Korean J. Chem. Eng., **27**(2), 504-510 (2010)
DOI: 10.1007/s11814-010-0081-2

Settling state detection of SBR based on DO profile analysis using dynamic time warping

Yejin Kim*,†, Hyeon Bae, Hyosoo Kim*, Jungphil Shin***, and Changwon Kim***

*Department of Environmental Engineering, Pusan National University, Busan, Korea
**Busan Techno-Park, Busan, Korea
***Aizu International University, Japan
(*Received 24 February 2009 • accepted 19 August 2009*)

Abstract−Settleability of activated sludge is one of the most important variables for stable solid-liquid separation of the biological wastewater process. Moreover, effective decanting is a sensitive work at sequencing batch reactor (SBR) which has a settleability fault, such as filamentous/non-filamentous bulking, deflocculation and sludge rising. It is not easy to monitor sludge settleability directly without any specified measurement system, but the values of settling phase can be measured by installing basic measuring instruments for monitoring the process in the reaction stage of SBR. In this study, patterns of DO profiles measured at settling phase showing significant difference according to the process status were used to explore whether a problem occurs or not. To use this information, an online algorithm was developed to detect and diagnose the settling fault. A dynamic programming method that is one of the pattern recognition methods was used to detect and classify the patterns of the DO profiles. Based on the discriminant function made by dynamic time warping results and an extracted variable from DO profiles, the classification rules were generated. With the discriminant function, the settleability fault was detected and classified successfully.

Key words: Detection, Diagnosis, Discriminant Analysis, DO, Dynamic Time Warping, SBR, Settleability

INTRODUCTION

The primary objective of the activated sludge process is to guarantee good effluent quality though successful pollutant removal and good solid-liquid separation [1]. Settleability, usually checked as SVI (sludge volume index), is a very important monitoring variable to control solid-liquid separation at the secondary settling tank. At the large scale WWTPs (wastewater treatment plants), which have skilled operators for management, SVI is measured everyday and the settleability, initial settling velocity and turbidity of supernatant are observed by human operators' eyes. With results of checking at an SVI test, the operator judges the presence of settleability problems and infers causes of settleability faults based on the operator's heuristics. To aid the operator's daily decision on facing settleability, many researchers have developed settling models or decision supporting tools [2-4]. For the case of a sequencing batch reactor (SBR), the decanting process is a very sensitive work especially with some settleability problems. The operators should examine the decanting efficiency at every discharging.

Usually, SBR has basic measurements such as DO, pH, and ORP because of its dynamic characteristics of reaction. The basic sensors perform on-line measuring and many strategies for effluent quality control have been suggested [5-8]. Although they have been used for a variety purposes, there is no trial to extract valuable information for settleability using the basic sensors. In this research, operational knowledge that the DO profile's pattern is different between good settling and bad settling was obtained by long-term operation. Sometimes the operational knowledge and the operator's heuristics can be the most valuable source for diagnosis system development [9,10]. The clearly different patterns of DO made it possible to automate the observation and inference procedure performed by a human operator. To make an automated algorithm for detection and isolation of the settling state, dynamic time warping (DTW) was used to recognize pattern type and discriminant analysis was used to isolate the settling state.

†To whom correspondence should be addressed.
E-mail: yjkim@pusan.ac.kr

MATERIALS AND METHODS

1. Sequencing Batch Reactor

A pilot scale SBR was operated to obtain DO patterns at various

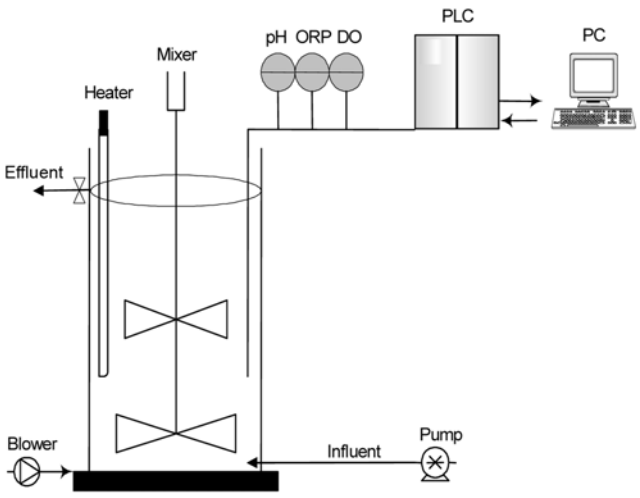

Fig. 1. Schematic of SBR.

Table 1. Operating mode of SBR

Title	Time	Operating	
Sub-cycle 1,2,3,4	4 hr	Anoxic 2 hr Oxic 2 hr	Influent feeding 5 min at starting
Sub-cycle with external carbon	4 hr	Anoxic 2 hr Oxic 2 hr	External carbon feeding 5 min at starting
Settling	3 hr	-	
Idle	1 hr	-	

settleabilities (Fig. 1). Reactor volume was 60 L and temperature was controlled as 30 °C. Synthetic wastewater was given to the reactor which was made glucose, acetate and lactose with a ratio of 4 : 2 : 2 and other nutrients. The influent loading was 0.035 NH_4^+-N kg/m^3/subcycle and C/N ratio 3. The influent and external carbon source as methanol was injected to SBR using peristaltic pumps. The pumps, a mixer and a blower were operated automatically using PLC and PC. DO, pH and ORP measurement system were installed and their measured values were stored on a connected personal computer every 30 seconds. The operating schedule of the SBR is described in Table 1. The removal efficiency was over 90% for 100 days operation for COD and N removal. The initial DO concentration of settling phase was from 5 ppm to 7 ppm because the last process before settling was aeration without feeding.

2. Settling Properties and DO Profiles

The DO probe was located at 2/3rd of the total height from the bottom of the SBR. When a sludge settleability fault occurred, the DO probe was exposed to the unsettled sludge and the DO profile reflects the DO consuming reaction of the sludge directly, showing steep slope to the DO concentration zero (Fig. 2(a)). On the other hand, with no settleability problem, the DO probe was exposed to supernatant for a short time, due to fast settling, and the DO profile reflects slow dispersion of DO being consumed at the settled sludge. It makes the DO profiles have a more slowly decreasing pattern than the case of settleability fault (Fig. 2(b)). Therefore, the patterns of the two cases could be distinguished easily by the human eye. The DO profiles at various settling state were obtained at SVI values from 20 to 200.

3. Dynamic Time Warping (DTW)

The DTW algorithm is known as a general method for extraction of properties within the pattern; it has been applied in the fields such as recognizing handwritten characters or signatures.

The DTW algorithm removes the time difference between patterns by warping the time-axis of the pattern until one pattern matches the best to the target-pattern [11]. Every pattern-vector is warped to the reference-pattern vector of the same category with the equal number of characteristic-vector (feature vector). As the reference pattern A and target vector B are given below:

$$A=a_1, a_2, \ldots, a_i, \ldots, a_K$$
$$B=b_1, b_2, \ldots, b_j, \ldots, b_M \quad (1)$$

Where, A is reference pattern, B is the pattern-vector that should be warped to A.

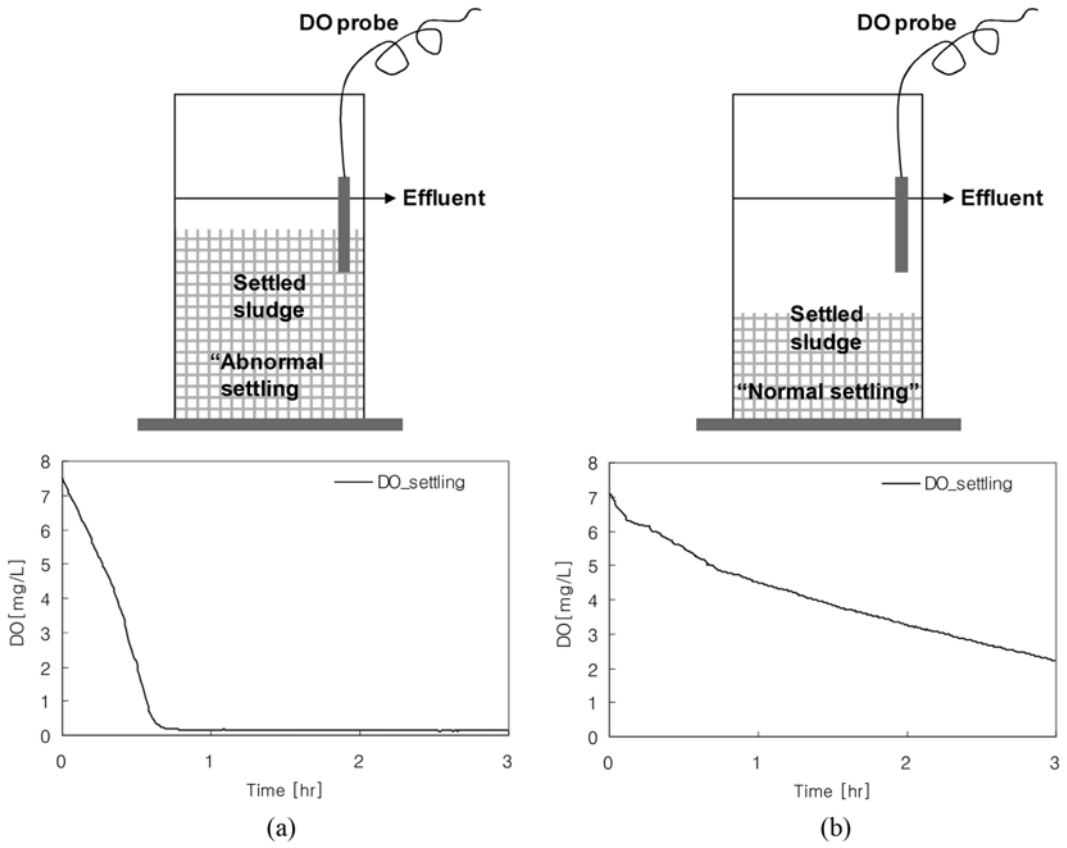

Fig. 2. DO profiles at (a) bad settleability (b) good settleability.

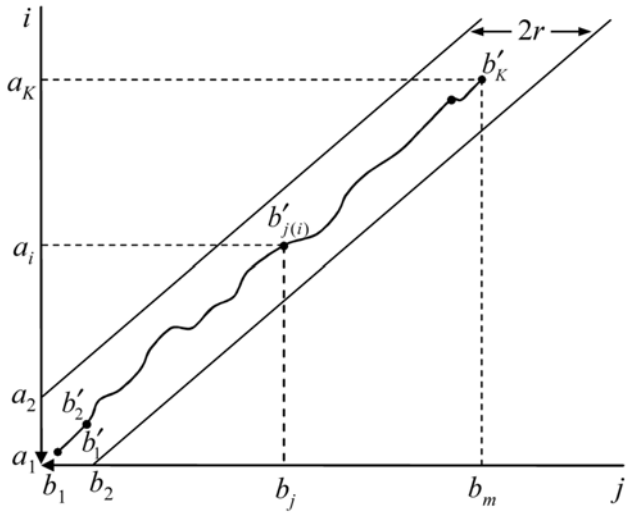

Fig. 3. Warping function and adjustment window.

Let A be the reference pattern and B be the pattern vector to be aligned against A. Fig. 3 shows A and B developed against the i and j axes.

Consider a warping function F between the input pattern time j and the reference pattern time i, where

$$j=j(i) \tag{2}$$

A measure of the difference between the two feature vectors a_i and b_j is the distance

$$d(i,j)=\|a_i-b_j\| \tag{3}$$

When the warping function is applied to B this distance becomes

$$d(i,j(i))=\|a_i-b'_j\| \tag{4}$$

where b'_j is the jth element of B after the warping function has been applied.

The weighted summation of these distances on the warping function is

$$E(F)=\sum_{i=1}^{K} d(i,j(i))*w(i) \tag{5}$$

where w(i) is a nonnegative weighting coefficient.

E reaches a minimal value when the warping function is determined to optimally align the two pattern vectors.

The minimum residual distance between A and B is the distance still remaining between them after minimizing the timing differences. The time normalized difference is defined as the following equation. Because this value is based on the summation of the distances between standard elements and warped elements, it can be affected by noise of the signal. If the target experiments produce noisy signals, a proper noise-filtering method should be applied to the signals before applying dynamic time warping.

$$D(A,B)=\min_{F}\left[\frac{\sum_{i=1}^{K} d(i,j(i))*w(i)}{\sum_{i=1}^{K} w(i)}\right] \tag{6}$$

Applying dynamic programming principles to the simplified time normalization equation gives the following algorithm for calculating the minimal value of the summation:

The dynamic programming equation is

$$g(i,j(i))=\min[g_{i-1}(i-1,j(i-1))+d(i,j(i))*w(i)] \tag{7}$$

The time normalized distance is

$$D(A,B)=\frac{1}{N}g_K(i(K),j(K)) \tag{8}$$

The initial condition is

$$g_1(1,1)=d(1,1)*w(1)=d(1,1) \tag{9}$$

In this study, p was set by 2 and all these steps were established with MATLAB ver. 7.1. For P=2 the dynamic programming equation will be

$$g_i(i,j)=d(i,j)+\min\begin{cases} g(i-1,j-1) \\ g(i-1,j-2) \\ g(i-1,j-3) \end{cases} \tag{10}$$

The permissible paths through which the warping functions may move under this slope constraint are shown in Fig. 4. Fig. 5 shows a sample application including the weights.

4. Discriminant Analysis

Discriminant analysis is a statistical technique that allows the researcher to study the differences between two or more groups of objects with respect to several variables simultaneously. It can be

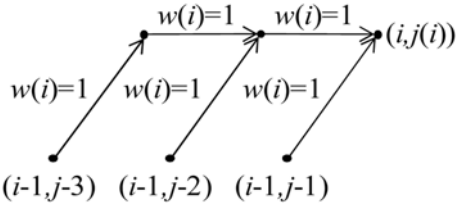

Fig. 4. Possible warping function path.

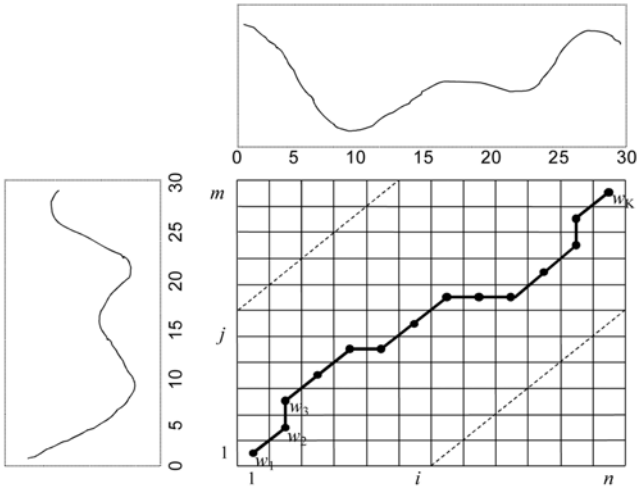

Fig. 5. Dynamic programming path using template pattern and test pattern.

used to classify the dependent variables into nominal scale (good or bad) using metric independent variables. Because this method has the advantage of showing the significant difference within each group based on considering meaningful variables, the method can provide more understandable results than that by using only individual variables [12]. The ratio of the variance is used to well describe the difference within groups in discriminant analysis, and the linear combination of the independent variables to maximize the ratio is called the discriminant function.

$$Y=B_0+B_1X_1+B_2X_2+\cdots+B_pX_p \quad (11)$$

where, Y, B_0, B_1-B_p, and X_1-X_p are discriminant score, discriminant constant, discriminant function coefficients, and independent variables used in the discriminant function, respectively.

The above equation is called Fisher's linear discriminant function. The function coefficients are fitted to maximize the discriminant score for data classification. In this study, the settleability was isolated by using the discriminant score (Y) that is calculated based on pattern matching between templates (standard profiles) and object signals (test profiles) using SPSS ver. 14.0.

RESULTS

1. Collected DO Profiles

To extract the features, it is necessary to generate the template signal. Fig. 6(a) shows DO profiles with various SVI indices, which occurred at the same MLSS concentration and Fig. 6(b) shows the profiles, which were recalculated based on "DO value at the beginning of settling phase - DO(t)" to equalize the initial value before conducting DTW.

2. Standard Profiles

Data sets like Table 2 among the data obtained by controlling SVI randomly were prepared. Because the experiment was per-

Table 2. Training and validation data set

Name		SVI [ml/g]	MLSS [mg/L]	Case
Training	T1	93.75	8,000	Abnormal
	T2	168.4	5,000	Abnormal
	T3	36.4	5,500	Normal_slow
	T4	38.8	5,410	Normal_slow
	T5	51.8	7,140	Normal_slow
	T6	29.5	5,755	Normal_fast
	T7	37.63	6,910	Normal_slow
	T8	<30	7,000	Normal_fast
Validation	V1	38.9	5,405	Normal_slow
	V2	32.4	7,000	Normal_slow
	V3	<30	7,000	Normal_fast
	V4	>100	7,000	Abnormal

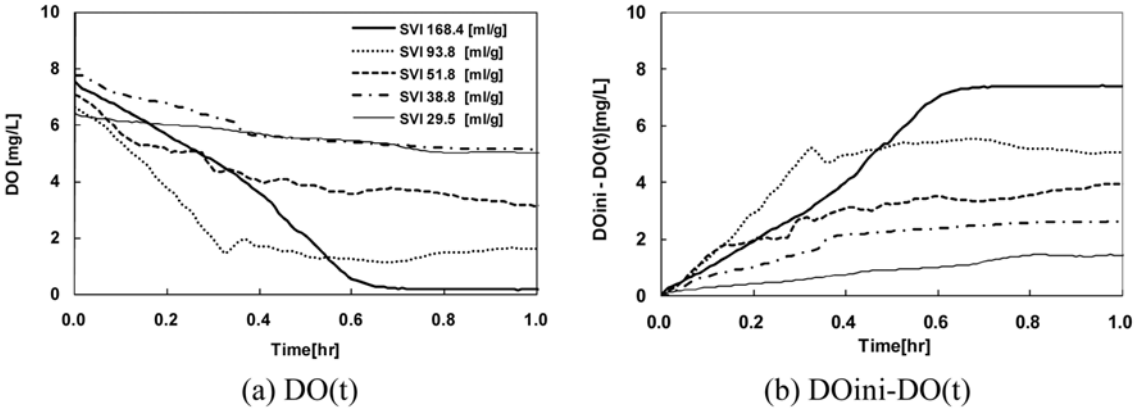

Fig. 6. DO(t) and DOini-DO(t) for the SVI.

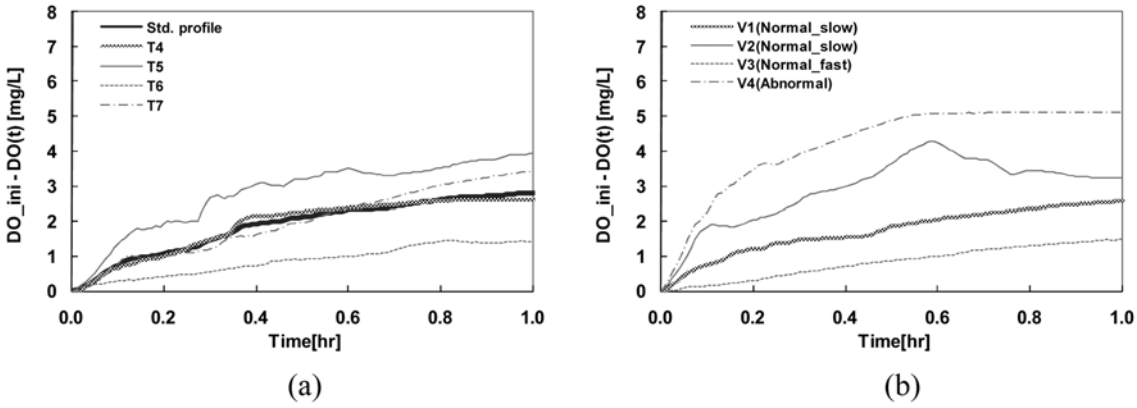

Fig. 7. Pattern profiles of training (a) and validation (b) data set.

formed by dividing SVI into three regions, the settling pattern could be divided into three regions. First one is the high SVI region (T1-T2; SVI higher than about 80) that the settling velocity is becoming lower. Second one is the middle SVI region (T3-T5, T7; SVI 30-80) with normal settling velocity. Third one is the low SVI region (T6, T8 and V1; SVI below 30). Among these SVI regions, the high SVI region was defined as abnormal settling due to the largest risk at discharging of supernatant of SBR. Medium and low SVI region can be considered as normal settling but the settling velocities are different from each other. Therefore, three cases were defined as Abnormal, Normal_slow and Normal_fast. As shown in Fig. 7, the time-based standard profile for normal settling, which is a template profile of the dynamic time warping, was generated by calculating the average value of DOini-DO(t) of T4-T7.

3. Detection Algorithm Based on Dynamic Time Warping

Dynamic time warping was performed between the standard settling profile and the target data T1-T8 as shown in Fig. 8. The distance D, called the overall matching score between the both patterns, was obtained as shown in Table 3.

If the D values resulting from pattern matching are used to detect the settleability fault in the case that the target profile is similar to the standard profile, the discriminant result guarantees high accuracy. However, if the target profile has quite a difference with the abnormal patterns used in this study or the distribution of the standard profile is shifted, it is difficult to guarantee the correct detection result by using the D values by itself. When using dynamic time warping method to compare two patterns from an experiment, it should be noted that the standard patterns may not explain every case of the patterns from the experiments. Therefore, in this study, a variable for pattern features, DOini(t)-DO(t) at 1 hour is added to improve the performance as shown in Table 3.

Table 3. Input data for model and rule generation

No.	Case	D value	DO_{ini}-DO(t) at 1 hr [mg/L]
T1	Abnormal	121.320	5.079
T2	Abnormal	156.750	7.363
T3	Normal_slow	5.317	2.570
T4	Normal_slow	3.862	2.607
T5	Normal_slow	34.927	3.926
T6	Normal_slow	7.983	3.431
T7	Normal_fast	87.731	1.435
T8	Normal_fast	70.474	1.557

As shown in the results of Table 3, in normal settling, there is a significant discrepancy between the D value by matching normal settling patterns to the standard profile and the D value by matching abnormal settling patterns to the standard profile. In addition, the DOini(t)-DO(t) at 1 hour shows the larger values for abnormal cases than those for normal cases. To extract the features and make

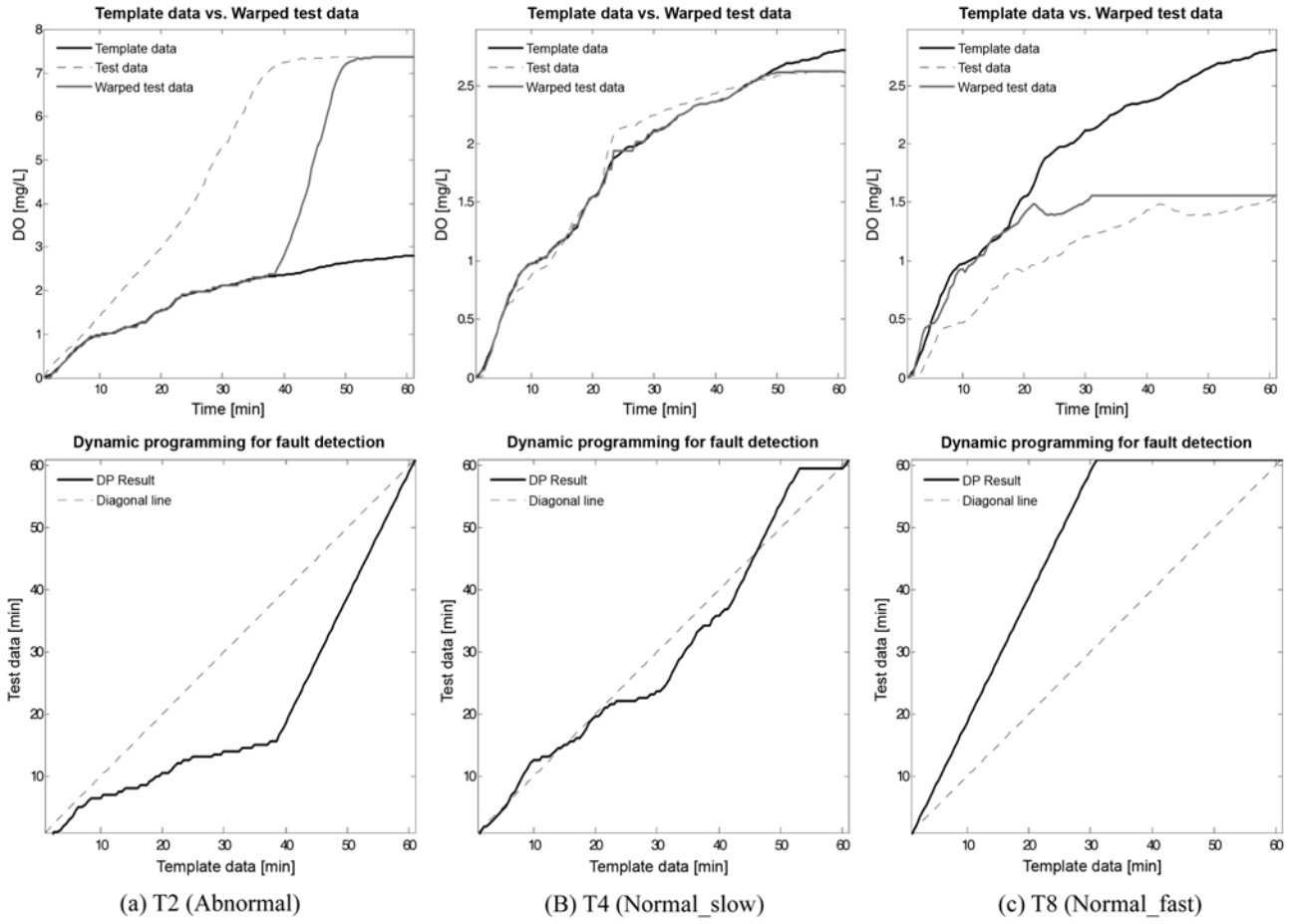

Fig. 8. Dynamic programming of test pattern to the standard pattern.

Table 4. Linear discriminant function for status

Variable	Y1 (Abnormal)	Y2 (Normal_slow)	Y3 (Normal_fast)
Constant	−33.92325	−16.26760	−21.65382
x1 (DP_value)	0.50244	−0.52184	0.74158
x2 (DO_slop at 1 hr)	−0.32312	12.55168	−10.26275

Table 5. Validation results using the developed discriminant function

Data	D	DO_{int}-DO(t) at 1 hr [mg/L]	Y1	Y2	Y3	Decision
V1 (Normal_slow)	5.48	2.60	−32.01	**13.51**	−44.27	Normal_slow
V2 (Normal_slow)	12.81	3.24	−28.53	**17.71**	−45.41	Normal_slow
V3 (Normal_fast)	162.65	1.48	47.32	−82.57	**83.78**	Normal_fast
V4 (Abnormal)	140.75	7.36	**34.42**	2.66	7.19	Abnormal

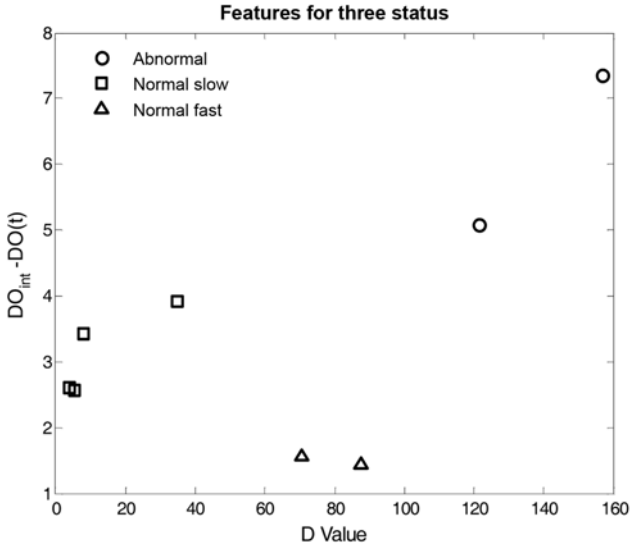

Fig. 9. Extracted features for three status of settling.

an objective rule to identify the test pattern's class, the discriminant equation was generated and the settleability fault was isolated by applying the discriminant equation to two features, D and DOini(t)-DO(t) at 1 hour.

4. Fault Isolation Based on Discriminant Analysis

As mentioned above, the input features for fault isolation are the distance by using DTW and DOini(t)-DO(t) at 1 hour values. The two input features include enough information to classify the settling fault, so it is possible to detect and isolate the normal and abnormal conditions using the features.

For classification using the coefficients of the discriminant function in this study, first, similarity is given by calculating the distance of each class. The final discriminant function consists of the coefficients as shown in Table 4. The discriminant function was generated for the three classes, so the results can be expressed as the three equations. For validation, four sets of data set (V1-V4) were applied to the developed algorithm, and using their D and DOini(t)-DO(t) at 1 hour of data it was proven that they can be classified to each group successfully (Table 5). This is because those two variables are distributed well for a clear classification to the each correct class as shown in Fig. 9.

CONCLUSION

A fault detection and isolation algorithm is proposed to monitor the settleability fault based on the heuristic knowledge of the operator. The proposed system can detect and isolate abnormal settleability caused by unstable settling that results overflow of solids with the effluent water when draining supernatant. First, dynamic programming was used to extract the feature pattern from DO profiles that are able to identify the settleability fault when setting is not completed. Next, for fault isolation, discriminant analysis that is a statistical method was applied. The two input features used in discriminant function generation and pattern feature extraction can well describe settling status, so the features can provide a good classification result based on the final discriminant function. Even though the cases applied in this study are not sufficient, the proposed methodology can imitate human visual inspection and inference processing to act on behalf of the operator's roles automatically.

ACKNOWLEDGMENT

This study was financially supported by grant from Pusan National University in program (Post-Doc. 2007) and Korea Ministry of Environment (MOE) as 『Human resource development Project for Waste to Energy』 and 『I²WaterTech of Eco-STAR project (I²WATER-TECH 04-5)』.

REFERENCES

1. D. Jenkins, M. G. Richard and G. T. Daigger, *Manual on the causes and control of activated sludge bulking and foaming*, 2nd edn., Lewis Publishers, Boca Raton, USA (1993).
2. T. Sekine, H. Tsugura, S. Urushibara, N. Furuya, E. Fujimoto and S. Matsui, *Wat. Res.*, **23**, 361 (1989).
3. C. K. Yoo, S. W. Choi and I. B. Lee, *Korean J. Chem. Eng.*, **19**(3), 377 (2002).
4. J. De Clercq, M. Devisscher, I. Boonen, P. A. Vanrolleghem and J. Defrancq, *Wat. Sci. Tech.*, **47**(12), 105 (2003).
5. G. Andreottola, G. Bortone and A. Tilche, *Wat. Sci. Tech.*, **35**(1), 113 (1997).

6. H. Bae, D. W. Choi, S. P. Cheon, S. Kim and Y. Kim, *Korean Institute of Intelligent Systems*, **15**(4), 431 (2005).
7. K. M. Poo, J. H. Im, J. H., Ko, Y. J. Kim, H. J. Woo and C. W. Kim, *Korean J. Chem. Eng.*, **22**(5), 666 (2005).
8. D. S. Kim, N. S. Jung and Y. S. Park, *Korean J. Chem. Eng.*, **25**(4), 793 (2008).
9. J. Baeza, D. Gabriel and J. Lafuente, *Environmental Modelling & Software*, **14**, 383 (1999).
10. M. B. Beck, A. Latten and R. M. Tong, *Modelling and operational control of the activated sludge process in wastewater treatment*, Professional Paper, 78-10, International Institute for Applied Systems Analysis, Laxenberg, Austria (1978).
11. H. Sakoe and S. Chiba, *IEEE Transactions on Acoustics Speech and Signal Processing*, **26**(1), 43 (1978).
12. S. Sharma, *Applied multivariate techniques*, John Wiley & Sons, USA (1996).