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

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