

인공신경망 패턴분석을 이용한 국내도시하수처리장 운전 연구

백병천[†] · 주치훈*

전남대학교 환경시스템공학과

*전남녹색환경지원센터

A study on the Operation of Domestic Municipal Sewage Treatment Plants using Neural Network Pattern Analysis

Byeong-Cheon Paik[†] · Chi-Hoon Ju*

Department of Environmental System Engineering, Chonnam National University

**Jeonnam Green Environment Center*

Abstract

Neural network pattern analysis were used for the efficient operation of three municipal sewage treatment plants. The nonlinear relationship of variables in each the sewage system was investigated in this study by a pattern analysis. SOM (Self Organizing Maps) Technique was used for the pattern analysis of correlations between operating parameters and effluent quality of CA, GY, YS municipal sewage treatment plants operating activated sludge processes, respectively.

Average flowrate of the influent sewage to the CA municipal sewage treatment plant was 141,148 m³/day. Average influent concentrations of CA municipal sewage treatment plant were BOD 145.3 mg/L, COD 88.6 mg/L, SS 157.5 mg/L, T-N 35.157 mg/L, T-P 5.737 mg/L, respectively. Average flowrate of the influent sewage into the GY municipal sewage treatment plant was 22,040 m³/day. Average influent concentrations of GY municipal sewage treatment plant were BOD 81.6 mg/L, COD 40.7 mg/L, SS 61.4 mg/L, T-N 23.968 mg/L, T-P 2.664 mg/L, respectively. Average flowrate of the influent sewage into the YS municipal sewage treatment plant was 11,408 m³/day. Average influent concentrations of YS municipal sewage treatment plant were BOD 79.2 mg/L, COD 52.9 mg/L, SS 66.1 mg/L, T-N 22.071 mg/L, T-P 2.078 mg/L, respectively.

Rainfall, flowrate, BOD volume loading, and F/M ratio influenced significantly to discharge effluent quality of CA and YS municipal sewage treatment plants. Rainfall, flowrate, SRT, and RAS influenced significantly to the discharge effluent quality of GY municipal sewage treatment plant. Operating parameters such as BOD volume loading, F/M ratio, SRT, and RAS influenced significantly to the efficient operation of the plants, but the rainfall and flowrate have strongly influenced to the discharge effluent quality of the three municipal sewage treatment plants.

Keywords : SOM, Pattern analysis, Sewage treatment plant, Operating parameters, Effluent quality

[†]Corresponding author E-mail: bpaik@jnu.ac.kr

1. 서론

우리나라는 경제개발 5개년 계획을 성공적으로 수행함으로써 놀라운 경제성장과 산업화를 이룩하였다. 경제성장은 생활수준의 향상을 가져왔지만 환경을 고려하지 않은 급속한 산업화와 개발 그리고 인구의 도시집중은 환경의 오염을 가중시켜 결국 심각한 환경문제에 봉착하게 되었다. 수질을 오염시키는 생활하수와 산업폐수 등 각종 오염원이 증가되어 왔음에도 불구하고 수질오염방지시설의 확충은 이를 뒤따르지 못하고 있다. 특히, 우리나라의 많은 도시에 하수도 시설이 보급되어 있는 실정이지만, 기술과 경험이 축적되지 않은 상태에서 짧은 기간에 처리장을 건설하여 운전하는 과정에서 여러 가지 문제점이 도출되었다. 하수도는 인구의 증가와 도시화의 급속한 확산으로 인해 인간 생활에 있어 가장 필수적인 시설로써 인식되고 있으며 단순히 지역적인 역할을 떠나 광역적인 측면에서의 환경보전을 위한 중요한 시설이 되고 있다.

하수처리장 설계 시 생물학적 하수처리시스템에서 오염물질의 거동 및 실제 처리장의 운전 시 발생하는 비용이나 시간의 낭비 및 효율적인 처리를 위해 IAWQ에서 개발된 활성슬러지모델(ASM Model Series)을 이용하여 하수처리장을 설계하고 있다. 활성슬러지모델은 수학적 모델에 기초를 둔 분석기술로서 다양한 조건에서의 하수처리장의 현상을 재현하는데 좋은 수학적 모델로 볼 수 있다. 하지만, 이 모델을 통해 실제 운영되는 활성슬러지의 하수처리장의 문제를 분석하고 통제하기란 난해한 문제이다. 특히, 원인과 영향의 관계 묘사, 다양한 파라미터에 대한 정의, 운영조건에 따른 파라미터의 조정과 같은 어려움 때문이다. 이런 어려움은 하수의 특성상 시간에 따른 유입유량의 변화와 유입수질의 변화, 지역적인 요소 등의 영향인자에 따라 성상이 다르기 때문이다¹⁾.

최근 인공지능망을 이용한 패턴인식 연구가 활발히 진행되고 있고 또한 좋은 성과를 올리고 있다. 인공지능망 모델은 다음과 같은 이유로 기존의 접근 방법에서 나타나는 패턴인식 문제를 해결하는데 매우 적합함을 보이고 있다.

첫째, 패턴인식에 필요한 많은 양의 데이터(data)를 병렬처리 할 수 있다. 둘째, 패턴인식과 같은 수학적 알고리즘의 적용이 곤란한 문제를 학습 방법에 의해 효과적으로 처리할 수 있다. 셋째, 패턴인식에서 많이 발생하는 노이즈(noise) 및 애매모호한(fuzzy) 데이터 처리, 일반화(generalization) 등에 효율적으로 대응할 수 있다²⁾.

SOM(Self-Organizing Map)은 Kohonen(1982)이 제안한 신경망 모델의 한 분야이며, 여러 가지 장점들을 가지고 있다. 첫째, 이 네트워크는 구조상 수행이 상당히 빠른 모델이다. 이 네트워크는 오류 역전파 모델과는 달리 여러 단계의 피드백이 아닌, 단 하나의 전방향 흐름(feed forward flow)을 사용한다. 그러므로 SOM은 잠재적으로 실시간 학습 처리할 수 있는 모델이다. 둘째, 이 네트워크는 연속적인 학습이 가능하다. 그러므로 만약 입력데이터의 통계적 분포가 시간에 따라 변하면 SOM은 자동적으로 이러한 변화에 적응하게 된다. 셋째, SOM은 자기조직화를 통한 정확한 통계적 모델이다³⁾. 따라서, 본 연구에서는 인공지능망에 기초를 둔 SOM을 이용한 패턴분석을 이용하여 활성슬러지공법으로 실제 운전되고 있는 국내 도시하수처리장을 분석하여 도시하수처리장의 효율적인 운전을 위한 운전인자들의 관계를 연구하고자 한다.

2. 연구방법

2.1. SOM(Self Organizing Map)

개요 및 특성

SOM은 Kohonen(1982)이 제안한 신경망 모델

의 한 분야이며 자기조직화지도(Self-Organizing Map) 또는 Kohonen Map 이라고도 불린다.

2.1.1. SOM의 개념

SOM은 자율학습(Unsupervised learning)을 사용한 신경망의 일종으로 경쟁 학습 모형을 이용하고 있다. 기대되는 출력 타입 없이 입력 데이터만 주어지는 자율 학습에서는 유사한 입력 값이 들어올 경우 그들을 같은 군집으로 묶기 위해 스스로 연결 강도를 조절해 나간다. 즉, 입력된 데이터 집합의 통계적 특성을 추출해서 유사한 데이터를 같은 군집으로 묶어 가는 학습 알고리즘이 진행된다. SOM은 다음과 같이 입력층과 출력층의 두 개의 층으로 구성되어 있다^{4,5)}.

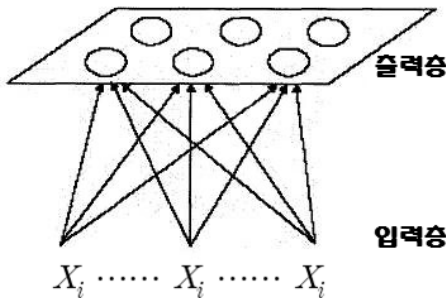


Fig. 1. Structure of SOM.

2.1.2. SOM 알고리즘

SOM의 알고리즘은 6단계로 구분하여 정의할 수 있다⁶⁻⁹⁾.

1) 뉴런의 숫자와 지도의 형태 정의

뉴런의 수는 일반적인 지도화와 주변 뉴런과의 크기 조정을 완만하게 할 수 있도록 가능한 큰 수를 선택한다.

2) 가중치 초기화

가중치의 값을 작은 값의 무작위 수의 작은 값으로 초기화한다.

3) 입력벡터 X 선택

4) BMN(Best Matching Neuron) mc 선택

입력벡터 X와 가장 인접한 뉴런 mc 선택은 다음 과정을 통해 구한다.

$$\|X - m_c\| = \min_i \|X - m_i\| \quad (1)$$

여기서 $\| \cdot \|$ 는 유클리디안 거리인 유클리디안 놈을 의미한다. c는 BMU인 승자뉴런을 의미하며, 입력벡터 X와 가장 유사한 가중치로 정의된다. 입력벡터와의 유사성은 전형적인 유클리디안의 거리를 통해 정의한다.

5) $m_c(t)$ 와 $N_c(t)$ 정의

$m_c(t)$ 와 $N_c(t)$ 를 X(t)에 따라 다음과 같이 정의한다.

$$\begin{aligned} m_i(t+1) &= m_i(t) + \alpha(t) [X(t) - m_i(t)] \quad \forall i \in N_c(t) \\ m_i(t+1) &= m_i(t) \quad \forall i \notin N_c(t) \end{aligned} \quad (2)$$

여기서 $\alpha(t)$ 는 학습률(Learning Rate)로서 스칼라값으로 $0 \leq \alpha(t) \leq 1$ 의 범위를 갖으며, 전형적인 값은 0.2-0.5이다. $N_c(t)$ 는 승자뉴런 c에 이웃한 뉴런들의 집합을 나타낸다. Fig. 2에서 왼쪽 그림은 격자상에서 $m_c(t)$ 와 $N_c(t)$ 를 표현한 것이며, 오른쪽 그림은 샘플입력자료에 따른 주변뉴런과 코드벡터를 표현한 것이다. $m_c(t)$ 와 $N_c(t)$ 의 뉴런의 거의 절반이상을 포함하며, 되풀이를 반복할수록 포함되는 뉴런의 수는 감소한다.

6) “5)” 과정 반복

이런 알고리즘을 통해 학습이 끝나면 코호넬의 출력노드 평면에 여러 Cluster가 생기게 된다. 각 Cluster는 한 노드를 중심으로 영역으로 나타난다. 어떤 Cluster는 서로 겹치는 부분이 있게 되고, 어떤 Cluster는 완전히 독립적인 영역으로 나타난다. 영역을 갖게 하

면 할수록 영역간에 겹치는 부분이 줄어들게 되므로 각 Cluster를 구별짓게 되는 것이다. 따라서, 코호네프트는 통계적인 성격이 강하므로 각 Cluster간의 겹치는 부분을 작게 하기 위해서는 코호네프트의 출력층의 크기가 적당히 커야한다. 작으면 서로 겹치는 부분이 많이 생기게 된다.

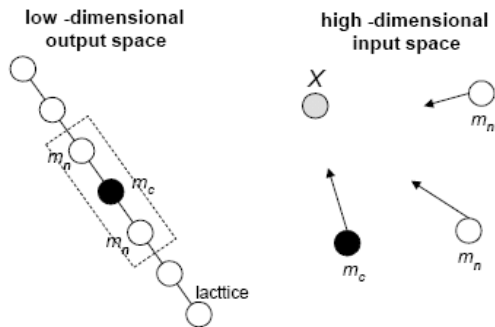


Fig. 2. Illustration of the SOM learning algorithm.

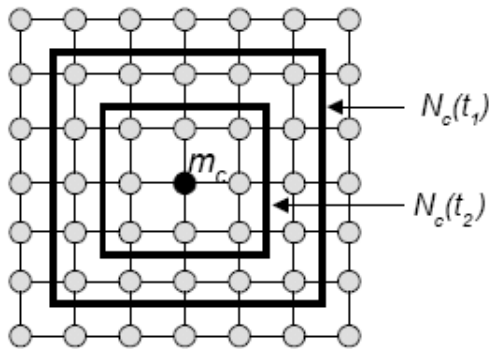


Fig. 3. Topological neighbourhood $NC(t)$ around the best-matching neuron $mC(t)$ at different training time steps.

2.2. 자료 분석방법

SOM을 이용하여 효율적인 하수처리장의 운전을 위해 대상하수처리장의 수질과 운전인자와의 관계를 분석하고자 한다.

CA, GY 및 YS 하수처리장의 2005~2007년도 하수처리장 운전 자료를 바탕으로 운전인

자와 처리수의 관계를 비교하였으며, 주요 운전인자는 강수량, 유입유량 BOD 용적부하, F/M비, SRT, RAS로 하였고, 처리수의 항목은 BOD, COD, SS, T-N, T-P로 하였다.

3. 결과 및 고찰

3.1. CA 하수처리장

CA하수처리장의 평균 하수유입량은 141,148 m^3/d 이며, 유입수질은 BOD 145.3 mg/L, COD 88.6 mg/L, SS 157.5 mg/L, T-N 35.157 mg/L, T-P 5.737 mg/L이었으며, 주요 운전인자는 BOD 용적부하는 평균 0.50 kg/d, F/M비는 평균 0.11, SRT는 평균 10.56 d, RAS는 평균 45%인 것으로 조사되었다. CA하수처리장의 강수량은 좌측 상단에 많은 양이 유입된 것으로 나타났다. 강수량이 많을 때는 비교적 낮은 농도가 유입되어 처리수는 강수량의 변화와 상반되는 패턴이 나타나 처리수에 영향을 미치는 것으로 나타났다. CA하수처리장의 유입유량은 좌측 상단에 많은 양이 유입된 것으로 나타났다. 유입유량이 많을 때는 비교적 낮은 농도가 유입되어 처리수는 유입유량의 변화와 상반되는 패턴이 나타나 처리수에 영향을 미치는 것으로 나타났다. CA하수처리장의 BOD 용적부하는 좌측 상단에 높은 값이 나타났다. BOD_{eff} , COD_{eff} , SS_{eff} , T- N_{eff} , T- P_{eff} 와 상반되는 패턴이 나타남으로써 BOD, COD, SS, T-N, T-P의 처리에 영향을 미치는 것으로 분석되었다. CA하수처리장의 F/M비는 좌측 상단에 높은 값이 나타났다. BOD_{eff} , COD_{eff} , SS_{eff} , T- N_{eff} , T- P_{eff} 와 상반되는 패턴이 나타남으로써 BOD, COD, SS, T-N, T-P의 처리에 영향을 미치는 것으로 분석되었다. CA하수처리장의 SRT는 상단에 일부 높은 값이 나타났다. 상반된 패턴이 없는 것으로 보아 처리수에 영향을 미치지 않는 것으로 분석되었다. CA하수처리장의 RAS는 하단에 높은 값으로 나타났다. BOD_{eff} , COD_{eff} , SS_{eff} ,

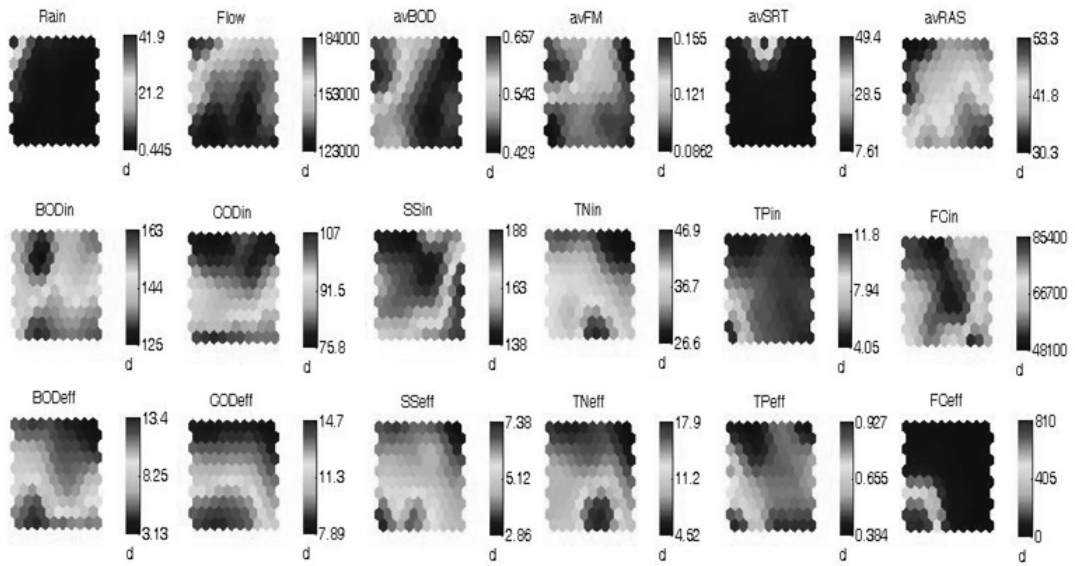


Fig. 4. CA하수처리장의 운전인자와 처리수의 관계

T- N_{eff} , T- P_{eff} 와 유사한 패턴이 나타나 BOD, COD, SS, T-N, T-P의 처리에 영향을 미치지 않는 것으로 분석되었다.

3.2. GY하수처리장

GA하수처리장의 평균 하수유입량은 22,040 m^3/d 이며, 유입수질은 BOD 81.6 mg/L, COD 40.7 mg/L, SS 61.4 mg/L, T-N 23.968 mg/L, T-P 2.664 mg/L이었으며, 주요 운전인자는 BOD 용적부하는 평균 0.15 kg/d, F/M비는 평균 0.14, SRT는 평균 11.80 d, RAS는 평균

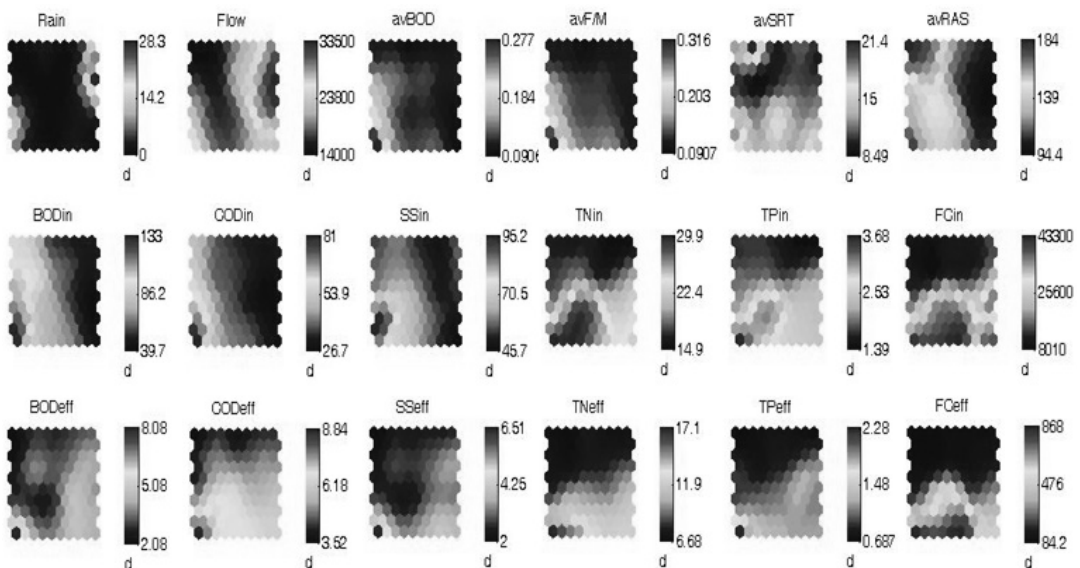


Fig. 5. GY하수처리장의 운전인자와 처리수의 관계

131.48%인 것으로 조사되었다. GY하수처리장의 강수량은 우측에 많은 양이 유입된 것으로 나타났다. 강수량이 많을 때는 비교적 낮은 농도가 유입되어 처리수는 강수량의 변화와 상반되는 패턴이 나타나 처리수에 영향을 미치는 것으로 나타났다. GY하수처리장의 유입유량은 우측에 많은 양이 유입된 것으로 나타났다. 유입유량이 많을 때는 비교적 낮은 농도의 유입되어 처리수는 유입유량의 변화와 상반되는 패턴이 나타나 처리수에 영향을 미치는 것으로 나타났다. GY하수처리장의 BOD 용적부하는 좌측 하단에 높은 값이 나타났다. BOD_{eff}, COD_{eff}, SS_{eff}, T-N_{eff}, T-P_{eff}와 유사한 패턴이 나타나 BOD, COD, SS, T-N, T-P의 처리에 영향을 미치지 않는 것으로 분석되었다. GY하수처리장의 BOD 용적부하는 좌측 하단에 높은 값이 나타났다. BOD_{eff}, COD_{eff}, SS_{eff}, T-N_{eff}, T-P_{eff}와 유사한 패턴이 나타나 BOD, COD, SS, T-N, T-P의 처리에 영향을 미치지 않는 것으로 분석되었다. GY하수처리장의 SRT는 좌측 상단에 높은 값이 나타났다. BOD_{eff}, COD_{eff}, SS_{eff}, T-N_{eff}, T-P_{eff}와 상반되는 패턴이 나타남으로써 BOD,

COD, SS, T-N, T-P의 처리에 영향을 미치는 것으로 분석되었다. GY하수처리장의 RAS는 좌측 상단에 높은 값이 나타났다. BOD_{eff}, COD_{eff}, SS_{eff}, T-N_{eff}, T-P_{eff}와 상반되는 패턴이 나타남으로써 BOD, COD, SS, T-N, T-P의 처리에 영향을 미치는 것으로 분석되었다.

3.3. YS하수처리장

YA하수처리장의 평균 하수유입량은 11,408 m³/d이며, 유입수질은 BOD 79.2 mg/L, COD 52.9 mg/L, SS 66.1 mg/L, T-N 22.071 mg/L, T-P 2.078 mg/L이었으며, 주요 운전인자는 BOD 용적부하는 평균 0.44 kg/d, F/M비는 평균 0.10, SRT는 평균 7.44 d, RAS는 평균 128.06%인 것으로 조사되었다. YS하수처리장의 강수량은 좌측 하단에 많은 양이 유입된 것으로 나타났다. 강수량이 많을 때는 비교적 낮은 농도가 유입되어 처리수는 강수량의 변화와 상반되는 패턴이 나타나 처리수에 영향을 미치는 것으로 나타났다. YS하수처리장의 유입유량은 좌측에 하단에 많은 양이 유입된 것으로 나타났다. 유입유량이 많을 때는 비교적 낮은 농도의 유입되어 처리수는 유입유량

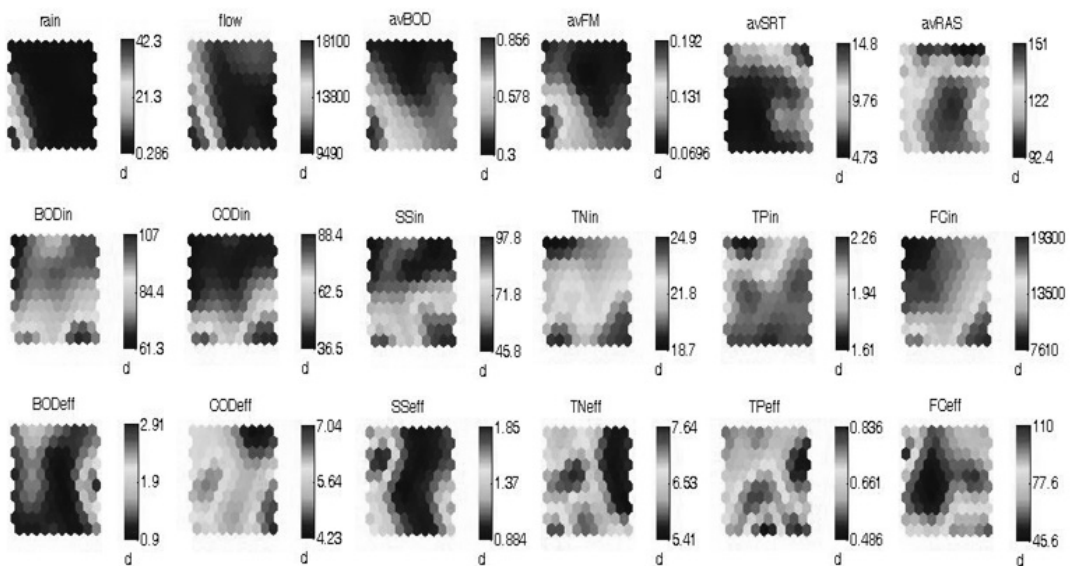


Fig. 6. YS하수처리장의 운전인자와 처리수의 관계

의 변화와 상반되는 패턴이 나타나 처리수에 영향을 미치는 것으로 나타났다. YS하수처리장의 BOD 용적부하는 좌측 하단에 높은 값이 나타났다. BOD_{eff}와 상반되는 패턴이 나타남으로써 BOD의 처리에 영향을 미치는 것으로 분석되었다. YS하수처리장의 F/M비는 좌측 하단에 높은 값이 나타났고, BOD_{eff}와 상반되는 패턴이 나타남으로써 BOD의 처리에 영향을 미치는 것으로 분석되었다. YS하수처리장의 SRT는 우측 상단에 높은 값이 나타났고, 상반된 패턴이 없는 것으로 보아 BOD, COD, SS, T-N, T-P의 처리에 영향을 미치지 않는 것으로 분석되었다. GY하수처리장의 RAS는 전반적으로 높은 값이 나타났고, 상반된 패턴이 없는 것으로 보아 BOD, COD, SS, T-N, T-P의 처리에 영향을 미치지 않는 것으로 분석되었다.

4. 결 론

- 1) CA하수처리장의 평균 하수유입량은 141,148 m³/d이며, 유입수질은 BOD 145.3 mg/L, COD 88.6 mg/L, SS 157.5 mg/L, T-N 35.157 mg/L, T-P 5.737 mg/L이었으며, 주요 운전인자는 BOD 용적부하, F/M비, SRT, RAS 으로 주요 운전인자와 처리수의 관계를 분석한 결과 처리수는 하단에 고농도가 분포한 경향이 나타났고, 강수량, 유입유량, BOD 용적부하, F/M비는 상단에 높은 값이 분포한 경향이 나타나 처리수와 상반되는 패턴을 보임으로써 강수량, 유입유량, BOD 용적부하, F/M비가 처리장 운전에 영향을 미치는 것으로 분석되었다.
- 2) GA하수처리장의 평균 하수유입량은 22,040 m³/d이며, 유입수질은 BOD 81.6 mg/L, COD 40.7 mg/L, SS 61.4 mg/L, T-N 23.968 mg/L, T-P 2.664 mg/L이었으며, 주요 운전인자는 BOD 용적부하,

F/M비, SRT, RAS 으로 주요 운전인자와 처리수의 관계를 분석한 결과 처리수는 좌측 하단에 고농도가 분포한 경향이 나타났고, 강수량과 유입유량은 우측 상단, SRT와 RAS는 좌측 상단에 높은 값이 분포한 경향이 나타나 처리수와 상반되는 패턴을 보임으로써 강수량, 유입유량, SRT, RAS가 처리장 운전에 영향을 미치는 것으로 분석되었다.

- 3) YA하수처리장의 평균 하수유입량은 11,408 m³/d이며, 유입수질은 BOD 79.2 mg/L, COD 52.9 mg/L, SS 66.1 mg/L, T-N 22.071 mg/L, T-P 2.078 mg/L이었으며, 주요 운전인자는 BOD 용적부하, F/M비, SRT, RAS 으로 주요 운전인자와 처리수의 관계를 분석한 결과 처리수는 좌측 하단에 고농도가 분포한 경향이 나타났고, 강수량, 유입유량, BOD 용적부하, F/M비는 우측 상단에 높은 값이 분포한 패턴이 나타나 처리수와 상반되는 패턴을 보임으로써 강수량, 유입유량, BOD 용적부하, F/M비가 처리장 운전에 영향을 미치는 것으로 분석되었다.
- 4) 인공신경망에 기초를 둔 패턴분석을 이용하여 활성슬러지공법으로 실제 운전되고 있는 도시하수처리장은 BOD 용적부하, F/M비, SRT, RAS도 하수처리장 운전에 영향을 미치지만, 선정된 도시하수처리장은 공통적으로 강우량과 유입유량이 유출수 수질에 크게 영향을 미치는 것으로 분석되었다.

References

1. 이상진, 어성욱, 이상득, 김홍수(2007), 충청남도 공공하수처리시설의 운영실태 분석 및 관리방안, 충남발전연구원.
2. 양동석(1997), 변형패턴 인식을 위한 스키마 이론이 결합된 신경망 모델에 관한 연구, 박사학위논문, 전남대학교.

3. 조영임(2003), 인공지능시스템, 홍릉과학출판사.
4. 전성해, 전홍석, 황진수(2002), 자기조직화 지도를 위한 베이지안 학습, 응용통계연구, 15(2), pp. 251-267.
5. 허명희(2003), 주성분 자기조직화 지도 P C-SOM, 응용통계연구, 16(2), pp. 321-333.
6. Yoon-Seok Timothy Hong, Michael R. Rosen, Rao Bhamidimarri.(2003), Analysis of a municipal wastewater treatment plant using a neural network-based pattern analysis, Water Research, 37(7), pp. 1608-1618.
7. Ruei-Shan Lu, Shang-Lien Lo.(2002), Diagnosing reservoir water quality using self-organizing maps and fuzzy theory, Water Research, 36(9), pp. 2265-2274.
8. SOM Toolbox Team.(2000), SOM Toolbox for Matlab 5, Helsinki University.
9. Byoung-Hwa Lee, Miklas Scholz.(2006), Application of the self-organizing map (SOM) to assess the heavy metal removal performance in experimental constructed wetlands, Water Research, 40(18), pp. 3367-3374.