

人工智慧應用在七日水庫入流量模擬

Application of an Artificial Intelligence Technique for Predicting a 7-Day Reservoir Inflow

國家災害防救科技中心 坡地洪旱組

Slopland and Hydrology Division, National Science and Technology Center for Disaster Reduction

摘要

本研究主要應用人工智慧技術於日入庫流量模擬，期能有助於台灣水資源調度分配之重要參考。本研究以石門與翡翠為示範水庫，應用人工智慧技術中的類別梯度提升迴歸(categorical gradient boosting regression，簡稱 CGBR)，建構示範水庫之入庫日流量推估模式。為獲得未來 1 至 7 天之入庫流量推估，本研究結合 CGBR 與多輸出架構(multiple output framework)，進行模式訓練與驗證。本研究亦採用線性水庫(linear reservoir，簡稱 IR)進行示範水庫之入庫日流量推估，並將推估結果與 CGBR 的結果，進行比較探討。結果顯示，CGBR 獲得的總入庫流量誤差(total inflow-discharge error，簡稱 TIE)，整體小於 IR 的 TIE，代表 CGBR 獲得較佳的入庫日流量推估表現。

關鍵字：人工智慧、日入庫流量、類別梯度提升迴歸、線性水庫

ABSTRACT

This study presents the application of the artificial intelligence technique to predict 7-day reservoir inflows. Based on combining the categorical gradient boosting regression (CGBR) model with the multiple output framework, the daily reservoir-inflow estimation model was proposed and applied to the Shihmen and Feitsui Reservoirs. The capability of the proposed model in predicting the reservoir inflow with lead times within 1 to 7 days was investigated. To further evaluate the performance of the proposed model, the results achieved from the CGBR model were compared to those obtained from the linear reservoir (IR) model. The results showed that the proposed model achieved minor total inflow-discharge errors compared to the IR model. Accordingly, the proposed model provided a reasonable and adequate performance in predicting the 7-day reservoir inflows, which would be helpful for efficient reservoir operation management.

Key Words: artificial Intelligence, daily reservoir inflow, categorical gradient boosting regression, linear reservoir.

一、前言

在台灣水資源有限之前提下，入庫日流量之準確推估，可實質地幫助台灣水資源調度分配之重要參考。近年已有許多研究進行水庫入流量之推估，國外水庫之研究方面，Yang et al. (2017)利用神經網路(artificial neural network, 簡稱 ANN)、支援向量(support vector machine, 簡稱 SVM)以及隨機森林(random forest, 簡稱 RF)，進行美國加利福尼亞 Trinity Lake 與中國丹江口水庫(Danjiangkou dam)之月入庫流量推估；Latif et al. (2021)應用 ANN、SVM 以及長短期記憶(long short-term memory, 簡稱 LSTM)，預測位於馬來西亞半島(Peninsular Malaysia)之 Durian Tunggal 水庫入庫流量；Maddu et al. (2022)應用 K-鄰近迴歸(K-nearest neighbors regression)、RF 以及 LSTM，進行印度 Bhadra 水庫之日入庫流量推估研究。

在國內水庫研究方面，Huang et al. (2022)應用七種人工智慧技術，進行石門水庫之未來 1 至 6 小時入庫流量預測研究；Guo et al. (2023)應用 CGBR，結合誤差修正，進行石門與翡翠水庫未來 1 至 24 小時之入庫流量預測研究。由文獻可得知，不同人工智慧技術已成功應用於國內外水庫入流量推估，包含不同未來時間尺度之預測表現探討；然而，石門與翡翠水庫之日入庫流量推估則較少探討。因此，本研究以未來 1 至 7 天為目標，進行石門與翡翠水庫之日入庫流量推估研究，期能有助於水庫水資源決策分配之重要參考。

本研究蒐集石門與翡翠水庫歷年之逐日水文資料，包含水庫集水區平均降雨量及入庫流量。透過資料之前處理與資料劃分，進行 CGBR 模式的訓練與驗證。此外，為進一步探討 CGBR 模式的表現，本研究採用 IR 模式進行石門與翡翠水庫於代表事件下之日入庫流量推估，並將模擬結果與 CGBR 結果進行比較，分析兩模式(CGBR 與 IR)之表現差異。

二、人工智慧技術

1. 模式架構

本研究所採用之 CGBR 模式架構可表示如下(Guo et al., 2023)：

$$\mathbf{Y} = f(\mathbf{X}) \quad (1)$$

式中， \mathbf{Y} 為模式的輸出向量結果， \mathbf{X} 為模式的輸入向量， f 為 CGBR 功能函數，用以學習輸入與輸出之間的非線性關係。

本研究所採用之輸入向量因子，主要考量有前期雨量、未來雨量以及前期入庫流量，可表示如下：

$$\mathbf{X} = [R_{t-L}, \dots, R_t, R_{t+1}, \dots, R_{t+L}, Q_{t-L}, \dots, Q_t] \quad (2)$$

式中， t 為時間， R 為水庫集水區之平均降雨量， Q 為入庫流量， L 為前置時間(lead time)。

(1)式中的輸出結果向量 \mathbf{Y} 可表示：

$$\mathbf{Y} = [\bar{Q}_{t+1}, \bar{Q}_{t+2}, \dots, \bar{Q}_{t+L}] \quad (3)$$

式中， \bar{Q} 為入庫流量之預測結果。本研究探討未來 1 至 7 天，故 L 為 7。

2. CGBR 技術

CGBR 主要係利用梯度提升決策樹(gradient-boosted decision tree)，以有效進行時間序列問題的迴歸預測，及避免過度擬合(overfitting)，其特徵值之變換(feature transformation values)可表示(Prokhorenkova et al., 2018)：

$$X_i^k = \frac{\sum_{j=1}^n \varphi(X_j^k = X_i^k) y_j + \alpha P}{\sum_{j=1}^n \varphi(X_j^k = X_i^k) + \alpha} \quad (4)$$

式中， n 為資料之總數據量($j=1, 2, \dots, n$)， X_i^k 為特徵值的轉換變量， y_j 為第 j 筆時序列資料之輸出結果， φ 代表指標函數(indicator function)， P 為先

驗值(prior value)， α 為先驗值數據資料之權重。

三、未來 7 天水庫入流量模擬與驗證

1. 研究區域與資料概述

石門與翡翠水庫之基本資料，如表 1 所示，包含集水區面積、設計有效容量、水文觀測資料範圍以及資料數據總量。水文觀測資料之主要來源為經濟部水利署北區水資源分署與臺北翡翠水庫管理局，資料種類為歷年逐日之水庫集水區平均降雨量與入庫流量數據。以翡翠水庫之資料數據最多，年份自 1988 年 1 月 1 日到 2020 年 12 月 31 日，合計 12,054 筆。

2. 模擬結果與表現評估

本研究採用文獻上最常使用之資料劃分比例，數據之 70%作為模式訓練，30%則作為模式驗證。模式訓練上，本研究發現兩示範水庫之觀測值與推估結果，在未來 1 至 7 天獲得之相關係數均可大於 0.9，代表模式訓練良好。在驗證之部分，圖 1(a)與圖 1(b)分別顯示 CGBR 模式於石門與翡翠水庫之未來第 7 天驗證結果，顯示兩示範水庫之 R 平方值均可大於 0.8，代表 CGBR 可獲得良好的日入庫流量推估表現。

3. 與線性水庫結果進行比較

為進一步探討 CGBR 與 IR 之表現差異，本研究選取五場代表事件，進行 TIE 之評估：

$$\text{TIE}(\%) = \frac{Q_t^{\text{simulated}} - Q_t^{\text{measured}}}{Q_t^{\text{measured}}} * 100\% \quad (5)$$

式中， $Q_t^{\text{simulated}}$ 與 Q_t^{measured} 分別代表總入庫流量之模擬結果與觀測值。

圖 2 與圖 3 分別顯示為石門與翡翠水庫之 CGBR 與 IR 模擬結果之比較，結果顯示 CGBR 整體可獲得較接近實際觀測入庫流量之預測表現。表 2 進一步彙整評估結果，明顯顯示 CGBR 獲得的 TIE 均比 IR 來的小，代

表 CGBR 可獲得較佳的入庫流量推估表現。

四、 結論與建議

本研究利用人工智慧之 CGBR 技術，透過歷年水文觀測資料數據，建構能夠推估未來 1 至 7 天的入庫流量模式。透過與 IR 模式之比較發現，CGBR 模式確實獲得較佳的日入庫流量推估表現。本研究目前僅以水庫集水區平均降雨量作為模式的輸入，隨著雨量觀測技術之精進，未來可嘗試將不同空間雨量測站之資訊納入考慮，應可增加模式的預測表現。

五、 致謝

本研究感謝臺北翡翠水庫管理局提供翡翠水庫水文觀測資料。此外，本研究使用經濟部水利署北區水資源分署提供的石門水庫水文觀測資料，在此一併致謝。

六、 參考文獻

- Guo, W. D., Chen, W. B., & Chang, C. H. (2023). Prediction of hourly inflow for reservoirs at mountain catchments using residual error data and multiple-ahead correction technique. *Hydrology Research*, 54(9), 1072-1093.
- Huang, I. H., Chang, M. J., & Lin, G.F. (2022). An optimal integration of multiple machine learning techniques to real-time reservoir inflow forecasting. *Stochastic Environmental Research and Risk Assessment*, 36(6), 1541-1561.
- Latif, S. D., Ahmed, A. N., Sathiamurthy, E., Huang, Y. F., & El-Shafie, A. (2021). Evaluation of deep learning algorithm for inflow forecasting: a case study of Durian Tunggal Reservoir, Peninsular Malaysia. *Natural Hazards*, 109(1), 351-369.
- Maddu, R., Pradhan, I., Ahmadisharaf, E., Singh, S. K., & Shaik, R. (2022). Short-range reservoir inflow forecasting using hydrological and large-scale atmospheric circulation information. *Journal of Hydrology*, 612, 128153.
- Prokhorenkova, L., Gusev, G., Vorobev, A., Dorogush, A. V., & Gulin, A. (2018). CatBoost: unbiased boosting with categorical features. *Advances in Neural Information Processing Systems*, 2018-December, 6638-6648.

Yang, T., Asanjan, A. A., Welles, E., Gao, X., Sorooshian, S., & Liu, X. (2017). Developing reservoir monthly inflow forecasts using artificial intelligence and climate phenomenon information. *Water Resources Research*, 53(4), 2786-2812.

表 1 石門與翡翠水庫基本資料

水庫	集水區面積(km ²)	水庫設計有效容量 (10,000 m ³)	資料範圍	數據總量
石門	763.4	25,188	2009.1.1~2021.12.31	4,748
翡翠	303	34,409.3	1988.1.1~2020.12.31	12,054

表 2 CGBR 與 IR 於石門與翡翠水庫之 TIE(%)表現比較

水庫	事件編號(名稱)	時間	CGBR	IR
石門	事件 1 (2019 年米塔颱風)	9.26 ~ 10.9	-5.63	-13.44
	事件 2 (2021 年烟花颱風)	7.16 ~ 8.2	-2.38	5.03
	事件 3 (2021 年圓規颱風)	10.7 ~ 10.20	3.30	25.52
翡翠	事件 4 (2015 年杜鵑颱風)	9.24 ~ 10.8	-14.43	-25.98
	事件 5 (2017 年 10 月豪雨)	10.5 ~ 10.26	-0.58	-17.79
	事件 1 (2019 年米塔颱風)	9.23 ~ 10.11	-2.09	-24.34

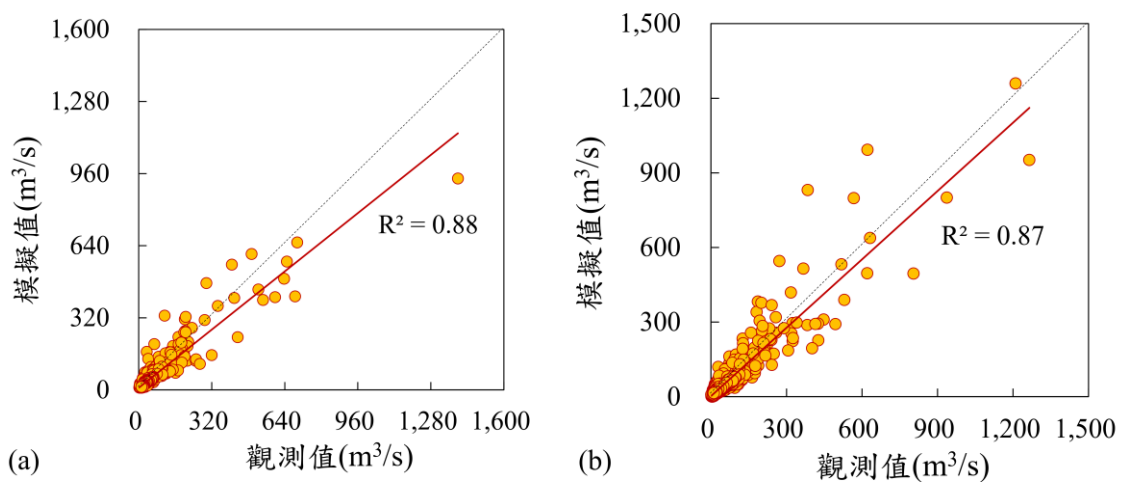


圖 1 CGBR 模式於(a)石門與(b)翡翠水庫之驗證結果

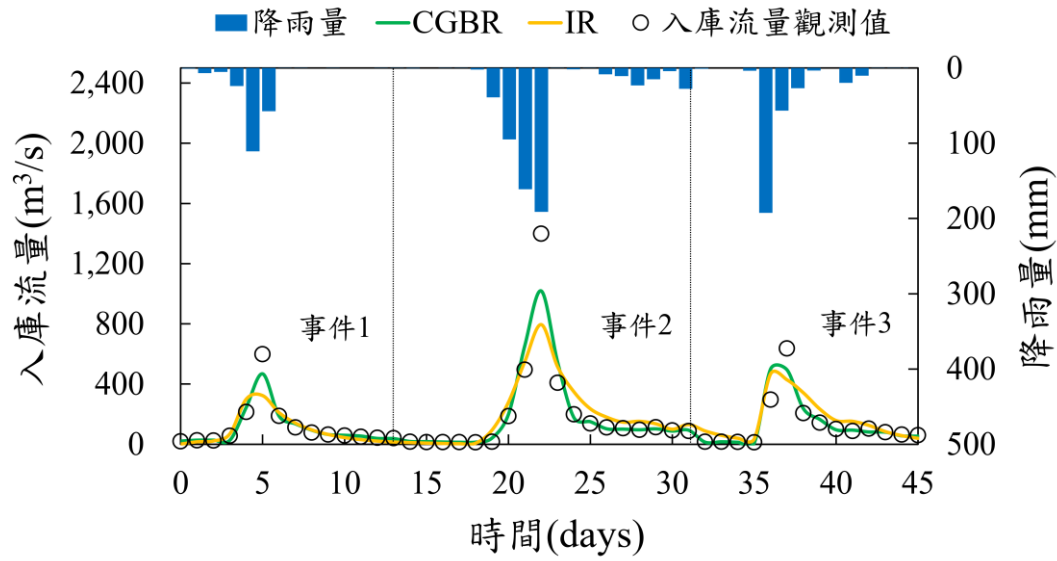


圖 2 CGBR 與 IR 模式於石門水庫代表事件之日入庫流量模擬結果比較

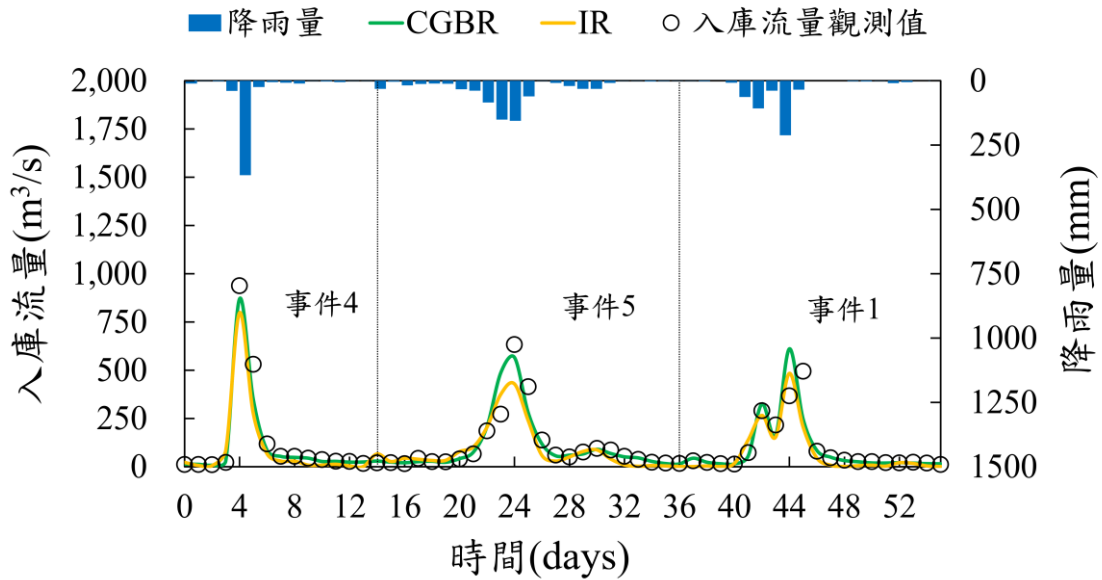


圖 3 CGBR 與 IR 模式於翡翠水庫代表事件之日入庫流量模擬結果比較