proceedings

MDPI

Proceedings

Online State Prediction of S. Cerevisiae Cultivation Purely Based on Ethanol Gas Sensors and an Observer †

Olivier Paquet-Durand *, Abdolrahim Yousefi Darani and Bernd Hitzmann

University of Hohenheim, Stuttgart, Germany; rahim@uni-hohenheim.de (A.Y.D.); bernd.hitzmann@uni-hohenheim.de (B.H.)

* Correspondence: o.paquet-durand@uni-hohenheim.de

† Presented at the 1st International Electronic Conference on Processes: Processes System Innovation, 17–31 May 2022; Available online: https://ecp2022.sciforum.net.

Abstract: To control a bioprocess, the determination of the current state is necessary. Most state variables require substantial effort and time to measure or sometimes are not measurable at all, a direct measurement is not always an option. Instead, an indirect chemometric approach based on some other easier to measure variable such as spectroscopy is commonly used to estimate the state of a bioprocess. In this contribution we present another much cheaper solution for S. cerevisiae cultivations where the only direct measurement were ethanol measurements in the headspace of the bioreactor based on metal oxide gas sensors. For the current state prediction, a process model and an unscented Kalman filter as observer was used. The basic idea is to apply the model to predict the process state, and then use the ethanol measurements to correct and change the model prediction online. The main advantage of this approach is, that metal oxide gas sensors are dead cheap and in contrast to spectroscopic approaches, no expensive calibration is required. The knowledge required is the process model and a rough estimation of the kinetic parameter values.

1. Introduction

The ability to monitor major process state variables, such as biomass, substrate and product concentrations accurately is essential for automatic control of bioprocesses. However, due to the unavailability of inexpensive or dependable measuring systems, rapid on-line measurements of said state variables are often not feasible. So, the development of chemometric software sensors that are capable of achieving rapid and accurate estimation of said process states, is of great interest [1-3].

One example for such software sensors that received a lot of interest lately, are Kalman filters and their non-linear extensions. They can be used for continuous and accurate estimation of the state of bioprocesses. In general, the Kalman filters combine available general knowledge in the shape of a process model and the already available process information such as on-line measurements to an estimation of the true state of the process. Various nonlinear extensions for the Kalman filter are available. They mostly differ in the way how the approximation of the prediction uncertainty is performed. Lisci and Tronci [4] appkied an extended Kalman filter (EKF) to predict the state of a fed-batch cultivation of baker's yeast. The variables of interest were temperature, dissolved oxygen amount and the substrate concentration. Another EKF implementation presented by Popova et al. [5] showed the estimation of product, substrate and biomass concentrations based on the measurements of glucose and ethanol during a *S. cerevisiae batch* cultivation. A nonlinear extension to the Kalman filter is the so called unscented Kalman filter (UKF) in which the uncertainty or covariance of the predicted state is approximated by using the unscented transform [6].

In recent years, several authors demonstrated the application of the UKF for the on-line estimation of state variables and parameters in various processes. For example, Jianlin et al. [7] demonstrated an approach based on a UKF for the prediction of biomass and substrate in a fed-batch cultivation of *S. cerevisiae* based on the measurements of dissolved

Citation: Lastname, F.; Lastname, F.; Lastname, F. Title. *Proceedings* **2022**, *69*, x. https://doi.org/10.3390/xxxxx

Academic Editor: Firstname Lastname

Published: date

oxygen and carbon dioxide. Krämer and King [8] used a Kalman filter in conjunction with the cultivation of *S. cerevisiae* for noise removal from their measured state variables. The measurements themselves were based on near infra-red spectra taken from the fed-batch cultivations.

In this publication the application of an unscented Kalman filter in *S. cerevisiae* batch cultivations is shown for estimation of biomass, glucose and ethanol concentration. Also, the kinetic parameters or growth rates are estimated on-line. The only measurements required are infrequent ethanol measurements in the gas phase. To assess the reliability the suggested method, it was tested on three *S. cerevisiae* batch cultivations that differed slightly in the initial substrate concentrations.

2. Material and Methods

2.1. Yeast cultivation & offline measurements

In total three cultivations of S. cerevisiae, C1, C2, and C3, were carried out. In all three batch cultivations, 100 mL Schatzmann medium [9] was inoculated with 5 g of yeast. The batches were then, after 10 min of shaking, put into a stirred tank reactor (Minifors, Inifors HT, Bottmingen, Switzerland). The glucose concentration in the media used for the batch cultivations was 2.85, 5 and 9 g/L for C1, C2, and C3 respectively. Also 1 mL/L trace elements solution was added. All cultivations were performed at a temperature of 30 °C and a controlled pH at 5. The aeration and agitation rates were also kept at 3.5 L/min and 450 rpm, respectively.

2.2. Ethanol gas sensors

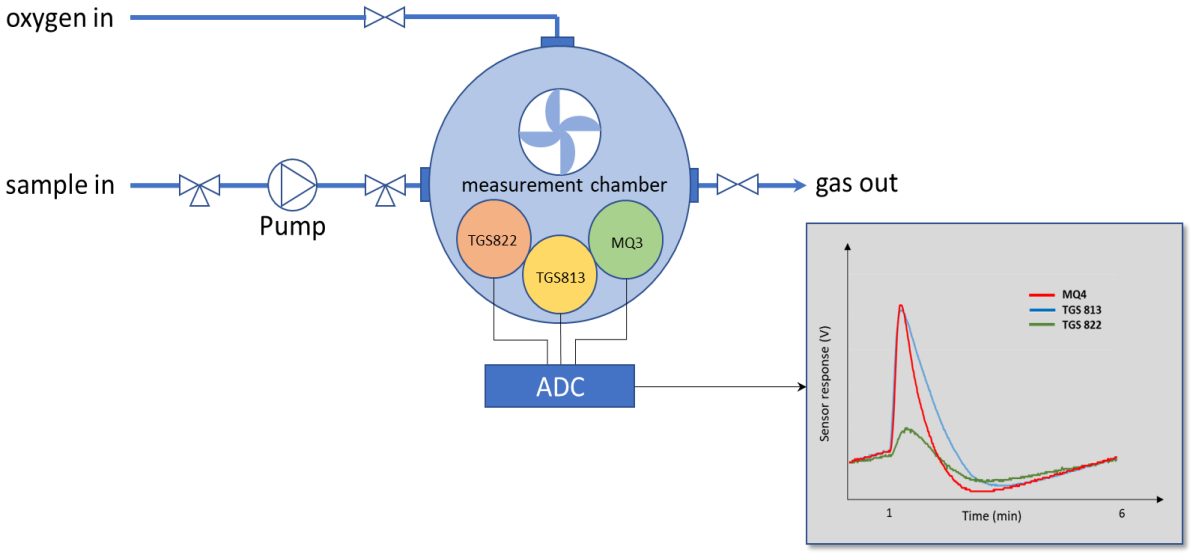

Figure 1. Gas sensor with sampling system and sensor array consisting of three sensors (TGS822, TGS813, MQ3) and an example raw output over a single measurement cycle.

As shown in **Error! Reference source not found.** the airtight measurement chamber with a volume of ~250 mL contained an array of three different commercially available and reasonably cheap tin oxide gas sensors (MQ3, TGS 813 and TGS 822). During a measurement a gas pump (Schwarzer Precision, Essen, Germany) is pumping a continuous sample gas stream with 400 mL/min into and through the measurement chamber for 10 s. Then the sensors need to be regenerated with oxygen. This is done by flushing the measurement chamber with either pure oxygen or compressed air for 90 or 120 s respectively. Then the chamber is sealed by closing all valves. Proper mixing inside the chamber is achieved by a 40 mm fan, running at 4000 rpm. After 3 minutes, the regeneration is finished. One measurement cycle therefore takes about 5 minutes and during the entire cycle

the output of the three gas sensors is captured by a 10-bit ADC and send to a computer for processing. The three 5 minutes long data streams from the three MOS sensors are then evaluated by a Matlab program and through a PLSR model the ethanol concentration in the sample gas stream can then be determined once every 5 minutes. For calibration of said PLSR model, cultivation C1 was used.

2.3. Dynamic process model

When applying a Kalman filter, a process model is required. To do this, the cell growth kinetic is estimated by two Monod terms. The first (main) substrate is glucose and then, after glucose is consumed, ethanol becomes the secondary substrate and therefore growth-limiting factor. The process itself is modelled as a batch process in an ideal stirred tank reactor:

$$\frac{\mathrm{d}X}{\mathrm{d}t} = \mu_G X + \mu_E X \qquad \mu_G = \frac{\mu_{max,G} \cdot \mathrm{G}}{K_G + \mathrm{G}}$$

$$\frac{\mathrm{d}G}{\mathrm{d}t} = -\frac{\mu_G X}{Y_{\frac{X}{G}}} \qquad \mu_E = \frac{\mu_{max,E} \cdot \mathrm{E}}{K_E + \mathrm{E}} \cdot \left(1 - \frac{\mu_G}{\mu_{max,G}}\right)^2$$

$$\frac{\mathrm{d}E}{\mathrm{d}t} = \frac{\mu_G X}{Y_{E/G}} - \frac{\mu_E X}{Y_{X/E}}$$

$$\frac{\mathrm{d}\mu_{max,G}}{\mathrm{d}t} = 0$$

$$\frac{\mathrm{d}\mu_{max,E}}{\mathrm{d}t} = 0$$

Here X, G and E represent the main state variables, namely biomass, glucose and ethanol concentration, respectively. $Y_{X/G}$, $Y_{E/G}$ and $Y_{X/E}$ are the three conversion factors (yields) that describe the conversion ratio from glucose to biomass, glucose to ethanol and ethanol to biomass. μ_G and μ_E stand for the actual specific cell growth rates on glucose and ethanol. They are computed from the maximum specific growth rates $\mu_{max,G}$ and $\mu_{max,E}$ based on Monod kinetics. The values for $\mu_{max,G}$ and $\mu_{max,E}$ are treated as state variables as well so that they are estimated with the Kalman filtering.

2.4. State Estimation

Here an unscented Kalman filter (UKF) was implemented to continuously estimate biomass, glucose and ethanol concentrations. Also, the two main kinetic parameters, the maximum growth rates were continuously estimated and corrected. Like all Kalman filter variants, the UKF estimates the most likely state of a system by weighing a simulated process state, obtained from the process model, and actual measurements. The weights are chosen based on measurement error and model uncertainty. A more detailed explanation of the Kalman implementation used here can be found at [10].

2.5. Offline measurements

For determination of the concentrations of biomass, ethanol and glucose, samples from the bioreactor were taken regularly. Cell dry mass was determined by centrifugation of 1.5 mL of sample at 14,000 rpm for 10 min at 4 °C. The wet cells were placed in a drying cabinet for 24 h at 103 °C. After cooldown for 30 min, the dry mass was weighted. The remaining supernatant was analyzed by HPLC to measure the glucose and ethanol concentrations. For evaluation of the UKF algorithm, the root-mean square error (RMSE) was calculated from the UKF estimated concentrations and the off-line measured concentrations. Also, from the RMSE, the percentage standard errors (SE) were calculated with respect to the highest concentrations:

$$RMSE = \sqrt{\sum_{i=1}^{N} \frac{(\hat{Y}_i - Y_i)^2}{\mathrm{N}}}$$

$$SE\ (\%) = \frac{\sqrt{\sum_{i=1}^{N} \frac{(\hat{Y}_i - Y_i)^2}{N}}}{Y_{max}} \times 100\%$$

$\hat{Y}_i$ stands for the estimated concentration by the UKF method, Y_i stands for the off-line concentration determined by HPLC analysis, N represents the number of measurements and Y_{max} represents the highest concentration in the corresponding off-line value.

3. Results and Discussion

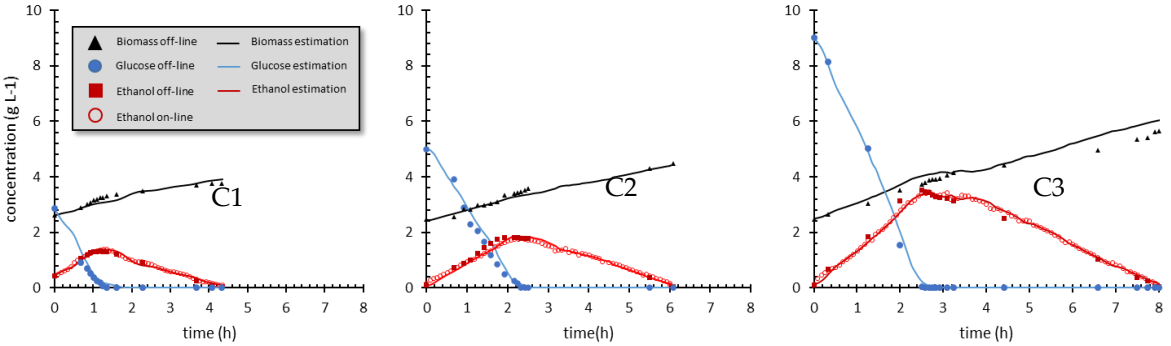

Figure 2. Ethanol measurement and Kalman state estimation for all three cultivations. The offline values have not been used and are only shown for reference.

In figure 2 the UKF estimated concentrations of biomass, glucose and ethanol (solid lines) in the bioreactor are shown. The gas sensor on-line measured ethanol concentrations (hollow red circles) and the HPLC off-line ethanol and glucose as well as biomass concentrations (solid forms) are presented as well. These off-line are only shown as a reference to demonstrate that the estimated values are in fact accurate.

When a difference in the on-line measured and simulated values is detected, the values estimated by the Kalman filter are corrected to be more inline with the measured ones. In cultivation C2, from 2 to 3 h time, there is a deviation between the reported gas sensor value and the off-line reference values. The reason for this could be various factors like fluctuations in temperature or electrical noise influencing the sensor electronics.

The accuracy the Kalman Filter regarding the estimation of state variables was evaluated by calculating the RMSEP and SEP between the estimated ethanol, biomass and glucose concentration and the measured offline concentrations, and the results are presented in Table 1:

Table 1. RMSEP and SEP of off-line measured values and their estimated concentrations by the UKF algorithm.

	Ethanol		Biomass		Glucose	
Experiment	**RMSEP [g/L]**	**SEP [%]**	**RMSEP [g/L]**	**SEP [%]**	**RMSEP [g/L]**	**SEP [%]**
C1	0.15	4	0.29	9	0.13	1.7
C2	0.08	4.5	0.09	5	0.16	4
C3	0.09	4.5	0.1	5	0.16	4

The standard error of estimated ethanol concentration is below 5 % for all three cultivations which is a reasonable result. The error in cultivation C1 is slightly smaller as this cultivation was used to calibrate the chemometric model that was used to feed the UKF. It was also possible to estimate the concentrations of biomass and glucose as well, although they were not measured at all. However, the biomass estimation showed quite a large error of almost 10 % in case of C1.

Figure 3 shows the estimated maximum specific growth rates with respect to the consumption of glucose $\mu_{max,G}$ and ethanol $\mu_{max,E}$ and the specific growth rates itself (μ_G and μ_E).

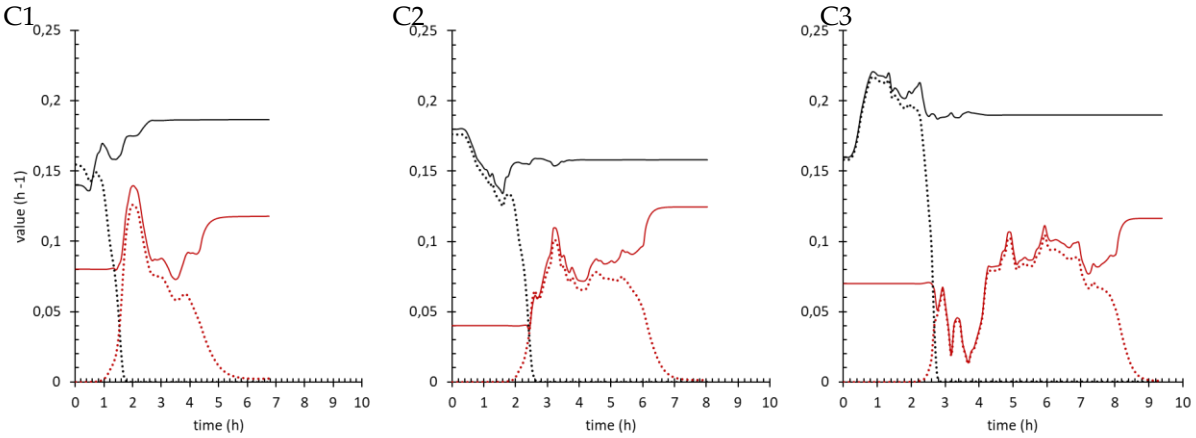

Figure 3. maximal specific growth rates $\mu_{max,G}$ and $\mu_{max,E}$ (solid lines) as well as the actual growth rates μ_G and μ_E (dashed lines) as they are estimated by the Kalman Filter algorithm.

As it can be seen in Figure 3, different starting values for μ_{max} values were chosen for each cultivation. These initial values are based on experience with previous cultivations ($\mu_{max,G} \approx 0.2\ h^{-1}$, $\mu_{max,E} \approx 0.05\ h^{-1}$) and then varied slighty and randomly to see whether the UKF algorithm is capable to correct these values. In C1 the $\mu_{max,G}$ is increasing shortly after the inoculation starts, this indicates that the chosen starting value was lower than the actual value, therefore the UKF algorithm converges to the true value. When the glucose is almost depleted, the metabolic change from glucose to ethanol consumption takes place, therefore $\mu_{max,E}$ would start to increase. However, shortly before glucose is completely depleted, $\mu_{max,G}$ increases which results in the decrease of μ_E, therefore the UKF increases the $\mu_{max,E}$ to compensate for the underestimation. In C2, the specific glucose based growth rate is decreasing slightly. Therefore, it can be assumed, that the selected starting values for the kinetic parameters are close to their actual values. Similarly, in C3, the specific glucose based growth rate is increasing slightly after the inoculation and then decreasing again. This points to the fact, that the selected starting values for the kinetic parameter is lower than its actual value but the initial guess was close. Nonetheless, with the used UKF algorithm the values converge to reasonable values quickly.

4. Conclusion

Here, a dynamic non-linear process model was used in combination with an unscented Kalman filter algorithm for the estimation of kinetic parameter and biomass, glucose and ethanol concentrations of a batch fermentation of *S. cerevisiae*. The algorithm only required on-line data in form of infrequent ethanol measurements from a MOS gas sensor to achieve this.

Three *S. cerevisiae* fermentations with slightly differing initial substrate/glucose concentrations were performed to analyze the behavior and capability of the proposed algorithm. The result indicated that the presented UKF based algorithm was capable of estimating and correcting the specific growth rates on-line. It was also possible to estimate the concentrations of biomass, glucose and ethanol continuously on-line with good accuracy while only actually measuring the ethanol concentration in the headspace with a cheap gas sensor array. The proposed method and algorithm can therefore be used in combination with low-cost gas sensors for monitoring of batch fermentation processes of baker's yeast.

References

1. Dochain, D., Bastin, G. (1990). Adaptive control of fed batch bioreactors. *Chemical Engineering Communications*, *87*(1), 67-85.
2. Dochain, D. (2003). State and parameter estimation in chemical and biochemical processes: a tutorial. *Journal of process control*, *13*(8), 801-818.
3. Bernard, O., Chachuat, B., Hélias, A., Rodriguez, J. (2006). Can we assess the model complexity for a bioprocess: theory and example of the anaerobic digestion process. *Water science and technology*, *53*(1), 85-92.
4. Lisci, S., Grosso, M., Tronci, S. (2020). A Geometric Observer-Assisted Approach to Tailor State Estimation in a Bioreactor for Ethanol Production. *Processes*, *8*(4), 480.
5. Popova, S., M. Ignatova, V. Lyubenova (2013). State and Parameters Estimation by Extended Kalman Filter for Studying Inhomogeneous Dynamics in Industrial Bioreactors.
6. Julier, S., Uhlmann, J., Durrant-Whyte, H. F. (2000). A new method for the nonlinear transformation of means and covariances in filters and estimators. *IEEE Transactions on automatic control*, *45*(3), 477-482.
7. Jianlin, W., F. Xuying, Z. Liqiang and Y. Tao (2010). Unscented transformation based robust Kalman filter and its applications in fermentation process. Chinese Journal of Chemical Engineering 18(3): 412-418.
8. Krämer, D. and R. King (2019). A hybrid approach for bioprocess state estimation using NIR spectroscopy and a sigma-point Kalman filter. Journal of Process Control 82: 91-104.
9. Schatzmann, H. (1975). *Anaerobes Wachstum von Saccharomyces cerevisiae: regulatorische Aspekte des glycolytischen und respirativen Stoffwechsels* (Doctoral dissertation, ETH Zurich).
10. Yousefi-Darani, A., Paquet-Durand, O., Babor, M., Hitzmann, B. (2021). Parameter and state estimation of backer's yeast cultivation with a gas sensor array and unscented Kalman filter. *Engineering in Life Sciences*, 21(3-4), 170-180.