APPLICATION OF HIGH DIMENSIONAL MODEL REPRESENTATIONS (HDMR) AND INTERPOLATION TO OPTIMIZE NUTRIENT REMOVAL UNDER UNCERTAINTY

F. Jiang[1] and M.B. Beck[2]

AUTHORS: Greduate stantent[/1],Professor[/2], Warnell School of Forestry and Natural Resources, University of Georgia. Athens, Georgia 30602.
REFERENCE: *Proceedings of the 2007 Georgia Water Resources Conference*, held March 27–29, 2007, at the University of Georgia.

Abstract. In order to circumvent difficulty caused by model uncertainty, high dimensional model representation (HDMR) and interpolation were employed to approximate model outputs with different combinations of manipulated variables and model parameters. The results showed that HDMR and interpolation could be successfully applied to optimize nutrient removal under uncertainty.

INTRODUCTION

Uncertainty is inevitable when using mathematical model to simulate real process that subject to both anthropentic and natural disturbance. If uncertainty associated with a model is neglected, an optimal solution induced from this model may be far from optimal when applied to reality. Some control strategies have been proposed to take model uncertainty into account. The most popular method is to perform Mante Carlo simulations. However, with so many parameters in ASM, the computation cost is prohibitively high. On the other hand, high dimensional model representation (HDMR) is a fast algorithm that can circumvent the apparent exponential difficulty of high-dimensional mapping problem. It has been successfully applied in atmospheric chemistry modeling (Li et al, 2000). Interpolation is an algorithm used to estimate function values between data points. Here we adopted HDMR and interpolation to approximate model outputs under different combinations of manipulated variables and model parameters.

PROCESS AND MODEL DESCRIPTION

The basic process investigated here was an activated sludge process (AS) to which an anaerobic tank and anoxic tank were added to enhance nutrient removal (Figure 1). The dimensions of the units were listed in Table 1.

Table 1. Main dimensions of units

	Construction	Item	Dimension
Bioreactor	Anaerobic tank	Volume	884 m^3
	Anoxic tank	Volume	1768 m^3
	Aerobic tank	Volume	6375 m^3
Clarifier		Area	500 m^2
		Height	4m

Table 2. Flux-based average influent characterization

Parameters	Unit	Dimension
Total COD	mgCOD/l	404
Ortho-P	mgP/l	1.94
Total phosphorus	mgP/l	5.43
NH_4-N	mgN/l	12.72
TSS	mgSS/l	261

INFLUENT LOAD

The influent data were collected in Athens No. 2 wastewater treatment plant (WWTP) of Georgia in 1998 (Liu, 2000; Liu and Beck, 2000). Generally, the influent quality can be classified as medium. Its main characteristics were listed in Table 2. The influent COD was fractioned into its components as in ASM No. 2d (Henze et al, 1999). The range of components listed in Table 3 was induced from the results of previous model calibration.

MODEL AND SIMULATION DESCRIPTION

ASM 2d was selected as it includes both nitrogen and phosphorus removal. The model was calibrated with the data collected in Athens No. 2 WWTP in 1998. Totally, there were 41 parameters in this model. We selected 16 parameters according to their sensitivity, and included the fractionation of influent COD in the framework. Thus, totally 24 parameters were adjusted in each simulation. All simulations were performed on WEST simulation platform (Hemmis nv, Kortrijk, Belgium). The implementation of the process was shown in Figure 2.

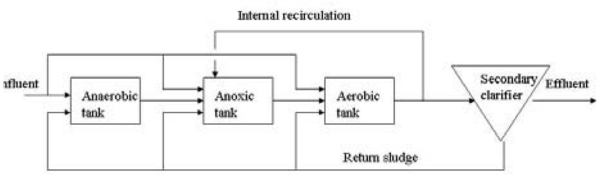

Figure 1. Flowchart of the process

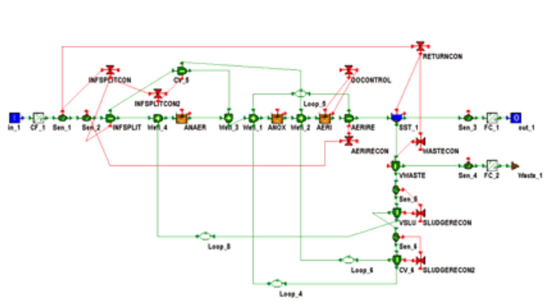

Figure 2. Implementation of process in WEST

Table 3. Characterization of influent COD (Ratio of total COD)

Component	Definition	Range (%)
S_I	Inert soluble organic material	4.33~5.75
S_A	Fermentation products	4.59~6.31
S_F	Fermentable, readily biodegradable organic substrates	8.08~11.81
X_I	Inert particulate organic material	25.4~28.7
X_S	Slowly biodegradable substrates	36.5~51.8
X_H	Heterotrophic organisms	5.13~9.81
X_{AUT}	Nitrifying organisms	0.39~0.61
X_{PAO}	Phosphate-accumulating organisms	0.20~0.34
X_{PP}	Poly-phosphate	0
X_{PHA}	A cell internal storage product of PAO	0.08~0.15

CONTROL STRATEGY DESCRIPTION

The control strategy used here was essentially stochastic optimization of manipulated variables under model uncertainty. The values of manipulated variables were chosen such that the expected objective function assumed a minimum (Infanger, 1993):

$$z = \min E\, f(x, \omega)$$
$$\text{s/t } x \in C = \cap_{\omega \in \Omega} C_\omega$$

Where x – manipulated variables;
ω - model parameters;
Ω - set of possible realizations of ω;
f(x, ω) – objective function;
C – set of feasible solutions.

The optimal solution represented the realistic solution of the stochastic optimization problem.

$$X^* \in \arg\min \{E f(x, \omega) \mid x \in \cap_{\omega \in \Omega} C^{\omega}\}$$

Table 4. Range of model parameters

Parameter	Definition	Unit	Range
μ_{PAO}	Maximum growth rate of PAO	d^{-1}	1.62~1.79
q_{PP}	Rate constant for storage of X_{PP}	$gX_{pp}g^{-1}X_{PAO}\,d^{-1}$	1.98~2.16
Y_{PHA}	PHA requirement for PP storage	$gCOD\,g^{-1}P$	0.06~0.16
K_{PS}	Saturation coefficient for phosphorus in storage of PP	gPm^{-3}	0.08~0.12
η_{NO3_HET}	Reduction factor for denitrification	-	0.33~0.49
K_{O2_AUT}	Saturation/inhibition coefficient for oxygen for X_{AUT}	gO_2m^{-3}	0.29~0.45
μ_{AUT}	Maximum growth rate of X_{AUT}	d^{-1}	0.87~1.04
b_H	Rate constant for lysis and decay	d^{-1}	0.33~0.45
b_{PP}	Rate for lysis of X_{PP}	d^{-1}	0.0019~0.10
K_{NH4_AUT}	Saturation coefficient for ammonia for X_{AUT}	gNm^{-3}	0.6~1.03
K_A	Saturation coefficient for acetate	$gCOD\,m^{-3}$	3~5
K_F	Saturation coefficient for growth on S_F	$gCOD\,m^{-3}$	3~5
μ_H	Maximum growth rate on substrate	d^{-1}	4.93~5.96
f_{XI}	Fraction of inert COD generated in biomass lysis	-	0.10~0.15
K_X	Saturation coefficient for particulate COD	$gCOD\,m^{-3}$	0.11~0.15
V_0	Settling velocity of sludge	m/d	617~916

For this research, 6 manipulated variables were employed, i.e. allocation of influent and return sludge among the three tanks, dissolved oxygen in aerobic tank, internal recirculation, outer recycle, and waste sludge. For each manipulated variable, we selected 5 values. Thus, totally we had 5^6 = 15,625 different combinations of manipulated variable values. Obviously, the simulations of all these combinations cannot be finished in a reasonable period of time. Alternatively, we used HDMR (Li et al, 2000) to approach model outputs as the following

$$f(x) \cong f_0 + \sum_{i=1}^{n} f_i(x_i) + \sum_{1 \le i \angle j \le n} f_{ij}(x_i,x_j)$$

Where

$f_0 = f(\overline{x})$

$f_i(x_i) = f(x_i, \overline{x}^{i}) - f_0$

$f_{ij}(x_i,x_j) = f(x_i,x_j, \overline{x}^{ij}) - f_i - f_j - f_0$

$(x_i, \overline{x}^{i}) = (\overline{x}_1, \ldots, \overline{x}_{i-1}, x_i, \overline{x}_{i+1}, \ldots, \overline{x}_n)$,

$(x_i,x_j, \overline{x}^{ij}) = (\overline{x}_1, \ldots, \overline{x}_{i-1}, x_i, \overline{x}_{i+1}, \ldots, \overline{x}_{j-1}, x_j, \overline{x}_{j+1}, \ldots, \overline{x}_n)$

The resultant computational effort to determine the expansion function will scale polynomically with n rather the traditional view of it being exponential (Saltelli, Chan, and Scott, 2000). By using HDMR we dramatically reduced the number of simulations needed. For example, only 265 simulations were needed to evaluate different combinations of 6 manipulated variable values.

In addition, for each combination of manipulated variable values (X), we need to find the model outputs when model parameters take different values (ω). Here we selected 120 combinations of model parameter values (i.e. $\omega_1, \omega_2, \ldots, \omega_{120}$). If we did not use HDMR, the computation cost would be very high. Even if we use HDMR, we still need to do 265*120 = 31,800 simulations. Obviously, we need some way to reduce simulations further. Then we plotted the model output with one combination of manipulated variables (e.g. X_1) against that with another combination of manipulated variables (e.g. X_2) under different combinations of parameter values (i.e. $\omega_1, \omega_2, \ldots, \omega_{120}$), we found that there was a strong relationship between them (see Figure 4). The correlation coefficients of effluent total phosphorus (TP) concentration, TP load, and operation cost (energy cost plus sludge disposal cost) were all over 0.87, which implied that we could use the model outputs of at some data points to estimate or interpolate model outputs at other points. The details of interpolation can be found in Burden and Faires (2001).

In short, with interpolation and HDMR, the number of simulation needed was reduced to 2,232, which was only about 1/840 of the original work (15625*120 = 1,875,000).

APPLICATION OF OPTIMIZATION STRATEGY

For nutrient removal process, model-based control can find many applications. At least two possible targets can be used to optimize manipulated variables: one is to produce best effluent quality for a given plant configuration; the other is to give lowest operation cost for given plant configuration and effluent limit. Here we optimized the manipulated variable values to meet the following limits: effluent TP mean < 1.25 mg/l, effluent TP 95th percentile < 2 mg/l, and effluent TP load < 23 kg/day. As there was uncertainty with model parameters, we wanted to be 95% confident that the limits above can be met with lowest operation cost. Then we used HDMR to predict the model outputs of all combinations of manipulated variables and model parameters, and selected manipulated variable values that can meet our targets with minimum operation cost.

As expected, for the optimal combination of manipulated variables, all influent was allocated to anaerobic tank because the carbon compound in the influent can be decomposed into short chain volatile fatty acid (VFA) needed by PAOs. The levels of internal recirculation, outer recycle, and waste sludge were shown in Figure 5. The results showed that the internal recirculation was kept at low level (equal to influent flow), which enhanced energy saving and TP removal. The outer cycle was kept at low level (20% of influent flow), whereas the waste sludge was kept at relatively higher level (17.5% of outer cycle), resulting in more TP removal through sludge disposal. The DO set-point was 2 mg/l, and the aeration level was shown in Figure 6.

Generally aeration level was relatively low, which was beneficial to energy saving and fermentation in the anaerobic tank. The effluent TP profile was shown in Figure 7, and the cumulative probability of TP mean, TP 95th percentile, and TP load were shown in Figure 8, Figure 9, and Figure 10, respectively. From these results, it was found that the probability levels that TP mean exceeded 1.25 mg/l, TP 95th percentile exceeded 2 mg/l, and TP load exceeded 23 kg/d were respectively 5 %, 0 %, and 5 % respectively.

These results were equal to or better than our preset targets (violation probability of 5%), indicating that HDMR and interpolation could be used to control complicated process such as nutrient removal.

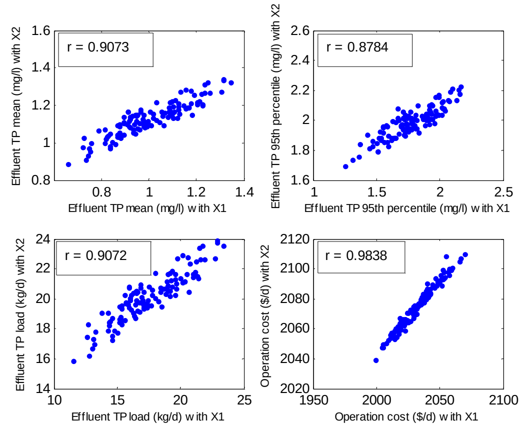

Figure 4. Correlation of effluent TP concentration, effluent TP load and operation cost with two combinations of manipulated variables (X_1 and X_2) under 120 combinations of model parameters ($\omega_1 \sim \omega_{120}$).

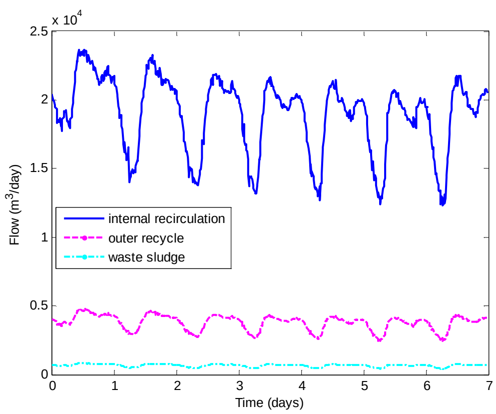

Figure 5. The levels of internal recirculation, outer recycle, and waste sludge

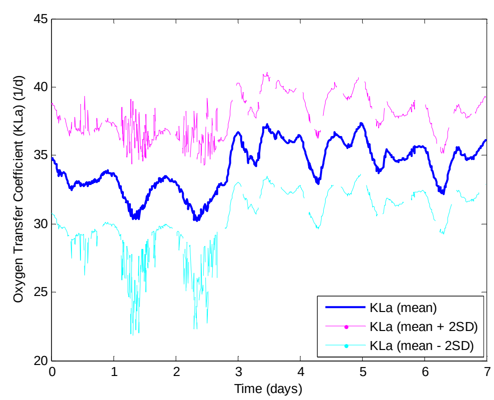

Figure 6. The level of aeration (Kla) in aerobic tank

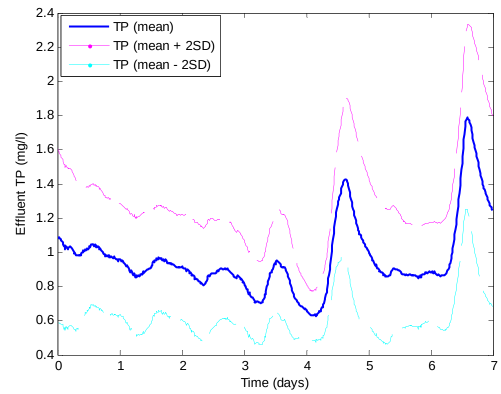

Figure 7. The effluent TP profiles

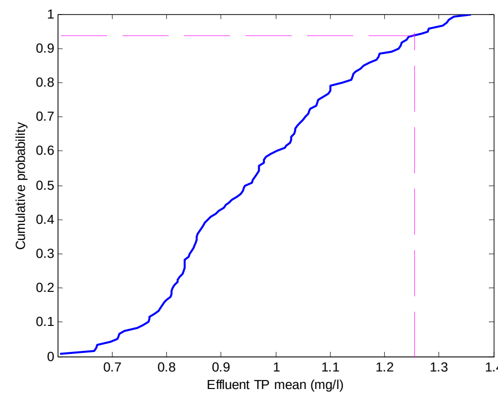

Figure 8. The cumulative probability of effluent TP mean

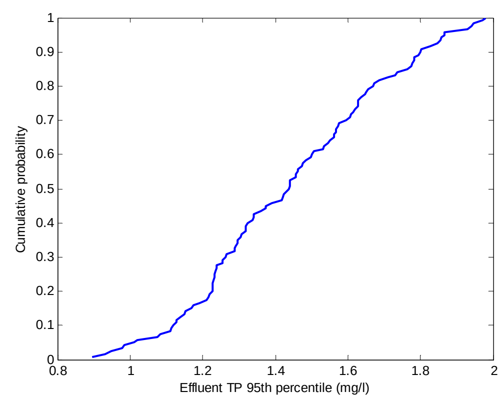

Figure 9. The cumulative probability of effluent TP 95^{th} percentile

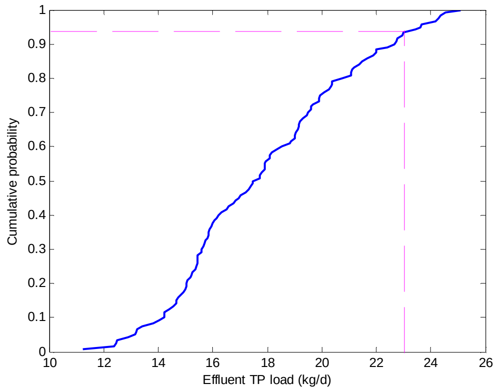

Figure 10. The cumulative probability of effluent TP load

CONCLUSIONS

In this research, HDMR and interpolation were used to approximate model outputs at different combinations of manipulated variable levels and model parameter values. According to the approximation, the levels of manipulated variables were optimized to meet different control targets in a reasonable time. The agreement between the approximate results and actual simulation outputs indicated that these algorithms could be used to control complicated process such as nutrient removal.

LITERATURE CITED

Burden, R.L. and Faires, J.D. (2001). Numerical analysis (7th ed.) BROOKS/COLE Pacific Grove, CA.

Henze, M.; Gujer, W.; Mino, T.; Wentzel, M.C.; Marais, G.v.R.; Loosdrecht M. C. M. V.

(1999). Activated sludge model No.2d, ASM2d. Wat. Sci. Tech. **39**(1), pp165-182.

Infanger, G. (1993). Planning under uncertainty: solving large-scale stochastic linear programs. South-Western Publishing Co. Danvers, Massachusetts.

Li, G., Wang. S. and Rabitz, H.(2000). High dimensional model representations (HDMR): concepts and applications. http://www.ima.umn.edu/reactive/abstract/li1.html

Liu, R. (2000) Monitoring, modeling, and control of nutrient removal in the activated sludge process. Ph.D. dissertation, University of Georgia, Athens, Georgia.

Liu, R.; Beck, M.B. (2000) Solute and particulate transport characterization in the activated sludge process. In: Preprints, IWA WATERMATEX 2000, Gent, September 18-19, pp 8.33-8.39.

Maybeck, P.S. (1982). Stochastic models, estimation, and control. ACADEMIC PRESS, INC. New York.

Saltelli, A., Chan, K. and Scott, E.M. (2000). Sensitivity analysis. John Wiley & Sons, Ltd. 605, Third Avenue, New York.