Desalination and Water Treatment
www.deswater.com
doi: 10.5004/dwt.2022.29069

279 (2022) 61–67
December

A combined water quality classification model based on kernel principal component analysis and machine learning techniques

Smail Dilmi

Laboratory of Analysis of Signals and Systems, Department of Electrical Engineering, University of M'sila, 28000 M'sila, Algeria, email: dilmid06@gmail.com

Received 7 June 2022; Accepted 2 October 2022

ABSTRACT

Water quality monitoring plays an essential role in environmental management and the protection of water resources. However, the increasing risks of pollution make the process of monitoring using conventional methods more complex and costly. Currently, the use of automated processes based on artificial intelligence and machine learning techniques has become necessary in the field of water quality control to achieve quality control and reduce operating costs. This paper presents a comparative study of three machine learning techniques, namely K-nearest neighbors (KNN), decision tree (DT), and support vector machine (SVM), for the water quality classification of Tilesdit Dam (Algeria). Furthermore, the kernel principal component analysis (KPCA) technique was utilized to choose the important variables for water quality classification. The models were trained and tested based on historical data collected from the dam monitoring station for 3 y (2016–2018). The results of the study indicated that a combination of KPCA and DT techniques gave the best performance, with a classification accuracy of 99.68%.

Keywords: Water quality classification; Machine learning techniques; Kernel principal component analysis; Variables selection

Presented at the First International Congress of Energy and Industrial Process Engineering (ICEIPE'22) 23–25 May 2022, Algiers, Algeria