

藉由相隔 15 年的兩組數據探索蘇鐵蕨物種分布模型 之時間可轉移能力

郭慶津¹ 邵寶嬋² 羅南璋³ 黃凱易^{4*}

摘要

本研究旨在探討不同演算法對物種分布模型 (species distribution model, SDM) 時間可轉移能力 (transferability) 的影響，以及使用深度學習法建立 SDM 的可能性。研究以蘇鐵蕨為目標物種，獲取該物種兩組調查時間相隔 15 年之數據，以最大熵值法 (maximum entropy, MAXENT)、隨機森林 (random forest, RF)、支持向量機 (support vector machine, SVM) 和深度學習法 U-net 進行試驗。結果顯示，MAXENT 和 SVM 有最佳的時間轉移能力，而 U-net 也有機會獲得甚佳的成果。顯示深度學習具研究潛力，後續研究有必要採納更多種類的深度學習法，並持續試驗。惟在環境變數方面，單獨使用地形因子可能限制了模型的時間轉移性，需尋找與物種更具直接因果關係的生態因子提高可轉移性。

關鍵詞：時間可轉移性、深度學習、物種分布模型、蘇鐵蕨

1. 前言

物種分布模型 (species distribution model, SDM) 是一種結合遙測 (remote sensing, RS)、地球空間資訊系統 (geospatial information system, GIS)、全球衛星導航系統 (global navigation satellite system, GNSS) 及人工智慧 (artificial intelligence, AI) 的技術。其主要由三個部分組成，物種的分布數據 (主要是出現數據)、環境變數，以及估計物種對環境偏好程度的建模方法 (李國慶等, 2013)。

此技術可預測物種潛在生育地，多年來已被應用在各式領域如造林樹種選擇 (Jensen *et al.*, 2021)、預測氣候變遷對物種分布之影響 (Brzozowski *et al.*, 2022、Glad & Mallard, 2022)、劃設保護區 (Maciel *et al.*, 2021、Yousefi *et al.*, 2022) 或是入侵生物管理 (Baer & Gray, 2022、El-Barougy *et al.*, 2021、Wang *et al.*, 2022)。

理想情況下，SDM 必須要具備時間可轉移 (透

通) 性 (transferability)。所謂時間轉移性指的是模型在模擬目標物種從過去至未來可能的空間分布變化，由此便能分析某一物種可能的遷徙路徑，或是族群未來可能的擴展位置。SDM 的時間轉移性對於外來物種的防除工作至關重要，若能正確預測其高風險擴散區，防除人員便能提前部署，在其入侵的早期階段變進行移除，避免其落地生根。

氣候變遷對物種分布的影響也往往是屬於這方面的研究。近年來，全球氣候變遷對生態系造成莫大的衝擊，許多學者想藉由 SDM 來分析當氣溫或降雨量改變，物種的生育地將如何變化。Brzozowski *et al.* (2022) 便以 SDM 預測一種稀有輪藻 (*Lychnothamnus barbatus*) 於現今、過去及未來的潛在分布，並比較不同氣候情境下的預測結果，找出其未來可能的生育地。

SDM 的時間轉移性給予生態研究一個機會來觀察和分析生態系統在不同時間尺度上的變化。但此類研究往往存有極大之限制，即我們無法得知物

¹ 國立中興大學森林學系 碩士生

² 林業及自然保育署南投分署埔里工作站 技士

³ 國立中興大學農業暨自然資源學院實驗林管理處育林組 組長

⁴ 國立中興大學森林學系 教授

* 通訊作者, E-mail: kyhuang@dragon.nchu.edu.tw

收到日期：民國 112 年 10 月 02 日

修改日期：民國 112 年 10 月 29 日

接受日期：民國 112 年 11 月 27 日

種在未來的實際生存點位，因而無法即刻證實模型的時間轉移能力。

而若要回溯物種過往的分布位置，儘管可自標本館或博物館得到累積數十年的資料，但這些數據難免存有位置偏差，或者是在衛星定位技術普及前採集的，因此只能從其文本描述推斷出概略的地理位置 (Feeley & Silman, 2010)。以這些品質堪虞的數據建立與驗證模型，將損及模型性能，影響後續的決策。

若能克服上述問題，另一個要面對的即是建模方法的選擇。經過長年的發展，目前已經有許多的演算法被提出並用於建立 SDM。而不同建模方法也會影響模型的時間可轉移性，這即是本研究之探討目標之一。此外，過往的 SDM 研究幾乎都是使用機器學習法來建立模型，鮮少有應用深度學習者。此乃因山高水深、層巒聳翠的自然環境甚難實地採集資料，故樣本數少有符合此新法之需求。

且前人研究指出機器學習法在模型的空間外推表現上可能不盡理想 (陳厚昌, 2013、Su *et al.*, 2012)。陳厚昌 (2013) 探討木荷於惠蓀林場空間分布，由其相距 5-8 km 的兩大集水區以包括 MAXENT、DOMAIN、決策樹等八種機器學習法相互空間外推，其準確度評估指標 κ 與真實技術統計 (true skill statistic, TSS) 落在 0.38-0.67，前三者居前且在伯仲

之間，惟表現不甚理想。Su *et al.* (2012) 模擬預測薯豆之物種空間分布，案例架構相近，所獲之結論亦同。

而深度學習具有學習能力強及優秀的泛化能力 (generalization ability)，這正是 SDM 所需要的特性。儘管本研究探討的是時間而非空間轉移性，但這兩者皆須藉良好泛化能力的演算法才能獲得較佳之推估成果。因此本研究之另一課題即是要克服樣本數的限制，以深度學習法建立 SDM，並與常見的機器學習法比較其時間可轉移能力。

2. 資料來源與研究方法

2.1 研究區域

本研究之試驗區位於南投縣仁愛鄉，為一涵蓋中興大學蓋惠蓀實驗林場之矩形區域，面積約為 18,000 公頃 (Ha)。惠蓀林場的海拔範圍落於 449 - 2,420 m 之間，整體地勢北低南高。圖 1 顯示了林場中的主要山峰與水系。最高峰為位於林場南界的守城大山 (2,420 m)，而林場中之兩大流域為關刀溪及東峰溪，兩溪皆由南向北流，並匯入林場北界的北港溪。惠蓀林場包含了低、中高海拔的生態環境，故其植物種類眾多，約達 1,100 種以上，為台灣中部極具代表性之森林 (羅南璋等, 2011)。

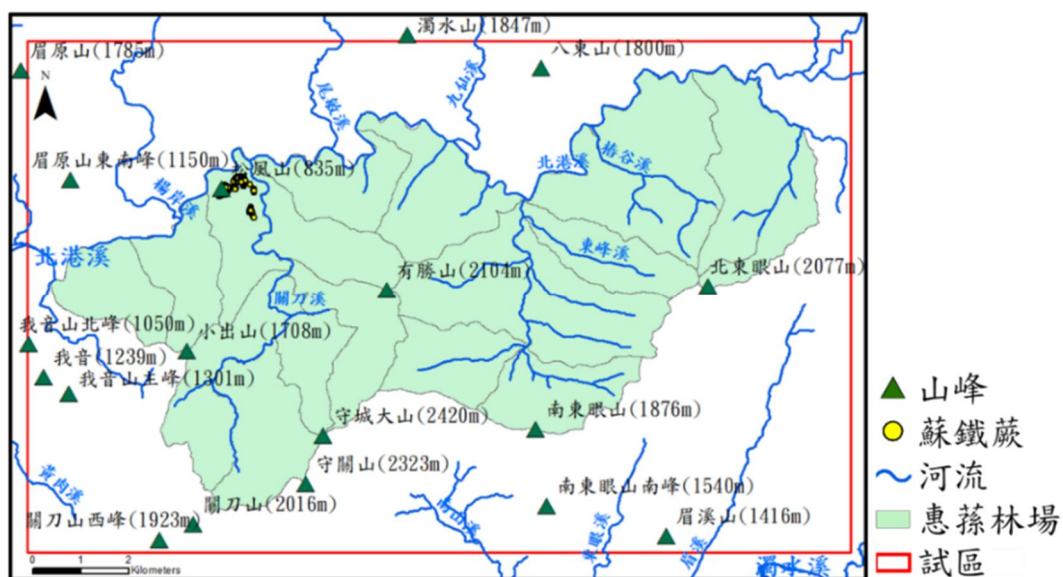


圖 1 研究區之主要水系與山峰

2.2 研究物種—蘇鐵蕨

本研究之目標物種為稀有植物蘇鐵蕨 (*Brainea insignis*)。其主要分布在台中和南投的中低海拔山區，以惠蓀林場的松風山一帶為其大本營，其他地方較為罕見 (郭城孟, 2001、許天銓等, 2019)。根據長年出入於惠蓀林場的中興大學實驗林管理處職員所觀察，若生育地環境適宜，此植物可存活至 30 年以上。

蘇鐵蕨為陽性植物，在光照良好的環境下生長旺盛，若於鬱閉處則成長受抑制。此物種是一種火燒適存植物，其頂芽具層層覆蓋之鱗片，能有效隔絕熱源，故在經歷林火後能快速復原 (蔡長宏, 1997)。而火焰的高溫能刺激其著生孢子 (呂金誠, 1990)，因此火燒是蘇鐵蕨傳宗接代的重要因子。

此外，詹雅婷、鄒裕民 (2021) 發現，蘇鐵蕨的出現與否和土壤稀土元素具密切關聯。有蘇鐵蕨生長的土壤稀土元素含量明顯高於無蘇鐵蕨分布處，而檢驗的 16 種稀土元素中以鈰 (Ce)、釹 (Nd)、鐳 (La) 為百分比含量最高，銩 (Tm)、鐳 (Lu) 為最低。惟此研究僅對惠蓀林場之族群進行分析，因此仍應當在其他蘇鐵蕨生育地做進一步的研究，充分理解和驗證蘇鐵蕨與稀土元素間的關聯性。

蘇鐵蕨樣本是本研究室歷經多年調查所獲得的資料。最早是在 2006 年於惠蓀林場內的松風山、杜鵑嶺及關刀山林道獲得 246 筆樣本，而 2021 年末和 2022 年初又分別在杜鵑嶺與松風山蒐集到 83 株蘇鐵蕨之資料。樣株的選擇方式是挑選經數年生長，已長出直立莖的植株，並以衛星定位儀測量其基部坐標。

早年調查採用的定位方式是動態後處理技術 (post processed kinematic, PPK)，使用儀器為 Trimble Pro XR GPS (簡稱 Pro XR) 並搭配 5 m 伸縮加長天線及雷射測距儀；近年調查之儀器則是使用 Trimble R12 衛星定位儀 (簡稱 R12)，並採用內政部國土測繪中心提供的 e-GNSS 即時動態定位系統 (real time kinematic, RTK)，透過無線網路連線產生虛擬參考站 (virtual reference station, VRS)，以 RTK VRS 進行定位。兩筆資料的定位誤差皆維持

在 ± 1 m 以內，具有相當可靠的位置情資 (location intelligence)，因此能有效降低位置偏差對模型的影響。而本研究將以 2006 年蒐集的樣本建立模型，並藉 2021 及 2022 年蒐集到的新樣本來驗證模型的時間可轉移性。



圖 2 蘇鐵蕨之外觀

2.3 環境變數

環境因子與各物種之間有複雜的反應關係。以植物而言，其可區分為氣候、土壤、生物及地形 (位置) 這四大類因子 (蘇鴻傑, 1987)，前三者一般可視為直接作用因子，若能作為預測變數，便有機會確保模型的透通性。然而這類資料多需藉氣象站之觀測數據或實地採集，數量遠不符所需，得再施以空間內插獲得相關網格圖層，惟其解析度仍然太低，且更含有難以預料之內插誤差。

地形因子雖屬於間接作用因子，但其是四者中唯一可憑遙測技術輕易獲取大範圍且高精度的資料，故其與物種之間的關係應是研究前期的首要核心。本研究共使用八種地形相關變數，分別為海拔、

坡度、坡向、曲率、太陽總輻射量 (global solar radiation, GSR)、地形位置指標 (topographic position index, TPI)、地形粗糙度指標 (terrain ruggedness index, TRI) 和地形濕度指標 (topographic wetness index, TWI)。這些因子皆是藉由內政部地政司衛星測量中心的 1 m 解析度空載光達數值高程模型 (digital elevation model, DEM) 求導計算而出。

2.3.1 海拔

海拔圖層係使用 DEM 所紀錄之高程資訊。一般而言，隨著海拔高度的提升，氣溫將逐漸降低。植物受其影響，植群呈垂直分布變化，低、中和高海拔地區的天然植物組成截然不同。

2.3.2 坡度、坡向與曲率

坡度 (slope) 會影響土壤水分和養分含量。坡度大的地區，土壤易被侵蝕，而在坡度較小的地區，土壤與水分相對容易積聚，環境通常顯得較為潮濕。坡向 (aspect) 為某一位置最陡下坡所面對的羅盤方向，不同的坡向，將導致溫度、日照、濕度與土壤水分出現差異 (蘇鴻傑，1987)。以北半球而言，南向與西南向坡通常較北向與東北向坡來的溫暖乾燥。本研究參考邵寶燁 (2020) 的做法，將坡向進行八方位轉換，各類別的角度範圍如表 1 所示。

曲率 (curvature) 可判斷地表的某一部分是凸還是凹，並可描繪水流侵蝕和徑流形成的過程。曲率數值越大，代表在凸坡上，負值越小代表在凹坡上，零值則在平坦面。坡度、坡向與曲率三圖層皆是使用 ArcGIS 10.3 的空間分析模組 (spatial analyst) 分別從 DEM 求導而出。

表 1 坡向類別的代碼 (邵寶燁，2020)

類別代碼	坡向	角度範圍
0	平坦	坡度為 0° 平坦地無坡向
1	北方	0° - 22.5°; 337.5° - 360°
2	東北方	22.5° - 67.5°
3	東方	67.5° - 112.5°
4	東南方	112.5° - 157.5°
5	南方	157.5° - 202.5°
6	西南方	202.5° - 247.5°
7	西方	247.5° - 292.5°
8	西北方	292.5° - 337.5°

2.3.3 太陽總輻射量

太陽輻射左右植物之光合作用，是植物生長過程中的重要因素之一。而到達地面之輻射量可分為直接輻射 (direct radiation) 和漫射輻射 (diffuse radiation)，這兩種輻射量之加總就是太陽總輻射量 (GSR)。本研究試使用 ArcGIS 10.3 的 Area Solar Radiation 模組自 DEM 推導出試區的 GSR 圖層。該資料的單位為瓦特小時每平方米 (Wh/m²)。

2.3.4 地形位置指標

地形位置指標 (TPI) 能夠描述某一地點之位置與當地地形起伏之相對關係，如山頂、谷底、山腰或平原。TPI 與植物的生長有密切關連，通常越接近稜線，周遭環境之土壤、養分及水分会較為缺乏，而在這些物質較易積累於山谷處，光照量則會顯得相對較少。

TPI 的計算是根據一個指定大小的核心視窗 (kernel size)，在這個視窗內，對每一個像元的高程減去所有像元高程的平均值，所得到的差值就是該像元的 TPI 值，其公式如下 (De Reu *et al.*, 2013):

$$TPI = z_0 - \bar{z} \dots\dots\dots(1)$$

$$\bar{z} = \frac{1}{n_R} \sum_{i \in R} z_i \dots\dots\dots(2)$$

z_0 為中心像元的高程； \bar{z} 是平均海拔； R 為核心視窗之半徑； n_R 是視窗內的像元總數； z_i 則是第 i 個像元的高程。如果一個地區的 TPI 值為負數，說明該地區相對於周圍環境比較低，若是正數，則該地相對高於周圍環境。而出現 0，則表示該地區與周圍環境的高度基本相等，在坡度緩的地方則為平地，但若是在坡度陡峭的山區，此值表示該地位於山腰或以上處。

此外，核心視窗的大小將大幅影響相同地點的 TPI 數值。小的視窗半徑將呈現細緻的微地形，大的視窗半徑則呈現大尺度特徵。本研究是使用 SAGA GIS 的模組進行試區 TPI 的計算，此模組還可使用與中心網格的距離作為權重，在設定時本研

究選擇距離倒數冪次加權法 (inverse distance to a power)，權重使用其預設值 1。視窗半徑的部分，本研究分別設置為 100、200 及 400 pixels，以三種 TPI 圖層 (TPI_100、TPI_200、TPI_400) 測試不同空間尺度特徵對蘇鐵蕨模型的影響性。

2.3.5 地形粗糙度指標

地形會影響濕度和陽光照射，以及侵蝕和沈積等過程，地形粗糙度指標 (TRI) 即可用於測量地形的異質性。TRI 透過計算以對應地點為中心的正方形內子網格高度的標準偏差來量化表面的粗糙度，詳細公式如下 (Zhou *et al.*, 2019)：

$$TRI = \sqrt{\frac{1}{n} \sum (H_{ij} - \bar{H})^2} \dots\dots\dots (3)$$

其中， n 表示視窗範圍內之網格數； H_{ij} 為各點之高度值； \bar{H} 則是視窗範圍內各點高度值之平均。此指標同樣是 DEM 衍伸的地形相關變數，而本研究是使用 Qgis 的 GDAL 外掛模組進行 TRI 圖層的生產。

2.3.6 地形濕度指標

土壤水分控制著環境作用與物種的分布，然而卻很難進行大範圍的實地測量與空間內插。因此從 DEM 推導出的地形濕度指標 (TWI) 通常會被作為土壤濕度與植物可用水分的替代指標 (Kopecký *et al.*, 2021)。TWI 數值較高的區域相對上會比數值低的區域來的濕潤。本研究是使用 GRASS GIS 的 r.topidx 模組進行 TWI 的運算，此指標之計算公式如下 (賴韋愷，2016)：

$$TWI = \ln\left(\frac{\alpha}{\tan\beta}\right) \dots\dots\dots (4)$$

其中， α 為該單位等高線上游集流面積； β 為該網元之坡度。上凸坡形因為位處上坡處，會因重力流入下坡而不易累積水分；下凹坡形則相反。

2.4 取樣設計

本研究獲得的兩筆蘇鐵蕨樣本調查時期相差近 15 年，因此可藉由兩種取樣設計法 (sampling design, SD) 探討模型的時間可轉移性。

SD1：訓練與測試樣本皆為 2006 年之蘇鐵蕨植株，經抽取出四倍的背景樣本後，以 70% 的資料為訓練集，剩餘的 30% 為測試集。

SD2：訓練集使用的是與 SD1 相同的數據，但測試樣本改成使用 2021 和 2022 年所獲得的資料，旨在測試 SDM 預測未來蘇鐵蕨空間分布的能力。

2.5 建模方法

本研究以最大熵值法 (maximum entropy, MAXENT)、隨機森林 (random forest, RF) 和支持向量機 (support vector machine, SVM) 這三種常見的機器學習法以及深度學習法 U-net 建立 SDM。

用於建立蘇鐵蕨 SDM 的軟體來源有四，MAXENT 使用線上免費軟體 MAXENT 3.4.4 版本進行操作，此軟體可免費自其官方網站下載 (https://biodiversityinformatics.amnh.org/open_source/maxent/)。SVM 與 RF 使用的軟體 EnMAP-Box 3 是開源 GIS 軟體 Qgis 的外掛模組，只要在 Plugins\Not installed 介面中輸入該軟體之名稱後即可下載使用。U-net 部分則是使用 ArcGIS pro 2.8 版本進行模式的建立。

2.5.1 最大熵值法

最大熵值法 (MAXENT) 的原理是在一組約束條件的限制下 (我們不完全瞭解目標物種的實際分布狀態)，通過找到最大熵的機率分布 (即最分散或最接近均勻的機率分布) 來估計目標物種可能的分布情形 (Phillips *et al.*, 2006)。此演算法對於小樣本數和應對存在數據的取樣偏差時比起其他建模方式更加具有穩健性 (robustness) (Bushi *et al.*, 2022)。也因此，其是目前使用最廣泛的 SDM 演算法之一，尤其受到林業研究者的青睞 (Pecchi *et al.*, 2019)。

在進行 MAXENT 的參數設定時，本研究僅依照主體樣本的不同調整背景樣本的數目大小，其餘參數皆使用預設值。輸出之格式選擇邏輯轉換 (logistic) 來產生物種之分布機率圖，而在轉換為二元分類圖則選取 maximum training sensitivity plus specificity 為閾值，因在測試後發現此閾值最能顯現蘇鐵蕨的潛適分布。

2.5.2 隨機森林

隨機森林 (RF) 是一種集成式學習 (ensemble learning) 的機器學習法。此演算法的主要思想是將多個決策樹 (decision tree, DT) 集成在一起，每個 DT 在一個不同的隨機樣本和特徵集上訓練。最後，將所有 DT 的預測結果進行投票，採用多數決的方式指定所屬類別。

在參數的設定上，本研究參考邵寶燁 (2020) 的作法，將決策樹數量 ($n_estimators$) 設定為 500，並調整最大樹深度 ($max\ tree\ depth$)、母節點 ($parent\ node$) 與子節點 ($child\ node$)。其餘參數皆使用預設值。

2.5.3 支持向量機

支持向量機 (SVM) 是一種無母數統計分類法，其優勢在於其不需要過多的訓練樣本即可進行運算，也能處理高維度的數據。此演算法的主要原理是在特徵空間中尋求具最大邊界的分割超平面 (hyperplane) 以區分不同的二元類別 (陳承昌、史天元, 2007)。非線性支持向量機透過核函數 (kernel function) 將原始資料轉換到另一個高維度的特徵空間中，然後在特徵空間上執行線性分類。常用的核函數有線性函數 (linear)、多項式函數 (polynomial)、放射型核函數 (radial basis function,

RBF) 及 S 型核函數 (sigmoid) (黃凱翔, 2008)。本研究是採用模組的預設值，以放射型核函數進行運算。

2.5.4 U-net

U-net 是一種全卷積神經網絡 (fully convolutional networks, FCN)，由於它能夠對輸入影像進行像元級別的分類，因此也適合用於衛星影像的分類 (Flood *et al.*, 2019)。而建立 SDM 基本上就是在進行影像分類的動作，也因此本研究選擇此演算法來模擬蘇鐵蕨的空間分布。

不同於前面三種機器學習法，U-net 是以網格式態的影像切片作為訓練樣本，因此不能直接使用蘇鐵蕨樣本 (向量的點型態) 進行模式的訓練。而本研究的作法是先以 MAXENT、SVM 與 RF 獲得蘇鐵蕨在試區的潛勢分布圖 (圖 3 (a))，接著挑選一塊包含訓練上述模式時所使用的蘇鐵蕨樣本之矩形區塊作為 U-net 的訓練區 (圖 3 (b))，當完成訓練後再將成果外推至整個研究區。

在參數設定上，本研究基本上沿用預設值，但會統一將批次大小 (batch size) 設為 4，max epochs 則在每次訓練時都進行調整，直到得到最佳訓練成果。而在訓練時僅有 90% 的樣本會用於訓練，剩下的 10% 用於驗證。

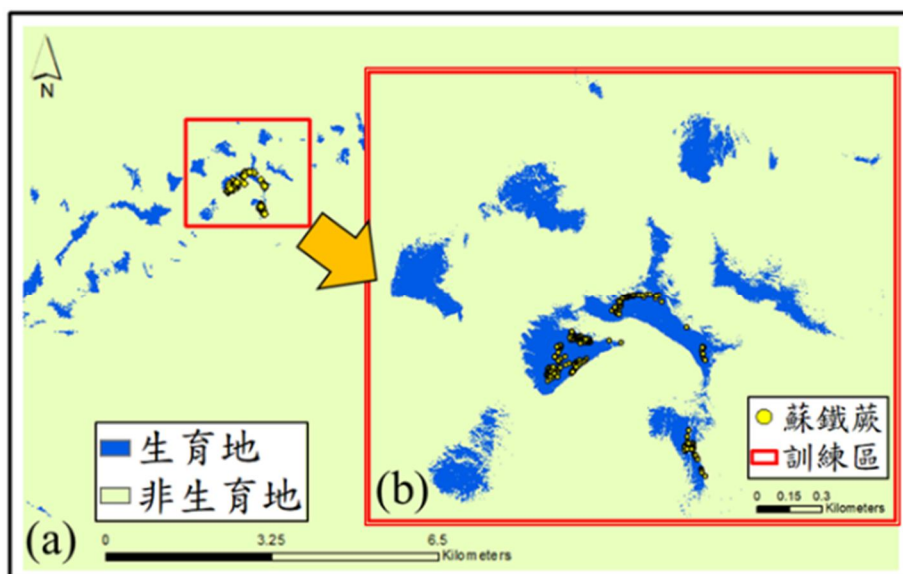


圖 3 生產 U-net 訓練樣本之示意圖 (a) 以機器學習法 (MAXENT、RF、SVM) 獲得蘇鐵蕨之潛勢分布圖、(b) 選擇出一塊矩形訓練區

2.6 模式驗證

模式的驗證非可以確認我們所構建的模型的可靠程度，以及對新樣本的預測能力。本研究採用分割樣本驗證，以 70% 的樣本用於訓練，剩餘的 30% 為測試樣本。訓練與測試樣本是完全相互獨立，因此測試結果會客觀且可靠。U-net 則與其他三種常見的機器學習法不同，一般會將數據及切割成訓練集、驗證集與測試集三個部分 (邵寶嬋, 2020)，在介紹 U-net 參數設定時有提及，訓練集與驗證集的比例為 90 : 10，測試集的部分則是選用和其他三個模式相同的測試樣本，如此一來再進行精度評估時才能有相同的比較基準。而若有測試樣點落於 U-net 的訓練區中，本研究會另外將其標記於 IGNORE_CLASSES 欄位中，避免被納為模型的訓練樣本。

本研究使用的評估指標分別是實際和預測類別相符的一致性係數之 $kappa$ 值，以及真實技術統計 (TSS)。 $kappa$ 值的範圍是在 -1 到 1 之間，但通常會介於 0 和 1，此指標的優點包括其簡單性，以及同時考慮誤授誤差 (commission error) 與漏授誤差 (omission error) (Manel *et al.*, 2001)。

TSS 於 2006 年由 Allouche 等人所提出，該指標的一大特色是其改善 $kappa$ 對盛行率 (prevalence; 即主體與背景的比例) 過於敏感的問題，但仍保有其原有的優點。也因此，TSS 被建議使用為 SDM 研究的評估指標 (陳厚昌, 2013)。TSS 的值域同樣是介於 -1 到 1，當數值為 1 時表示預測結果與實際情

形完全一致而值為 0 或小於 0 則表示性能不比隨機好 (Allouche *et al.*, 2006)。

3. 結果討論

3.1 環境因子敘述統計

表 2 為蘇鐵蕨及試區的環境因子統計值。在海拔的部分，蘇鐵蕨僅分布在惠蓀林場的低海拔地區，平均生長在海拔 763 m 之區域。由於其屬於陽性植物，故其 GSR 平均值明顯高於試區，充分展現其需光亮高的特性。也因此，TPI 顯示，其主要生長在山腰以上之區域，並且也能出現在山脊或山頂的位置。蘇鐵蕨多半出現在東南坡偏南向坡，儘管有少數生長在北向坡的植株，但這些植株之生長位置接近山脊，故仍然能接受充沛的日照。而坡度及 TRI 說明，過於崎嶇陡峭地形，導致土壤含水量及養分偏低的情況下，蘇鐵蕨可能不易生長，因 TWI 顯示，適度的水分才有助於此物種的生存。

3.2 各因子對模式之相對重要性

在執行 SDM 的運算時，各個環境變數對不同的模式將有不同的貢獻度。本研究使用 MAXENT 的折疊刀法試驗 (jackknife test) 對各變數的重要性進行評價。折疊刀法試驗會依序移除每個變數，使用剩餘變數建立模型，再單獨以每個變數建立模型，最後與納入所有變數的模型以橫條圖的方式進行比較。

表 2 蘇鐵蕨之環境因子敘述統計

DEM (1 m)	蘇鐵蕨				全區 (背景)			
	平均值	最大值	最小值	眾數	平均值	最大值	最小值	眾數
海拔 (m)	763.6	823.4	696.8	764.7	1294.3	2420.2	440.0	1735.6
坡度 (°)	32.4	58.6	2.7	39.4	38.3	76.2	0.8	48.6
坡向 (°)	4.8	8	1	4	4.8	8	1	7
曲率 (m ⁻¹)	4.7	171.0	-70.0	-2.0	0.5	370.0	-263.0	0.0
TWI	3.66	6.61	-0.36	3.07	3.84	10.07	-1.26	3.35
GSR (Wh/m ²)	1,546,422	1,732,420	643,142	1,584,920	1,362,241	1,998,860	324,450	1,401,920
TRI	0.51	1.29	0.05	0.47	0.68	3.15	0.01	0.58
TPI_100	1.32	3.43	-1.03	1.28	0.03	3.22	-4.36	0.36
TPI_200	1.57	3.52	-0.89	1.65	0.00	3.10	-3.37	0.03
TPI_400	1.54	3.12	-0.11	1.81	0.00	3.06	-2.86	0.26

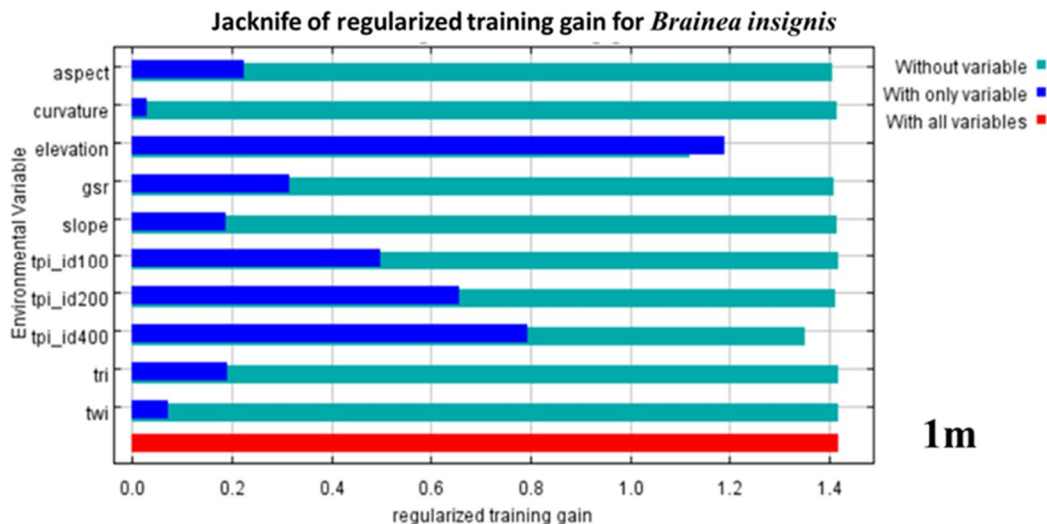


圖 4 MAXENT 之折疊刀法試驗結果

圖 4 為 MAXENT 之折疊刀法試驗結果。該圖顯示，海拔是所有變數中重要性最高者，若缺少此變數，模型精度將大幅下降。而繼海拔之後，三種 TPI 將能獲得次佳的訓練增益 (training gain) 效果，其中又以 TPI_400 的貢獻度最高。其餘變數中以 GSR 的重要性較高，惟曲率和 TWI 對模型則幾乎無貢獻，因此在進行模式的運算時會移除這兩個變數。

3.3 U-net 訓練區

圖 5 為取樣設計 SD1 的 U-net 訓練區位置圖，範圍涵蓋 Pro XR 在惠蓀林場的所有樣本。此訓練區是一個 2435×2070 m 的矩形，代表在 1 m 解析度時總共有 5,040,450 個像元能作為訓練樣本。當然此訓練區無可避免的會將 R12 測量之樣點包含在內，為避免這些點被納入訓練樣本，必須將這些資料統一為單獨的類別，並標記在模型的 IGNORE_CLASSES 中。

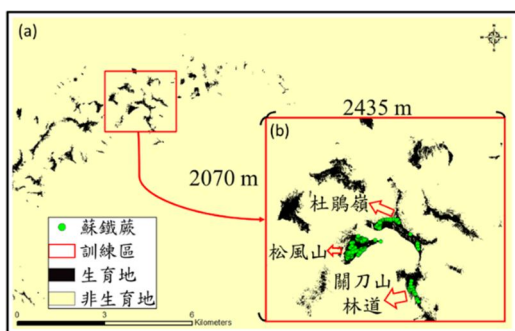


圖 5 U-net 訓練區之位置 (a) RF 預測分布圖、(b) U-net 訓練區

3.4 模型表現

3.4.1 以相同時期的資料訓練與測試模型

表 3 為取樣設計 SD1 之模式精度。當使用 2006 年調查到的所有樣本作為訓練與測試樣本時，獲得的準確度評估指標數值皆顯示模型表現相當優異，*kappa* 及 TSS 均在 0.90 以上。值得一提的是，U-net 大多能獲得良好的訓練成果。雖部分與原始的演算法相比雖有劣化，但 MAXENT 模式在經過 U-net 的深度學習後產出的 SDM 具有更高的生產者精度，故模型整體精度有所提升，因而成為所有模型中表現最佳者。證實以常見的機器學習法，提供更多的訓練樣本給深度學習進行 SDM 的構建確實可行。

圖 6 為 MAXENT_U-net 模式之預測分布圖，其推測蘇鐵蕨適合生長在試區的西北角，特別是北港溪兩側的山嶺中。而此模式幾乎都能正確預測到蘇鐵蕨在三個生育地的分布位置，僅有在杜鵑嶺北側有少許漏授樣本。

表 4 顯示了各模式預測出的蘇鐵蕨潛在生育地面積。其顯示蘇鐵蕨的生育地不會超過研究區總面積的 2%。此結果符合其為稀有植物的特徵，表明適合其生長的區域十分有限。若比較 U-net 與其他機器學習法之結果可發現，除 MAXENT_U-net 小幅增加面積之外，其餘的 U-net 模型都能夠有效減少蘇鐵蕨的預測生育地面積。由於這些模型的精度皆

相當優良，預測生育地減少表示有助縮減需要實地調查的區域，讓研究者能鎖定蘇鐵蕨的高潛勢分布地。

表 3 各模式於取樣設計 SD1 之模型精度

模式	SD1	
	<i>kappa</i>	TSS
MAXENT	0.97	0.97
RF	0.95	0.94
SVM	0.97	0.97
MAXENT_U-net	0.98	0.98
RF_U-net	0.93	0.90
SVM_U-net	0.93	0.92

表 4 各模式於取樣設計 SD1 推估出的蘇鐵蕨生育地面積

模式	面積 (ha)	試區百分比 (%)
MAXENT	176.8	1.0
RF	160.1	0.9
SVM	236.9	1.3
MAXENT_U-net	184.9	1.0
RF_U-net	145.5	0.8
SVM_U-net	185.8	1.0

3.4.2 模型的時間轉移能力

表 5 呈現了取樣設計 SD1 以及利用 R12 樣本進行驗證的結果。很明顯的，各模式的精度皆下降。雖然 MAXENT、SVM 和 MAXENT_U-net 之 TSS 值仍可保持在 0.85 以上，但 RF 下滑至 0.74，SVM_U-net 更是僅剩 0.66。因時間與資源因素，2021 及 2022 年的調查仍是在已確認有蘇鐵蕨分布的松

風山與杜鵑嶺進行樣本的採集。然而即便如此，模式的精度仍然顯著下滑，說明模型時間轉移性的不易。

對於蘇鐵蕨 SDM，造成及時間透通性不佳的原因很可能是兩期樣本的環境變數 TPI 特徵不同所造成。TPI 數值越高，表示越接近山頂，而圖 7 顯示，2006 年獲得的樣本主要分布在山腰以上至山頂之區域，僅有部分位於松風山之植株是出現在山腰。然而 15 年後，以 R12 定位的樣本雖有部分是生長在山頂之區域，但也有為數不少之植株是出現在杜鵑嶺及松風山的山腰處，甚至在杜鵑嶺更有生長在略低於山腰者。

TPI 屬於蘇鐵蕨 SDM 的重要變數，因此當演算法僅接收到 Pro XR 樣本之訊息，便無法正確拿捏蘇鐵蕨於坡面位置之分布下限，認定山腰以下之區域不適合蘇鐵蕨生長。時間轉移性表現較好者如 MAXENT (圖 7 (b)) 與 SVM 雖能多能正確預測到山腰的植株，但仍會遺漏杜鵑嶺山腰以下之樣本。至於 RF_U-net 與 SVM_U-net (圖 7 (c)) 對於分布下限的認定較為嚴格，山腰之區域多半被視為非生育地，因此在松風山東北角及杜鵑嶺東南邊出現許多遭遺漏的樣本。此現象說明單靠屬於間接作用因子的地形變數建立 SDM 時，其解釋能力可能不足夠，應尋找其他更具因果關係的因子如植生指標，才能提升 SDM 的時間通透性。

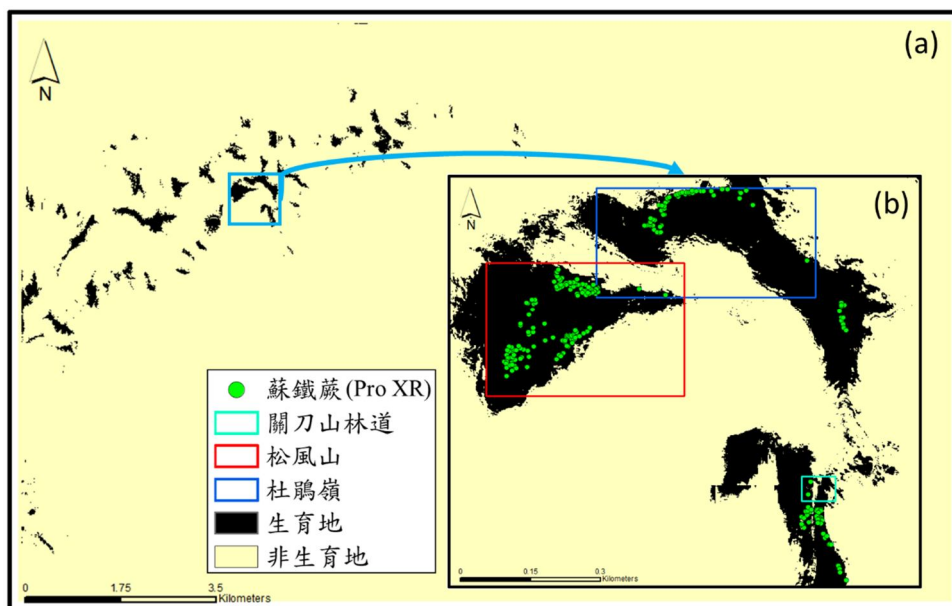


圖 6 (a) 以 MAXENT_U-net 建立的蘇鐵蕨 SDM、(b) 該模式於松風山、杜鵑嶺及關刀山林道的預測結果

至於蘇鐵蕨往山腰分布的原因，本研究推測與降水型態的改變有關。根據水利署惠蓀雨量站所紀錄 39 年來之降水資訊，雖年平均降雨並未明顯變化，但降雨型態有轉為雨季短且雨量集中、旱季長的趨勢。植物的有效降水減少造成林中枯立木明顯增加。而這種現象反倒有利於蘇鐵蕨生長，乃因此植物耐旱，能適應此種降雨特徵。枯立木的出現則增加森林的透光率，讓蘇鐵蕨能充分接收陽光的區域變多，使其族群能向山腰擴展。

隨著時間的推移，蘇鐵蕨的空間分布有逐漸朝向山腰分布的趨勢。而若只進行一次的現場調查是不可能觀察到此現象的。這即指出精良 SDM 的建立是一個迭代 (iterative) 的過程，完成首次模擬後必須重返現場，檢核第一輪模型與物種實際分布情形的相符程度，此即第二次現場驗證。當蒐集更多

的樣本後再進行第二次的模擬，並重複此流程，反覆修正直到獲得最佳模式。

本研究之成果同步證明二次現場驗證之重要性。單憑一輪的模擬容易高估 SDM 的預測性能，研究者必須由虛擬數位世界的預測成果中返回實際世界檢核模型的可靠性。若條件許可，甚至要進行更多次的現場實證。

表 5 取樣設計 SD1 和 SD2 之模型精度

模式	SD1		SD2	
	kappa	TSS	kappa	TSS
MAXENT	0.97	0.97	0.91	0.86
RF	0.95	0.94	0.82	0.74
SVM	0.97	0.97	0.91	0.86
MAXENT_U-net	0.98	0.98	0.90	0.85
RF_U-net	0.93	0.90	0.76	0.68
SVM_U-net	0.93	0.92	0.75	0.66

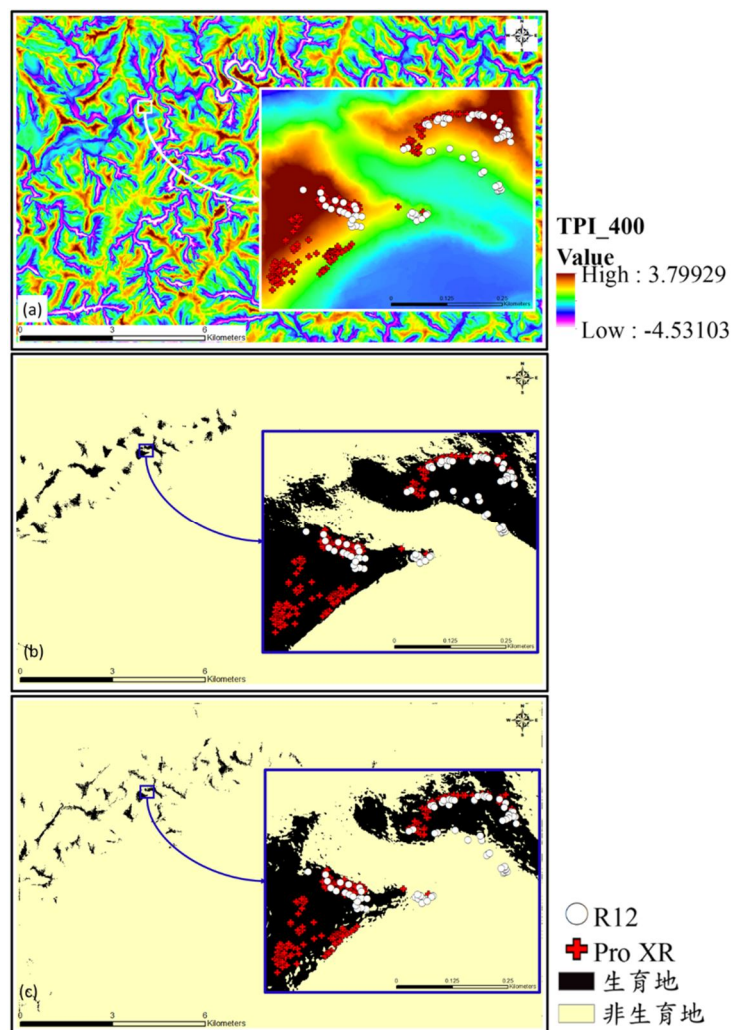


圖 7 環境變數 TPI_400 對模式在取樣設計 SD1 和 SD2 造成的影響 (a) R12 與 Pro XR 樣本於 TPI_400 之數值、(b) MAXENT 之預測分布圖、(c) SVM_U-net 之預測分布圖

4. 結論與建議

自然資源領域多需要在崇山峻嶺的環境實地採集資料，數據得來艱困且數量通常相當有限。然而藉由將機器學習與深度學習兩相結合後，本研究證實，此操作可在多數情況下成功訓練出可靠的深度學習模型，為 SDM 研究帶來新的可能性。對於建構蘇鐵蕨的 SDM，U-net 算是相當優異的演算法，其預測精度不差，推估出的潛在生育地面積又小，對於將有限的資源投入到新族群的探索有極大助益。利用深度學習建立 SDM 仍有許多值得探討的地方，本研究僅為初步嘗試，後續研究有必要採納更多種類的深度學習法，並持續進行試驗。

在本研究中，若訓練與測試樣本皆來自同一時期之調查成果，SDM 幾乎都能表現得相當優異。然而在預測 15 年後的蘇鐵蕨分布時，各模式之精度明顯下降，顯示模型的時間轉移性仍有待改善。四種演算法中以 MAXENT 和 SVM 有最佳的時間透過性，而 U-net 也有機會獲得頗佳的成果。而在環境變數方面，單靠地形因子使模式的轉移能力受限，故仍需找尋物種與環境有直接因果關係的生態因子，方有可能改善時間轉移性之困境。

可靠的 SDM 需藉由不斷的反覆迭代修正來生成。若條件許可，後續仍需多次往返虛擬預測結果與實體現場進行比對驗證，藉此提升 SDM 的應用價值，並能更好的預測未來蘇鐵蕨族群分布的變化。而未來也會將試區範圍擴大至惠蓀林場外，以測試評估深度學習法及其餘三種常用機器學習法在空間外推之效能。

參考文獻

呂金誠，1990。野火對臺灣主要森林生態系影響之研究，國立中興大學植物學系博士論文。[Lyu, J.C., 1990. Effects of wildfires on the main forest ecosystems in Taiwan, Doctoral Dissertation, National Chung Hsing University, Taiwan, ROC. (in Chinese)]

李國慶、劉長成、劉玉國、楊軍、張新時、郭柯，

2013。物種分佈模型理論研究進展，生態學報，33(16) : 4827-4835 , DOI : 10.5846/stxb201212031735。[Li, G.Q., Liu, C.C., Liu, Y.G., Yang, J., Zhang, X.S., and Guo, K., 2013. Advances in theoretical issues of species distribution models, Acta Ecologica Sinica, 33(16): 4827-4835, DOI: 10.5846/stxb201212031735. (in Chinese)]

邵寶燁，2020。地形屏障效應對不同海拔之樹種物種分布模擬預測之影響，國立中興大學森林學系碩士論文。[Shao, B.H., 2020. The impact of the terrain-shelterbelt effects on species distribution modeling for tree species at various altitude, Master Thesis, National Chung Hsing University, Taiwan, ROC. (in Chinese)]

許天銓、陳正為、Ralf Knapp、洪信介，2019。台灣原生植物全圖鑑第八卷(上)：蕨類與石松類，台北市：貓頭鷹出版社。[Hsu, T.C., Chen C.W., Knapp, R., and Hung, H.C., 2019. Illustrated Flora of Taiwan: Ferns and Lycophytes I, Owl Publishing House Co., LTD., Taiwan, ROC. (in Chinese)]

郭城孟，2001。蕨類圖鑑：台灣三百多種蕨類生態圖鑑，台北市：遠流出版公司。[Kuo, C.M., 2001. Ferns of Taiwan, Yuan-Liou Publishing Co., Ltd, Taiwan, ROC. (in Chinese)]

陳承昌、史天元，2007。支持向量機應用於水稻田辨識之研究，航測及遙測學刊，12(3)：225-240，DOI：10.6574/JPRS.2007.12(3).3。[Chen C.C., and Shih, T.Y., 2007. Rice paddy identification using the support vector machine, Journal of Photogrammetry and Remote Sensing, 12(3): 225-240, DOI: 10.6574/JPRS.2007.12(3).3. (in Chinese)]

陳厚昌，2013。物種分布模式預測功效影響因子之評估，國立中興大學森林學系碩士論文。[Chen, H.C., 2013. Evaluation of the factors affecting predictive performance of species distribution models, Master Thesis, National Chung Hsing

- University, Taiwan, ROC. (in Chinese)]
- 黃凱翔，2008。利用支持向量機於機載高光譜感測影像之分類，國立中興大學土木工程學系碩士論文。[Huang, K.S., 2008. Application of support vector machines to airborne hyperspectral image classification, Master Thesis, National Chung Hsing University, Taiwan, ROC. (in Chinese)]
- 詹雅婷、鄒裕民，2021。惠蓀林場蘇鐵蕨與土壤稀土元素分布之相關性，國立中興大學農業暨自然資源學院實驗林管理處補助計畫-成果報告，pp.217-233。[Chan, Y.T., and Tzou, Y.M., 2021. Rare earth elements (REEs) in soil and *Brainea insignis* at Huisun Experimental Forest Station, Project Report, National Chung Hsing University, Taiwan, ROC, pp.217-233. (in Chinese)]
- 蔡長宏，1997。關刀溪森林生態系干擾地更新之研究，國立中興大學植物學系碩士論文。[Tsai, C.H., 1997. Studies on the regeneration of disturbed sites at Guandaushi forest ecosystem, Master Thesis, National Chung Hsing University, Taiwan, ROC. (in Chinese)]
- 賴韋愷，2016。以地形濕度指標及太陽輻射指標搭配長用地形變數於首輪迭代物種分布模擬，國立中興大學森林學系碩士論文。[Lai, W.K., 2016. Modeling species distribution on an iterative basis with topographic wetness indices and solar radiation index along with common terrain variables in initial rounds, Master Thesis, National Chung Hsing University, Taiwan, ROC. (in Chinese)]
- 羅南璋、張偉顛、黃凱易，2011。應用 3S 技術及多變量統計於薯豆及卡氏槲潛在生育地之推估，林業研究季刊，33(3)：55-70，DOI：10.29898/SHBQ.201109.0005。[Lo, N.C., Chang, W.I., and Huang, K.Y., 2011. Application of 3S and multivariate statistics to predict the potential habitat of *Elaeocarpus japonicus* and *Castanopsis carlesii*, Quarterly Journal of Forest Research, 33 (3): 55-70, DOI: 10.29898/SHBQ.201109.0005. (in Chinese)]
- 蘇鴻傑，1987。森林生育地因子及其定量評估，中華林學季刊，20(1)：1-14。[Su, H.J., 1987. Forest habitat factors and their quantitative assessment, Quarterly Journal of Chinese Forestry, 20(1): 1-14. (in Chinese)]
- Allouche, O., Tsoar, A., and Kadmon, R., 2006. Assessing the accuracy of species distribution models: Prevalence, kappa and the true skill statistic (TSS), Journal of Applied Ecology, 43(6): 1223-1232, DOI: 10.1111/j.1365-2664.2006.01214.x.
- Baer, K.C., and Gray, A.N., 2022. Biotic predictors improve species distribution models for invasive plants in Western U.S. Forests at high but not low spatial resolutions, Forest Ecology and Management, 518: 120249, DOI: 10.1016/j.foreco.2022.120249.
- Brzozowski, M., Pelechaty, M., and Bogawski, P., 2022. A winner or a loser in climate change? Modelling the past, current, and future potential distributions of a rare charophyte species, Global Ecology and Conservation, 34: e02038, DOI: 10.1016/j.gecco.2022.e02038.
- Bushi, D., Mahato, R., Nimasow, O.D., and Nimasow, G., 2022. MaxEnt-based prediction of the potential invasion of *Lantana camara* L. under climate change scenarios in Arunachal Pradesh, India, Acta Ecologica Sinica, 43(4): 674-683, DOI: 10.1016/j.chnaes.2022.08.004.
- De Reu, J., Bourgeois, J., Bats, M., Zwertvaegher, A., Gelorini, V., De Smedt, P., Chu, W., Antrop, M., De Maeyer, P., Finke, P., Van Meirvenne, M., Verniers, J., and Crombé, P., 2013. Application of the topographic position index to heterogeneous landscapes, Geomorphology, 186: 39-49, DOI: 10.1016/j.geomorph.2012.12.015.
- El-Barougy, R.F., Dakhil, M.A., Halmy, M.W., Gray, S.M., Abdelaal, M., Khedr, A.-H.A., and Bersier,

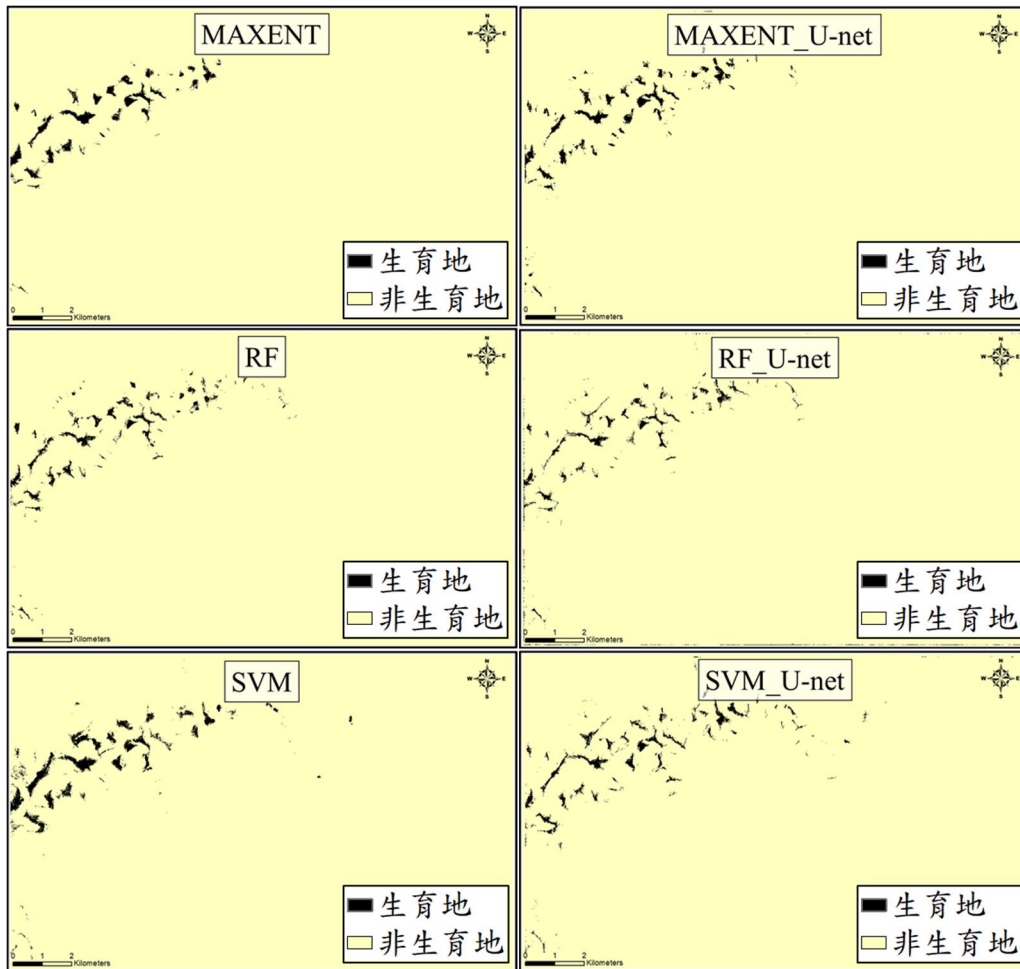
- L.-F., 2021. Invasion risk assessment using trait-environment and species distribution modelling techniques in an arid protected area: Towards conservation prioritization, *Ecological Indicators*, 129: 107951, DOI: 10.1016/j.ecolind.2021.107951.
- Feeley, K.J., and Silman, M.R., 2010. Modelling the responses of Andean and Amazonian plant species to climate change: The effects of georeferencing errors and the importance of data filtering, *Journal of Biogeography*, 37(4): 733-740, DOI: 10.1111/j.1365-2699.2009.02240.x.
- Flood, N., Watson, F., and Collett, L., 2019. Using a U-net convolutional neural network to map woody vegetation extent from high resolution satellite imagery across Queensland, Australia, *International Journal of Applied Earth Observation and Geoinformation*, 82: 101897, DOI: 10.1016/j.jag.2019.101897.
- Glad, A., and Mallard, F., 2022. Alpine marmot (*Marmota marmota*) distribution evolution under climate change: The use of species distribution models at a local scale in the western Pyrenees massif (France), *Ecological Informatics*, 69: 101646, DOI: 10.1016/j.ecoinf.2022.101646.
- Jensen, D.A., Rao, M., Zhang, J., Grøn, M., Tian, S., Ma, K., and Svenning, J.-C., 2021. The potential for using rare, native species in reforestation—A case study of yews (Taxaceae) in China, *Forest Ecology and Management*, 482: 118816, DOI: 10.1016/j.foreco.2020.118816.
- Kopecký, M., Macek, M., and Wild, J., 2021. Topographic Wetness Index calculation guidelines based on measured soil moisture and plant species composition, *Science of the Total Environment*, 757: 143785, DOI: 10.1016/j.scitotenv.2020.143785.
- Maciel, E.A., Oliveira-Filho, A.T., Sobral-Souza, T.S., Marimon, B.S., Cupertino-Eisenlohr, M.A., José-Silva, L., and Eisenlohr, P.V., 2021. Climate change forecasts suggest that the conservation area network in the Cerrado-Amazon transition zone needs to be expanded, *Acta Oecologica*, 112: 103764, DOI: 10.1016/j.actao.2021.103764.
- Manel, S., Williams, H.C., and Ormerod, S.J., 2001. Evaluating presence-absence models in ecology: The need to account for prevalence, *Journal of Applied Ecology*, 38(5): 921-931, DOI: 10.1046/j.1365-2664.2001.00647.x.
- Pecchi, M., Marchi, M., Burton, V., Giannetti, F., Moriondo, M., Bernetti, I., Bindi, M., and Chirici, G., 2019. Species distribution modelling to support forest management. A literature review, *Ecological Modelling*, 411: 108817, DOI: 10.1016/j.ecolmodel.2019.108817.
- Phillips, S.J., Anderson, R.P., and Schapire, R.E., 2006. Maximum entropy modeling of species geographic distributions, *Ecological Modelling*, 190(3-4): 231-259, DOI: 10.1016/j.ecolmodel.2005.03.026.
- Su, S.Y., Lo, N.J., Chang, W.I., and Huang, K.Y., 2012. Two way extrapolation and validation on ecological patterns of *Elaeocarpus japonicus* between main watersheds in Huisun of central Taiwan, *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 39: 357-362, Melbourne, Australia, DOI: 10.5194/isprsarchives-XXXIX-B8-357-2012.
- Wang, A., Melton, A.E., Soltis, D.E., and Soltis, P.S., 2022. Potential distributional shifts in North America of allelopathic invasive plant species under climate change models, *Plant Diversity*, 44(1): 11-19, DOI: 10.1016/j.pld.2021.06.010.
- Yousefi, M., Naderloo, R., and Keikhosravi, A., 2022. Freshwater crabs of the Near East: Increased extinction risk from climate change and underrepresented within protected areas, *Global*

Ecology and Conservation, 38: e02266, DOI: 10.1016/j.gecco.2022.e02266.
Zhou, Y., Guan, H., Huang, C., Fan, L., Gharib, S., Batelaan, O., and Simmons, C., 2019. Sea breeze

cooling capacity and its influencing factors in a coastal city, Building and Environment, 166: 106408, DOI: 10.1016/j.buildenv.2019.106408.

附錄

(1) 取樣設計 SD1 之蘇鐵蕨預測分布圖



Exploring the Temporal Transferability in the Species Distribution Model of *Brainea insignis* Using Two Sets of Data Separated by 15 Years

Ching-Jin Kuo¹ Bao-Hua Shao² Nan-Chang Lo³ Kai-Yi Huang^{4*}

Abstract

This study aims to investigate the impact of different algorithms on the temporal transferability of species distribution model (SDM) and the feasibility of using deep learning techniques to build SDM. The study focuses on *Brainea insignis* as the target species and utilizes two sets of samples collected with a 15-year interval. Experiments were conducted using the maximum entropy (MAXENT), random forest (RF), support vector machine (SVM), and U-net—a deep learning approach. The results indicate that MAXENT and SVM exhibit the best temporal transferability, while U-net also shows promising results. This highlights the research potential of deep learning, and future studies should consider incorporating a wider range of deep learning methods and continue experimentation. However, concerning environmental variables, relying solely on topographic factors may constrain the model's transferability, necessitating the identification of ecological factors with more direct causal relationships with the species to enhance transferability.

Keywords: Temporal Transferability, Deep Learning, Species Distribution Model, *Brainea insignis*

¹ Master Student, Department of Forestry, National Chung Hsing University

² Associate Technical Specialist, Pu-Li Workstation, Nantou Branch,
Forestry and Nature Conservation Agency

³ Division Chief, Silviculture Division, Experimental Forest Management Office, National Chung Hsing University

⁴ Professor, Department of Forestry, National Chung Hsing University

* Corresponding Author, E-mail: kyhuang@dragon.nchu.edu.tw

Received Date: Oct. 02, 2023

Revised Date: Oct. 29, 2023

Accepted Date: Nov. 27, 2023