多種多時刻資料驅動模式於入庫流量模擬應用-以翡翠水庫為例

Various Multistep-Ahead Data-Driven Models for Simulating Reservoir Hourly Inflow: A Case Study of Feitsui Reservoir, Taiwan

國家災害防救科技中心

副研究員	研究員	研究員
郭文達	陳偉柏	張志新
Wen-Dar Guo	Wei-Bo Chen	Chih-Hsin Chang

摘　要

近年來，隨著觀測資料豐富累積及人工智慧技術穩健發展，使得多種資料驅動模式在水文水理分析課題應用上，逐漸獲得可靠且良好的預測成果。雖然文獻上多有探討不同資料驅動模式於水文領域的評估研究，但大多係分析各不同模式之誤差表現，較少探討不同模式間之通用能力差異比較。為此，本研究旨在結合資料驅動與時間序列多時刻預測架構，發展九種入庫流量資料驅動模式，包含：支援向量 MO-SVR、隨機森林 MO-RFR、極限梯度提升 MO-XGBR、輕量梯度提升 MO-LGBMR、類別梯度提升 MO-CGBR、多層感知機 MO-MLPR、門控遞迴神經網路 ED-GRU、長短期記憶 ED-LSTM、以及卷積神經網路架構之長短期記憶 CNN-LSTM，藉由歷史實測水文數據，進行模式驗證及通用能力表現評估。

本研究以翡翠水庫為示範區域，首先蒐集彙整 2006 年至 2019 年期間之豪雨及颱風事件水文資料，包含水庫上游集水區 6 處雨量站及水庫入流量逐時數據。本研究所發展九種入庫流量資料驅動模式中之 MO-SVR、MO-RFR、MO-XGBR、MO-LGBMR、MO-CGBR 以及 MO-MLPR 模式，係基於多輸出迴歸架構；ED-GRU、ED-LSTM 以及 CNN-LSTM 模式則係採用編碼解碼架構。為了解訓練好的模式是否也能準確預測未來新數據變化趨勢，本研究採用納什效率係數 NSE 與通用能力 GA 指標，以探討九種模式對於翡翠水庫入庫流量之模擬表現能力；結果顯示，MO-SVR、MO-RFR、MO-XGBR、MO-LGBMR、MO-CGBR、MO-MLPR、ED-GRU、ED-LSTM 以及 CNN-LSTM 模式，在未來 1 至 6 小時之平均 NSE 值均大於 0.9，未來 1 至 6 小時之平均 GA 值則分別為 0.9、1.4、1.5、1.5、1.1、0.8、1.1、1.2 以及 1.4。研究結果顯示，九種模式可透過水庫上游集水區之降雨量，快速模擬颱風或豪雨可為水庫帶來多少入庫流量；其中以 MO-SVR、MO-CGBR 以及 ED-GRU 模式可獲得不錯的翡翠水庫入庫流量模擬通用性，研究成果可提供未來入庫流量推估預警之重要參考。

關鍵詞：資料驅動模式，入庫流量，編碼解碼架構，多輸出迴歸，通用能力

Abstract

With sufficient observed data and advanced development of artificial intelligence, various data-driven models have gradually yielded reliable and appropriate results in applying hydrological studies. Although several studies have presented the evaluations of different data-driven models for hydrological problems, few studies have used the concept of GA (generalization ability) to assess the models' performance comprehensively. Therefore, this study employed the multistep-ahead prediction framework to propose nine data-driven models, namely the multi-output support vector regression (MO-SVR), the multi-output random forest regression (MO-RFR), the multi-output extreme gradient boosting regression (MO-XGBR), the multi-output light gradient boosting machine regression (MO-LGBMR), the multi-output categorical gradient boosting regression (MO-CGBR), the multi-output multilayer perceptron regression (MO-MLPR), the encoder-decoder gated recurrent unit (ED-GRU), the encoder-decoder long short-term memory (ED-LSTM), and the combined convolutional neural network with LSTM (CNN-LSTM) models, for simulating reservoir hourly inflow during typhoons or storms.

In this study, the inflow hydrographs in Feitsui Reservoir were simulated using nine multistep-ahead data-driven models. Historical hourly rainfall and reservoir inflow were collected from 2006 to 2019, in which 70% and 30% of data were used for model training and validation. Based on multi-output regression, six data-driven models, namely the MO-SVR, MO-RFR, MO-XGBR, MO-LGBMR, MO-CGBR, and MO-MLPR, were extended to achieve the multi-hour ahead predictive ability of reservoir inflow. Three data-driven models, namely the ED-GRU, ED-LSTM, and CNN-LSTM, were extended to the multistep-ahead data-driven models using encoder-decoder framework. This study used Nash-Sutcliffe efficiency (NSE) and GA to evaluate the performance of nine data-driven models. The results indicate all models obtain the good NSE value of over 0.9 for predictions with 1–6-h lead times. In addition, the lead-time averaged GA using MO-SVR, MO-RFR, MO-XGBR, MO-LGBMR, MO-CGBR, MO-MLPR, ED-GRU, ED-LSTM and CNN-LSTM models were 0.9, 1.4, 1.5, 1.5, 1.1, 0.8, 1.1, 1.2 and 1.4, respectively. The results show that the MO-SVR, MO-CGBR, and ED-GRU models are better generalizable for inflow simulations in Feitsui Reservoir. The findings from this study provide a meaningful basis for reservoir inflow forecasting in future research.

Keywords: Data-driven model, Reservoir inflow, Encoder-decoder framework, Multi-output regression, Generalization ability.