

FUZZY-SUPPORTED ESTIMATOR FOR HPO-AS PROCESS

By Mark T. Yin,¹ Weibo Yuan,² and Michael K. Stenstrom,³ Fellow, ASCE

ABSTRACT: Lack of real-time measurements is a major problem in the operation and control of the high-purity oxygen activated-sludge process. A real-time estimator using dissolved oxygen measurements in each stage and liquid flow rates was developed to estimate biomass and substrate concentrations, biomass growth, and decay rates. A fuzzy algorithm was used to estimate unmeasured variables, such as influent substrate and recycle biomass concentrations. The convergence of the algorithms used for the estimator is fast and stable, even with a large range of initial inputs and noisy dissolved oxygen measurements. The estimated results compare well with both plant data and synthetic results produced by a process model and corrupted with white noise. The difference in the predictions of single substrate and structure models is demonstrated. The estimator was tested successfully for certain types of process upsets, such as shock hydraulic loading and high diluted sludge volume index.

INTRODUCTION

One of the most challenging problems to operators of the activated sludge (AS) process is the lack of real-time measurements. This leads to poor observability and poor system performance. The high-purity oxygen (HPO) AS process, which is characterized by HPO feed (90–98% oxygen) and covered aeration tanks in series, is more complex to operate than conventional open-air AS processes. More quantitative controls are needed. It is crucial for the operator to know the state of the ongoing process when making operational decisions. The known states and parameters of the process also can be utilized for process control. A real-time estimator for the HPO-AS process would be valuable because it can estimate many unmeasured states.

Constructing a real-time estimator is a common practice in the control engineering field (Aborhey and Williamson 1978; Ljung 1979). A number of successful applications of this technique for fermentation have been reported (Stephanopoulos and San 1984; Ramirez 1987; Shimizu and Takamatsu 1989; Dochain 1992; and Dochain et al. 1992). In the fermentation process, such as a fed-batch bioreactor, the input substrate usually is known, and biomass is not recycled. For the AS process, the influent substrate concentration is highly stochastic and unknown, and biomass is recycled to maintain biomass concentration in the aeration tank. The recycled biomass concentration usually is not measured real time. This greatly increases the complexity of applying real-time estimation techniques to the AS process. Meditch and Hostetter (1974) developed an algorithm for systems with unknown inputs, but the algorithm can be applied only to constant-coefficient linear systems.

Previous applications of real-time state and parameter estimation techniques for the AS process also can be found. Holmberg and Olsson (1985) presented a simultaneous estimation scheme for $K_L a$ and oxygen uptake rate (OUR) based on a linear Kalman filter, taking advantage of the differing timescales of the two variables. Marsili-Libelli (1990) constructed a real-time estimator to predict $K_L a$ and OUR using linear approximation. The estimator was coupled with a self-tuning proportional-integral-derivative regulator. Their results

confirm that useful estimates can be obtained in all cases. Ayesa et al. (1991) developed a real-time state and parameter estimation algorithm for identifying the International Association on Water Quality model parameters. They employed a recursive nonlinear, extended Kalman filter to simulate the behavior of a specific AS process under the steady and transient conditions. The results showed rapid convergence and accurate state estimation even with noisy data.

The purpose of this study was to develop a method to estimate important, but unmeasurable variables in the HPO-AS process using available real-time measurements and process models. We employed an asymptotic algorithm (Dochain et al. 1992) to estimate biomass and substrate concentrations in each stage using dissolved oxygen (DO) measurements. The specific growth and decay rates were estimated simultaneously based on the estimated biomass and substrate concentrations using a recursive least-squares algorithm (Young 1984; Dochain and Bastin 1990). The OURs for each stage also were estimated. An unstructured or single-substrate HPO-AS model was used for the real-time estimator (Lawrence and McCarty 1970; Stenstrom et al. 1989). A real-time fuzzy logic algorithm (Pedrycz 1993) was used to estimate the unmeasured influent substrate and recycle biomass concentrations. The recycle biomass concentration was estimated using the effluent total suspended solid (TSS) with 30 fuzzy rules and a mass balance on the secondary clarifier. In this way, the off-line measurement of dilute sludge-volume index (DSVI) was incorporated.

The estimator was calibrated using pilot plant data (Samstag 1989; Stenstrom 1990; Yuan et al. 1993). The performance and simulated results of the estimator were compared with the pilot plant data and process data from a full-scale HPO-AS plant. A structured model (Stenstrom 1990) also was used to simulate the process for verification. A complete description of the model, controller, and fuzzy sets and rules is available elsewhere (Yin 1995).

HPO-AS PROCESS AND MODEL

The HPO process is different from the conventional open-air AS process because of its use of HPO (90–98%) and the process configuration. Fig. 1 shows a typical HPO-AS process. The aeration tanks are covered and arranged in series (usually 3–6 stages) to increase treatment efficiency and oxygen utilization. The oxygen normally is fed to stage 1, which operates at slightly higher pressure than atmospheric (usually approximated 1.008 atm or 5–15 cm water column). The mixed liquor DO concentrations in each stage are higher (6–10 mg/L) than commonly found in open-air AS processes.

Primary effluent to the aeration tanks can be manipulated depending on operational needs. The primary effluent can be fed to stage 1 in the conventional way or can be fed to stage 2 to establish a reaeration mode or fed directly to stage 4 to

¹Asst. Res. Engr., Envir. R & D, Shell Devel. Co., Houston, TX 77251-1380.

²Engr., Marshal Industries, El Monte, CA.

³Prof., Civ. and Envir. Engrg. Dept., UCLA, Los Angeles, CA 90095-1593.

Note. Associate Editor: M. T. Suidan. Discussion open until July 1, 1999. To extend the closing date one month, a written request must be filed with the ASCE Manager of Journals. The manuscript for this paper was submitted for review and possible publication on March 4, 1998. This paper is part of the *Journal of Environmental Engineering*, Vol. 125, No. 2, February, 1999. ©ASCE, ISSN 0733-9372/99/0002-0137-0145/\$8.00 + \$.50 per page. Paper No. 17902.

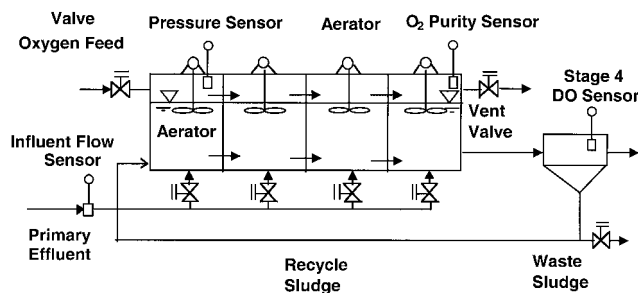


FIG. 1. Flow Diagram for Typical Four-Stage HPO-AS Process

TABLE 1. Summary of Some Measurements for HPO-AS Process

Variable name (1)	Measured on-line (2)	Measured off-line (3)	Unmeasured (4)
Influent flow rate	×		
Influent substrate concentration		×	×
Recycle flow rate	×		
Recycle biomass concentration		×	×
Oxygen partial pressure	×		
DO concentration	×		
Sludge wasting flow rate	×		
Effluent TSS		×	×
DSVI		×	
Effluent BOD ₅ (or chemical oxygen demand)		×	×

Note: × denotes the most likely method(s) of measurement.

protect sludge inventory when hydraulic shock loading is encountered (Torpey 1948; Clift et al. 1983). The estimator was operated in the reaeration mode to conform to the pilot plant operation (Samstag 1989). The estimator can work in any feed mode and we demonstrate this later by simulating a hydraulic shock load.

To construct the real-time estimator for liquid phase operation, it was necessary to examine the on- and off-line measurements used in the HPO-AS process. Table 1 summarizes some of the possible measurements that can be used in estimator design. Stage DO concentration is one major measurement required for the estimator that can be measured easily in most HPO-AS plants. Inexpensive, reliable DO and oxygen partial pressure instruments are available. The estimator also requires the real-time measurements of influent, recycle biomass and waste biomass flow rates.

Another important parameter to run the estimator is the oxygen transfer coefficient ($K_L a$). In most cases, $K_L a$ should be proportional to the impeller speed (surface aerator) or gas recirculation rate (submerged turbine aerator). Various empirical equations may be used to estimate stage $K_L a$, and the following equation is used in this research:

$$K_L a = 0.143(P)^{0.9} \quad (1)$$

where P = stage propeller (kW). This equation is site and process specific and should not be applied elsewhere. Using this equation, $K_L a$ becomes a known parameter for the estimator.

A conventional Monod-type dynamic model developed by Stenstrom et al. (1989) was used as the governing equation for both state and parameter estimators. The liquid phase model takes the form

$$\frac{d}{dt} \begin{bmatrix} DO \\ S \\ X \\ CO \\ NH \\ DN \end{bmatrix}_i = \begin{bmatrix} K_1 & K_2 \\ K_3 & K_4 \\ K_5 & K_6 \\ K_7 & K_8 \\ K_9 & K_{10} \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \mu_g \\ \mu_d \end{bmatrix}_i - \begin{bmatrix} DO \\ S \\ X \\ CO \\ NH \\ DN \end{bmatrix}_i D_i + \begin{bmatrix} DO_0 \\ S_0 \\ X_0 \\ CO_0 \\ NH_0 \\ DN_0 \end{bmatrix}_i D_{in} + \begin{bmatrix} OTR \\ 0 \\ 0 \\ CTR \\ 0 \\ NTR \end{bmatrix}_i \quad (2)$$

where DO = dissolved oxygen concentration; S = substrate concentration; X = the biomass concentration; CO = dissolved carbon dioxide concentration; NH = ammonia concentration; DN = dissolved nitrogen gas concentration; μ_g and μ_d are the specific growth rate and decay rates, respectively; subscripts o and i = influent and stage number, respectively; coefficients K_1 – K_{10} = various kinetic and stoichiometric coefficients; D_i and D_{in} = dilution ratios for influent flow rate plus recycle flow rate and influent flow rate, respectively; and OTR, CTR, and NTR = gas transfer rates that include the mass transfer coefficients and driving force terms. The gas phase balance is not presented but has been published previously (Stenstrom et al. 1989).

Eq. (2) can be generalized as follows:

$$\frac{d\xi_i}{dt} = K\mu_i(\xi_i) - D_i \xi_i + F_i + TR_i(\xi_i) \quad (3)$$

where ξ_i = state vector consisting of DO, substrate, biomass, carbon dioxide, ammonia, and dissolved nitrogen concentra-

TABLE 2. Definition and Values of Estimator Parameters

Element (1)	Formula (2)	Explanation (3)	Value (4)
K_1	$-(1 - Y)Y_{O21}/Y$	Y , cell yield, mass X /mass S	$Y = 0.55$
K_2	$-Y_{O22}$	Y_{O21} , mass O_2 /mass S	$Y_{O21} = 1.42$
K_3	$-1/Y$	Y_{O22} , mass O_2 /mass X	$Y_{O22} = 1.42$
K_4	0		
K_5	1		1
K_6	-1		-1
K_7	$(1 - Y)Y_{CO21}/Y$	Y_{CO21} , mass CO_2 produced/mass S converted	$Y_{CO21} = 1.37$
K_8	Y_{CO22}	Y_{CO22} , mass CO_2 produced/mass X oxidized	$Y_{CO22} = 1.95$

Note: All time units are in hours (after Tzeng 1992).

TABLE 3. Definition and Values of Structured Model Parameters

Parameter (1)	Value (2)	Description (3)
bci	0.012	Decay coefficient (h^{-1})
bsstor	0.405	Transfer coefficient
bstor	0.5	Transfer coefficient
f_{cstrm}	0.6	Maximum fraction (m/m)
K_{cstrm}	0.05	Saturation coefficient (m/m)
K_{oex}	1.42	O_2 stoichiometric coefficient (m/m)
K_{O2sol}	1.10	O_2 stoichiometric coefficient (m/m)
K_{O2str}	1.10	O_2 stoichiometric coefficient (m/m)
K_{SO2}	2.0	O_2 saturation coefficient (mg/L)
m_{sol}	0.006	Maximum growth rate (h^{-1})
m_{stor}	0.75	Maximum growth rate (h^{-1})
Y_{1sol}	0.4	Active mass yield (m/m)
Y_{1str}	0.4	Active mass yield (m/m)
Y_2	0.15	Inert mass yield (m/m)

Note: All time units are in hours (after Tzeng 1992).

tions; $\mathbf{K}$ = stoichiometric and yield coefficient matrix; $\boldsymbol{\mu}_i$ is the biomass growth and decay rate vector, which is a function of the states; $\mathbf{F}_i$ is mass input vector in the liquid phase; and $\mathbf{TR}_i$ = mass transfer rate vector, which also is a function of the process states. The elements of $\mathbf{K}$ are defined and described in Tables 2 and 3. Tables 2 and 3 also shows the values of the stoichiometric and yield coefficients used in the estimator. It should be noted that the $\mathbf{K}$ matrix is known and constant.

ESTIMATION METHODOLOGIES

Real-Time Estimator Schematic

The real-time estimator consists of three major parts: fuzzy estimation of influent substrate and effluent TSS concentrations, process state estimation, and model parameter estimation. Because influent substrate and TSS usually are not measured real time, their estimation is an important function. An asymptotic algorithm was employed for state estimation, and recursive least-squares algorithm was applied for parameter estimation.

Fig. 2 shows the overall structure of the estimator. The estimator was incorporated into an existing single-substrate, HPO-AS simulation program (Stenstrom et al. 1989) and runs in parallel with the simulator. The estimator and simulator share the same gas phase model (O_2 feed, pressure, etc). The estimator starts with the fuzzy estimation of influent substrate and effluent TSS. These estimates along with the DO measurements, initial conditions, and other measurements initiate both the estimator and the simulator. The estimated substrate and biomass concentrations are used for estimating biomass growth and decay rates. Finally, the estimated states and parameters are displayed graphically to the operator. To illustrate the success of the estimator, we compare these estimates with

both plant data and simulation results from the more advanced structured model.

Asymptotic and Recursive Least-Squares Algorithms

An asymptotic observer algorithm (Dochain et al. 1992) was employed to estimate biomass and substrate concentrations in each aeration stage using DO measurement only. It partitions (2) into two parts (a and b), which correspond to the measured and unmeasured states, respectively. The partitioned states are combined linearly to eliminate the nonlinear vector $\boldsymbol{\mu}_i$ by introducing an auxiliary state vector $\mathbf{Z}_i$, as follows:

$$\frac{d\hat{\mathbf{Z}}_i}{dt} = -D_i\hat{\mathbf{Z}}_i + \mathbf{A}_0(\mathbf{F}_{i,a} + \mathbf{TR}_{i,a}) + (\mathbf{F}_{i,b} + \mathbf{TR}_{i,b}) \quad (4)$$

$$\hat{\boldsymbol{\xi}}_{i,b} = \mathbf{A}_2^{-1}(\hat{\mathbf{Z}}_i - \mathbf{A}_1\hat{\boldsymbol{\xi}}_{i,a}) \quad (5)$$

where $\mathbf{A}_1$, $\mathbf{A}_2$, and $\mathbf{A}_0$ = coefficient matrices and can be obtained from the $\mathbf{K}$ matrix; $\hat{\boldsymbol{\xi}}_{i,a}$ and $\hat{\boldsymbol{\xi}}_{i,b}$ = measured and estimated states, respectively.

This algorithm allows the partial use of (2) to perform the estimation. In this case the first three states in (2) were used and partitioned. Because not all the values in the $\mathbf{A}_2$ matrix are known, it cannot be inverted. To overcome this problem, we used the estimated $\hat{\boldsymbol{\mu}}_d$ in the parameter estimator at time $t - 1$ to approximate $\boldsymbol{\mu}_d$ in the $\boldsymbol{\mu}_i$ vector of (2) (time increments generally were 1 hr). Therefore, the coefficient matrices in (4) and (5) have the following formats:

$$\mathbf{A}_0 = \mathbf{A}_1 = -\frac{1}{\mathbf{K}_1} \begin{bmatrix} \mathbf{K}_3 \\ \mathbf{K}_5 \end{bmatrix} \quad \text{and} \quad \mathbf{A}_2 = \mathbf{I} \quad (6a,b)$$

$$\hat{\boldsymbol{\xi}}_{i,a} = \text{DO}_i \quad \text{and} \quad \hat{\boldsymbol{\xi}}_{i,b} = [\hat{S} \quad \hat{X}]_i^T \quad (7a,b)$$

We used a recursive least-squares algorithm (Young 1984; Dochain and Bastin 1990) for biomass growth and decay rate estimation. This algorithm is a recursive least-squares algorithm obtained by applying a linear regression technique. Based on (2) the algorithm can be written as

$$\begin{aligned} \hat{\boldsymbol{\mu}}_{i,t+1}^m &= \hat{\boldsymbol{\mu}}_{i,t}^m + T\Gamma_{i,t}\mathbf{K}\mathbf{H}_{i,t}(\hat{\boldsymbol{\xi}}_t)\{\hat{\boldsymbol{\xi}}_{t+1} - \hat{\boldsymbol{\xi}}_t \\ &\quad - T[\mathbf{K}\mathbf{H}_{i,t}(\hat{\boldsymbol{\xi}}_t)\hat{\boldsymbol{\mu}}_{i,t}^m - D\hat{\boldsymbol{\xi}}_t + \mathbf{TR}_i + \mathbf{F}_i]\} \end{aligned} \quad (8)$$

$$\begin{aligned} \Gamma_{i,t+1} &= \frac{\Gamma_{i,t}}{\lambda} \{\mathbf{I} - T^2\mathbf{H}_{i,t}^T(\hat{\boldsymbol{\xi}}_t)\mathbf{K}^T[\lambda\mathbf{I} \\ &\quad + T^2\mathbf{K}\mathbf{H}_{i,t}(\hat{\boldsymbol{\xi}}_t)\Gamma_{i,t}\mathbf{H}_{i,t}^T(\hat{\boldsymbol{\xi}}_t)\mathbf{K}^T]^{-1}\mathbf{K}\mathbf{H}_{i,t}(\hat{\boldsymbol{\xi}}_t)\Gamma_{i,t}\} \end{aligned} \quad (9)$$

where $\hat{\boldsymbol{\mu}}_{i,t+1}^m$ and $\hat{\boldsymbol{\mu}}_{i,t}^m$ = estimated maximum growth and decay rates at times $t + 1$ and t for the i th stage, respectively; $\mathbf{K}$ = coefficient matrix; T and $\mathbf{I}$ = integration time step and identity matrix, respectively; $\Gamma_{i,t}$ = auxiliary matrix at time t , which is updated every integration step; λ = forgetting factor, which controls the stability and the tracking properties of the estimator; and $\mathbf{H}_{i,t}$ = a matrix that reflects our knowledge of the biokinetics.

To compensate for the unmeasured substrate and biomass concentrations in the aeration tanks, the estimated substrate and biomass concentrations along with the DO measurement in the parameter estimator were used. The first two states in (2) were used to construct the estimator. The matrices in (8) and (9) have the following values:

$$\mathbf{K} = \begin{bmatrix} \mathbf{K}_1 & \mathbf{K}_2 \\ \mathbf{K}_3 & \mathbf{K}_4 \end{bmatrix} \quad (10)$$

$$\hat{\boldsymbol{\mu}}_{i,t} = \hat{\boldsymbol{\mu}}_{i,t}^m\mathbf{H}_{i,t} = \begin{bmatrix} \boldsymbol{\mu}_{\max,i} & 0 \\ 0 & \mathbf{K}_{d,i} \end{bmatrix} \mathbf{H}_{i,t} \quad (11)$$

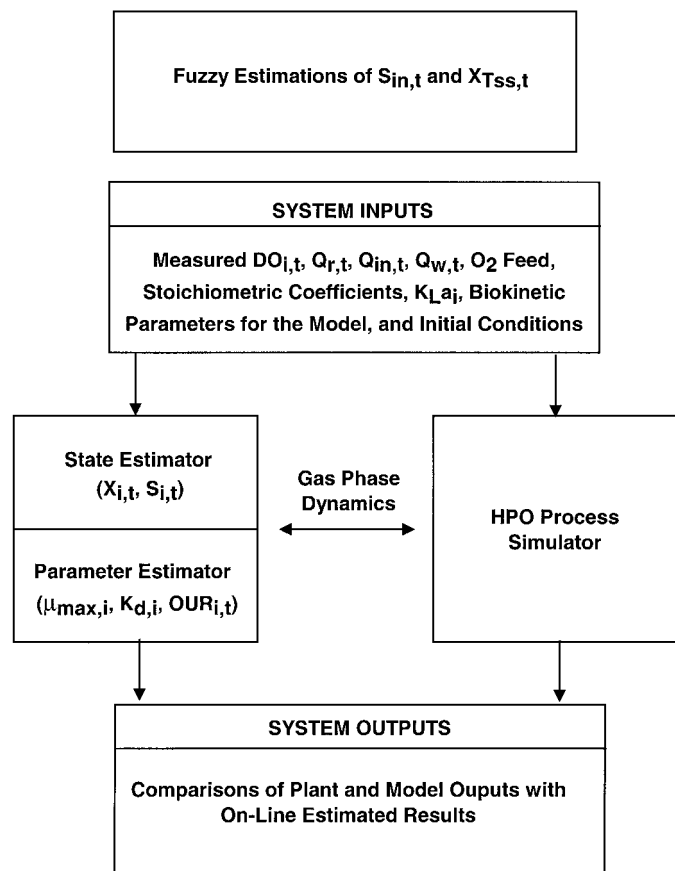


FIG. 2. On-Line Estimator Schematic

$$\mathbf{H}_{i,t} = \begin{bmatrix} X_{i,t} \left(\frac{S_{i,t}}{K_{S_i} + S_{i,t}} \right) \left(\frac{DO_{i,t}}{K_{DO} + DO_{i,t}} \right) & 0 \\ 0 & X_{i,t} \left(\frac{DO_{i,t}}{K_{DO_i} + DO_{i,t}} \right) \end{bmatrix} \quad (12)$$

where $\mu_{\max_i}$ and K_{d_i} = *i*th stage biomass maximum growth and decay rates, respectively; and K_{S_i} and K_{DO_i} = half-saturation coefficients for substrate and DO concentrations, respectively. Incorporating Monod kinetics into $\mathbf{H}_{i,t}$ is very important to obtain good performance of the estimator. The biomass growth is subject to the limitation of biomass, substrate, and DO concentrations. The endogenous respiration is also limited by the DO concentration. This is especially true in stage 1 when the process is operated in the reaeration mode, where less substrate is present and endogenous respiration becomes significant.

Fuzzy Logic

Because influent substrate and recycle biomass concentrations are not measured in real time, estimating these two variables becomes crucially important to support the estimator. We used a fuzzy logic algorithm to make these predictions. Three sets of fuzzy rules have been developed and implemented into the estimator.

The first set consists of 14 fuzzy rules used for estimating the influent substrate concentration. For municipal wastewater, the substrate-loading pattern is correlated strongly to living habits and the characteristics of the sewer system in the service area, such as the length of the sewer and travel time of flow. The peak substrate loading at treatment plant usually lags the peak substrate discharge because there is a travel lag between the plant and the residential area. The 14 rules were formulated based on the observed loading pattern corresponding to the time.

The other set of 30 fuzzy rules estimates the effluent TSS concentration. We considered the two operational parameters that may have significant influence on effluent TSS: influent flow rate and DSVI. Olsson and Stephenson (1985) have correlated high influent flow rate and high effluent TSS concentration. The DSVI represents the sludge-settling characteristics (Koopman and Cadée 1983; Hultman et al. 1991) where high DSVI value usually indicates poor sludge-settling rates. Such conditions usually result in high effluent TSS concentration.

Fig. 3 shows the fuzzy rule relations among influent flow rate, DSVI, and effluent TSS. The horizontal axis is the influent flow rate that changes from very low to extremely high. Similarly, DSVI changes from very low to very high on the vertical axis. Each box inside the matrix represents the pre-

dicted effluent TSS. This matrix reflects the empirical knowledge among the three parameters: higher values of flow rate and DSVI correspond to higher effluent TSS concentration. It should be noted that the estimator only works on each horizontal axis period, because DSVI is not measured real time. Whenever the estimator receives a new DSVI, the estimator creates new values along the vertical axis corresponding to the current DSVI. In this way, the off-line measurements are used for real-time estimation.

After obtaining the effluent TSS, the recycle biomass concentration can be obtained by making a mass balance around the clarifier as follows:

$$X_{R,t} = \frac{(Q_{in,t} + Q_{R,t})X_{4,t} - (Q_{in,t} - Q_{W,t})X_{TSS,t}}{Q_{R,t} + Q_{W,t}} \quad (13)$$

where $Q_{in,t}$, $Q_{R,t}$, and $Q_{W,t}$ = influent, recycle, and waste sludge flow rates at time *t*, respectively; $X_{4,t}$ is the stage 4 biomass concentration, estimated from the state estimator; and $X_{TSS,t}$ = effluent TSS concentration obtained using the fuzzy logic algorithm described earlier.

For unusual conditions, such as sludge bulking or wet-weather flow, fuzzy TSS estimation is needed for the abnormal state. A third set of rules (10 rules) was developed to allow the fuzzy estimate to adapt to abnormal conditions. When predefined flow rates or DSVI limits are exceeded, the rule set is invoked and the scaling factors are obtained. These scaling factors then are applied to the support sets for flow rate, DSVI, and effluent TSS concentration. In this way, the scale and shape of the membership functions are changed, to establish a new working estimation state.

RESULTS AND DISCUSSION

Plant Data, Model, and Estimator Inputs

The pilot plant data shown in Table 4 were used to confirm and evaluate the performance of the estimator. These data are average values and do not represent the process dynamics, especially when the process is under transient conditions. To create dynamic conditions, a structured, dynamic HPO-AS process model (Stenstrom 1990) was used to simulate dynamic conditions. The results of this model were compared with the estimator's predictions for the same input conditions.

The structured dynamic HPO-AS process model is much more advanced than the simple, single-substrate model used in the estimator. The structured model was calibrated and val-

TABLE 4. Pilot Plant Data

Parameter (1)	Value (2)	Parameter (3)	Value (4)
Liquid phase volume	7.84 m ³	Stage 1 OUR	63 mg O ₂ /L-h
Gas phase volume	1.13 m ³	Stage 2 OUR	96 mg O ₂ /L-h
Gas phase pressure	3 cm w.c.	Stage 3 OUR	48 mg O ₂ /L-h
Average flow rate	6 m ³ /h	Stage 4 OUR	41 mg O ₂ /L-h
Recycling ratio	52%	Average stage 1 DO	7.6 mg/L
Sludge waste rate	0.48 m ³ /h	Average stage 2 DO	5.2 mg/L
Net yield	0.6–0.85	Average stage 3 DO	5.5 mg/L
MLSS	1,346 mg/L	Average stage 4 DO	5.0 mg/L
MLVSS	1,171 mg/L	Stage 1 O ₂ purity	93.7%
Recycle sludge conc.	3,577 mg/L	Stage 2 O ₂ purity	82.8%
Recycle sludge conc. (VSS)	3,112 mg/L	Stage 3 O ₂ purity	71.0%
Influent total BOD ₅	88 mg/L	Stage 4 O ₂ purity	65.6%
Influent soluble BOD ₅	39 mg/L	O ₂ flow in	0.62 m ³ /h
Influent total chemical oxygen demand	217 mg/L	O ₂ flow out	0.07 m ³ /h

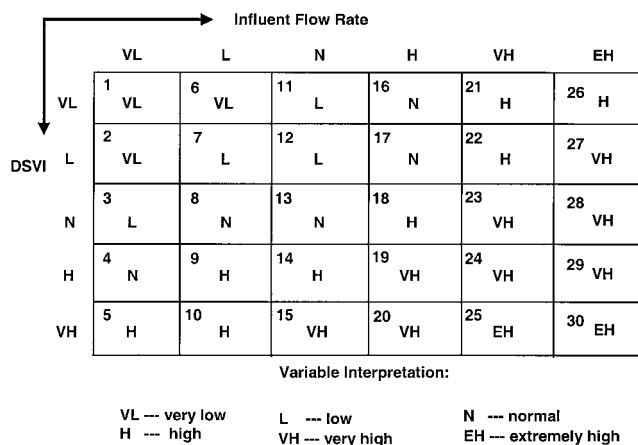


FIG. 3. Fuzzy Rule Relation Matrix for Flow Rate and DSVI versus Effluent TSS

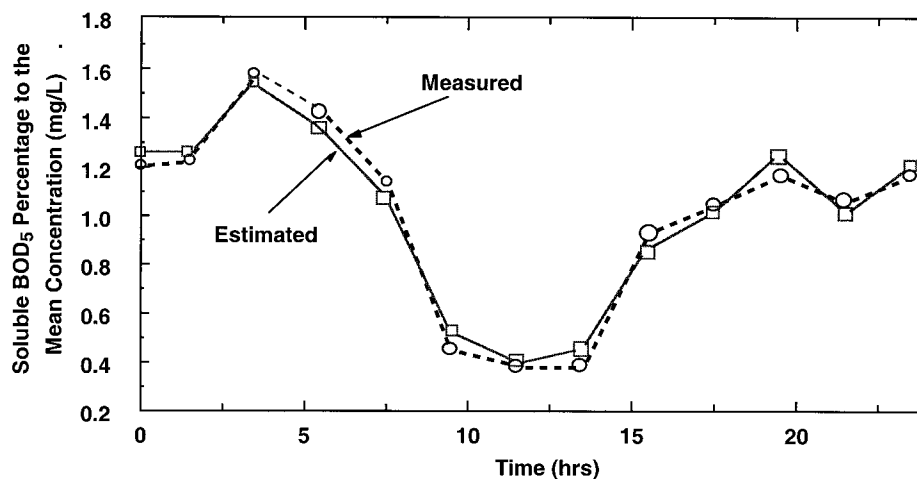


FIG. 4. Comparison of Fuzzy Estimated and Measured Soluble BOD₅ for a Typical Day

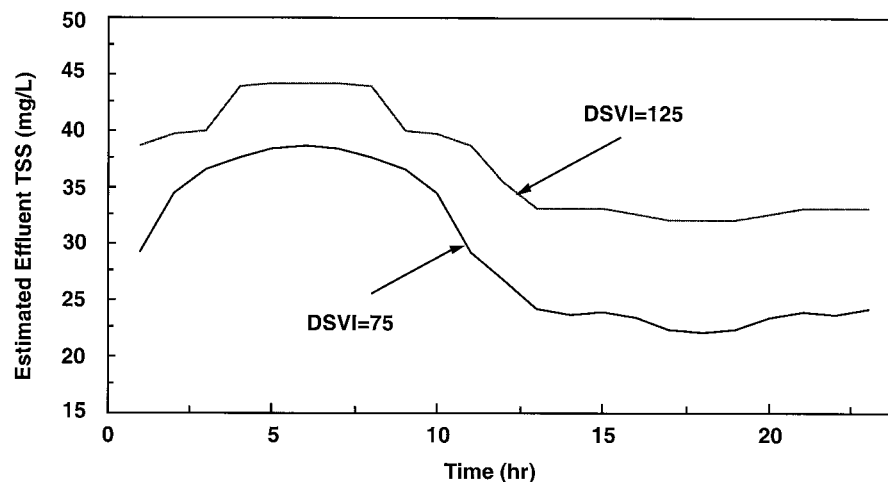


FIG. 5. Fuzzy Estimation of Effluent TSS with Different DSVIs

idated by Tzeng (1992) and Yuan (1994) based on this pilot plant data and data from a full-scale HPO-AS plant. Good agreement between model predictions and plant data was obtained. Table 3 summarizes the biokinetic parameters used in the model. The stoichiometric and yield coefficients used in the estimator are shown in Table 2.

The same input data were used for both simulator and estimator inputs. The pilot plant had four stages in series and was operated in reaeration mode (primary effluent fed to the second stage). The plant used surface aerators. The diurnally varying influent flow rate was approximated by a sinusoid. The DO measurements in each stage were assumed in a sinusoidal pattern because of the flow rate pattern, with a mean equaling to the measured average plant data (Table 4), and were corrupted with a white noise.

Fuzzy Estimation of Influent Substrate and Effluent TSS

Fourteen fuzzy rules were used to estimate the influent substrate pattern. All rules use time as the rule antecedent. Fig. 4 shows the estimated BOD₅ for a typical day. The number and defined ranges (support sets) of rules will vary from plant to plant, depending on the characteristics of the sewer system and the service area. One should create these rules based on the observed substrate-loading pattern to obtain estimates for different sites.

Fig. 5 shows the estimated effluent TSS using the fuzzy rules described in the Fuzzy Logic section. As the influent flow rate undergoes a periodic change, the TSS changes accord-

ingly; more TSS is lost through the effluent when flow rate is high, and less TSS is present in the effluent when flow rate is low. Another important parameter that greatly influences effluent TSS is the DSVI value. Two DSVI conditions were simulated, as shown in the graph. Both DSVIs are partially within the "normal (N)" range, but one is partially in the "high (H)" range (DSVI = 125) and another is in the "low (L)" range (DSVI = 75). As expected, the results show increased TSS with increased DSVI. This is exactly the knowledge that is implemented in the rule matrix (Fig. 3).

A crucial step in making this estimation is the design of the fuzzy relation matrix (Fig. 3), which should represent the best empirical knowledge among the three parameters. Correctness of the rules determines the overall trend of the estimates (the shape of the curves shown in Fig. 5). The shape of the curve will be completely different if the relation matrix is defined differently from that in Fig. 3. The number of the rules and appropriately defined ranges of the linguistic variables, such as high, low, normal, are responsible for making an accurate prediction. In principle, increasing the number of rules can make the prediction more accurate (smoother curves) but also greatly increases the required knowledge and complexity of the estimation. To define properly the range of each linguistic variable, one should use trial and error to fine-tune the fuzzy estimator based on specific plant data.

State Estimation

To validate the asymptotic state estimator, the waste biomass flow rates and concentrations measured in the pilot plant

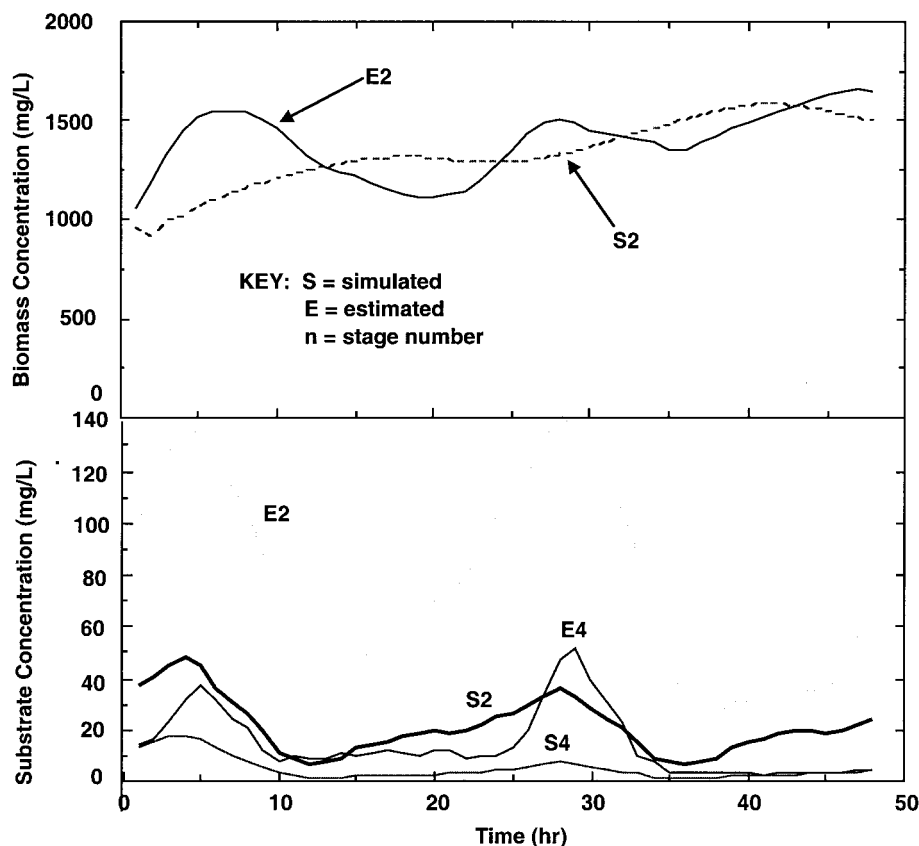


FIG. 6. Comparison of Estimated and Simulated Biomass and Substrate Concentrations

were used as inputs to the estimator. The convergence and stability of the estimator with a large range of initial values of the auxiliary states (Z 's) were evaluated. A sensitivity analysis of the Z_1 and Z_2 vectors was performed. Initial values of both vectors were varied for both biomass (± 500 mg/L) and substrate (± 50 mg/L) in all four stages. Even with this large range of initial estimates, the convergence was fast and stable for both states. The convergence for biomass estimation is faster than for substrate (approximately 2 h). The relative error between the plant measured and estimated biomass concentration was 6%. Substrate estimation accuracy rapidly improved over 3 h.

To evaluate the estimator's ability to predict reactor biomass and substrate concentrations, the estimator was used with influent flow rate and substrate concentration functions and measured DO concentrations. The influent flow rate and substrate concentration were approximated by sinusoidally varying functions with means equal to the plant data and with an amplitude equal to the expected variations in the plant data. White noise producing variations in the mean by approximately 10% also was added for realism. The estimated effluent substrate could not be compared with pilot plant effluent substrate concentrations, which were not measured. Instead, estimated effluent substrate was compared with simulations from the structured model. The estimated influent substrate and recycle biomass concentrations generally were used as inputs to the structured model. Fig. 6 shows the estimated and simulated biomass and substrate concentrations. The estimated biomass concentration in stage 2 tracks the simulated biomass well, except in the initial 10 h of the simulation (differences related to initial conditions and startup convergence). The same pattern of biomass concentration changes can be observed for the rest of the stages. The oscillation of biomass concentration is caused by the sinusoidally varying inputs.

Fig. 6 reveals significant differences between the simulated and estimated substrate concentrations in stage 2. The reason

for this is the difference between the structured and conventional models. The simulator uses the concept of stored substrate and mass; the estimator does not use this more advanced concept and is based on a simple, single-substrate model (Lawrence and McCarty 1970). In the structured model, particulate substrate is stored rapidly (or entrapped in bioflocs) and is biodegraded more slowly. This accounts for the rapid substrate uptake when the influent is contacted initially with the mixed liquor. This also causes the maximum oxygen demand to lag the maximum loading. The differences in predicted and estimated substrates is much less in stage 4, where the stored substrate is largely consumed and more endogenous conditions exist.

Conventional Monod kinetics have difficulty in simulating this phenomenon. The ability to simulate rapid substrate uptake and lagging OUR has been an important goal of structured model development (Jacquart et al. 1973; Busby and Andrews 1975; Stenstrom and Andrews 1979; Dold et al. 1980; Clift and Andrews 1981). The International Association on Water Quality AS model also has this ability (Henze et al. 1987). To include this ability in the estimator, structured model mechanisms must be used. These mechanisms can be included but would increase greatly the complexity of the algorithm and would be more time-consuming. If an estimator is to be used in situations where rapid substrate uptake rate is important, then an estimator based on the structured model should be used.

Parameter Estimation

Parameter estimation is performed using (8)–(12). The maximum specific growth rate, decay rate, and OUR were estimated. The biomass, substrate, DO concentrations in each stage, and the fuzzy estimated influent flow rate and recycle biomass concentrations are used in the parameter estimation.

Fig. 7 shows the estimated specific growth rates and simu-

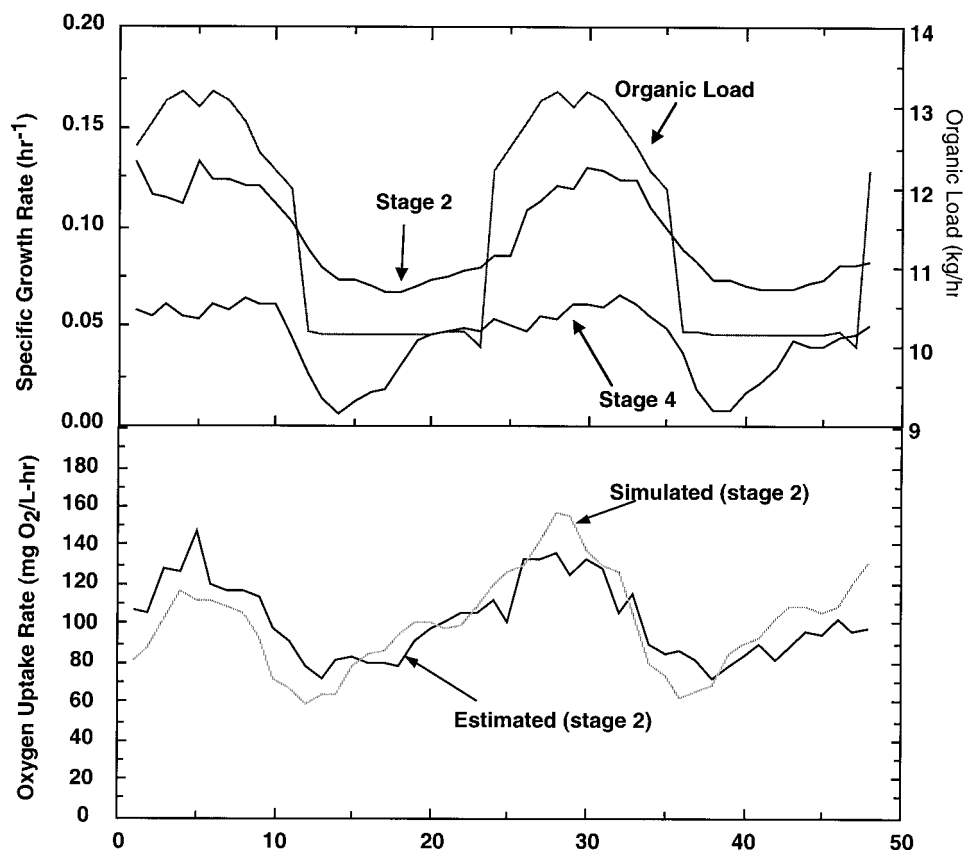


FIG. 7. Estimated Results for Biomass Specific Growth Rate and OURs in Stages 2 and 4

TABLE 5. Summarized Simulation Results and Comparison among Estimated, Simulated, and Measured Data

Parameter (1)	Stage 1 (2)	Stage 2 (3)	Stage 3 (4)	Stage 4 (5)
$K_L a$ used (h^{-1})	2.0	5.0	3.0	3.0
K_s used (mg/L)	1.0	20.0	20.0	20.0
K_{DO} used (mg/L)	0.5	0.5	0.5	0.5
Average biomass maximum growth rate (h^{-1})	0.013	0.133	0.091	0.081
Average biomass decay rate (h^{-1})	0.010	-0.005	-0.003	0.006
Average biomass specific growth rate (h^{-1})	0.012	0.095	0.055	0.043
Average estimated OUR (mg O_2 /L-h)	82.1	100.9	54.2	51.5
Average simulated OUR (mg O_2 /L-h)	80.6	100.2	63.4	45.7
Average measured OUR (mg O_2 /L-h)	63.0	96.0	48.0	41.0
Average estimated biomass concentration (MLVSS, mg/L)	3,447	1,042	1,055	1,070
Average simulated biomass concentration (MLSS, mg/L)	3,597	1,334	1,333	1,327
Average measured biomass concentration (MLVSS, mg/L)	3,112	1,171	1,171	1,171

lated and estimated OURs in stage 2 over a 48-h simulation period. As evidenced in the upper graph of Fig. 7, the dynamics of biomass specific growth rates in stages 2 and 4 are in phase with the organic loading (dotted line); the higher the organic loading, the larger the growth rate. The average growth rate in stage 2 is higher than that of stage 4 because of the higher substrate concentration in stage 2 (primary effluent is fed into stage 2).

The estimated OUR in stage 2 tracks the OUR (lower graph)

from the structured model well and the organic loading pattern (comparing lower and upper graphs). This good tracking ability is particularly caused by the correctly estimated specific growth rate. However, a time lag (2–3 h) between the estimated and simulated OURs was observed in stages 2, 3, and 4 (not shown). This is again caused from the use of different models in constructing the simulator and estimator. A simple but empirical alternative method to overcome this problem is to use a delay time ($t - \tau$, where τ is the delay time) in (8) and (9).

Table 5 summarizes the simulation results. The negative values of biomass decay rates in stages 2 and 3 indicate large biomass growth. The simulated and plant-measured data for OUR and biomass concentration reasonably agree. The high simulated and estimated OURs in stage 1 are caused by high biomass concentrations.

State Estimation under Abnormal Operation

Process upsets, such as sludge bulking and high flow caused by storm water usually are experienced in wastewater-treatment plants. Under these abnormal operating conditions, it is very important for the operator to be aware of the important process states, in order to change the operation mode to accommodate the upsets. A process estimator may be useful in these circumstances. In this section we present two abnormal cases (high hydraulic loading and high DSVI value) using the real-time estimator to predict biomass concentrations in each stage. The simulation inputs are the same as those previously presented in the state and parameter estimation sections, except influent flow rate and DSVI. They are changed to simulate the process upsets.

The upper graph of Fig. 8 shows the estimated biomass concentrations of stages 1 and 4 when high influent flow rate occurs at 36 h. Because of the higher flow rate, the biomass in the clarifier (caused by high TSS in effluent) washes out,

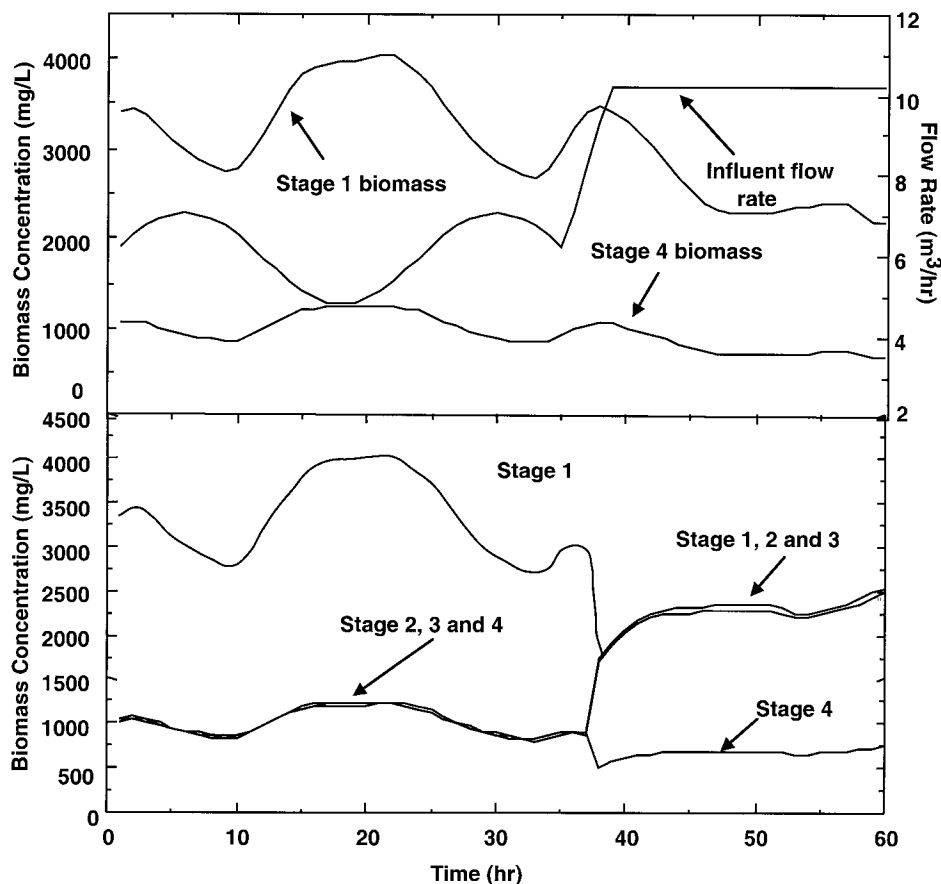


FIG. 8. Estimated Biomass Concentrations under Hydraulic Shock Loading

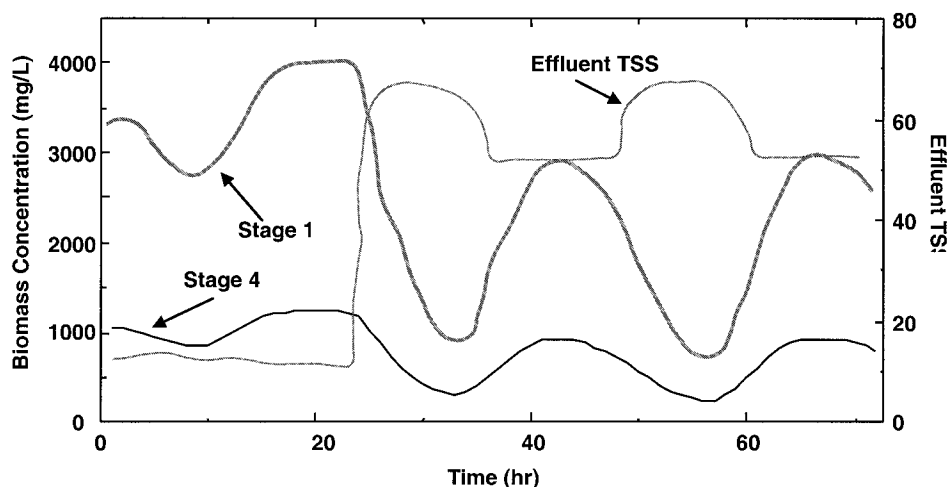


FIG. 9. Estimated Biomass Concentrations for Stages 1 and 4 with High DSVI

increasing effluent TSS and reducing recycle biomass concentration. The average stage 4 biomass concentrations before and after the occurrence of high flow rate are 1,033 and 803 mg/L, respectively. A 22% reduction of biomass concentration is estimated. Similar reductions are observed for the other stages. If an even higher flow rate was used, further reductions of biomass would be expected in each stage. This may cause treatment process failure.

An alternative strategy dealing with hydraulic shock loading is step feed (Torpey 1948). In using this strategy, primary effluent is fed directly to the later or last stage of the process. In this way, the biomass in stages 1, 2, and 3 can be conserved. The lower graph of Fig. 8 shows the results of applying this strategy. When the feed point is changed from stage 2 to stage 4, the fuzzy logic sets the biomass concentrations in stages 2

and 3 equal to the stage 1 biomass concentration. As compared with the reaeration mode (upper graph), this strategy can increase the total sludge masses from 8.0 to 11.7 kg in all four stages, which is a 46% increase over reaeration operation. This change allows the plant to retain its biomass inventory and greatly helps the plant recover after the high hydraulic loading ends.

High DSVI implies poor biomass settling ability or sludge bulking and large amounts of biomass may be lost through the effluent. To properly apply the estimator in this situation, the estimation of effluent TSS becomes critically important. As indicated earlier, a separate set of rules was used to readjust the rule sets that are used in normal operation. Fig. 9 shows the performance of the adaptive rules for the TSS estimation (dotted line) when high DSVI occurs at time equal to 24 h.

The effluent TSS increases from 13.5 to 58 mg/L. This results in reduction of biomass concentration for all stages. As shown in Fig. 9, the average reduction of biomass concentration for each stage is approximately 40%, which may jeopardize the treatment process and cause process failure or permit violations.

CONCLUSIONS

A fuzzy logic-supported real-time state and parameter estimator were developed for the HPO-AS process. For state estimation, an asymptotic algorithm was employed. A recursive least-squares method was used for parameter estimation. To compensate for unknown influent substrate and recycle biomass concentrations, a fuzzy logic algorithm was used to predict these unmeasured process variables. The estimator can predict the unmeasured substrate and biomass concentrations in each stage using DO measurements alone. It also can estimate other important parameters, such as maximum, specific growth and decay rates, and OURs. Knowledge of these stimulated states and parameters can greatly help the operator in making decisions, as well as providing quantitative supports for advanced process control.

The convergence of the algorithms is fast and stable even with a large range of initial values and noisy DO measurements. The estimated process parameters agree reasonably well with both plant and structured model simulations. The estimator can closely track the simulated OURs under transient conditions. It also can be used for state and parameter estimation when certain types of process upsets occur.

A crucial step in developing such an estimator is fuzzy estimation of effluent TSSs, because TSS has great influence on the biomass inventory. Rule accuracy is the key for the whole estimation. Influent flow rate and DSVI are the two major parameters influencing TSS. The estimator utilizes the off-line measurement (DSVI) for real-time estimation purposes. The fuzzy rules developed for the estimator also have the ability to detect process upsets, such as high hydraulic loading (storm water) and high DSVI, and accommodate the upsets by invoking modified sets of rules.

APPENDIX. REFERENCES

- Aborhey, S., and Williamson, D. (1978). "State and parameter estimation of microbial growth processes." *Automatica*, 14, 493–498.
- Ayesa, E., Florez, J., Garcia-Heras, J. L., and Larrea, L. (1991). "State and coefficients estimation for the activated sludge process using modified Kalman filter algorithm." *Water Sci. and Technol.*, 24(6), 235–247.
- Busby, J. B., and Andrews, J. F. (1975). "Dynamic modeling and control strategies for the activated sludge process." *J. Water Pollution Control Fed.*, 47, 1055–1080.
- Cliff, R. C., and Andrews, J. F. (1981). "Predicting the dynamics of oxygen utilization in the activated sludge process." *J. Water Pollution Control Fed.*, 53(7), 1219–1232.
- Cliff, R. C., Andrews, J. F., and Garrett, M. T. (1983). "Minimization of energy consumption in the oxygen activated sludge process." *IAWPRC Specialized Conf. on Energy Savings in Water Pollution Control*, International Association for Water Pollution Research, Paris.
- Dochain, D. (1992). "Adaptive control algorithm for nonminimum phase nonlinear bioreactor." *Comp. Chem. Engrg.*, 16(5), 449–462.
- Dochain, D., and Bastin, G. (1990). "Adaptive control of fedbatch bioreactors." *Chem. Engrg. Communications*, 87, 67–85.
- Dochain, D., Perrier, M., and Ydstie, B. E. (1992). "Asymptotic observers for stirred tank reactors." *Chemical Engrg. Sci.*, 47(15/16), 4167–4177.
- Dold, P. L., Ekama, G. A., and Marais, G. v. R. (1980). "A general model for the activated sludge process." *Prog. Water Technol.*, 12(6), 47.
- Henze, M., Grady, C. P. L., Guger, W., Marais, G. v. R., and Matuso, T. (1987). "Activated sludge model no. 1." *IAWPRC Scientific and Tech. Rep. No. 1*, International Association on Water Quality, London.
- Holmberg, A. (1986). "Adaptive dissolved oxygen control and on-line estimation of oxygen transfer and respiration rates." *AIChE Annu. Meeting*, American Institute of Chemical Engineers, New York.
- Holmberg, U., and Olsson, G. (1985). "Simultaneous on-line estimation of oxygen transfer rate and respiration rate." *IFAC modeling and control of biotechnical processes*, International Federation of Automatic Control, Noordwijkerhout, 205–209.
- Hultman, B., Lowen, M., Karlsson, U., Li, P. H., and Molina, L. (1991). "Prediction of activated sludge sedimentation based on sludge indices." *Water Sci. and Technol.*, 24(7), 33–42.
- Jacquart, J. C., Lefort, D., and Lovel, J. M. (1973). "An attempt to take account of biological storage in the mathematical analysis of activated sludge behavior." *Advances in water pollution research*, S. H. Jenkins, Ed., Pergamon, Tarrytown, New York.
- Koopman, B., and Cadee, K. (1983). "Prediction of thickening capacity using diluted sludge volume index." *Water Res.*, 17(10), 1427–1431.
- Lawrence, A. W., and McCarty, P. L. (1970). "Unified basis for biological treatment design and operation." *J. Sani. Engrg. Div., ASCE*, 96(3), 757–778.
- Ljung, L. (1979). "Asymptotic behavior of the extended Kalman filter as a parameter estimator for linear systems." *IEEE Trans. Aut.*, 24, 36–50.
- Marsili-Libeli, S. (1990). "Adaptive estimation of bioactivities in the activated sludge process." *Proc. IEEE*, 137(6), 349–356.
- Meditch, J. S., and Hostetter, G. H. (1974). "Observer for system with unknown and inaccessible inputs." *Int. J. Control*, 19(3), 473–480.
- Olsson, G., and Stephenson, J. P. (1985). "The propagation of hydraulic disturbances and flow rate reconstruction in activated sludge plants." *Envir. Technol. Letters*, 6, 536–545.
- Pedrycz, W. (1993). *Fuzzy control and fuzzy systems*. Wiley, New York.
- Ramirez, W. F. (1987). "Optimal state and parameter identification. An application to batch fermentation." *Chemical Engrg. Sci.*, 42(11), 2749–2756.
- Samstag, R. (1989). "West Point treatment plant secondary treatment facilities—high purity oxygen design test facility draft report." Metropolitan of Seattle, Seattle.
- Shimizu, H., and Takamatsu, T. (1980). "An algorithm approach to constructing the on-line estimation system for the specific growth rate." *Biotechnol. and Bioengr.*, 33, 354–364.
- Stenstrom, M. K. (1990). "Westpoint treatment plant oxygen process modeling." *UCLA ENG Rep. No. 90-17*, University of California, Los Angeles, Calif.
- Stenstrom, M. K., and Andrews, J. F. (1979). "Real-time control of the activated sludge process." *J. Envir. Engrg. Div., ASCE*, 105, 245–260.
- Stenstrom, M. K., Kido, W., Shanks, R. F., and Mulkerin, M. (1989). "Estimating oxygen transfer capacity of a full-scale pure oxygen activated sludge plant." *J. Water Pollution Control Fed.*, 61(2), 38–50.
- Stephanopoulos, G., and San, K. Y. (1984). "Studies on on-line bioreactor identification." *Biotechnol. and Bioengr.*, 26, 1176–1188.
- Torpey, W. N. (1948). "Practical results of step aeration." *Sewage Works J.*, 20, 781–788.
- Tzeng, C.-J. (1992). "Advanced dynamic modeling of the high purity oxygen activated sludge process," PhD dissertation, Civil and Environmental Engineering Department, University of California, Los Angeles, Calif.
- Yin, M. T. (1995). "Developing a decision support system for operation and control of the high purity oxygen activated sludge process," PhD dissertation, Civil and Environmental Engineering Department, University of California, Los Angeles, Calif.
- Yin, M. T., Yuan, W., and Stenstrom, M. K. (1994). "A decision support system for wastewater treatment plant operations." *Proc., 1st ASCE Congr. on computing in Civ. Engrg.*, ASCE, Reston, VA.
- Young, P. (1984). *Recursive estimation and time-series analysis, an introduction*. Springer, New York.
- Yuan, W. (1994). "Dynamic modeling and expert systems for the activated sludge process," PhD dissertation, Civil and Environmental Engineering Department, University of California, Los Angeles, Calif.
- Yuan, W., Okrent, D., and Stenstrom, M. K. (1993). "Model calibration for the high-purity oxygen activated sludge process—algorithm development and evaluation." *Water Sci. and Technol.*, 28(11–12), 3–171.