Comparison of different state estimator algorithms applied to a simulated anaerobic digestion reactor

Shadi Attar[1] Finn Aakre Haugen [2]

[1,2] Faculty of Technology, Natural Sciences and Maritime Sciences, University College of Southeast Norway, Norway
shadi.attar@usn.no, Finn.Haugen@usn.no

Abstract

This study deals with a simulator-based comparison of different state estimators of an anaerobic digestion process. A simulated biogas reactor based on the AM2 model is considered. Extended Kalman Filter, Unscented Kalman Filter, Particle Filter and Moving Horizon Estimator are four state estimators studied. The investigation is on both states and parameters estimation. The maximum number of parameters can be estimated equals the number of the measurement.

Keywords: Anaerobic digestion, state estimation, Kalman filter, particle filter, moving horizon estimator, simulation

1 Introduction

Anaerobic digestion (AD) is a biological process to produce biogas consisting of methane and carbon dioxide in the absence of oxygen. There are several biogas models to represent the behaviour of the AD reactor. In this study, the Anaerobic Digestion Model number 2 (AM2) (Bernard, et al., 2001; Dochain, 2008) with 6 states and 13 parameters is applied. The model is implemented in MATLAB.

The main goal of this study is states and parameters estimation for a simulated AD reactor based on the AM2 model. Estimation of states and parameters is useful for monitoring and control purposes. Four states estimation methods consisting of Extended Kalman Filter (EKF), Unscented Kalman Filter (UKF), Particle Filter (PF) and Moving Horizon Estimator (MHE) are considered.

State estimation for large-scale wastewater treatment plants is studied in (Busch, et al., 2013). UKF state estimation is surveyed for a biogas rector (Haugen, et al., 2014). A state estimator is developed using a calibrated simulation model of a full-scale biogas plant, which is based on the Anaerobic Digestion Model N0.1. (Gaida, et al., 2012). A number of software sensor design methods, including extended Kalman filters, receding-horizon observers is studied in (Bogaerts & Vande Wouwer, 2003).

The outline of the article is as follows Section 2 includes the process description and the AM2 model equations and observability analysis method. Section 3 explains about state estimation consists of four methods and comparison their results. Parameter estimation are described in Section 4 regarding different number of measurement for estimating both environmental parameters and model parameters. Section 5 presents and discusses about the simulation results. Conclusions are given in Section 6, and future work is suggested in Section 7.

2 Materials and methods

2.1 Process description

The biogas reactor model applied in this article is the AM2 model with parameter values as presented in (Bernard, et al., 2001). The reactor has the form of a cylindrical tank with effective volume of 948 litres. The temperature of reactor will be constant due to a temperature control system.

2.2 Mathematical model

The mathematical model which is the basis of this simulator, is generated from mainly mass balance and bacterial kinetics equations. Moreover, ionic balance, affinity constant, bacterial kinetics, ideal gas law and Henry's law are considered in the model (Dochain, 2008).

2.2.1 Material balances

In Anaerobic digestion process, organic material is converted by microorganisms in two phases are called acidogenesis and methanization. In the first phase, the acidogenic bacteria (X_1) consume the organic substrate (S_1) and produce CO_2 and volatile fatty acids (S_2). In the second phase, the population of methanogenic bacteria (X_2) uses for produce the methane. Z and C are the total alkalinity and total inorganic carbon, respectively.

Mass balances gives the following differential equations, which constitute a state space model of the AD reactor:

$$\frac{dX_1}{dt} = [\mu_1 - \alpha D]X_1 \quad (1)$$

$$\frac{dX_2}{dt} = [\mu_2 - \alpha D]X_2 \quad (2)$$

$$\frac{dZ}{dt} = D(\mathrm{Z_{in}} - Z) \quad (3)$$

$$\frac{dS_1}{dt} = D(S_{1\mathrm{in}} - S_1) - k_1\mu_1 X_1 \quad (4)$$

$$\frac{dS_2}{dt} = D(S_{2\mathrm{in}} - S_2) + k_2\mu_1 X_1 - k_3\mu_2 X_2 \quad (5)$$

$$\frac{dC}{dt} = D(\mathrm{C_{in}} - C) - q_C + k_4\mu_1 X_1 + k_5\mu_2 X_2 \quad (6)$$

D is the dilution rate defined as:

$$D = \frac{F_{\mathrm{feed}}}{V} \quad (7)$$

where F_{feed} is the infeed flowrate and V is the effective volume of medium in the reactor. D represents a normalized flow. $Z_{\mathrm{in}}, S_{1\mathrm{in}}, S_{2\mathrm{in}}, C_{\mathrm{in}}$ in (3) - (6) are respectively concentration of the inflow of alkalinity, substrate, VFA and dissolved inorganic carbon.

q_C represents the flow of inorganic carbon form liquid phase to gas phase is calculated accordance to (8) - (10):

$$q_C = k_L a(C + S_2 - Z - K_H P_C) \quad (8)$$

where $k_L a$ and K_H are liquid-gas transfer constant and Henry's constant, respectively. P_C is partial carbon dioxide pressure and it can be calculated as:

$$P_C = \frac{\emptyset - \sqrt{\emptyset^2 - 4K_H P_T(C + S_2 - Z)}}{2K_H} \quad (9)$$

with

$$\emptyset = C + S_2 - Z + 2K_H P_T + \frac{k_6}{k_L a}\mu_2 X_2 \quad (10)$$

Methane flow is directly related to the methanogenic rate (μ_2):

$$q_m = k_6\mu_2 X_2 \quad (11)$$

The model equation for the pH is:

$$pH = -\log\left(K_b \frac{C + S_2 - Z}{Z - S_2}\right) \quad (12)$$

where K_b is an affinity constant.

2.2.2 Bacterial kinetics

The models to describe the growth of microorganism are assumed Monod-type and Haldane-type (Dochain, 2008). The growth of acidogenic bacteria, μ_1, is considered based on Monod type kinetics:.

$$\mu_1 = \mu_{1\max}\frac{S_1}{S_1 + K_{S_1}} \quad (13)$$

The growth of methanogenic bacteria, μ_2, is assumed based on Haldane kinetics:

$$\mu_2 = \mu_{2\max}\frac{S_2}{S_2 + K_{S_2} + \frac{S_2^{\,2}}{K_{I_2}}} \quad (14)$$

where $\mu_{1\max}$ and $\mu_{2\max}$ are respectively the maximum growth rate of acidogenic biomass and methanogenic biomass.

2.3 Observability

The standard method to check the observability of a linear model is used in this study. At first, the AM2 model is linearized and the model coverted to (15):

$$\frac{dx}{dt} = Ax + BD$$
$$y = Cx + ED \quad (15)$$

The observability matrix for (15) is computed according to (16)

$$OB = \begin{bmatrix} C \\ CA \\ CA^2 \\ \vdots \\ CA^{n-1} \end{bmatrix} \quad (16)$$

n is the number of elements in the state vector. If the rank of the observability matrix is less than n, the system is non-observable, i.e. there are state variables that are non-observable.

3 State estimation

3.1 Introduction

In the practical systems, we are not able to measure some of variables which have effects on control and monitoring purpose. To have reliable monitoring and control, we need to estimate these variables as far as possible. A suitable state estimator is essential for this aim. As shown in Figure 1, the state estimation is used in the control system. However, the controller is not described in the present article.

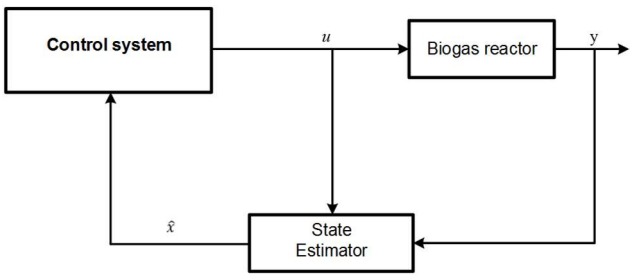

Figure 1. A state estimator scheme.

The state vector of the AM2 model is:

$$x = [\, X_1,\, X_2, Z\,, S_1\,, S_2\,, C\,] \quad (17)$$

It is decided to estimate some of parameters such as $S_{1\text{in}}, S_{2\text{in}}$ as the environmental variables and k_6 as the model parameter in the parameter estimation section. Four states estimator methods consisting of EKF, UKF, PF (Simon, 2006) and MHE (Boegli, 2014) are implemented based on the AM2 model.

The dilution rate, D is as an input variable, and q_m is considered as process measurement.

$$y = q_m \tag{18}$$

Where y is measurement vector.

The observability of the state vector is tested regarding the standard method of observability analysis and only four state variables are observable.

Initial values of the states for the simulated model are based on a steady state condition and shown in Table 1. Initial value of all state estimators are determined with considering a large initial estimation errors in X_1, X_2, Z to check the convergence rate of the methods. See Table 1.

We need to determine P, Q and R which are state estimation error covariance, process noise covariance and measurement noise covariance, respectively. They are computed as:

$$\begin{aligned} P &= \text{diag}(K_P\ {x_{init}}^2) \\ Q &= \text{diag}(K_Q\ {x_{init}}^2) \\ R &= \text{diag}(K_y\ {y_{init}}^2) \end{aligned} \tag{19}$$

where x_{init} is assumed as an initial values of augmented state vector, y_{init} is an initial values of measurement vector. K_P, K_Q and K_y are coefficients equal 0.1, 0.01 and 1 respectively, select by trial and error.

Table 1. The initial values for the simulated model and the state estimators

Initial values (x_{init})	*Simulated model*	*State estimators*	*Unit*
X_{1init}	0.39	1.39	g/L
X_{2init}	0.85	1.85	g/L
Z_{init}	62.2	72.2	mmol/L
S_{1init}	1.17	3.17	g/L
S_{2init}	2.78	2.78	mmol/L
C_{init}	67.9	70.9	mmol/L

The amount of biogas depends on the dilution rate. During the simulation, dilution rate would be changed to check its effect and it is assumed that the dilution rate is always known.

The model parameters are known and considered based on (Bernard, et al., 2001).

Noise is added to the simulated measurements.

3.2 Extended Kalman Filter

This method is based on linearizing the non-linear system around a nominal state set. The linear model is only used for calculation of Kalman gain.

The procedure of implementation of EKF method in the simulation follows the principle of the discrete-time extended Kalman filer presents in (Simon, 2006). The outcomes of this method for 15 days is shown in Figure 2 with the magenta curves.

3.3 Unscented Kalman Filter

The EKF is based on linearization to propagate the mean and covariance of states. UKF is an enhancement of the Kalman filter to reduce the linearization error of the EKF.

The procedure of implementation of UKF method in the simulation follows the principle of the discrete-time unscented Kalman filer resent in (Simon, 2006). Figure 2 presents the results of UKF state estimator for all six state variable in AM2 model with the black curves.

3.4 Particle Filter

This method is based on statistical approach. The particle filter is a probability-based estimator as an enhancement on Bayesian estimator (Simon, 2006). The number of particles is assumed to be 100. The results of particle filter in the same condition on UKF and EKF are shown in Figure 2 with the green curves.

3.5 Moving Horizon Estimator

This method is an optimization problem over the specific horizon, N, which is shifted in terms of time.

The procedure of implementation of MHE method in the simulation follows the principles in (Boegli, 2014). The horizon, N, in this simulation equals to 10.

The objective function to minimize is:

$$\begin{aligned} \min_x \quad & \sum_{k=t-N}^{t} \|x_{k+1} - f(x_k, D_k)\|^2{}_Q \\ & + \sum_{k=t-N}^{t} \|y_k - h(x_k, D_k)\|^2{}_R \\ & + \|x_{t-N} - \hat{x}_{t-N}\|^2{}_P \end{aligned} \tag{20}$$

subject to: $C - S_2 - Z > 0$

y_k is the process measurements.

Since $C - S_2 - Z$ equals to CO_2, this term should be positive. The first to third sums in (20) represents dynamical system mismatch, measurement mismatch and arrival cost, respectively.

The arrival cost term represents the error of the state estimate at the "end" of the horizon. In our implementation, the state at k = t - N is estimated, thereby implicitly minimizing the arrival cost term. Therefore, we have omitted the arrival cost term as an explicit term.

The results of MHE are shown in Figure 2 with the red curves.

3.6 Comparison of the various state estimators

All four state estimators are executed on the same conditions and with the same initial values. It is assumed that dilution rate, D, changes after one day from 0.34 to 1 d^{-1} and after 7.5 days, D is changed to 0.5 d^{-1}.

Two criteria are supposed for comparison among of the state estimation methods. The first criterion is the Mean Squared Error (MSE), and the second criterion is computational time is represented by the simulation execution time for each state estimation method.

MSE relative to EKF for four state estimators is shown in Table 2. Regarding the relative MSE, the minimum MSE belongs to MHE methods so the MHE method have fittest curve by smallest MSE. However, the computational time has remarkable larger computational burden of other methods, refer to Table 3.

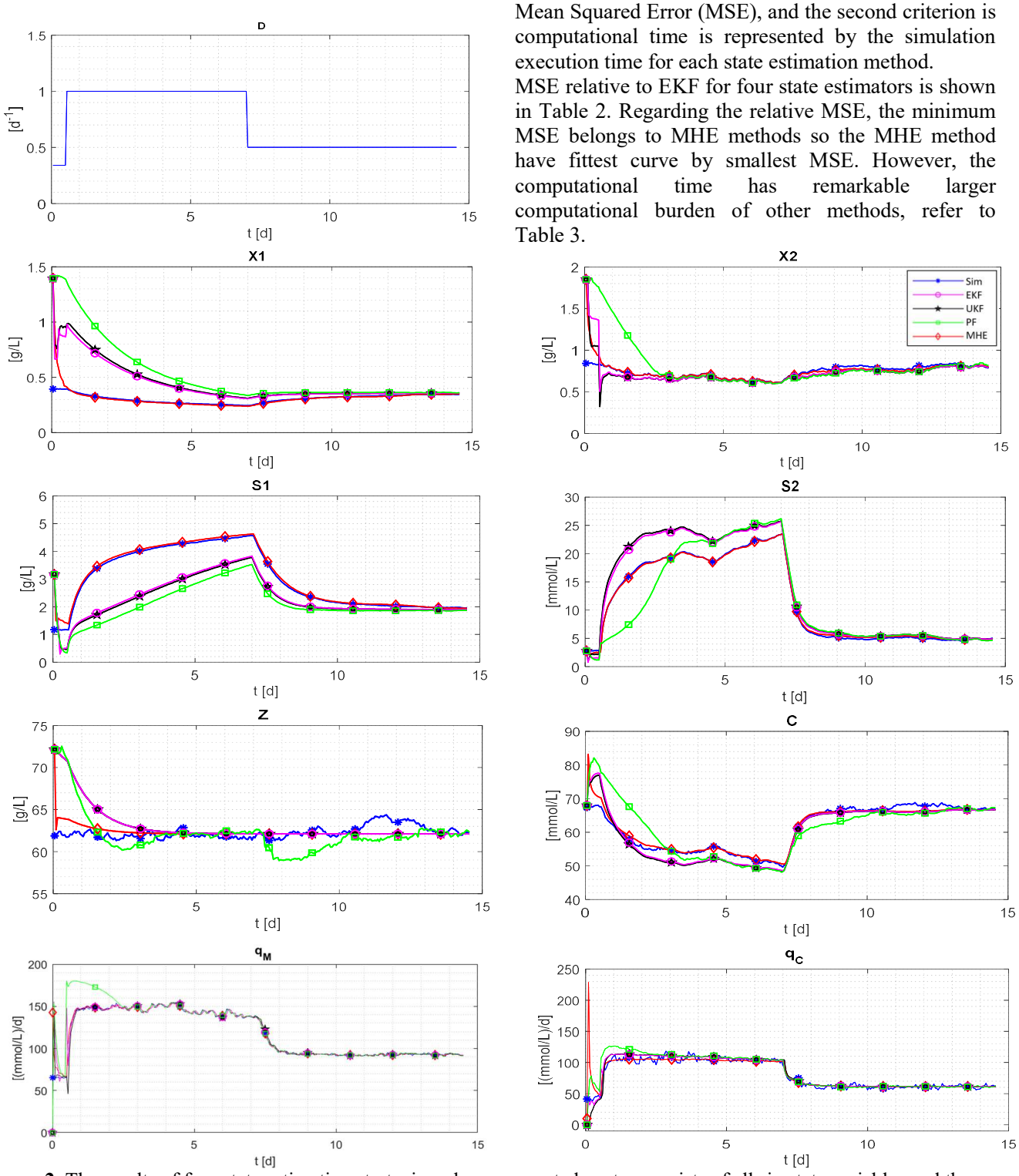

Figure 2. The results of four state estimation strategies when augmented vector consists of all six state variables and the methane gas flow rate is the only measurement.

The convergence rate for the MHE strategy is fastest comparing with other methods in all estimated states. EKF has faster computation but the Jacobians need to be calculated.

Table 2. MSE relative to EKF for four state estimation methods regarding one measurement and six state variables

Relative MSE	*EKF*	*UKF*	*PF*	*MHE*
X_1	1	1.48	4.10	0.16
X_2	1	0.5	3.82	0.27
Z	1	1	1.02	0.54
S_1	1	1.28	2.06	0.03
S_2	1	1.35	1.60	0.02
C	1	1	2.25	0.44
q_m	1	0.55	3.08	0.48
q_c	1	1.36	2.83	0.83

Table 3. The relative computational burden to EKF measured with simulation time.

	EKF	*UKF*	*PF*	*MHE*
Computational time	1	13	62	14892

4 Parameters estimation

Some parameters are not known nor measured while they are essential parameters for control and monitoring. Such parameters should therefore be estimated.

4.1 One measurement and one estimated parameter

The organic substrate concentration of the inflow, $S_{1\text{in}}$, is considered as a first parameter for estimation. The augmented state vector is:

$$x = [\, X_1,\, X_2, Z\, , S_1\, , S_2\, , C\, , S_{1\text{in}}] \qquad (21)$$

The parameter state variables are here denoted augmentation state variables.

Initial values of $S_{1\text{in}}$ for the simulated model and state estimators are 9.5 and 14.5 g/L, respectively. The methane gas flow rate is assumed as measurement, so measurement vector is (18). All conditions are the same as mentioned in Section 3. When one measurement is considered, two states are non-observable.

Table 4 shows the relative MSE to the EKF for all four methods. Based on the relative MSE, the MHE method is not reliable approach to estimate the parameters.

Since the error is remarkably large with MHE, this method is ignored to be able to compare other estimators. We focus on the result for S_1 and S_2 state variables and the assumed parameter variable in Figure 3. However, there are two non-observable states based on the rank test, EKF, UKF and PF methods can estimate the states. It is shows that standard observability analysis are not effective for non-linear systems.

Table 4. MSE relative to EKF for four state estimation methods for one measurement and six state variables plus $S_{1\text{in}}$ as the estimated parameter

Relative MSE	*EKF*	*UKF*	*PF*	*MHE*
X_1	1	9.82	9.54	4
X_2	1	2.43	9.73	0.87
Z	1	1	1.02	7.48
S_1	1	2.06	2.58	718.758
S_2	1	0.67	4.44	31.84
C	1	1	1.14	7.55
$S_{1\text{in}}$	1	0.67	1.48	2.34
q_m	1	0.867	1.829	8.280
q_c	1	1.099	1.99	58.13

Table 4 shows the EKF and the UKF have better results for the parameters estimation. In the following sections, we focus on just the EKF and UKF.

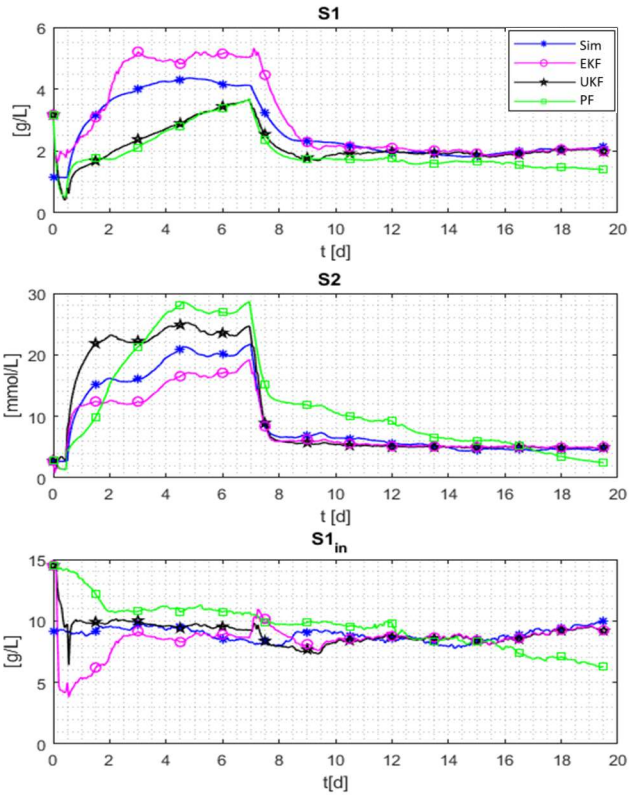

Figure 3. The results of the state estimators consisting of EKF, UKF and PF for an augmented state vector with 7 elements and one measurement.

4.2 One measurement and two estimated parameters

The VFA concentration of the inflow, $S_{2\text{in}}$, is considered as second parameter for estimation. So the augmented state vector is:

$$x = [\, X_1,\, X_2, Z\, , S_1\, , S_2\, , C\, , S_{1\text{in}}, S_{2\text{in}}] \qquad (22)$$

All conditions are the same as mentioned in Section 4.1 plus the initial values of $S_{2\text{in}}$ for the simulated model and state estimators are 93.5 and 133.5 mmol/L, respectively. Three states are non-observable when there is just one measurement.

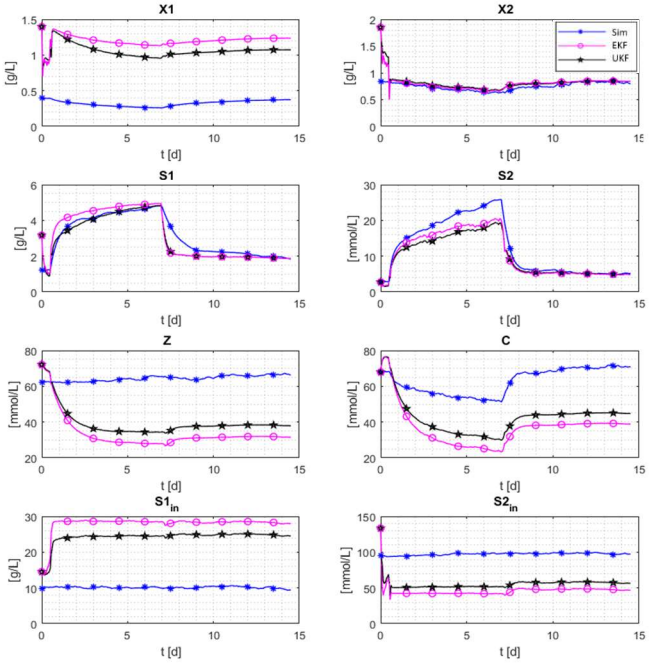

Figure 4. The results of the state estimators consisting of EKF , UKF for a augmented state vector with 8 elements and one measurement.

The results states that it is not possible to estimate two parameters with one measurement. This lack the measurement ruins also the state estimation, refer to Figure 4. The states X_1, Z, C are not possible to estimate.

In the Section 4.5, it will be checked the possibility of the estimation of two parameters with two measurement.

4.3 Two measurements and one estimated parameter

It is of interest to investigate the case when the number of measurement is greater than the number of estimated parameters. Carbon dioxide gas flow rate is considered as the second process measurement. so the measurement vector is :

$$y = [q_m\, , q_c\,] \qquad (23)$$

One state is non-observable based on the rank test. The results are shown in Figure 5.

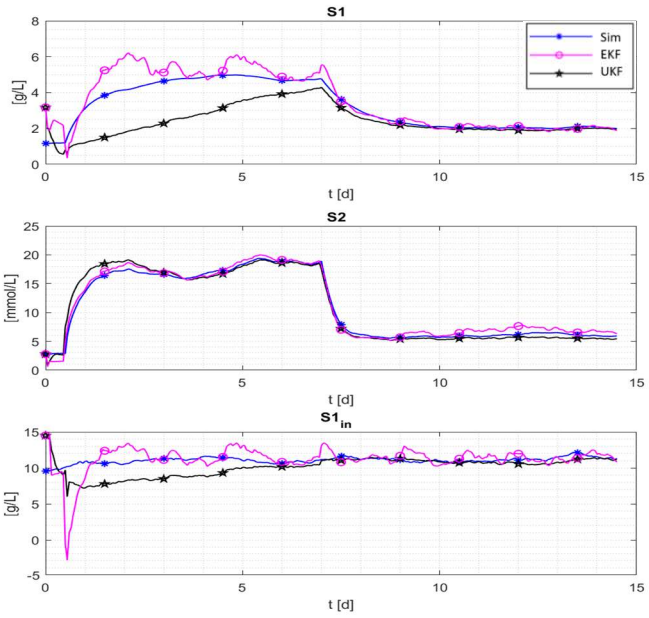

Figure 5. The results of the EKF, UKF state estimators for an augmented state vector with 7 elements and one measurement.

4.4 Three measurements and one estimated parameter

All conditions are the same as the mentioned condition in Section 4.3 except considering pH of the reactor as third measurement. The measurement vector is:

$$y = [q_m\, , q_c\, , pH] \qquad (24)$$

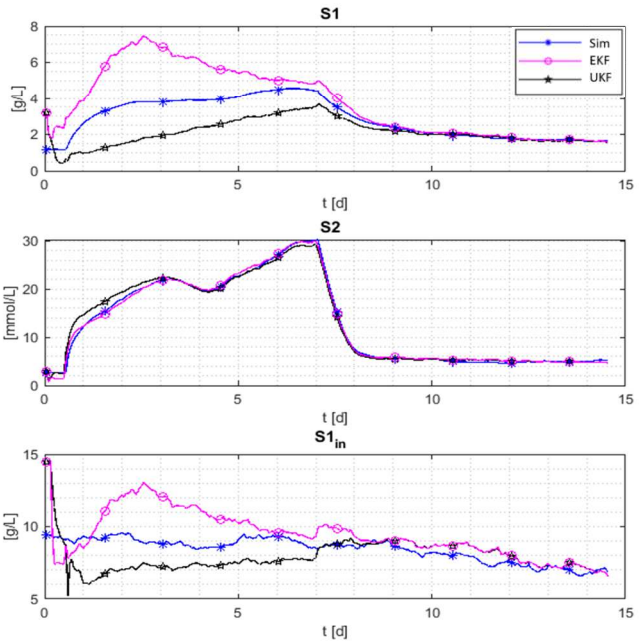

Figure 6. The results of the EKF, UKF state estimators a augmented state vector with 8 elements and three measurement.

All state become observable. The results are shown in Figure 6 for the EKF and the UKF.

Regarding the results in Section 4.3 and 4.4, we can estimate parameters if the number of measurements is larger than the number of the estimated parameters.

4.5 Two measurements and two estimated parameter

It is supposed the augmented state vector and the measurement vector are according to equation (22) and (23), respectively. There is one non-observable state based on the rank test of the observability matrix.

The results of the simulation are shown in Figure 7. Regarding to the results, two measurements are required at least to estimate two parameters.

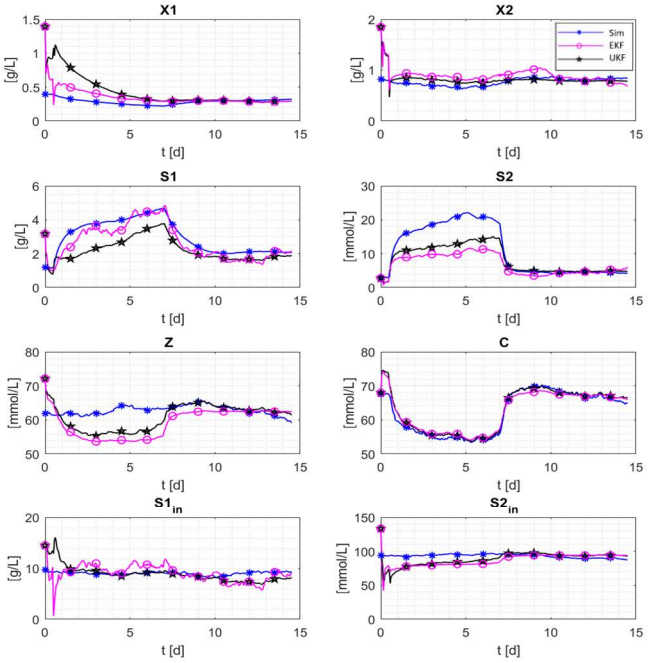

Figure 7. The results of the EKF, UKF state estimators for an augmented state vector with 8 elements and two measurement.

4.6 Two measurements and three estimated parameter

Until now, the environmental parameters are estimated, namely inflow concentrations $S_{1\text{in}}$ and $S_{2\text{in}}$. It is assumed that third estimated parameter is selected among of model parameters. One of the most important model parameters which has a direct effect on the methane production is k_6. The augmented state vector is

$$x = [\, X_1,\, X_2, Z\,, S_1\,, S_2\,, C\,, S_{1\text{in}}, S_{2\text{in}}\,, k_6] \qquad (25)$$

The initial values of k_6 for the simulated model and state estimators are 435 and 300 mmol/L. the measurement vector is assumed as (23). All states are observable. The results are shown in Figure 8. Since the number of measurements is less than the number of estimated parameters, it is not possible to estimate the parameters.

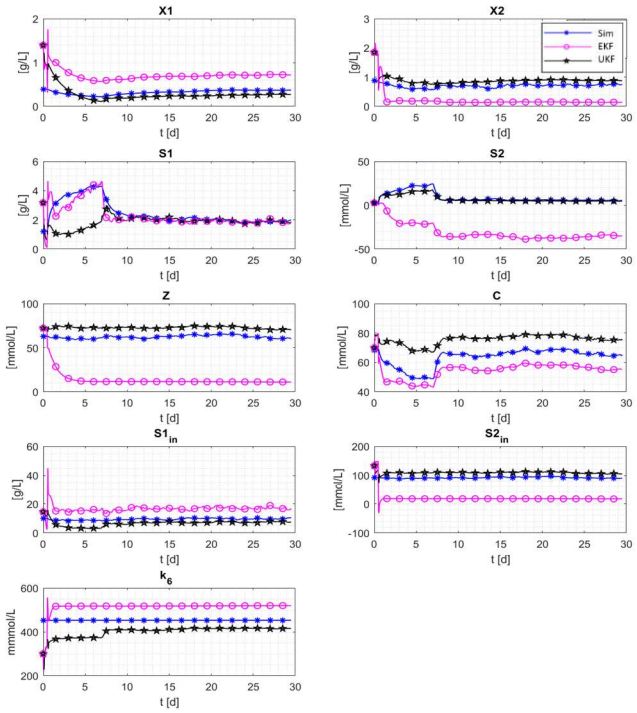

Figure 8. The results of the EKF , UKF state estimators for an augmented state vector with 9 elements and two measurement.

4.7 Three measurement and three estimated parameter

The augmented state vector and the measurement vector are according to equation (25) and (24), respectively.

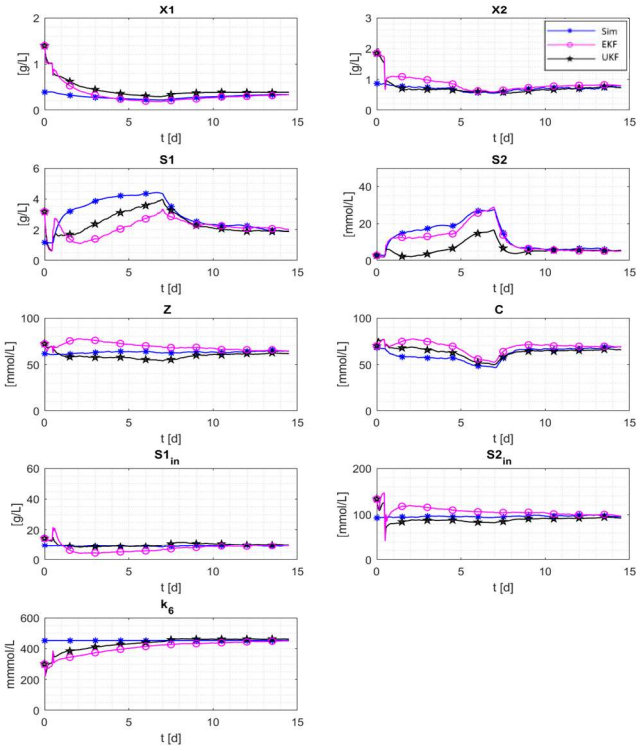

Figure 9. The results of the EKF, UKF state estimators for an augmented state vector with 9 elements and three measurements.

All state are observable based on the rank test on observability matrix.

The results shows in Figure 9 and presents that three parameters can estimate with three measurements.

5 Discussions

The results shows that all four methods can estimate all states of AM2 model for a biogas reactor.

In this case, the MHE is the best state estimation method to estimate the all six states, but computational times is excessive comparing with other methods. However, the MHE is not suitable for estimating the parameters, we assume this is because the optimization problem is not well conditioned. We have calculated the Hessian of the objective function during the estimation. The Hessian is positive definite, but with a large ratio of its eigenvalues, and hence the MHE has found a minimum. However, this may well be a local minimum.

Both the EKF and the UKF give a good estimation with lower implementation effort. Both methods are good for state and parameter estimation. For the EKF, there is an extra computation demand related to calculation of the Jacobians needed for calculation of the Kalman gain. While the UKF straightforward approach and Jacobian free algorithm.

In this study, PF can estimated both states and parameters, but the accuracy is lower comparing to EKF and UKF.

In EKF and UKF, the maximum number of parameter can be estimated is related to the number of the measurements. The maximum number of parameters can be estimated equals the number of the measurement. See Table 5.

Table 5. Assessment of the state estimation based on the number of measurements and estimated parameters

Meas#	1	1	1	2	2	3	2	3
Orig. states#	6	6	6	6	6	6	6	6
Aug. states #	0	1	2	1	2	1	3	3
Non-ob. state#	2	2	3	1	1	0	0	0
Assessment	*P*	*P*	*F*	*P*	*P*	*P*	*F*	*P*

**P= Pass, F=Fail*

6 Conclusions

In this study, the four model-based state estimation methods regarding the AM2 model are investigated. These methods are EKF, UKF, PF and MHE which are evaluated in the simulation case study.

Further parameter estimations are verified regarding to different number of measurement and different number of parameter to estimate.

7 Future work

Plans for future work are: Observability analysis for nonlinear models; Analysis of robustness of the estimators; Application to real AD reactor data.

Acknowledgements

Thanks to USN for financial supporting.

Nomenclature

C	total inorganic carbon concentration (mmol/L)
D	dilution rate (d^{-1})
k_1	yield for substrate degradation
k_2, k_3	VFA yields (mmol/L)
k_4, k_5	CO2 yields (mmol/L)
k_6	yield for CH4 production (mmol/L)
K_b	Affinity constant (mol/L)
K_H	Henry's constant (mmol/L per atm)
k_La	liquid-gas transfer constant (d^{-1})
K_{I_2}	inhibition constant (mmol/L)
K_{S_1}	half-saturation constant (g/L)
K_{S_2}	half-saturation constant (mmol/L)
P_T	total pressure (atm)
q_C, q_m	CO_2 and CH_4 flow rates (mmol/L per d)
S_1	organic substrate concentration (g/L)
S_2	volatile fatty acids concentration (mmol/L)
X_1, X_2	concentration of acidogens and methanogs (g/L)
V	effective volume of medium in the reactor (L)
Z	total alkalinity (mmol/L)
α	fraction of bacteria in the liquid phase
μ_{max}	maximum specific growth rate (d^{-1})

References

Bernard, O. et al., 2001. Dynamical model development and parameter identification for an anaerobic wastewater treatment process. *Biotechnology and bioengineering,* 75(4), pp. 424-438.

Boegli, M., 2014. *Real-Time Moving Horizon Estimation,Application to Friction State and Parameter Estimation,* Leuven: s.n.

Bogaerts, P. & Vande Wouwer, A., 2003. Software sensors for bioprocesses. *ISA Transaction,* Volume 42, pp. 547-558.

Busch, J., Elixmann, D., Kuhl, P. & Gerkens, C., 2013. State estimation for large-scale wastewater. *Water research,* Volume 47, pp. 4774-4787.

Dochain, D., 2008. *Bioprocess Control.* London: Wiley.

Gaida, D., Wolf, C., Meyer, C. & Stuhlsatz, A., 2012. State estimation for anaerobic digesters using the ADM1. *water Science & Technology ,* Volume 66, pp. 1088-1095.

Haugen, F. A., Bakke, R. & Lie, B., 2014. State estimation and model-based control of pilot Anaerobic digestion reactor. *Journal of Control Science and Engineering.*

Simon, D., 2006. *Optimal state estimation.* Hoboken, NewJersey: Wiley Insterscience.