

資料不確定下之地質岩性推估 — 以臺北盆地為例

Uncertainty Analysis of Lithological Classification in Taipei Basin

國立臺灣大學
生物環境系統工程學系
碩士

吳孟庭*
Meng-Ting Wu

國立臺灣大學
生物環境系統工程學系
博士

林遠見
Yuan-Chien Lin

國立臺灣大學
生物環境系統工程學系
教授

余化龍
Hwa-Lung Yu

摘 要

在工程應用領域和環境科學分析等研究中，對於研究區域的地質組成架構有基本的掌握，對於地質岩性基本的分佈情形有所了解，能夠對分析地質、水文、地下水或是洪旱災等，環境相關研究進行時有所幫助，使分析有所依據和參考。但由於許多實際觀測上的限制，例如先天的環境困境或是後天採樣的屏障，使得實際能夠使用的採樣和觀測資料有限，因此許多研究進行時僅能使用有限的資料進行分析。在此情況之下為有效提高地質空間的解析度，並了解未知空間點之岩性分佈，傳統上也有許多空間統計方法應用于此，而本研究應用的貝氏最大熵法(Bayesian Maximum Entropy, BME method)同時考慮時空間之相關性，並且同時處理確定和不確定性資料。以貝氏條件機率概念為基礎，其理論建立于完整的數理哲學概念之上，本研究嘗試應用迭代演算法進行模式收斂，並於收斂過程中考慮模式限制式條件所含之不確定性範圍。而針對台北盆地地質岩性分類之推估，即是屬於離散分佈的類別型資料，因此本研究整體應用方法稱為類別型貝氏最大熵法(Categorical BME)，望能透過有限的地質鑽探之岩心資料，建立完整的三維類別型貝氏最大熵法岩性推估模式，推估台北盆地之三維岩性分佈情形，期望可以供台北盆地地質相關研究和實務建設之參考。

關鍵詞：類別型貝氏最大熵法，迭代演算法，限制式條件正規化，地質岩性分類，水文地質架構。

*通訊作者，國立臺灣大學生物環境系統工程學系碩士，10617 台北市中正區羅斯福路四段 1，b98602046@ntu.edu.tw

ABSTRACT

In environmental or other scientific applications, we must have a certain understanding of geological lithological composition. Because of restrictions of real conditions, only limited amount of data can be acquired. To find out the lithological distribution in the study area, many spatial statistical methods used to estimate the lithological composition on unsampled points or grids. This study applied the Bayesian Maximum Entropy (BME method), which is an emerging method of the geological spatiotemporal statistics field. The BME method can identify the spatiotemporal correlation of the data, and combine not only the hard data but the soft data to improve estimation. The data of lithological classification is discrete categorical data. Therefore, this research applied Categorical BME to establish a complete three-dimensional Lithological estimation model. Try to regularize the maximum entropy density estimation. Apply the limited hard data from the cores and the soft data generated from the geological dating data and the virtual wells to estimate the three-dimensional lithological classification in Taipei Basin.

Keywords: Categorical bayesian maximum entropy method, Iterative scaling, Regularization, Lithological classification, Hydrogeological setting.

一、前 言

1.1 研究緣起

在環境科學研究等領域中，由於操作上的限制、經費不足或先天條件限制等原因導致採樣資料有限，使得空間解析度不足。例如在水文領域中，從降雨觀測、河川流量、地下水井觀測、水文地質參數等資料，皆是空間中的有限點觀測資料。為提高空間解析度並了解未知空間點的資料情形，我們必須透過許多空間統計方法來對未知點或未知網格點進行推估。例如距離反比加權法、克利金法(Kriging method)、貝氏最大熵法(Bayesian Maximum Entropy, BME method) (Christakos, 1990, 1991; Christakos & Li, 1998; Christakos & Serre, 2000)。

地理統計方法的應用中，較常使用的方法是克利金法推估，然而它具有資料常態分佈及線性推估等假設且只能考慮確定性的資料(hard data)，而貝氏最大熵法(BME)為一新興時間空間地理統計方法，此方法可結合數值模式方法與資料導向方法，並可以同時考慮空間與時間相關

性，不受到高斯分佈及線性推估的限制，且更具有推估準確性(Bogaert & D'Or, 2002; Christakos & Li, 1998; D'Or, Bogaert, & Christakos, 2001)。

1.2 研究目的

貝氏最大熵法近年被學者進一步發展為針對離散資料或類別型資料進行空間推估的「類別型貝氏最大熵法(Categorical BME)」(Bogaert, 2002)。自然界中，有很多資料都屬於離散的類別型資料，例如土壤型態、地質分類、土地利用、植被等，這些資料在進行空間推估時就必須採用類別型貝氏最大熵法進行推估。為分析未知點之整體水文地質參數，需要對研究區域的岩性有通盤的了解，而岩性即是一種離散的類別型資料，本研究欲透過有限的地質鑽探所取得之岩心資料和定年資料，加入其他多項地質相關研究成果資料，如虛擬井資料和經驗正交函數法 EOF 之結果，多筆資料加以整合後輸入本研究模式用於推估未知點。

進行完整的三維類別型貝氏最大熵法岩性推估模式建立，需同時考量樣本資料的限制式條

件，依據 Dudik, Phillips, and Schapire (2004)提出的理論，加以放鬆限制式的範圍。最後將前處理分析後的多筆參考資料輸入本研究模式，並以限制式不確定性範圍進行收斂後，推估各類別間機率的相關性以及未知點的岩性類別，完整建立台北盆地的三維岩性分佈，以供其他相關研究或實際建設參考。

二、理論概述

2.1 類別型離散資料

類別型資料(categorical data)為一種離散型資料，過去有許多針對分析連續型資料的研究，但對於類別型資料分析的研究則相對較少(D'Or & Bogaert, 2004)。而實際上許多現地實驗和研究的結果，都是用類別型的離散資料做呈現，例如土壤型態、地質分類、土地利用、植被等統計分類的結果。因此針對類別型資料的分析，需先建立在離散分佈下的類別型隨機場，定義其中相關的各项變數和參數，以及二階定常性假設下的限制式條件(Bogaert & Fasbender, 2008)。

傳統上應用於處理類別型資料的地理統計方法(Journel & Huijbregts, 1978)，例如 Indicator kriging (IK)、Indicator cokriging (Journel, 1989)和 neural networks (Venables & Ripley, 1994)等方法經常被使用，其建立在二階定常性的隨機場下為線性的高斯分佈，期望值(均值)需符合 best linear unbiased estimator (BLUE)的無偏差假設。當實際的資料並非常態分佈時，就不再符合線性假設，因此傳統的統計方法缺乏足夠合理的理論基礎，也存在許多使用上的限制(Goovaerts, 1997)。需運用其他統計方法來找到研究資料的條件機率分佈，而貝氏最大熵法能夠處理非線性的狀況，並解決傳統方法存在的問題(Bogaert & Fasbender, 2008)。

2.2 貝氏最大熵法 Bayesian Maximum Entropy (BME)

貝氏最大熵法(BME)為一新興時間空間地理統計方法，此方法可結合數值模式方法與資料導

向方法，並可以同時考慮空間與時間相關性。除此，BME 法不受到高斯分佈及線性推估的限制，並以貝氏條件機率概念結合物理知識或其他不確定性資料(soft data)增強推估資訊，同時對序率(stochastic)資料與空間(spatial)資料進行推估(Christakos, 1990, 1991; Christakos & Li, 1998; Christakos & Serre, 2000)。自 1991 年被提出以來，貝氏最大熵法已經在學術上被廣泛應用在不同的科學與工程領域並且被證實較傳統空間統計方法(如各種克利金法)更具有推估準確性(Allard, 2011; Bogaert & D'Or, 2002; Christakos & Li, 1998; D'Or *et al.*, 2001)。

在貝氏最大熵法(BME)的推估過程中，水文地質特性可代表一個時空隨機場(spatial-temporal random field)。令 $X_p = X_{s,t}$ ，表示在時空隨機場一項水文地質屬性，而其中向量 $P = (s,t)$ 是一個時空的點(s 為空間中地理座標位置， t 為時間)。隨機場模式提供了普遍和數學嚴謹架構去調查水文地質的現象。

這樣可以提高在複合空間-時間域的預測能力。在機率場模式的機率密度函數(probability density function, PDF)， f_{KB} 可表示如下：

$$P_{KB}[\chi_1 \leq X_{p_1} \leq \chi_1 + d\chi_1, \chi_2 \leq X_{p_2} \leq \chi_2 + d\chi_2, \dots] = \int f_{KB}(X_{p_1}, X_{p_2}, \dots) d\chi_1 d\chi_2 \dots \quad (1)$$

其中，下標的 KB 是指使用知識庫(knowledge base, KB)去構成機率密度函數。

BME 方法定義的特徵包括了(1)一般知識(general knowledge base)，又稱為 G-KB。(2)特殊位置(site specific knowledge base)的知識，又稱 S-KB。全部的知識庫就是 $k = G \cup S$ 。BME 的基礎方程式如下：(細節請參照(Christakos & Olea, 2005; Christakos & Serre, 2000))

$$\left. \begin{aligned} \int d\chi(g - \bar{g}e^{\mu^T g} = 0 \\ \int d\chi \xi_s e^{\mu^T g} - Af_k(X_p) = 0 \end{aligned} \right\} \dots \dots \dots (2)$$

其中， g 為 $g_\alpha (\alpha = 1, 2, \dots)$ 方程式中的一個向量，描述已知的 G-KB。 μ 為 $\mu_\alpha (\alpha = 1, 2, \dots)$ 所組成的向量，為 g_α 的係數，隨空間時間座標改變而改

變。 ξ_s 代表 S-KB。 A 是一個標準化的參數 (normalization parameter), f_k 是水文地質特性在時空每一點的機率密度函數, 下標 K 的意思是指混合 G-KB 跟 S-KB。

G-KB 參考整個研究範圍 p , 構成了時空點向量 p_k 的水文地質參數估計。點向量 p_{data} 是特殊位置的有用資訊。在 G-KB 中包括了時空相關的統計學知識, 例如平均值(mean)、共變異係數(covariance)、多點統計(multiple-point statistics)與連續階(continuity orders)等。其中平均值與共變異係數是在研究中最常被使用。重要的參數都是在隨機場 X_p 中的特性, 過去在連續變數的推估上, 一般採用共變異係數(covariance)作為空間關係的限制條件; 而針對離散型或類別型變數則需改以上節所述之 indicator variograms 作為推估時空間關係的描述。在 S-KB 中主要是由確定性資料(hard data)的點向量 p_{hard} 和不確定性資料(soft data)的點向量 p_{soft} 所構成的, 整合 G-KB 和 S-KB 在機率密度函數中的 f_{KB} 在式 1 中, 我們就可以描述推估點 p_k 水文地質的時間空間分佈。

2.3 類別型貝式最大熵法 Categorical BME

貝氏最大熵法(BME)的概念原則(Christakos, 1990), 也被引用到地理統計方面, 對類別型資料進行空間統計推估(Allard, 2011; Bogaert, 2002), 特別是應用貝氏最大熵法到土壤資料分析的相關研究(D'Or *et al.*, 2001; D'Or & Bogaert, 2004)。因此本研究針對土壤岩性的類別進行推估分析, 參照前人相關的類別型空間推估研究, 並著重在應用迭代演算法部分的改良, 將其應用在臺灣當地的土壤地質分析。此方法可以有效的將所有已知的信息綜合起來, 並且無偏見的處理所有的不確定性, 取得最佳的機率密度函數, 加以推估分析其他未知點。

2.4 迭代演算法 Iterative scaling

過去由於電腦的運算能力還不足以處理最大熵的演算, 最大熵法的概念還無法實際應用, 但近年來逐漸克服電腦運算上的限制, 貝氏最大熵法才逐漸廣為應用, 且相關的演算法也有諸多

改良和演進。最大熵法的基本概念, 就是要在符合所有限制條件之下, 盡可能找到一個最均勻(uniform)的機率分佈, 使得模式分析的結果能夠涵蓋所有已知的資訊, 並且避免在未知的資訊上作過多的主觀推論(Adam L. Berger, 1996)。

在無法直接求解的情況之下, 為符合最大熵法的概念找到最趨近均勻的分佈, 需要運用到演算法進行最佳化求解, 而其中迭代演算法(Iterative scaling)則是針對離散資料推估的一種演算法。此演算法之發展、演進和改良的相關研究已行之有年, 其基本參數定義如表 1。

表 1 最大熵法基本參數定義

x	已知的任一資訊(例如觀測井的某岩性類別)
y	同時出現的其他資訊(例如觀測井的其他岩性類別)
(x_N, y_N)	N 個已知的樣本資訊組合(可能是任意兩類別的組合)
$p(y x)$	在已知資訊 x 的條件下, 同時出現資訊 y 的條件機率

當我們採樣的樣本數有 N 個, 樣本組合為 $(x_1, y_1), (x_2, y_2), \dots, (x_N, y_N)$, 則任意兩個資訊 (x, y) 同時出現的樣本機率定義如下

$$\tilde{p}(x, y) \equiv \frac{1}{N} \times (x, y) \text{ 在樣本中出現的次數} \quad (3)$$

類別型資料要推估兩種資訊類別同時出現的機率, 因此首先需給定 Indicator function 為

$$f(x, y) = \begin{cases} 1, & \text{當 } x \text{ 和 } y \text{ 同時出現時} \\ 0, & \text{其他} \end{cases} \quad \dots\dots\dots (4)$$

此時, 樣本機率 $\tilde{p}(f)$ 和理論機率 $p(f)$ 定義為

$$\tilde{p}(f) \equiv \sum_{x, y} \tilde{p}(x, y) f(x, y) \quad \dots\dots\dots (5)$$

$$p(f) \equiv \sum_{x, y} \tilde{p}(x, y) p(y|x) f(x, y) \quad \dots\dots\dots (6)$$

令機率期望值相等 $p(f) = \tilde{p}(f)$ 為限制條件, 則由(式 5)和(式 6)得

$$\sum_{x,y} \tilde{p}(x)p(y|x)f(x,y) = \sum_{x,y} \tilde{p}(x,y)f(x,y) \dots(7)$$

在已知資訊 x 的條件下，同時出現資訊 y 的條件機率 $p(y|x)$ ，即是欲推估的目標函數，熵可定義如下

$$H(p) \equiv -\sum_{x,y} \tilde{p}(x)p(y|x)\log p(y|x) \dots\dots(8)$$

需找到一個機率密度函數 $p(y|x)$ 使得熵 $H(p)$ 最大，此過程即是在限制式條件下的最佳化求解過程，定義其 Lagrangian 為

$$\Lambda(p, \lambda) \equiv H(p) + \sum_i \lambda_i(p(f_i) - \tilde{p}(f_i)) \dots\dots(9)$$

應用兩兩類別 (x,y) 下的 Indicator function(式 4, $f(x,y)$) 視為不同 feature(f_i)，並給定每個 f_i 一個乘數 λ_i (a Lagrange multiplier)，因此熵的最佳化求解，可由 $\Lambda(p, \lambda)$ 偏微分求得每個 f_i 對應的參數 λ_i 。或應用最大似法(maximum-likelihood approach)求解參數 λ_i ，最大熵法和最大似法兩者為對偶函數(dual function)，都可用來求解參數 λ_i 進而找到目標函數的機率分佈，詳見(Adam L. Berger, 1996)。

由於此最佳化求解無法找到解析解，因此需要應用數值分析方法中的迭代演算法求解，最原始的最大熵模型訓練方法為「通用迭代演算法(generalized iterative scaling, GIS)」(Darroch & Ratcliff, 1972)，之後較常被應用的是改良過的「改進式迭代演算法(improved iterative scaling, IIS)」(Stephen Della Pietra, 1997)。並且將此方法加以應用在類別型資料的推估演算 (Bogaert, 2002; D'Or & Bogaert, 2004)，嘗試將類別型資料和連續型資料做結合 (M.A. Wibrin, 2006; P. Bogaert, 2004)。

而過去的迭代演算法是對於限制式條件百分之百信任，但實際上樣本的均值和變異函數並不能完全等同於現地真實狀況的理論值，可能由於採樣的不同就會有所改變進而產生誤差。因此迭代演算法仍需考量到限制式的不確定性誤差範圍，應用 Regularization 調整最大熵的機率密度函數推估(Dudik *et al.*, 2004)。

2.5 最大熵推估正規化 Regularization

迭代演算法的演算過程也是一種機器學習(Machine Learning)，過去在 ML 領域的許多理論都有提到，不能忽視演算法中限制式所含的不確定性。因此近年來有許多研究著重在如何於演算法中，同時考量限制式不確定性的範圍，也有許多研究論文針對這個問題提出相關的數學演算改良(林軒田, 2013)。

因此，本研究欲應用 Bootstrap 算出限制式條件的不確定性範圍後，參照 Dudik *et al.* (2004) 對於最大熵機率密度函數推估正規化之理論(Regularization)，延伸應用到類別型貝式最大熵法(Categorical BME)。在演算過程中放鬆每個限制式條件，同時運用蒙地卡羅馬可夫鏈法(MCMC)加以訓練，使經驗分佈的推估誤差趨於最小。

原先的限制式條件為樣本機率 $\tilde{p}(f)$ 和真實理論機率 $p(f)$ 相等，

$$p(f_i) = \tilde{p}(f_i) \quad \text{for } i \in \{1, 2, \dots, n\} \dots\dots\dots(10)$$

其中針對每個 f_i 所對應的模型參數 λ_i ，可將機率目標函數 $p_i(y|x)$ 表示為

$$p_i(y|x) \equiv \sum_{i=1}^n q_{\lambda_i}(y|x) = \frac{1}{Z_{\lambda}(x)} \exp\left(\sum_{i=1}^n \lambda_i f_i(x, Y)\right) \dots\dots\dots(11)$$

where

$$q_{\lambda_i}(y|x) = \frac{1}{Z_{\lambda}(x)} \exp(\lambda_i \times f_i(x, y)) \quad \text{for } i \in \{1, 2, \dots, n\}$$

並定義樣本機率 $\tilde{q}_{\lambda_i}(f_i)$ 和實際理論機率 $q_{\lambda_i}(f_i)$ 之間存在 β_i 的誤差範圍，

$$|q_{\lambda_i}(f_i) - \tilde{q}_{\lambda_i}(f_i)| \leq \beta_i \quad \text{for } i \in \{1, 2, \dots, n\} \dots(12)$$

令式 12 為新的限制式，以此為限制條件應用最大熵法的理論，找到滿足限制式且最趨於均勻的理論機率密度函數分佈。接著參照(Dudik *et al.*, 2004) 理論調整最初的最大熵公式，並應用 Lagrangian 和 maximum-likelihood approach 等方法轉換推導，最後求解的目標函式可改寫為 $Fi(\lambda, \delta)$ 解其最小值，

$$F_i(\lambda, \delta) = -\delta \tilde{q}_i[f_i] + q_i[f_i](e^\delta - 1) + \beta_i(|\lambda_i + \delta| - |\lambda_i|) \quad \text{for } i \in \{1, 2, \dots, n\} \quad \dots (13)$$

式 13 中 β_i 為每個 λ_i 的誤差範圍，本研究由 Bootstrap Method 計算取得；而 $\delta = \Delta\lambda$ 為 λ_i 迭代演算時的移步量， δ 可以定為式 14，依據 β_i 的大小擇其一使用。

$$\ln \left(\frac{(\tilde{q}_i[f_i] - \beta_i)(1 - q_i[f_i])}{(1 - \tilde{q}_i[f_i] + \beta_i)q_i[f_i]} \right) \quad \dots (14)$$

or $\ln \left(\frac{(\tilde{q}_i[f_i] + \beta_i)(1 - q_i[f_i])}{(1 - \tilde{q}_i[f_i] - \beta_i)q_i[f_i]} \right)$

最後同樣應用迭代演算法依照前述設定，解 $F_i(\lambda, \delta)$ 最小值之步驟如下，詳細請參照(Dudik *et al.*, 2004)：

- 輸入：特徵值 f_i ；樣本經驗分佈 $\tilde{q}_i[f_i]$ ；每個 λ_i 的誤差範圍 β_i
- 輸出：最佳化 λ_i^* ；最佳化機率分佈 $\tilde{q}_{i^*}[f_i]$ for $i \in \{1, 2, \dots, n\}$
 - I. 初始條件： $\lambda_i = 0$ for all $i \in \{1, 2, \dots, n\}$
 - II. 令 (λ, δ) 為式 13 的解，根據式 14 計算 δ ，調整參數 $\lambda_i = \lambda_i \leftarrow \lambda_i + \delta$
 - III. 重複步驟 II. 直到 λ_i 收斂

對每個 λ_i 都要分別執行上述步驟，直到整個模式都收斂。

三、研究區域之資料蒐集與概況分析

3.1 研究區域介紹

本研究以台灣北部淡水河流域之台北盆地為研究區域(圖 1)，台北盆地為台灣之首都，高度開發的建設遍佈整個區域，使得建設規劃和研究進行時，無法在每個位置直接經由鑽探的方式得到相關的地質資料。因此如能透過岩性推估模式，通盤了解整體的地質岩性結構，有助於地下水、水文和地質研究之進行和都市建設之規劃。

3.2 資料蒐集與前處理

本研究資料主要分為確定性資料(hard data)和不確定性資料(soft data)，每筆資料都要先經過前處理後，再放入岩性推估模式中一起對未知點



圖 1 台北盆地區域地質圖(中央地質調查所, 2014)

做岩性推估。而確定性資料為觀測井的岩性資料，不確定性資料目前處理到包含定年資料、虛擬井資料以及 EOF 地下水位變動分析資料，未來如有其他相關資料可作為參考，同樣可以加入模式之中進行推估。

根據中央地質調查所的岩心資料細分為 13 種岩性，每種岩性對應一個類別編號，本研究著重在台北盆地的水文地質架構建置，期望推估出整個區域礫、砂和泥的空間分布，因此整合岩性資料為礫、砂和泥三大類別進行分析，其中第一類為泥、第二類為砂、第三類為礫石，且通常泥層為阻水層，砂層和礫石層為透水層，則可以藉此大致判斷含水層和阻水層的分佈。

每筆資料的處理方式略有不同，但最終都要整理成坐標位置點 (包含經緯度的 XY 坐標以及深度的 Z 坐標)，及每個位置對應的岩性類別機率，整合如表 2 資料前處理整理格式範例(坐標系統為 TWD97)，再以此機率相關性資料整合進整體岩性的模式推估中。

3.2.1 確定性資料

本研究資料蒐集位於台北盆地 49 口地質觀測井的岩心鑽探資料，其空間分布如圖 2 所示，每口井都是經過實地地質鑽探分析後取得的岩心，可確切知道在此岩心的每個深度及其岩性類別，因此該筆資料對於岩性是百分之百確定的，則將其設定為確定性資料。

表 2 地調所岩性分類對照表

資料類型		粒徑												
		細 ←								→ 粗				
地調所	編號	1	2	3	4	5	6	7	8	9	10	11	12	13
	類別岩性	黏土	泥	粉沙	極細砂	細砂	中砂	粗砂	極粗砂	細礫	中礫	粗礫	巨礫	極巨礫
本研究	編號	1			2					3				
	類別岩性	泥層			砂層					礫石層				
地下水層		阻水層			含水層									

表 3 資料前處理整理格式範例

坐標位置點			岩性各類別機率 (總和機率為 1)		
Easting	Northing	Depth(m)	泥(類別 1)	砂(類別 2)	礫石(類別 3)
297460	2768120	1	1	0	0
295060	2770390	72	0	1	0
303430	2773750	107	0	0	1
296100	2771150	180	0.3	0.5	0.2

【註】坐標點第一列為泥層、第二列為砂層、第三列為礫石層，三筆資料皆為確定性資料；坐標點第四列為不確定性資料，每個類別都有機率值，且三個類別機率總和為 100%。

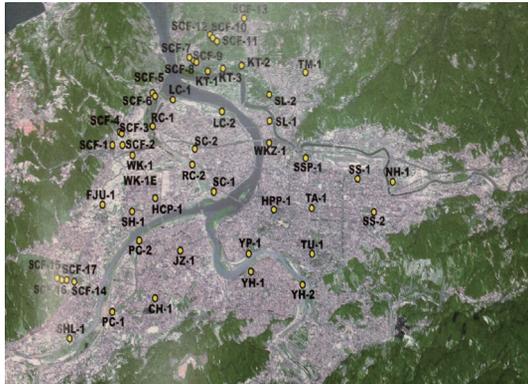


圖 2 研究區域地質井空間分布圖(費立沅, 2011)

其中每個位置點對應的岩性機率為 100%，如表 3-2 資料前處理整理格式範例中的前三列，為三個不同位置下的岩心資料，及其對應深度的岩性。以第一列的位置點資料為例，因為是泥層(類別 1)的確定性資料，則其類別 1 的岩性機率為 $100\% = 1$ ，其餘兩類別出現在這個位置點的機率為零，則砂(類別 2)和礫石(類別 3)之機率值為 0。

其中為使資料同時考量到岩層厚度的影

響，因此嘗試對岩心資料的深度進行不同厚度的切割之試驗，最後顯示原始資料以 1 m 為厚度進行分層能夠有較優的推估結果，因此本研究原始資料的深度將以 1 m 厚度做均勻裁切。

3.2.2 不確定性資料

本研究輸入之不確定性資料有三筆，第一筆資料為定年資料，該資料取自於台北盆地的 6 口井，其坐標點位置如表 4 所示，為圖 3 臺北地圖上的紅、藍點，每口井都是經過地質鑽探分析後取得的岩心資料，再經過碳 14 定年法估計其地質岩層形成的時間。因此可確切得到岩心的每個深度其岩性類別形成年份，因該筆資料對於岩性的形成時間是確定的，可藉此時間定年資料推估岩性生成時間的相關性。

第二筆資料為虛擬井資料，除了採樣得到的觀測井資料之外，採用中央地質調查所過去針對臺北盆地的幾個剖面，打虛擬井點後推估可能岩性分佈的相關成果，圖 3 中綠色線狀密集打點則為虛擬井資料位址。其岩性類別是用 3.2.1 的實測岩心資料所推估而來，從原先的 6 口觀測地質井資料延伸推估為 205 口虛擬井資料。

表 4 定年資料採樣站點位址

編號	Easting (97)	Northing (97)	井深 (m)	站名
010103G1	297461.67	2768118.37	180	江子翠
010501G1	295055.53	2770387.3	272	新莊
630401G1	303425.82	2773745.44	107	新生公園
630501G1	302049.8	2770640.48	263	二二八公園
630601G1	301520.2	2774320.26	181.3	衛工處
630701G1	300908.33	2768112.15	233.6	青年公園

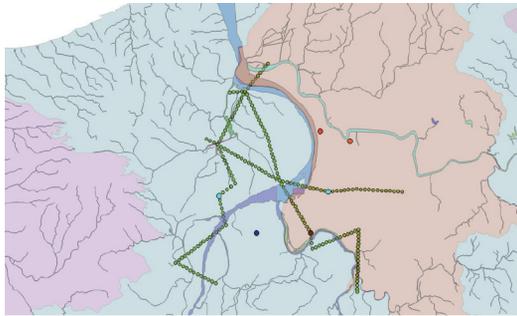


圖 3 臺北盆地定年資料和虛擬井資料分布情形

第三筆不確定性資料為 EOF 地下水位變動分析資料，本研究團隊應用經驗正交函數法 (EOF)，針對地下水日水位觀測資料，進行時空變動特徵分析之結果，結合實際岩性觀測之分佈特性後，做為不確定性資料輸入本岩性推估模式之中。

四、岩性推估模式建置

1. 研究流程

透過中央地質調查所在本研究區域，所執行之岩心鑽探取得的岩性資料和定年資料，首先整理以鑽探所得各分層之不同岩性資料為確定性資料 (hard data)，並以地質沈積的定年資料為不確定性資料 (soft data)，假設地質定年年份越相近且空間距離相近，其岩性相關性越高。由此，透過類別型貝氏最大熵法 (Categorical BME) 進行全區三維岩性推估，其中類別型貝氏最大熵法的推估

流程，主要根據 (Christakos, 2000) 所提出的三大主要步驟：

(1) 第一階段 (prior stage)：

要找出具有最大熵 (maximum entropy) 及遵守一般限制 (general constraints) 的聯合機率密度函數 (joint probability distribution)。如先驗所知的機率集合及研究方法中所述之一般知識 (G-KB)，實務上會根據 (cross-)covariance 或 (cross-)variograms 來得到，在離散或類別變數則是以 indicator (cross-)covariance 或 (cross-)variograms 為推估時空關係的限制。

(2) 第二階段 (meta-prior stage)：

找出研究區域所述之特殊知識 (S-KB) 並轉換成相對應的數學關係。

(3) 第三階段 (integration stage)：

整合並遵守上述 prior stage 之 maximum entropy distribution 和 meta-prior stage 之資訊下所計算之後驗條件分布 (posterior conditional distribution)。

本研究詳細流程如圖 4 所示：

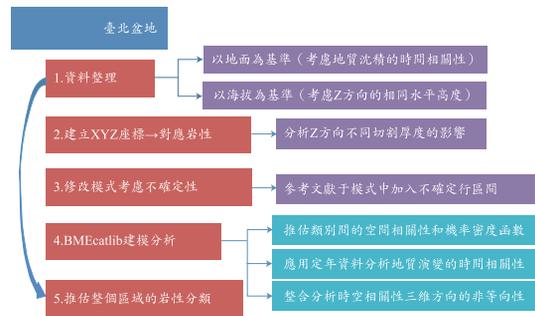


圖 4 類別型貝氏最大熵法三維岩性推估流程

4.2 前置分析

首先地質岩心資料整理以兩種深度基準進行分析，如考慮地質沈積的時間相關性，則以地面為基準平面，分析不同深度的岩性組成；如考慮 Z 方向為相同水平高度，則以海拔為基準平面進行分析。將 Z 方向依照不同厚度做切割後，建立三維 XYZ 坐標點和對應岩性，用以建模推估

兩兩類別間的空間相關性，計算不同距離下的機率密度函數。根據本研究前置分析結果顯示以地面為基準平面，其推估結果能夠呈現較好的相關性和準確率，可能因為以地面為基準，能夠同時考量到地質沈積、褶皺等變化，因此本研究以地面為基準面進行分析。

原始資料將岩性分成十三類，本研究將十三種細分的岩性類別，統合以泥、砂和礫石三大岩性類別進行分析，能夠更清楚看出三大岩性的分層，並且可延伸應用於分析阻水層和含水層的分佈。

本研究岩性推估主要分為兩大類型資料進行，分別是「確定性資料(hard data)」和「不確定性資料(soft data)」，確定性資料取自觀測井資料的實測岩性，不確定性資料則可以取多筆和地質相關的其他研究成果，本研究目前取用定年資料、虛擬井資料數化資料和 EOF 分析結果資料，未來可加入更多參考資料當作不確定性資料以修正模式。

分析類別型貝氏最大熵前，需要先分析每筆資料在整體研究區域中的相關性機率，並建立每個類別之間的距離和機率之相關性。而在地質研究中的距離，可從兩個距離面向定義，分別為「空間相關性」和「時間相關性」，兩者的相關性在地質學上都有一定的意義，因此本研究從兩個面向分別進行分析。

空間相關性主要是以鑽探所得之不同位置的岩性資料為確定性資料，依據兩點之間的空間距離對應出各類別可能出現的機率。空間距離在地質學上的意義即是，相似的岩性分佈在空間上會有相對應的距離，例如同樣岩性可能會距離較近，如果找到空間距離和岩性類別出現的相關性機率，則可用以推估地質空間相關的資料。

兩個位置點時間相關性則是以定年推估資料為不確定性資料，依據定年距離對應出各類別可能的機率。假設定年年份越近或空間距離越近，岩性相關性越高，則本研究以空間相關性為主，同時考慮時間相關性的影響，進行岩性推估模式之建立。

4.2.1 岩性空間相關性分析

首先是空間相關性的部分，以四十九口井的岩心資料進行分析，初步考慮水平和垂直方向之非等向性，分別以水平 XY 方向和垂直 Z 方向進行機率密度函數推估，分別繪製不同方向之機率密度函數圖。

最後再去除各方向的非等向性，計算出三大岩性類別間的機率密度函數，統合推估出整體臺北盆地三類岩性(泥、砂、礫石)之機率密度函數圖(如圖 5)，而圖 5 之結果呈現的變化趨勢，可大致看出隨距離改變之機率變化，接著應用圖 5 之空間相關性結果，依據兩點之間的距離計算出各類別可能的機率，以進行後續岩性推估之建模。

4.2.2 岩性時間相關性分析

岩性時間相關性的分析，應用已有的六口觀測井定年資料並參考地質相關專業知識(中央地質調查所，2014；鄧屬予 *et al.*，民國 88 年 1 月)，於 3.2.2 章節的定年資料前處理後，可推估得到整個台北盆地的定年分佈，此定年分佈雖不是整體確定的真實情況，但可藉由其他的研究資料，例如過去的地質勘查或台北盆地地層沈積演進的過程等資料，對此推估結果做局部修正，提升定年資料前處理分析的準確性，以供後續建模參考。

分析臺北盆地整體的定年資料分布情形後，以此定年分佈資料為時間基底，計算各位置間的時間距離，並應用定年資料推估各定年時間距離下，台北盆地地質岩性各類別間的時間相關性(如圖 6)，並應用此時間相關性推估各點可能的類別機率，最後將此推估結果取為不確定性資料，應用於後續整體岩性推估建模。

4.3 三維岩性推估整合模式建置

最後應用前置分析的空間相關性和時間相關性推估結果，結合所有觀測井的空間岩性類別資料(hard data)，以及其他地質研究相關的不確定性資料(soft data)，包含定年資料、虛擬井資料和 EOF 研究分析等資料，進行整合性的三維岩性類別推估，並且參照(Dudik *et al.*, 2004)之理論，針對其中限制式所含之不確定性進行模式修改。

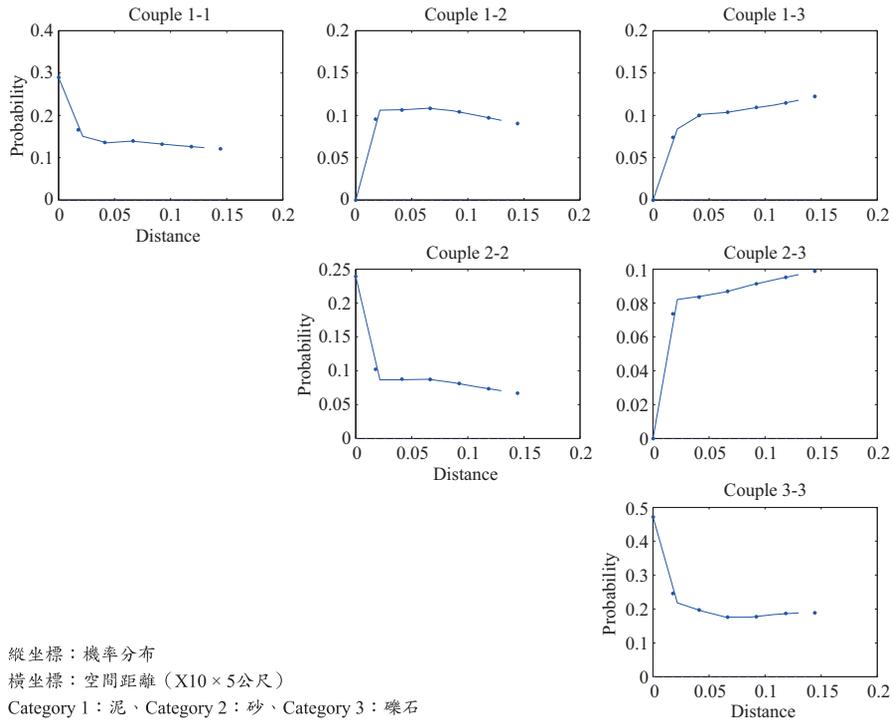


圖 5 臺北盆地岩性之空間相關性推估結果

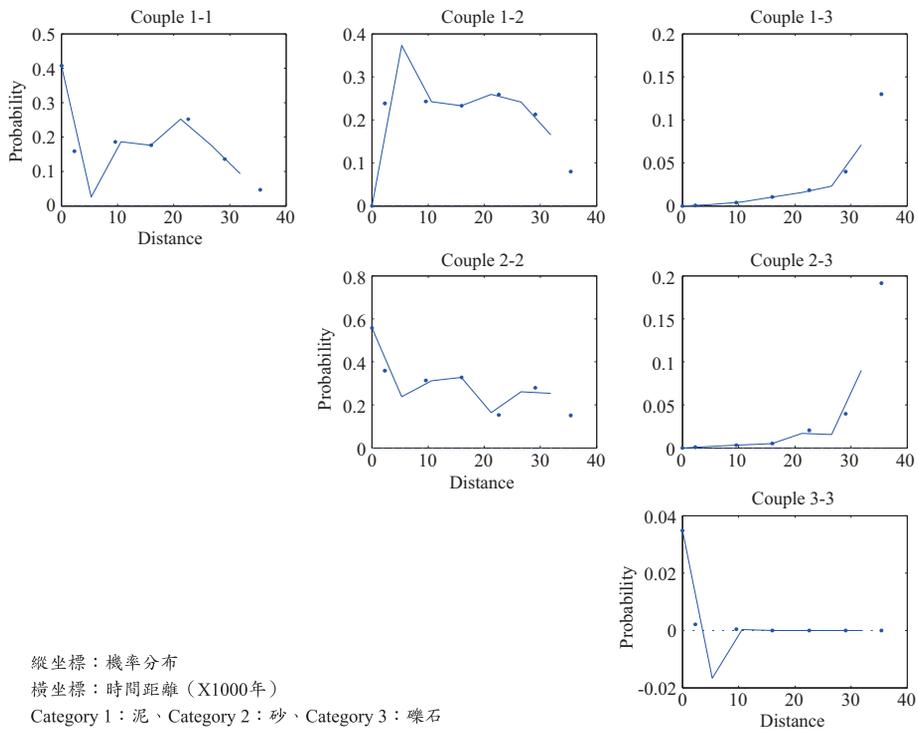


圖 6 臺北盆地岩性之時間相關性推估結果

4.3.1 確定和不確定性資料輸入之推估參數設定

本研究應用貝氏最大熵法等理論為基礎進行建模，參數設定如下表 5，利用前述確定性資料(hard data)和多項不確定性資料(soft data)相關性之結果，計算其距離差距對應類別的條件機率加總運算，對未知點進行類別型資料的整合推估，同理也可針對特定座標點進行推估。

表 5 本研究基本參數定義

x	已知的任一觀測井其地質岩性類別 (可為類別 1、2、3)
y	同時出現於其他位置的岩性類別 (可為類別 1、2、3)
(x_N, y_N)	樣本已知的 N 組資料 (任意兩岩性類別組合)
$p(y x)$	在已知岩性 x 的條件下，同時出現岩性 y 的條件機率
$f(x,y)$	兩兩岩性類別依照不同距離，切出不同的特徵值
λ	不同距離之特徵值 feature 對應的權重
β	模式限制式的誤差範圍

如式 11 所列，其中 $p_2(y|x)$ 即為欲求之條件機率密度函數，需用不同距離的特徵值 feature 分開，進行類別型資料的機率計算。針對不同距離特徵值 feature 其權重 λ ，分別進行迭代演算法更新 λ 並以此 λ 更新 $p_2(y|x)$ ，直到每個 λ 都收斂後，即可得到機率分佈 $p_2(y|x)$ 收斂的結果。

在模式收斂過程中所使用的限制式條件，參照章節 2.5 最大熵推估正規化 Regularization 調整限制式誤差範圍，式 12 為限制式調整放鬆之原則。依照樣本岩性資料，應用自助抽樣法 Bootstrap Method 進行抽樣，總抽樣次數 M 次介於 200 次~1000 次之間，使抽樣樣本組數量足夠，統計量趨於穩定即可。

抽樣完畢後可以得到近似於原樣本的分佈，因自助抽樣法 Bootstrap Method 的理論是從原樣本中隨機抽取後放回，如此重複抽取 N 個後，即得到一組隨機抽樣樣本，如需要 M 組樣本則重複上述步驟 M 次。應用在本研究即是，當前的樣本數量有 N 個岩心資料，每次從 N 筆資料中

抽取一個資料再放回，所以每次抽到每筆資料的機率都是一樣的 $1/N$ ，如此重複抽取 N 個即完成一組隨機抽樣樣本，如需要 M 組抽樣樣本，則重複執行抽樣 M 次每次抽 N 個。

由於自助抽樣法理論設定為抽取後放回，使每筆資料抽到的機率一樣，且理論允許重複抽到同一個資料，藉此仿照實際採樣情況的隨機打點，可能會因為採樣地點不同或重複相似而略有不同，使得原始觀測資料之理論值可能會有不同。

抽樣完畢後的 M 組隨機抽樣樣本，都可以計算出和原採樣資料相同的統計量，例如平均值和變異函數，進而可以得到 M 個實地採樣時可能產生的樣本平均值。M 個平均值再經由同樣方式隨機抽樣計算，可以得到本研究區域的採樣樣本，其存在的可能平均值分佈情形和機率，同時也可以算出樣本平均值在研究區域中的誤差範圍。

本研究藉此抽樣結果，建立研究區域中各類別間機率的誤差範圍，取其機率均值的 95% 信賴區間，以此區間為機率誤差範圍 β_i (機率均值的上下界)。同時建立每個類別機率，對應其距離差異之下，可能的機率誤差範圍表 Beta table (如表 6)，將此應用於後續模式推估改良，而本研究兩兩類別相關之 Beta table (如圖 7)。

表 6 空間距離對應類別機率及其誤差範圍 β 上下界 - 以泥對應泥為例(類別 1 對應類別 1)

空間距離	Beta 上界	平均值	Beta 下界
0.00000	0.28990	0.28969	0.28937
0.02166	0.15067	0.15041	0.15004
0.04333	0.13516	0.13500	0.13461
0.06499	0.13964	0.13944	0.13910
0.08665	0.13369	0.13353	0.13317
0.10831	0.12831	0.12821	0.12782
0.12998	0.12369	0.12359	0.12324

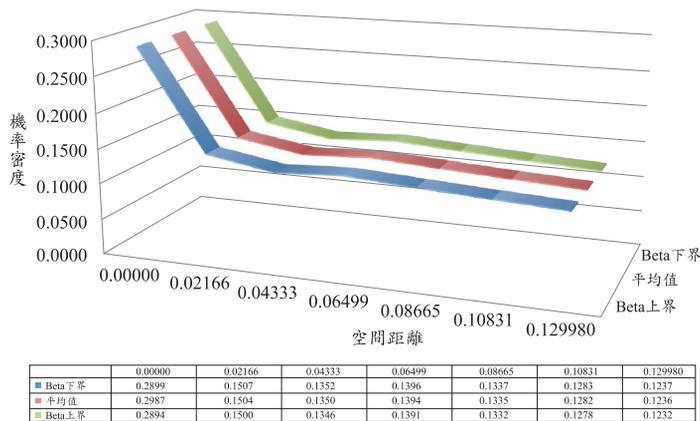


圖 7 空間距離對應類別機率及其誤差範圍 β 上下界 - 以泥對應泥[Category1-1]

4.3.2 整合推估模式建置

本研究應用貝氏最大熵法等理論為基礎進行建模後，利用空間類別相關性之結果(圖 8)，可以得知兩兩類別之間，距離和機率的相關性。因此欲推估某個未知岩性類別的點時，可先計算該點和周遭已有的觀測資料之距離，並計算每個點在如此距離之下，每筆資料的岩性機率對該點的影響各是多少。

分別將每筆資料對該點影響的機率，加總後可得各類別的機率，即可得知該未知點最有可能是哪個岩性類別及各別的機率，最後可對未知點進行類別型資料的推估，同理也可針對特定座標點進行推估。

接著依據時間相關性的結果(圖 9)，分析各時間距離下的臺北盆地岩性分布，推估方法和空間距離的步驟相同，只是距離長度由前述的實際空間距離，更改為時間上的年份差距，本研究將其定為時間距離，並加以計算時間和機率之相關性。將時間相關性所推出的各類別機率結果，取為不確定性資料，最後應用時間相關性的推估結果(soft data)，結合所有井的空間類別資料(hard data)，以及其他地質研究的不確定性資料，做整合性的三維岩性推估，則可對任一空間點做岩性類別分析，並可以得到每個空間位置點各個類別的機率。

由於過去研究中的類別型最大熵法推估模

式，僅使用最基本的迭代演算方法(GIS 和 IIS)，因此本研究多加入了限制式不確定性的理論(Dudik *et al.*, 2004)。在每次兩點之間計算類別機率時，都需使用迭代演算法來收斂參數，但在收斂的過程中，不將限制式當作百分之百正確的限制條件，依據 2.5 的理論方法對限制式範圍做修正。而在限制式的收斂範圍加上如 4.3.1 所分析的 Beta table，藉此不確定性的上下界能夠放寬收斂的限制範圍，期望對本研究的準確率能有所提升。

由於整個台北盆地面積大，但可用的推估點有限，因此同時應用交叉驗證(cross validation)確認本研究方法是否能夠有效推估出實際的地質參數。將每個地質井的資料分別獨立後，運用其他位置點對此點進行推估，並將推估結果和實際岩性的類別做比對，計算各座標位置點推估之準確率。本研究將三維空間進行均勻劃分後，推估出整個台北盆地各點的可能岩性，並比對各岩性類別在空間中的大致分布情形。

五、結果與討論

本研究應用貝氏最大熵法等理論為基礎進行建模後，利用類別相關性之結果，可對未知點進行類別型資料的推估，同理也可針對特定座標點進行推估。因此，先對三維空間進行均勻劃分後，推估出整個台北盆地各點的可能岩性，可以

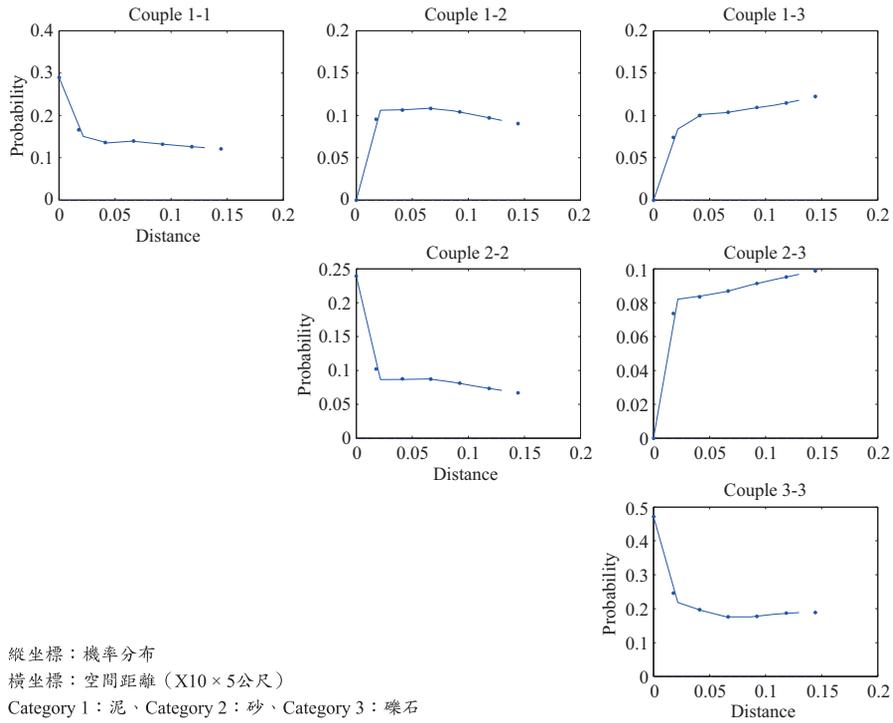


圖 8 臺北盆地岩性之空間相關性推估結果

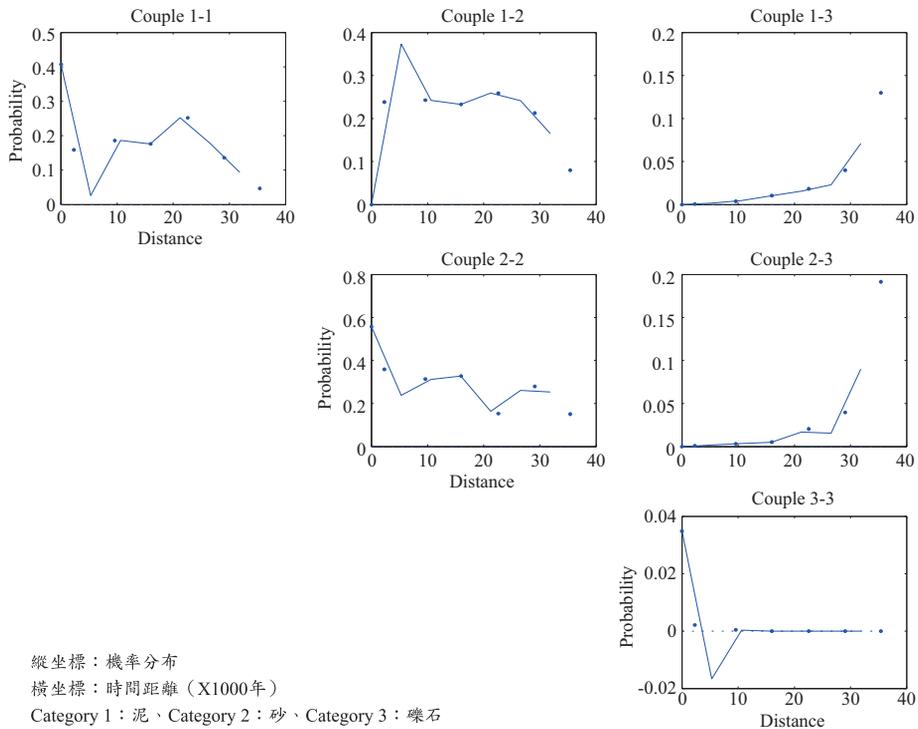


圖 9 臺北盆地岩性之時間相關性推估結果



圖 10 SC-1 站點位址

表 7 岩性推估結果和原始資料比對表

本研究	原始岩性資料			
	類別	泥(1)	砂(2)	礫石(3)
岩性 推估 結果	泥(1)	49	9	34
	沙(2)	4	0	2
	礫石(3)	23	0	179

看出各類別在空間中的大致分布情形，且各推估點可以確實對應到實際地質剖面的岩性劃分，以確認推估結果之準確率。

由於整個台北盆地面積不小，但可用的推估點資料有限，因此本研究同時應用交叉驗證(cross validation)確認本研究方法是否能夠有效推估出實際的地質參數。將每個地質井的資料分別獨立後，運用其他位置點對此點進行推估，並將推估結果和實際岩性的類別做比對，計算各座標位置點推估之準確率，其結果顯示座標點周圍的資料夠豐富，則準確率可以高達 60%~80%。

驗證推估結果的準確率，分別將每口井資料抽掉後，以其他資料對已有的井資料進行推估，再和原始資料做交叉驗證 (cross validation)，例如以 SC-1 站點資料為例(如圖 10 紅圈處)，原始資料共計 300 筆資料，其中有 228 筆資料推估正確，準確率為 76%，推估結果的岩性類別和原始資料的比對情形如表 7 所示。

針對每口井進行交叉驗證後，可將推估結果

繪製於過去地質研究的剖面圖之上(如圖 8)，比對本研究的岩性推估結果，和過去地質學家的推測研究成果是否相似。

由交叉驗證的結果(表 8)可看出，本研究方法可有效應用周圍有限的類別型資料，對三維座標點進行推估，藉由推估整個三維空間的類別資料分布結構，進而提供其他研究或都市規劃做參考，增加研究基本參考資料的豐富度。同時能夠比較多方考量各種不同的地質相關性資料，以多項 soft data 對於研究結果進行優化，以期未來能夠持續強化本模式，使本研究方法加以應用在各個領域，以解決許多研究資料不足的窘境，提供相關研究更多可參考的資訊。

六、致謝

感謝經濟部中央地質調查所計畫「地下水水文地質與補注模式研究-補注區劃設與資源量評估(2/4)」的補助及資料提供，使本研究得以順利進行。

七、參考文獻

1. 中央地質調查所(2014)，地下水水文地質與補注模式研究-補注區劃設與資源量評估計畫(2/4)。
2. 林軒田(Producer) (2013)，機器學習基石 (Machine Learning Foundations). Retrieved from <http://www.coursera.org/course/ntumlone>
3. 費立沅(2011)，臺北盆地的地質與防災 = *The geology and geohazard prevention for the Taipei basin*. 新北市：經濟部中央地質調查所。
4. 鄧屬予、袁彼得、陳培源、彭志雄、賴典章、費立沅 & 劉桓吉(民國 88 年 1 月)，臺北盆地堆積層的岩性地層。經濟部中央地質調查所特刊，第十一號，41~66。
5. Adam L. Berger, S. A. D. P., Vincent J. and Della Pietra (1996). *A Maximum Entropy Approach to Natural Language Processing*. Paper presented at the Association for Computational Linguistics.

表 8 各井資料點之交叉驗證結果

			交叉驗證之準確率結果(分別取用不同資料組合進行岩性推估)					
井編號	井名	總資料點(個)	空間資料	空間加定年資料	空間加虛擬井資料	三筆資料	四筆資料	加入 Beta Range
HCP-1	新莊一	540	0.4278	0.4444	0.7019	0.7037	0.7000	0.6574
KT-1	關渡一	515	0.4796	0.4738	0.5806	0.5825	0.5476	0.4233
KT-2	關渡二	231	0.4805	0.4719	0.4848	0.4719	0.5455	0.2900
KT-3	關渡三	331	0.5438	0.5468	0.5227	0.5257	0.5529	0.3535
LC-1	蘆洲一	298	0.4799	0.4933	0.4933	0.4933	0.4933	0.4161
PC-1	板橋一	84	0.7262	0.6786	0.6310	0.6667	0.9048	0.6429
PC-2	板橋二	290	0.3241	0.3241	0.6310	0.6310	0.6483	0.5897
SC-2	三重二	285	0.6105	0.6175	0.7614	0.7579	0.7368	0.6491
SCF-1	山腳一	60	0.3833	0.3667	0.3833	0.3667	0.4000	0.4000
SCF-3	山腳三	68	0.3235	0.3235	0.3676	0.3676	0.4118	0.5000
SCF-4	山腳四	47	0.2553	0.2553	0.2553	0.2766	0.2979	0.3617
SCF-5	山腳五	180	0.4944	0.5000	0.4944	0.5000	0.5222	0.2667
SCF-6	山腳六	400	0.4375	0.4500	0.6100	0.6125	0.6175	0.2800
SCF-7	山腳七	90	0.3556	0.3556	0.3556	0.3556	0.3556	0.4111
SCF-8	山腳八	173	0.2486	0.2486	0.2486	0.2486	0.2486	0.4046
SCF-9	山腳九	367	0.3569	0.3597	0.3488	0.3433	0.3433	0.5123
SCF-10	山腳十	128	0.5625	0.5625	0.5625	0.5625	0.5625	0.2188
SCF-11	山腳十一	507	0.3590	0.3590	0.3393	0.3393	0.3393	0.1913
SCF-12	山腳十二	125	0.5040	0.5040	0.5040	0.5040	0.5040	0.0000
SCF-13	山腳十三	500	0.5960	0.6220	0.5960	0.6220	0.6220	0.0080
SCF-14	山腳十四	198	0.6162	0.6313	0.6313	0.6364	0.6364	0.2929
SCF-15	山腳十五	77	0.7403	0.7662	0.7662	0.7662	0.7922	0.3766
SCF-16	山腳十六	159	0.6478	0.6478	0.6541	0.6541	0.6541	0.2956
SCF-17	山腳十七	244	0.6270	0.6189	0.6311	0.6311	0.6393	0.2705
SL-1	士林一	225	0.5867	0.5911	0.5867	0.5911	0.3644	0.4044
SS-1	松山一	98	0.6429	0.6429	0.6429	0.6327	0.9388	0.4898
SS-2	松山二	148	0.6757	0.6351	0.6554	0.6014	0.4932	0.4189
TA-1	大安一	214	0.4766	0.5000	0.5093	0.5421	0.6121	0.5888
TU-1	台大一	280	0.5250	0.5107	0.5786	0.5750	0.5714	0.4929
YH-1	永和一	173	0.4509	0.4335	0.8497	0.8497	0.9480	0.7168
RC-1	二重一	109	0.3394	0.3394	0.3303	0.3303	0.3303	0.0367
SL-2	士林二	108	0.5370	0.5185	0.5370	0.5185	0.5278	0.3981
CH-1	中和一	35	0.4286	0.4286	0.4571	0.4000	0.9143	0.3429
NH-1	內湖一	49	0.8367	0.7551	0.8367	0.7551	0.7755	0.7959
TM-1	天母一	38	0.1053	0.2368	0.1053	0.2368	0.5526	0.4737
YH-2	永和二	21	0.3810	0.3333	0.1905	0.1905	0.1905	0.1905
FJU-1	輔大一	64	0.4063	0.4219	0.5000	0.5000	0.5156	0.6094
SHL-1	樹林一	39	0.5128	0.4615	0.5641	0.5385	0.9744	0.7949
LC-2	蘆洲二	102	0.5098	0.5196	0.3627	0.3529	0.2647	0.5000
010103G1	江翠	148	0.1757	0.1757	0.1892	0.1892	0.9662	0.2838
010501G1	新莊	243	0.2840	0.2675	0.2881	0.2881	0.3045	0.2346
630401G1	新生公園	85	0.3529	0.4118	0.3529	0.4118	0.3412	0.2471
630501G1	二二八公園	248	0.3185	0.3185	0.2903	0.2903	0.3145	0.2661
630601G1	衛工處	154	0.4416	0.4351	0.4416	0.4351	0.2532	0.2338
630701G1	青年公園	228	0.4605	0.5044	0.4693	0.4693	0.4737	0.4518
		average	0.4673	0.4681	0.4954	0.4959	0.5489	0.3952
		max	0.8367	0.7662	0.8497	0.8497	0.9744	0.7959
		min	0.1053	0.1757	0.1053	0.1892	0.1905	

(三筆資料：包含空間、定年及虛擬井資料)

(四筆資料：包含空間、定年、虛擬井資料及 EOF 推估結果)

(加入 Beta Range：考量資料不確定性，放鬆限制式條件後的推估結果)

6. Allard, D., D'Or, D., & Froidevaux, R. (2011). An efficient maximum entropy approach for categorical variable prediction. *European Journal of Soil Science*, 62(3), 381-393.
7. Bogaert, P. (2002). Spatial prediction of categorical variables: the Bayesian maximum entropy approach. *Stochastic Environmental Research and Risk Assessment (SERRA)*, 16(6), 425-448. doi: 10.1007/s00477-002-0114-4
8. Bogaert, P., & D'Or, D. (2002). Estimating soil properties from thematic soil maps: the Bayesian maximum entropy approach. *Soil Science Society of America Journal*, 66(5), 1492-1500.
9. Bogaert, P., & Fasbender, D. (2008). Nonlinear Spatial Prediction with Non-Gaussian Data: A Maximum Entropy Viewpoint. *geoENV VI – Geostatistics for Environmental Applications*, 445–455.
10. Christakos, G. (1990). A Bayesian Maximum-Entropy View to the Spatial Estimation Problem. *Mathematical Geology*, 22(7), 763-777. doi: Doi 10.1007/Bf00890661
11. Christakos, G. (1991). Some Applications of the Bayesian, Maximum-Entropy Concept in Geostatistics. *Maximum Entropy and Bayesian Methods //*, 43, 215-229.
12. Christakos, G. (2000). *Modern Spatiotemporal Geostatistics*: Oxford University Press, USA.
13. Christakos, G., & Li, X. Y. (1998). Bayesian maximum entropy analysis and mapping: A farewell to kriging estimators? *Mathematical Geology*, 30(4), 435-462. doi: Doi 10.1023/A:1021748324917
14. Christakos, G., & Olea, R. A. (2005). New space-time perspectives on the propagation characteristics of the Black Death epidemic and its relation to bubonic plague. *Stochastic Environmental Research and Risk Assessment*, 19(5), 307-314.
15. Christakos, G., & Serre, M. L. (2000). BME analysis of spatiotemporal particulate matter distributions in North Carolina. *Atmospheric Environment*, 34(20), 3393-3406.
16. D'Or, D., Bogaert, P., & Christakos, G. (2001). Application of the BME approach to soil texture mapping. *Stochastic Environmental Research and Risk Assessment*, 15(1), 87-100.
17. D'Or, D., & Bogaert, P. (2004). Combining Categorical Information with the Bayesian Maximum Entropy Approach. *geoENV IV – Geostatistics for Environmental Applications*, 295-306.
18. Darroch, J. N., & Ratcliff, D. (1972). Generalized Iterative Scaling for Log-Linear Models. *Annals of Mathematical Statistics*, 43(5), 1470- &. doi: Doi 10.1214/Aoms/1177692379
19. Dudik, M., Phillips, S. J., & Schapire, R. E. (2004). Performance guarantees for regularized maximum entropy density estimation. *Learning Theory, Proceedings, 3120*, 472-486. doi: Doi 10.1007/978-3-540-27819-1_33
20. Goovaerts, P. (1997). *Geostatistics for natural resources evaluation*. New York: Oxford University Press.
21. Journel, A. G. (1989). *Fundamentals of geostatistics in five lessons*. Washington, D.C.: American Geophysical Union.
22. Journel, A. G., & Huijbregts, C. J. (1978). *Mining geostatistics*. London ; New York: Academic Press.
23. M.A. Wibrin, P. B. and D. Fasbender. (2006). Combining categorical and continuous spatial information within the Bayesian maximum entropy paradigm. *Stoch Environ Res Risk Assess*, 20, 423-433. doi: 10.1007/s00477-006-0035-8
24. P. Bogaert and M.-A. W. (2004). Combining categorical and continuous information using Bayesian Maximum Entropy. Université catholique de Louvain: dept. of Environmental Sciences and Land Use Planning –

Environmetry and Geomatics.

York: Springer-Verlag.

25. Stephen Della Pietra, V. D. P., John Lafferty.
(1997). *Inducing Features of Random Fields*.
Paper presented at the IEEE TRANSACTIONS
ON PATTERN ANALYSIS AND MACHINE
INTELLIGENCE.
26. Venables, W. N., & Ripley, B. D. (1994).
Modern applied statistics with S-Plus. New

收稿日期：民國 105 年 5 月 27 日

修正日期：民國 105 年 7 月 25 日

接受日期：民國 105 年 7 月 26 日