Journal of
Hydroinformatics

Journal of Hydroinformatics Vol 25 No 5, 1667 doi: 10.2166/hydro.2023.001

Deep learning model on rates of change for multi-step ahead streamflow forecasting

Woon Yang Tan [a], Sai Hin Lai [a,b,*], Kumar Pavitra[c], Fang Yenn Teo[d] and Ahmed El-Shafie[a]

[a] Department of Civil Engineering, Faculty of Engineering, University of Malaya, Kuala Lumpur 50603, Malaysia
[b] Department of Civil Engineering, Faculty of Engineering, Universiti Malaysia Sarawak (UNIMAS), 94300, Kota Samarahan, Sarawak, Malaysia
[c] Department of Geography and Planning, University of Liverpool, Liverpool L69 3BX, UK
[d] Faculty of Science and Engineering, University of Nottingham Malaysia, Semenyih, Selangor 43500, Malaysia
*Corresponding author. E-mail: laish@um.edu.my

WYT, 0000-0002-2506-0021; SHL, 0000-0002-7143-4805

ABSTRACT

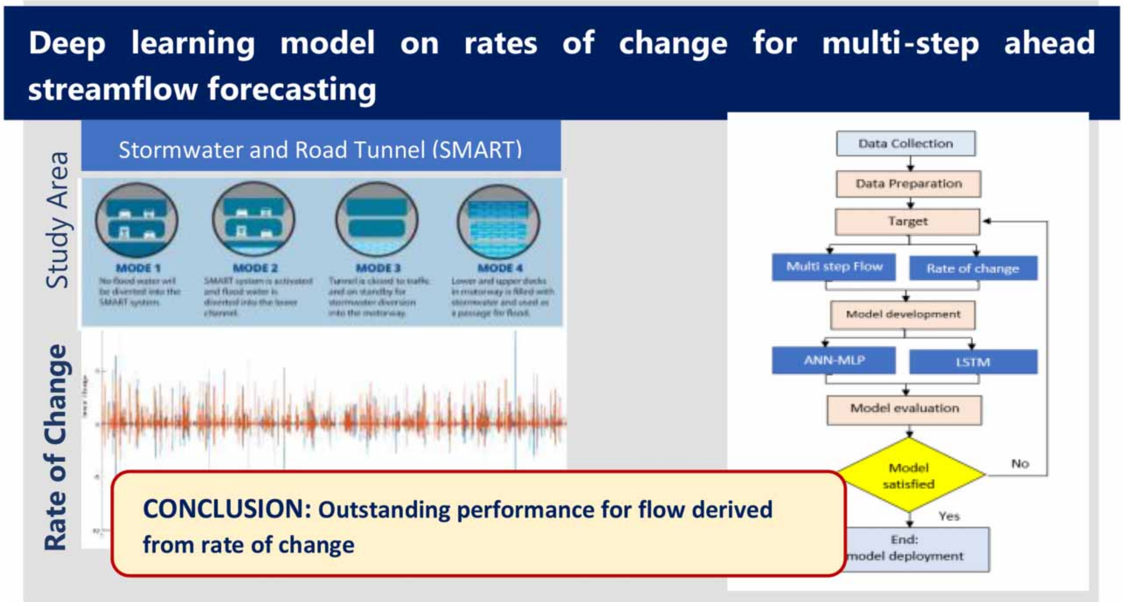
Water security and urban flooding have become major sustainability issues. This paper presents a novel method to introduce rates of change as the state-of-the-art approach in artificial intelligence model development for sustainability agenda. Multi-layer perceptron (MLP) and deep learning long short-term memory (LSTM) models were considered for flood forecasting. Historical rainfall data from 2008 to 2021 at 11 telemetry stations were obtained to predict flow at the confluence between Klang River and Ampang River. The initial results of MLP yielded poor performance beneath normal expectations, which was $R = 0.4465$, MAE $= 3.7135$, NSE $= 0.1994$ and RMSE $= 8.8556$. Meanwhile, the LSTM model generated a 45% improvement in its R-value up to 0.9055. Detailed investigations found that the redundancy of data input that yielded multiple target values had distorted the model performance. Q_t was introduced into input parameters to solve this issue, while $Q_{t+0.5}$ was the target value. A significant improvement in the results was detected with $R = 0.9359$, MAE $= 0.7722$, NSE $= 0.8756$ and RMSE $= 3.4911$. When the rates of change were employed, an impressive improvement was seen for the plot of actual vs. forecasted flow. Findings showed that the rates of change could reduce forecast errors and were helpful as an additional layer of early flood detection.

Key words: artificial intelligence, neural network, rates of change, streamflow forecasting

HIGHLIGHTS

- A highly accurate flood forecasting system based on deep learning has been developed.
- Multi-lead ahead forecasting for streamflow has been investigated.
- The novel architecture model has been successfully applied for controlling streamflow in SMART tunnel.
- The proposed new model architecture could be applied to forecast river streamflow in different hydrological areas

GRAPHICAL ABSTRACT



1. INTRODUCTION

In recent years, climate change has seriously affected the ecosystem (Han *et al.* 2022). Climate change in the form of unpredictable precipitation, temperature and evaporation patterns not only affects the surrounding ecosystem but can disrupt the hydrological trait of streamflow. Streamflow is a major element of the hydrological cycle. The attribute of streamflow is highly associated with climate and land-use conditions (Masrur Ahmed *et al.* 2021). Alterations made by human activities can accelerate instabilities in the temperature and rainfall patterns resulting in adverse conditions such as sea level rise and extreme weather (Adikari *et al.* 2021). Similarly, one of the critical issues when land cover is being altered is the loss or reduction of allowable areas for infiltration. The change can result in more runoff into the river. Failure to mitigate or adapt this change will cause overflow at the riverbanks and a major flood.

The global threat of urban flooding to water security and the economy is too enormous to be ignored. Major flooding can cause severe damage to the infrastructure. The mudflow often ruins belongings. Thus, a resilient approach is necessary to mitigate this peril. A good and reliable multi-step ahead forecasting model can be a potential risk management solution for better flood management and disaster preparedness to allow sufficient evacuation and asset protection (Kao *et al.* 2021; Nanditha & Mishra 2021).

1.1. Process-driven model vs. data-driven model

There are two approaches to developing a multi-step ahead forecasting model, the process-driven model and the data-driven model (Huang *et al.* 2021).

The process-driven model, also known as a physically based model, is derived from the physics mass conservation theory and momentum preservation. Physical data, such as precipitation, river alignment and hydraulic structures, such as culverts, weirs and dams, or other geological past evidence are required from physical sites through observation and numerous ground surveys, interviews, satellite and aerial photography. Various assumptions can bring uncertainties to these data (Teng *et al.* 2017). Specific parameters cannot be obtained directly and must be substituted with default parameters. The watershed in the model will be represented as lumped, semi-distributed or fully distributed (Cai & Yu 2022). It concurrently will be defined under the range of complexity from conceptual to physically specified (Bourdin *et al.* 2012). Forecasting activity will be developed based on the principle of flood formation, considering physical features, hydrology, hydrodynamics and other theoretical aspects (Chen *et al.* 2021).

Data-driven model has gained considerable interest in recent years and is regarded as an alternative to the process-based model. Data-driven model does not require the simulation of physical processes. The model employs historical values or multivariate models to forecast the future (Zhao *et al.* 2021). The model can retrieve information from