Modelling time-varying performance of green infrastructures through deep learning methods

Y. Yang[1] & T. F. M. Chui[1,*]

[1]*Department of Civil Engineering, The University of Hong Kong, Pokfulam, Hong Kong SAR, China*

**Corresponding author email: maychui@hku.hk*

Highlights

- Rainfall-runoff model built for a small site achieved a test Nash–Sutcliffe efficiency (NSE)=0.88.
- Proposed deep network structure allows physical parameters to be incorporated in modelling.
- The impact of seasonality and infrastructure age on GIs' hydrological performance is quantified.

Introduction

Green infrastructures (GIs), also known as sustainable urban drainage systems (SuDS) or water sensitive urban design (WSUD), are eco-friendly supplements to conventional urban drainage networks. In recent years, many numerical models have been developed to predict GIs' hydrological performances under different rainfall conditions and to study the involved hydrological processes (Elliott and Trowsdale, 2007). Most of the models are process-driven, where the hydrological processes are characterized using physically-based or empirical equations. However, these models are sometimes inapplicable due to lack of information on the physical properties of the study site or the involved hydrological processes. The performances of GIs may vary in different seasons and can be affected by infrastructure age, among many other factors (Lewellyn et al., 2016). The time-varying performances of GIs can be hard to simulate as many current models are developed for stationary systems.

Data-driven models do not explicitly require information on the physical properties or the involved processes of the studied system. They can be trained to learn the correlations between the state variables of a system given only observation data. For instance, Yang and Chui (2018) built data-driven models to predict water level and overflow occurrence of GIs using inflow and weather data. One type of data-driven models, deep learning models are especially suitable for these tasks, as they excel in modelling high-dimensional data, such as high-resolution rainfall and runoff time series. However, excluding the physical parameters from modelling can have some negative consequences. For example, it can be difficult to assess the impact of infrastructure aging, as stationary correlations are normally assumed between the input and output variables of interests. This problem may be resolved by incorporating time-varying parameters into modelling. This research aims to show that time-varying parameters can be used as inputs in deep learning models by adopting deep neural networks of specific topology. This research also attempts to illustrate that the inclusion of time-varying parameters allows the time-varying properties of the GIs to be reflected in modelling results.

Methodology

The long short-term memory neural networks (LSTM) are used in this study, as they are can capture the long-term dependency between input and output variables (Gers et al., 1999). As shown in Figure 1, the model maintains a hidden state at each time step (h_0 through h_4), which can be modified by rainfall and decoded to runoff through some functions in chronological order. The value of the state affects how the state is updated at the current time step. The initial state, therefore, may have continuing effects on the update scheme. The context variables, such as rainfall occurring season and infrastructure age, are assumed to influence the overall state update scheme and are therefore used to infer the initial state. The functions for initializing, updating, and decoding the hidden states are all represented by neural networks, whose parameters are optimized through training.

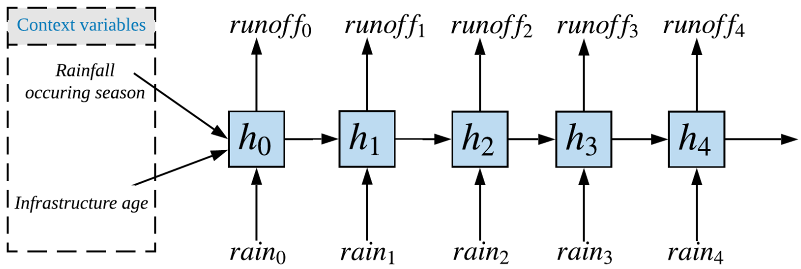

Figure 1. Structure of the recurrent neural network model. Arrows indicate flow of information and represent transformations that map random variables to other variables or hidden values. h_0 through h_4 are vectors corresponding to hidden states of the system.

Case study

The study site locates in Washington Street, Geauga County, Ohio, the U.S., where green roofs, porous pavements, and bioretention systems were implemented to treat stormwater runoff generated by a commercial lot (Darner and Dumouchelle, 2011). Rainfall and runoff at this site were monitored by the United States Geological Survey (USGS) since 2009. Hydrological property of this site was found to vary over time (Darner and Dumouchelle, 2011, Darner et al., 2015). The rainfall occurring month and the accumulated rainfall depth since 2009 (which reflects the overall workload of infrastructure) are used as context variables in this study. The 10-min rainfall and runoff data recorded between 2009 and 2013 and the context variables are divided into the training, validation (for monitoring and guiding training processes), and test sets to evaluate prediction accuracy. Individual rainfall events are extracted from continuous data using a 24-hour dry spell threshold (Guo and Senior, 2006). The influence of context variables is measured by the variations among the predicted outflow hydrographs, which are obtained by deliberately varying the values of the context variables while fixing the rainfall input.

Results and discussion

The predicted outflow hydrograph for the test set is shown in Figure 2a. The low to moderate flows were accurately modelled, and relatively small variations can be observed for the high flows. The rising and recession limps were also precisely modelled. The performance metrics also confirm the excellent prediction accuracy: NSE = 0.88 and R^2 = 0.89 for the test set, and NSE = 0.94 and R^2 = 0.94 for training set.

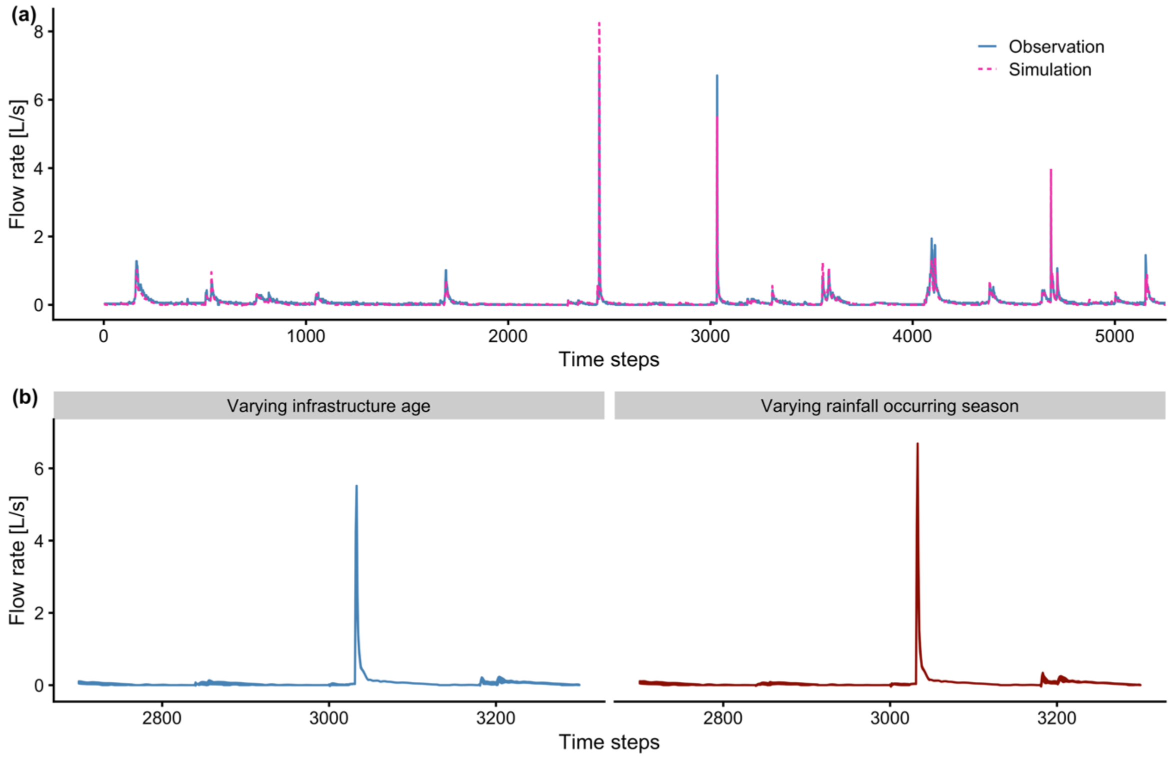

Figure 2. (a) Outflow hydrograph of the study site for the test set, and (b) variations of the predicted hydrographs when varying the context variations and fixing the rainfall input. For clarity, only part of the result is shown, and runoff events occurring in different time between 2009 and 2013 are plotted close to each other.

The variations of the predicted hydrographs are obtained by varying infrastructure age from small to large values or rainfall occurring season through all the possible values, while fixing the rainfall time input (as shown in Figure 2b). The results suggest that the context variables can have considerable impacts on predictions, especially for the low flows. Rainfall occurring season seems to have a higher impact on outflow predictions for having larger prediction variations when comparing to infrastructure age. Other context variables may be used as input and can be incorporated into neural network models in different ways (besides initializing the hidden state), these topics are worthy of further investigation. For example, the infiltration rates of GIs measured over time may be incorporated into modelling, and 1-D convolutional neural networks maybe used instead of LSTM.

Conclusions and future work

This study shows that deep learning methods can be useful tools for predicting hydrological responses of green infrastructures, especially when process-based models are inadequate due to insufficient knowledge of the studied system or the system's time-varying properties. Context information, such as infrastructure age and rainfall occurring season, can be included as input variables of deep neural network models by carefully designing the network topology. Incorporating context information in modelling potentially increases prediction accuracy and also allows the time-varying performance of green infrastructures to be reflected in the modelling results.

Similar to process-based models, the states of the modelled system at each time step are also encoded using numerical values in deep learning models. Inspecting the hidden state evolution through time and comparing the state update schemes of different types of models can potentially improve the understanding of the studied system. Further investigations on the effectiveness of adopting different neural network structures are also recommended.

Acknowledgements

This study was partly supported by a grant from the Research Grants Council of the Hong Kong Special Administrative Region, China (Project No. HKU17255516), and partly supported by the RGC Theme-based Research Scheme (Grant No: T21-711/16-R) funded by the Research Grants Council of the Hong Kong Special Administrative Region, China. We would like to thank Robert A. Darner from the U.S. Geological Survey for providing site information of the study site.

References

Darner R. A. and Dumouchelle D. H. (2011). Hydraulic Characteristics of Low-impact Development Practices in Northeastern Ohio, 2008-2010, US Department of the Interior, US Geological Survey.

Darner R. A., Shuster W. D. and Dumouchelle D. H. (2015). Hydrologic Characteristics of Low-Impact Stormwater Control Measures at Two Sites in Northeastern Ohio, 2008–13.

Elliott A. and Trowsdale S. A. (2007). A review of models for low impact urban stormwater drainage. Environmental modelling & software, 22, 394-405.

Gers F. A., Schmidhuber J. and Cummins F. (1999). Learning to forget: Continual prediction with LSTM.

Guo Y. and Senior M. J. (2006). Climate model simulation of point rainfall frequency characteristics. Journal of Hydrologic Engineering, 11, 547-554.

Lewellyn C., Lyons C. E., Traver R. G. and Wadzuk B. M. (2016). Evaluation of seasonal and large storm runoff volume capture of an infiltration green infrastructure system. Journal of Hydrologic Engineering, 21, 04015047.

Yang Y. and Chui T. F. M. (2018). Hydrologic Performance Simulation of Green Infrastructures: Why Data-Driven Modelling Can Be Useful? International Conference on Urban Drainage Modelling. Springer, 480-484.