Performance comparison of outlier detection methods in smart water grid: A case study of YeongJong Island, Korea

KANG MIN KOO[1], KUK HEON HAN[2], KYUNG SOO JUN[1], KYUNG TAEK YUM[*1]

[1]*Sungkyunkwan University*, [2]*National Smart Water Grid Research Group*

(a) Purpose of study or research hypothesis

Comparison of outlier ?detection methods for water consumption data collected from smart water meter in Smart Water Grid (SWG)

(b) Key issue(s) or problem(s) addressed

- Water consumption data is the core for SWG operation.
- But the oulier contained in the dataset could affect demand forecasting, making it less reliable and less accurate
- To increase the reliability of the data, the Outlier detection method suitable for water consumption big data must be selected

(c) Methodology or approach used

- Selected 8 models most commonly used in outlier detection: Grupps's test (Generalized Extreme Studentized Deviate), Gaussian distribution model, Markov chain-based model, Isolation model (Isolation forest), Neighbor-based model (K-NN), and Neural networks (One-class SVM, k-mean clustering, Long Short-term Memory)
- Applied these models to hourly accumulated water consumption data collected in 2018-2019 (2 years) from 527 consumers of Yeongjong Island, Korea
- Evaluation indices: Statistical normality test (p-value), rate of ouliers, detection time, and kurtosis and skewness

(d) Results or conclusions derived from the project

- It is possible to present a reasonable model for optimal water supply system (WSS) management objectives and to adequately detect outliers in water consumption data sets.

?

(e) Implications of the project relevant to congress themes

- Secure the reliability of big data, which is the most important to increase the efficiency of the WSS using SWG technology

Keywords : Outlier detection, SWG, Water consumption, YeongJong Island