

연안도시 하수처리장의 유입 BOD예측을 위한 DNN 모델 개발에 관한 연구

이상민 · 정용현[†]

국립부경대학교(대학원생) · [†]국립부경대학교(교수)

A Study on the Development of a DNN Model for Predicting the Inflow BOD in Sewage Treatment Plants

Sang-Min LEE · Yong-Hyun CHUNG[†]

Pukyong National University(student · [†]professor)

Abstract

The aim of this study is to predict BOD (Biochemical Oxygen Demand) more accurately in sewage treatment plants, allowing for the suggestion and implementation of various control strategies. This would enable the proper regulation of organic load rates, ultimately enhancing the efficiency of the treatment process. For this purpose, the study was conducted based on 2,100 daily data measurements from three sewage treatment plants in City B collected between January 1, 2015, and September 30, 2020. To predict the complex behavior of the influent characteristics that determine the BOD of the influent, a DNN(Deep Neural Network) model composed of nonlinear functions was used. To validate these results, model verification was conducted using R2 and RMSE with MLR using stepwise selection and principal component analysis. Additionally, the water quality characteristics from the immediate past are closely correlated with the BOD of the influent on the current day because of the continuous inflow nature of the influent. Consequently, derived variables of the water quality characteristics from the previous day and two days prior were added and applied to the model. To verify the validity of these derived variables, the predictive power of the model with the derived variables was compared and evaluated against a model using a dataset without these derived variables. The results demonstrated that the DNN model outperformed the existing sewage treatment plant prediction methods and the MLR model. The application of a nonlinear model to sewage treatment plants showed the potential to improve the process efficiency. Additionally, it was verified that the performance of both the DNN and MLR models improved when the derived variables were added, compared to when they were not, confirming the validity of the derived variables.

Key words : DNN, MLR, BOD, Sewage treatment plants

I. 서론

국외에서 주로 쓰레기 발생량의 예측과 하수처리장 등을 대상으로 신경망을 이용한 연구가 지속적으로 보고되고 있다.

Zaed et al.(2007)는 2004년 ~ 2007년까지 이란

마샤드 지역의 주별 쓰레기 발생량 데이터를 이용하여 한 주 뒤의 쓰레기 발생량을 예측하였다. 과거 12주까지의 주별 쓰레기 발생량 데이터와 예측 한 주 전의 쓰레기를 운반하는데 이용된 트럭 수를 설명변수로 사용하였다. 데이터를 정규화하고 은닉층 내 노드 수를 4 ~ 26개로 변화시

[†] Corresponding author : 051-629-6543, chungyh@pknu.ac.kr

켜 가장 성능이 좋은 모델을 선정하였다. 연구 결과 은닉층 내 10개의 노드로 구성된 모델의 결정계수가 0.75로 가장 높은 예측력을 가졌다. 이를 통해 신경망 내 노드 수에 따라 성능이 변화하는 것을 증명했다.

Noori et al.(2009)는 이란의 테헤란을 대상으로 한 주 뒤의 쓰레기 발생량을 예측하기 위해 예측한 주 전의 쓰레기를 운반하는데 이용된 트럭 수를 설명변수로 하고 22개의 노드로 구성된 단일 ANN 모델을 개발하여 3개의 주성분으로 구성된 다중선형회귀 모델과의 예측값을 비교하였다. 연구 결과 ANN 모델의 결정계수가 0.837로 선형회귀 모델보다 우수한 예측력을 가진 것을 증명하였다.

Abdoli et al.(2011)는 2001년~2010년까지 120개의 월별 인구수, 최고 기온, 가구 수입 데이터를 설명변수로 사용하여 2011년~2032년까지의 장래 폐기물 발생량을 예측하였다. 전체 데이터 중 84개를 학습 데이터로 사용하였고 18개를 검증 데이터로 사용하여 신경망 모델을 개발하여 선형회귀 모델과의 예측력을 비교하였다. 연구 결과, 신경망 모델이 쓰레기 발생량의 경향(Trend)과 파동을 더 정확하게 예측하였다.

Dogan et al.(2008)는 2005년의 364개 일별 COD, N, P, SS, 유량(Inflow) 데이터를 활용하여 하수처리장 유입수의 당일 BOD 농도를 예측하였다. 데이터 중 244개는 모델의 학습에 사용하였고 120개는 모델을 검증하는 데 사용하였다. 다중선형회귀 모델과 ANN 모델에 각각 적용하여 선형모델과 비선형모델의 예측 성능을 비교하였다. 연구 결과 은닉층내 노드가 3개인 ANN 모델의 결정계수가 0.919로 다중회귀분석(0.906)보다 우수한 예측력을 가지는 것을 증명했다.

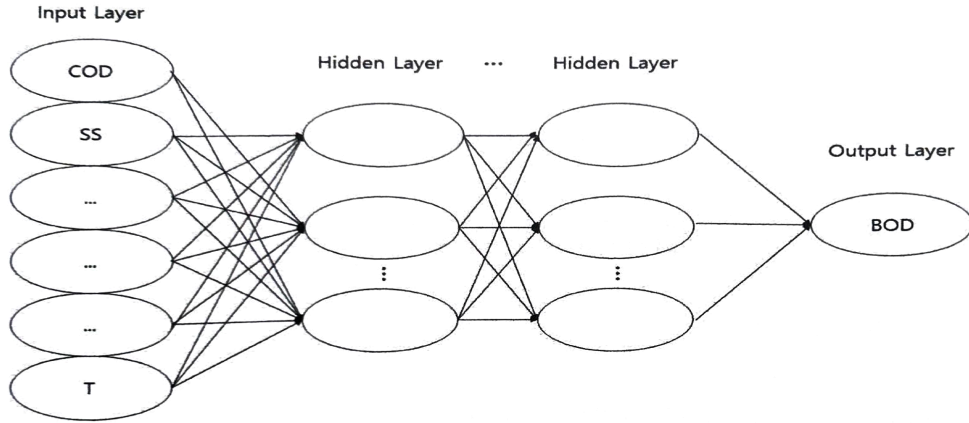
Nasr et al.(2012)는 이집트 EL-AGAMY 하수처리장을 대상으로 COD, BOD, SS 데이터를 설명변수로 사용하여 COD, BOD, SS를 각각 예측하는 ANN 모델을 연구했다. 본 연구는 1년간의 일별 데이터를 활용하여 60%는 모델의 학습 데이터로 사용하였고 20%는 검증 데이터로 사용하였

다. 나머지 20%로 검증된 모델을 최종 테스트했을 때 모델의 결정계수가 0.86으로 ANN의 예측력을 검증했다.

Hamed et al.(2004)는 10개월간 하수처리장의 일별 BOD, SS 데이터를 설명변수로 사용하여 당일 BOD, SS를 각각 예측하는 두 개의 모델을 개발했다. 모델 학습에 필요한 훈련 데이터를 각각 60%(92개), 70%(108개), 80%(123개), 90%(138개)로 나누었고 나머지를 검증 데이터로 사용했다. BOD를 예측하는 모델은 세 개의 노드가 들어간 ANN 모델을 개발했고, SS를 예측하는 모델은 두 개의 노드가 들어간 ANN 모델을 개발했다. 연구 결과, BOD를 예측한 모델의 결정계수는 0.63 ~ 0.81로 높은 예측력을 나타냈으나 SS를 예측한 모델의 결정계수는 0.45 ~ 0.65로 다소 떨어지는 예측력을 나타냈다. 이를 통해 훈련 데이터와 검증 데이터를 나누는 비율에 따라서 모델 예측력이 달라지는 것을 증명했다.

Hamoda et al.(1999)는 1995년 9월 13일 ~ 1997년 1월 1일까지 쿠웨이트의 Ardiya 하수처리장의 228개의 일별 기온, 유량, BOD, SS, 첫 번째 공정 이후의 BOD, SS 데이터를 설명변수로 사용하여 하수처리장 내 두 번째 공정 후의 BOD와 SS, 방류수의 BOD, SS 농도를 예측하였다. 본 연구는 노드가 25개로 구성된 은닉층이 1개와 2개인 두 개의 신경망 모델의 예측력을 비교하였는데 은닉층이 1개인 신경망 모델의 결정계수가 더 높았다. 본 연구를 통해 하수처리장을 대상으로 한 신경망 모델의 예측력을 검증하였고 은닉층 수의 증가가 과적합을 일으킬 수 있다는 것을 검증했다.

국내 하수처리장 유입수의 BOD 농도 예측 방법과 신경망을 통한 예측 방법에 관한 비교 연구가 국내에는 많이 부족한 실정이다(Jeong et al., 2006). 또한, 선행연구 대부분 2년 미만의 데이터를 사용해 예측력을 검증하였는데 이는 유입수의 복잡하고 비선형적 거동을 예측하기에 데이터 신뢰성이 떨어질 수 있는 데이터 크기가 판단된다(Hong, 2003).



[Fig. 1] Structure of DNN Model.

따라서 본 연구는 5년 이상의 일별 데이터를 사용하여 모델의 신뢰도를 높이고, 유입수 특성상 처리시설로 연속해서 들어오는 점을 고려하여 과거 수질 데이터와 상관성이 높다고 판단, 하루 전과 이틀 전의 수질 데이터를 파생변수로 추가하여 예측력을 높이고자 하였다.

본 연구를 통해 하수처리장에서 보다 정확한 BOD를 사전에 예측하여 다양한 제어 전략을 제시하고 대응할 수 있으며 유기물 부하율을 적절히 조절하여 공정 과정의 효율성을 높이고자 하는데 목적이 있다.

II. 연구 방법

1. 자료수집

연구에 사용된 데이터는 2015년 1월 1일부터 2020년 9월 30일까지 B시 3개 하수처리장에서 측정된 2,100개의 유입 수질 데이터인 BOD, COD, SS, TN, TP, 유량(Q), 대장균(E.coli)과 기상 데이터인 기온(T), 강수량(P)을 공공 개방 포털에서 확보하였다. 확보된 데이터는 전처리(Preprocessing) 후 예측 모델 개발을 위한 자료로 사용되었다.

2. DNN(Deep Neural Network) 모델

본 연구에 사용되는 하수처리장 유입 BOD예측을 위한 DNN 모델은 [Fig. 1]과 같은 구조로 되어 있으며, 입력층은 설명 변수에 해당하는 수질과 기온 변수를 입력시켜 여러 개의 은닉층 지나며 연산 과정을 거친 후 최종적으로 출력층에서 예측 BOD 값이 출력되도록 하였다. 본 연구의 은닉층에 사용된 활성화 함수는 ReLU(Rectified Linear Unit)로 입력값이 0보다 작으면 0으로 출력, 0보다 크면 입력값으로 출력된다.

ReLU 함수는 출력값을 특정 범위 내로 한정하지 않기 때문에 기존 시그모이드(Sigmoid) 함수에서 발생할 수 있는 기울기 소실(Vanishing gradient) 문제를 해결할 수 있으며 함수식이 간단하여 다른 비선형 함수보다 연산속도가 빠른 장점이 있다(Choi et al., 2020).

모델의 예측력을 높이기 위해선 데이터셋에 맞는 최적의 모델 구조를 설정해야 함으로 사용자는 이를 위해 은닉층 수와 노드 수를 적절하게 조절하여야 하며, 만약 노드의 수나 은닉층을 깊게 쌓지 않으면 데이터의 복잡한 변동을 예측하지 못하는 과소 적합(Underfitting) 문제가 발생할 수 있다. 반면 노드의 수가 지나치게 많고 은닉층을 너무 깊게 쌓으면 훈련에 사용되는 데이터

의 변동에만 치우쳐져 훈련 오차는 매우 낮지만 새로운 데이터가 들어 왔을 때 오차가 높게 발생하는 과대 적합 문제가 발생할 수 있으므로 사용자는 다양한 구조로 훈련된 모델의 오차를 비교하여 평가해야 한다. 따라서 역전파 알고리즘에서 매우 중요한 개념인 경사 하강법(Gradient descent)으로 신경망 훈련 시 사용되는 가중치를 편미분하여 오차 함수의 기울기가 낮은 쪽으로 이동시켜 극값을 찾은 뒤 이것을 최적 가중치로 확정하였다. 모든 변수를 동일 범위로 스케일링(Scaling)하기 위해 데이터 정규화(Normalization)를 실행했으며, 정규화된 데이터는 0~1 사이의 값을 가지도록 하였다.

가. DNN의 학습

DNN 모델의 경우 전체 데이터(2,100개) 중 무작위로 추출한 70%(1,470개)를 훈련 데이터로 사용하여 예측 모델을 개발했다. 개발된 모델의 성능을 검증하기 위해 나머지 30%(630개)를 검증(Validation) 데이터로 사용하였다. 본 연구의 DNN 모델은 반복 횟수를 2,000회, 배치 사이즈를 32로 설정하였고 과도한 학습으로 인한 과대 적합을 방지하기 위해 학습 중 100번 이상 오차가 연속해서 감소하지 않으면 학습이 종료되도록 설정하였다. 정규화를 통해 데이터가 0과 1사이 값을 가지기 때문에 학습률은 0.001로 설정하였다. 신경망은 초기 가중치를 임의의 난수 값을 생성하여 넣게 되는데, 연구의 동일한 결과를 얻기 위해 Random state의 초깃값을 0으로 고정하였다. 은닉층마다 드롭아웃(Dropout)을 0.3으로 설정하였는데, 이는 과대 적합 문제를 해결하는 방법으로 은닉층 노드를 30% 무작위로 제거된 상태에서 학습을 수행, 실제 검증 시 제거했던 노드를 복원하고 가중치를 연산하여 과대 적합을 방지한 유입 BOD를 예측하였다. 이는 파생변수의 유효성을 검증하기 위해 파생변수를 추가하지 않은 데이터셋과 추가한 데이터셋을 적용하여 총 6개의 예측 모델을 개발하였다. 파생변수를 추가하

지 않은 데이터셋은 8개(COD, SS, TN, TP, 유량, 대장균, 강수량, 기온)의 설명변수를 사용하였다. 파생변수를 추가한 데이터셋은 하루 전과 이틀 전의 수질 성상을 추가로 한 18개(BODt-1, BODt-2, COD, CODt-1, CODt-2, SS, SSt-1, SSt-2, TN, TNt-1, TNt-2, TP, TPt-1, TPt-2, 유량, 대장균, 강수량, 기온)의 설명변수를 사용하였다.

나. 다중선형회귀 모델 MLR(Multi linear regression)

DNN의 예측력을 비교 및 평가하기 위해 단계 선택법과 PCA를 이용한 다중선형회귀 모델을 개발하였다. 이들은 모두 선형결합으로 이루어진 회귀 모델이며 다중공선성 문제로부터 자유로워서 DNN 모델의 비선형성과 비교 및 평가하기 적합하다고 판단했다. 두 다중선형회귀 모델 역시 파생 변수를 추가한 데이터셋과 추가하지 않은 데이터셋 각각에 적용하여 파생 변수의 유효성을 검증하였다.

PCA를 적용한 다중선형회귀 모델의 경우 데이터 변동의 누적 설명 비율을 95% 이상이 되는 최소 주성분 수로 설정하였다. 모델에 사용되는 변수들은 표준화(Standardization) 작업을 통해 변수 간 크기 범위를 같이 스케일링하였다

다. 모델 검증

DNN 모델의 정량적 평가를 위해 예측값과 실제값의 오차 범위를 나타낼수 있는 평균 제곱근 오차(Root mean square error, RMSE)를 사용하였다(식 1). 또한, 모델의 설명력을 검증하기 위해 수정된 결정계수(Coefficient of determination, R2)를 사용하여 예측력이 가장 우수한 모델을 선정하였다(식 2).

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (\hat{y}_i - y_i)^2}{n}} \dots\dots\dots (1)$$

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \bar{y})^2}{\sum_{i=1}^n (y_i - \hat{y}_i)^2} \dots\dots\dots (2)$$

여기서, n : the number of values, i : ith value,
 $i = 1, 2, 3, \dots, n$
 y_i : measured value
 \bar{y}_i : average of measured values
 \hat{y}_i : output value by predicted model

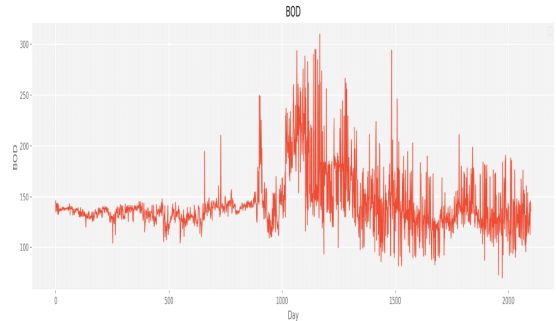
입수 내 농도 변화에 영향을 주지 않음을 의미한다.

Ⅲ. 연구 결과

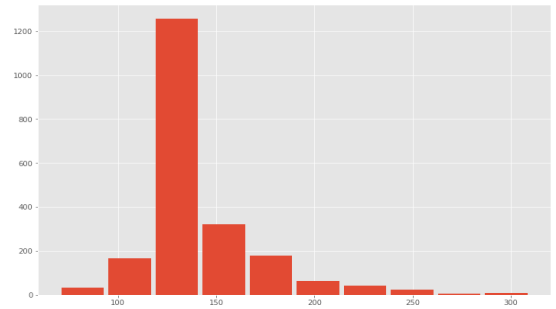
1. 상관 분석

2015년 1월 1일부터 2020년 9월 30일까지 A 하수처리장의 일별 데이터는 <Table 1>과 같다. A 하수처리장의 BOD 변동은 [Fig. 2]와 같다. 2015년부터 2017년까지 100~150ppm으로 수렴하는 양상을 보이다가 2018년부터 최대 300ppm까지 변동 폭이 커지는 것을 확인할 수 있다. 2019년부터는 변동 폭이 최대 100~200ppm 정도로 일별 최댓값이 감소하였다. 일별 BOD 데이터를 히스토그램으로 나타냈을 때 특정 농도 범위에 데이터가 125~150ppm에 치우쳐 있는 것을 [Fig. 3]을 통해 알 수 있다. 데이터가 특정 범위에 집중되어 있으면 학습 시, 해당 범위로 가중치가 조절되어 해당 범위 밖의 데이터가 들어왔을 때 오차가 커지는 과대 적합 문제가 발생할 수 있다.

중속변수 BOD와 설명변수 간 상관관계를 분석한 결과는 <Table 2>와 같다. BOD는 COD와 약 0.79의 매우 강한 양의 상관관계를 나타내고 있었고 SS, TP, TN 순으로 BOD와 강한 양의 상관관계를 나타내고 있었다. 반면 BOD와 유량의 상관관계는 0.02로 매우 낮았는데, 이는 유량의 변화가 유



[Fig. 2] Fluctuation of Daily BOD.



[Fig. 3] Histogram of Daily BOD.

다만 유량은 기온과 0.49의 강한 양의 상관관계를 나타내고 있는데 이는 기온이 높아지면서 지역 물 사용의 증가로 볼 수 있다.

강수량은 BOD와 상관계수가 -0.06으로 비가 올 때 하수처리장으로 유입되는 BOD 농도를 미미하게 감소시키는 것으로 분석된다. 이는 하수처리장은 완전 분류식 하수관거를 사용하기 때문이라고 판단된다.

<Table 1> Dataset information

	BOD	COD	SS	TN	TP	Q	E.coli	P	T
mean	142.46	80.52	130.66	34.76	4.00	8.43	1.18	4.64	12.35
std	29.02	14.44	32.56	6.90	1.21	1.39	7.03	16.82	8.68
min	70.10	45.20	66.00	11.24	1.24	5.08	1.30	0.00	-10.20
25%	128.10	72.60	112.00	30.23	3.23	7.52	7.70	0.00	5.20
50%	136.40	78.60	120.00	33.55	3.81	8.28	9.20	0.00	13.40
75%	145.20	85.20	140.00	38.43	4.40	8.97	14.08	0.20	19.40
max	309.60	217.70	400.00	78.18	14.75	14.46	96.00	264.10	28.30

<Table 2> Pearson correlation coefficient in A WTPP variables

	BOD	COD	SS	TN	TP	Q	E.coli	P	T
BOD	1.00	0.79	0.64	0.55	0.57	0.02	0.37	-0.06	-0.20
COD	0.79	1.00	0.63	0.60	0.64	-0.02	0.30	-0.07	-0.12
SS	0.64	0.63	1.00	0.54	0.57	0.16	0.51	-0.03	-0.01
TN	0.55	0.60	0.54	1.00	0.67	-0.01	0.27	-0.07	-0.08
TP	0.57	0.64	0.57	0.67	1.00	-0.05	0.30	-0.06	-0.06
Q	0.02	-0.02	0.16	-0.01	-0.05	1.00	0.19	0.29	0.49
E.coli	0.37	0.30	0.51	0.27	0.30	0.19	1.00	-0.01	-0.00
P	-0.06	-0.07	-0.03	-0.07	-0.06	0.29	-0.01	1.00	0.16
T	-0.20	-0.12	-0.01	-0.08	-0.06	0.49	-0.00	0.16	1.00

2. 다중선형회귀

파생변수를 추가하지 않은 데이터셋을 대상으로 단계 선택법을 적용했을 때 8개 변수 중 TP, 강수량 변수가 다중공선성 문제를 일으켜 소거되었다.

PCA를 통한 8개 주성분 중 6개 주성분을 사용하였을 때 전체 데이터 변동의 95.7%를 설명할 수 있었다. 이에 따른 MLR 모델의 설명변수는 <Table 3>과 같다.

<Table 3> Dataset without derived variables

MLR model	Input variables
Using Stepwise method	COD, SS, TN, Q, E.coli, T
Using PCA	PCs 1 ~ PCs 6

파생변수를 추가한 데이터셋을 대상으로 단계 선택법을 적용했을 때 18개 변수 중 SSt-1,

SSt-2, TNt-1, TP, 강수량 변수가 다중공선성 문제를 일으켜 소거되었다. PCA를 통한 18개 주성분 중 12개 주성분을 사용하였을 때 전체 데이터 변동의 95.0%를 설명할 수 있었다. 이에 따른 파생변수를 추가한 MLR 모델의 설명변수는 <Table 4>와 같다.

<Table 4> Dataset with derived variables

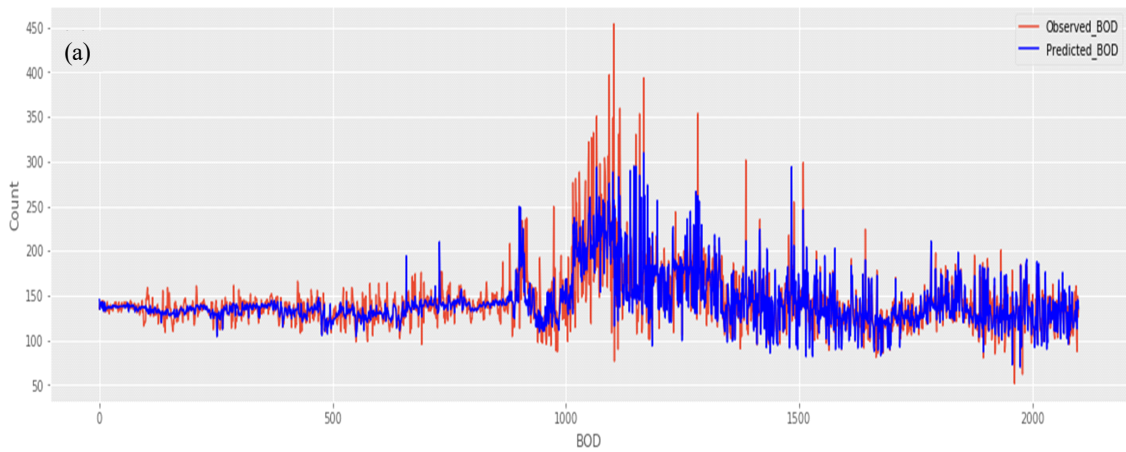
MLR model	Input variables
Using Stepwise method	BODt-1, BODt-2, COD, CODt-1, CODt-2, SS, TN, TNt-2, TPt-1, TPt-2, Q, E.coli, T
Using PCA	PCs 1 ~ PCs 12

3. 모델 비교 및 평가

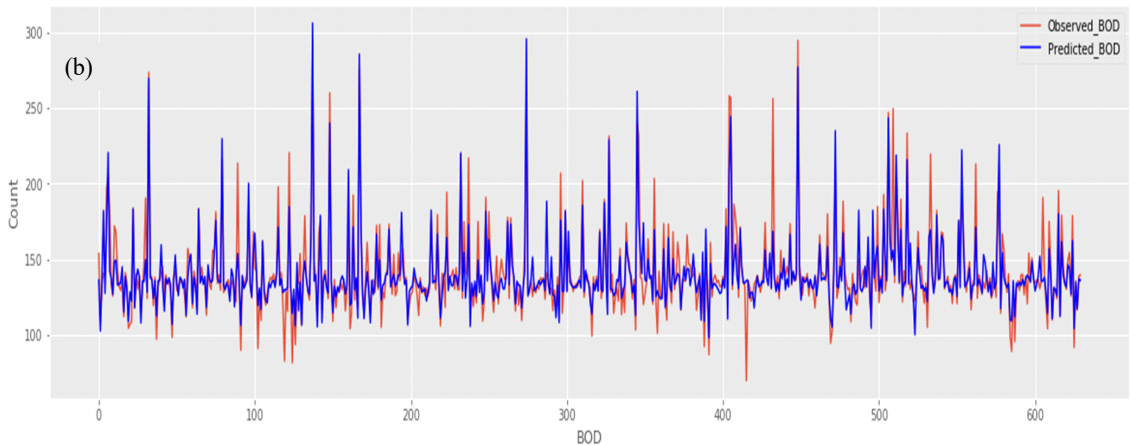
하수처리장의 기존 예측 방법을 포함한 7개의 예측 모델의 예측값과 실측값에 따른 R²와 RMSE를 통한 정량적 오차 검정을 <Table 5>와 [Fig. 4a-4g]로 나타내었다. 연구 결과, 파생 변수를 추가한 DNN 모델이 가장 뛰어난 예측력을 보였으며 예측 모델 모두 기존 하수처리장의 예측 방법보다 높은 예측력을 보였다. 또한 파생변수를 추가했을 때와 하지 않았을 때 모두 DNN 모델이 MLR 모델보다 높은 예측력을 보였다. 이를 통해 A 하수처리장의 유입 성상을 예측하는데 있어 선형결합 모델보다 비선형적 모델이 더 뛰어난 것을 확인하였다. 다만 파생변수를 추가한 DNN 모델이 파생변수를 추가하지 않은 DNN 모델보다 성능 개선은 약 3.5%로 이는 파생변수 중 과거 BOD 변수 외에는 유의미한 파생변수가 없기 때문이라 판단된다.

<Table 5> Comparison models

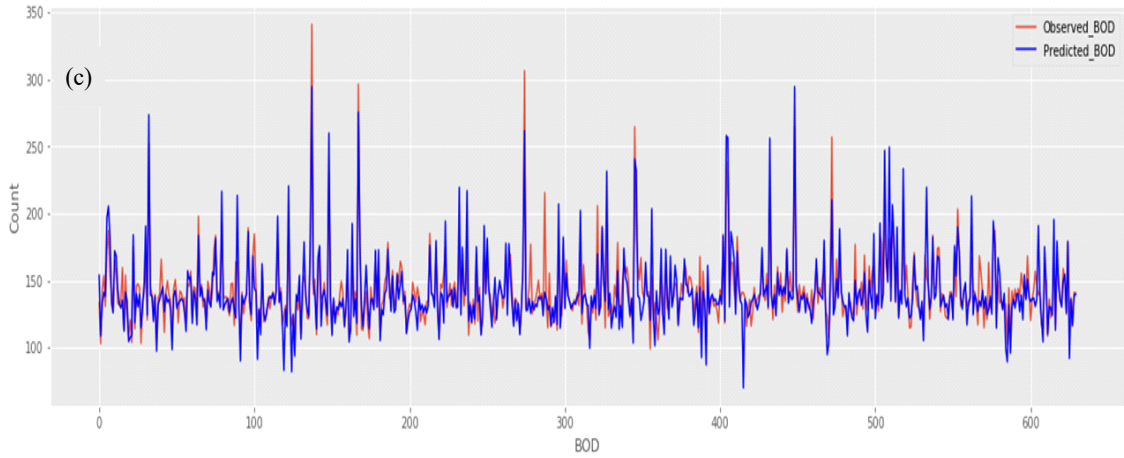
Forecasting model	R ²	RMSE
Original prediction method	0.632	20.71
DNN without derived variable	0.762	13.63
MLR without derived variable(Stepwise method)	0.712	15.60
MLR without derived variable(PCA)	0.700	15.68
DNN using derived variable	0.797	13.23
MLR using derived variable(Stepwise method)	0.775	13.55
MLR using derived variable(PCA)	0.736	14.74



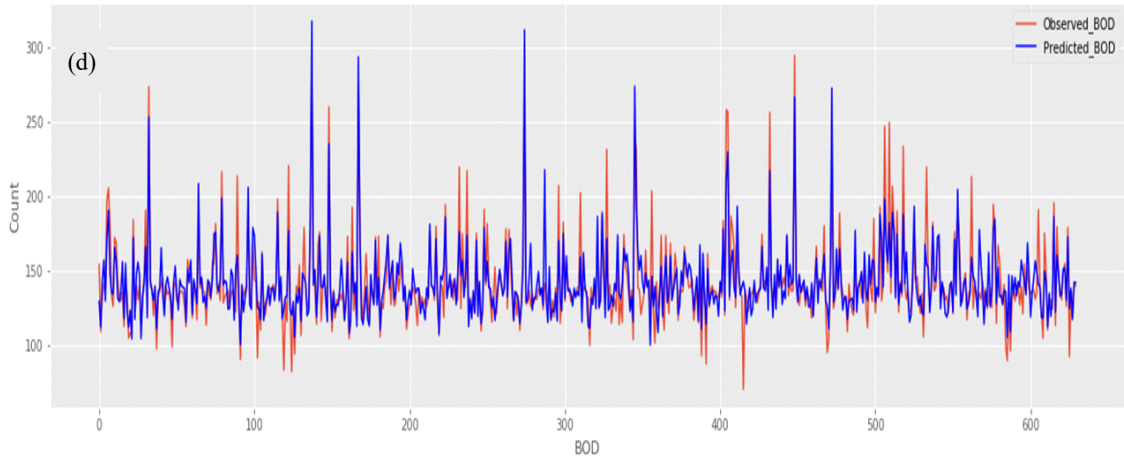
[Fig. 4a] Original prediction method.



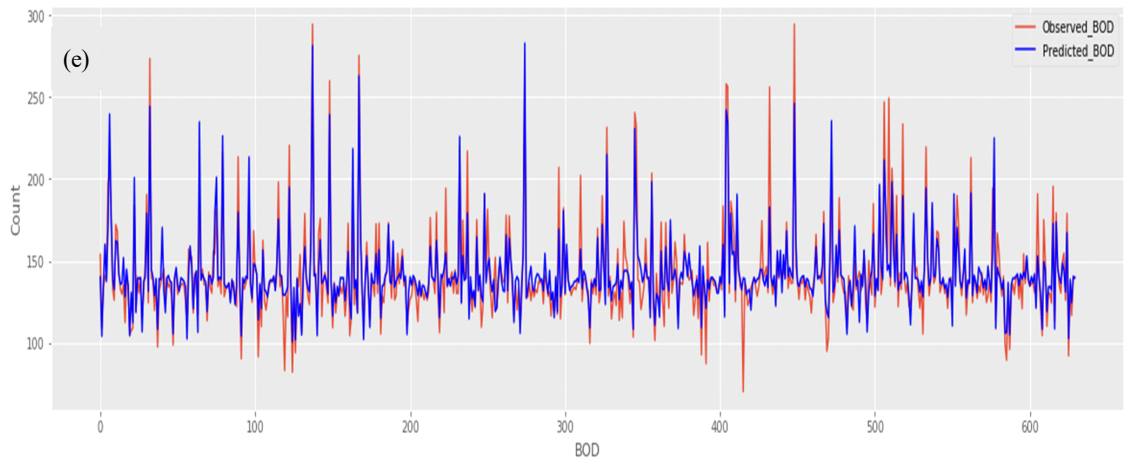
[Fig. 4b] DNN without derived variable.



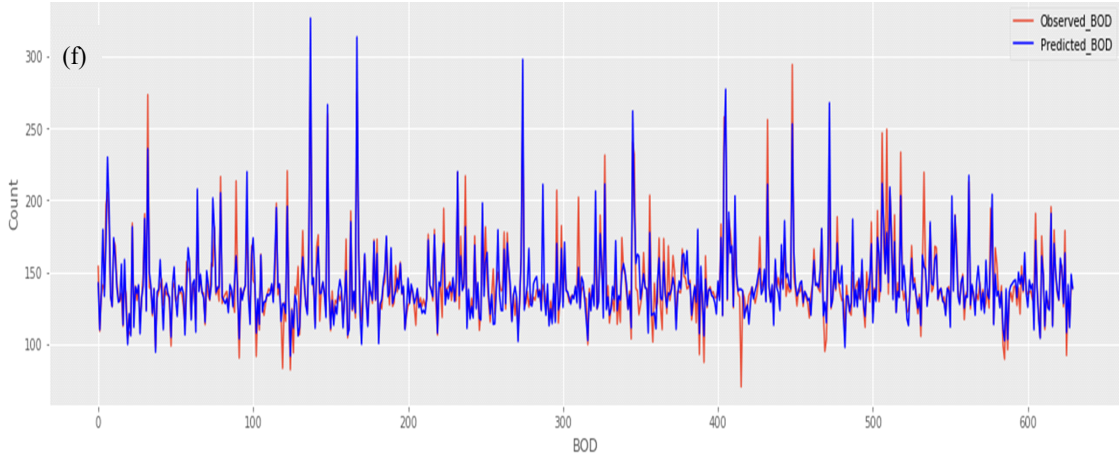
[Fig. 4c] MLR without derived variable(Stepwise method).



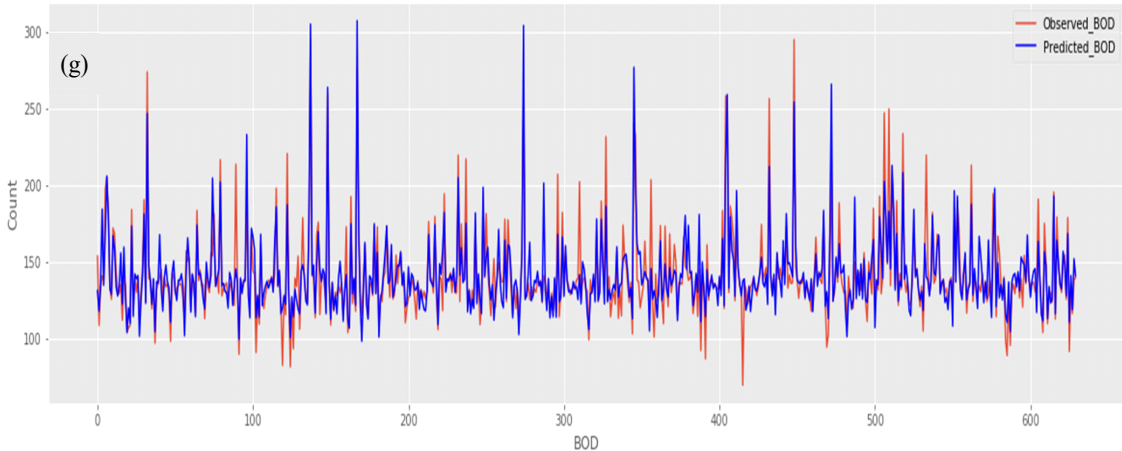
[Fig. 4d] MLR without derived variable(PCA).



[Fig. 4e] DNN using derived variable.



[Fig. 4f] MLR using derived variable(Stepwise method).



[Fig. 4g] MLR using derived variable(PCA).

IV. 결론

본 연구는 하수처리장을 대상으로 유입수의 BOD를 당일 예측하는 모델을 개발, 기존 하수처리장에서 사용하는 예측 방법보다 높은 예측력을 바탕으로 하수공정처리의 다양한 제어 전략을 사전에 제시하고 대응하여 공정과정의 효율성을 높이고 사후대처가 아닌 사전 감지 및 예방을 통한 법적 기준 초과 방류 문제를 해결할 수 있는 도구로 이용하고자 하였다. 이를 위해 2015년 1월 1일부터 2020년 9월 30일까지 B시 3개 하수처리장에서 측정된 2,100개의 일별데이터를 바탕

으로 연구를 수행하였다. 유입수의 BOD를 결정짓는 유입 성상들의 복잡한 거동을 예측하기 위해 비선형함수로 이루어진 DNN 모델을 이용하였고 이를 검증하기 위해 단계 선택법과 주성분 분석을 각각 적용한 MLR 모델과의 R^2 과 RMSE를 통한 모델 검정을 실시하였다. 또한, 유입수가 연속해서 들어오는 성질을 통해 직전 과거의 수질 성상이 당일 유입수의 BOD와 밀접한 상관관계가 있다고 판단되었으며, 이에 따라 하루 전과 이틀 전의 수질 성상을 파생변수로 추가하여 모델에 적용하였고, 파생변수의 유효성을 검증하기 위해 파생변수를 추가하지 않은 데이터셋을 적용

한 모델과의 예측력을 비교 및 평가하였다.

연구 결과, DNN 모델이 기존 하수처리장 예측 방법 및 MLR 모델보다 뛰어난 성능을 보이는 것을 증명하여 하수처리장에 비선형 모델이 적용 되었을 때 공정 효율을 높일 수 있다는 가능성을 제시했다. 또한, 과생변수를 추가하였을 때가 추가하지 않았을 때보다 DNN, MLR 모델 모두 성능이 개선된 것을 확인하여 과생변수가 유효하다는 것을 검증했다.

References

- Abdoli MA, Nezhad MF, Sede RS and Behboudian S(2011). Longterm Forecasting of Solid Waste Generation by the Artificial Neural Network, *Environmental Progress & Sustainable Energy*, 31(4), 628~636.
<https://doi.org/10.1002/ep.10591>
- Choi GI and JeongYJ(2020). Efficient iris recognition using deep-learning convolution neural network(CNN), 15(3), 521~526.
<https://doi.org/10.13067/JKIECS.2020.15.3.521>
- Dogan E, Ates A, Yilmaz EC and Eren B(2008). Application of Artificial Neural Networks to Estimate Wastewater Treatment Plant Inlet Biochemical Oxygen Deman, *Environmental Progress*, 27(4), 439~446.
<https://doi.org/10.1002/ep.10295>
- Hamed MM, Khalafallah MG and Hassanien EA (2004). Prediction of wastewater treatment plant performance using artificial neural networks, *Environmental Modelling & Software*, 19(10), 919~928.
<https://doi.org/10.1016/j.envsoft.2003.10.005>
- Hamoda MF, Al-Ghusain IA and Hassan AH(1999). Integrated wastewater treatment plant performance evaluation using artificial neural networks, *Water Science and Technology*, 40(7), 55~65.
[https://doi.org/10.1016/S0273-1223\(99\)00584-3](https://doi.org/10.1016/S0273-1223(99)00584-3)
- Hong YST, Rosen MR and Bhamidimarri R(2003). Analysis of a municipal wastewater treatment plant using a neural network based pattern analysis, *Water Research*, 37(7), 1608~1618.
[https://doi.org/10.1016/S0043-1354\(02\)00494-3](https://doi.org/10.1016/S0043-1354(02)00494-3)
- Jeong HS, Lee SH, Shin HS and Song EY(2006). Analysis and Prediction of Sewage Components of Urban Wastewater Treatment Plant Using Neural Network, *Journal of Korean Society of Environmental Engineers*, 28(3), 308~315.
- Nasr MS, Moustafa MAE, Sief H AE and Kobrosy GEI(2012). Application of Artificial Neural Network (ANN)for the prediction of EL-AGAMY wastewater treatmentplant performance-EGYPT, *Alexandria Engineering Journal*, 51(1), 37~43.
<https://doi.org/10.1016/j.aej.2012.07.005>
- Noori R, Abdol iMA , Ghazizade MJ and Samieifard R(2009). Comparison of Neural Network and Principal ComponentRegression Analysis to Predict the Solid Waste Generation in Tehran, Iranian *Journal of Public Health*, 38(1), 74~84.
- Zade JG and Noori MR(2008). Prediction of Municipal Solid Waste Generation by Use of Artificial Neural Network: A Case Study of Mashhad, *International Journal of Environmental Research*, 2(1), 13~22.

-
- Received : 16 January, 2026
 - Revised : 09 February, 2026
 - Accepted : 16 February, 2026