Desalination and Water Treatment
www.deswater.com

doi: 10.5004/dwt.2023.29771

303 (2023) 48–58
August

Development of a water quality prediction model using ensemble empirical mode decomposition and long short-term memory

Sukmin Yoon[a], Chi Hoon Park[b], No-Suk Park[c], Beomsu Baek[d], Youngsoon Kim[e,*]

[a]*Gyeongsang National University Office of Academy and Industry Collaboration and Engineering Research Institute, 501, Jinju-Daero, Jinju 52828, Republic of Korea*
[b]*Department of Energy Engineering, Future Convergence Technology Research Institute, Gyeongsang National University, 501, Jinju-Daero, Jinju 52828, Republic of Korea*
[c]*Department of Civil Engineering and Engineering Research Institute, Gyeongsang National University, 501, Jinju-Daero, Jinju 52828, Republic of Korea*
[d]*Department of Information and Statistics, Gyeongsang National University, 501, Jinju-Daero, Jinju 52828, Republic of Korea*
[e]*Department of Information and Statistics and Department of Bio & Medical Bigdata (BK21 Four Program), Gyeongsang National University, 501, Jinju-Daero, Jinju 52828, Republic of Korea, email: youngsoonkim@gnu.ac.kr*

Received 24 May 2023; Accepted 14 July 2023

ABSTRACT

Water distribution systems consistently supply high-quality water at suitable pressure and volume for human and industrial consumption. Meticulous water quality management is vital to these systems. South Korea, having established legal standards for water distribution in 1963, operates the National Auto Water Quality Monitoring System for real-time water quality monitoring and contamination warnings when levels exceed legal thresholds. The U.S. Environmental Protection Agency (EPA) points out that fixed thresholds can trigger an abundance of false-positive alarms, causing irregular hydraulic changes, and false-negative errors. This could potentially lead to a failure in detecting initial instances of pollution or micropollution that fall below the established threshold. To address this, our study developed an proactive contamination warning method for South Korea's monitoring system, utilizing long short-term memory (LSTM) for water quality prediction. We also employed ensemble empirical mode decomposition (EEMD) in feature engineering to enhance LSTM's prediction performance. Additionally, we devised an optimal water quality prediction model development methodology by comparing short- and long-term prediction performances. Our findings revealed that using EEMD for feature engineering improved the stability and reduced the prediction lag of LSTM, outperforming traditional methods. This refined approach offers a more reliable and efficient means of monitoring and managing water quality in distribution systems.

Keywords: Contamination warning; Ensemble empirical mode decomposition; Feature engineering; Long short-term memory; Water distribution system; Water quality

* Corresponding author.