

# 混合水理模式與機器學習之淹水潛勢分析

## Flood Susceptibility Mapping based on a Hybrid Hydraulic-Machine Learning Modeling Approach

國立中興大學土木工程學系

研究生  
吳韋峰  
Wei-Feng Wu

副教授  
陳佳正  
Chia-Jeng Chen

### 摘要

氣候變遷會導致洪水事件越見頻繁，且受洪水影響之區域將更具不確定性。而機器學習因其效率及準確性，已被廣泛利用於評估淹水潛勢。本研究欲利用水理模式 Two-dimensional Runoff Inundation Toolkit for Operational Needs (TRITON)於彰化地區進行淹水模擬，並混合三種機器學習方法包含卷積神經網路(Convolutional Neural Networks, CNN)、隨機森林(Random Forest, RF)與支援向量機(Support Vector Machine, SVM)進行淹水潛勢評估。TRITON 採用圖形處理器進行高效能運算，藉由模擬多種不同淹水情境取得易淹水點位，並蒐集淹水影響因子，再與機器學習模型結合建立淹水潛勢評估模式。預期在此混合模式架構下，除能解決稀缺之歷史淹水點位紀錄外，亦可藉由此模式架構繪製淹水潛勢地圖，提供作為防災參考依據。

關鍵字: 淹水潛勢、TRITON、卷積神經網路、隨機森林、支援向量機

### Abstract

Climate change will induce more frequent flood events, and areas prone to floods will become more and more uncertain. Machine learning methods have been widely used in flood susceptibility assessments due to their efficiency and accuracy. This study aims to propose a hybrid modeling framework that uses a hydrodynamic model referred to as the Two-dimensional Runoff Inundation Toolkit for Operational Needs (TRITON) and three machine learning methods, namely Convolutional Neural Networks (CNN), Random Forest (RF), and Support Vector Machine (SVM), for the flood susceptibility assessment in Changhua. TRITON, which makes use of graphics processing units to achieve high-performance computing, simulates various flood scenarios to acquire flood susceptible sites that are used as training and validation points for the machine learning methods to establish the flood susceptibility model. It is expected that the hybrid modeling framework can not only deal with the scarcity of historical flood records but also produce flood susceptibility maps for disaster preparedness.

Keywords: Flood susceptibility; TRITON; Convolutional neural network; Random forest; Support vector machine

## 一、 前言

近年來，氣候與環境變遷的影響，加上都市發展持續增加，極端降雨的情況頻頻發生，導致洪水事件愈發頻繁，使洪水成為全球最具破壞性的天災之一，為因應其流動性與複雜性，並有效地管理與制定防災策略，辨別易受淹水影響之區域顯得更加重要。為了精確地呈現研究區域之淹水情形，許多水理模式已被應用於模擬特定的降雨或流量情境，如美國陸軍工兵部隊所開發的 HEC-RAS 用於模擬潰壩事件(Dasallas et al., 2019)，亦或與地理資訊系統(Geographic Information System, GIS)結合呈現洪水事件前後之差異(Khalfallah & Saidi, 2018)，鑒於水理模式僅能應用於模擬特定降雨或流量情境，無法有效呈現研究區域淹水之災害潛勢，致淹水潛勢評估模式有建立的必要性。機器學習近年來在各領域顯示具有巨大的潛力，因其效率及準確性已被廣泛利用於評估災害潛勢，包括辨別森林野火潛勢(Bjånes et al., 2021)、山洪易發地點(Islam et al., 2021)和坡地災害潛勢(Rong et al., 2021)，而淹水潛勢則基於研究區域地質環境條件等因素辨別易受淹水影響之區域(Zhao et al., 2018)，本研究欲使用機器學習建立淹水潛勢評估模式，以精確地預測淹水之可能性。

利用機器學習建立淹水潛勢評估模式前，需給定淹水點位與蒐集研究區域各項影響淹水因子，由於過往現場調查的歷史淹水點位紀錄稀少，故本研究欲混合水理模式與機器學習以解決稀缺之歷史淹水點位紀錄(Fang et al., 2022)，藉由 TRITON (Two-dimensional Runoff Inundation Toolkit for Operational Needs)模擬多種不同淹水情境，相對於其他傳統水理模式因其架構複雜需較長的計算時間，而 TRITON 特色為採用圖形處理器進行高效能運算，可快速地模擬多種不同淹水情境以取得易淹水點位，並分別混合三種機器學習方法，其中包含卷積神經網路(Convolutional Neural Networks, CNN)、隨機森林(Random Forest, RF)與支援向量機(Support Vector Machine, SVM)進行淹水潛勢評估，並透過各項評估指標與淹水因子的特徵貢獻度比對，以作為不同機器學習方法在淹水潛勢應用之表現進行評估。

## 二、 研究區域與方法

### 2.1 研究區域與資料來源

本研究區域為位於台灣中部的彰化地區(如圖 1)，東邊以八卦山與南投縣所組成，西邊為台灣海峽，南邊以濁水溪與雲林溪為界，北邊則以烏溪與台中市為界，水文方面除濁水溪與烏溪之外，亦有洋仔厝溪、鹿港溪與舊濁水溪流經彰化地區，氣候則以夏季多雨冬季乾燥的亞熱帶季風氣候為主，冬季因東北季風遭中央山脈的阻擋使氣候乾燥，夏季因西南風與海風、颱風侵襲等關係，使雨量極為豐沛。

本研究使用降雨資料以及土地利用資料運用於模擬降雨事件使用，降雨資料為經濟部水利署公布之 24 小時降雨設計雨型，可用於產製多種不同降雨量情境，土地利用資料則使用內政部國土測繪中心於 2015 年所公布之資料，將土地利用現況進行系統性地分類及統計，運用高解析航遙測影像，配合地籍圖、臺灣通用電子地圖等各式參考圖資及部分外業調查，全面及持續性進行調查土地利用現況變化情形，合理規劃土地利用。

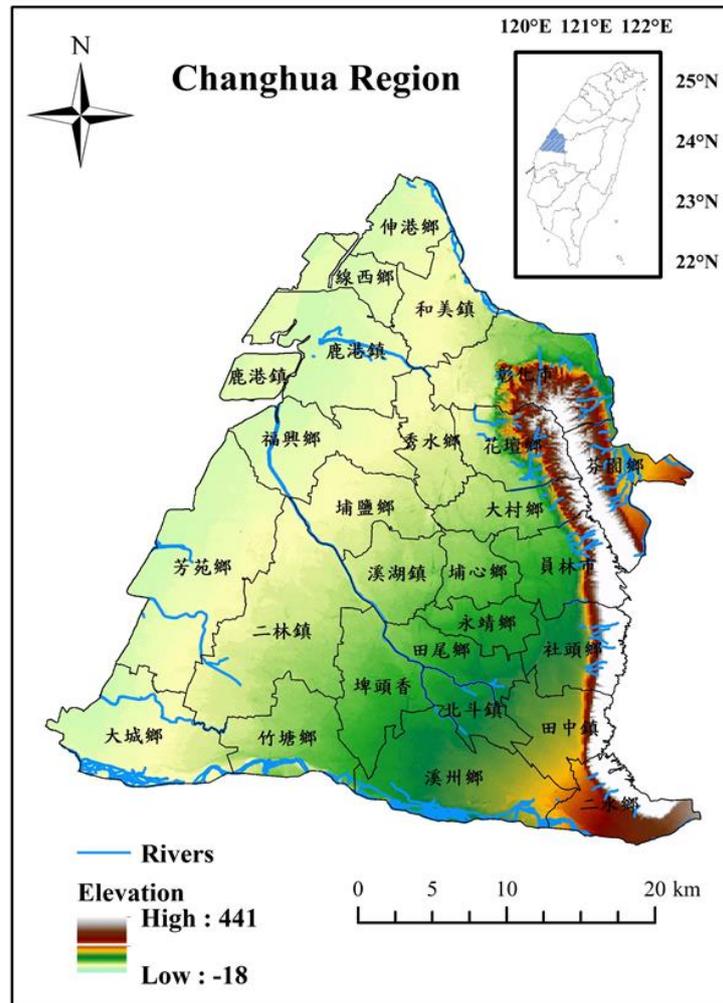


圖 1. 研究區域

## 2.2 TRITON 模式

TRITON 模式為橡樹嶺國家實驗室與田納西理工大學所共同開發的二維洪水模型，其特色為採用圖形處理器進行高效能運算，使大幅減少模擬降雨事件所需時間，與傳統水理模式如 HEC-RAS 相比，其模擬時間減少數倍有餘。TRITON 以二維淺水波方程式為主要控制方程式，運用一階精確有限體積法求解非線性雙曲線淺水波方程式，在水深上集成了水平動量和連續方程式（式 2-1 及式 2-2）。

$$\frac{\partial U}{\partial t} + \frac{\partial F}{\partial x} + \frac{\partial G}{\partial y} = S_b + S_f \quad (2-1)$$

$$U = (h \ q_x \ q_y) \quad F = \left( q_x \frac{q_x^2}{h} + \frac{1}{2} g h^2 \frac{q_x q_y}{h} \right) \quad G = \left( q_y \frac{q_x q_y}{h} + \frac{1}{2} g h^2 \right)$$

$$S_b = \begin{pmatrix} 0 \\ -gh \frac{\partial z}{\partial x} \\ -gh \frac{\partial z}{\partial y} \end{pmatrix} \quad S_f = \begin{pmatrix} 0 \\ -\frac{gn^2}{h^3} q_x \sqrt{q_x^2 + q_y^2} \\ -\frac{gn^2}{h^3} q_y \sqrt{q_x^2 + q_y^2} \end{pmatrix} \quad (2-2)$$

其中，U 是一包括水深 h 的守恆變量之向量， $q_x$  與  $q_y$  為 x 和 y 方向的單位流量，F 和 G 是通量的向量， $S_b$  和  $S_f$  是底床坡度和底床摩擦項，z 是河床高度，g 是重力加速度，n 是曼寧粗糙度係數。

TRITON 於模擬降雨事件時，其主要輸入數據包括數值高程模型(DEM)、曼寧因子(Manning's coefficient)、有效降雨量及研究區域邊界條件，並透過 GIS 工具進行輸入數據的預處理，以及後續輸出數據的後處理，使快速明瞭地呈現研究區域之最大淹水深度。

### 2.3 淹水因子

洪水災害通常與地形氣候等自然條件密不可分，本研究根據過往文獻與數據可得性收集了 10 項影響淹水潛勢之淹水因子，以用於淹水潛勢評估模式之開發，包括高程、坡向、曲率、坡度、土地利用、與河川的距離、地形濕度指數(Topographic Wetness Index, TWI)、逕流強度指數(Stream Power Index, SPI)、泥砂輸送指數(Sediment Transport Index, STI)和地形粗糙指數(Topographic Ruggedness Index, TRI)。高程與土地利用資料為使用內政部國土測繪中心公布之五米資料，其中高程為影響洪水的重要因子之一，通常與洪水發生之機率成反比(Dodangeh et al., 2020)，因此彰化八卦山山腳一帶相較於其他地勢較平緩地區淹水可能性較高；坡向與坡度會影響洪水的流動方向與速度，其中洪水流量與流速則隨著坡度的增加而增加；曲率會影響著低流速的洪水，平坦且凹面的區域使容易發生洪水；與河川的距離之資料為使用經濟部水利署公布之河川河道資料與 GIS 工具進行計算，且河川附近往往容易發生洪災，反之，距離愈遠其風險則愈低；地形濕度指數(TWI)表示研究區域對泥沙流失的承受能力，其概念為當坡度較陡峭時，土砂不易堆積、使水流不易入滲與聚集(式 2-3)；逕流強度指數(SPI)代表該地的輸砂與侵蝕能力，隨著流域面積與坡度上升，使洪水風險上升(式 2-4)；泥砂輸送指數(STI)代表地形對於侵蝕之影響，當河床泥砂堆積，降低河道蓄水能力，導致洪水容易氾濫(式 2-5)；地形粗糙指數為該網格高程與相鄰網格的差異量。

$$TWI = \ln \left( \frac{A_s}{\tan \beta} \right) \quad (2-3)$$

$$SPI = A_s \times \tan \beta \quad (2-4)$$

$$STI = \left( \frac{A_s}{22.13} \right)^{0.6} \left( \frac{\sin \beta}{0.0896} \right)^{1.3} \quad (2-5)$$

其中， $A_s$  為該網格流域面積， $\beta$  為該網格坡度。並藉由 GIS 工具將各淹水因子的之網格解析度統一調整為 50 米，以匯入機器學習模型進行訓練。

## 2.4 機器學習模型

### (一) 卷積神經網路

卷積神經網路(Convolutional Neural Networks, CNN)在機器學習中，以識別影像的能力較為出色，隨著計算能力的進步，亦可應用於其他方面與災害潛勢上。其架構由一層輸入層、多層隱藏層與一層輸出層所組成，而 CNN 特色為隱藏層中由一個或多個卷積層(Convolution Layer)及池化層(Pooling Layer)與最後的全連接層(Fully Connected Layer)所組成。在卷積層中利用卷積核(Kernel)針對輸入影像進行卷積運算，其目的為特徵擷取，並依序滑動與內積，再於池化層進行池化，常見池化方式有最大池化以及平均池化，其目的為壓縮擷取的特徵，萃取重要特徵以縮小尺寸減少計算時間，最後連接至全連接層以提取經過池化後的特徵；為有效地將數據集進行訓練，以精確地預測淹水潛勢，本研究以 10 項淹水因子作為輸入數據於卷積層與池化層，及連接至最後的全連接層，其中於卷積核層中設置 16 個大小為 2 的卷積核進行卷積運算，與在池化層中進行大小為 2 的最大池化，並在最後輸出層的神經元設置 2 作為卷積神經網路的架構，以實現二元分類問題(如圖 2)。

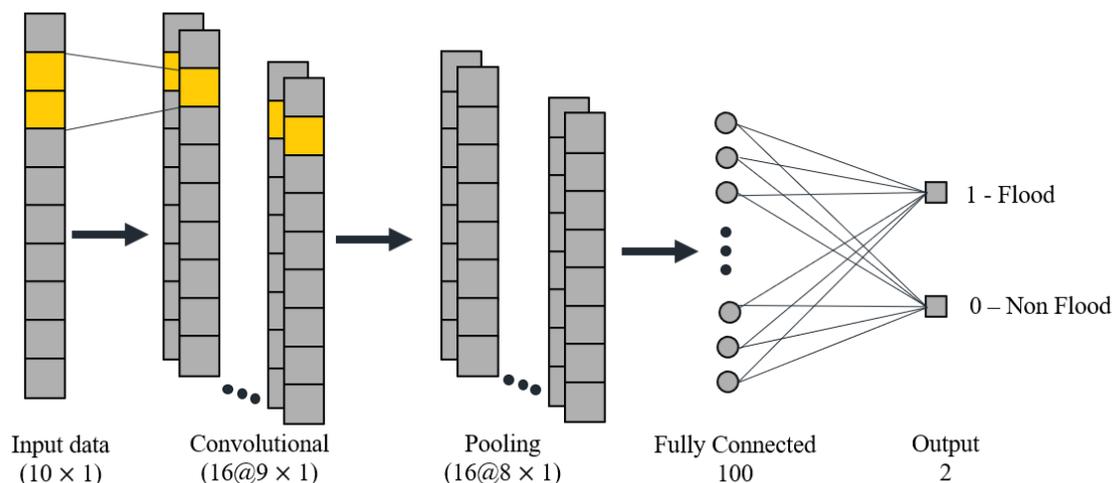


圖 2. 卷積神經網路架構

### (二) 隨機森林

隨機森林(Random Forest, RF)為許多不同的決策樹所組成的一個學習器，並隨機分配輸入之訓練資料，並以多數決的方式整合分類結果，其生成方法為先從訓練集中隨機選取  $n$  筆資料取後放回，利用該筆資料隨機選取  $p$  個特徵進行分割節點訓練決策樹，並重複  $m$  次以訓練出  $m$  顆決策樹，整合其預測結果以多數決決定分類結果；為精確地預測淹水潛勢，本研究設置 1000 顆決策樹於建立 RF 模型，並利用所有淹水因子與使用熵(Entropy)公式決定分支方式以建立決策樹(式 2-6)。

$$Entropy = \sum_{i=1}^c -P_i \times \log_2(P_i) \quad (2-6)$$

其中， $c$  為分類類別的數量， $P_i$  為數據集中屬於該類別的比例。

### (三) 支援向量機

支援向量機(Support Vector Machine)為一種線性分類器，基於統計學理論為基礎之機器學習模型，除了線性分類之外，亦可應用於小樣本、非線性與多維度等方面。其特色為將低維度且線性不可分的數據集投射至高維度空間，並找到一個超平面將樣本有效地切割分類，並使兩類樣本之間的邊界(Margin)最大化，將其區隔開來使提高預測能力；除此之外，核函數的選擇對於SVM至關重要，其中SVM模型的核函數包括線性核(Linear Kernel)、多項式核(Polynomial Kernel)、徑向基函數核(Radial Basis Function, RBF)與 Sigmoid 核(Sigmoid Kernel)，根據測試與過往研究提及(Islam et al., 2021; Zhao et al., 2020)，使用 RBF 函數在淹水潛勢上成效優於其他核函數，故本研究使用 RBF 作為 SVM 核函數(式 2-7)。

$$RBF = \exp(-\gamma\|x_i - x_j\|^2) \quad (2-7)$$

其中， $\|x_i - x_j\|^2$  為兩類別特徵向量之間的歐幾里得距離， $\gamma$  為 gamma 參數，其參數決定了函數作用範圍，隨著參數的增大而減弱，使 RBF 對數據集中的雜訊有較好的抗干擾能力。

### 2.5 驗證方法與評估方式

為避免模型訓練發生過度擬合，與將數據集充分利用與學習，本研究使用 K 折交叉驗證(K-fold Cross-Validation)作為驗證方法，藉由將數據集劃分為 K 等份，每次選擇其中一份作為驗證集，其餘 K-1 份作為訓練集，重複 K 次以獲取 K 組不同的分組結果，進行平均以減少誤差，藉由評估指標以檢視模型成效。

為方便檢視模型成效，本研究使用評估指標包括準確度(Accuracy)、精確度(Precision)、靈敏度(Sensitivity)與特異度(Specificity)，其公式如下。

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (2-8)$$

$$Precision = \frac{TP}{TP + FP} \quad (2-9)$$

$$Sensitivity = \frac{TP}{TP + FN} \quad (2-10)$$

$$Specificity = \frac{TN}{TN + FP} \quad (2-11)$$

其中真陽性(True Positive, TP)表示預測正確的淹水樣本數、真陰性(True Negative, TN)表示預測正確的非淹水樣本數、偽陽性(False Positive, FP)表示預測錯誤的淹水樣本數及偽陰性(False Negative, FN)表示預測錯誤的非淹水樣本數。並藉由靈敏度與 1-特異度繪製 ROC 曲線(Receiver operating characteristic curve)，計算 AUC 值(Area under roc curve)以評估模型預測之能力。

### 三、 初步結果

在運行 TRITON 二維淹水模擬前，藉由經濟部水利署公布之台中市 24 小時定量降雨 650mm 淹水潛勢圖，作為驗證對象進行模式驗證，其淹水潛勢圖為利用 SOBEK 淹水模式模擬，由於 TRITON 二維模擬缺乏入滲機制等，以及未考慮防洪排水等設施，經模擬測試，將降雨量扣掉 50%降雨損失則可較接近 SOBEK 淹水模式模擬的情形(如圖 3)，故本研究使用 50%降雨量作為有效降雨量進行淹水模擬。藉由 TRITON 模擬多種不同降雨量的淹水情境，取得研究區域易淹水與非淹水區域，並蒐集研究區域淹水影響因子，與混合三種機器學習模型建立淹水潛勢評估模式，並使用 K 折交叉驗證法與各評估指標，檢視模型訓練成效好壞，進而繪製淹水潛勢地圖。

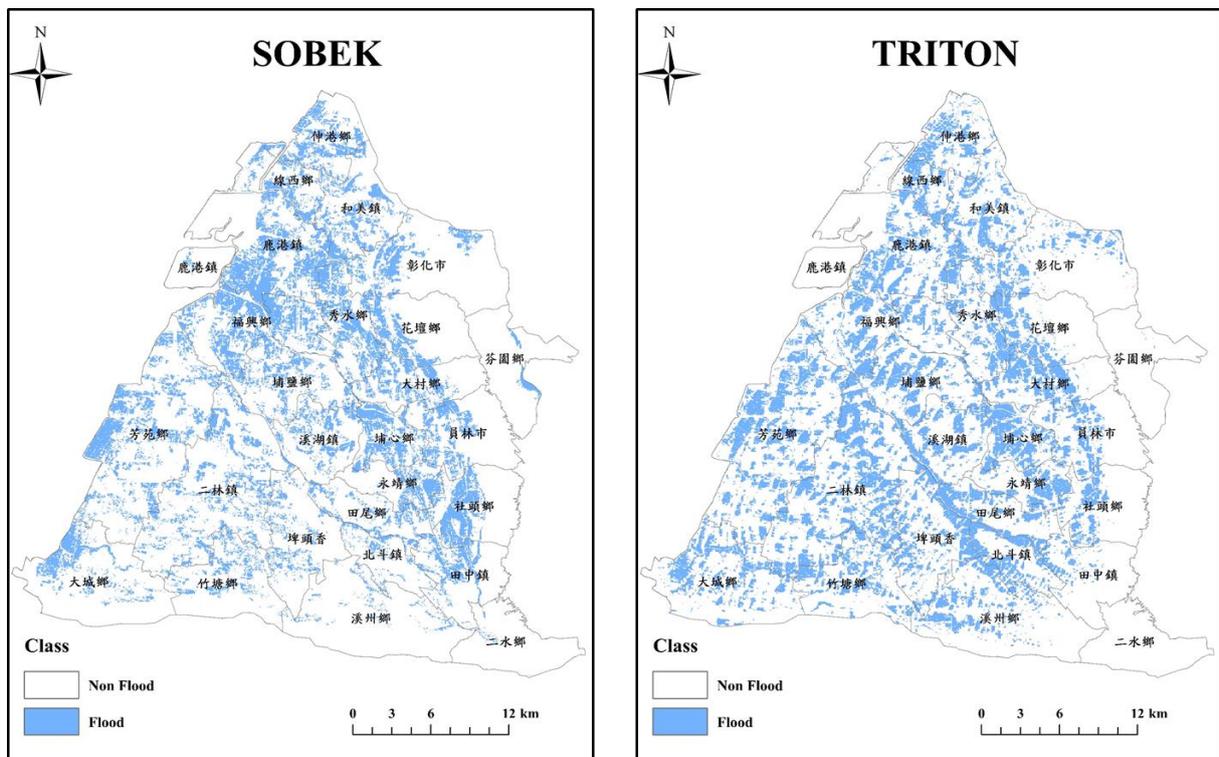


圖 3. 24 小時延時 650mm SOBEK 及 TRITON 淹水模擬情形

各模型訓練的驗證結果如表 1 至 3 所示，紀錄各模型於不同點位數量之評估指標，可看出各模型於點位數量愈多時，無論在各評估指標上，或所繪製的 ROC 曲線與計算的 AUC 值(如圖 4)，成效隨著點位數量增多而提升。圖 5 為各模型於各點位數量之特徵貢獻度，(a)至(c)依序分別為 CNN、RF 與 SVM，可看出 CNN 與 SVM 於各點位數量皆以高程的貢獻度為最高，而 RF 除 1 萬個點外，以曲率的貢獻度為最高，初步判定因 RF 決策樹建立機制所致，使與其他模型不同。圖 6 為各模型於各點位數量之淹水潛勢圖，(a)至(c)依序分別為 CNN、RF 與 SVM，各模型於 1 萬個點進行訓練的驗證效果雖為最佳，但所產製的彰化淹水潛勢圖在大多地區呈現極高淹水潛勢，有失其真實性，其他測試之驗證效果雖相對較差，但淹水潛勢圖的呈現較為合理，需持續嘗試不同的降雨情境與點位數量使淹水潛勢圖更加合理及精準。

表 1. CNN 於各點位數量之評估指標

Points	CNN			
	10000	5000	3000	1000
Accuracy	0.90	0.86	0.84	0.81
Precision	0.91	0.83	0.81	0.78
Sensitivity	0.90	0.90	0.90	0.86
Specificity	0.91	0.82	0.79	0.76

表 2. RF 於各點位數量之評估指標

Points	RF			
	10000	5000	3000	1000
Accuracy	0.92	0.85	0.84	0.82
Precision	0.95	0.85	0.86	0.85
Sensitivity	0.90	0.84	0.83	0.79
Specificity	0.95	0.85	0.86	0.86

表 3. SVM 於各點位數量之評估指標

Points	SVM			
	10000	5000	3000	1000
Accuracy	0.88	0.84	0.83	0.78
Precision	0.88	0.84	0.82	0.77
Sensitivity	0.90	0.86	0.84	0.79
Specificity	0.87	0.83	0.82	0.76

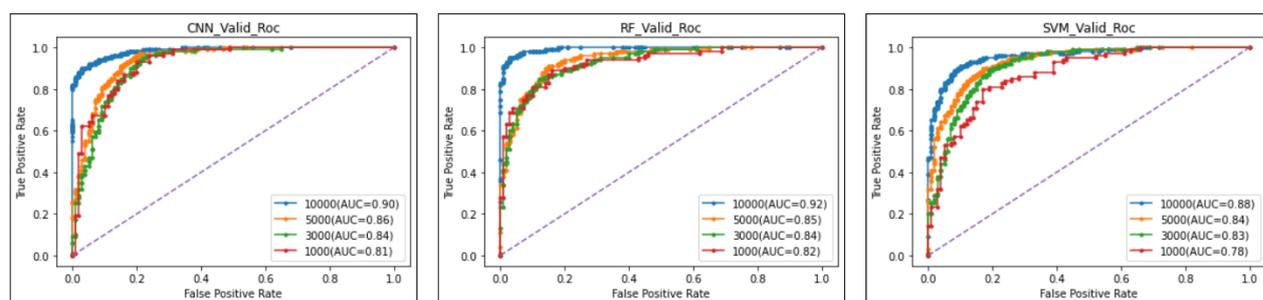
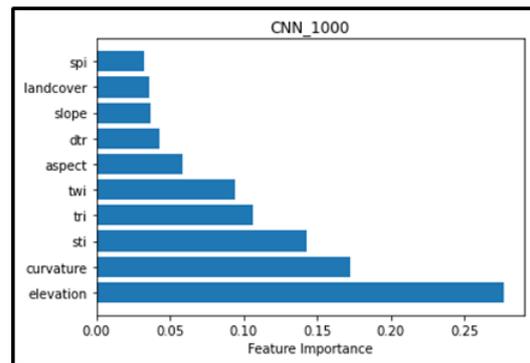
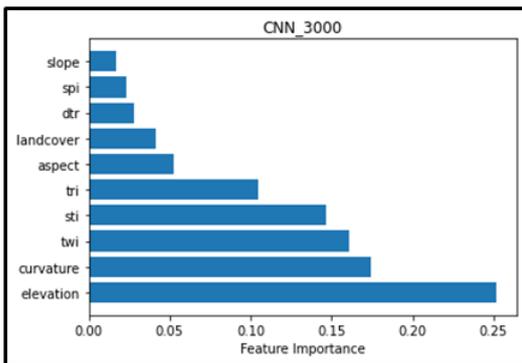
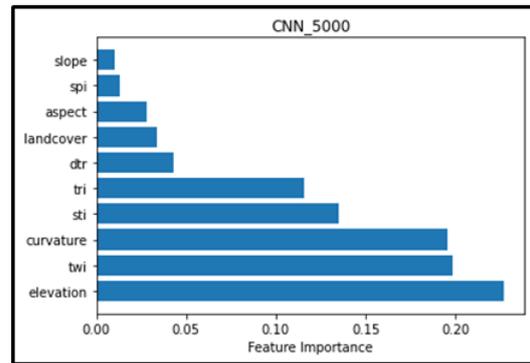
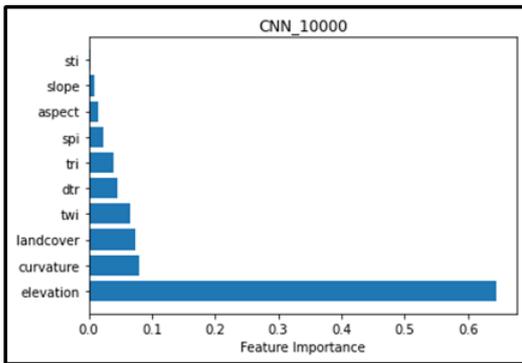
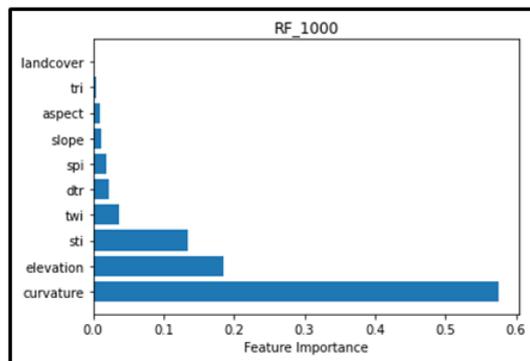
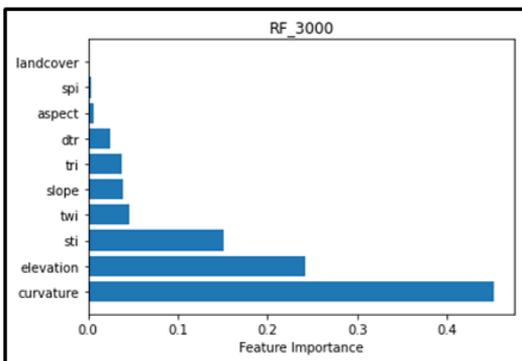
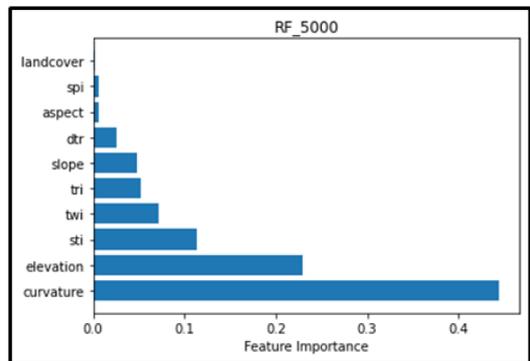
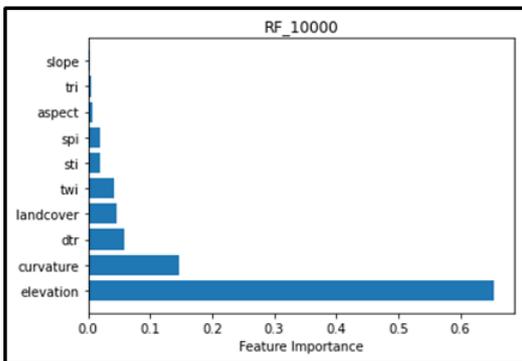


圖 4. 各模型於各點位數量之 ROC 曲線及 AUC 值

(a)



(b)



(c)

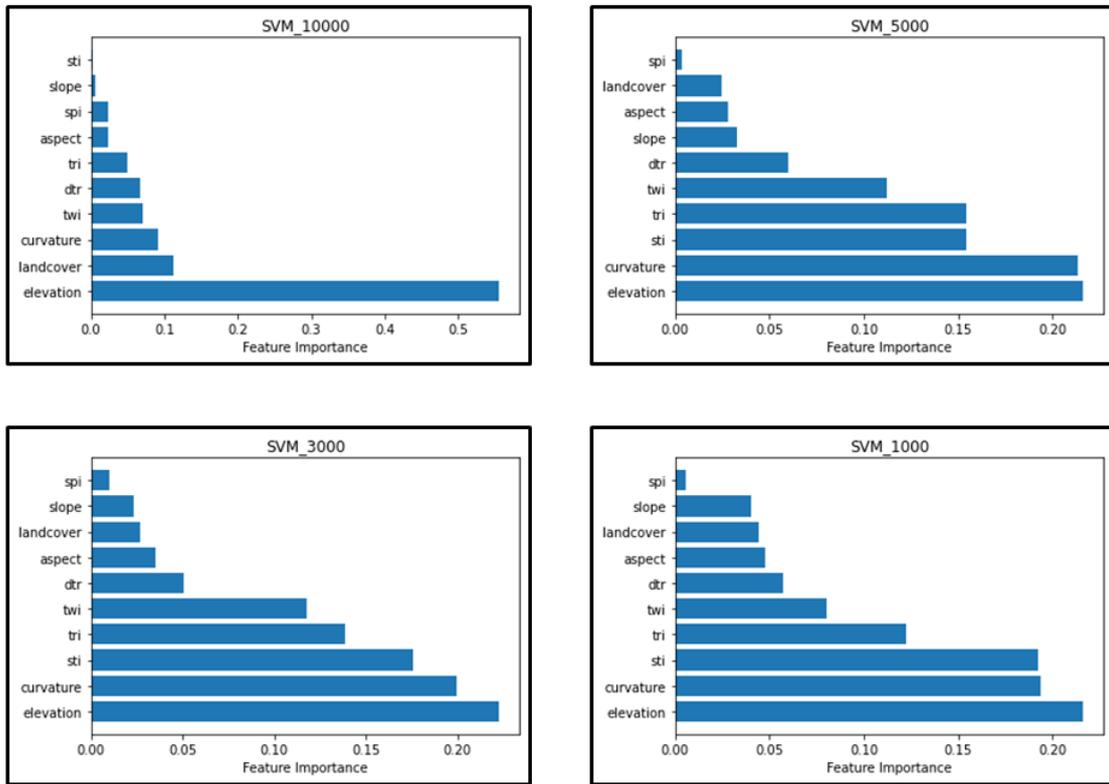
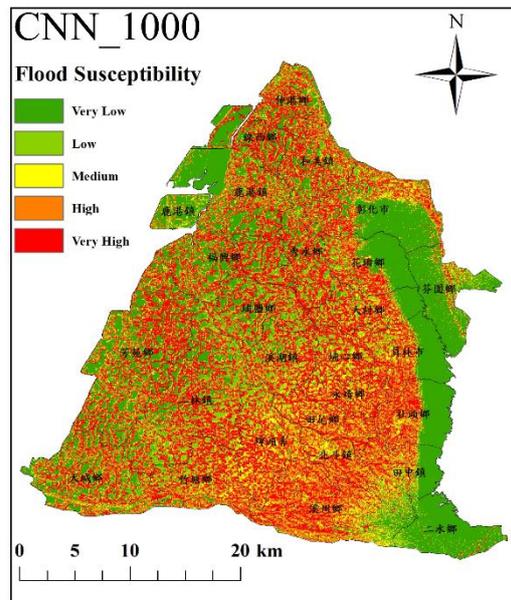
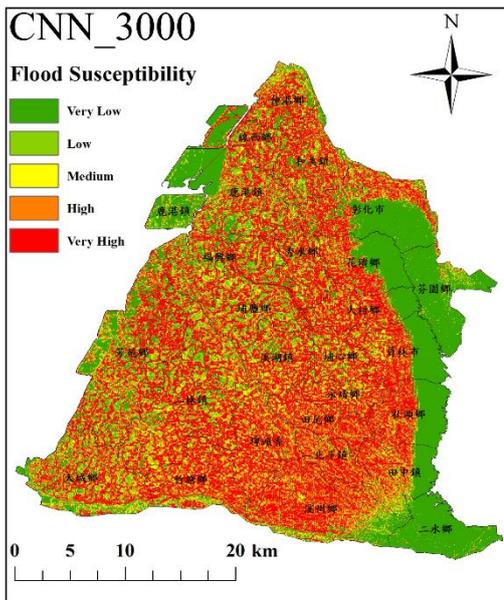
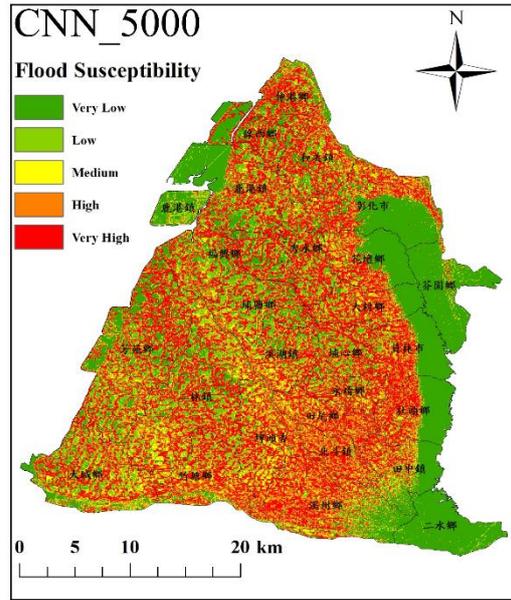
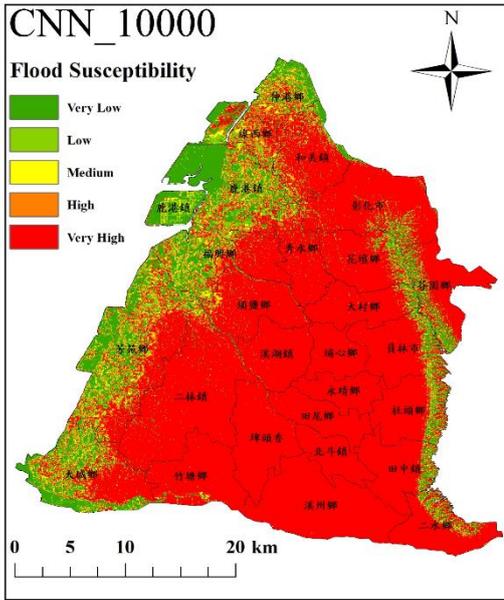
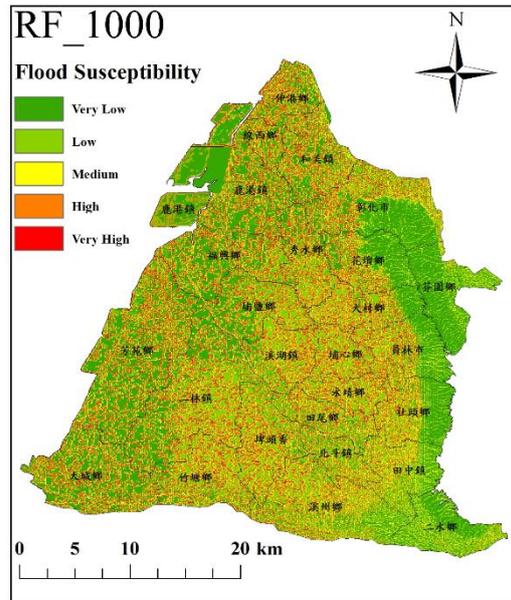
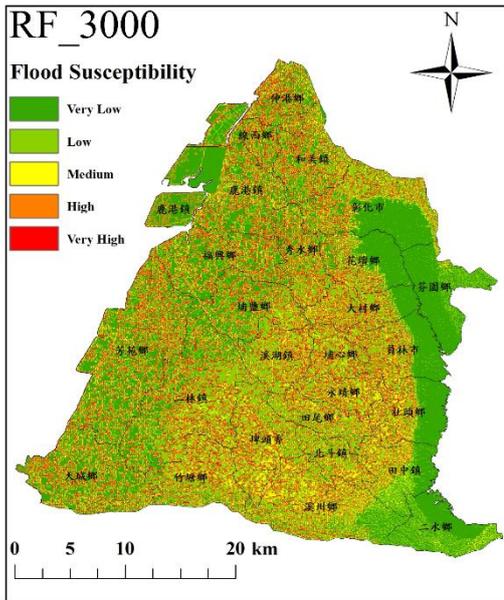
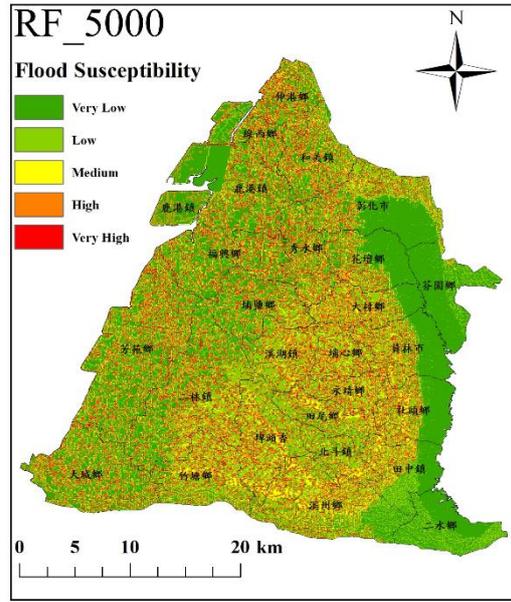
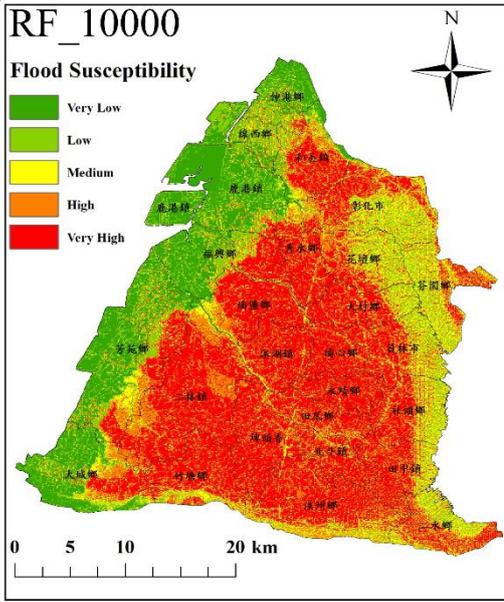


圖 5. 各模型於各點位數量之特徵貢獻度

(a)



(b)



(c)

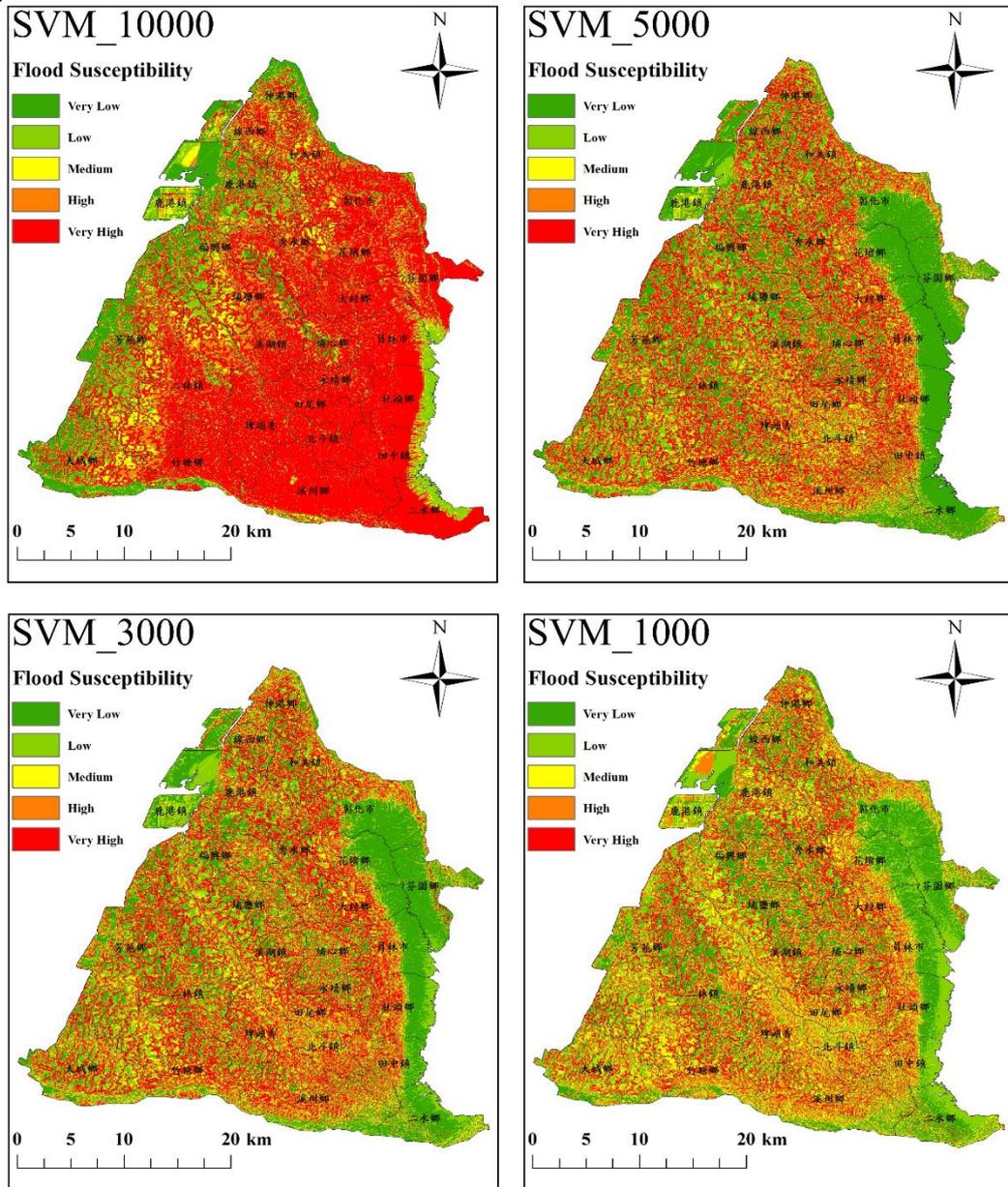


圖 6. 各模型於各點位數量之淹水潛勢圖

#### 四、 小結與後續工作

執行至目前，水理模式部分，透過 TRITON 模擬測試案例，瞭解其模式本身性能，並藉由經濟部水利署所公布之台中市 24 小時定量降雨 650mm 淹水潛勢圖，作為驗證對象進行模式驗證，以利爾後模擬之準確度；機器學習部分，雖已混合 TRITON 模式與機器學習模型，但所產製的研究區域淹水潛勢圖存在許多不確定性，混合模式架構亦尚未完成。以下列出本研究將繼續進行之後續工作，以達獲取更精準之淹水潛勢之目的。

1. 運行 TRITON 模式於彰化地區，採用空間上非均勻的降雨形式，測試多種不同的降雨情境，並比較各情境下之淹水情形。
2. 應用 TRITON 模式各情境下所輸出結果(如易淹水與非淹水區域)，並收集研究區域之淹水因子，與機器學習模型混合以建立淹水潛勢評估模式。
3. 測試數據集點位數量，與調整機器學習模型架構，以完成混合模式的實現與完善結果，產製更精準之淹水潛勢圖。
4. 在此混合模式架構下，除解決稀缺之歷史淹水點位紀錄，亦可透過評估指標與淹水因子的特徵貢獻度，比較不同機器學習方法在應用淹水潛勢上之差異。

## 五、 參考文獻

1. Bjånes, A., De La Fuente, R., & Mena, P., "A deep learning ensemble model for wildfire susceptibility mapping," *Ecological Informatics*, 65, 101397, 2021.
2. Dasallas, L., Kim, Y., & An, H., "Case study of HEC-RAS 1D-2D coupling simulation: 2002 Baeksan flood event in Korea," *Water*, 11(10), 2048, 2019.
3. Dodangeh, E., Choubin, B., Eigidir, A. N., Nabipour, N., Panahi, M., Shamshirband, S., & Mosavi, A., "Integrated machine learning methods with resampling algorithms for flood susceptibility prediction," *Science of the Total Environment*, 705, 135983, 2020.
4. Fang, L., Huang, J., Cai, J., & Nitivattananon, V., "Hybrid approach for flood susceptibility assessment in a flood-prone mountainous catchment in China," *Journal of Hydrology*, 612, 128091, 2022.
5. Islam, A. R. M. T., Talukdar, S., Mahato, S., Kundu, S., Eibek, K. U., Pham, Q. B., Kuriqi, A., & Linh, N. T. T., "Flood susceptibility modelling using advanced ensemble machine learning models," *Geoscience Frontiers*, 12(3), 101075, 2021.
6. Khalfallah, C. B., & Saidi, S., "Spatiotemporal floodplain mapping and prediction using HEC-RAS-GIS tools: Case of the Mejerda river, Tunisia," *Journal of African Earth Sciences*, 142, 44-51, 2018.
7. Rong, G., Li, K., Su, Y., Tong, Z., Liu, X., Zhang, J., Zhang, Y., & Li, T., "Comparison of tree-structured parzen estimator optimization in three typical neural network models for landslide susceptibility assessment," *Remote Sensing*, 13(22), 4694, 2021.
8. Zhao, G., Pang, B., Xu, Z., Peng, D., & Zuo, D., "Urban flood susceptibility assessment based on convolutional neural networks," *Journal of Hydrology*, 590, 125235, 2020.
9. Zhao, G., Pang, B., Xu, Z., Yue, J., & Tu, T., "Mapping flood susceptibility in mountainous areas on a national scale in China," *Science of the Total Environment*, 615, 1133-1142, 2018.