

## 研究報告

空間尺度及資料解析度：空間外推物種生態樣式之  
兩難以香桂與蘇鐵蕨為例羅南璋<sup>1</sup> 王文巧<sup>2</sup> 張偉顥<sup>3</sup> 黃凱易<sup>4</sup>

【摘要】分析物種與環境的關係一直是生態學的中心議題。生態與環境問題，如生物多樣性損失及全球暖化等，是在大區域或於長期時空尺度下運作，但生態研究實測資料通常是於短暫期間且在小區域採集，兩者尺度不相稱。「空間外推」一直是生態學的部份，但 20 世紀下半變成必要技術且為應用生態學研究焦點。直接作用因子外推大空間尺度生態關係，雖只有粗解析度資料，且預測欠精確，但具普遍性。相對的，間接作用因子外推小空間尺度生態關係，雖有細解析度資料，且預測精確，卻難適用在更廣大地區，此即生態學空間外推的兩難困境。過去數十年，遙測及 GIS 技術創新，可在更大空間尺度及更細緻層次描繪空間樣式。本研究因取樣設計需要，選香桂與蘇鐵蕨為研究對象。藉由 GIS 疊合兩物種與地形變數及 SPOT-5 影像植生指標圖層，協同抉擇樹 (DT) 及區別分析 (DA)，預測兩者於惠蓀林場試區空間樣式。建、驗模採三種取樣設計，是分別取自東峰溪與關刀溪流域不同樣本組合而成。準確度評估顯示，DT 遠優於 DA，而兩者執行效率相當。重要的是 DT 於首次模擬，大幅縮小實地調查面積，節省可觀經費及人力，故更適用於兩物種適生育地模擬。植生指標改善模式預測能力效用很小，乃因其光譜及空間解析度皆不足，無法分辨散生香桂與蘇鐵蕨。兩種統計法建立「東峰模式」皆未通過關刀驗模組檢測，凸顯僅含地形變數模式無法跨越空間外推無建模樣本區域。未來研究將從高空間、高光譜解析度遙測資料萃取物種光譜資訊作建模用變數，期使模式能普遍應用於大空間尺度。

【關鍵詞】：香桂、蘇鐵蕨、空間外推、地理資訊系統、遙感探測、數值高程模型、抉擇樹、區別分析。

---

1. 國立中興大學實驗林管處育林組組長

Specialist, Experimental Forest Management Office, Chung-Hsing University.

2. 國立中興大學森林學系 碩士班研究生

The Graduate Student for Master Degree

3. 農委會林務局森林企劃組組長

Chief, Forest Planning Section, Forest Bureau, Council of Agriculture.

4. 國立中興大學森林學系 教授，通訊作者

Professor, Department of Forestry, National Chung-Hsing University, Corresponding Author, e-mail: kyhuang@dragon.nchu.edu.tw

## Research paper

# Spatial Scale and Data Resolution: The Dilemma of Spatially Extrapolating Ecological Pattern of Species - A Case Study of *Cinnamomum subavenium* and *Brainea insignis*

Nan-Chang Lo<sup>1</sup> Wen-Chiao Wang<sup>2</sup> Wei-I Chang<sup>3</sup> Kai-Yi Huang<sup>4</sup>

**【Abstract】** The analysis of species-environment relationship has always been a central issue in ecology. Ecological and environmental problems, such as global warming and biodiversity loss, operate over very large areas or over extended periods of time, but the field data that characterize ecological research are typically collected over relatively small areas during studies of short duration. The scales with data collection mismatch. “Spatial extrapolation” has always been a part of ecology, but it became a *sine qua non* and a major research focus in applied ecology in the latter half of the 20<sup>th</sup> century. The model using direct parameters for extending ecological patterns is more general and applicable over larger areas although it has only coarse resolution and low accuracy data available and thus is inaccurate. On the contrary, the model using indirect parameters can be applied within a limited geographical extent although it has fine resolution and high accuracy data available and thus is accurate. This is the dilemma of spatial extrapolation in ecology. Technological innovations over the last few decades, especially in the fields of remote and GIS, have greatly enhanced scientists' ability to describe patterns over broader spatial scales and at a greater level of detail. Randaishan cinnamons (*Cinnamomum subavenium* Miq.) and cycad-ferns (*Brainea insignis*) were chosen as target for this study because their locations and distributions meet the requirements of sampling designs. GIS technique was applied to overlay the sample layers of the two species on the layers of topographic variables and vegetation indices derived from SPOT-5 images for modeling the species' suitable habitat. Decision tree (DT) and discriminant analysis (DA) models were developed to predict and map the species' suitable sites in the study area, and to determine the optimum one in terms of accuracy and efficiency. Three sampling designs derived from different combinations of samples taken from Tong-Feng and Guan-Dau watersheds were used for model development and validation. Accuracy assessment showed that the accuracy of DT was much better than that of DA; and the two models were highly efficient in implementation of model development and validation. More importantly, DT can be applied to predict the species' suitable habitat because they greatly reduced the area of field survey at the first stage. Vegetation indices could not improve the predicting ability of models for the widely distributed species because of SPOT imagery lacking fine spectral resolution and spatial resolution. “Tong-Feng models” developed from two methods failed to pass validation by Guan-Dau test samples despite passing validation by Tong-Feng test samples. The outcome emphasized that the models only based on topographic variables could not perform spatial extrapolation accurately from a smaller area with training data to a larger area without any training data. Follow-up studies will attempt to extract spectral information associated with species from high spatial, spectral resolution remotely sensed data and use it as variable for model development so that models are more general and applicable over larger areas.

**【Key words】** Randaishan cinnamon, Cycad-fern, Spatial Extrapolation, Geographic Information System (GIS), Remote Sensing, Digital Elevation Model (DEM), Decision Tree (DT), Discriminant Analysis (DA).

## 一、前言

分析物種與環境的關係一直都是生態學的中心議題。Miller *et al.* (2004) 指出科學家(狹義為生態學家)愈加需要因應外界要求,從大空間尺度來解釋生態與環境問題,而與此焦點相關的現象,如全球暖化、環境污染、生物多樣性損失及土地利用變遷,是在廣大區域且(或)於長時期的時、空尺度下運作,但是生態研究現場實測資料通常是於短暫研究期間內在相對小區域所採集,前後二者的尺度兩不相稱。科學家面臨最艱鉅的挑戰之一就是解決此兩不相稱尺度的缺失。若考慮到大空間尺度資料採集在既定後勤支援、財務與技術的限制下,面對此挑戰往往取決於科學家能否以手邊現有資料做可靠的預測。

當預測是根據現有的知識,那麼它就可更精準地稱為「空間外推」(spatial extrapolation)。「空間外推」是透過基於假設未知與已知之間的連續、一致或其他相似性之推測,從而投射、延伸或擴展已知資料或經驗至某一未知或未經歷的地區,以便獲取該未知地區的知識(Gove and Merriam-Webster, 1986)。此一定義涵蓋「按一定比例增加或放大」(scaling up)或根據較小空間尺度採集資料導出可應用在大空間尺度的推論與規則之過程。該定義也包含從一處延伸某一生態關係(或樣式)到具幾近相同空間尺度的另一處(Turner *et al.*, 1989)。後者,即第二個定義的空間外推可在原範圍(即觀測範圍)之外,或在其範圍之內,一如空間內插(spatial interpolation)(Miller *et al.*, 2004)。

Miller *et al.* (2004) 明確點出,不論是上述那一類型的「空間外推」,它一直都是生態學的一部份,但在 20 世紀下半「空間外推」就變成必要的技術(*sine qua non*),這反映出生態學領域已逐漸轉變並深化為預測科學。「空間外推」已成為應用生態學主要研究焦點,且目前雖有多樣化方法應用於多種情境背景,但已浮現出數種常見型式。在最可靠的外推之中,

應變數與最密切關聯的環境特徵,可運用遙測(Remote Sensing)予以精確測繪。在過去數十年,技術創新,特別是遙測及地理資訊系統(Geographic Information System, GIS),讓科學家具備更強的技能,可在更大空間尺度及較過去更詳細層次來描繪空間樣式,大幅提昇他們面對挑戰的能力。Lillesand *et al.* (2008) 強調多階遙測觀念(multistage remote sensing concept),融合分析地面實測、低高空航空資料及太空衛星資料,實施空間外推。換言之,分析者可藉由不同層次的遙測資料,從小空間尺度下有地面實測樣點的較小地區,外推並擴展至大空間尺度下卻無地面實測樣點的廣大地區。生態學空間外推係將過往研究所辨認某一位置的生態關係(或樣式)朝外延伸到另一新的位置或擴展到更大的空間尺度下的涵蓋區域。此「空間外推」包含了三個主要步驟:定義目標與概念模式、建立模式與預測、驗證模式與評估。然而空間外推的限制之一是對許多生態關係的機制與回饋欠缺充分瞭解,故其相互關係僅適合在狹窄的空間和時間尺度之範圍內,不利於在大空間尺度、長期狀態下實施空間外推(Miller *et al.*, 2004)。

概念模式的定義是根據前人文獻與實地調查的描述資料或應變數(擬預測的樣式或作用)與預測變數的統計關係,稱為經驗式(empirical)、統計式或現象學式(phenomenological)(Decoursey, 1992; Korzukhin *et al.*, 1996; Sharpe and Rykiel, 1991; Pickett *et al.*, 1994; Leary, 1985)之模式。這類模式的數學式的主要目標是濃縮經驗事實,並不期望能夠描述模式參數(預測變數)與應變數間實際的「因果關係」,也不能明確示知基本生態功能與機制(Wissel, 1992)。另一類模式設計切合實際且具普遍性,它們稱為機械論式(mechanistic)(如Prentice, 1986)、生理學式(physiological)(如Leersnijder, 1992)、因果關係式(如Decoursey, 1992)或作用過程模式(process models)(Korzukhin *et al.*, 1996),而

這些模式的預測都是依據實際的因果關係。所以，它們被視為具普遍性的，乃因它們所依據的因果關係是生物作用或機能的 (Woodward, 1987)。從理論或機械論的觀點來看，這些生態參數被認為是生物實體分布與豐富度的因果關係驅動力，根據它們來預測生物實體之分布方能切合實際 (Guisan and Zimmermann, 2000)。

植群預測模式在本質上通常是經驗式的；但是，它們可使用具生理意義之參數來建立，且能夠比僅以地形參數所建經驗模式形塑為更具因果關係之機械論式 (Lenihan, 1993；Prentice *et al.*, 1992；Moore *et al.*, 1991；Burke *et al.*, 1989)。一般植群預測模式使用解釋 (或自) 變數可分為植群生育地直接及間接作用因子，而其中屬於直接作用因子，如日照、降水、氣溫及土壤等，對植物生長產生直接影響 (王子定, 1974；Guisan and Zimmermann, 2000)。直接作用因子的資料通常由數目有限的測站或樣點收集，如氣象站等，是屬於點資料，不利顯現與生物相關之微氣候，且都需經空間推估產生面資料，以供 GIS 運用分析，難免引入內插誤差產生的空間不確定性，故其資料為粗解析度且低精確度 (Guisan and Zimmermann, 2000)。

另一方面，間接作用因子如海拔、坡度、坡向及坡面位置等地文因子 (地形變數) 需藉由對直接因子發生作用而後影響植物的生長 (王子定, 1974)。間接作用因子的資料可藉由遙測技術獲取，產生網格型資料層，如數值高程模型 (Digital Elevation Model, DEM)，通常不需經空間推估即可為 GIS 所運用，故其資料解析度及精確度相對較上述者高出甚多 (Guisan and Zimmermann, 2000)。因此，以部份或全部高解析度的地文因子發展出山地植群預測模式，就毋需感到驚訝 (Fischer, 1990；Moore *et al.*, 1991；Brzeziecki *et al.*, 1993；Brown, 1994；Guisan *et al.*, 1998；Guisan *et al.*, 1999)。在小空間尺度下，涵蓋數個集水區的均質氣候區

內，緯度相差有限，局部氣候常是植群分布主要影響因子，而局部氣候的差異則是由地文因子所引起，故一般植群分布均以地文因子進行模擬 (蘇鴻傑, 1987；Guisan *et al.*, 1999)。相反的，大空間尺度預測模式僅僅以生物物理變數 (即直接作用因子) 來建立，乃因地形變數在此粗解析度下就不再具有預測能力 (Box, 1981；Prentice *et al.*, 1992；Lenihan, 1993；Huntley *et al.*, 1995；Neilson, 1995)。

上述使用間接參數的缺點是模式僅能應用在有限的地理範圍內而不產生顯著誤差，但無法精確地普遍應用在大空間尺度下的更廣大地區，乃因在不同地理區內相同地形位置可以呈現不同直接作用因子的組合 (Guisan and Zimmermann, 2000)。Walter and Walter (1953) 稱此為「相對位置恆常定律」 (the law of relative site constancy)。此定律描述事實：「物種有傾向藉由選擇類似的微位置來改變它們所在地形位置，以補償在氣候狀態下的區域差別」。這必然也意味著使用直接作用因子作為預測變數 (參數)，可以確保模式普遍應用在大空間尺度下較廣大地區。尤有甚者，直接作用因子有助於將空間模式納入植群動態演替預先鋪路，如 Solomon and Leemans (1990) 或 Halpin (1994) 所提議。確切來說，選擇直接作用因子外推大空間尺度生態關係雖具普遍性，但只有粗解析度與低精確度資料可供使用，而預測結果欠缺精確性；相對的，選擇間接作用因子外推小空間尺度生態關係，有細解析度與高精確度資料可供使用，預測結果雖具精確性，卻難以適用在大空間尺度下的廣大地區。這就是生態學空間外推所遭遇左右兩難的困境 (dilemma)，亦即模式應變數與預測變數之選擇受到模擬目標系統之內有無資料的限制，此為空間外推限制之二 (Miller *et al.*, 2004)。

為了印證上述僅含地形 (地文) 變數的預測模式無法普遍應用在大空間尺度之外推上，本研究將藉由多變量統計，建立僅含地形變數的預測模式，並透過三種不同取樣設計，進行

空間外推，預測涵蓋惠蓀林場研究區內香桂與蘇鐵蕨的適生育地 - 空間樣式。本研究採用的分類迴歸樹（抉擇樹）及區別分析，除了參考 Guisan and Zimmermann (2000) 及 Wang and Ni (2006) 之外，主要是依據 Miller *et al.* (2004) 評論參數式線性迴歸、空間統計克利金法及非參數式分類迴歸樹（抉擇樹）等三類多變量統計方法之利弊所挑選出。預測模式採用的自變數是由 DEM 導出海拔、坡度、坡向、坡面位置，加上衛星影像導出植生指標，藉由多變量統計軟體篩選出最有效用之地形變數來建立模式。本研究採用這兩種植物是著眼於不同空間尺度搭配的取樣設計，蘇鐵蕨三個取樣區皆在關刀溪流域內，三者相距小於 1 km，而香桂兩個取樣區為分別在惠蓀林場東西側的關刀溪與東峰溪流域，二者緯度相同，屬同一氣候區，且相距約 5 km 上下。本研究取樣設計是考量測點的空間自相關 (spatial autocorrelation)，亦即測點之間的統計關係，其關係是測點間距離與方向的函數。對於實際世界內觀察連續變化的現象（例如海拔、溫度等），在空間內事物的位置越相近，其性質較那些相隔越遠者越相似 (Jensen, 2005)。

Miller *et al.* (2004) 於其結論指出，空間外推是評斷生態關係堅實與否之手段，故其建議「以多個模式應用在同一資料檔之比較」或「應用單一模式在不同空間尺度描繪生態樣式的資料檔」。本研究同時採用這兩種評定方式，運用 GIS 平台，將關刀溪流域及東峰溪流域 GPS 定位香桂及蘇鐵蕨樣株圖層與海拔、坡度、坡向、坡面位置及衛星影像導出植生指標等圖層加以疊合並抽取樣本，再代入抉擇樹及區別分析，建立預測模式，從而模擬及測繪整個試區蘇鐵蕨及香桂的適生育地。因此，除了從推測準確度與執行效率來評斷以這兩種統計方法建立預測模式之優劣所在，更重要的是本研究在兩種空間尺度下，將「關刀溪流域松風山、杜鵑嶺及關刀山林道蘇鐵蕨樣株（三取樣區相距小於 1 km）」及「東峰溪與關刀溪兩流域

香桂樣株（兩取樣區相距約 5 km 上下）」（見圖 1 所示）分別加以分割與組合，形成三種取樣設計，從而評估這三者對模式空間外推能力之影響。具體而言，在蘇鐵蕨的反差對比下，本研究在印證僅含地形變數且不納入關刀溪流域香桂建模樣本的「東峰模式」，無法將東峰溪流域之香桂生態樣式正確地外推到相隔 5 km 的關刀溪流域。

## 二、研究區域

惠蓀林場行政區域隸屬南投縣仁愛鄉，為中興大學四個實驗林場之一，位置約在南投縣東北角，如圖 1 紅點所示，座標範圍東經 121°0' -121°8'，北緯 24°2' -24°7'，略呈東北至西南走向，海拔由 454 m 至 2,419 m。境內有五個主要集水區，從西到東分別為蘭島溪 (629 ha)、關刀溪 (2,216 ha)、東峰溪 (2,061 ha)、椿谷溪 (976 ha) 及合水溪 (715 ha)，流向均由南向北，匯流入北港溪，加上數個小集水區，總面積約 7,477 ha。全區氣候溫潤，植物種類繁多，約有 1,100 種，為台灣中部地區極具代表性之森林。本研究為配合使用衛星影像，故採矩形區域，涵蓋整個不規則形狀之惠蓀林場，總面積約為 17,136.0 ha，分析模擬係根據該矩形區域實施。蘇鐵蕨的地面定位調查區域是在關刀溪流域左上角的松風山、杜鵑嶺及關刀山林道，而香桂的地面定位調查區域是在東峰溪及關刀溪流域，圖 1 除了顯示這兩個集水區之相關位置，也顯示在兩個不同空間尺度下的蘇鐵蕨取樣區及香桂取樣區。

香桂 (*Cinnamomum subavenium* Miq., Randaishan cinnamon) 又名巒大桂，為樟科常綠大喬木，因其樹皮及葉均具肉桂香味，因而得名 (劉業經等, 1994)。分布華南地區及台灣，台灣多分布於中、南部海拔 500-2,000 m 之森林中，數量甚多，耐陰性強，在大型孔隙或崩塌地區，少發現其小苗；但在林下其小苗數量眾多，且在擾動較少及土壤層較厚的森林中，常成為大喬木 (樹高可超過 20 m)，佔據森

林的上層。香桂在惠蓀林場為優勢樹種之一，於海拔 650 m (松風山區) 即可發現它的蹤跡，但主要出現在海拔 900-2,000 m 環境較為潮濕的森林中。根據羅南璋 (1992) 於東峰溪流域設置 49 個樣區進行植群分析調查分布於東峰溪流域、北東眼山區的香桂，多為胸高直徑 60 cm 以上，樹高 20 m 以上之大樹，然分布於關刀流域者胸高直徑多不及 30 cm，且樹高低於 20 m 以下，在形態上因兩個流域的生長環境不同，而有明顯的差異存在。

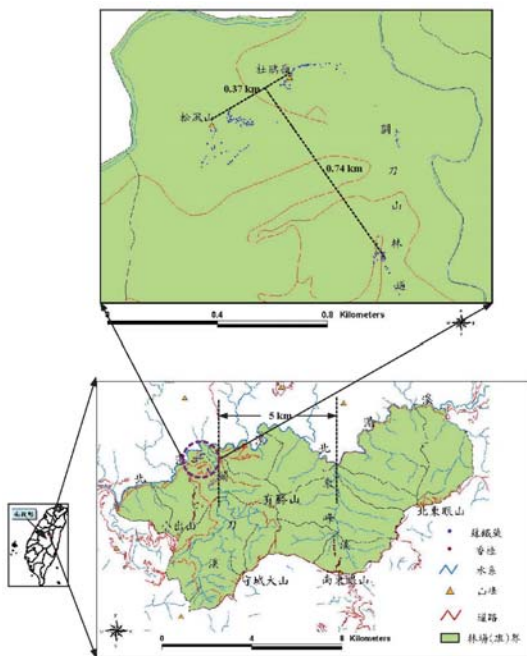


圖1. 惠蓀林場研究區域位置圖

Fig. 1. Location map of the Huisun study area

蘇鐵蕨 (*Brainea insignis* (Hook.) J. Sm) 屬烏毛蕨科 (Blechnaceae) 的蕨類植物。在農委會林務局自然資源與生態資料庫網站上，蘇鐵蕨保育等級列為瀕臨滅絕 (Endangered, EN)，其棲地之範圍、面積及品質持續下降、減少，且屬於小而且狹隘分布之族群，生育地點不超過五處。蘇鐵蕨顧名思義，形狀似蘇鐵，頗具觀賞價值，但其移植困難，不易成活。蘇鐵蕨屬全世界僅有蘇鐵蕨一種。分布泰國、馬來亞、印尼、菲律賓、中國南部及台灣。在台灣分布

於中部一帶山區，特別是惠蓀林場及東卯山一帶，數量較多，以惠蓀林場為主要之生育地。由於其生態幅度較窄，分布範圍受限，成為亟待保護之稀有植物之一 (呂金誠等, 1986)，蘇鐵蕨分布於東卯山海拔高約 1,250 m 之台灣二葉松林下，土壤為多石塊之砂質壤土；在惠蓀林場松風山之生育垂直分布大約在 700 m 至 825 m，能耐旱，生長於向陽坡面，土壤排水良好的地區，多與台灣二葉松、楓香、栓皮櫟混生，在人工造林地之林緣也有所見 (呂金誠等, 1986)。

### 三、材料與方法

本研究共分為六個大項目的主要步驟，茲將其分項條列並搭配圖 2 之研究流程詳述如後。

#### (一) 資料蒐集

資料蒐集的項目，包括數值高程模型 (Digital Elevation Model, DEM) 資料檔，農航所製作像片基本圖 (1/10,000)、赴現場以 GPS 搭配 5m 伸縮型加長天線桿及雷射測距系統測繪蘇鐵蕨及香桂樣株資料。

另外，向中央大學太空及遙測研究中心選購九幅 SPOT-5 衛星影像，採第二級正射糾正處理，並將座標系統設定為 TWD67 基準及橫麥卡托投影經差二度分帶 (TM 2°)，最後以最近鄰法重取樣為 5 m × 5 m。本研究由九幅影像之中再篩選出品質最佳兩幅，作為求導夏秋兩季季差植生指標之用。這兩幅影像之雲覆蓋量皆要求在 15 % 以下，入射角或視角皆在 ± 15° 以內，其影像的基本資料如表 1 所示。

#### (二) 資料處理

##### 1. 稜谷線數化

使用 Calcomp Drawing Board III (A0 尺寸) 的高精度數化板將研究區內所涵蓋 1/10,000 像片基本圖上之稜谷線加以數化及建檔，而其數化結果將用於後續求導與建立坡面位置圖層。

##### 2. 地文因子圖層建立

應用 ERDAS Imagine 影像處理及分析軟

體模組從 DEM 資料產生海拔圖層，並求導與建立坡度、坡向圖層。方位角採 360° 為圓周式循環數列，而非直線數列，不利於比較，故在納入資料分析前，應予以轉換處理 (Chang, 2010)。本研究實施行前預備測試，分別採用 360° 方位角、坡向分為八級 (每級 45°)、坡向度數轉換為餘弦等三種計量方式，所得結果實質上均無差異，故決定坡向採用 360° 方位角的計量方式，以維持坡向與海拔、坡度皆為比

率或間距式數值變數之本質 (比率或間距變數的資料含量大於級序與類別變數)。

### 3. 坡面位置圖層建立

一般而言，生育地在山坡上之相對位置，對於土壤性質及林木生長深具關係 (劉崇瑞等, 1983)。求導及建立坡面位置圖層，首先需將數化的稜脊線與山谷線分別轉換為網格檔，並依據 Skidmore (1990) 所提相對位置計算式子，計算每一網元至最鄰近稜線與谷線的歐

表1. 惠蓀林場研究區 SPOT-5 衛星影像基本資料

Table 1. The basic information of SPOT-5 satellite images over the study area

影像編號	季節	接收日期	視角 (°)	平均雲量 (%)	方位角 (°)	高度角 (°)
I0009406	夏	2004/07/10	-7.8	11	90.7	70.8
I0009409	秋	2005/11/11	-13.2	3	155.9	45.4

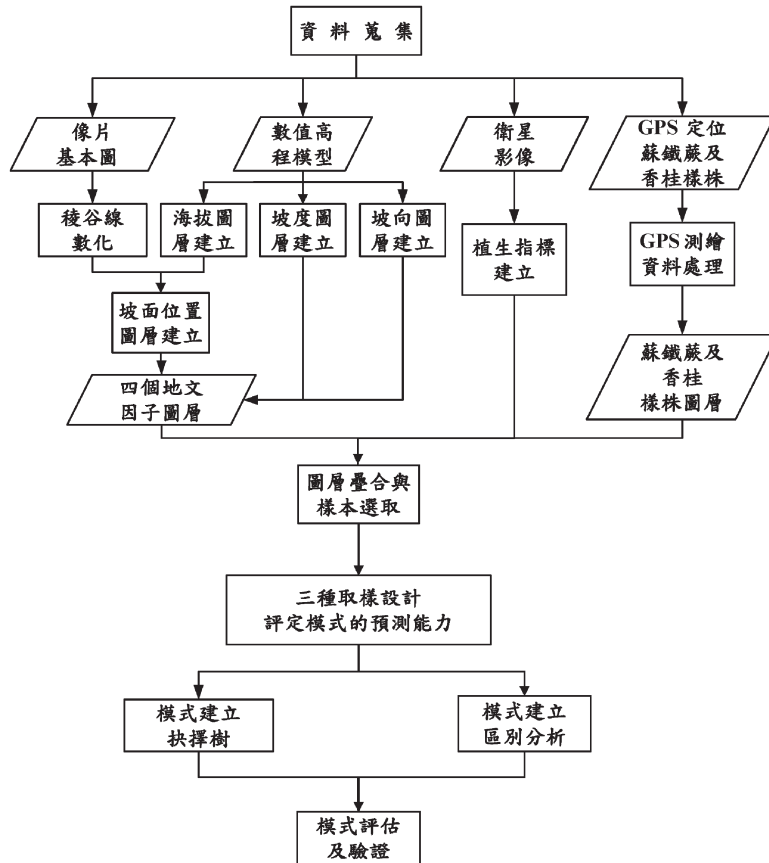


圖2. 本研究的作業流程圖

Fig. 2. The operational flowchart of this study.

幾里得距離 (Euclidean distance)，並求出該網格點之相對位置比例  $P_{ij}$ ，其公式如下所示：

$$P_{ij} = PV / (PV + PR) \dots\dots\dots (1)$$

$PV = P$  點到與最鄰近山谷線的歐幾里得距離；

$PR = P$  點到與最鄰近稜脊線的歐幾里得距離；

上式中的  $P$  為某一檢定點 (網元)， $P_{ij}$  為在  $i$  列， $j$  行之相對位置比例。本研究分別以稜脊線及山谷為最高及最低坡面位置，到山谷線之  $P_{ij} = 0.0$ ，而到稜脊線之  $P_{ij} = 1.0$ ，再將其間分為六級，連同稜脊線、山谷線共分為八級 (謝立忻等，2005)。

#### 4. 植生指標圖層建立

Hoffer (1978) 指出植物在夏、秋兩季，近、中紅外光反射因葉子內部結構及含水量而有明顯之差異，因此可作為分辨不同種類植物的重要依據。植生指標原始構想除了源自於植物、土壤於近、中紅外光譜反射原理及特性，再加上植物物候之特性，在夏、秋季節葉片細胞壁結構及含水量有甚大差異，使這兩季的近、中紅外波段差值比可以作為樹種之間的區別的依據。本研究由夏 (2004/07/10) 和秋 (2005/11/11) 兩期 SPOT-5 衛星影像求導出植生指標，其求導的公式如下所示：

$$[(NIR - MIR)_{秋} / (NIR - MIR)_{夏}] \dots\dots\dots (2)$$

#### 5. GPS測繪資料處理

本研究赴試區現場獲取香桂及蘇鐵蕨樣株資料，係使用 Trimble Pro XR 移動站搭配 5 m 可伸縮加長天線予以定位，另視衛星訊號實際接收狀況，輔以雷射測距系統協助定位。於調查結束返回中興大學後，再以設置於森林系館建築物頂層 Trimble 基站資料融合移動站資料，做後續差分改正，經改正後樣株位置之二維平均定位誤差在 1 m 以內。最後，將其轉換為 ArcGIS 所能讀取及處理的 shapefile 圖層資料，且座標系統設定為 TWD67 及 TM 2° 分帶，「東峰溪及關刀溪流域香桂樣株」及「關刀溪流域松風山、杜鵑嶺及關刀山林道蘇鐵蕨樣株」的分布分別如圖 1 內暗紅色 (香桂) 及藍色 (蘇鐵蕨) 斑點所示。

#### (三) 圖層疊合及樣本選取

本研究將香桂樣株及上述五項因子圖層之座標系統皆設定為 TWD67 與 TM 2° 分帶，並以 ERDAS Imagine 軟體中 Image Interpreter\ Utilities\ Layer Stack 疊合形成 GIS 資料庫，供後續取樣分析與模擬之用。接著再以 ERDAS Imagine 軟體中 Image Interpreter\ Utilities\ Subset 分別切取目標樹種區塊 (即主體) 與該樹種之互補區塊 (即背景)。依據 Sperduto and Congalton (1996) 之研究，將背景 (非目標物種生育地) 樣本對主體 (目標物種生育地) 樣本以至少三倍的比例來選取背景樣本。同時依據 Pereira and Itami (1991) 從東峰溪與關刀河流域採亂數隨機分散抽取背景樣本，以降低樣本空間自相關至最小，然後分別計算主體區塊和背景區塊在這五項因子的相關統計值。本研究調查香桂樣株共有 149 株，其中 108 株來自東峰溪流域，而其餘的 41 株來自關刀溪流域；蘇鐵蕨樣本數總共為 182 株，其中松風山及杜鵑嶺樣本為 144 株，而關刀山林道的樣本數為 38 株。本研究採用分割樣本建模與驗模之取樣設計方法 (Chang, 2010)，選取全數樣本中佔某一比例的一部份 (例如 2/3 或 1/2) 作為模式建立用之訓練樣本，而剩餘的另一部份 (例如 1/3 或 1/2) 則作為後續模式驗證用之檢測樣本，詳見下面第 (四) 項的說明。

#### (四) 模式空間外推能力評估

模式空間外推能力需考量多項因素，由於這些因素組合變化繁複，計算量極為龐大，為化繁為簡，本研究分別以最早完成調查「東峰溪流域之香桂樣本」及「關刀溪流域松風山與杜鵑嶺之蘇鐵蕨樣本」作為建模與驗模起始點，「以兩種模式應用在同一資料檔之比較」；同時「應用單一模式在兩不同空間尺度描繪生態樣式的資料檔之比較」

##### 1. 取樣設計

本研究香桂的樣株係來自先完成調查的東峰溪流域及後完成調查的關刀溪流域，兩個樣區分隔約 5 km 上下；蘇鐵蕨樣株係來自關刀



溪流先完成調查的松風山與杜鵑嶺及後完成調查的關刀山林道，三個樣區相鄰且分隔小於 1 km；兩者皆藉由 GPS 地面定位調查所獲得。在模式建立使用的訓練樣本及模式驗證使用的檢核樣本採取不同數量及分布的配置，以評估與瞭解「香桂樣株數量及分布」及「蘇鐵蕨樣株數量及分布」對模式空間外推能力的作用所在。本研究採用三種取樣設計：

#### (1) 取樣設計-A

香桂：「東峰模式」建模用之「主體」訓練樣本與驗模用之「主體」檢核樣本皆來自東峰流域，且都不含關刀流域香桂樣本。建模用之「背景」訓練樣本與驗模用之「背景」檢核樣本皆從從惠蓀林場試區（包含東峰溪與關刀溪兩流域）以亂數隨機抽取出來。全數樣本（含主體與背景）中之三分之二（2/3）作為模式建立之用，而其餘之三分之一（1/3）則作為模式驗證之用。因東峰香桂樣株最先完成調查，將此樣本配置用於基準模式之建模與驗模，以作為其他取樣設計相互比較的基準。

蘇鐵蕨：「松風杜鵑模式」建模用之「主體」訓練樣本與驗模用之「主體」檢核樣本皆來自松風山與杜鵑嶺，且都不含關刀山林道蘇鐵蕨樣本。建模用之「背景」訓練樣本與驗模用之「背景」檢核樣本皆從從惠蓀林場試區（包含東峰溪與關刀溪兩流域）以亂數隨機抽取出來。全數樣本（含主體與背景）中之三分之二（2/3）作為模式建立之用，而其餘之三分之一（1/3）則作為模式驗證之用。因松風山與杜鵑嶺之蘇鐵蕨樣株最先完成調查，將此樣本配置用於基準模式之建模與驗模，以作為其他取樣設計相互比較的基準。

#### (2) 取樣設計-B：

香桂：此一設計仍然使用與取樣設計 -A 相同的「東峰模式」，即建模使用與取樣設

計 -A 相同的一套主體與背景樣本（建模不納入任何關刀流域香桂樣本），「背景」驗模樣本也相同，而與取樣設計 -A 唯一的差別在「主體」驗模樣本。此設計的驗模樣本是將全數關刀流域之「主體」樣本及來自兩流域的「背景」樣本（佔全數「背景」樣本的三分之一）用於驗證「東峰模式」，以評估推測錯誤究竟發生在「關刀驗模樣本」，或者是發生在「東峰驗模樣本」。此一設計也在瞭解模式（預測變數僅使用地文因子）在空間外推之能力，能否跨越空間上兩相分隔的流域（東峰流域及關刀流域相隔約 5 km 上下）。

蘇鐵蕨：此一設計仍然使用與取樣設計 -A 相同的「松風杜鵑模式」，即建模使用與取樣設計 -A 相同的一套主體與背景樣本，（建模不納入任何關刀山林道蘇鐵蕨樣本），「背景」驗模樣本也相同，而與取樣設計 -A 唯一的差別在「主體」驗模樣本。此設計的驗模樣本是將全數關刀山林道之「主體」樣本及來自全試區的「背景」樣本（佔全數「背景」樣本的三分之一）用於驗證「松風杜鵑模式」，以評估推測錯誤究竟發生在「關刀山林道驗模樣本」，或者是發生在「松風杜鵑驗模樣本」。此一設計也在瞭解模式（預測變數僅使用地文因子）在空間外推之能力，能否跨越空間上兩相分隔的樣區（松風山杜鵑嶺樣區及關刀山林道樣區相隔小於 1 km）。

#### (3) 取樣設計-C：

香桂：建模用之「主體」訓練樣本與驗模用之「主體」檢核樣本皆來自東峰流域與關刀流域等兩個樣區，且建模用之「背景」訓練樣本與驗模用之「背景」檢核樣本皆從從惠蓀林場試區（包含東峰溪與關刀溪兩流域）以亂數隨機抽取出來。全

數樣本（含主體與背景）中之三分之二作為模式建立之用，而其餘之三分之一則作為模式驗證之用。確切來說，取樣設計 -C 的「背景」建模及驗模樣本與前兩種取樣設計完全相同，而與它們唯一差別是在「主體」建模與驗模樣本，取樣設計 -C「主體」建模與驗模樣本都混合兩個流域的香桂樣本，前兩種取樣設計建模樣本都不含任何關刀河流域的香桂樣本。「主體建模樣本」是三分之二的東峰溪流域香桂樣本且加上三分之二的關刀溪流域香桂樣本；「主體驗模樣本」三分之一的東峰溪流域香桂樣本且加上三分之一的關刀溪流域香桂樣本。

蘇鐵蕨：建模用之「主體」訓練樣本與驗模用之「主體」檢核樣本皆來自松風山、杜鵑嶺及關刀山林道等三個樣區，且建模用之「背景」訓練樣本與驗模用之「背景」檢核樣本皆從惠蓀林場試區（包含東峰溪與關刀溪兩流域）以亂數隨機抽取出來。全數樣本（含主體與背景）中之三分之二作為模式建立之用，而其餘之三分之一則作為模式驗證之用。確切來說，取樣設計 -C 的「背景」建模及驗模樣本與前兩種取樣設計完全相同，而與它們唯一差別是在「主體」建模與驗模樣本，取樣設計 -C「主體」建模與驗模樣本都混合三個樣區的蘇鐵蕨樣本，前兩種取樣設計的建模樣本都不含任何關刀山林道的蘇鐵蕨樣本。「主體建模樣本」是三分之二的松風山與杜鵑嶺蘇鐵蕨樣本且加上三分之二的關刀山林道蘇鐵蕨樣本；「主體驗模樣本」三分之一的松風山與杜鵑嶺蘇鐵蕨樣本且加上三分之一的關刀山林道蘇鐵蕨樣本。

## 2. 兩種統計法建立模式

本研究使用 SPSS 統計分析軟體進行兩種多變量統計模式之運算，茲分述如下：

### (1) 抉擇樹

分類迴歸樹 (Classification and Regression Trees, CART) 演算法為 (Breiman *et al.*, 1984) 所開發，Verbyla (1987) 首次運用於生態研究上，其為抉擇 (決策) 樹 (Decision Tree, DT) 的演算法之一。此方法因在統計上無常態分布的假設，故屬於無母數統計方法之一 (鄧家駒, 2004)。抉擇樹同時也是一種資料探勘 (data mining) 領域中兼具分類與預測功能的技術，藉由分析給定的已知原始資料，來建立適當的分類規則，再根據此規則對未知類別的資料進行預測，以一個單一解釋 (預測) 變數函數，來建構一個二分式抉擇樹。樹狀圖形頂端的節點稱為根節點 (root node)，最下方的節點則稱為葉節點 (leaf node)，每個葉部節點擁有一個類別標記 (class label)，如圖 3 所示。圖中 X, Y, Z 為分支，即測試之結果，中節點代表測試條件，即輸入之解釋變數 (如海拔、坡度等環境因子)，葉節點分為「0」與「1」，在本研究即分別為香桂 (或蘇鐵蕨) 的「非生育地」與「生育地」。當給定的原始資料從根部節點進入抉擇樹後，可應用各種不同的演算法，來決定每筆資料該進入下一層的哪一個子節點 (child node)，此過程不斷地反覆，直到所有的資料到達葉部節點，而停止分割的情況有六種，詳見林震岩 (2008)。

從根部到每一個葉部節點都有一條獨特的路徑，每一條路徑皆表示一個用來分類資料的規則。當一筆未知的資料進入抉擇樹時，該筆資料將會依其符合的分類規則，而得到預測的結果 (O'Brien *et al.*, 2005)。統計軟體 SPSS 13.0 版之後才具有 CART 分析功能，放置於 Analyze\ Classify\ Tree 功能之下的演算法。CART 的建樹過程，必須藉由分析訓練資料「集中」的資料特性，反覆地找出最適當的分類屬性及其分岔點 (splitting point)，將資料集予以分類。分散度 (diversity) 或稱混雜度 (impurity)，則是用來評估一個分隔屬性變數的衡量標準，對於一組資料的分散度指標以 Gini (1912) 開

發吉尼指標 (Gini index) 計算其公式如下所示：

$$Gini\ index = 1 - \sum c^2 \dots\dots\dots (3)$$

上式之中的  $c$  表示每一組資料的反應比例，而最好的分隔變數是能夠降低一個資料組最多的分散度 (De'ath and Fabricius, 2000)。模式準確度評估 (accuracy assessment) 是將樣本分割為建模組與驗模組資料，而以驗模組樣本進行評估 (Breiman et al., 1984)。因此，在建構 CART 模式之前，必須先選定已分類好的原始資料來源，而資料來源通常包含了許多的紀錄，每筆紀錄擁有相同數目的屬性及一個目標屬性。依照資料的特性和資料探勘的目標，必須將原始資料分割成建模組與驗模組資料，提供 CART 使用。此外，為了避免過度學習 (over-fitting) 而使 CART 樹狀結構出現過於冗長複雜的情形，可利用 CART 驗模組模式修正其結構。其原理係利用建模資料組建立預測模式後，再將驗模資料組代入模式進行驗證，可經由總資料組之中不斷選取部分資料作為建模組與驗模組，並觀察不同樹狀結構大小分類錯誤率之變化，來反覆修正的過程 (Breiman et al., 1984)。

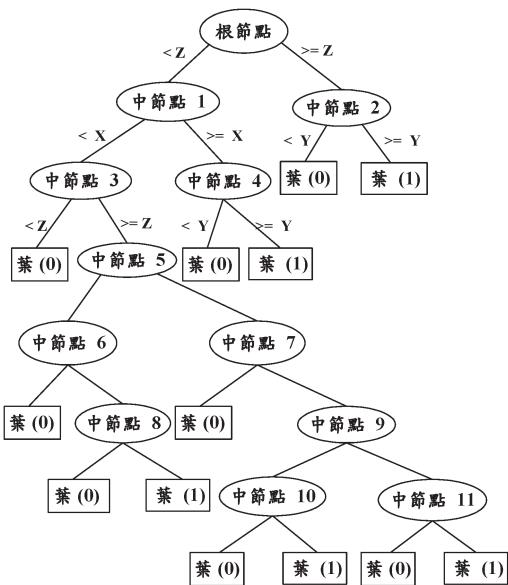


圖3. 抉擇樹之樹狀結構示意圖

Fig. 3. The tree structure of a decision tree

(2) 區別分析

區別分析，或稱判別分析、鑑別分析 (Discriminant Analysis, DA) 為 Fisher (1936) 所創，最早應用於貸款信用評分模型以區分好壞之貸款者；Williams (1983) 首次將此演算法應用在生態上，便對於多個植物族群的分類獲得很好的分類效果。區別分析是一種劃分群體之技術，事先將已知類別的觀察值，選取具有分類效果的樣本，利用類別變數 (grouping variable, g類) 當反應變數，多個計量的區別變數 (discriminant variable) 當解釋變數，建立區別函數 (discriminant function)，公式如下所示：

$$y = b_0 + b_1x_1 + b_2x_2 + \dots + b_px_i + \dots + b_nx_n \dots\dots (4)$$

$y$  為區別函數值 (或稱區別分數 discriminant score),  $x_i$  為區別變數,  $b_i$  為區別係數 (discriminant coefficient or weight), 利用區別函數將新觀察值進行適當分類 (林震岩, 2008)。如圖 4 所示，假設有兩個變數  $X_1$  和  $X_2$ ；兩橢圓代表兩個族群 I 和 II 有部份相互重疊影響。族群 I 可能為香桂 (或蘇鐵蕨) 分布區域、族群 II 可能為其它鄰近香桂 (或蘇鐵蕨) 分布區域之樹種，其生態環境是交疊重複難以區分開的。區別分析的原理便是找出能劃分 I、II 兩族群的交點所連成之直線 DF，與另一條引自原點而與直線 DF 垂直的直線 Y 交於 C 點。直線 Y 即是區別函數圖形，線上任何一點觀測值可由  $X_1$  和  $X_2$  兩個變數轉換而得；直線 Y 上的 C 點，稱為分界點，將 Y 值分為兩部份，作為區分 I、II 兩族群之依據。兩個族群的  $X_1$  和  $X_2$  變異數不同，要使組間變異數對組內變異數比值達最大時方能找到分界點；就可以得到每個區別變數的權重係數，其意義與迴歸分析類似。標準化後的區別係數值愈大，代表此區別變數對群體的區別能力愈高 (黃俊英, 2007)。本研究採用統計軟體 SPSS 之 Analyze\ Classify\ Discriminant 運算資料。 $(b_0 + b_1x_1 + \dots + b_nx_n)$  項乃由統計軟體 SPSS 自動疊代產生之線性判別式。

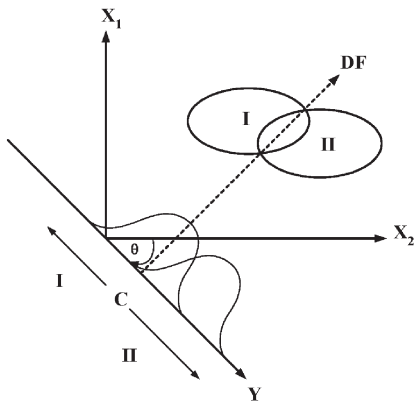


圖4. 區別分析的幾何圖形觀念

Fig. 4. The concept of geometry for discriminant analysis

區別分析在應用上需遵守三個前提假設，詳見林震岩 (2008) 所列。由於赴現場調查相當耗工耗時，在進行統計分析時先暫時略過前提假設檢定，改以事後檢定的方式，採多組獨立樣本回推驗證模式的可靠性。此外區別分析不只停留在描述各類組間變異上的關係，還能夠對未知所屬之類組加以區別分類，故本研究較重視後者預測香桂空間分布的情形，與前提假設違反並無直接相關。但為避免複共線性會有不利分類的影響，採逐步估算法 (step-wise estimation)，逐漸加入新的解釋變數，將不適合的剔除直到模式合理為止。

#### (五) 模式評估及驗證

模式準確度評估是將推測結果與實際情況相互比較，並計算兩相符合之正確百分比率，其中應包含整體準確度、 $kappa$ 相符係數 (coefficient of agreement)、主體 (目標) 推測準確度、背景推測準確度。惟因採全區逐格做像元對像元之比對殊為不可能，故採抽樣方式兩兩比對。由全區抽取樣本除須符合統計取樣原則外，所選總樣本通常分為建模用之訓練樣本與驗模用之檢核樣本兩組，再分別實施模式準確度評估，此方法普遍應用在多光譜遙測影像分類之準確度評估，旨在避免建模樣本自身評估所隱含之偏頗與降低外推 (extrapolation) 之

可能風險 (Jensen, 2005; Lillesand *et al.*, 2008)。

SPSS 統計軟體 DT 與 DA 模組具準確度評估功能，「建模組」訓練樣本於輸入樣本建立模式之時一併完成準確度評估。DT 與 DA 的「驗模組」與「建模組」的準確度評估是於 SPSS 軟體 DT 與 DA 模組執行時一併完成；此二模組於繪製目標植物的適生育地全區預測圖皆需藉由 ERDAS Imagine 軟體 Spatial Modeler 的協助，來完成繪圖工作。

#### 四、結果與討論

本研究之試區為矩形，其左上角之 TM 2° 分帶座標為 (248,845.500, 2,668,192.500)，右下角為 (264,590.500, 2,657,317.500)，涵蓋面積約為 17,136.0 ha (包含惠蓀林場不規則區域面積 7,477 ha)。本研究所用之 DEM 網格資料，是根據前列矩形區域的座標，以 ERDAS Imagine 的 Interpreter\Utilities\Subset 功能切取出來，其範圍大小為 3,150 行 (columns) × 2,176 列 (rows)，總共有 6,854,400 個像元，像元尺寸為 5 m × 5 m。

因「東峰溪流域香桂樣株」及「關刀溪流域松風山與杜鵑嶺蘇鐵蕨樣株」最早完成調查，本研究將前面第三節材料及方法所敘述的「取樣設計 -A」用來建模作為比較的基準模式。海拔、坡度、坡向、坡面位置及植生指標為本研究建立模式的五個預測變數。這五個變數有 30 種組合，剔除五取一的單一變數模式、五取二的雙變數模式之後，仍有五取三、五取四、五取五等之組合，總共還有 15 種組合，計算工作量仍然相當龐大。所以，本研究使用之 SPSS 套裝軟體的 DT 及 DA 模組，可讓分析者從輸入的預測變數自動篩選出最有效用之變數，大幅減少計算工作量，從而節省工作時間。SPSS 套裝軟體兩種多變量統計模組均篩選出坡面位置、坡度、海拔為前三名最有效用的變數，而坡向及 SPOT-5 衛星影像植生指標則殿後。SPOT-5 影像近紅外波段空間解析度為 10 m，而中紅外波段則為 20 m。香桂冠

徑約 5-10 m，蘇鐵蕨冠徑通常約 1-2 m；植冠展開面積都不足一個像元。因此，本研究認為 SPOT-5 影像無法完整解析單株香桂或蘇鐵蕨，且無法與周遭其他植物分隔開。同等重要的是，上列兩個波段皆為寬頻波段，光譜解析度亦不夠細緻，難以區別香桂或蘇鐵蕨與周遭其他植物之光譜反射的差別。因此，本研究以海拔、坡度、坡面位置等三個地形變數來分別建立這兩種模式，且以它們所得之推測準確度作為比較基準。

(一) 兩種模式預測能力之比較

本研究採用海拔、坡度及坡面位置三項變數，來建立基準模式，進一步來評比 DT 及 DA 方法之優劣。表 2 顯示這兩種模式於香桂適生育地推測準確度之比較。先從整體精度來看，DT「建模組」及「驗模組」整體精度為 96% 及 91%，分別高於 DA 的 81% 及 84%；DT 於這兩組的平均值為 94%，高於 DA 的 83%。更重要的是 DT 在建模與驗模之 *Kappa* 平均值 0.68 遠高於 DA 的 0.40。從整體精度及 *Kappa* 值而言，DT 在香桂適生育地的預測能力上遠優於 DA。

其次，再就表 2 內兩種模式的生產者精度與使用者精度來比較兩者之優劣所在。為了能更清楚瞭解表中數字的意義，本研究以建模與驗模平均值來簡化說明。DT 在「非生育地」的

建模與驗模生產者精度平均值 97% 遠高於 DA 的 81% 約 16%，但 DT 在「生育地」的建模與驗模生產者精度平均值 70%，反而低於 DA 的 87% 約 17%。再深入檢視 DA 數據，DA「生育地」推測之漏授 (100% - 生產者精度平均值 = 100% - 87%) 僅 13%，但 DA「生育地」推測之誤授 (100% - 使用者精度平均值 = 100% - 41%) 卻高達 59%，與 DT 的誤授 27% (100% - 73%) 相較，幾達 2.2 倍。由此可推斷 DA 預測誤差主要是發生在將「非生育地」(背景) 誤授為「生育地」(主體或目標物種 - 就表 2 而言是香桂)，大幅犧牲背景之準確度，來換取提高香桂「生育地」之準確度。這裡要特別強調的，是背景數量在整個試區所佔的比例通常遠大於主體數量所佔的比例，因此模式在預測「非生育地」時，應儘可能優先維持背景的高準確度，最好能在 90% 以上。惟 DA 統計方法產生的線性判別式，似乎有將「非生育地」與「生育地」兩群體均等分割的傾向，從而放寬「生育地」判定之標準，導致不少的「非生育地」被誤判為「生育地」。這正是 DA 方法的缺點所在，而此缺點更不利於找出模式尚未考量到之預測變數。因此，從推測準確度的角度來看，DT 模式確實遠優於 DA 模式。

表 3 顯示蘇鐵蕨 DT 及 DA 模式推測準確度之比較。表 3 所呈現蘇鐵蕨結果與表 2 所呈

表 2. 海拔、坡度及坡面位置建立兩種模式預測試區香桂適生育地之誤差矩陣表

Table 2. The error matrix for the two models based on elevation, slope, and terrain position predicting the suitable habitat of Randaishan cinnamons in the study area

模式	整體精度 (%)	<i>Kappa</i>	類別	驗證	生產者精度 (%)	使用者精度 (%)	
DT	建模	96	生育地	建模	78	84	
	驗模	91		驗模	62	70	62
	平均	94	0.68	建模	98	97	97
				驗模	95	97	95
DA	建模	81	生育地	建模	90	41	
	驗模	84		驗模	85	87	41
	平均	83	0.40	建模	80	98	98
				驗模	83	81	98

平均：分別是指建模及驗模的整體精度、*Kappa*、生產者精度、使用者精度的平均值。

現香桂結果甚為相近，亦即不論從建模、驗模整體精度或 *Kappa* 來看，DT 都遠優於 DA。其次，DA 在預測蘇鐵蕨的適生育地上與它在預測香桂的適生育地都顯現相同的缺點，即 DA 放寬了適生育地判定的標準，將「非生育地」誤授為「生育地」，以換取提高「生育地」的推測準確度。不過，這兩個表所呈現結果唯一的差異是表 3 中這兩模式預測蘇鐵蕨適生育地的各項數據都分別高於表 2 中兩模式預測香桂適生育地的各項數據，尤以 DT 更為明顯。蘇鐵蕨松風山與杜鵑嶺樣區相距小於 1 km，蘇鐵蕨樣株分布範圍較窄，其間距最大不超過 1 km，而在東峰溪流域樣區香桂樣株分布範圍較廣，其間距最大可達 5-6 km。從「空間自相

關」的觀念來看，蘇鐵蕨在空間連續性或相似性較香桂為高。因此，DT 與 DA 模式預測蘇鐵蕨適生育地之準確度高於香桂。

表 4 顯示兩種模式推測香桂及蘇鐵蕨適生育地面積與百分比之統計。在香桂部份，DT 推測所得面積佔試區總面積的 4.3%，而 DA 則為 11.7%；在蘇鐵蕨部份，DT 為 1.9%，而 DA 為 9.1%，這兩模式推測香桂及蘇鐵蕨全區適生育地的結果是一致的。這兩模式首次推測香桂適生育地即可從全區面積約 17,136.0 ha 篩選出約 4-12% 的高潛力區，而蘇鐵蕨為 2-9%，大幅減少後續地面調查之面積，從而降低所需投入經費、時間與人力。綜合表 2、表 3、表 4 的結果，不論是香桂或蘇鐵蕨，DT 模式

表3. 海拔、坡度及坡面位置建立兩種模式預測試區蘇鐵蕨適生育地之誤差矩陣表

Table 3. The error matrix for the two models based on elevation, slope, and terrain position predicting the suitable habitat of cycad-fern in the study area

模式	整體精度 (%)	<i>Kappa</i>	類別	驗證	生產者精度 (%)	使用者精度 (%)		
DT	建模	98	生育地	建模	92	平均	95	平均
	驗模	96		驗模	93	92	82	89
	平均	97	非生育地	建模	99	平均	98	平均
		0.85		驗模	96	97	99	99
DA	建模	86	生育地	建模	98	平均	55	平均
	驗模	84		驗模	98	98	48	52
	平均	85	非生育地	建模	84	平均	100	平均
		0.59		驗模	81	82	100	100

平均：分別是指建模及驗模的整體精度、*Kappa*、生產者精度、使用者精度的平均值。

表4. 海拔、坡度、坡面位置建立三種模式預測試區香桂及蘇鐵蕨適生育地之分布統計

Table 4. The distribution statistics of the potential habitat of Randaishan cinnamons and cycad-ferns in the study area predicted by the models based on elevation, slope, and terrain position

物種	香 桂				蘇 鐵 蕨			
	DT		DA		DT		DA	
預測類別	面積 (ha)	百分比 (%)	面積 (ha)	百分比 (%)	面積 (ha)	百分比 (%)	面積 (ha)	百分比 (%)
生育地	739.45	4.3	2,013.38	11.7	322.00	1.9	15,76.51	9.1
非生育地	16,396.55	95.7	15,122.62	88.3	16,814.00	98.1	15,559.49	90.9
合計	17,136.00	100.0	17,136.00	100.0	17,136.00	100.0	17,136.00	100.0

在推測兩物種的適生育地上均遠較 DA 更為嚴謹且精確。

接著是這兩模式在執行效率上之比較。三者在 SPSS 軟體裏有各自的模組，DT 與 DA 在點選模式各輸入參數後，皆可高度自動化執行模式之建立與驗證，產生各式的統計參數報表及誤差矩陣表，供模式效能評比之用，兩者可謂旗鼓相當。這兩模式在最後繪製香桂及蘇鐵蕨全區的適生育地分布圖時，都須跨軟體平台作業，轉移到 ERDAS Imagine 軟體的 Spatial Modeler 模組上來執行。DT 模式在 Spatial Modeler 模組下建立與執行繪圖程序較 DA 模式略為耗工費時。

(二) 三種取樣設計對模式空間外推能力之影響評估

本研究採用的取樣設計有三種，已詳述於第三節材料及方法。此處再強調的是本研究是以蘇鐵蕨三種取樣設計為比較基準（松風山、杜鵑嶺及關刀山林道樣區相距小於 1 km），來凸顯香桂三種取樣設計（東峰溪流域及關刀溪流域兩樣區相距約 5 km 上下）對於空間外推能力影響的反差。擬瞭解香桂「東峰模式」於建模時只納入東峰樣本而不納入關刀樣本，在空間阻隔狀況下，所謂「東峰模式」究竟是否可以跨越空間阻隔，精確地預測於關刀溪流域

的香桂適生育地？確切來說，就是在印證前言所提：僅含地形變數之預測模式在空間外推上不具有大空間尺度的普遍應用性。

表 5 顯示的結果為採東峰與關刀建模及驗模樣本的三種取樣設計對於兩種香桂模式預測準確度之影響評估。首先，從表 5 分別就兩種模式相互比較三種取樣設計之整體精度。DT 由取樣設計 -A 的整體精度 91%，明顯下降至取樣設計 -B 之 83%，降幅達 8%，再回升至取樣設計 -C 的 91%，而相對上 DA 由取樣設計 -A 的整體精度 84% 小幅下降至取樣設計 -B 的 81%，降幅僅 3%，再回升至取樣設計 -C 的 82%。在取樣設計 -C 建模及驗模皆採東峰與關刀混合樣本之下，兩模式整體精度又回升到取樣設計 -A 的水準，DT 尤為明顯。表 6 顯示的結果是採松風山、杜鵑嶺與關刀山林道建模及驗模樣本的三種取樣設計對於兩種蘇鐵蕨模式預測準確度之影響評估。表 6 呈現蘇鐵蕨整體精度升降與表 5 香桂的變化相似，惟其升降幅度不若香桂明顯，DT 小有升降，而 DA 幾無變化。這樣的結果表示從取樣設計 -A 換到取樣設計 -B 對於 DT 推測準確度的影響超過對於 DA 的影響，也似乎意味著 DA 在跨越空間分隔之推測能力優於 DT，但是否果真如此？

從表 5 與表 6 整體精度的升降變化上，可

表5. 三種取樣設計對於兩種香桂適生育地模式預測準確度影響之評估

Table 5. Assessment of the effects of three sampling designs on the accuracies of the two models for predicting the suitable habitat of Randaishan cinnamons

模式	整體精度 (%)	類別	取樣設計	生產者精度 (%)	使用者精度 (%)
DT	取樣設計-A	生育地	A	62	62
			B	0	0
			C	60	76
	取樣設計-B	非生育地	A	95	95
			B	95	87
			C	96	93
DA	取樣設計-A	生育地	A	85	41
			B	63	34
			C	85	47
	取樣設計-B	非生育地	A	83	98
			B	83	94
			C	82	97

註：漏授 = 100% - 生產者精度；誤授 = 100% - 使用者精度。

以看出有兩道力量在影響其變化，一道為 DT 及 DA 兩種模式，另一道為三種實驗設計。因此，在這裡就得先釐清兩道力量影響的差別所在。由表 5 先分別就兩種香桂模式各自比較三種取樣設計「生育地」的生產者精度，再分別就兩種香桂模式各自比較三種取樣設計「非生育地」的生產者精度。先就「生育地」的生產者精度來看，DT 從取樣設計 -A 於「生育地」之生產者精度為 62%，劇跌至取樣設計 -B 之 0%，再劇升至取樣設計 -C 的 60%；相對地，DA 從取樣設計 -A 的 85% 明顯下降至取樣設計 -B 之 63%，降幅達 22%，再回升至取樣設計 -C 的 85%。兩種香桂模式「生育地」的生產者精度都受到取樣設計 -B 建模不含關刀溪香桂樣本的影響，而都有明顯下跌，惟 DT 所受影響更大，劇降至 0%，將關刀溪香桂驗樣本全部推測錯誤。再就「非生育地」的生產者精度來看，DT 在三種取樣設計於「非生育地」之生產者精度幾近相等，分別是 95%、95%、96%；DA 也是如此，分別為 83%、83%、82%。值得注意的是，兩種模式「非生育地」生產者精度雖未受三種取樣設計差別之影響，但正如表 2 所指出的，DA 模式的缺點是放寬「生育地」判定之標準，導致不少的「非生

育地」被誤判為「生育地」，所以 DA「非生育地」之生產者精度在 83%，明顯低於 DT 的 95%，達 12%。這就是 DA 犧牲「非生育地」之生產者精度，來換取提高「生育地」的生產者精度的代價。表 6 顯示蘇鐵蕨的結果與表 5 香桂相近，兩種模式在「非生育地」的生產者精度差距甚大，DT 明顯高於 DA，兩模式也都受到三種取樣設計差異之影響而有升降，惟其變化幅度不若香桂明顯。這一點已在表 2 與表 3 比較時予以指出，應是香桂及蘇鐵蕨樣本分布與間距於空間自相關之差異所造成。

接著，由表 5 先分別就兩種香桂模式各自比較三種取樣設計「非生育地」的使用者精度，再分別就兩種香桂模式各自比較三種取樣設計「生育地」的使用者精度。先就「非生育地」使用者精度來看，DT 從取樣設計 -A 於「非生育地」之使用者精度 95% 下跌至取樣設計 -B 之 87%，下跌 8%，再回到取樣設計 -C 的 93%；相對地，DA 從取樣設計 -A 的 98% 下降至取樣設計 -B 之 94%，降幅 4%，再回到取樣設計 -C 的 97%。兩模式「非生育地」的使用者精度都受到取樣設計差別的影響而小有升降，雖 DA 變化幅度略小一些，但兩模式仍維持在高準確度的狀態。再就「生育地」使用

表6. 三種取樣設計對於兩種蘇鐵蕨適生育地模式預測準確度影響之評估

Table 6. Assessment of the effects of three sampling designs on the accuracies of the two models for predicting the suitable habitat of cycad ferns

模式	整體精度 (%)	類別	取樣設計	生產者精度 (%)	使用者精度 (%)
DT	取樣設計-A	生育地	A	93	82
			B	44	63
			C	98	78
	取樣設計-B	非生育地	A	96	99
			B	93	93
			C	94	100
DA	取樣設計-A	生育地	A	98	48
			B	100	42
			C	98	56
	取樣設計-B	非生育地	A	82	99
			B	81	100
			C	82	99

註：漏授 = 100% - 生產者精度；誤授 = 100% - 使用者精度。



者精度來看，DT 從取樣設計 -A 於「生育地」之使用者精度 62%，劇跌至取樣設計 -B 之 0%（陰影區），再劇升至取樣設計 -C 的 76%；相對地，DA 從取樣設計 -A 的 41% 下降至取樣設計 -B 之 34%，降幅 7%（陰影區），再回升至取樣設計 -C 的 47%。兩模式「生育地」的使用者精度都受到取樣設計之差別影響而有大幅升降，尤其是 DT 受到取樣設計 -B 無關刀溪流域香桂建模樣本的影響遠大於 DA，使 DT 於取樣設計 -B「生育地」的使用者精度低於 DA，惟 DT 在取樣設計 -A 及取樣設計 -C 的使用者精度仍明顯皆高於 DT。同時，表 6 顯示蘇鐵蕨的結果與表 5 香桂相近，兩種模式在「非生育地」的使用者精度差距甚小，然在「生育地」的使用者精度則差距甚大，DT 皆明顯高於 DA；兩模式都受到三種取樣設計差異之影響而有升降，惟其變化幅度不若香桂明顯。這一點已在表 2 與表 3 比較時予以指出，應是香桂及蘇鐵蕨樣本分布與間距於空間自相關之差異所造成。

綜合上面所做之討論來看，在取樣設計 -A 及取樣設計 -C 下，DT 模式推測準確度明顯優於 DA 模式，的確毋庸置疑，然而在取樣設計 -B 下，DA 卻優於 DT。惟實際上是 DA 統計模式對「生育地」判定之標準遠較 DT 模式來的寬鬆，導致不在少數的「非生育地」被誤判為「生育地」，大幅犧牲「非生育地」的推測準確度，以換取提昇「生育地」的推測準確度所造成，此正所謂「歪打正著」。因此，在大多數狀況下，DT 模式確實優於 DA 模式。現在將兩種模式優劣差異，即主要力道之一卸除之後，剩下的是另一主要力道，即三種取樣設計下香桂與蘇鐵蕨樣本分布與數量差異之影響。

在這裡就回歸前言所提：在蘇鐵蕨的反差對比下，本研究印證僅含地形變數且不納入關刀溪流域香桂建模樣本的「東峰模式」無法將東峰溪流域之香桂生態樣式正確地外推到相隔 5 km 的關刀溪流域。從表 6 顯示蘇鐵蕨的三種取樣設計結果之反差對比下，「松風杜鵑

模式」僅含海拔、坡度、坡面位置地形變數，並僅藉由松風山與杜鵑嶺的蘇鐵蕨樣本代入模式，可以正確外推蘇鐵蕨生態樣式至相距小於 1 km 的關刀山林道，惟短距離之空間相隔仍然會造成推測準確度略微下降。相對地，從表 5 香桂的三種取樣設計結果顯示，特別是取樣設計 -B，「東峰模式」僅含三個相同地形變數，並僅藉由東峰香桂樣本代入模式，雖然可以正確推測東峰溪流域之香桂適生育地，但是無法跨越空間阻隔正確地外推關刀溪流域之香桂適生育地。總而言之，僅含地形變數的預測模式無法普遍應用在大空間尺度之空間外推（前言所提兩個定義），於焉證實。

## 五、結論與建議

根據不同準確度指標評估，DT 的表現明顯優於 DA。DA 的缺點是將背景「非生育地」誤授為主體「生育地」，導致 DA 預測能力明顯不及 DT。就節省地面調查所需的經費及人力而言，DT 的成效明顯優於 DA。就執行效率來看，DA 略優於 DT，惟差距有限。整體而言，DT 在香桂及蘇鐵蕨適生育地預測上均優於 DA，較 DA 更加適用。

採用衛星影像原本是為跨越空間上的阻隔，提昇模式空間外推的能力，然而此次所採用之 SPOT-5 影像導出植生指標幾無提昇模式預測能力之效果，乃因 SPOT-5 的近、中紅外波段皆為寬頻，其光譜解析度不足以分辨樹種間細微光譜反射差異，而且空間解析度也不足以分辨台灣複雜、散生之林相，再加上大氣效應、光照與感測儀觀測幾何效應及地形效應所導致光譜變異之影響所致。

本研究從上述第四節的討論可獲得下面四點推論。第一，惠蓀林場內關刀溪與東峰溪兩個流域雖毗鄰相連，但它們香桂樣株所在地文特性卻無法互通共用，表示兩者的微地形仍有顯著差異存在。第二，因香桂所在兩個流域微地形之差異，造成這兩種統計模式無法僅藉由地文因子跨越空間阻隔，從一個流域將其香

桂生態樣式正確地外推至另一個流域（前言所提及空間外推的第二個定義）。第三，在目標物的空間外推上應嘗試從它找出具備跨越空間阻隔之環境特性作為模式的預測變數，例如目標物與背景之間光譜反射的細微差異，是未來後續研究探討的重點之一。第四，這兩種模式在香桂「生育地」的使用者精度在 40-76% 範圍之間，這表示模式可能還遺漏某些重要的生育地因子未納入考量中，諸如微地形、微氣候、土壤及周遭生物等因子都有可能。

影響模式推測能力並非僅有使用的統計方法，也包含資料的品質、預測變數和目標樹種樣株的數量和分布等，其中遙測資料的品質、所萃取的光譜預測變數及目標樹種樣株數量及分布均甚為重要，都是後續研究的重點。在未來之後續研究將考慮使用高空間解析度影像、高光譜影像及光達資料，期能改善模式於空間外推能力，並減輕生態學空間外推所遭遇之兩難困境。

## 六、參考文獻

- 王子定 (1974) 理論育林學 (上)。國立編譯館，第48頁。
- 呂金誠、劉業經、陳明義 (1986) 紀蘇鐵蕨 (*Brainea insignis*) 在臺灣之一新分布及其植物社會組成。中華林學季刊19 (1)：121-126。
- 林震岩 (2008) 多變量分析：SPSS的操作與應用。智勝文化。第489-530頁。
- 黃俊英 (2007) 多變量分析，第七版。華泰文化。第137-163頁。
- 劉棠瑞、蘇鴻傑 (1983) 森林植物生態學。臺灣商務印書館，第462頁。
- 劉業經、呂福原、歐辰雄 (1994) 台灣樹木誌。國立中興大學農學院。第440頁。
- 鄧家駒 (2004) 多變量分析。華泰文化。第369-394頁。
- 謝立忻、羅南璋、黃凱易 (2005) 應用3S地球空間技術於植群空間分布型態之探討。林業

研究季刊27 (4)：37-46。

- 蘇鴻傑 (1987) 森林生育地因子及其定量評估。台灣省林業試驗所研究報告季刊20 (1)：1-14。
- 羅南璋 (1992) 惠蓀實驗林場東峰溪集水區植群分析 國立中興大學森林研究所碩士論文。
- 羅南璋 (2010) 3S技術結合多變量統計模擬測繪林木之潛在生育地。國立中興大學森林研究所博士論文。
- Box, E. O. (1981) Macroclimate and Plant Forms: An Introduction to Predictive Modeling in Phytogeography, Junk, The Hague, 258 pp.
- Breiman, L., J. H. Friedman, R. A. Olsen, and C. G. Stone (1984) Classification and Regression Trees. Chapman and Hall, New York, USA., 357 pp.
- Brown, D. G. (1994) Predicting vegetation types at treeline using topography and biophysical disturbance variables. Journal of Vegetation Science, 5: 641-656.
- Brzeziecki, B., F. Kienast, and O. Wildi (1993) Modeling potential impacts of climate change on the spatial distribution of zonal forest communities in Switzerland. Journal of Vegetation Science, 4: 499-508.
- Burke, I. C., W. A. Reiners, and R. K. Olson (1989) Topographic control of vegetation in a mountain big sagebrush steppe. Vegetation, 84: 77-86.
- Chang, K. T. (2010) Introduction to Geographic Information Systems, 5<sup>th</sup> ed., The McGraw-Hill Companies, New York., pp. 276-356..
- De'ath, G. and K. E. Fabricius (2000) Classification and regression trees: a powerful yet simple technique for ecological data analysis, Ecology, 81 (11): 3178-3192.
- Decoursey, D. G. (1992) Developing models with more detail: do more algorithms give more

- truth? *Weed Technol.*, 6: 709-715.
- Fisher, R. A. (1936) The use of multiple measurements in taxonomic problems. *Annals of Eugenics*, 7: 179-188.
- Fischer, H. S. (1990) Simulating the distribution of plant communities I an alpine landscape. *Coenoses*, 5: 37-43.
- Gini C. (1912) Variabilità e mutabilità, Studi economico-giuridici pubblicati per cura della Facoltà di Giurisprudenza della Regia Università di Cagliari, Anno III, parte 2, also reproduced in C. Gini (1939) op. cit.
- Gove, P. B. and Merriam-Webster (eds.) (1986) Webster's Third New International Dictionary of the English Language, Unabridged. Springfield (MA): Merriam-Webster.
- Guisan, A. and N. E. Zimmermann (2000) Predictive habitat distribution models in ecology. *Ecological Modeling*, 135: 147-186.
- Guisan, A., J. P. Theurillat, and F. Kienast (1998) Predicting the potential distribution of plant species in an alpine environment. *Journal of Vegetation Science*, 9: 65-74.
- Guisan, A., S. B. Weiss, and A. D. Weiss (1999) GLM versus CCA spatial modeling of plant species distribution. *Plant Ecology*, 143: 107-122.
- Halpin, P. N. (1994) GIS analysis of the potential impacts of climate change on mountain ecosystems and protected areas. In: Price, M. F. and D. I. Heywood (Eds.), *Mountain Environments and Geographic Information Systems*. Taylor and Francis, London.
- Hoffer, R. M. (1978) Biological and physical considerations in applying computer-aided analysis techniques to remote sensor data. In: Swain, P. H. and S. M. Davis (Eds.), *Remote Sensing: The Quantitative Approach*, McGraw-Hill, Inc., New York, pp.227-289.
- Huntley, B., P. M. Berry, W. Cramer, and A. P. McDonald (1995) Modeling present and potential future ranges of some European higher plants using climate response surfaces. *Journal of Biogeography*, 22: 967-1001.
- Jensen, J. R. (2005) *Introductory Digital Image Processing-A Remote Sensing Perspective*, 3rd ed., Pearson Education, Inc, New Jersey.
- Korzukhin, M. D., M. T. Ter-Mikaelian, and R. G. Wagner (1996) Process versus empirical models: which approach for forest ecosystem management. *Canadian Journal of Forest Research*, 26: 879-887.
- Leary, R. A. (1985) *Interaction Theory in Forest Ecology and Management*. Nijhoff, Dordrecht, The Netherlands.
- Leersnijder, R. P. (1992) Pinogram: a pine growth area model. *Ecological Modeling*, 61: 1-139.
- Lenihan, J. M. (1993) Ecological responses surfaces for north American tree species and their use in forest classification. *Journal of Vegetation Science*, 4:667-680.
- Lillesand, T. M., R. W. Kiefer, and J. W. Chipman (2008) *Remote Sensing and Image Interpretation*, 5th Edition. John Wiley & Sons. Inc., New York.
- Miller, J. R., M. G. Turner, E. A. H. Smithwick, C. L. Dent, and E. H. Stanley (2004) Spatial extrapolation: the science of predicting ecological patterns and processes. *BioScience*, 54 (4): 310-320.
- Moore, D. M., B. G. Lees, and S. M. Davey (1991) A new method for predicting vegetation distributions using decision tree analysis in a geographic information system. *Environment Management*, 15: 59-71.
- Neilson, R. P. (1995) A model for predicting continental-scale vegetation distribution and water balance. *Ecol. Appl.*, 5: 362-385.

- O'Brien, C. S., S. S. Rosenstock, J. J. Hervert, J. L. Bright, and S. R. Boe (2005) Landscape-level models of potential habitat for Sonoran pronghorn. *Wildlife Society Bulletin*, 33 (1): 24-34.
- Pereira, J. M. C. and R. M. Itami (1991) GIS-based habitat modeling using logistic multiple regression: A study of the Mt. Graham red squirrel. *Photogrammetric Engineering & Remote Sensing*, 57 (11): 1475-1486.
- Pickett, S. T. A., G. Kolasa, and C. G. Jones (1994) *Ecological Understanding: the Nature of Theory and the Theory of Nature*. Academic Press, New York.
- Prentice, I. C. (1986) Some concepts and objectives of forest dynamics research. In: Fanta, J. (Ed.), *Forest Dynamics Research in Western and Central Europe*. PUDOC, Wageningen, pp. 32-41.
- Prentice, I. C., W. Cramer, W., S. P. Harrison, R. Leemans, R. A. Monserud, A. M. Solomon (1992) A global biome model based on plant physiology and dominance, soil properties and climate. *Journal of Biogeography*, 19: 117-134.
- Sharpe, P. J. A. and E. J. Rykiel, Jr. (1991) Modeling integrated response of plants to multiple stress. In: Mooney, H. A., W. E. Winner, and E. J. Pell (Eds.), *Response of Plants to Multiple Stress*. Academic Press, San Diego, CA, pp. 205-224.
- Skidmore, A. K. (1990) Terrain position as mapped from a gridded digital elevation model. *International Journal of Geographical Information Science*, 4: 33-49.
- Solomon, A. M. and R. Leemans (1990) Climatic change and landscape-ecological response: issues and analysis. In: Boer, M. M. and R. de Groot (Eds), *Landscape-ecological impact of climatic change*, Proceedings of a European Conference, Lunteren, The Netherlands, pp. 293-317.
- Sperduto, M.B. and R. G. Congalton (1996) Predicting rare orchid (small whorled Pogonia) habitat using GIS. *Photogrammetric Engineering & Remote Sensing*, 62 (11): 1269-1279.
- Turner, M. G., V. H. Dale, and R. H. Gardner (1989) Predicting across scales: Theory development and testing. *Landscape Ecology*, 3: 245-252.
- Verbyla, D. L. (1987) Classification trees: a new discrimination tool. *Canadian Journal of Forest Research*, 17: 1150-1152.
- Walter, H. and E. Walter (1953) Das Gesetz der relativen Standortskonstanz. *Das Wesen der Pflanzengesellschaften*. *Ber. Dtsch. Bot. Ges.*, 66: 228-236.
- Wang, J. and J. Ni (2006) Review of modeling the distribution of plant species. *Journal of Plant Ecology*, 30 (6): 1040-1053.
- Wissel, C. (1992) Aims and limits of ecological modeling exemplified by island theory. *Ecological Modeling*, 63: 1-12.
- Williams, B. K. (1983) Some observations of the use of discriminant analysis in ecology. *Ecology*, 64: 1283-1291.
- Woodward, F. I. (1987) *Climate and Plant Distribution*. Cambridge University Press, Cambridge, 174 pp.