

다항식 방사형기저함수 신경회로망을 이용한 ASP 모델링 및 시뮬레이터 설계

Design of Modeling & Simulator for ASP Realized with the Aid of Polynomial Radial Basis Function Neural Networks

김 현 기* · 이 승 주** · 오 성 권†
(Hyun-Ki Kim · Seung-Joo Lee · Sung-Kwun Oh)

Abstract - In this paper, we introduce a modeling and a process simulator developed with the aid of pRBFNNs for activated sludge process in the sewage treatment system. Activated sludge process(ASP) of sewage treatment system facilities is a process that handles biological treatment reaction and is a very complex system with non-linear characteristics. In this paper, we carry out modeling by using essential ASP factors such as water effluent quality, the manipulated value of various pumps, and water inflow quality, and so on. Intelligent algorithms used for constructing process simulator are developed by considering multi-output polynomial radial basis function Neural Networks(pRBFNNs) as well as Fuzzy C-Means clustering and Particle Swarm Optimization. Here, the apexes of the antecedent gaussian functions of fuzzy rules are decided by C-means clustering algorithm and the apexes of the consequent part of fuzzy rules are learned by using back-propagation based on gradient decent method. Also, the parameters related to the fuzzy model are optimized by means of particle swarm optimization. The coefficients of the consequent polynomial of fuzzy rules and performance index are considered by the Least Square Estimation and Mean Squared Error. The descriptions of developed process simulator architecture and ensuing operation method are handled

Key Words : Fuzzy c-means clustering method, Particle swarm optimization, Weighted least square estimator, Fuzzy inference system

1. 서 론

최근 환경보존 및 에너지 절약에 대한 관심이 높아지면서 하수처리 시설 또한 이러한 사회변화에 맞춰 시설을 효율적으로 운영 및 관리하는데 관심이 커지고 있다. 하지만 현재 하수처리 시설을 효율적으로 운영 및 관리하는데 많은 어려움이 있다. 특히 하수처리 시설에서 가장 보편적인 공정인 활성오니 공정은 하수와 폐수를 처리하는데 있어서 가장 보편적인 공정이며 이 공정은 생물학적 특성(NH₃, NO₃, DO, PH, ORP, 온도, MLSS 등의 다수 계층과 공기량, 반송 슬러지량, 내부 반송량을 제어하는데 상호 복합적으로 운영)으로 인해 예측이 매우 어려우며 운영 효율 또한 매우 떨어진다.

국내 하수처리장은 대부분이 송풍기는 DO값에 의해 운전되고 슬러지 반송 펌프는 MLSS값에 의해 운전되도록 설계되고 있으나 생물학적 메커니즘의 특성으로 다수의 수질 측정값이 필요하여, 부하 변화가 심하고 공정의 특성이 복잡하

여 제어량과 조작량의 선택이 어렵고 다수의 제어기를 조작자가 계산하기에는 상당히 어려움에 직면하여 대부분 수동으로 운영되고 있다[1][6]. 점차 증가된 오폐수처리를 위한 국내 하수처리의 최적 설비구축 및 공정개선에 대한 연구가 활발히 추진되고 있다[11]. 최근에는 외국에서 도입한 수질 모델링 소프트웨어를 이용한 효율적 운용을 위해 설비 운전 최적화를 꾀하고 있으며[10], 한국형 인공지능 하수처리공정 모델 기법개발과 적용에 보다 관심을 가지게 되었다.

본 논문은 위와 같은 예측이 불가능하고 그에 따라 운영이 비효율적인 기존의 활성오니처리 공정을 보다 효율적인 운영을 돕는 지능형 공정시뮬레이터를 구현을 목표로 한다. 특히 퍼지뉴로 알고리즘인 FCM 클러스터링을 접목한 다항식 방사형 기저함수 신경회로망으로 활성오니공정시스템에 대한 지능형모델링 설계를 통해 그 시뮬레이터를 구축하고자 한다. 또한 시스템을 모델링하는데 있어서 최적화 알고리즘 중의 하나인 PSO 알고리즘을 사용하여 성능이 최적화된 모델을 설계한 후 적용 가능한 실 공정 시뮬레이터를 개발하고자 한다.

2. 하수처리 시스템

우선 활성오니 공정을 모델링하기 위해서는 공정의 특성 및 운영 방식을 알아야한다. 그리고 활성오니 공정을 이해하기 위해서는 보다 큰 범위의 고도처리 공정을 알아야 한

† Corresponding Author : Dept. of Electrical Engineering, The University of Suwon, Korea

E-mail : ohsk@suwon.ac.kr

* Dept. of Electrical Engineering, The University of Suwon, Korea

** Dept. of Electrical Engineering, The University of Suwon, Korea

Received : May 17, 2012; Accepted : January 17, 2013

다[1].

그림 1은 기본적인 고도처리 공정을 보여주며 생물학적 처리 시스템과 화학적 처리 시스템이 통합되어 운영되는 공정이다. 고도처리 공정이란 일반적인 2차 처리 공정의 활성오니 공정뿐만 아니라 후속 시설로 3차 처리 공정이 추가된 공정을 의미한다.

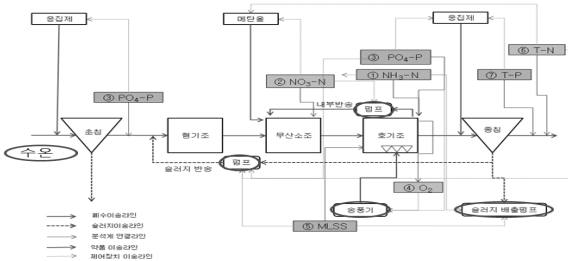


그림 1 고도처리 공정 구성
Fig. 1 Constitution of advanced wastewater treatment.

일반적인 2차 처리공정인 활성오니 공정은 BOD, SS등을 90% 정도를 제거하는 반면 질소와 인은 미생물에 필요한 영양소로만 제거되어 질소는 10~30%, 인은 10~30% 정도만 제거된다. 따라서 이러한 문제를 극복하기 위하여 질소와 인을 처리하기 위한 3차 처리 공정인 고도처리 시설을 추가 설치한다. 하지만 2차 처리 공정의 후속 시설로만 고도처리 공정을 설치할 필요가 없으므로 인하여 그림 1은 2차 공정과 3차 공정의 통합 공정이다.

2.1 생물학적 처리 시스템

그림 2는 그림 1의 고도처리 공정에서 화학 약품 처리를 분리해 놓은 부분이며, 그림 2의 각 조의 역할은 다음과 같다.

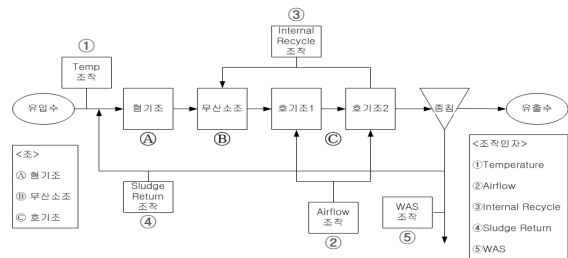


그림 2 활성오니 공정의 생물학적 시스템
Fig. 2 Biological system of activated sludge treatment.

① 혐기조는 용존산소가 없으며 유기물의 인 방출량이 증가한다. ② 무산소조는 혐기조와 같이 용존산소가 없는 조로서 호기조에서 내부 반환하여 유입되는 질산성 질소를 질소 가스로 환원시켜 대기 중으로 방출시켜 질소를 제거하기 위한 부분으로 미생물이 유기물을 분해할 때 용존산소대신 질산염(NO3)을 이용하여 질산염(NO3)을 질소(N2)가스로 처리하여 기능적인 면에서 탈 질소조라고도 부른다. ③ 호기조는 무산소조와 혐기성조에 미처리된 유기물을 제거하고 암모니아성 질소를 질산성 질소로 산화시키기 위하여 호기성 상태를 유지한다. 암모니아(NH3)=>질산염(NO3)으로 질

산화 시키는 미생물은 호기성 미생물이기 때문에 호기 상태로 유지시켜주고, 질산염(NO3) 혼합액을 무산소조로 내부 반환시켜 탈질이 일어나도록 한다. 또한 인(P)을 섭취한다. 마지막으로 중첩은 물속의 부유물을 가라앉히는 곳에서 부유물을 제외한 정화된 물만 배출 할 수 있는 구조를 가지고 있다.

2.1.1 활성오니 공정 인자 선정

본 논문에서는 생물학적 처리 공정의 모델링을 수행하였다. 고도처리 공정의 기본 모델은 화학적 처리의 경우 투입되는 약품의 증가에 의해서 결정되고 계속 유지할 경우 슬러지의 처리 비용도 크게 증가하여 생물학적 처리 공정의 효율적인 운영을 위하여 선정하였다.

첫 번째로 시스템을 모델링을 하기 위해서 사용 할 기준이 되는 수질 인자의 선택이 우선적으로 이루어져야 하며 선택되는 모델의 인자가 시스템의 모델에 관계가 밀접 할수록 모델링 시 실제 모델과 최대한 유사한 성능을 기대할 수 있다.

표 1은 공공하수처리시설의 방류수질기준으로 2011년 1월 1일 적용되는 기준이다. 아래의 표를 기반으로 인자를 선택 시 유입되는 유량은 기본으로 하며 BOD, 부유물질, 총 질소와 총 인이 된다. 하지만 하수처리 시스템의 경우 서론에서 기술한 바와 같이 공정 특성이 복잡하고 질소와 인의 경우 상관관계가 정확하지 않아 총 질소와 총 인을 세분화하여 총 질소는 암모니아(NH3)와 질산염(NO3)의 농도를 총 인에는 인산(PO4)의 농도를 인자를 추가한다.

표 1 방류 수질 기준

Table 1 Standard of discharge water of sewage.

| 구분 | 1일 하수처리 용량 50m2 이상 | 1일 하수처리 용량 50m2 미만 |
|-------------------------|--------------------|--------------------|
| 생물화학적 산소요구량 (BOD)(mg/L) | 10이하 | 10이하 |
| 화학적 산소요구량 (COD)(mg/L) | 40이하 | 40이하 |
| 부유물질 (SS)(mg/L) | 10이하 | 10이하 |
| 총 질소(T-N)(mg/L) | 20이하 | 40이하 |
| 총 인(T-P)(mg/L) | 2이하 | 4이하 |
| 총 대장균수(개/mL) | 3,000이하 | |
| 생태독성 (TU) | 1 | |

둘째로 시뮬레이터 개발 시 조작 인자를 선택해야 한다. 조작 인자는 그림 2에서 보는 바와 같이 ①Temp, ② Airflow, ③Internal Recycle, ④Sludge Return, ⑤WAS로 이루어진다. ①Temp는 유입수의 온도를 나타낸다. 수온은 물 속의 용존산소량과 미생물에 많은 영향을 미친다. ② Airflow는 호기조1과 호기조2로 불어넣는 공기량으로 미생물을 통하여 암모니아성 질소를 질산성 질소로 산화시키는 산화제 역할을 한다. ③Internal Recycle은 호기조1과 호기조2에서 만들어진 질산성 질소를 질소 가스로 변화시키기 위해 무산소조로 이동시키는 물의 양이며, ④Sludge Return은 중첩부의 가라앉은 부유물에 농축된 미생물을 일정 부분을 공정의 초기 부분으로 순환시키는 부분이다. 끝으로 ⑤WAS는 중첩부의 가라앉은 부유물의 폐기시키는 양을 뜻하는데

이는 미생물의 양에 영향을 미친다.

결론적으로 위에서 말한 총 13가지 인자(BOD, 부유물질(TSS), 총 질소(T_N), 암모니아(NH3) 질산염(NO3), 총 인(T_P), 인산(PO4), Temp, Airflow, Internal Recycle, Sludge Return, Was)는 생물학적 처리 공정에서 미생물의 생성과 유지에 직접적으로 영향을 미치기에 모델링의 위한 인자로 선택한다.

3. 다항식 RBF 뉴럴네트워크 모델

본 장에서는 활성오니 공정의 모델링을 하기 위한 다항식 방사형 기저함수 신경회로망(polynomial Radial Basis Function Neural Networks : p-RBFNNs)의 구조와 FCM클러스터링을 통한 방사형 기저함수의 중심값 선정, 후반부 연결가중치를 학습하기 위해 사용한 오류역전파 알고리즘을 설명한다.

3.1 다항식 방사형 기저함수 신경회로망

본 논문에서 사용된 다항식 방사형 기저함수 신경회로망의 모델은 다중입력 다중출력(Multi-Input Multi-Output : MIMO)의 모델구조를 가지고 있다.

추론 시스템은 퍼지 규칙 기반 시스템으로써 (1)과 같은 규칙으로 구성되어있으며 조건부의 활성함수는 방사형 기저함수로서 가우시안 함수(Gaussian Function)를 사용하였고, 가우시안 함수를 통해서 생성된 는 다음 (2)와 같은 형태를 이루며 v_i 와 σ 값은 FCM 클러스터링을 사용하여 선정한다.

$$R_j : \text{IF } x_1 \text{ and } x_2 \text{ and } \dots \text{ and } x_n \text{ is } R_j \\ \text{THEN } y_j = f_j(x_1, x_2, \dots, x_n) \quad (1)$$

$$R_i(\mathbf{x}) = \exp\left(-\frac{\|\mathbf{x} - \mathbf{v}_i\|^2}{2\sigma_i^2}\right) \quad (2)$$

3.1.1 Fuzzy C-Means 클러스터링을 통한 가우시안 함수의 중심 값 선정

클러스터링 알고리즘이란 데이터의 분류를 위해서 사용되는 것으로 데이터의 내부가 비슷한 패턴, 속성, 형태 등의 기준을 통해 데이터를 분류하여 내부의 구조를 찾아내는 방법으로, machine learning, data mining, pattern recognition, image analysis 그리고 bioinformatics와 같은 분야에 널리 사용되어지고 있다. 본 절에서는 본 논문에서 pRBFNNs 시스템의 은닉층 함수로 사용되어지는 가우시안 함수의 중심 값을 설정을 Fuzzy C-Means clustering (FCM) 방법을 사용하여 설정한다[2-4]. FCM 클러스터링 방법은 일반적으로 클러스터링 방법 중의 하나로써 데이터 전처리에 넓게 사용되어지고 있으며, 주어진 데이터의 특징을 해석하여 파악된 데이터의 정보를 기반으로 입력 데이터를 분할함으로써 기존의 데이터 전처리 방법들보다 더 효율적이고 유기적으로 주어진 데이터의 특성을 파악하는 장점을 가지고 있다.

FCM 클러스터링 방법은 n개의 벡터 x_i ($i=1, \dots, n$) 집합을 c개의 클러스터로 분할하고, 목적함수가 최소가 일 때 생

성된 각 클러스터에서 중심 값을 찾는 방법으로, FCM과 K-means 클러스터링 방법의 중요한 차이점은 FCM 클러스터링 방법은 0과 1사이의 소속정도에 의해서 나타난 소속감의 정도를 가지고, 주어진 데이터들을 몇 개의 클러스터로 나눌 수 있는지를 판별하기 위해 퍼지 분할을 사용한다는 점이다. 즉 퍼지 분할을 적용하기 위해서, 소속 함수 U는 0과 1사이의 값을 가진 요소들로 나타낸다. 그러나 데이터 집합에 대한 소속감 정도의 합은 (3)과 같이 항상 1이다.

V라 표시되는 클러스터의 중심 값을 구하기 위한 FCM 클러스터링 방법의 알고리즘은 다음과 같다.

[단계 1] 클러스터의 개수 $c(2 \leq c \leq n)$ 을 정하고 퍼지화 계수 $m(1 < m < \infty)$ 을 선택한다. 초기 partition matrix $U(r)$ 을 초기화 한다. 알고리즘 반복 횟수를 $r(r=0, 1, \dots)$ 로 표시한다.

$$U(r) = \left\{ u_{ik} \in [0, 1], \sum_{i=1}^c u_{ik} = 1 \forall k, 0 < \sum_{k=1}^n u_{ik} < n \forall i \right\} \quad (3)$$

[단계 2] (4)-(7)를 이용하여 FCM 클러스터 중심 $v_i(i=1, 2, \dots, c)$ 을 계산한다.

FCM 클러스터링에 대한 목적함수는 다음과 같은 형태를 가지고 있다.

$$J(u_{ik}, \mathbf{v}_i) = \sum_{i=1}^c \sum_{k=1}^n u_{ik}^m (d_{ik})^2 \quad (4)$$

여기서, u_{ik} 는 0과 1사이의 소속을 나타내는 정도의 값으로 i번째 클러스터에 속해져 있는 x_k 의 k번째 데이터의 소속정도를 나타낸다. v_i 는 i번째 클러스터 중심 벡터이다. m은 partition matrix의 애매함의 정도에 대한 영향을 나타내는 퍼지화 계수이다. 이 값은 $m \in [1, \infty)$ 와 같은 범위를 가지고 있다.

(4)의 목적함수를 최소화하기 위해서 먼저 v_i 와 u_{ik} 와 관련된 식의 목적함수를 다음과 같이 분화시켜야 한다.

$$\mathbf{v}_i^{(r)} = \left\{ v_{i1}^{(r)}, \dots, v_{ic}^{(r)} \right\}, v_{ij}^{(r)} = \frac{\sum_{k=1}^n (u_{ik})^m \cdot x_{ki}}{\sum_{k=1}^n (u_{ik})^m} \quad (5)$$

$$u_{ik} = \frac{(1/\|\mathbf{x}_k - \mathbf{v}_i\|^2)^{1/m-1}}{\sum_{j=1}^c (1/\|\mathbf{x}_k - \mathbf{v}_j\|^2)^{1/m-1}} = \frac{1}{\sum_{j=1}^c \left(\frac{d_{jk}}{d_{ik}}\right)^{2/m-1}} \quad (6)$$

(4)에서 주어진 목적함수를 만족하기 위하여 (5)과 (6)에 의해서 구해진 값들을 이용하여 각각의 클러스터 중심과 데이터와의 거리를 (7)의 거리 계산식을 통해 새로운 소속행렬을 생성한다.

• Euclidean distance[39] :

$$d_{ik} = d(\mathbf{x}_k - \mathbf{v}_i) = \left[\sum_{j=1}^s (x_{kj} - v_{ij})^2 \right]^{1/2} \quad (7)$$

[단계 3] 다음과 같이 새로운 소속행렬 $U(r+1)$ 을 계산한다.

$$u_{ik}^{(r+1)} = \frac{1}{\sum_{j=1}^c \left(\frac{d_{jk}^{(r)}}{d_{ik}^{(r)}} \right)^{2/(m-1)}} \quad (8)$$

[단계 4] 다음 (9)을 계산해서 만일 $\Delta > \epsilon$ 이면 $r=r+1$ 로 정하고 [단계 2]로 가서 다시 알고리즘을 반복 수행하고 그렇지 않고 $\Delta \leq \epsilon$ 이면 알고리즘을 종료한다. 여기서, ϵ 는 임계값이다.

$$\Delta = \|U^{(r+1)} - U^{(r)}\| = \max_{i,k} |u_{ik}^{(r+1)} - u_{ik}^{(r)}| \quad (9)$$

여기서, v_i 와 d_i 는 I번째 노드의 가우시안 함수의 중심 값과 폭을 의미한다. 가우시안 함수의 중심 값은 FCM 클러스터링을 통하여 구하여진 각 클러스터의 중심 값을 사용하며 폭은 클러스터의 중심 값과 데이터의 거리를 사용한다. 따라서 pRBFNNs를 구축하고 있는 노드의 수는 FCM 클러스터링의 클러스터 수와 같으며, 가우시안 함수의 중심 값은 클러스터의 중심 값과 같다.

3.1.2 오류역전파 알고리즘을 사용한 후반부 추론의 학습

본 RBFNNs의 결론부의 다항식 $f(x)$ 는 3가지 형태의 다항식 구조를 가진다. 이 3가지 형태의 다항식은 오류역전파 알고리즘을 사용하여 학습을 진행한다.

네트워크의 학습은 후반부 다항식 계수인 연결가중치를 가장 적절한 값으로 조정해 나가면서 이루어진다. 이는 원 출력과 모델 출력간의 차, 즉 오차를 바탕으로 이루어진다.

간략추론, 선형추론 및 변형된 2차식 추론의 경우 연결가중치의 조절은 기존의 값에 오차로부터 구해진 조절량만큼의 변화를 더해 주며, 간략 추론의 경우 다음과 같다.

$$E_p = \frac{1}{2} (y_p - \hat{y}_p)^2 \quad (10)$$

$$w_{kc}^0(p+1) = w_{kc}^0(p) + \Delta w_{kc}^0 \quad (11)$$

$$\Delta w_{kc}^0 = \eta \left(- \frac{\partial E_p}{\partial w_{kc}^0} \right) \quad (12)$$

$$- \frac{\partial E_p}{\partial w_{kc}^0} = - \frac{\partial E_p}{\partial \hat{y}_p} \frac{\partial \hat{y}_p}{\partial h_k} \frac{\partial h_k}{\partial a_{kc}} \frac{\partial a_{kc}}{\partial w_{kc}^0} \quad (13)$$

$$\Delta w_{kc}^0 = \eta (y_p - \hat{y}_p) \hat{f}_{kc} \quad (14)$$

또한 모멘텀 계수를 포함한 연결가중치의 변화량은 다음과 같다.

$$\Delta w_{kc}^0 = \eta (y_p - \hat{y}_p) \hat{f}_{kc} + \alpha (w_{kc}^0(p) - w_{kc}^0(p-1)) \quad (15)$$

선형추론인 경우와 변형된 2차식 추론인 경우에도 같은 방법으로 유도되며, 각각 다음과 같다.

$$\Delta w_{kc}^1 = \eta (y_p - \hat{y}_p) \hat{f}_{jk} x_i + \alpha (w_{kc}^1(p) - w_{kc}^1(p-1)) \quad (16)$$

$$\Delta w_{kc}^2 = \eta (y_p - \hat{y}_p) \hat{f}_{kc} x_i x_j + \alpha (w_{kc}^2(p) - w_{kc}^2(p-1)) \quad (17)$$

네트워크의 학습은 후반부 다항식 계수인 연결가중치를 가장 적절한 값으로 조정해 나가면서 이루어진다.

4. 다항식 RBF 뉴럴네트워크 모델의 최적화

본 장에서는 시뮬레이터 개발을 위해 제안된 다항식 RBFNNs 모델의 최적화를 위하여 입자 군집 최적화 (Particle Swarm Optimization: PSO) 알고리즘을 적용하였다[8-9]. 최적의 해를 찾아가는 PSO 알고리즘 시행 순서는 다음과 같다.

[단계 1] Initialization

n개의 particle을 탐색 공간 내에서 랜덤하게 생성한다. 이 초기 particle들은 'pbest'로 설정하고 이들이 초기 swarm을 구성한다. 각 particle의 각 velocity(이동속도)를 [-Vmax Vmin] 내에서 랜덤하게 생성한다. 초기 particle중 최적의 적합도를 갖는 particle을 'gbest'로 선정한다.

[단계 2] Inertia weight & velocity updating

(18)에 의해 관성가중 값을 계산하고, (19)에 의해 j번째 particle velocity를 계산한다. 제한조건[-Vmax Vmin]을 확인한다.

$$w(t) = W_{\max} - \frac{W_{\max} - W_{\min}}{\text{iter}_{\max}} \times t \quad (18)$$

$$v_{jk}(t+1) = w(t) \cdot v_{jk}(t) + c_1 \cdot r_1 \cdot \{pbest_{jk}(t) - x_{jk}(t)\} + c_2 \cdot r_2 \cdot \{gbest_k(t) - x_{jk}(t)\} \quad (19)$$

[단계 3] Position updating

다음 (20)에 의해 particle의 위치 정보를 조절된다. 여기서, $x_k^{\min} \leq x_{jk}(t) \leq x_k^{\max}$ 이다.

$$x_{jk}(t) = v_{jk}(t) + x_{jk}(t-1) \quad (20)$$

[단계 4] Individual & global best updating

새로운 particle들의 적합도를 계산한다. 이전 'pbest'와 비교되며 'pbest'를 재설정한다. 최적의 위치정보를 갖는 'pbest'를 'gbest'로 재설정한다.

[단계 5] Stopping criteria

종료 조건을 만족하면 탐색과정을 종료하고, 그렇지 않으면 [단계 2]부터 반복한다.

[단계 6] Optimal parameter

최종적으로 생성된 'gbest'는 최적의 위치정보를 갖는다.

5. 시뮬레이터 설계 및 구현

본 장에서 제안된 모델을 사용하여 하수처리 활성오니 공정의 모델링 및 시뮬레이터를 구현하도록 하겠다. FCM클러스터링을 사용하여 pRBFNNs 은닉층의 가우시안 활성화함수의 중심 값과 폭을 결정하며, 오류역전파 알고리즘을 통하여

연결가중치를 학습시키고, PSO 알고리즘을 사용하여 FCM 알고리즘의 클러스터수와 후반부 다항식의 형태, 오류역전과 알고리즘의 학습률을 선택한다.

5.1. PSO 알고리즘을 통한 파라미터 최적화

표 2는 모델의 최적화를 위해 사용된 PSO알고리즘과 FCM 클러스터링 및 모델의 파라미터들을 나타낸 것이다. 초기 사용된 수치적 값들은 학습과 성능의 수렴에 따른 페이스에 따라서 실험적 값과 모니터링의 운행에 의해서 시행착오를 통해 선택하였다.

표 2 PSO 알고리즘 및 모델 구축을 위한 정보

Table 2 Values of parameters of PSO algorithm and the proposed model

| | | |
|----------------|---------------------------|---------------------|
| PSO | Generation size | 20 |
| | Swarm size | 20 |
| | vmax | 20% of search space |
| | [wmin wmax] | [0.4 0.9] |
| | c1, c2 | 2.0 |
| FCM clustering | Number of clusters | 15 ~ 35 |
| | Fuzzification coefficient | 2 |
| 제안된 모델 | Polynomial type | 1 ≤ T ≤ 3 |
| | Weighting factor(θ) | 0.5 |
| | Division of data | 50% : 50% |
| | Learning rate | 0.00001 ~ 1 |
| 여기서, T는 다항식 구조 | | |

제안된 모델의 성능을 평가하기 위하여 총 674개 13입력-7출력의 데이터를 50% : 50%로 나누어 구성하였다. 입력 변수는 유입되는 물의 양, BOD, TSS, 총 질소(T_N), 암모니아(NH3), 질산염(NO3), 총 인(T_P), 인산(PO4), Temp, Airflow, Internal Recycle, Sludge Return과 Was 총 13개로 구성되어 있으며, 출력 변수는 BOD, TSS, 총 질소(T_N), 암모니아(NH3), 질산염(NO3), 총 인(T_P)와 인산(PO4) 총 7개로 구성되어 있다. 모델의 성능 지수는 MSE를 사용하여 모델을 평가하였다.

MPI는 Training data의 PI와 Test data의 EPI를 (21)를 통하여 나타냈으며, 여기서 n은 총 데이터의 수, $y_{training}^i$ 는 Training data의 실험치, $\hat{y}_{training}^i$ 은 Training data의 모델치를 나타낸다.

$$PI = \frac{1}{n} \sum_{i=1}^n (y_{training}^i - \hat{y}_{training}^i)^2 \tag{21}$$

$$EPI = \frac{1}{n} \sum_{i=1}^n (y_{testing}^i - \hat{y}_{testing}^i)^2$$

(22)는 모델 최적화를 위하여 목적함수를 정의한 부분으로 θ 값은 표 4를 통하여 확인 할 수 있다.

$$MPI = PI \times \theta + EPI \times (1 - \theta) \tag{22}$$

표 3 탐색된 최적화 파라미터

Table 3 Tuned optimization parameter

| Parameter | Value |
|--------------------|-------|
| Number of clusters | 34 |
| Polynomial type | 2 |
| Learning rate | 0.422 |

PSO 알고리즘을 통하여 찾아낸 FCM클러스터의 수, 후반부 연결가중치의 형태 그리고 오류역전과 알고리즘의 학습률은 표 3과 같다.

표 4 제안된 모델의 성능지수 값

Table 4 Values of Performance index of the proposed model

| | BOD | TSS | T_N | NH3 | NO3 | T_P | PO4 |
|-----|--------|--------|--------|--------|--------|--------|--------|
| PI | 0.0096 | 0.0097 | 0.0244 | 0.0311 | 0.0144 | 0.0497 | 0.0619 |
| EPI | 0.0042 | 0.0034 | 0.0159 | 0.0353 | 0.0172 | 0.0505 | 0.0638 |

표 4는 설계된 모델의 최적화 적용 후 출력되는 각 인자의 성능을 나타낸 것이다. MSE(Mean Square Error)의 특성상 소수점 아래의 오차의 경우, 제곱으로 인하여 더 작아지는 형태를 보이는 것을 가만할 때, 평균적인 오차는 최대 0.2대의 오차를 보임을 확인 할 수 있으며, 실제 하수처리장에서 시스템의 수동 조작 시 소수점 첫 자리까지 인정함으로 실 시스템에서 적용하기에 충분히 가능한 성능이라 평가 할 수 있다.

시뮬레이터를 개발하기 위해 엔지니어링 업체로부터 데이터를 습득하는 과정에서 캐나다의 Hydromantis사에서 개발한 GPS-X(General Purpose Simulator)라는 하수처리시설 설계프로그램을 사용하였으나, 이 프로그램은 설계의 목적으로 사용되는 시뮬레이터이기 때문에 본 시뮬레이터 개발에 사용된 데이터는 실제 하수처리장 보다 수질의 변화가 급격한 경우를 고려하여 설계되어 있기 때문이다.

이러한 데이터를 여러 개를 결합하여 모델링함으로 인해 유입되는 물의 일정한 수질을 갖는 경우의 데이터의 범위가 작고 또한 범위의 양쪽 끝의 표현하는 데이터의 양이 작기 때문에 그래프의 각 첨단에서의 오차가 상대적으로 많이 발생하였음을 볼 수 있다.

하지만 실제 하수처리장의 경우 각 계절별로 유입되는 하수는 주변 환경의 시설물이 급격하게 변화하지 않음으로 수질이 유사한 패턴을 보이며 이로 인하여 시뮬레이터는 실제 하수처리장을 튜닝 후 적용 시 충분히 극복이 가능한 문제이다.

5.2. 하수처리 시뮬레이터 구현

그림 3과 그림 5는 활성오니 공정의 모델링 및 방류수질 예측을 위해 구현된 시뮬레이터이다. 그림 3의 시뮬레이터는 하수처리 시설에 변경에 대응하여 모델링을 하기 위한 시뮬레이터로서 개발된 하수처리 활성오니 공정 뉴로 퍼지 모델이 내장되어있다. 시뮬레이터의 (a)는 모델의 튜닝 및 최적화를 위하여 PSO의 세대수와 Swarm의 사이즈, 학습

를, 학습수, 탐색할 은닉층의 범위를 설정하는 부분이며, (b)는 (a)의 선정 값으로 모델의 튜닝 및 최적화를 진행하여, 최종적인 모델의 은닉층 및 학습률, 선택된 추론 방식을 출력하고, 각 출력 인자의 최종적인 성능은 (c)에 성능지수로 나타난다. (d)는 각 출력 인자 중 선택되는 인자의 값을 (e)의 그래프에 나타낸다. 시뮬레이터를 통한 모델의 생성 후에는 모델의 파라미터가 저장된 파일이 생성되며, 이 파일을 그림 5의 시뮬레이터에서 불러오므로써 그림 5의 시뮬레이터가 완성된다.

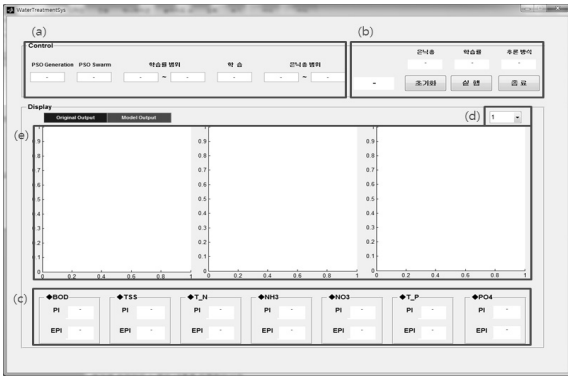


그림 3 활성오니 공정 모델링을 위한 뉴로 퍼지 시뮬레이터의 초기 화면
 Fig. 3 Initial display of Fuzzy-neuro simulator for modeling activated sludge process

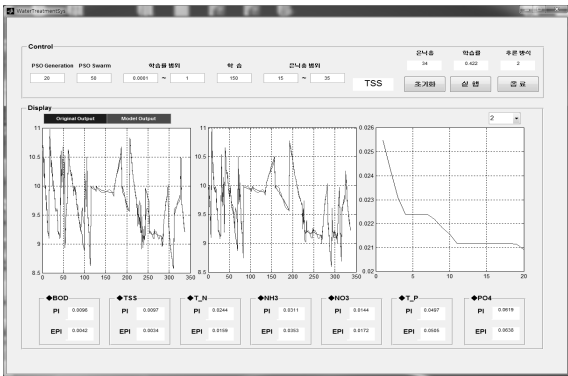


그림 4 활성오니 공정 모델링을 위한 뉴로 퍼지 시뮬레이터
 Fig. 4 Fuzzy-neuro simulator for modeling activated sludge process

그림 5의 모델링된 활성오니 공정의 방류수질을 예측하기 위한 시뮬레이터로 (a)유입수, (b)조작량, (c)유출량, (d)설정 부분으로 구성되며 (c)유출량은 유출수, 법적기준치, 기준치 달성 부분으로 구분된다. 법적 기준치는 표 1의 공공 하수처리 시설 방류 수질 기준을 사용한다.

유입수와 조작량의 변화에 따라 유출수가 변화하는 시뮬레이터이다. (a)유입수는 활성오니 공정에 유입되는 물의 수질을 나타내며 (b)조작부는 Temp, Airflow, Internal Recycle, Sludge Return, Was의 조작인자의 값을 조작하는 부분이다. 그리고 (1)-(5)는 조작부의 각 인자를 조작할 수

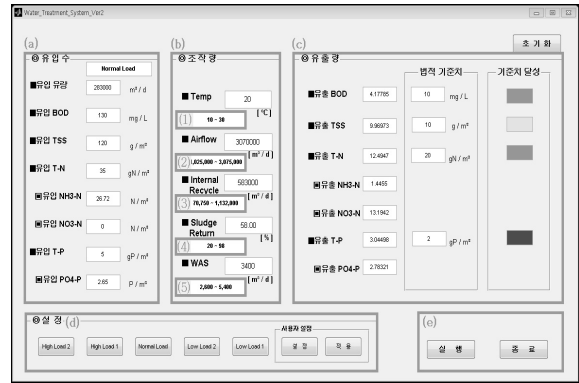


그림 5 활성오니 공정의 시뮬레이터 설계
 Fig. 5 Simulator design of activated sludge process.

있는 범위를 나타낸다. (c)유출량은 유입수와 조작량으로 추론되어 계산된 값이며 이때 각 유출수의 수질은 법적 기준치와 비교를 통하여 수질의 적합성을 기준치 달성을 3가지 색으로 나타낸다. 첫 번째로 녹색은 법적 기준치의 90%미만으로 안정적인 상황이며 두 번째로 황색은 법적 기준치의 90~100% 사이의 상황으로 관리자의 주의 필요하다. 마지막 세 번째로 적색은 방류수질기준(법적기준치)에 부합하지 못하는 물이 유출되는 상황을 나타내 관리자가 조작량을 변경해야 하는 상황을 나타낸다. (d)설정은 활성오니 공정에 유입되는 수질을 나타내는 것으로 다음과 같은 각 버튼 High Load2, High Load1, Normal Load, Low Load2, Low Load1은 각각은 오염된 수질의 정도를 설정해 놓은 것이며 사용자 설정은 임의로 사용자가 유입되는 수질을 입력이 가능하도록 만들어진 곳이다.

본 하수처리 시뮬레이터의 조작방법은 설정(d)의 원하는 수질을 선택 또는 사용자 설정 버튼을 눌러 임의로 수질을 입력한 후, (e)의 실행 버튼을 눌러 시뮬레이션을 실행한 후, 유출량(c)에 출력되는 값들을 확인하며 법적 기준치에 적합한지 확인한다.

구현된 하수처리 시뮬레이터의 조작방법 절차는 다음과 같다.

1. 설정(d)의 원하는 수질을 선택하거나 사용자 설정 버튼을 눌러 임의로 수질을 입력한다.
2. 사용자가 원하는 각 조작량(b)의 5가지를 선정한다.
3. (e)의 실행 버튼을 눌러 시뮬레이터를 실행한다.
4. 유출량(c)에 출력되는 값들을 확인하며 법적 기준치에 적합한지 확인한다.
5. 법적 기준치에 적합하지 않는 경우, 2번 단계로 돌아가 다시 수행한다.

5.3. 하수처리 시뮬레이터 동작

그림 6은 활성오니 공정의 시뮬레이터를 구동 영상을 캡처한 것으로 (a)는 조작인자를 잘못 설정하여 출력된 값이 법적 기준치에 부합하지 못하는 것이며 (b)는 조작인자의 조정을 통하여 법적 기준치에 부합하는 유출수를 시뮬레이션 한 것이다.



(a) 법적 기준과 불일치



(b) 법적 기준과 일치

그림 6 실제 활성오니 공정의 시뮬레이터
Fig. 6 Actual simulator of activated sludge process

그림 6의 시뮬레이터의 경우 논문상으로 동작 영상을 보여 줄 수 없는 제약이 따른다. 그리하여 표 5를 준비하였다. 표 5는 실제 데이터 상의 값과 시뮬레이터를 통하여 유추된 값을 비교 하고자 Normal Load의 수질 상에서 실제 데이터와 같은 수온으로 변화 시켜가며 시뮬레이터 출력 값과 실제 출력 값을 표로 나타낸 것이다. 수온은 약 5도 정도의 변화량을 보이며 정확하게 온도의 변화가 균일한 간격을 갖지 못하는 이유는 실제 데이터의 온도의 데이터 값의 간격이 아래 표와 같기 때문이다.

표 5를 보는 바와 같이 실제 출력 값과 모델의 출력 값이 매우 흡사한 것을 볼 수 있으며, 따라서 모델링 및 시뮬레이터의 제작이 성공적으로 이루어졌음을 볼 수 있다.

표 5 실 출력과 모델 출력의 비교

Table 5 Compare of real data and model output

| 온도 | 구분 | BOD | TSS | T_N | NH3 | NO3 | T_P | PO4 |
|-------|--------------|------|------|-------|------|-------|------|------|
| 12.63 | Real Output | 4.89 | 9.99 | 14.37 | 2.69 | 10.92 | 0.89 | 0.49 |
| | Model Output | 4.86 | 9.97 | 14.32 | 2.74 | 10.83 | 0.88 | 0.48 |
| 21.11 | Real Output | 4.02 | 9.78 | 12.64 | 0.71 | 11.24 | 1.46 | 1.09 |
| | Model Output | 4.01 | 9.78 | 12.67 | 0.77 | 11.22 | 1.47 | 1.09 |
| 25.15 | Real Output | 3.63 | 9.68 | 12.18 | 0.56 | 10.98 | 1.79 | 1.45 |
| | Model Output | 3.61 | 9.67 | 12.2 | 0.66 | 10.88 | 1.74 | 1.39 |
| 29.99 | Real Output | 3.18 | 9.56 | 11.78 | 0.46 | 10.72 | 2.35 | 2.05 |
| | Model Output | 3.17 | 9.55 | 11.81 | 0.57 | 10.63 | 2.32 | 2.01 |

6. 결론

본 논문에서는 하수처리 시스템의 활성오니 공정의 모델링 및 시뮬레이터 구현을 목표로 하였으며, 시뮬레이터의 구현 과정은 (a) 활성오니 공정의 모델링을 위한 인자를 총 13가지[(수질인자 : 유입되는 유량, BOD, 부유물질, 총 질소, 총 인, 암모니아(NH3), 질산염(NO3), 인산(PO4)) (조작인자 : Temp, Airflow, Internal Recycle, Sludge Return, Was)]를 선택하였으며 (b) 이를 바탕으로 총 674개, 입력 13개(유입수의 수질인자 + 조작인자), 출력 7개(유출수의 수질인자)의 데이터를 구성하였다. (c) 모델의 데이터 구성을 완료 후 pRBFNNs를 사용하여 모델링을 실시하였으며 조건부 은닉층의 가우시안 함수의 중심 값은 데이터에 FCM 클러스터링 알고리즘을 적용하여 구하였으며 최적화 알고리즘 중의 하나인 PSO알고리즘을 사용하여 오류역전과 알고리즘의 학습율과 은닉층의 수, 후반부 추론식의 종류를 선택하였다. (d) MSE를 사용하여 모델의 성능평가를 실시하였다. 그리고 (e) 완성된 모델을 바탕으로 Matlab을 이용한 시뮬레이터를 구현하였다.

본 연구의 목적은 (a) 수학적으로 매우 복잡한 활성오니 공정의 모델링을 위한 인자를 선택하는 것, (b) 최적화된 모델링을 통하여서 시뮬레이터를 구성할 수 있을 만큼 안정적으로 구성되는 것, (c) 하수처리 시뮬레이터의 설계 시 유출수를 법적 기준치와 비교를 통하여 수질의 판단이 가능하게 한 것, (d) 하수처리 시뮬레이터의 UI가 사용자의 하수처리장 운영에 편리 할 수 있도록 구성된 것, (e) 추후 제어기 추가 시 자동화 시스템이 가능한 것에 초점을 맞춰 설계되었다.

물론 데이터 구성의 어려움으로 인하여 모델링 시 각 데이터의 결합부분에서 실제 출력 값과 모델링의 출력 값의 오차가 발생하긴 하였지만 연간 데이터의 구성 시 극복될 수 있는 문제로 그 외의 부분에서는 우수한 성능을 확인할 수 있었다.

감사의 글

이 논문은 2012년도 정부(교육과학기술부)의 재원으로 한국연구재단의 지원을 받아 수행된 기초연구사업임 (NRF-2012-003568) 및 경기도의 경기도지역협력연구센터사업(GRRC 수원2012-B2, U-city 보안감시 기술협력센터)에서 지원하여 연구하였음.

References

[1] Hyun-Ki Kim, Sung-Kwun Oh, Hee-Soo Hwang, Kwang-Bang Woo, "Fuzzy Modeling for Activated Sludge Process and Control Systems" Trans, KIEE. Vol. 40, No. 10, pp. 1072- 1083, 1991.

[2] J. N. Choi, Y. I. Lee, and S. K. Oh, "Fuzzy Radial Basis Function Neural Networks with Information Granulation and Its Genetic Optimization," Lecture Notes in Computer Science, Vol. 5552, pp. 127-134, 2009.

[3] L.A Zadeh, "Toward a theory of fuzzy information granulation and its centrality in human reasoning and fuzzy logic", Fuzzy Sets Syst, Volume 90, pp. 111 - 117, 1997

[4] Y. Qian, J. Linag, C. Dang, "Knowledge structure knowledge granulation and knowledge distance in a knowledge base", Int. J. Approx. Reason., Volume 50 No. 1 , pp. 174 - 188, 2009

[5] Y. Qian, J. Linag, "Combination entropy and combination granulation in rough set theory", Int. J. Uncertain Fuzzy Knowledge-Based Syst., Volume 16 No. 2, pp. 179 - 183, 2008

[6] S.K. Oh, 2002, "FUZZY MODEL & CONTROL SYSTEM", Naeha Publishing.

[7] L. Zhao, Y. Yang, Y. Zent, "Eliciting compact T-S fuzzy models using subtractive clustering and coevolutionary particle swarm optimization," Neurocomputing, Volume 72, Issues 10-12, pp. 2569-2575, 2009.

[8] J. Kennedy, R. Eberhart, "Particle Swarm Optimization", in: Proceedings of the IEEE International Conference on Neural Networks, 4 (1995) 1942 - 1948

[9] J. Kennedy, "The Particle Swarm: Social adaptation of Knowledge", in: Proceedings of the IEEE International Conference on Evolutionary Computation, 1997, 303 - 308.

[10] GPS-X 6.1 Version User Manual, Hydromantis, Inc. 2012.

[11] MinJeong Kim, MinHan Kim, YongSu Kim and ChangKyo Yoo "Design of a Wastewater Treatment Plant Upgrading to Advanced Nutrient Removal Treatment Using Modeling Methodology and

Multivariate Statistical Analysis for Process Optimization" Korean Chem. Eng. Res., Vol. 48, No. 5, October, 2010, pp. 589-597

저 자 소 개



김 현 기 (金賢起)

1977년 연세대학교 전기공학과 졸업, 동 대학원 석사(1985), 박사.(1991). 1997~2009년 경기지역 산학연 기술지원단 단장, 1998~2009 산업기술정책연구소 심의위원, 2000~2009 사단법인 산학연 컨소시엄 전국협의회 회장, 1989~현재 수원대학교 전기공학과 교수, 관심분야는 시스템 자동화 및 감시제어, 지능형모델링 및 제어등.
Tel : 031-220-2666
Fax : 031-220-2667
E-mail : hkkim@suwon.ac.kr



이 승 주 (李承周)

2011년 수원대학교 전기공학과 졸업, 2011년 ~ 현재 동 대학원 석사과정, 관심분야는 퍼지추론 시스템, 뉴럴 네트워크
Tel : 031-222-6544
E-mail : inpneuma@suwon.ac.kr



오 성 권 (吳聖權)

1981년 연세대학교 전기공학과 졸업, 동 대학원 석사(1983), 박사(1993). 1983-1989년 금성산전연구소(선임연구원). 1996-1997년 캐나다 Manitoba 대학 전기 및 컴퓨터공학과 Post-Doc. 1993-2005년 원광대학교 전기전자 및 정보공학부 교수. 2005~현재 수원대학교 전기공학과 교수, 2002~현재 대한전기학회, 제어로봇시스템학회, 퍼지 및 지능시스템학회 편집위원. 관심분야는 퍼지 시스템, 퍼지-뉴럴 네트워크, 자동화 시스템, 고급 computational intelligence, 지능 제어 등.
Tel : 031-229-8162
Fax : 031-220-2667
E-mail : ohsk@suwon.ac.kr