

應用數據驅動技術發展新式水位-流量率定曲線 Development of a Novel Rating Curve Using Data-Driven Approaches

國家災害防救科技中心 坡地洪旱組

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

摘要

本研究應用各類數據驅動模型及統計回歸模型，針對海陸域複合洪災之即時預警系統研發，於巨量數據流量與水位之時間序列關係分析。研究中以水利署橫溪水位站歷年實測水位及流量數據為基礎，深入探討支持向量回歸、二次多項式回歸、隨機森林、人工神經網絡等多種數據驅動及統計回歸模型，對建立新式河川水位-流量率定曲線的表現。透過數種數據驅動模型的對比分析結果，本研究評估各模型於不同年份及洪水事件之預測表現，並呈現不同模型於高低流量預測上之準確度變異。分析結果指出，模型準確度之提升依賴於大量且精確的實測資料，以及實測資料時間解析度之細化。本研究所獲得的成果，可為洪災預警及未來發展新示水文模型提供參考方向。

關鍵字：數據驅動模式、水位-流量率定曲線、水文模擬、河川水位、河川流量

ABSTRACT

This study applies various data-driven and statistical regression models to analyze the temporal relationship between massive data for river discharge and water level time series, i.e., rating curve. Based on multi-year field measurements at the Hengxi station of the Tahan River in Taiwan, the research extensively investigates several data-driven and statistical regression models, including support vector regression, quadratic polynomial regression, random forest, and artificial neural networks. Through comparative analysis, this study evaluates the predictive performance of each model across different years and flood events, revealing variations in accuracy in predicting high and low flow conditions. The results indicate that the improvement in model accuracy depends on an increase in the amount of measured data and the refinement of hydraulic models. The findings of this study can serve as a reference for flood warning systems and the future development of hydrological models.

Key Words: data-driven approaches, river discharge, water level, rating curve

一、前言

隨著氣候變遷影響日益加劇，極端天氣事件之頻發使得準確預測河川水位與流量變化越來越重要。河川水位-流量率定曲線之建立為水文學中一項基本而關鍵之工作，因其直接關聯降雨逕流模式參數校驗之正確與否、洪水預警系統之有效運作及水資源管理之決策參考。傳統之水位-流量關係建立依賴於長期之現場測量，但此種方式在資料收集上耗時且成本高昂，且在極端事件下往往難以獲得高質量之數據(Wu and Yang, 2008)。而隨氣候變遷及極端天氣事件之增加，準確之河川水位-流量率定關係對於洪水災害管理及水資源規劃變得更加重要。近年來數據驅動模型之進步，更可在傳統模式之外提供不同之參考。本研究以橫溪集水區作為研究區域，收整橫溪站之歷史水位與流量數據，使用不同之統計回歸方法及數據驅動方法，以建立新式水位-流量率定曲線。由於不同之統計回歸方法及數據驅動方法均有其不同之不確定性及限制，研究最後以多模式集成(multi-model ensembles)方式結合這些不同模型之預測結果，期可得到更全面之水位-流量預測範圍。

二、研究方法

本研究旨在採用水利署實測流量數據，結合水位站觀測記錄，發展以數據驅動模型建構之新式水位流量關係曲線。並使用模型之預測流量與水利署現行之水位流量率定曲線進行對比與驗證。研究中使用以下數據驅動與回歸模型進行分析：支持向量回歸 (Support Vector Regression, SVR)、二次多項式回歸 (Quadratic Polynomial Regression)、線性回歸 (Linear Regression)、隨機森林回歸 (Random Forest Regression)、人工神經網絡 (Artificial Neural Network, ANN)、嶺回歸 (Ridge Regression)、Lasso 回歸 (Lasso Regression)、彈性網回歸 (Elastic Net)、梯度提升回歸 (Gradient Boosting)、K 最近鄰回歸 (K-Nearest Neighbors Regression)、決策樹回歸 (Decision Tree Regression) 及雙對數率定曲線 (double logarithmic rating curve)。在這些方法中，二次多項式回歸、線性回歸以及雙對數率定曲線屬於傳統統計回歸模型範疇，其餘均為較現代之數據驅動模型，目的在探索並改善傳統水文預測方法之準確性與可靠性。

三、亮點成果

圖 1 及圖 2 為 2008 年之各種統計回歸模型及數據驅動模式之模型生成流量

與水利署率定流量之散佈圖及年流量序列比較圖。其中與水利署率定流量最接近之前三模型為 SVR、雙對數率定曲線及二項式回歸，最差之模型為彈性網模型。

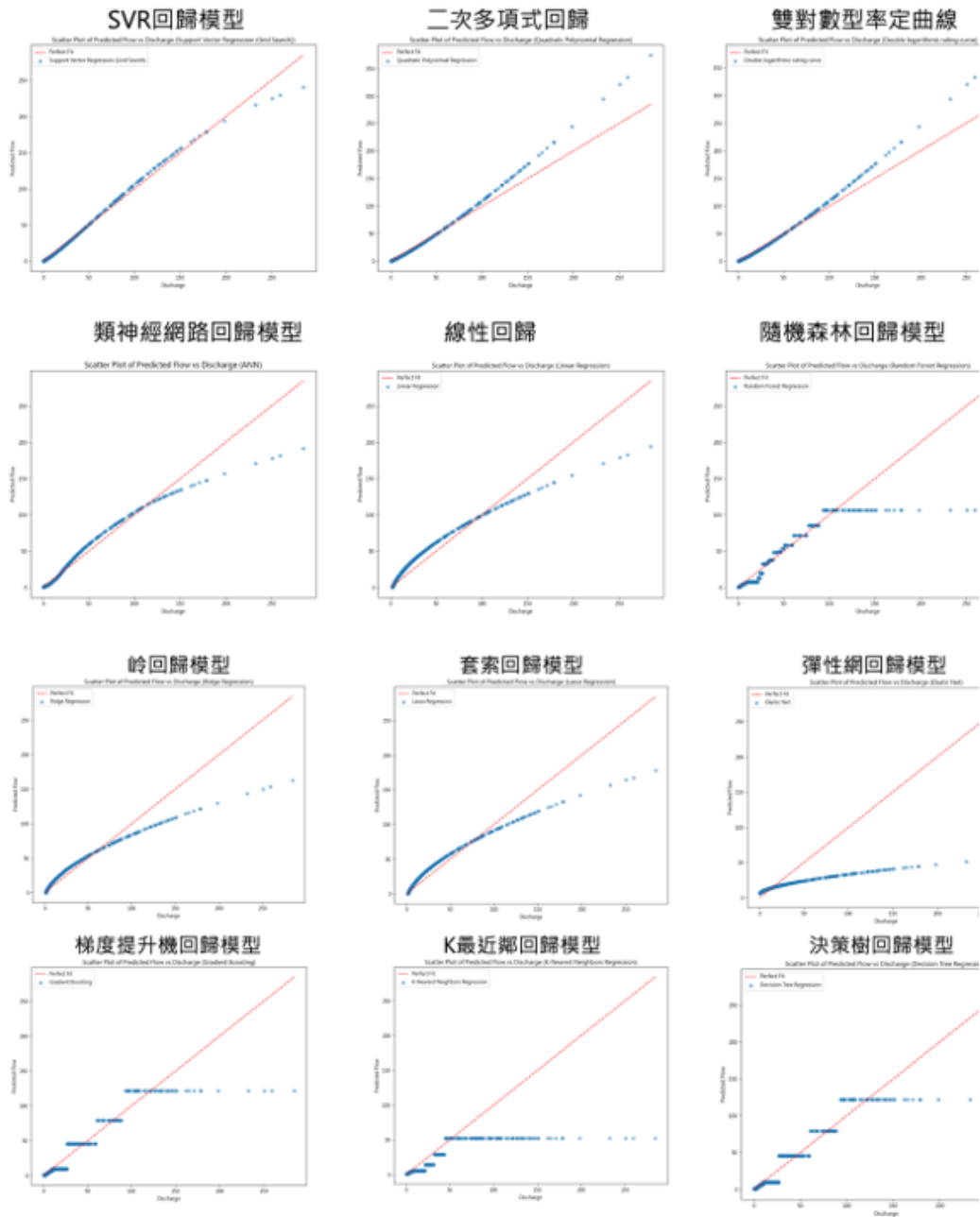


圖 1. 2008 年橫溪站各種統計回歸模型及數據驅動模式之模型生成流量與水利署率定流量之散佈圖

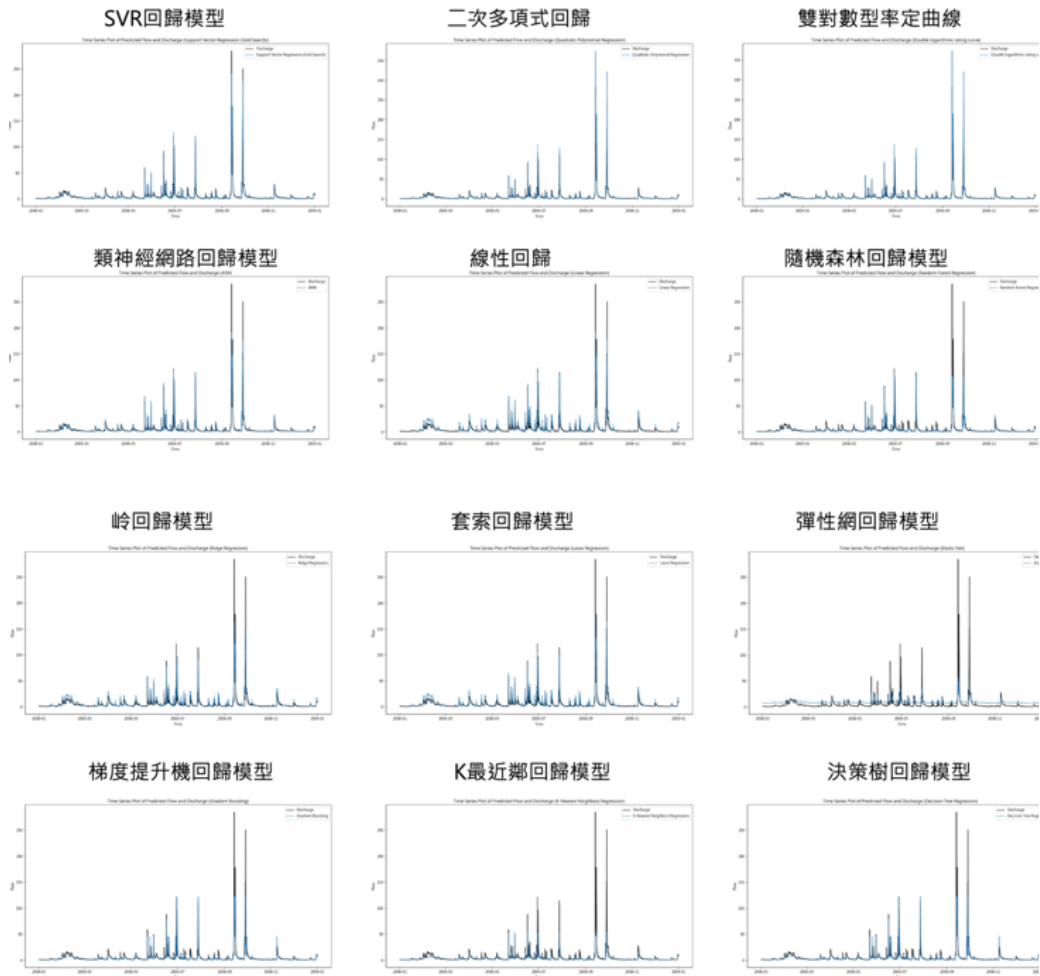


圖 2. 2008 年橫溪站各種統計回歸模型及數據驅動模式之模型生成流量與水利署率定流量之年流量時間序列

四、 結論與建議

本研究以橫溪集水區作為示範區，使用水利署水位觀測資料以及現地觀測流量資料，以不同統計回歸模式及數據驅動模式建立其河川水位-流量率定曲線，並將模式結果與水利署之率定流量進行比較驗證。研究中並嘗試以不同之資料分類方式進行相關性分析，希冀得到合理之河川水位-流量率定曲線建立之資料分類方式。研究結果顯示，乾濕季或月份分類之方式均無法得到良好之水位-流量相關性，以年份為單位之水位-流量相關性最高，因此本研究以逐年方式進行橫溪集水區之水位-流量率定曲線之建置。研究中使用：支持向量回歸 (Support Vector Regression, SVR)、二次多項式回歸 (Quadratic Polynomial Regression)、線性回歸 (Linear Regression)、隨機森林回歸 (Random Forest Regression)、人工神經網絡 (Artificial Neural Network, ANN)、嶺回歸 (Ridge Regression)、Lasso 回

歸 (Lasso Regression)、彈性網回歸 (Elastic Net)、梯度提升回歸 (Gradient Boosting)、K 最近鄰回歸 (K-Nearest Neighbors Regression)、決策樹回歸 (Decision Tree Regression) 及雙對數率定曲線 (double logarithmic rating curve) 等統計回歸及數據驅動回歸模式進行水位-流量率定曲線之建置。研究結果顯示使用資料之代表性將直接影響率定曲線之建置結果。但相較於數據驅動模式，傳統之統計回歸模式在使用同樣低品質之資料時，以水利署之率定流量作為評估標的時其表現優於數據驅動模式之結果。但若資料品質佳時，數據驅動模式有可能其表現優於傳統之統計回歸模式結果。研究最後以多模式集成 (multi-model ensemble) 之區間方式呈現不同模式所得水位-流量率定曲線之成果。結果顯示在洪峰時刻或流量變化快時多模式集成之預測區間將變大，顯示其不確定性升高，但在低流量或流量變化慢處之預測區間則較小，顯示其不確定性較低。

五、參考文獻

Wu, R.; Yang, J. An Analytical Method of Stage–Fall–Discharge Rating. *Hydrol. Process.* **2008**, *22*, 2959–2973, doi:10.1002/hyp.6867.