ELSEVIER

Chemical Engineering Journal 106 (2005) 53–58

Chemical Engineering Journal

www.elsevier.com/locate/cej

Dynamic modeling of biogas production in an UASB reactor for potato processing wastewater treatment

E.M.P. Barampouti, S.T. Mai, A.G. Vlyssides*

Chemical Engineering School, National Technical University of Athens, 9 Heroon Polytechniou Street,
Zographou 15700, Greece

Received 19 November 2003; received in revised form 26 June 2004; accepted 26 June 2004

Abstract

In this work dynamic mathematical models for the prediction of biogas production in a UASB reactor were developed. The dynamic modeling technique was applied successfully to a 2-year data record from a potato wastewater treatment plant. The technique used included regression analysis by residuals. Seventeen parameters were examined including the following: wastewater's flow rate, reactor's temperature and pH, total and soluble influent COD, wastewater's temperature and pH, total and soluble effluent COD, volatile fatty acids, alkalinity, biogas production rate and each parameter with a time lag of up to 10 days. Finally, after all parameters and all time lag trials three models were the best fitted models that were developed. The models' adequacy was checked by χ^2 test for a data record of the same UASB reactor but at a different time period and proved to be satisfactory. Additionally, a comparison among the three models was conducted as far as their ability to predict and to control biogas production rate is concerned. Through these models various aspects of the process can be enlighten, such as the fact that the hydrolysis of starch requires a resident time of 9 days.

Keywords: Dynamic model; Anaerobic digestion; Upflow anaerobic sludge blanket; Potato processing wastewater

1. Introduction

Wastewater in the potato processing industry contains high concentrations of biodegradable components such as starch and proteins [1–3]. This wastewater can be considered as complex wastewater because of a rather high concentration of suspended solids, a high content of insoluble fraction of COD and significant quantities of potential foaming substances, such proteins and fats (Table 1) [4].

In contrast to a typical wastewater anaerobic treatment plant where the biogas production rate is directly related to the consumption of the organic load, in a potato processing wastewater treatment plant, this is not the case. This fact can be attributed to the presence of suspended organic solids (starch) whose hydrolysis duration render the prediction of biogas production more complicated. Therefore, the operational parameters' fluctuation with time has a strong impact on the UASB reactor's control. Taking into consideration that there is a great number of measurable parameters (handlable and not) that affect the efficiency of UASB, a dynamic mathematical model of biogas production rate based on time series analysis of these parameters is of major importance for the plant control.

The aim of this work is the application of a suitable methodology, so as to derive a dynamic mathematical model for the control of an operating industrial anaerobic plant. The methodology chosen is the regression analysis by residuals, whose main advantages are:

- The model's construction only needs data of routine determinations usually performed in any industrial plant.
- The derived model takes into account all the particularities of the specific plant thus can successfully control plant's operation.

---

[*] Corresponding author. Tel.: +30 210 7723268; fax: +30 210 7723269.
E-mail address: avlys@tee.gr (A.G. Vlyssides).

1385-8947/$ – see front matter © 2004 Elsevier B.V. All rights reserved.
doi:10.1016/j.cej.2004.06.010

Nomenclature

$A_{t=I}$	parameter A with time lag $t = i$ days
ALK	bicarbonate alkalinity (mequiv./L)
COD	soluble effluent COD (mg/L)
$\mathrm{COD_{in}}$	soluble influent COD (mg/L)
pH	reactor's pH
$\mathrm{pH_{in}}$	wastewater's pH
Q	wastewater's flow rate (m^3/day)
Q_B	biogas production rate (m^3/day)
$\hat{Q}_B$	predicted value of biogas production rate (m^3/day)
Q_{COD}	$Q \times$ COD/1000, soluble effluent COD flow (kg/day)
$Q_{\mathrm{COD_{in}}}$	$Q \times \mathrm{COD_{in}}$/1000, soluble influent COD flow (kg/day)
Q_{TCOD}	$Q \times$ TCOD/1000, total effluent COD flow (kg/day)
$Q_{\mathrm{TCOD_{in}}}$	$Q \times \mathrm{TCOD_{in}}$/1000, total influent COD flow (kg/day)
R^2	correlation coefficient
Res	$Q_B - \hat{Q}_B$, residuals are calculated by substituting data-set values into the regression equation of the previous stage and subtracting them from the corresponding measurement of biogas (m^3/day)
T	reactor's temperature (°C)
T_{in}	wastewater's temperature (°C)
TCOD	total effluent COD (mg/L)
$\mathrm{TCOD_{in}}$	total influent COD (mg/L)
TSS	total suspended solids (mg/L)
UASB	upflow anaerobic sludge blanket
VFA	volatile fatty acids (mequiv./L)
VFA/ALK	buffering capacity of the medium
VSS	volatile suspended solids (mg/L)

2. Methodology

2.1. Dynamic model

The dynamic model used in this study was developed from measurements recorded at equally spaced time intervals. If the response at time t is denoted by Y_t, the model will contain terms of the form:

$$Y_{t-1}, Y_{t-2}, \dots, Y_{t-n}$$

where Y_{t-1} is the response one sampling period in the past, Y_{t-2} the response two sampling periods in the past, and so on.

Additionally, for variables, X_j, which act as inputs, terms of the following type will appear in the model:

$$X_{j,t}, X_{j,t-1}, X_{j,t-2}, \dots, X_{j,t-m}$$

where $X_{j,t}$ is the current measurement of variable j at time t.

The model form, which is linear in the coefficients, is

$$\begin{aligned} Y_t &= k_0 + A_1 Y_{t-1} + A_2 Y_{t-2} + \cdots + A_n Y_n \\ &\quad + k_{10} X_{1,t} + \cdots + k_{1m} X_{1,t-m} + k_{20} X_{2,t} \\ &\quad + \cdots + k_{2n} X_{2,t-n} + \cdots \end{aligned}$$

This model is called a lagged regression model because the variables that are the “independent variables” are current values or values at previous times or “lags” [5].

Table 1
Main characteristics of wastewater

Q (m^3/day)	110–1047
TSS (mg/L)	1300–5340
VSS (mg/L)	1040–4272
$\mathrm{TCOD_{in}}$ (mg/L)	4800–9025
$\mathrm{BOD_5}$ (mg/L)	1200–3650
T (°C)	12–29
pH	6.14–6.9

2.2. Residual analysis

Building the regression models by residual analysis will be presented in this article. The method consists of the following steps:

Step 1: Choose the variable best correlated with the Y-variable, transform it as necessary to produce a straight line, and perform a least-squares regression with the dependent variable (Y-variable to be predicted with a correlation coefficient R_0). The result will be an equation of the form:

$$\hat{Y} = b_0 + b_1 f(X_1)$$

where $\hat{Y}$ is the predicted value of Y-variable, b_0 and b_1 the constants, and X_1 the variable.

Step 2: Calculate “residuals” as follows:

$$Z_i = Y_i - [b_0 + b_1 f(X_{1,i})]$$

where Z_i is the residuals, Y_i the data for variable to be predicted and $X_{1,i}$ the data for variable X_1.

Step 3: Choose the best-correlated X-variable. Transform the X-variable, if necessary, to yield a linear plot.

Step 4: Add the new, transformed variable to the regression model, and perform a least-squares fit by computer, resulting in:

$$\hat{Y} = b_0 + b_1 f(X_1) + b_2 g(X_2)$$

Step 5: Calculate residuals and repeat the process until all variables have been added. Each time the correlation coefficient R of the model is

$$R = R_0 + R_1(1 - R_0) + \cdots$$

Each term of the equation expresses the participation of each variable in the final correlation coefficient.

Step 6: Check the goodness-of-fit of the model. Moderate deviations from a straight line may not be serious [6]. The adequacy of a theoretical model implies the difference between the observed and the expected results. This was checked by and χ^2 test.

2.3. Case study

In this study, multiple linear regression was used to develop a discrete dynamic model for a UASB reactor. For the construction of the dynamic model a 2-year historical data record from a UASB reactor of a potato processing wastewater treatment plant was used. The reactor's volume was 600 m^3 and the mean hydraulic retention time of the wastewater was about 1 day. A highly variable waste load, in terms of concentration, flow and operating conditions, exists because of the many batch type processes used in the manufacturing plant.

In order to control the reactor's operation and efficacy, various parameters such as wastewater's flow, temperature and pH, UASB reactor's temperature and pH and biogas produc-

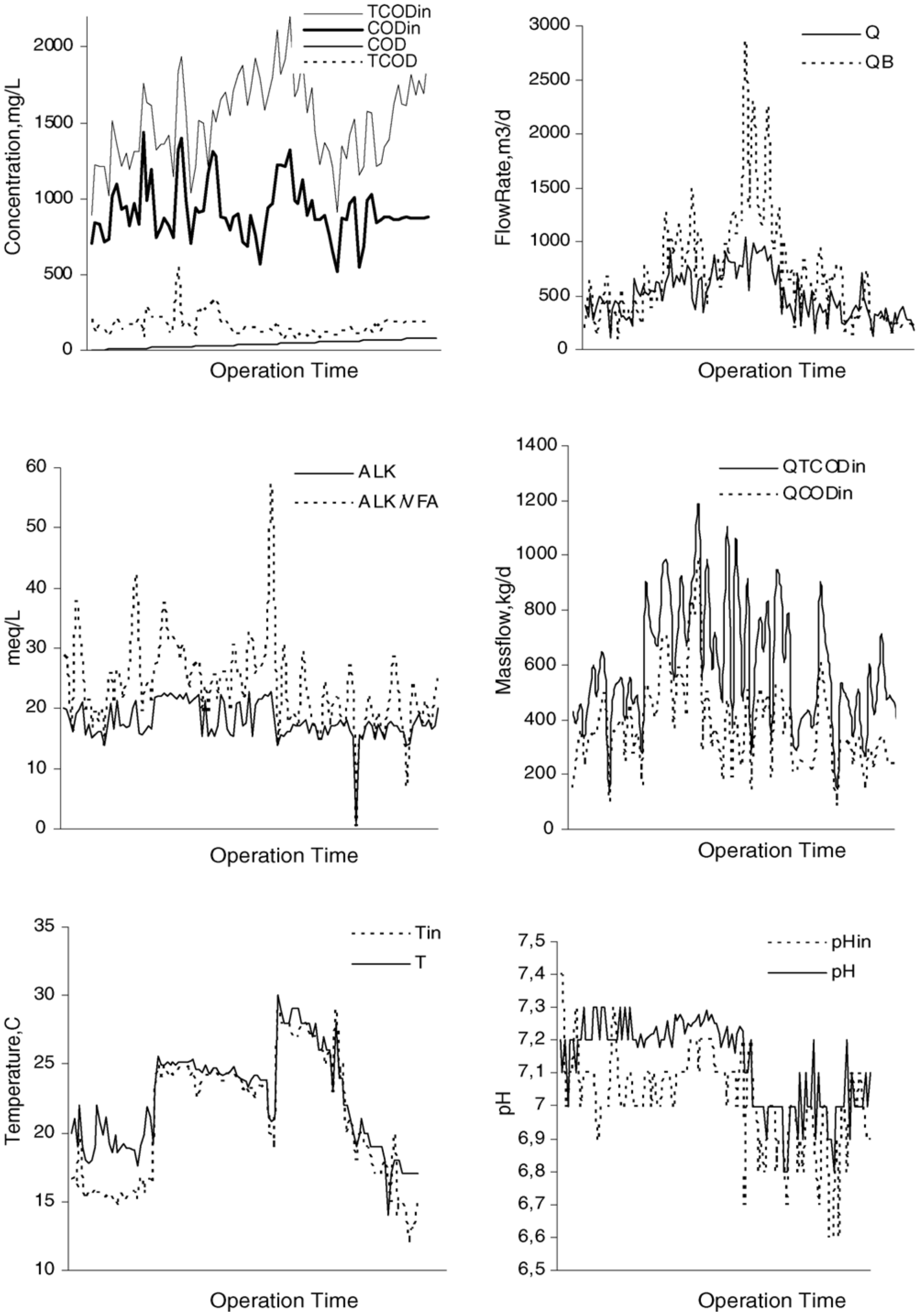

Fig. 1. Fluctuation of wastewater's characteristics.

tion rate are on-line measured and so daily measurements of these parameters are available. Apart from these, total and soluble influent COD, total and soluble effluent COD, volatile fatty acids and alkalinity are frequently measured in the plant's laboratory. Subsequently, a data record of 12 independent variables (Fig. 1) was accessible so as to build a prediction model.

This study's main objective was to correlate the biogas production rate with the independent variables of the data record with a time lag up to 10 days. The correlation was achieved using the technique mentioned above, regression by residual analysis.

During the analysis, apart from these 12 measured variables, correlation took place examining the following parameters: total and soluble influent COD flow, total and soluble effluent COD flow and the queen ratio. The correlation attempt included the equations:

$$Y = A + BX$$

$$Y = A\,\mathrm{e}^{BX}$$

$$Y = A + B \log X$$

$$Y = A + B\sqrt{X}$$

$$Y = A + \frac{B}{X}$$

3. Results and discussion

Using the methodology mentioned above and the data of Fig. 1, three models almost equivalent, as far as correlation coefficient is concerned, were developed.

3.1. Model 1

The variables that were strongly correlated with biogas production rate were:

- wastewater's flow rate, Q (m^3/day), with time lag $t = 0$ day;
- total influent COD concentration, TCOD$_{\mathrm{in}}$ (mg/L), with time lag $t = 9$ days;
- soluble influent COD mass flow, QCOD$_{\mathrm{in}}$ (kg/day), with time lag $t = 0$ day;
- soluble effluent COD concentration, COD (mg/L) with time lag $t = 9$ days.

The dynamic model developed to relate biogas production rate to these variables was

$$\begin{aligned}\hat{Q}_{\mathrm{B}} &= 1.587524 Q_{t=0} + 0.533289 \mathrm{TCOD}_{\mathrm{in},t=9} \\ &\quad - 27,88853\sqrt{Q_{\mathrm{COD_{in}},t=0}} + 1.134708 \mathrm{COD}_{t=9} \\ &\quad - 663.6038 \quad (R^2 = 0.976) \qquad (1)\end{aligned}$$

Table 2
Levels of the regression by residual analysis of Model 1

Level	Best-fitted variable	Variable's participation % in the final R^2
First	$Q_{t=0}$	74.0
Second	TCOD$_{\mathrm{in},t=9}$	15.2
Third	$Q_{\mathrm{COD_{in}},t=0}$	6.5
Fourth	COD$_{t=9}$	1.9

The variables of the consecutive levels of regression analysis are shown in Table 2.

3.2. Model 2

The variables that were strongly correlated with biogas production rate were:

- biogas production rate, Q_{B} (m^3/day), with time lag $t = 5$ days;
- ratio VFA/ALK (buffering capacity) with time lag $t = 2$ days.

The dynamic model developed to relate biogas production rate to these variables was

$$\begin{aligned}\hat{Q}_{\mathrm{B}} &= 0.816816 Q_{\mathrm{B},t=5} - \frac{22.5839}{\mathrm{VFA/ALK}_{t=2}} \\ &\quad + 572.0485 \quad (R^2 = 0.962) \qquad (2)\end{aligned}$$

In this case, the variables of the consecutive levels of regression analysis are shown in Table 3.

3.3. Model 3

The variables that were strongly correlated with biogas production rate were:

- biogas production rate, Q_{B} (m^3/day), with time lag $t = 1$ day;
- wastewater's flow rate, Q (m^3/day), with time lag $t = 7$ days;
- biogas production rate, Q_{B} (m^3/day), with time lag $t = 10$ days.

The dynamic model developed to relate biogas production rate to these variables was

$$\begin{aligned}\hat{Q}_{\mathrm{B}} &= 0.519448 Q_{\mathrm{B},t=1} + 0.592813 Q_{t=7} \\ &\quad + 0.259759 Q_{\mathrm{B},t=10} - 171.33448 \quad (R^2 = 0.958) \qquad (3)\end{aligned}$$

Table 3
Levels of the regression by residual analysis of Model 2

Level	Best-fitted variable	Variable's participation % in the final R^2
First	$Q_{\mathrm{B},t=5}$	69.0
Second	VFA/ALK$_{t=2}$	27.2

Table 4
Levels of the regression by residual analysis of Model 3

Level	Best-fitted variable	Variable's participation % in the final R^2
First	$Q_{B,t=1}$	69.8
Second	$Q_{t=7}$	13.8
Third	$Q_{B,t=10}$	12.2

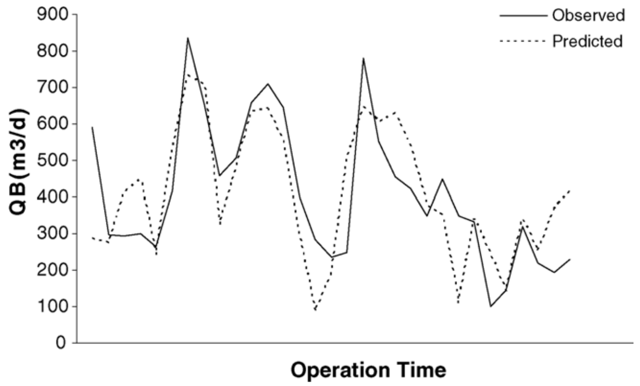

Fig. 2. Biogas production rate, predictions and observed values for the new time period (Model 1).

The variables of the consecutive levels of regression analysis of the third model are shown in Table 4.

3.4. Goodness-of-fit test

The correlation coefficient R that was calculated for the resultant model cannot give sufficient information for its adequacy. In other words, it cannot predict how the model will react in an unknown data range. In order to check the models, χ^2 test were conducted for a data record of the same UASB reactor but at a different time period.

Figs. 2–4 compare the observed values to the predicted values of Models 1–3 respectively for the new study period. They are plots of predicted biogas production rate based on previous actual waste and system operating data. The models predict the biogas production rate in all cases adequately well.

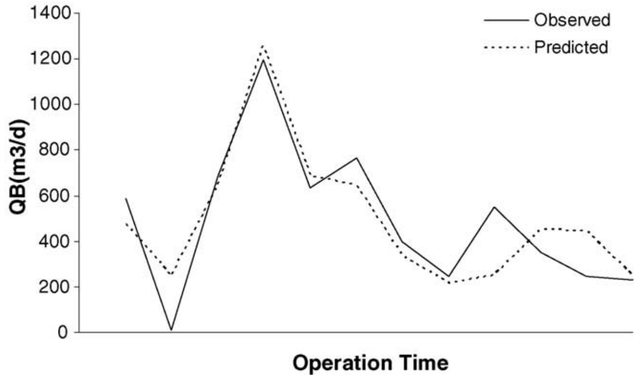

Fig. 3. Biogas production rate, predictions and observed values for the new time period (Model 2).

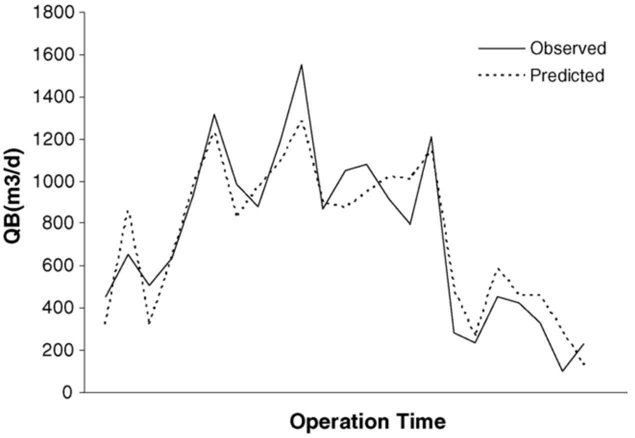

Fig. 4. Biogas production rate, predictions and observed values for the new time period (Model 3).

Table 5
Goodness-of-fit using χ^2 test

Model	Degrees of freedom	χ^2	Results
1	4	0.28	Perfect
2	4	1.5	Perfect
3	4	3.9	Perfect

The results of test χ^2 that are shown in Table 5 reveal that the models can be a satisfactory prediction tool for the specific plant.

4. Conclusions

The methodology of regression analysis by residuals for the construction of a dynamic model proved to be very satisfactory. The three models that arose from data of routine determinations in an industrial plant can be used as a powerful tool for the plant's control.

Despite the fact that the three models have the same effectiveness to estimate the biogas production rate, they are quite different regarding their ability to predict and control biogas production rate. This ability is based on how handlable the parameters are and how long is their time lag. Model 1 has a diminished ability for prediction and thus control due to the variables Q and $Q_{\mathrm{COD_{in}}}$ that have time lag $t = 0$ day. Moreover, although $\mathrm{COD_{in}}$ and COD have a time lag of $t = 9$ days, they cannot easily be manipulated, especially when the plant does not include by-pass. Model 2 has a satisfactory ability of prediction because of the time lags of the parameters ($Q_{B,t=5}$, $\mathrm{VFA/ALK}_{t=2}$) but it cannot control the process since they are not at all handlable. Regarding Model 3, its ability to predict is very satisfactory but has little to offer as far as control is concerned. Unfortunately it is proved that pH and temperature, which are handlable parameters, do not have any participation in any of the models developed.

Another worth notice conclusion that came up is that the long time lags of Models 1 and 3 (9 and 10 days) are possibly due to the required hydrolysis time of suspended solids of starch.

Acknowledgement

This work was supported by TASTY FOOD S.A.

References

[1] U. Abeling, C.F. Seyfried, Anaerobic–aerobic treatment of potato starch wastewater, Water Sci. Technol. 28 (2) (1993) 165–176.

[2] I. Hadjivassilis, S. Gajdos, D. Vanco, M. Nicolaou, Treatment of wastewater from the potato chips and snacks manufacturing industry, Water Sci. Technol. 36 (2/3) (1997) 329–335.

[3] G.R. Zoutberg, Z. Eker, Anaerobic treatment of potato processing wastewater, Water Sci. Technol. 40 (1) (1999) 297–304.

[4] S. Kalyuzhnyi, L. Estrada de los Santos, J. Rodrí guez-Martínez, Anaerobic treatment of raw and preclarified potato-maize wastewaters in a UASB reactor, Bioresour. Technol. 66 (1998) 195–199.

[5] J.L. Hansen, A.E. Fiok, J.C. Hovious, Dynamic modeling of industrial wastewater treatment plant data, J. WPCF 52 (7) (1980) 1966–1975.

[6] R.M. Ingels, How to use the computer to analyze test data, Chem. Eng. (1980) 145–156.