DOI 10.5162/SMSI2020/B3.2

Adaptive Soft Sensor for Bioprocess Monitoring

Manuel Siegl, Vincent Brunner, Dominik Geier, Thomas Becker
Chair of Brewing and Beverage Technology, Weihenstephaner Steig 20, 85354 Freising, Germany
Corresponding Author's e-mail address: dominik.geier@tum.de

Summary:
Soft sensors can be used to predict variables that cannot be measured directly. However, even these soft sensors are subject to errors that reduce the accuracy of the prediction. One way to overcome this is to predict a target quantity redundantly using independent measurement systems for the input variables. This study reports on the development of an algorithmic system for combining the redundant submodels to one reliable soft sensor. The proof of concept was conducted with a *Pichia pastoris* bioprocess.

Keywords: bioprocess monitoring, soft sensor, fault detection, model maintenance

Background, Motivation, and Objective

In biotechnological processes, process variables exist that cannot be measured directly in real-time and therefore have to be determined offline in tedious laboratory analyses. To determine these variables online, a so-called soft sensor can be utilized. However, the sensors contributing inputs to the soft sensors are often subject to interference and measurement errors. The typical error types of sensors can be classified as follows: bias (intermittent, stepwise, drift-wise or cyclic deviation), a reduction in precision, and a temporary or complete failure of sensors [1, 2]. The reasons can be attributed to damaged sensors, connection problems and inadequate calibration [3]. In order to be able to calculate an accurate prediction despite these limitations, more complex systems are required. One way to identify these individual sensor errors is to combine the information from the entire sensor network [4].

In addition to the identification of the errors, automated recalibration is also desirable for some error types (e.g., drift). One possibility to maintain soft sensors is to use selected historical data points [5]. Another option is the direct implementation of new laboratory measurements for the maintenance of the soft sensors [5, 6]. However, especially in bioprocesses, laboratory measurements are often very time-consuming and cannot be integrated into the running process using these methods. For example, in the bioprocess investigated in this work, the cultivation of the yeast *P. pastoris*, the target value (dry cell weight concentration, hereinafter referred to as X) cannot be determined in less than 2 days. This study aims to predict the target value X using several separate submodels and then to statistically interconnect the prediction of the submodels. In addition, it was examined whether incorrect submodels can be identified and directly maintained.

Generation of the process data

Five cultivations were carried out in a Biostat® Cplus bioreactor with a working volume of 15 L at 30 °C and 500 mbar. The dissolved oxygen was maintained at 40 %. As cultivation medium FM22 with an initial glycerol concentration of 40 g L^{-1} was used. The pH of the batch cultivation was controlled to 5 with ammonium hydroxide, which additionally served as a nitrogen source. Data pre-processing and modeling were performed in MATLAB R2019a.

Development of the Adaptive Soft Sensor

Three different submodels were used to predict the biomass concentration X. The first submodel was used to predict the target value using a model based on the pH correction agents (*base submodel*). The second submodel is based on the exhaust gas measurements and includes the calculation of the carbon dioxide emissions rate (*CER submodel*). The third prediction is based on the measurement of a mid-infrared sensor (*MIR submodel*). To predict X using the mid infrared spectrum, a Savitzky-Golay filter was first applied and then the biomass was predicted using partial least squares regression (PLSR).

The sampling interval of the sensors made it possible to predict X every 30 s. For the subsequent interconnection, the three predictions were averaged within 5-minute intervals and their standard deviations were calculated (n = 10). A system based on a t-test with a subsequent minimum variance estimator was used to combine the submodel predictions to an adaptive soft sensor for biomass concentration (*mixed model*).

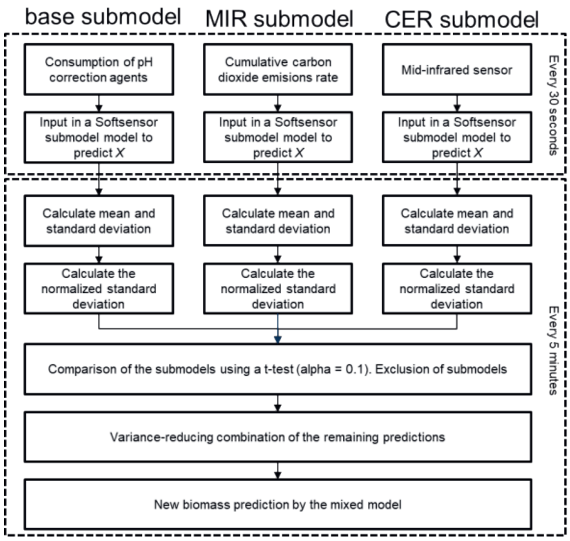

Figure 1: Structure of the adaptive soft sensor.

Performance of the Adaptive Soft Sensor

Three datasets were used to calibrate and validate the single submodels. The validation of the mixed soft sensor model was performed in two ways: on the one hand with the already used data with additional artificial errors and on the other hand with the remaining two data sets.

Figure 2 shows the validation of the system for artificially generated errors. Three errors were added: the intermittent malfunction of the logging of the pH correction agent (**I**); the intermittent increase of the measuring noise of the mid-infrared sensor (**II**) and a stepwise increase of the *CER* based submodel (**III**).

In all three cases, the algorithm of the mixed model was able to compensate for the sensor errors delivering a robust final prediction for *X*. In case of the fault type of stepwise increase (bias), recalibration would theoretically be possible using the two correct soft sensors or historical measurements. With the two datasets not used for calibration, malfunctions occurred mainly at the MIR measurement, which could be compensated for by the developed system (data not shown).

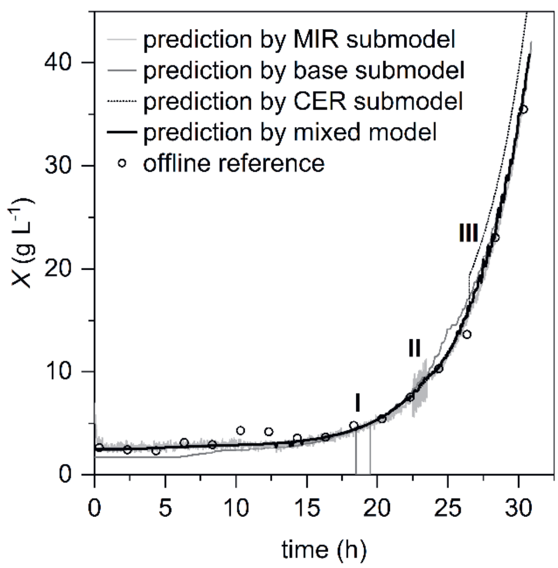

Figure 2: Comparison of the individual submodels with the adaptive soft sensor. Artificial errors were added to the underlying sensor data (I,II,III).

In the future, besides the automated recalibration due to stepwise biases, the behaviour of the system in case of multiple errors will be investigated and improved.

References

[1] Mehranbod, Nasir; Soroush, Masoud; Piovoso, Michael; Ogunnaike, Babatunde A. (2003): Probabilistic model for sensor fault detection and identification. In: *AIChE J.* 49 (7), S. 1787–1802. DOI: 10.1002/aic.690490716.

[2] Isermann, R. (2006). Fault-diagnosis systems: an introduction from fault detection to fault tolerance. Springer Science & Business Media. DOI: 10.1007/3-540-30368-5.

[3] Sharma, Abhishek B.; Golubchik, Leana; Govindan, Ramesh (2010): Sensor faults. In: *ACM Trans. Sen. Netw.* 6 (3), S. 1–39. DOI: 10.1145/1754414.1754419.

[4] Brunner, Vincent; Klöckner, Lukas; Kerpes, Roland; Geier, Dominik Ulrich; Becker, Thomas (2019): Online sensor validation in sensor networks for bioprocess monitoring using swarm intelligence. In: *Analytical and bioanalytical chemistry*. DOI: 10.1007/s00216-019-01927-7.

[5] Yuan, Xiaofeng; Ge, Zhiqiang; Song, Zhihuan (2016): Spatio-temporal adaptive soft sensor for nonlinear time-varying and variable drifting processes based on moving window LWPLS and time difference model. In: *Asia-Pac. J. Chem. Eng.* 11 (2), S. 209–219. DOI: 10.1002/apj.1957.

[6] Chen, Kuilin; Castillo, Ivan; Chiang, Leo H.; Yu, Jie (2015): Soft Sensor Model Maintenance: A Case Study in Industrial Processes∗ In: *IFAC-PapersOnLine* 48 (8), S. 427–432. DOI: 10.1016/j.ifacol.2015.09.005.