25th International Congress on Modelling and Simulation, Darwin, NT, Australia, 9 to 14 July 2023
mssanz.org.au/modsim2023

Exploratory analysis of water quality in a small urbanized watershed using deep learning

A. Ofosu and M.G. Erechtchoukova

School of Information Technology, York University, Toronto, Canada
Email: marina@yorku.ca

Abstract: Water is a life-sustaining resource for living organisms inside and outside of water bodies. Natural waters serve as municipal, industrial, agricultural irrigation water supply sources, homes for aquatic ecosystems, recreation, and other essential uses. The quality of water determines water uses. Therefore, it must be monitored, managed, and reported to help stakeholders in decision-making that can protect ecosystems of watersheds and improve measures to mitigate factors adversely affecting waterbodies. Water quality is represented by a set of parameters that describe specific characteristics or properties of water. These parameters are determined by measuring water's physical and chemical characteristics and concentration levels of various substances present in a water column with subsequent sample analysis in laboratories. This results in low frequencies of observations for water quality parameters compared to hydrometric and meteorologic data. Frequencies of observation adopted by many water quality monitoring systems vary between 4 to 12 samples per year, suggesting applying modelling techniques in support of decision-making.

The study is aimed at developing a data-driven computational tool for water quality modelling in a small highly urbanized watershed of the Don River, Ontario, Canada. The study focuses on major ions, namely, cations: calcium (Ca^{2+}), magnesium (Mg^{2+}), sodium (Na^{+}), and potassium (K^{+}) and anions such as bicarbonate (HCO_3^{-}), carbonate (CO_3^{2-}), chloride (Cl^{-}), and sulphate (SO_4^{2-}). These parameters are not affected significantly by the aquatic ecosystem. Their dynamics are mainly determined by the meteorological and hydrological processes. The study uses data from different monitoring systems belonging to Toronto and Region Conservation Authority and Environment Canada (Figure 1). The investigated data set consists of water quality parameters, hydrometric and meteorologic characteristics observed on the watershed over the period of 50 years. Concentrations of selected water quality parameters are modelled using deep neural networks. The data pre-processing framework for cleansing and integration of data observed with different frequencies from different locations is developed. The framework is applied for comparative analysis of neural networks of various configurations.

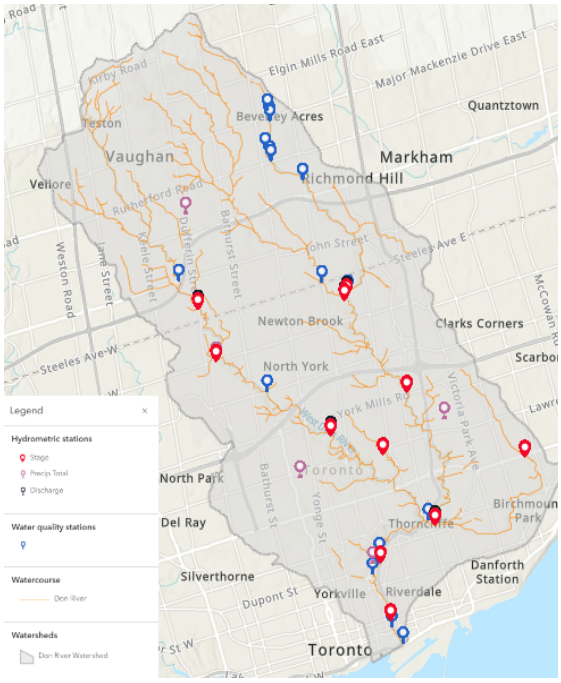

Figure 1. Monitoring sites on the Don River watershed

Two sets of computational experiments were conducted. In the first set of experiments integrated data from all monitoring stations from the Don River watershed were fed into the deep learning algorithms to train a neural network to predict the concentration of major ions for the upcoming month (t+1). The second set of experiments focuses on using upstream environmental parameters to train the model and predict the major ion concentrations in the lower basin of the watershed. The study investigates the performance of developed models in accurately predicting ion concentrations and provides insights into the relationship between environmental factors and water quality in the investigated watershed. The findings have practical applications for water resource management and pollution prevention efforts.

Keywords: *Water quality, integrated hydrology, major ions, deep learning*