

運用背景資料取樣策略增進物種分布模型的預測效能

Improving the predictive performance of species distribution models using background data sampling strategies

呂明倫

Ming-Lun Lu

農業部生物多樣性研究所 552005 南投縣集集鎮民生東路 1 號

Taiwan Biodiversity Research Institute, Nantou, Taiwan

通訊作者：alan@tbri.gov.tw

Corresponding author: alan@tbri.gov.tw

摘要

物種分布模型中，最大熵 (maximum entropy, MaxEnt) 的效能高度依賴輸入數據的質量，特別是物種出現紀錄的代表性，然而，實際研究中常因取樣偏差導致模型準確性和泛化能力下降。由於背景資料的選擇對於校正取樣偏差至關重要，本研究旨在比較不同背景資料取樣方法對 MaxEnt 模型性能的影響，並以

臺灣特有且易危的霧社櫻 (*Prunus taiwaniana*) 為研究對象，探討其分布預測的最佳策略。總共比較了五種背景資料取樣方法，包括目標群體密度、目標種密度、隨機取樣、受限距離及最小多邊範圍。結果顯示，目標群體密度和目標種密度方法表現佳，其 Boyce 指數都達到 0.9 以上，表示這兩種方法能有效校正取樣偏差並捕捉廣泛的環境條件。此外，從最佳模型的預測結果可知，霧社櫻偏好寒冷環境，並對特定的降水量與晝夜溫差等氣候條件具有高度敏感性，其高適宜棲地主要集中在臺灣中北部的中海拔山區。本研究強調合理選擇背景資料取樣策略的重要性，特別是對於稀有或受脅物種，適當的取樣方法不僅能顯著提升模型的準確性，還能为保育工作提供重要的科學依據。

關鍵詞：最大熵、取樣偏差、霧社櫻、目標群體密度、保育

Abstract

In species distribution models, the performance of maximum entropy (MaxEnt) highly depends on the quality of the input data, particularly the representativeness of species occurrence records. However, sampling bias in practical studies often leads to reduced model accuracy and generalization capabilities. Because the choice of background data is crucial for correcting sampling bias, this study aimed to compare the effects of different background data sampling methods on MaxEnt model performance, using the endemic and vulnerable plant species (*Prunus taiwaniana*) as the research subject to explore the best strategy for predicting its distribution. Five background data sampling methods were compared: target group density, target species density, random sampling, constrained distance, and minimum convex polygon. The results showed that target group density and target species density performed well, with Boyce index values exceeding 0.9, indicating that these two

methods effectively correct sampling bias and capture a wide range of environmental conditions. Furthermore, predictions from the best-performing model revealed that *P. taiwaniana* prefers cold environments and exhibits high sensitivity to specific climatic conditions, such as precipitation levels and diurnal temperature variation. The most suitable habitats are primarily concentrated in the mid-elevation mountainous regions of central and northern Taiwan. This study emphasizes the importance of selecting appropriate background data sampling strategies, particularly for rare or threatened species. Proper sampling methods not only significantly improve model accuracy but also provide crucial scientific support for conservation efforts.

Key words: maximum entropy, sampling bias, *Prunus taiwaniana*, target group density, conservation.

收件日期：2025 年 3 月 26 日

Received: March 26, 2025

接受日期：2025 年 9 月 30 日

Accepted: September 30, 2025

緒言

物種分布模型 (species distribution models, SDMs) 是生態學和生物地理學中用於預測物種潛在分布範圍的重要工具，這些模型利用環境資料與物種出現紀錄，建構物種與棲息環境之間的複雜關係，並為生物多樣性推估、保護區規劃、入侵物種管理及氣候變遷影響等應用提供科學依據 (Araújo *et al.* 2019; Velazco *et al.* 2020; Franklin 2023; Lu and Huang 2023)。SDM 的類型繁多，其中，最大熵 (maximum entropy, MaxEnt) 是一種基於機器學習的演算法，因其具備高效和穩定性，尤其適合處理小樣本數據，成為目前使用最廣泛的工具之一 (Morales *et al.* 2017)。MaxEnt 的核心理論源自資訊理論中的 Shannon 資訊熵 (Shannon 1948)，其基本假設是物種的分布應滿足對環境條件的最大化利用，同時最小化分布的不確定性。具體來說，MaxEnt 利用物種出現紀錄所在的環境條件，以及與研究區內所有可能環境條件 (由背景資料定義) 之間的相似性，推導出最

大化棲地適宜性的分布模型，從而生生成最接近真實分布的預測結果 (Phillips *et al.* 2006)。

MaxEnt 的效能高度依賴於輸入數據的質量，特別是物種出現紀錄的代表性。在實際研究中，物種出現數據往往受到取樣偏差的影響，即可能過度集中在易於到達的地區，如道路或城市附近，忽略了偏遠或難以抵達的區域 (Kadmon *et al.* 2004)，這種偏差會導致模型對某些環境條件的過度擬合，從而降低預測的準確度和泛化力 (Phillips *et al.* 2009; Kramer-Schadt *et al.* 2013)。為減少取樣偏差對模型的影響，背景資料 (background data) 的選擇變得至關重要。在 MaxEnt 模型中，背景資料被用來定義研究區內所有可能的環境條件，從而構建一個環境空間的參照框架 (Phillips and Dudík 2008)，幫助區分目標物種的實際分布與隨機環境條件下的分布型態，故背景資料的合理選擇可以確保模型涵蓋更廣泛的環境條件，而不僅僅局限於物種出現點附近的環境特徵。

背景點的取樣方式本身也可能引入新的偏差，例如，背景資料過於集中在某些特定環境條件下，可能會進一步放大抽樣偏差的影響，導致模型低估或高估物種對某些環境因子的偏好 (Barbet-Massin *et al.* 2012)。因此，如何設計合理的背景資料取樣策略，以平衡取樣偏差和環境異質性，成為提升 SDM 預測效能的關鍵挑戰之一。目前已有許多背景資料取樣方法被提出，如目標種密度、目標群體密度、受限距離及最小多邊範圍等 (Barber *et al.* 2022；Ahmadi *et al.* 2023；Sorbe *et al.* 2023；Schartel and Cao 2024)，每種方法都有其獨特的優勢和局限性，但至今尚無一種普遍適用的最佳背景點取樣策略。因此，比較不同背景資料取樣方法對模型性能的影響，對於提升 MaxEnt 模型的預測能力具有重要意義。

霧社櫻 (*Prunus taiwaniana*) 為一種臺灣特有的植物物種，主要分布在中北部中海拔山區，被收錄於國際自然保育聯盟瀕危物種紅色名錄，並被評定為易危等級。本研究以霧社櫻

為研究對象，探討不同背景資料取樣方法對 MaxEnt 模型性能與分布預測結果的影響，旨在尋找最適合該物種分布預測的背景資料取樣策略。目前國內針對此議題的研究相對較少，因此，本研究不僅有助於提升 MaxEnt 模型在特有物種分布預測中的應用效能，還能為其他類似物種的保育工作提供重要的科學依據。

材料與方法

一、物種出現紀錄與環境資料

本研究區範圍為臺灣本島，物種調查資料是來自全球生物多樣性資訊機構 (global biodiversity information facility, GBIF)，為與環境資料的時間間隔匹配，抽取 1993–2022 年間，霧社櫻的物種出現紀錄，共取得 69 筆，座標格式為經緯度，考量到地理座標的精確度，先刪除小數點不足 3 位的紀錄。由於蒐集方法多樣且非系統化，紀錄中可能存在空間偏差。為解決因調查點過於密集而可能產生的空間自相關

性問題，隨機選擇一個起始觀察點，並將其相鄰 1 km (符合預測變項之空間解析力) 的點移除，然後對剩餘的點重複此過程，使點的分布更加均勻，最終留下 44 個點做為建模的樣本 (圖 1)。

建模所使用之預測變項僅包含氣候、地形和土地覆蓋等 3 大類，氣候資料是從臺灣氣候變遷推估資訊與調適知識平台取得，本研究從平台上提取 1993–2022 年的溫度與降水量資料，進一步參考 Hijmans *et al.* (2005) 的計

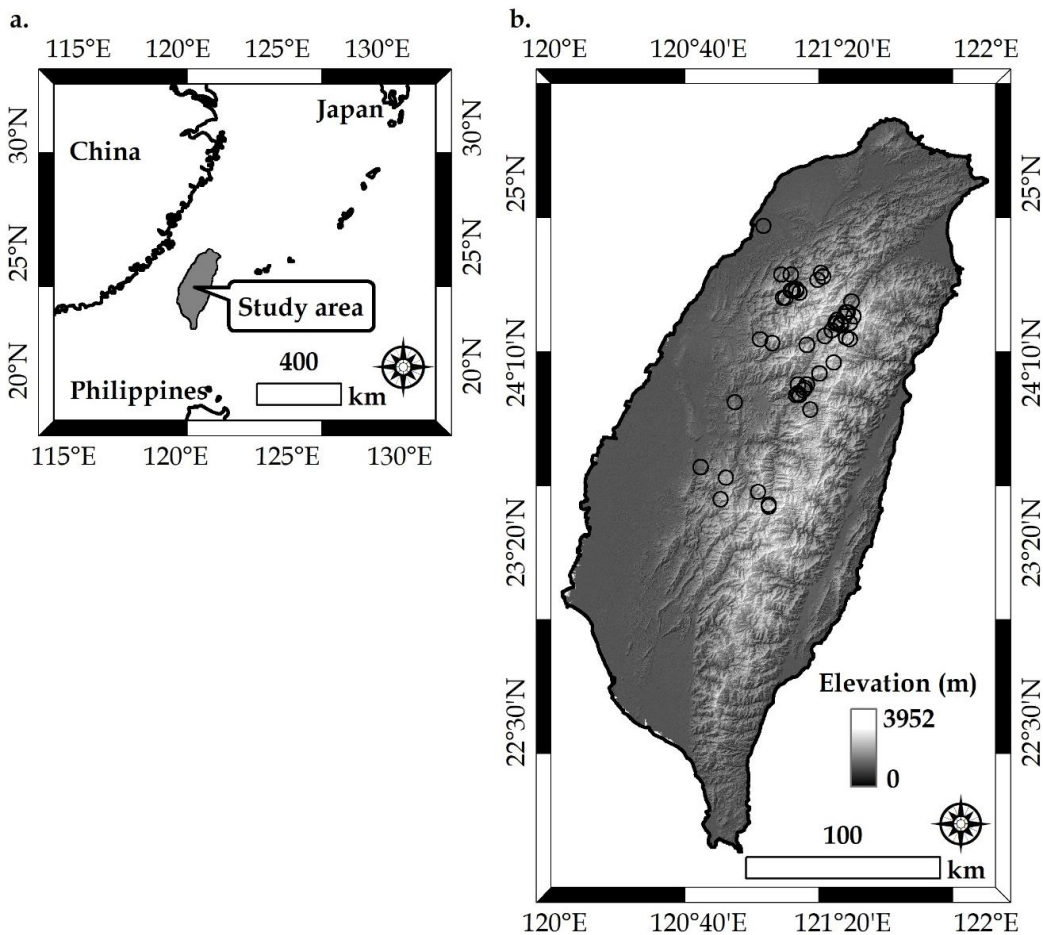


圖 1 (a) 研究區地理位置圖；(b) 物種出現點 (黑色標記； $n = 44$) 及數值高程模型 (digital elevation model)。

Fig. 1 (a) Geographical location of the study area; (b) Species occurrence records (black circle; $n = 44$) and digital elevation model.

算方法，產製 19 個空間解析力為 1×1 km 的生物氣候變項 (BIO1-BIO19)，其源自每月的溫度 ($^{\circ}\text{C}$) 和降水量 (mm)，包括年度趨勢、季節性變化及極端氣象特徵 (worldclim.org/data/bioclim.html)。地形方面，從內政部取得數值高程模型 (空間解析力為 20 m)，進一步利用 ArcGIS 10.6 軟體製作坡度與坡向，並將解析力重取樣為 1×1 km。土地覆蓋是取自 Zhang *et al.* (2023) 所建構的 2020 年全球土地覆蓋圖，包含森林、草地、農地、都市、荒地與水體等 6 種類型 (1×1 km 空間解析力)。總計共有 23 個預測變項 (附表 1)，為克服變項在統計上的多重共線性問題，以 R 軟件包「virtual species」執行 Pearson 相關分析 (Leroy *et al.* 2016)，排除相關性高的變項，當相關係數大於 0.7 以上時擇一保留 (Dormann *et al.* 2013)。

二、模型建構

利用 Phillips *et al.* (2006) 所開發的 MaxEnt 3.4.3k 建構 SDM，由於本研究透過不同的偏差校正方法進行比

較，因此，所有的超參數皆使用軟體的預設值，以減少人為調整對模型性能的潛在影響，從而更客觀地比較不同背景取樣方法的效能。將各霧社櫻出現的紀錄逢機選取 80% 做為訓練資料集，過程中使用 k 折交叉驗證法 (k-fold cross-validation) 確保模型的穩定性，重複進行 10 次 (10 折)，並取平均得到最後結果。MaxEnt 會產生值域介於 0–1 的概率圖，概率愈大者愈有可能是物種適宜的棲地，故可表示棲地適宜性。另一方面，剩餘 20% 的紀錄不參與訓練，做為測試資料集，用於評估模型。評估方法為 Boyce 指數 (Boyce index, BI) 和接受者操作特徵 (receiver operating characteristic) 所產生之曲線面積 (area under the curve, AUC)。BI 的計算使用 R 軟件包「ecospat」(Di Cola *et al.* 2017)，其為一種專門用於評估僅有物種出現紀錄模型的指標，透過比較 MaxEnt 輸出的概率與物種出現紀錄的關係來計算，值域介於 -1 到 1 之間，正值表示模型表現良好，0 表示模型與隨機預測無異，負值則表明

模型表現不佳 (Boyce *et al.* 2002)。AUC 則為 MaxEnt 內建產生，其值域介於 0.5–1，當值愈高代表模式推估出來的準確性愈佳。此外，模型解釋力可參考 MaxEnt 額外產生的變項置換重要性與反應曲線等訊息。

三、背景資料取樣

取樣方式包括隨機、目標種密度、目標群體密度、受限距離及最小多邊範圍等共計 5 種 (圖 2)，所有方法均產生 10000 個背景點 (MaxEnt 的預設值)，由於各種方法的取樣邏輯和空間分布的先驗假設截然不同。具體方法分述如下：

(一) 隨機取樣

本法為最傳統且最簡單的方法，即直接從研究區域內，以完全隨機的方式取樣。此法假設物種的取樣過程在區內是均勻的，不對任何環境條件或地理位置進行預設偏好，確保客觀性和可重複性 (圖 2a)。

(二) 目標種密度

本法旨在校正物種出現點的空間聚集偏差，依循 Elith *et al.* (2010) 的方法。首先，基於霧社櫻的訓練樣本，利用 ArcGIS 的核密度分析 (kernel density) 工具，推導表示「出現點密度」的空間圖層，該圖層的值越高，代表該區域的出現點越密集，並將網格值由小到大劃分成 20 個等距類別。隨後，以此密度圖層作為取樣的權重，利用 R 軟件包「terra」進行加權隨機取樣，在出現點密集的区域 (高密度)，被選為背景資料的機率也較高；反觀，在出現點稀疏的区域 (低密度)，被選中的機率較低 (圖 2b)。

(三) 目標群體密度

目標群體密度取樣法是上述目標種密度的擴展，旨在利用近緣物種的出現紀錄來定義一個更廣泛的生態相關環境，即假設梅屬 (*Prunus*) 物種的整體分布型，與霧社櫻具有相似生態需求的環境。與目標種密度不同，該法的密度分析基礎是從 GBIF 下載梅屬所有物種的出現紀錄，做為目標群體。接著，推導梅屬物種的核密度圖

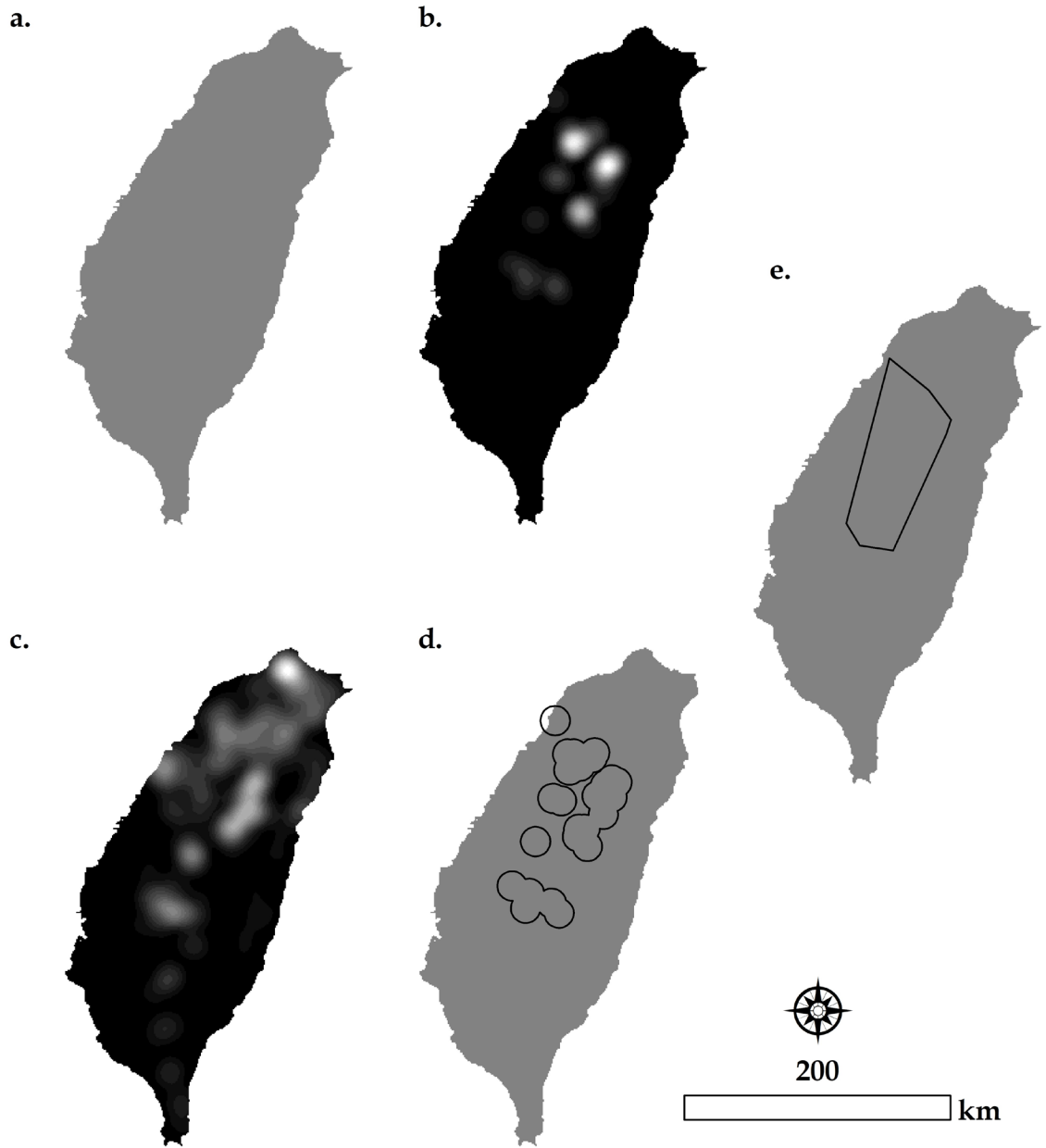


圖 2 不同背景資料的取樣方法，(a) 隨機取樣，(b) 目標種密度，(c) 目標群體密度，(d) 受限距離，(e) 最小多邊範圍。(b) 和 (c) 愈白處代表取樣強度愈高。

Fig. 2 Different background data sampling methods: (a) Random sampling, (b) Target species density, (c) Target group density, (d) Constrained distance and (e) Minimum convex polygon. (b) and (c): Whiter areas represent higher sampling intensity than other sampling sites.

層，再以此圖層作為權重，依循目標種密度的方法進行加權隨機取樣 (圖 2c)。

(四) 受限距離

本法將背景資料的取樣範圍限制在物種出現點的鄰近區域內，以避免涵蓋過多生態上不適宜的環境。具體而言，以每個用於訓練的霧社櫻出現點為圓心，畫出半徑為 10 km 的圓形緩衝區。所有背景資料的候選範圍被限定在這些緩衝區內，再進行完全隨機取樣。該法假設物種的潛在分布不會過度偏離其已知的出現地點 (圖 2d)。

(五) 最小多邊範圍

本法定義了一個由所有出現點所包圍的最小凸多邊形區域。首先，繪製霧社櫻訓練樣本的凸包 (convex hull)，形成一個涵蓋所有出現點的最小多邊形。所有背景資料的候選範圍被限定在此凸包範圍內，再進行完全隨機抽樣。該法假設物種的潛在分布不會超出其已知出現點所構成的最外圍邊界 (圖 2e)。

結果

各種背景資料取樣方法產生的模型，依據 BI 與 AUC 評估的結果顯示

表 1 不同樣本偏差校正方法的模型表現

Table 1 Model performance with different sample bias correction methods

Method	BI	AUC
Target group density	0.92	0.93
Target species density	0.91	0.91
Random sampling	0.89	0.91
Constrained distance	0.71	0.78
Minimum convex polygon	0.20	0.55

BI = Boyce index

AUC = Area under the receiver operating characteristic curve

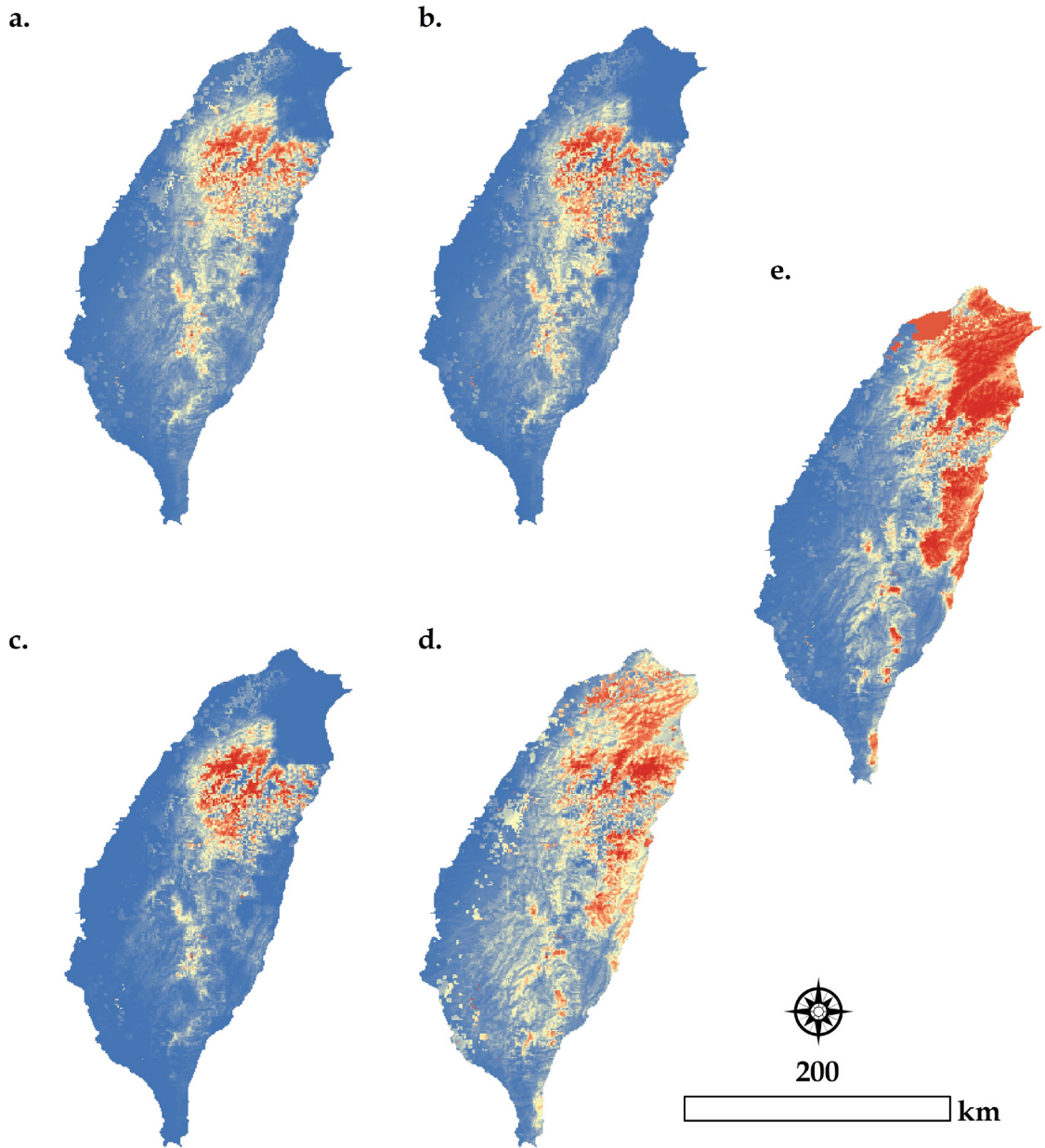


圖 3 不同背景資料取樣方法下霧社櫻 (*Prunus taiwaniana*) 的棲地適宜性預測結果，(a) 目標群體密度，(b) 目標種密度，(c) 隨機取樣，(d) 受限距離，(e) 最小多邊範圍。高適宜偏紅色，低適宜性偏藍色。

Fig. 3 Predicted habitat suitability for *Prunus taiwaniana* under different background data sampling methods: (a) target group density, (b) target species density, (c) random sampling, (d) constrained distance, and (e) minimum convex polygon. Higher and lower suitability is represented in red and blue, respectively.

(表 1)，不同取樣方法對模型預測能力有顯著差異，其中目標群體密度的效果最佳 (BI = 0.92；AUC = 0.93)，其次為目標種密度 (BI = 0.91；AUC = 0.91)，接續為隨機取樣 (BI = 0.89；AUC = 0.91)，而受限距離 (BI = 0.71；AUC = 0.91) 與最小多邊範圍的效果較不理想 (BI = 0.20；AUC = 0.55)。

從預測出的概率圖來看，目標群體、目標種密度與隨機取樣的空間分

布型態較相似 (圖 3a、b 及 c)，高適宜性區域主要集中在臺灣中北部的中海拔山區，這與霧社櫻的實際分布範圍趨同，但隨機取樣與其他兩種方法仍有差異，隨機取樣的預測結果略顯集中，空間細緻度較低，導致高適宜性的分布稍窄。受限距離與最小多邊範圍都局限了背景資料範圍，似乎無法涵蓋足夠的環境異質性，導致預測結果嚴重偏離實際情況 (圖 3d 和 e)。

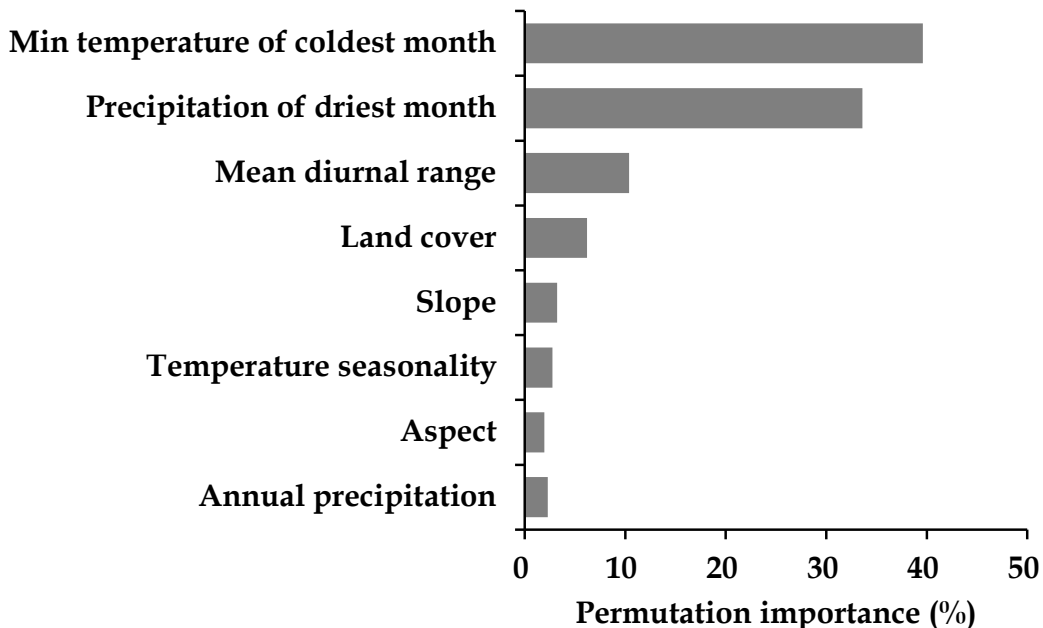


圖 4 MaxEnt 模型中環境變數對霧社櫻 (*Prunus taiwaniana*) 分布的重要性，基於目標群體密度取樣法。

Fig. 4 Importance of environmental variables in the MaxEnt model for *Prunus taiwaniana* distribution, based on the target group density sampling method.

由於目標群體密度的取樣方法可以獲得最佳模型性能，故以該法呈現最終結果。經相關分析篩選出的 8 個變項貢獻度分析顯示，霧社櫻的棲地適宜性主要受到少數幾個關鍵變項的影響。各變項中名列前三依序為最冷月最低溫、最乾月降水量與平均晝夜溫差，這些變項的重要性均高於 10%，合計貢獻了 84% 左右的模型解釋力，由此顯示，霧社櫻的分布型態主要由氣候條件決定。反觀，其他變項的重要性均低於 10%，對於模型的影響力相對有限，但仍可對模型的局部適宜性預測提供補充作用 (圖 4)。

圖 5 顯示主要影響霧社櫻分布變項的反應曲線，其中，最冷月最低溫度在 -5-5°C 間，棲地適宜性迅速上升並達到峰值，隨後隨著溫度升高，棲地適宜性逐漸下降圖 (5a)。最乾燥月降水量的最適範圍在 5-20 mm 的範圍內，隨後隨著降水量的進一步增加，棲地適宜性開始下降，超過 50 mm 明顯不適宜 (5b)。平均晝夜溫差在 8-10°C 為最適宜的環境條件，隨

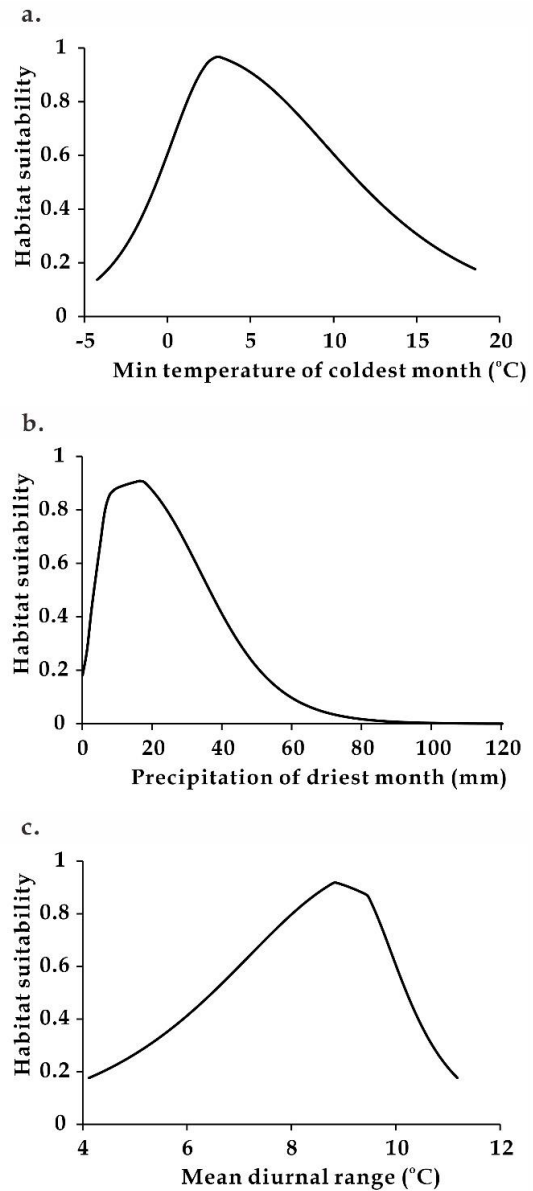


圖 5 MaxEnt 模型中霧社櫻 (*Prunus taiwaniana*) 對關鍵環境變數的反應曲線。

Fig. 5 Response curves of key environmental variables for *Prunus taiwaniana* in the MaxEnt model.

後曲線即呈現迅速下降 (5c)。綜合上述，霧社櫻對寒冷環境有高度的適應性，此外，特定的降水量與晝夜溫差表示，該物種對氣候條件的變化有高度敏感性。

討論

背景資料取樣策略的選擇會顯著影響 MaxEnt 模型的預測能力，從霧社櫻的棲地適宜性預測結果可知，目標群體密度和目標種密度方法的表現最佳，代表這兩種方法能夠涵蓋更廣泛的環境條件，有效校正取樣偏差，並建構高效的預測模型。目標群體與目標種密度都是基於物種本身的地理分布來構建參照框架，從而捕捉到與霧社櫻生態需求相似的環境特徵，本研究得到的結果與過往相關研究提出的論點一致，進一步驗證了這兩種方法在處理小樣本、空間偏差明顯的物種出現紀錄時的優勢 (Vollering *et al.* 2019; Barber *et al.* 2022; Ahmadi *et al.* 2023)，因此，目標群體密度和目標種密度是一種能有效抵消偏差的方

法，尤其適用於稀有物種的分布預測。相比之下，背景資料隨機取樣代表沒有考慮到樣本有偏差的問題，縱使在本研究中仍有不錯表現，但也存在引入過多雜訊的隱憂 (Kramer-Schadt *et al.* 2013)，即其背景點涵蓋了過多非適宜的環境，降低了模型的泛化能力，造成邊緣棲地的判斷趨於保守，故顯示隨機取樣法的性能不如前兩種方法。

受限距離和最小多邊範圍方法的表現相較不理想，這主要是因為這兩種方法過度局限於局部區域，無法充分反映霧社櫻對更廣泛環境的適應性。受限距離方法將背景點限制在訓練樣本半徑 10 km 的範圍內可能過於偏頗，進一步放大了取樣偏差的影響。另一方面，最小多邊範圍方法所制訂的背景範圍，可能因涵蓋過多超出霧社櫻生態幅度的區域 (如平原地區)，導致模型的泛化能力大幅下降。這些結果 VanDerWal *et al.* (2009) 與 Barber *et al.* (2022) 的試驗結果相符，都顯示過度限制背景資料範圍會顯著

降低模型的預測效能。因此，在實際應用中，應謹慎選擇背景資料的範圍，以平衡取樣偏差和環境異質性。

從目標群體密度取樣方法所建構的 MaxEnt 模型來看，氣候類變項對模型的影響力最為顯著。其中，最冷月最低溫是決定霧社櫻在空間上存在與否的關鍵因子。詳細的偏好棲地特徵顯示，霧社櫻與其他梅屬植物物種相似，都偏好低溫的寒冷環境。這是因為其生理上需要累積一定程度的冷量來突破自然休眠，因此，維持一定的低溫水平對其生長發育週期至關重要 (Gao *et al.* 2012 ; Zhang *et al.* 2023 ; Huang 2024)。此外，降水量多寡與晝夜溫差也都是重要因子，冬季降水超過 50 mm 或晝夜溫差極端變化的情況均不利於霧社櫻的生存，可能會對其帶來負面影響。這些結果進一步揭示了霧社櫻對氣候條件的敏感性，也解釋了其潛在分布範圍受限的原因，並且更意味著其對氣候變遷和人類活動的干擾可能更加敏感。根據 MaxEnt 模型的預測結果，霧社

櫻的高適宜棲地面積並不大，主要集中在臺灣北部的棲蘭山、雪山與太平山一帶，南部則較為零星分布，這與過去的植群調查報告基本一致 (歐 2009 ; 林 2009 ; 林 2013)，進一步驗證了模型的可靠性。

結論

本研究強調了背景資料取樣策略在 SDM 建模中的重要性，特別是對於像霧社櫻這樣的受脅物種，利用目標群體及目標物種密度的取樣方法，都能夠有效涵蓋更廣泛的環境條件，校正取樣偏差，並捕捉到與該物種生態需求相似的環境特徵，從而建構高效的預測模型。此外，最佳模型驗證了氣候條件在決定物種分布中的核心影響力，並揭示霧社櫻對寒冷環境的高度適應性，這為其保育和管理提供了重要的科學依據。

引用文獻

林世宗。2009。台灣原生櫻花樹種品

- 種蒐集與採種園建立(三)。行政院農業委員會林務局羅東林區管理處。
- 林世宗。2013。棲蘭山檜木林珍稀物種繁殖保育計畫 - 以原生櫻花及台灣檫樹為例。行政院國軍退除役官兵輔導委員會榮民森林保育事業管理處。
- 歐辰雄。2009。觀霧地區植群生態調查及植栽應用之研究。內政部國家公園署雪霸國家公園管理處。
- Ahmadi, M., M. R. Hemami, M. Kaboli and F. Shabani. 2023. MaxEnt brings comparable results when the input data are being completed; Model parameterization of four species distribution models. *Ecology and Evolution* 13: e9827.
- Araújo, M. B., R. P. Anderson, A. Márcia Barbosa, C. M. Beale, C. F. Dormann, R. Early, R. A. Garcia, A. Guisan, L. Maiorano, B. Naimi, R. B. O'Hara, N. E. Zimmermann and C. Rahbek. 2019. Standards for distribution models in biodiversity assessments. *Science advances* 5: eaat4858.
- Barber, R. A., S. G. Ball, R. K. Morris and F. Gilbert. 2022. Target-group backgrounds prove effective at correcting sampling bias in Maxent models. *Diversity and Distributions* 28: 128–141.
- Barbet-Massin, M., F. Jiguet, C. H. Albert and W. Thuiller. 2012. Selecting pseudo-absences for species distribution models: how, where and how many? *Methods in Ecology and Evolution* 3: 327–338.
- Boyce, M. S., P. R. Vernier, S. E. Nielsen and F. K. Schmiegelow. 2002. Evaluating resource selection functions. *Ecological Modelling* 157: 281–300.
- Di Cola, V., O. Broennimann, B. Petitpierre, F. T. Breiner, M. D'Amen, C. Randin, R. Engler, J. Pottier, D. Pio, A. Dubuis, L. Pellissier, R. G. Mateo, W. Hordijk, N. Salamin and A. Guisan. 2017. ecospat:

- an R package to support spatial analyses and modeling of species niches and distributions. *Ecography* 40: 774–787.
- Dormann, C. F., J. Elith, S. Bacher, C. Buchmann, G. Carl, G. Carré, J. R. García Marquéz, B. Gruber, B. Lafourcade, P. J. Leitão, T. Münkemüller, C. McClean, P. E. Osborne, B. Reineking, B. Schröder, A. K. Skidmore, D. Zurell and S. Lautenbach. 2013. Collinearity: a review of methods to deal with it and a simulation study evaluating their performance. *Ecography* 36: 27–46.
- Elith, J., M. Kearney and S. Phillips. 2010. The art of modelling range-shifting species. *Methods in ecology and evolution* 1: 330–342.
- Franklin, J. 2023. Species distribution modelling supports the study of past, present and future biogeographies. *Journal of Biogeography* 50: 1533–1545.
- Gao, Z., W. Zhuang, L. Wang, J. Shao, X. Luo, B. Cai and Z. Zhang. 2012. Evaluation of chilling and heat requirements in Japanese apricot with three models. *HortScience* 47: 1826–1831.
- Hijmans, R. J., S. E. Cameron, J. L. Parra, P. G. Jones and A. Jarvis. 2005. Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology* 25: 1965–1978.
- Huang, J. Y. 2024. Geographic range shifts of Taiwan's endemic plant species *Prunus transarisanensis* under climate change. *Taiwan Journal of Biodiversity* 26: 1–28.
- Kadmon, R., O. Farber and A. Danin. 2004. Effect of roadside bias on the accuracy of predictive maps produced by bioclimatic models. *Ecological Applications* 14: 401–413.
- Kramer-Schadt, S., J. Niedballa, J. D. Pilgrim, B. Schröder, J. Linden-

- born, V. Reinfelder, M. Stillfried, I. Heckmann, A. K. Scharf, D. M. Augeri, S. M. Cheyne, A. J. Hearn, J. Ross, D. W. Macdonald, J. Mathai, J. Eaton, A. J. Marshall, G. Semiadi, R. Rustam, H. Bernard, R. Alfred, H. Samejima, J. W. Duckworth, C. Breitenmoser-Wuersten, J. L. Belant, H. Hofer and A. Wilting. 2013. The importance of correcting for sampling bias in MaxEnt species distribution models. *Diversity and Distributions* 19: 1366–1379.
- Leroy, B., C. N. Meynard, C. Bellard and F. Courchamp. 2016. virtual-species, an R package to generate virtual species distributions. *Ecography* 39: 599–607.
- Lu, M. L. and J. Y. Huang. 2023. Predicting negative effects of climate change on Taiwan's endemic bumblebee *Bombus formosellus*. *Journal of Insect Conservation* 27: 193–203.
- Morales, N. S., I. C. Fernández and V. Baca-González 2017. MaxEnt's parameter configuration and small samples: are we paying attention to recommendations? A systematic review. *PeerJ* 5: e3093.
- Phillips, S. J. and M. Dudík. 2008. Modeling of species distributions with Maxent: new extensions and a comprehensive evaluation. *Ecography* 31: 161–175.
- Phillips, S. J., M. Dudík, J. Elith, C. H. Graham, A. Lehmann, J. Leathwick and S. Ferrier. 2009. Sample selection bias and presence-only distribution models: implications for background and pseudo-absence data. *Ecological applications* 19: 181–197.
- Phillips, S. J., R. P. Anderson and R. E. Schapire. 2006. Maximum entropy modeling of species geographic distributions. *Ecological modeling* 190: 231–259.
- Schartel, T. E. and Y. Cao. 2024.

- Background selection complexity influences Maxent predictive performance in freshwater systems. *Ecological Modelling* 488: 110592.
- Shannon, C. E. 1948. A mathematical theory of communication. *The Bell system technical journal* 27: 379–423.
- Sorbe, F., T. Gränzig and M. Förster. 2023. Evaluating sampling bias correction methods for invasive species distribution modeling in Maxent. *Ecological Informatics* 76: 102124.
- VanDerWal, J., L. P. Shoo, C. Graham and S. E. Williams. 2009. Selecting pseudo-absence data for presence-only distribution modeling: how far should you stray from what you know? *Ecological Modelling* 220: 589–594.
- Velazco, S. J. E., B. R. Ribeiro, L. M. O. Laureto and P. D. M. Júnior. 2020. Overprediction of species distribution models in conservation planning: A still neglected issue with strong effects. *Biological Conservation* 252: 108822.
- Vollering, J., R. Halvorsen, I. Auestad and K. Rydgren. 2019. Bunching up the background betters bias in species distribution models. *Ecography* 42: 1717–1727.
- Zhang, T., C. Cheng and X. Wu. 2023. Mapping the spatial heterogeneity of global land use and land cover from 2020 to 2100 at a 1 km resolution. *Scientific Data* 10: 748.
- Zhang, Y., K. Ma and Q. Li. 2023. Research on chilling requirements and physiological mechanisms of *Prunus mume*. *Horticulturae* 9: 603.

附錄 1 模擬物種分布之預測變項

Appendix 1 Predictor variables used for the model species distributions

Variable
Annual mean temperature
Mean diurnal range (Mean of monthly (max temp - min temp))
Isothermality (BIO2/BIO7) ($\times 100$)
Temperature seasonality (Standard deviation $\times 100$)
Max temperature of warmest month
Min temperature of coldest month
Temperature annual range (BIO5-BIO6)
Mean temperature of wettest quarter
Mean temperature of driest quarter
Mean temperature of warmest quarter
Mean temperature of coldest quarter
Annual precipitation
Precipitation of wettest month
Precipitation of driest month
Precipitation seasonality (Coefficient of variation)
Precipitation of wettest quarter
Precipitation of driest quarter
Precipitation of warmest quarter
Precipitation of coldest quarter
Altitude
Slope
Aspect
Land cover
