2nd International Joint Conference on Water Distribution
Systems Analysis & Computing and Control in the Water Industry
Valencia (Spain), 18-22 July 2022
doi: https://doi.org/10.4995/WDSA-CCWI2022.2022.14777

DEVELOPMENT OF PREDICTION MODEL OF OZONE DOSAGE AND RESIDUAL OZONE CONCENTRATION USING MACHINE LEARNING METHODS IN OZONE PROCESS OF DRINKING WATER TREATMENT PROCESS

J.Y.Kwon[1], J.S.Hyung[2], T.H.Kim[3], H.K.Park[4] and J.Y.Koo[5]

[1,2,3,4,5]Department of Environmental Engineering, University of Seoul
Seoul, Dongdaemun-gu, (Republic of Korea)

[1] *mattis94@naver.com*, [2] *gudwlstjr@uos.ac.kr*, [3] *kimth0712@uos.ac.kr*, [4] *phk940904@uos.ac.kr*,
[5] *jykoo@uos.ac.kr*

Abstract

The ozone process, which is the latter process of the water purification process, injects ozone to remove taste odor substances from tap water. Still, it is difficult to work the ozone process due to recent changes in water quality, such as taste and odor substances due to climate change. Therefore, this study developed an ozone injection rate determination model and a residual ozone concentration prediction model to properly remove flavor odor substances from raw water and proposed an operational diagnosis and optimal decision-making method for the ozone process in water purification. An ozone injection rate determination model and a residual ozone concentration prediction model were developed using data on water quality, flow rate, and operating conditions measured at Seoul's Y water purification plant. Two models were developed: the random forest and the MLP models. The performance difference between the two was verified by comparing the correlation coefficient and error index. Bayesian optimization, a global search method within a given composition space, was used to determine hyperparameters for each model. RMSE was selected as an objective function to determine the optimal hyperparameter through cross-validation. If the above model is applied to the ozone process, it is expected that an immediate response to changes in raw water quality and human error prevention will be possible.

Keywords
Machine Learning, Random Forest, MLP, Bayesian Optimization, Ozone Injection Rate, Residual Ozone Concentration.

1 INTRODUCTION

The water treatment process produces tap water essential for universal activities such as daily human life, economic activities, and industrial activities. Because of this, it should always produce safe tap water. However, due to abnormal phenomena such as torrential rain and algae caused by climate change, the water intake facility's water quality change has intensified. This has increased the difficulty in operating the water treatment process.

Due to an improper ozone injection following the inflow of taste odor substances, it may be possible for some taste odor substances to remain in tap water with the ozone process located at the rear of the water purification process. Therefore, this study developed an ozone injection rate determination model and a residual ozone concentration prediction model to properly remove flavor odor substances from raw water and proposed an operational diagnosis and optimal decision-making method for the ozone process in water purification.

2 MATERIAL AND METHOD

2.1 Current status of the subject

As shown in Figure 1, the Y water purification plant, which takes water from the P water intake plant and produces and supplies purified water, was selected as the study target area. In this plant, ozone is continuously injected at a concentration of 0.5 mg/L or less, and when the concentration of Geosmin and 2-MIB increases in such high temperatures in summer or low water temperatures in winter, ozone is injected in proportion to the concentration of odor substances.

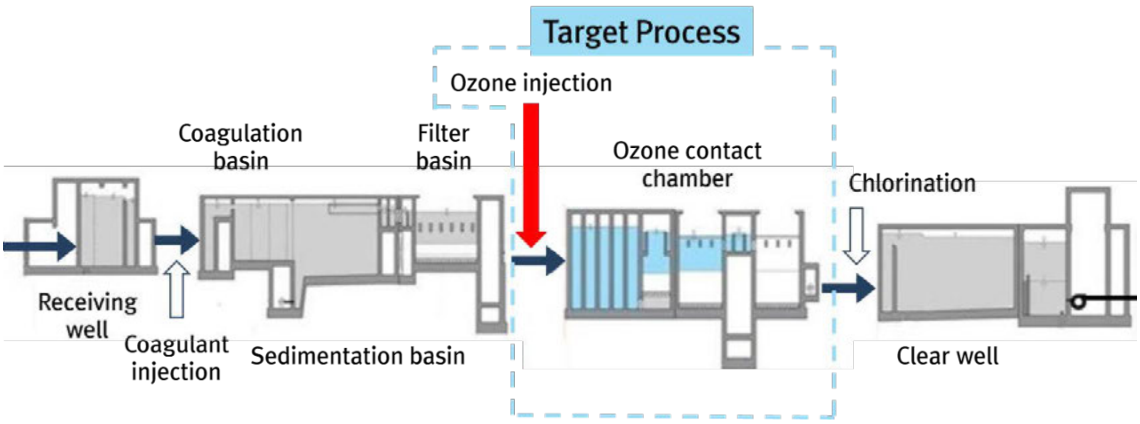


Figure 1. Target water treatment processes in Y water treatment plan

2.2 Configuration of collected data

The Y water purification plant data used in the study, which contains the data of ozone injection, residual ozone concentration, hydrogen peroxide injection, Geosmin injection, and 2-MIB injection, were collected from January 1, 2014, to December 31, 2015, and from January 1, 2017, to November 30, 2018. In the raw water of the study's target water purification plant, 2-MIB concentration reached 269.3 ng/L from 2014 to 2018, and Geosmin concentration reached 80 ng/L. The resulting ozone injection rate was operated up to 2.0 mg/L.

As shown in Figure 2, the change in ozone injection rate for the flavor odor substance concentration increases by injecting more ozone as the 2-MIB concentration and the Geosmin concentration increase. When the 2-MIB concentration increases, more ozone is injected than the Geosmin concentration. The decomposition rate of ozone varies depending on the water temperature, pH, and concentration of dissolved organic matter in water. This makes it difficult to control the proper injection rate and concentration. The ozone injection rate is determined by the concentration of the raw water's flavor odor substance. Still, excessive ozone may be injected to ensure safe water quality, thereby generating residual ozone. As residual ozone causes facility corrosion and impairment of the working environment, it must be controlled.

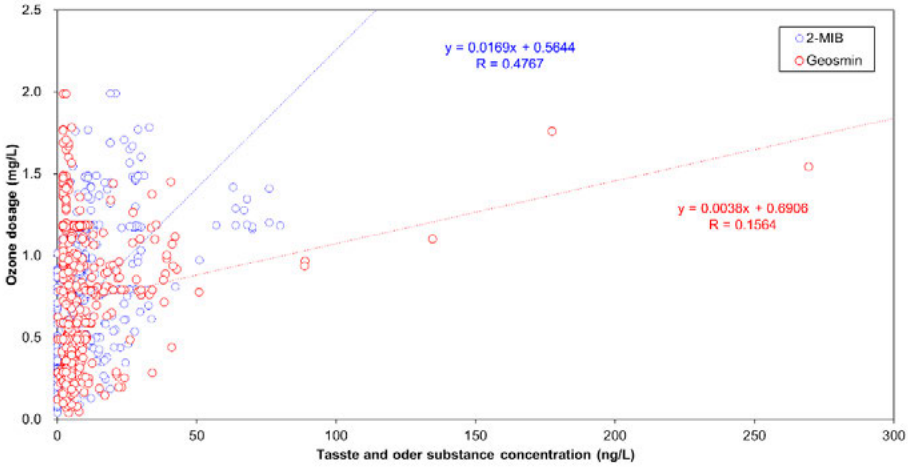


Figure 2. Trend analysis between taste and odor substansces

2.3 Establishment of ozone injection rate determination model and residual ozone concentration prediction model data

As input variables for developing the ozone injection rate model, the concentration of raw water flavored substances such as 2-MIB concentration and Geosmin concentration, raw water temperature, raw water TOC, filtered water turbidity, filtered water pH and ozone contact time, 2-MIB and Geosmin removal rate in the ozone process were considered as input variables. A model for determining ozone injection rate was developed using 32 data of 2-MIB concentration and Geosmin concentration collected during 2014, 2017, and 2018 for raw water, filtered water, ozone-treated water, and activated carbon filtered water. When constructing the ozone injection rate determination model data, learning data and verification data were sorted according to the size of the concentration of taste odor substances generated per day for the final prediction. The learning and verification data were learned with 32 datasets, and predictions were conducted with 6 datasets.

In the case of constructing residual ozone concentration prediction model data, 32 data were used to classify learning and verification data, as in the case of creating an ozone injection rate determination model.

Before developing an artificial intelligence model, a data preprocessing process, which removes and supplements outliers and missing values after data construction, was performed. Then, using the model developed in this study, a prediction period was selected to determine the rate of flavor odor removal by ozone injection, and the data for that period were utilized as prediction data. The rest of the period was used as data for learning and verifying the model.

2.4 Hyperparameter Optimization

As input variables for developing the ozone injection rate model, the concentration of raw water flavored substances In this study, Bayesian optimization was performed to select the optimal hyperparameters of the machine learning model. Bayesian optimization starts with any combination of hyperparameter composition spaces. Then, if the performance improves, change the hyperparameter combination to maintain the combination and, if not, revert to the previous combination. The most significant advantage of Bayesian optimization is that it determines hyperparameters by reflecting prior knowledge, and it generally outperforms random exploration (Bergstra et al., 2011; Bergstra et al., 2013; Falkner et al., 2018; Li et al., 2017).

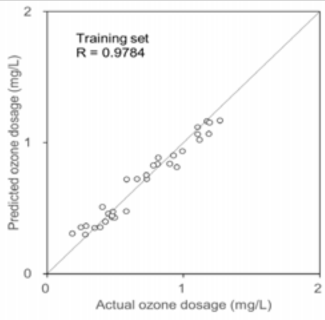
3 RESULTS AND DISCUSSION

3.1 Development of a Model for Determining Ozone Injection Rate

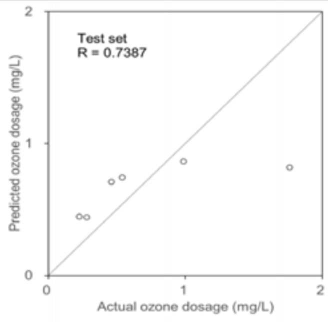
The ozone injection rate determination model was developed by applying optimal hyperparameters and dividing them into learning and verification data for 32 datasets. Based on 6 datasets, the developed model was compared with the actual ozone injection rate. The learning and verification results of the model are presented in Table 1, and the results of visualizing the correlation between the real value and the predicted value are provided in Figure 3. The correlation coefficient was shown as 0.9424 between learning data and model estimation data. In addition, 6 datasets not used for learning were used to predict the ozone injection rate and showed that the OZ-MLP Model had a higher number than the OZ-RF Model. Accordingly, it was found that the OZ-MLP Model is the optimal ozone injection rate determination model with the best performance.

Table 1. Performance evaluation results of ozone dosage determination model using RF, MLP

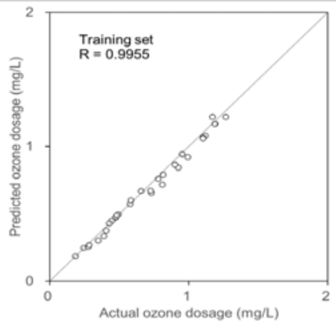
Model	PI	Model performance	
		Training (n=32)	Test (n=6)
OZ-RF	R	0.9784	0.7387
	MAE	0.06 mg/L	0.32 mg/L
OZ-MLP	R	0.9955	0.9424
	MAE	0.03 mg/L	0.15 mg/L



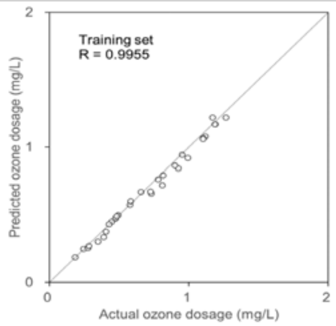
(a) RF Training result



(b) RF Test result



(c) MLP Training result



(d) MLP Test result

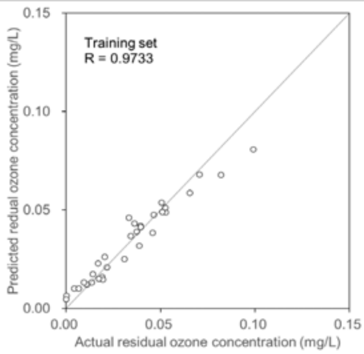
Figure 3. *Training and test results of ozone dosage determination model using RF and MLP*

3.2 Development of Residual Ozone Concentration Prediction Model

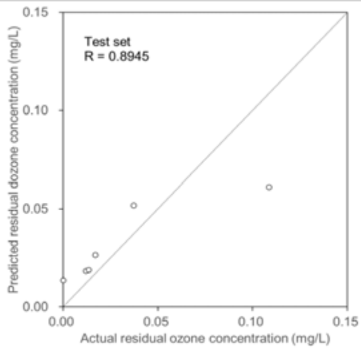
Like the ozone injection rate determination model, the residual ozone concentration prediction model was developed by dividing 32 datasets into learning and verification data by using optimal hyperparameters. Then, the developed model was used on 6 datasets and compared with the actual ozone injection rate. The learning and verification results of the model are shown in Table 2 below, and the results visualizing the correlation between the real and predicted values are shown in Figure 4. The correlation coefficient between the learning data and the estimated data was 0.9007. In addition, using 6 datasets not used for learning, the prediction of the ozone injection rate was made. It showed that the RO-MLP model had a higher number than the RO-RF model. Accordingly, it was derived that the RO-MLP model is the optimal residual ozone concentration prediction model with the highest performance.

Table 2. Performance evaluation results of ozone concentration prediction model using RF, MLP

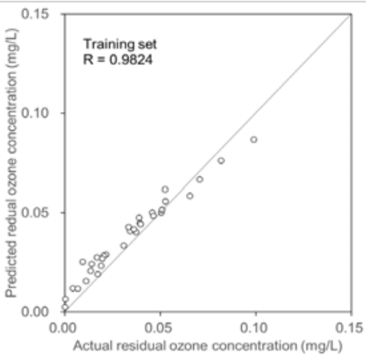
Model	PI	Model performance	
		Training (n=32)	Test (n=6)
RO-RF	R	0.9733	0.8945
	MAE	0.005 mg/L	0.016 mg/L
RO-MLP	R	0.9824	0.9007
	MAE	0.006 mg/L	0.028 mg/L



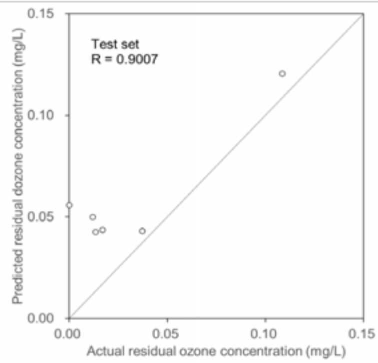
(a) RF Training result



(b) RF Test result



(c) MLP Training result



(d) MLP Test result

Figure 4. *Training and test results of residual ozone concentration prediction model using RF and MLP*

3.3 Comparison with previous studies

As for the ozone process, empirical modeling by experimental and statistical methods was conducted, as shown in Table 3. Still, there were limitations, such as not using data sets on raw water quality or considering Geosmin concentration among flavor-smelling substances.

Table 3. Ozone dosage and residual concentration modeling studies

Study	Input	Output	Model	Best performance
Cromphout et al. (2013)	CT	Ozone dosage	Empirical model	R2=0.86
Hyung et al. (2017)	Residual chlorine, temperature, conductivity, alkalinity, Geosmin concentration	Ozone dosage	Linear regression	(ordinary) R:0.858 (taste and odor substances) R:0.830
		residual ozone concentration		(ordinary) R:0.858 (taste and odor substances) R:0.856

Therefore, in this study, 2-MIB data, a taste-smelling substance other than Geosmin, was also considered, and raw water quality data according to seasonal changes were used to supplement the limitations. In addition, machine learning was applied instead of the regression model, which is a traditional model technique. The R values in previous studies were 0.83~0.858, in contrast to the R values that are 0.9007~0.9424. This increased the predictive power of this study.

4 CONCLUSION

This study developed a machine learning-based model that can perform the optimal operation and decision-making in the ozone process during water purification. The development model developed an ozone injection rate determination model with 2-MIB concentration, Geosmin concentration, water temperature, TOC, filtered water turbidity, filtered water pH, ozone contact time, 2-MIB and Geosmin removal rate in the ozone process as input variables, and a residual ozone concentration prediction model. The optimal ozone injection rate determination model is OZ-MLP Model whose correlation coefficient is 0.9424 based on the predicted data, MAE 0.028 mg/L. And, The optimal residual ozone concentration prediction model is RO-MLP Model which correlation coefficient is 0.9007 based on the predicted data, MAE 0.15 mg/L. For the high quality of tap water, the ozone injection rate for 100% removal of flavor odor substances was determined by the model through treatment up to the ozone process. Based on the model application, it was found that additional ozone should be injected at a minimum of 0.0971 mg/L and at a maximum of 0.2681 mg/L than the actual rate of ozone injection every day. Using the model and methodology developed in this research, it is judged that it is possible to determine the injection rate of hydrogen peroxide by predicting the appropriate ozone injection rate for removing flavor odor substances and the residual ozone concentration. When the above-developed model is applied within the actual process, it can be used as the basis for smart ozone process operation, such as supporting water supply operators' decision-making and responding to rapid changes in raw water. In addition, it is expected that the model's performance can be improved if additional filtered water data or additional taste odor substance measurement data after undergoing a water purification process on the ozone process leaflet are collected in the future.

5 ACKNOWLEDGEMENTS

This subject is supported by Korean Ministry of Environment as Global Top Project (2016002120006).

6 REFERENCES

[1] Bergstra, J., Bardenet, R., Bengio, Y. & Kegl, B, "Algorithms for hyper-parameter optimization," In Advances in neural information processing systems 24, 2011, pp. 2546-2554.

[2] Cromphout, J., Goethals, S. & Verdickt, L, ", Optimization of the ozone dosage at the drinking water treatment plant of Kluizen," Water Science & Technology : Water Supply, Vol.13, 2013, pp. 1569-1575.

[3] Hyung, J., Kim, K., Kim, M., Lee, I. & Koo, J, "A study on prediction method for ozone dosage and residual ozone concentration in advanced ozone water tratment," Water Practice & Technology, Vol.12, 2017, No.1 pp. 87-96.