Desalination and Water Treatment
www.deswater.com

doi: 10.5004/dwt.2017.20681

77 (2017) 156–162
May

Addressing water pollution hotspots in the tributary monitoring network using a non-linear data analysis tool

Seo Jin Ki[a], Sanghwan Song[b], Tae-Woo Kang[b], Sangdon Kim[b], Taegu Kang[b], Seung Gwon Baek[b], Jong Hun Baek[b], Joon Ha Kim[a,*]

[a]*School of Environmental Science and Engineering, Gwangju Institute of Science and Technology (GIST), 123 Cheomdangwagi-ro, Buk-gu, Gwangju 61005, Korea, email: joonkim@gist.ac.kr*
[b]*Yeongsan River Environment Research Center, National Institute of Environmental Research, 5 Cheomdangwagi-ro 208beon-gil, Buk-gu, Gwangju 61011, Korea*

Received 14 November 2016; Accepted 1 January 2017

ABSTRACT

Successful data analysis is an essential component of any environmental monitoring programs. This study introduces an effective data analysis method to identify water pollution hotspots as well as to drop redundant monitoring parameters and samples using a self-organizing map (SOM), which has a strong specialty in pattern extraction from complex monitoring data. A full data set consisted of nine parameters that were obtained on a monthly basis from 83 sites in various tributary streams along the Yeongsan River, Korea, from May 2011 to December 2015. The given data set was further partitioned into a number of subsets to examine their effect on variable importance and temporal pattern analysis. We found that water pollution hotspots were more clearly addressed in load-based SOM analysis than in concentration-based SOM analysis due to strong correlation between variables resulted from variability reduction by combining two variables into a single one for load analysis. In addition, the variables chemical oxygen demand and electrical conductivity and the parameters discharge and total nitrogen were found to participate most and least actively in describing spatial and temporal variation of the observed variables, respectively. About 35% of the sampling locations showed high similarity among the monthly data extending from November in the previous year to February in the following year. We believe that the proposed methodology can be useful in revising the upcoming water monitoring study by clarifying several issues related to monitoring parameters and frequency in the existing program.

Keywords: Self-organizing map; Tributary monitoring; Water pollution hotspots; Sampling frequency; Temporal variability; Variable selection

* Corresponding author.