III Sanitary Engineering and Sustainable Water Use

PREDICTION OF WATER QUALITY INDEX FOR PROCESSES IMPROVEMENT ON DRINKING WATER TREATMENT PLANT

GORAN VOLF[1], IVANA SUŠANJ ČULE[1], ELVIS ŽIC[1], SONJA ZORKO[2]

[1] University of Rijeka, Faculty of Civil Engineering; Croatia
e-mail: goran.volf@uniri.hr; isusanj@uniri.hr; elvis.zic@uniri.hr

[2] Istarski Vodovod d.o.o. Buzet; Croatia
e-mail: sonja.zorko@ivb.hr

Abstract

To improve drinking water treatment plant processes on the Butoniga reservoir in Istria (Croatia), the prediction of the Water Quality Index (WQI) is performed.

Based on parameters such as temperature, pH, turbidity, $KMnO_4$, NH_4, Mn, Al and Fe, the calculation of WQI is conducted, while for the prediction models of WQI along with mentioned parameters, additionally O_2, TOC and UV254 are used.

All data are pre-processed concerning modeling and research goal. For the WQI prediction models, the entire span of the measured daily data was used, from 2011 to 2020, while missing data were filled in by usage of cubic spline interpolation.

Four models were built to predict WQI with a time step of one-, five-, ten-, and fifteen days in advance to improve certain processes in the drinking water treatment plant regarding changes in raw water quality in the Butoniga reservoir which could influence the treatment processes. Therefore, these models can help optimize certain processes depending on the quality of raw water.

For the experiment, the machine learning algorithm M5P for rule-based models integrated into the WEKA modeling software was used. Predicted WQI (one, five, ten-, and fifteen-days' time step) was set as a target (dependant) variable, whereas water temperature, pH, turbidity, $KMnO_4$, NH_4, Mn, Al, Fe, O_2 TOC and UV254 were set as independent variables (descriptors) from which the predicted WQI was modeled. The above parameters were used because they represent the parts of the system (drinking water treatment plant) on top of which the target variables depend.

To achieve the highest correlation coefficient (R), and the optimal number of rules default values of parameters for building models were used in WEKA modeling software. The model that was performing most accurately, according to the validation method, was selected as a representative model for prediction purposes.

Obtained results from models showed that one-day prediction model for WQI has only one rule which is related to a linear equation predicting the WQI and a very high correlation coefficient of 0.93, while five-day prediction model is consists of 12 rules with a correlation coefficient of 0.81. Ten-day prediction model has 14 rules and a correlation coefficient of 0.79. Finally, fifteen-day prediction model has five rules and a correlation coefficient of 0.71. As expected, obtained correlation coefficient decreases as the number of prediction days increases, while the number of rules, and related linear equations depend on the parameters set in WEKA modeling software which give the highest values of the correlation coefficient. Also, all models have high accuracy compared to the measured data, with a good prediction of the peak values.

Therefore, obtained models through the prediction of WQI could help to manage certain drinking water treatment plants depending on the quality of raw water in the Butoniga reservoir.

Keywords: WQI, prediction model, machine learning, processes improvement, drinking water treatment plant, Butoniga.