

## Manuscript Preview

### Manuscript Information

Journal	Water
Article Type	Article
Title	Evaluating statistical machine learning algorithms for classifying dominant algae in Juam Lake and Tamjin Lake, Republic of Korea
Abstract	<p>South Korea's National Institute of Environmental Research (NIER) operates an algae alert system to monitor water quality at public water supply source sites. Accurate prediction of dominant harmful cyanobacteria genera, such as Aphanizomenon, Anabaena, Oscillatoria, and Microcystis, is crucial for managing water source contamination risks. This study utilized data collected between January 2017 and December 2022 from Juam Lake and Tamjin Lake, which are representative water supply source sites in the Yeongsan River and Seomjin River basins. We performed an exploratory data analysis on the water quality parameters monitored to understand overall fluctuations. Using data from 2017 to 2021 as training data and 2022 data as test data, we compared the dominant algae classification accuracy of 11 statistical machine learning algorithms. The results indicated that the optimal algorithm varied depending on the survey site and evaluation criteria, highlighting the unique environmental characteristics of each site. By predicting dominant algae in advance, we can better prepare for water source contamination accidents. Our findings demonstrate the applicability of machine learning algorithms as efficient tools for managing water quality in water supply source systems using monitoring data.</p>
Keywords	water quality; Yeongsan River; Seomjin River; correlation analysis; self-organizing map; statistical machine learning algorithm; classification
Number of Pages	35

[Go Back to Edit](#)

### Author Information

Number of Authors	10
	<b>Author 1</b>
E-Mail Address	hsyliark@korea.kr
Name	Seong-Yun Hwang
	<b>Author 2</b>
E-Mail Address	bchoi628@korea.kr
Name	Byung-Woong Choi
	<b>Author 3</b>
E-Mail Address	thanks@korea.kr
Name	Jong-Hwan Park
	<b>Author 4</b>
E-Mail Address	sds8488@korea.kr
Name	Dong-Seok Shin
	<b>Author 5</b>
E-Mail Address	jys7246@korea.kr
Name	Hyeon-Su Chung
	<b>Author 6</b>
E-Mail Address	miza03@korea.kr
Name	Mi-Sun Son

**Author 7**

E-Mail Address chaehong@korea.kr

Name Chae-Hong Lim

**Author 8**

E-Mail Address chm2022@korea.kr

Name Hyeon-Mi Chae

**Author 9**

E-Mail Address hahaha9909@korea.kr

Name Don-Woo Ha

**Author 10**

E-Mail Address happy3313@korea.kr

Name Kang-Young Jung \*

Author Contributions Conceptualization, Seong-Yun Hwang and Kang-Young Jung; Methodology, Seong-Yun Hwang and Byung-Woong Choi; Software, Seong-Yun Hwang; Validation, Kang-Young Jung; Investigation, Hyeon-Su Chung, Mi-Sun Son, Chae-Hong Lim, Hyeon-Mi Chae and Don-Woo Ha; Resources, Hyeon-Su Chung, Mi-Sun Son, Chae-Hong Lim, Hyeon-Mi Chae and Don-Woo Ha; Data curation, Seong-Yun Hwang; Writing – original draft, Seong-Yun Hwang; Writing – review & editing, Seong-Yun Hwang and Kang-Young Jung; Visualization, Seong-Yun Hwang and Kang-Young Jung; Supervision, Byung-Woong Choi, Jong-Hwan Park and Dong-Seok Shin.

[Go Back to Edit](#)

**Manuscript Files**

Manuscript (Word/ZIP) manuscript.v1.docx

Figures, Graphics, Images [figures.v1.zip](#)

Coverletter (PDF) coverletter.v1.pdf

Funding Information National Institute of Environmental Research (NIER): NIER-2023-01-01-043

[Go Back to Edit](#)