EGU2020-845
https://doi.org/10.5194/egusphere-egu2020-845
EGU General Assembly 2020

Improved real-time SWMM flow forecasts using two machine learning approaches

Everett Snieder[1], Arash Shahmansouri[2], Chi-Han Cheng[2], Yuanhao Ding[2], Edward Graham[2], and Usman Khan[1]
[1]York University, Civil Engineering, Toronto, Canada (esnieder@yorku.ca)
[2]Civica Infrastructure, Vaughan, ON, Canada

The Storm Water Management Model (SWMM) is a popular and widely used physics-based numerical model for water resource management and flow forecasting. Calibrating SWMM requires a large amount of geospatial and hydro-meteorological data that may be hard to collect, has high uncertainty associated with it, and are often non-stationary. These issues are compounded when modelling large watersheds with several sub-catchments, leading to thousands of parameters that need to be calibrated collectively. The calibration process is time consuming (and often conducted manually), and results in models that are biased, only tuned to specific events, and lead to high uncertainty in the flow forecasts, and thus, limiting their utility.

In this research, a two-stage machine-learning process is proposed to first calibrate a large-scale SWMM model using a genetic algorithm (GA), and second, to bias correct the flow forecast values using an artificial neural network (ANN) ensemble to improve real-time flow forecasts.

A SWMM model for the 14 Mile Creek Watershed in Ontario, Canada is used as a case study for the proposed method. The model contains 60 sub-catchments with 10 parameters each, and a total of 1144 elements that require calibration. The model is driven by a suite of numerical weather models and precipitation estimates (including the Global Environmental Multiscale - Local Area Model, the North American Mesoscale Forecast System, and the Rapid Refresh and High Resolution Rapid Refresh models). These models have a lead time of up to 36 hours at an hourly resolution. A GA approach was implemented in MATLAB to calibrate the watershed for both single- and multi-event scenarios using a multi-criteria optimisation approach for a suite of model performance metrics (the Nash-Sutcliffe Efficiency, peak flow difference, and relative error of the total runoff volume). Historical precipitation and flow data with an hourly time-step was used in the calibration procedure.

Next, an ANN is trained using recent (i.e., 1 to 24 hour lag time) observed flow, SWMM forecast flow, and observed precipitation, to predict the SWMM bias (the difference between SWMM forecasts and flow observations). The estimated bias is used to correct the real-time SWMM forecasts which are driven by the precipitation forecasts. This bias correction procedure implicitly minimizes the collective error associated with the radar forecasts, the SWMM parameter uncertainty, and the SWMM epistemic uncertainty. Ensemble methods are employed within the

ANN to quantify the uncertainty of the bias-corrected forecast flows.

Preliminary results indicate that GA-based calibration improved the NSE from 0 to 0.75; however, some single event-based GA calibration did not maintain acceptable performance (NSE > 0.65) when cross-validated with other events. Bias corrected forecasts further improve the NSE to 0.9 for some events. A comparison between the uncalibrated, GA-calibrated, bias-corrected, and pure ANN forecasts are presented.