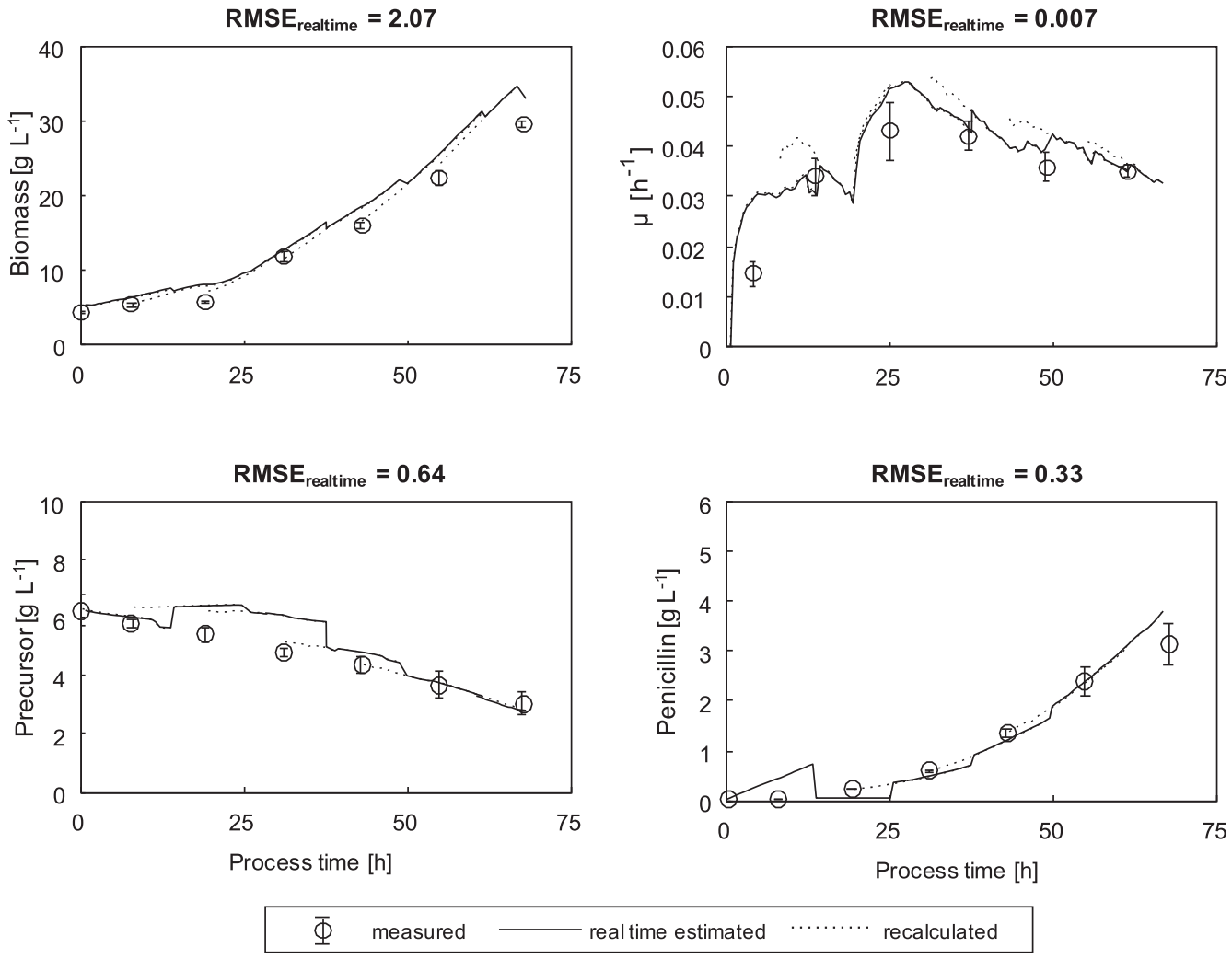

Fig. 4. Estimation of biomass, PEN and POX concentration and specific rate versus time for Exp. B with and without recalculation step using 6 h time delayed offline measurements.

challenging in practice, the recalculation step is important in order to improve overall and future estimations.

5. Conclusions

In this work a state observer for the real time monitoring of biomass, product and precursor concentration in a *P. chrysogenum* cultivation process has been developed. Combining several recently developed and used methods such as particle filtering, parameter estimation techniques and the effective incorporation of different information sources, namely both online and time delayed offline measurements, the implemented algorithm has shown satisfactory estimation results using two fed-batch experiments. Furthermore, the convincing estimation of not directly measurable variables of interest such as the specific growth rate highlights the strength of the algorithm additionally. Parameter estimation, especially improving estimation quality as has been demonstrated, could be used for elucidating unknown physiological mechanisms. An advantage of the presented approach is that it depends on available standard online and offline measurements only, and does not require expensive or hard-to-handle online devices, while providing comparable results. The unavoidable recalculation step that has to be performed due to time delayed data is expensive with respect to data storage and calculating time, but has a positive effect on estimation quality especially for greater time delays. Thus, the developed framework could be considered as alternative monitoring strategy throughout the product life cycle in industrial application, leading to additional process understanding and transparency. Furthermore, the information gained by the algorithm could be used for feedback and process control.

Acknowledgments

This work was supported by the Austrian research funding association (FFG) under the scope of the COMET program within the project "Industrial Methods for Process Analytical Chemistry – From Measurement Technologies to Information Systems (imPACts)" (contract# 843546) and the Christian Doppler Forschungsgesellschaft (Grant No. 171). Furthermore, the authors would like to thank Sandoz GmbH for its support, Daniela Ehgartner for providing experimental data and Fabian Pollesböck from Exputec GmbH for assisting with the computational real time environment.

References

Aehle, M., Kuprijanov, A., Schaepe, S., Simutis, R., Lübbert, A., 2011. Simplified off-gas analyses in animal cell cultures for process monitoring and control purposes. Biotechnol. Lett. 33 (11), 2103–2110.

Ali, J.M., Hoang, N.H., Hussain, M.A., Dochain, D., 2015. Review and classification of recent observers applied in chemical process systems. Comput. Chem. Eng. 76, 27–41.

Arulampalam, M.S., Maskell, S., Gordon, N., Clapp, T., 2002. A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking. IEEE Trans. Signal Process. 50 (2), 174–188.

Brun, R., Kühni, M., Siegrist, H., Gujer, W., Reichert, P., 2002. Practical identifiability of ASM2d parameters systematic selection and tuning of parameter subsets. Water Res. 36 (16), 4113–4127.

Crisan, D., Doucet, A., 2002. A survey of convergence results on particle filtering methods for practitioners. IEEE Trans. Signal Process. 50 (3), 736–746.

Doucet, A., De Freitas, N., Gordon, N., 2001. An introduction to sequential Monte Carlo methods. In: Sequential Monte Carlo methods in practice. Springer, New York, pp. 3–14.

Farza, M., Busawon, K., Hammouri, H., 1998. Simple nonlinear observers for on-line estimation of kinetic rates in bioreactors. Automatica 34 (3), 301–318.

Glassey, J., Gernaey, K.V., Clemens, C., Schulz, T.W., Oliveira, R., Striedner, G., Mandenius, C.F., 2011. Process analytical technology (PAT) for biopharmaceuticals. Biotechnol. J. 6 (4), 369–377.

Goffaux, G., Wouwer, A.V., 2005. Bioprocess state estimation: some classical and less classical approaches. In: Control and Observer Design for Nonlinear Finite and Infinite Dimensional Systems. Springer, Berlin, Heidelberg, pp. 111–128.

Goffaux, G., Wouwer, A.V., Bernard, O., 2009. Improving continuous-discrete interval observers with application to microalgae-based bioprocesses. J. Process Control 19 (7), 1182–1190.

Golabgir, A., Herwig, C., 2016. Combining mechanistic modeling and raman spectroscopy for real-time monitoring of fed-batch penicillin production. Chem. Ing. Tech. 88 (6), 764–776.

Golabgir, A., Hoch, T., Zhariy, M., Herwig, C., 2015. Observability analysis of biochemical process models as a valuable tool for the development of mechanistic soft sensors. Biotechnol. Prog. 31 (6), 1703–1715.

Gonzalez, J., Fernandez, G., Aguilar, R., Barron, M., Alvarez-Ramirez, J., 2001. Sliding mode observer-based control for a class of bioreactors. Chem. Eng. J. 83 (1), 25–32.

Gopalakrishnan, A., Kaisare, N.S., Narasimhan, S., 2011. Incorporating delayed and infrequent measurements in Extended Kalman Filter based nonlinear state estimation. J. Process Control 21 (1), 119–129.

Gordon, N.J., Salmond, D.J., Smith, A.F., 1993. Novel approach to nonlinear/non-Gaussian Bayesian state estimation. IEE Proceedings F (Radar and Signal Processing), vol. 140. IET Digital Library, pp. 107–113. No. 2.

Gudi, R.D., Shah, S.L., Gray, M.R., 1994. Multirate state and parameter estimation in an antibiotic fermentation with delayed measurements. Biotechnol. Bioeng. 44 (11), 1271–1278.

Guo, Y., Huang, B., 2015. State estimation incorporating infrequent, delayed and integral measurements. Automatica 58, 32–38.

Gustafsson, F., 2010. Particle filter theory and practice with positioning applications. IEEE Aerosp. Electron. Syst. Mag. 25 (7), 53–82.

Hermann, R., Krener, A., 1977. Nonlinear controllability and observability. IEEE Trans. Automat. Control 22 (5), 728–740.

Hol, J.D., Schön, T.B., Gustafsson, F., 2006. On resampling algorithms for particle filters. In: Nonlinear Statistical Signal Processing Workshop, 2006 IEEE, pp. 79–82.

Hulhoven, X., Wouwer, A.V., Bogaerts, P., 2006. Hybrid extended Luenberger-asymptotic observer for bioprocess state estimation. Chem. Eng. Sci. 61 (21), 7151–7160.

Johansen, A.M., Doucet, A., 2008. A note on auxiliary particle filters. Stat. Prob. Lett. 78 (12), 1498–1504.

Kalman, R.E., 1960. A new approach to linear filtering and prediction problems. J. Basic Eng. 82 (1), 35–45.

Kantas, N., Doucet, A., Singh, S.S., Maciejowski, J., Chopin, N., 2015. On particle methods for parameter estimation in state-space models. Stat. Sci. 30 (3), 328–351.

Kitagawa, G., 1998. A self-organizing state-space model. J. Am. Stat. Assoc., 1203–1215

Koch, C., Posch, A.E., Goicoechea, H.C., Herwig, C., Lendl, B., 2014. Multi-analyte quantification in bioprocesses by Fourier-transform-infrared spectroscopy by partial least squares regression and multivariate curve resolution. Analytica chimica acta 807, 103–110.

Krämer, D., King, R., 2016. On-line monitoring of substrates and biomass using near-infrared spectroscopy and model-based state estimation for enzyme production by *S. cerevisiae*. IFAC-PapersOnLine 49 (7), 609–614.

Krener, A.J., 2009. Nonlinear Observers. Control Systems, Robotics and Automation, Volume XIII: Nonlinear, Distributed, and Time Delay Systems - II, 153.

Kroll, P., Sagmeister, P., Reichelt, W., Neutsch, L., Klein, T., Herwig, C., 2014. Ex situ online monitoring: application, challenges and opportunities for biopharmaceuticals processes. Pharm. Bioprocess. 2 (3), 285–300.

Larsen, T.D., Andersen, N.A., Ravn, O., Poulsen, N.K., 1998. Incorporation of time delayed measurements in a discrete-time Kalman filter. Proceedings of the 37th IEEE Conference on Decision and Control, 1998, vol. 4. IEEE, pp. 3972–3977.

Leiter, E., Marx, F., Pusztahelyi, T., Haas, H., Pócsi, I., 2004. *Penicillium chrysogenum* glucose oxidase – a study on its antifungal effects. J. Appl. Microbiol. 97 (6), 1201–1209.

Li, T., Sun, S., Sattar, T.P., Corchado, J.M., 2014. Fight sample degeneracy and impoverishment in particle filters: a review of intelligent approaches. Expert Syst. Appl. 41 (8), 3944–3954.

Liu, J., West, M., 2001. Combined parameter and state estimation in simulation-based filtering. In: Sequential Monte Carlo methods in practice. Springer, New York, pp. 197–223.

Luoma, P., Golabgir, A., Brandstetter, M., Kasberger, J., Herwig, C., 2017. Workflow for multi-analyte bioprocess monitoring demonstrated on inline NIR spectroscopy of *P. chrysogenum* fermentation. Anal. Bioanalyt. Chem. 409 (3), 797–805.

Megee, R.D., Kinoshita, S., Fredrickson, A.G., Tsuchiya, H.M., 1970. Differentiation and product formation in molds. Biotechnol. Bioeng. 12 (5), 771–801.

Menezes, J.C., Alves, S.S., Lemos, J.M., de Azevedo, S.F., 1994. Mathematical modelling of industrial pilot-plant penicillin-G fed-batch fermentations. J. Chem. Technol. Biotechnol. 61 (2), 123–138.

Mondal, S.C., Ray, P.K., Maiti, J., 2014. Modelling robustness for manufacturing processes: a critical review. Int. J. Prod. Res. 52 (2), 521–538.

Montague, G.A., Morris, A.J., Ward, A.C., 1989. Fermentation monitoring and control: a perspective. Biotechnol. Genetic Eng. Rev. 7 (1), 147–188.

Nielsen, J., 1997. Physiological Engineering Aspects of *Penicillium chrysogenum*. World Scientific.

Oreshkin, B.N., Liu, X., Coates, M.J., 2011. Efficient delay-tolerant particle filtering. IEEE Trans. Signal Process. 59 (7), 3369–3381.

Patwardhan, S.C., Narasimhan, S., Jagadeesan, P., Gopaluni, B., Shah, S.L., 2012. Nonlinear Bayesian state estimation: a review of recent developments. Control Eng. Pract. 20 (10), 933–953.

Paul, G.C., Thomas, C.R., 1996. A structured model for hyphal differentiation and penicillin production using *Penicillium chrysogenum*. Biotechnol. Bioeng. 51 (5), 558–572.

Paul, G.C., Syddall, M.T., Kent, C.A., Thomas, C.R., 1998. A structured model for penicillin production on mixed substrates. Biochem. Eng. J. 2 (1), 11–21.

Pitt, M.K., Shephard, N., 1999. Filtering via simulation: auxiliary particle filters. J. Am. Stat. Assoc. 94 (446), 590–599.

Posch, A.E., Herwig, C., 2014. Physiological description of multivariate interdependencies between process parameters, morphology and physiology during fed-batch penicillin production. Biotechnol. Prog. 30 (3), 689–699.

Rathore, A.S., Winkle, H., 2009. Quality by design for biopharmaceuticals. Nat. Biotechnol. 27 (1), 26–34.

Revilla, G., López-Nieto, M.J., Luengo, J.M., Martin, J.F., 1984. Carbon catabolite repression of penicillin biosynthesis by *Penicillium chrysogenum*. J. Antibiotics 37 (7), 781–789.

Ruijter, G.J.G., Kubicek, C.P., Visser, J., 2002. Production of organic acids by fungi. In: Industrial Applications. Springer, Berlin, Heidelberg, pp. 213–230.

Schmitz, K., Peter, V., Meinert, S., Kornfeld, G., Hardiman, T., Wiechert, W., Noack, S., 2013. Simultaneous utilization of glucose and gluconate in *Penicillium chrysogenum* during overflow metabolism. Biotechnol. Bioeng. 110 (12), 3235–3243.

Simon, D., 2006. Optimal State Estimation: Kalman, H Infinity, and Nonlinear Approaches. John Wiley & Sons.

Soons, Z.I.T.A., Shi, J., van der Pol, L.A., van Straten, G., van Boxtel, A.J.B., 2007. Biomass growth and kLa estimation using online and offline measurements. IFAC Proc. Vol. 40 (4), 85–90.

Tulsyan, A., Huang, B., Gopaluni, R.B., Forbes, J.F., 2013. On simultaneous on-line state and parameter estimation in non-linear state-space models. J. Process Control 23 (4), 516–526.

United States Department of Health and Human Services, Food and Drug Administration, 2004. Guidance for Industry. PAT - A Framework for Innovative Pharmaceutical Development, Manufacturing, and Quality Assurance. <www.fda.gov/downloads/drugs/guidancecomplianceregulatoryinformation/guidances/ucm070305.pdf>. Version: May 8, 2017.

United States Department of Health and Human Services, Food and Drug Administration, 2009. Guidance for Industry. Q8 (R2) Pharmaceutical Development. <https://www.fda.gov/downloads/drugs/guidances/ucm073507.pdf>. Version: May 8, 2017.

Vojinovic, V., Cabral, J.M.S., Fonseca, L.P., 2006. Real-time bioprocess monitoring: Part I: In situ sensors. Sens. Actuat. B: Chem. 114 (2), 1083–1091.