巢狀超矩形學習模式於水資源系統之研究

A Study of the Nested Hyper-Rectangles Learning Model for Water Resources Systems

臺灣大學農工所博士班研究生

陳　莉

Li Chen

臺灣大學農工所教授

張斐章

Fi-John Chang

摘　要

預測及分類爲水文學者重要的工作。巢狀超矩形學習模式是一種「以範例爲基礎的學習」（Exemplar-based learning）模式，最主要的觀念是將過去發生的許多事件以一個事件爲一個點（Point）的型式貯存在歐基里得（Euclidean）n維空間(E^n)中，將來如果有一新的事件（一個新的點）加入模式的時候，就可以計算出與原來樣本空間中最接近的點，進而達成分類或預測的效果。

此模式因可經由新加入之範例而動態調整「距離計量」中的各項參數以回饋系統，故其準確度隨著所訓練樣本的增加而愈正確，即本模式可具有學習的能力。

爲說明此一模式，本研究先設定一簡單的數學函數並以此模式預估其理論值。之後，並實際應用於河川流量的分級預報，日雨量記錄的補遺，及年平均流量之延伸等工作，結果皆顯示其優越的預報及分類能力，值得進一步的研究及推展。

ABSTRACT

Prediction and categorization are important tasks of hydrologists. The nested hyper-rectangles learning model which is an examplar-based learning model is studied and applied to above tasks. The main idea of the model is "seeding" history data in Euclidean n-space, E^n, as examplars, then comparing new examples (data) to those seeding points, and finding the most similar example in memory.

A dynamic adjustment for model's parameters is proposed with the"distance metric" which is used to determine the similarity, so the simulated system can be represented or predicted more and more accurate through the feedack of added examples. That is the model has ability to learn.

In order to show the general characteristics of the model performances, a simple mathematical function is simulated by the model. It is then applied to three different hydrological systems.

They are: (1) forecasting streamflow categorizaton; (2) estimating the missing record of daily precipitation; and (3) extending the annual streamflow. The results demonstrate the power of hyper-rectangle learning model, and the model can be a very useful tool on hydrological system.

一、前　　言

在許多不同的領域中，預測工作常是非常重要的一環：舉例而言，預測股價(Value of stocks）或是對醫院病人情況的預估。而分類亦是一門重要的課題，如生物學上的分類。在水資源系統中，分類與預測則是水文工作者，最常面臨的問題及挑戰。

人類智慧的發展常是經由學習及經驗的累積以斷定事物之分類或對未來作預測。傳統分析方式爲藉由統計分析，如廻歸、主成份分析、聚類分類（Clusting Analysis），針對資料之結構歸納成一些函數或法則公式等，此一方式雖可能具有去蕪存菁的功用，然而更多的時候可能因公式化之故而忽略一些特例所具有的重要信息，或因函數的型態與實際資料結構並不十分吻合而造成失眞的情況。

本研究敍述一種新的「依範例學習」（Exemplar-based learning）方法，稱爲「巢狀推廣範例理論」(Nested Generalized Exemplar Theory, NGE) 藉此從事預測（Prediction）或分類（Categorization）工作，此一方法可改善統計分析中勉强「套配」及不具學習功能的缺點。本文首先將介紹此一理論，並以該理論之演算法推估一數學函數值，以了解其精確度，最後則用以推估或研判實際水資源系統的變化情形。

二、NGE 學習模式之重要特質

「依範例學習」之理論簡稱爲 EACH（Exemplar-Aided Constructor of Hyper-rectangles），其基本假設爲以相同領域中先前發生過的例子（Examples）爲基礎，作爲預測或分類之依據，經由過去的例子持續的增加，最初係以點（Point）貯存在歐氏 n 維空間（E^n）中，而 n 代表在一個例子中變數（Variables）或特徵（Features)的數目，NGE 理論又將上述各點形成超矩形(Hyper-rectangles)結構 (Medin and Schaffer, 1978)，當例子的個數增多時，則一些例外（Exceptions）可能在超矩形中產生，即所謂的「洞」（Holes），這些洞裏面又可能有其他的洞，結果形成了「巢狀」（Nested）的超矩形結構。

EACH 主要的概念爲新的樣本與先前發生過的例子做比對，至於如何決定那一個例子和新的樣本最爲近似，則係利用「類似度計量」(Similarity metric），也可稱爲「距離計量」（Distance metric），因爲它量測了各例子與新樣本間的距離。我們使用 "Exemplar" 表示在記憶中已貯存的「範例」，而以 "Example" 代表新加入系統中的「樣本」。每一個範例 (Examplar) 都存著輸出變數的值，並用於預測或分類。簡而言之，EACH 新樣本的預測或分類變數值將與最接近的範例中存在的值設爲相同，而所謂學習，是發生在系統經由預測而獲得一些回饋 (Feedback) 之後，使系統能依最新的信息，作必要的修正或適度的改善，使其更能與實際狀況相吻合，詳細的步驟將在後幾節中說明。

2-1. 知識的表現

NGE 學習模式並非產生規則 (Rules)，亦非一般的決策樹 (Decision trees)，取而代之的，它建立了一個佈滿範例 (Exemplars) 的記憶空間，其中有些爲繁衍 (Generalizations) 的超矩形，有些則從系統經驗中取得的獨立樣本 (Examples)，這種結構化的範例記憶 (Structured exemplar memory) 是學習程序的主要輸出。因爲EACH爲軸平行矩形(Axis-Parallel rectangle)，所以只需要記憶對角線上的兩點就可定義一個超矩形，其內部的一些範例點則可省略不記，這種資料結構節省了不少記憶。

繁衍範例可採用超矩形(Hyper-rectangles)的型式；換言之，即爲在 E^n 中的矩形實體(Rectangular-solids)，此外，EACH 考量各範例之重要性並引進權重的觀念以修正原始的「距離計

量」（Distance metric），用以測測量各範例之間的距離，同時以其預測的結果來回饋系統，此「距離計量」在學習程序裏扮演另一項很重要的角色。

EACH 演算法的另一個特點，就是一超矩形中可能形成巢狀結構，而在其內部的超矩形（可視爲大超矩形裏面的洞）即因例外（Exceptions）的情況所造成，這長久以來一直是機器學習演算法中一個深具意義的問題，從概念上來解釋即：

等級（Classes）中包括次等級（Sub-classes），此巢狀的超矩形提供了自然而簡明的表示法。

Aha 與 Kibler（1989）將此一學習模式和其他不同的「以範例學習」的方法做了各項比較，並顯示其優越的特性。

2-2. 學習策略

依範例爲基礎的學習策略是屬於「增加」（Incremental）學習的方式，簡單的說，就是指系統依據每一範例來修正內部的架構，有一項重要的實驗顯示 EACH 對於輸入的次序（Order）相當敏感（Michalski et al., 1983）。另外，有些學者提出「非增加」的學習方式，例如形成橢圓體（Ellipsoids）而非超矩形（Everitt, 1980），但該方式受限於無法產生巢狀結構，也無法對「距離計量」加以修正。

2-3. 離散與連續變數

EACH 需要掌握的變數型式可爲任何數值，從2元到連續變數，其處理所有變數的方法大致相同，此外，EACH 預測的變數也可爲離散或連續的，如爲連續變數，系統需設一「誤差寬容參數」（Error tolerance parameter）如設定爲 0.05，當一範例的代表值爲 5.0 的話，則在範圍〔4.95，5.05〕之間的任何數值均被視爲「相配」（Match）。如新的樣本其預測變數值爲 5.03，則此樣本與該範例「相配」，（EACH 不必在記憶中貯存這一個新的點，上述方法乃將連續的預測值化爲以離散的逼近結果。

三、巢狀推廣範例學習之演算法

(The nested generalized exemplar learning algorithm)

3-1. 初步工作

爲了做預測，EACH 必須有歷史性的範例（Exemplars）爲基本，在記憶體中首先隨機的選取例子（其最小的個數爲1）來「播種」（Seeding），此種「播種」方法乃簡單的將每一例子貯存在記憶中，一個例子即爲一個特徵（Features）的向量，而每一特徵可能包括許多特值，範圍從2元（Binary）特徵到無限（實數值），對於連續的變數，系統可定出「誤差寬容參數」，以指示兩個值之間的接近程度是否達到被視爲「相配」（Match）的標準。

3-2. 預 測

每一新加入的例子依照下述的「相配」程序，然後以最佳的「相配」來做預測：系統可決定新的樣本與那一個範例最接近，即爲「相配」的範例，並將二者當成相同的類別（類別可爲二元，離散或連續的）。

「相配」程序可說是本演算法的決策中心，此程序使用「距離計量」來測定新的資料點（An example）與一個範例（Exemplar，在 E^n 中的一點或一個超矩形）之間的距離（或相似度），設新的樣本爲E，而已存在的超矩形爲H。

系統藉著量測兩個物體間的歐氏距離來計算E和H間「相配」的分數，其最簡單的方程式便是假設H爲一點，距離即可由一般計算每一特徵空間（Feature space）在幾何上的距離，然後再加入一些其他的項目（各項參數權重）如後所述：

$$D_{EH} = Wh \left[\sum_{i=1}^{m} \left(Wi \frac{Efi - Hfi}{\max i - \min i} \right)^2 \right]^{0.5}$$

式中 Wh 爲範例H的權重，Wi 爲特徵 i 的權重，Efi 爲在樣本E中第 i 個特徵的值，Hfi 爲在範例H中第 i 個特徵的值，min i 和 max i 是某個特徵 i 所有出現的值中最小的和最大的值，而m爲E中可辨識特徵的個數。如果H代表一超矩形，則距離量測公式必須做些微的修正，即改爲計算最近的角（Corner）、邊（Edge）或邊的面（Face of edge）。

所謂最佳的「相配」是指有最短的距離，上述公式將距離沿著每一維度除以（maxi-mini）是爲了使其標準化在〔0,1〕區間內，對二元（Binary）特徵來說，距離計算就簡單許多：若特徵爲相同的，距離爲0，否則即爲1，同樣的計算法也可應用於一些離散或非數值的特徵。

在「距離計量」公式中有兩個權重參數，即

Wh 及 Wi，Wh 可量度利用範例H並使預測正確的機會，即爲在用到某一範例H的全部次數與其中爲正確預測次數的比例，換言之， Wh 說明了在每一個範例中有關「可靠性」的訊息，如一個超矩形H被使用過很多次，可是却經常做出錯誤的預測的話，權重 Wh 將會變得非常大，如此H將趨於不被選爲最接近的「相配」，即漸漸被淘汰，另外，如果H爲一擾動（Noisy）點，則它將由於Wh 的漸增亦被忽略，理論上 Wh 最小的值爲 1 （即預測完全正確）。

另一個權重 Wi 代表第 i 項特徵所佔的權重，權重之調整反映出並非所有的特徵對於分類的決定都具有相等重要的影響，應用 EACH 演算法顯示若這些權重的修正很緩慢則系統有較佳的表現，相反地，急速的調整將導致系統預測準確度的振動（Oscilations）。

3-3. 回 饋

當 EACH 對其預測之結果與眞實結果做比較時，可產生學習的效果，即如果系統做了正確的預測，超矩形H就被認爲在 E^n 中足以包括新的點E，此時便不對各項參數做任何修正。

如果系統做出錯誤的預測，則使用「第二機會」（Second chance）經驗法則以避免產生過多不必要的記憶（Salzberg, 1985、1986、1988），所以在產生一個新的例子之前， EACH 首先檢測在記憶中第二最佳的「相配」，若是第二最佳的「相配」也做了錯誤的預測，那麼系統就將新的樣本E視爲記憶中的一點，這個新的例子可能存在於範例H的內部，這種情形就好像在H中有一個例外（Exception）或「洞」（Hole）進而由E演變成在H中的一個超矩形，亦即巢狀組織。

當我們發現系統產生錯誤的預測時， EACE 將調整特徵 fi 的權重 Wi，乃藉由一個非常簡單的迴圈（Loop）來完成：先對每一 fi 設定各評斷標準值，若 Efi 與 Hfi「相配」（相差小於所設標準值），則權重 Wi 以 Wi=Wi(1+△f)漸增，而 △f爲整個特徵的調整率（一典型的 △f值設爲0.05），如此增加的權重將使兩個目標體距離拉遠，理由是既然因爲 EACH 做錯誤的預測是由於E和H太相近，所以應將它們分開一些，相反地，若錯誤是因 Ffi 與 Hfi 不十分「相配」，則 Wi 以 Wi=Wi (1−△f) 減少。

3-4. 模式演算法的流程

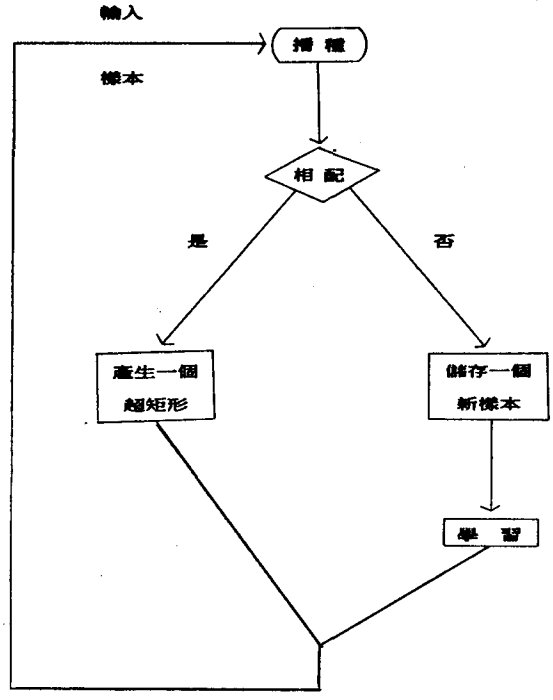

圖1 EACH 演算法之流程圖

圖1爲EACH 演算法之流程圖，其詳細內容如下：

1.播種（Seeding）：EACH 必須具有歷史性的範例，藉以爲預測的基礎。

2.相配（Matching）：使用「距離計量」。系統可預測這個新加入的樣本E將落入原來範例中那一最接近的羣，並計算是否在誤差容許範圍之內？若是則建立一個超矩形，若非則單獨記憶此一新樣本點。

3.回饋（Feedback）：

(1)如果系統做了正確的預測，則以H和E爲對角線形成一個新的超矩形。

(2)如果系統做了不正確的預測，則在記憶中將這個新的樣本E視爲一獨立的點而加以貯存。

4.學習（Learning）：如果發現系統預測錯誤時，EACH 可自動調整每一 fi 的權重 Wi：

(1)如果某項Efi太接近 Hfi，則設定 Wi=Wi (1+△f)

(2)如果某項Efi不接近 Hfi，則設定 Wi=Wi (1−△f)

3-5. 圖例說明

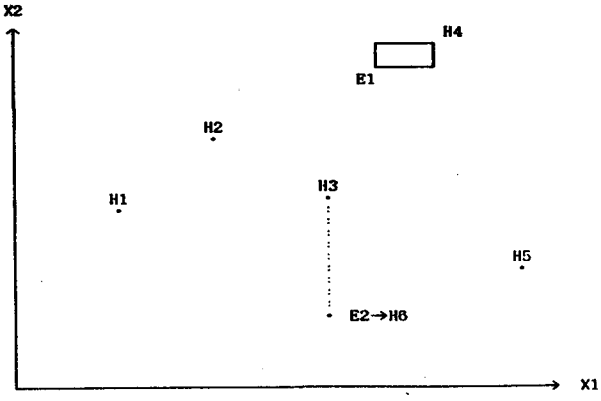

圖2 EACH 演算法於兩個變數情況下播種與相配之示意圖。

圖2爲 i=2 的情況，其演算的步驟與對應的幾何義意：

(0)首先在二維空間（平面）中「播種」，其個數假設爲5，即範例爲 H1 ，H2……H5 。

(1)現在加入一個新的樣本點 E1，可由「距離計量」分別計算 E1 與 H1……H5 的各個接近程度，結果若 E1 與 H4 最「相配」，且其預測值在誤差容許範圍之內，則以 E1H4 爲對角線形成一超矩形。

(2)模式再加入一個新的樣本點 E2，再由「距離計量」分別算出 E2 與 H1……H5 的各個接近程度，結果若 E2 與 H3 最「相配」，但是超出所設定的誤差範圍，則將 E2 視爲系統中一個新的範例 H6 。

系統不斷加入新的樣本點，以上述(1)或(2)的情況，如果在誤差容許範圍之內則形成超矩形，反之則存入一個新的點，重覆此種過程，使系統持續成長、學習，而其表現也愈來愈好。

四、前人研究成果：醫學診斷或生物學的分類 (Salzberg, 1988)

4-1. 癌症復發的預測

樣本爲 273 位曾動手術去除腫瘤的病人，以 EACH 預測他們在未來五年內是否會再發病，選擇9個變數（特徵）其中3個爲二元，其它6個爲實數或離散的變數。爲求適當的比較（測試方法與 Michalski et al., 1986 相同）先將樣本分成訓練集合與測試集合，隨機取70%的樣本來訓練後，針對剩餘 30%的樣本作預測，結果正確率爲 78%，如果以一位醫生來作診斷可得 64% 的正確度（Michalski et al., 1986），而隨機猜測應爲50%正確率。

4-2. 鳶尾花的分類

150個鳶尾花的樣本，（使用資料： Fisher, 1936），欲將其分成三類，4個變數爲：萼長、萼寬、瓣長、瓣寬，以樣本中的5個當作已知情況，再對其他 145個作預測，結果正確率爲135／145=93%。

4-3. 心臟病患的存活率

研究對象爲 119位病患，以統計迴歸相互比較。本預測試驗將病患分成兩類：是否可存活超過一年，資料包括9個變數，結果非線性迴歸預測正確率爲60%，以 EACH 做預測：若部分樣本事先未知（即將訓練集合與預測集合分開），則正確率爲57%，但若使用全部樣本（與迴歸相同）則可達100%的正確率。

五、應用實例

5-1. 例 題

已知一函數 $Y=X_1^2+X_2^2$，其中 $0\leq X_1\leq 1$，$0\leq X_2\leq 1$，首先隨機選取n組(X_1，X_2，Y)依序加入巢狀超矩形模式中，因其特徵變數只有 X_1，X_2 兩個，所以「距離計量」公式中的 i=2，而「誤差寬容」(Error tolerance) 設爲0.2，即預測的Y值與眞正的Y值相差在 0.2之內均可視爲同一超矩形，經過整個演算程序後，上述n個點逐漸形成數個不同的超矩形，而完成訓練 (Training) 工作；接着進行模式預測能力的試驗，例如任意選取100組 (X_1，X_2)，且不由函數求Y值，即把已知之函數放入黑盒 (Black box) 中當作未知的情況，而把這 100個點逐一加入已完成訓練的巢狀超矩形模式裏，此時對模式中的各項參數均不做任何調整，故可用於驗證模式預測的正確度；判定時誤差範圍在 0.2以內視爲預測正確。

爲明瞭訓練過程中不同的超矩形個數m對預測正確度的影響，本研究分別測試m= 5 、10、15、20及25，五種情況，在每個固定的m值下，每次進行對100組 (X_1，X_2) 的預測，並以模式執行 100次的平均值視爲超矩形個數爲 m 時的預測正確度 CP (Correct prediction %)，結果m=5，

CP＝56；m＝10， CP＝78%；m＝15，CP＝86%；m＝20；CP＝91%；m＝25，CP＝95%，如表1與圖3，由此可知當訓練的超矩形個數增多時，預測正確度隨之提高且變異性逐漸降低，當超矩形個數爲15時即有不錯的表現，而當其個數增至25時，已達95%的高正確度，故本模式的預測能力相當有潛力。

表1　超矩形個數與預測正確度之關形

超矩形個數	預測正確度	
	平均值(%)	標準偏差(%)
5	55.94	15.02
10	78.00	11.29
15	85.72	8.59
20	90.78	5.58
25	94.45	2.72

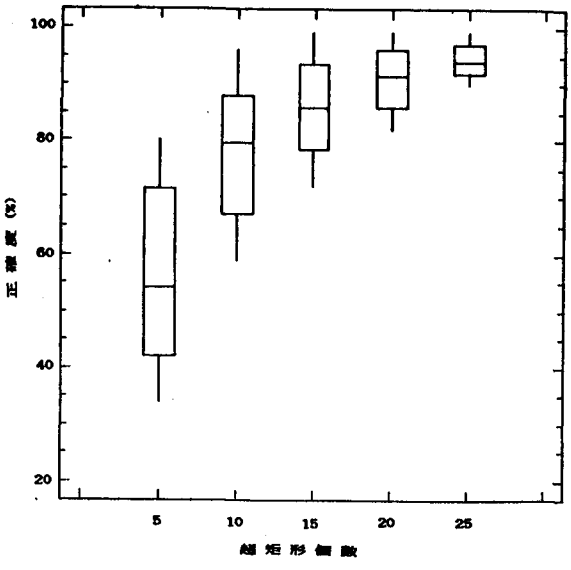

圖3　超矩形個數與預測正確度之多重鬚盒圖
(Box-and Whisker Plot)

5-2.　推估河川流量

以石門水庫上游高義站四月份流量爲分級預報，首先選取三個可能影響該流量的因子（特徵變數 i＝3），即三月份平均日流量（X_1），三月份最大日流量（X_2）及三月份平均日雨量（X_3）等三個預報因子，而四月份平均日流量（Y）分類爲充足、正常、或不足三種等級。採用資料年限自民國46年至民國74年，共計29年，並以75、76兩年進行預報驗證工作。

四月份日平均流量三個不同等級的區分如下：

y_1：流量不足　9CMS≧Y

y_2：流量正常　16CMS≧Y＞ 9CMS

y_3：流量充足　50CMS≧Y＞16CMS

將29年的資料（29個點）輸入巢狀超矩形學習模式，經過訓練後形成12個超矩形，先對此29年歷史記錄作檢驗，即由 X_1、X_2 及 X_3 來判別Y的分級爲 y_1、y_2 或 y_3，可達 100%完全正確率。若在沒有任何訊息可資提供的情況，以隨機猜測於三個區間的可能性，其答對率應爲33%，若以多變數常態假設，並以貝氏推論推判，乃爲51.7%（張斐章，徐國麟，1990），若以模糊集理論配合檢定過之從屬函數，判別率可達62%（張斐章，徐國麟，1990）。再以75、76兩年作實際預報，結果判定皆與實際記錄相吻合。

5-3.　雨量資料之補遺

以淡水河流域大漢溪上游地區雨量資料之補遺爲例。假設秀巒站（位於白石溪）的部分日雨量資料有缺失，乃以其附近周圍的4個雨量站：鎮西堡（位於泰崗溪）、白石（位於白石溪）、鞍部（位於玉峰溪）及玉峰（玉峰溪）的完整日雨量記錄來補足秀巒站的資料，示於圖4及表2。首先選取鎮西堡、白石、鞍部及玉峰四站的記錄年限爲民國44年至78年，共35年，以此所有日雨量資料爲四個特徵

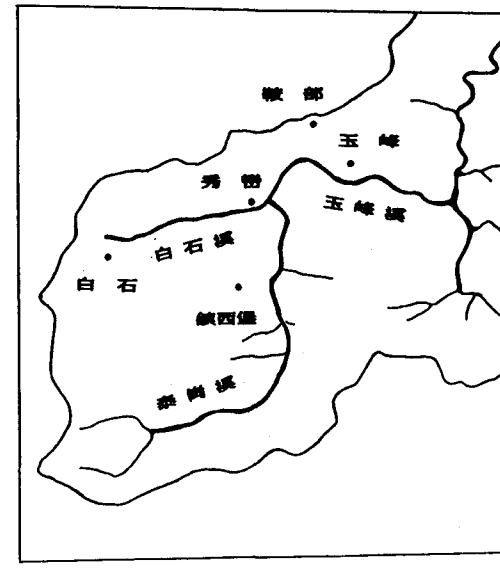

圖4　石門水庫上游雨量站位置圖

表 2　淡 水 河 雨 量 站 記 錄 表

站號	站名 站址	支流	東經 北緯	標高 (公尺)	機關儀器	起用年 詢止年	記錄年份
P001	鎮西堡 新竹縣尖石鄉秀巒村鎮西堡	泰崗溪	121 18 00 24 34 00	1630.00	省石門水庫 自記	43 7	43-
P003	白石 新竹縣尖石鄉秀巒村白石	白石溪	121 13 00 24 33 00	1630.00	省石門水庫 自記	29 9	29-30 43-
P005	鞍部 新竹縣尖石鄉玉峰村宅老	玉峰溪	121 16 0 24 40 0	1450.00	省石門水庫 自記	43 6	43-
P006	秀巒 新竹縣尖石鄉秀巒村控溪14號	白石溪	121 17 0 24 37 0	840.00	省石門水庫 自記	40 3	40-41 43-
P007	玉峰 新竹縣尖石鄉玉峰村5鄰		121 18 0 24 40 0	740.00	省石門水庫 自記	43 6	43-

（i＝4），即 X_1、X_2、X_3 及 X_4，而以欲補遺的秀巒站日雨量爲所預測的Y值（假設秀巒站有100 天日雨量記錄遺失），形成五維的歐氏空間，其中任一點可表示成（X_1、X_2、X_3、X_4、Y），在所有點中隨機抽選1000組做爲訓練（Training），將此1000點逐步輸入模式中，以形成許多不同的巢狀超矩形結構，然後以秀巒站欲補遺日期之其它四站資料（X_1、X_2、X_3、X_4）共100組來預測Y值，此過程不再調整模式中的任何參數值，模式將所預測出的Y值與其記錄中眞正發生的Y值做比較以判斷其是否正確。

由模式所得的預測結果與傳統的雨量補遺「內插法」，（王如意、易任，1985）相較，以正確度而言，二者大致相同，無明顯的優劣之分，原因乃爲周圍四站與秀巒站的日雨量呈現良好的線性關係（相關係數r均在 0.9左右），有利於使用內插法，且由於各站日雨量本身不具有一致性，即指相互接近的X值事件中其Y值差異相當大，舉例：以前發生過（X_1、X_2、X_3、X_4、Y）爲（0、0、0、0、0），而某次却出現（0、0、0、0、100）的特殊情況，所以由巢狀超矩形學習模式未能得到更理想的結果。

5-4.　平均年流量的補遺

水資源規劃時，常面臨流量站資料短缺的問題，如水文站記錄過短或某一時期因故資料遺落，故需加以補遺或延伸。以往常用的方式爲藉由鄰近的測站以迴歸的方法進行補遺或延伸的工作，迴歸分析主要的缺點爲需預先設定迴歸的函數型態，如線性迴歸，另外當資料個數少時，其迴歸所得結果之不確定(或變異數)問題極難克服，巢狀超矩形演算法提供了另一有效可行的替代方案，茲以美國Potomac 算流域中的四個流量站：Strausburg, Antietam, Point of Rocks 和 Cumberland，資料爲30年（1931 —1960）的平均年流量值示於表3與圖5（Salas et al., 1980）爲例。

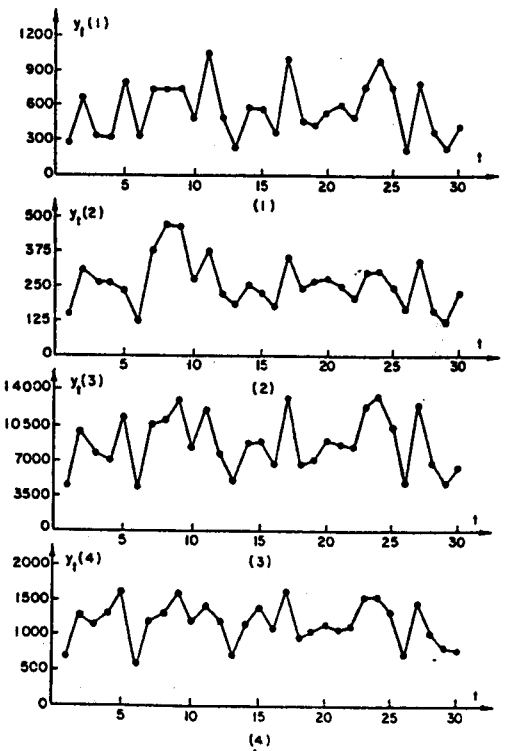

圖5 Potomac 流域四流量站：(1)Strausburg, (2)Antietam, (3)Point of Rocks, 與 (4) Cumber land從1931至1960之年流量 (cfs)

表 3 四站的年平均流量 (cfs) (1931-1960)

Station Strausburg (S_1)									
282	675	344	325	904	330	742	741	745	500
1047	490	241	592	567	366	998	461	330	545
610	494	769	984	767	226	790	378	232	402
Station Antietam (S_2)									
158	309	268	268	242	127	383	479	423	278
378	228	192	257	231	181	357	242	270	280
251	209	336	340	251	173	344	168	125	229
Station Point of Rocks (S_3)									
4642	10100	7767	7056	11350	4665	10840	11010	13030	8543
13210	7867	5220	8828	8925	6849	13480	6744	7317	9108
9002	8692	12670	13440	10670	4856	12700	6920	4920	6490
Station Cumberland (S_4)									
731	1314	1192	1344	1649	643	1237	1336	1609	1231
1440	1228	750	1172	1439	1134	1652	986	1087	1175
1113	1218	1574	1570	1356	760	1480	1060	852	799

我們假設四站中的某一站其最後10年的平均年流量未知，必須由其他三站這10年的平均年流量來推估，方法爲先將四個站前20年中某些年的平均年流量做爲訓練資料，輸入本模式以建立巢狀超矩形，其後即不修改參數值，並用於預測最後10年裏某一站未知的平均年流量值。

本研究即以四站前20年資料爲訓練集合，誤差寬容度定爲85，由本學習模式所得對最後10年的預測結果，其中分別探討第一站至第四站爲未知的四種不同情況，計算其預測之平均絕對誤差 (Mean Absolute Errors) 與訓練形成的超矩形分類個數。另外亦以相同的訓練集合（前20年資料）進行線性複迴歸分析定出四站中某一站爲未知的迴歸式，以此迴歸式求出某一站最後10年的平均年流量，並計算迴歸式的相關係數R與預測結果的平均絕對誤差，並與本學習模式一一對應比較，示於表4，結果都比迴歸方法的平均絕對誤差較小。

爲觀察不同的訓練個數對預測結果的影響，以第三站 (Point of Bocks) 爲例，分別用前5年、10年、15年至20年爲訓練，以預測其最後10年的平均流量，發現無論平均絕對誤差或平方均根誤差 (Root Mean Square Error) 都隨着訓練個數的增加而逐漸減少，表示所用的訓練個數愈多則預測的準確度就愈高，但愈來愈趨於平緩（改進逐漸減少），當所訓練的個數超過20時已無明顯的進步

表 4 流量四站最後十年之補遺結果比較

訓練年數：N＝20　　預測年數：m＝10

站　　名	平均絕對誤差	
	超矩形	線性迴歸
S_1: Strausburg	58.7	63.2
S_2: Antietam	43.8	53.8
S_3: Point of Rocks	488.2	537.5
S_4: Cumberland	63.9	128.6

註：四站之線性迴歸式與相關係數值

$S_1 = 1.153 - 0.698S_2 + 0.13S_3 - 0.325S_4$　　$R = 0.9345$

$S_2 = 56.339 - 0.397S_1 + 0.078S_3 - 0.204S_4$　　$R = 0.7517$

$S_3 = 310.522 + 6.343S_1 + 6.716S_2 + 3.051S_4$　　$R = 0.9768$

$S_4 = 268.44 - 1.14S_1 - 1.258S_2 + 0.22S_3$　　$R = 0.8488$

表5 "Point of Rocks" 站以5年訓練之補遺結果比較

年流量	線性廻歸		超矩形模式	
Y	Yr	|Yr-Y|	Yh	|Yh-Y|
9002	7716.906	1285.094	10100	1098
8692	8257.29	434.71	7767	925
12670	12482.675	187.325	11350	1320
13440	12345.286	1094.714	11350	2090
10670	9706.563	963.437	10100	570
4856	4108.31	747.69	4642	214
12700	11742.308	957.692	11350	1350
6920	6540.12	379.88	7767	847
4920	4390.884	529.116	4642	278
6490	4917.094	1572.906	4642	1848
絕對誤差總和＝8152.564			10540	
平均絕對誤差＝ 815.256			1054.0	
平方均根差＝ 915.241			1207.3	

表6 "Point of Rocks" 站以10年訓練之補遺結果比較

年流量	線性廻歸		超矩形模式	
Y	Yr	|Yr-Y|	Yh	|Yh-Y|
9002	8020.716	981.284	8543	459
8692	7195.013	1496.987	8543	149
12670	12482.659	172.659	13030	360
13440	13208.835	231.165	11350	2090
10670	9391.295	1278.708	10100	570
4856	3631.611	1224.389	4642	214
12700	12641.793	58.261	13030	330
6920	5129.597	1790.403	7767	847
4920	2728.671	2191.329	4642	278
6490	5622.558	867.442	4642	1848
絕對誤差總和＝10292.624			7145	
平均絕對誤差＝ 1029.262			714.5	
平方均根差＝ 1232.261			970.6	

表7 "Point of Rocks" 站以15年訓練之補遺結果比較

年流量	線性廻歸		超矩形模式	
Y	Yr	|Yr-Y|	Yh	|Yh-Y|
9002	8006.669	995.331	8828	174
8692	8303.243	388.757	7867	825
12670	11663.411	1006.589	13030	360
13440	11479.103	1960.897	13210	230
10670	9367.835	1302.165	10100	570
4856	5311.983	455.983	5220	364
12700	11148.123	1551.877	13030	330
6920	6976.627	56.627	7767	847
4920	5345.567	425.567	5220	300
6490	6008.633	481.367	4642	1848
絕對誤差總和＝8625.16			5848	
平均絕對誤差＝ 862.516			584.8	
平方均根差＝1035.763			753.8	

表8 "Point of Rocks" 站以20年訓練之補遺結果比較

年流量	線性廻歸		超矩形模式	
Y	Yr	|Yr-Y|	Yh	|Yh-Y|
9002	8640.187	361.813	8828	174
8692	7942.682	749.318	7867	825
12670	11626.095	1043.905	13030	360
13440	13004.5	435.5	13210	230
10670	10377.431	292.569	10100	570
4856	4603.624	252.376	5220	364
12700	11526.232	1173.768	13030	330
6920	6449.48	470.52	6849	71
4920	4600.006	319.994	5220	300
6490	6215.007	274.923	4642	1848
絕對誤差總和＝5374.686			4882	
平均絕對誤差＝ 537.469			488.2	
平方均根差＝ 624.3543			701.3	

，另外，也以線性複迴歸方法做相同的預測比較，結果在訓練個數N＝5時，本學習模式的兩種誤差均大於迴歸，而在N＝10、15時，已比迴歸的表現好，最後N＝20時，在平均絕對誤差方面本模式較低，而在平方均根誤差方面却較迴歸爲高，原因是預測的10年中前9年本學習模式的結果相當理想，可是最後1年的誤差却很大，因在20年的訓練資料中找不到接近的點，所以預測值不是很好，以上數據列於表5至表8，訓練個數與平均絕對誤差的相關圖繪於圖6。

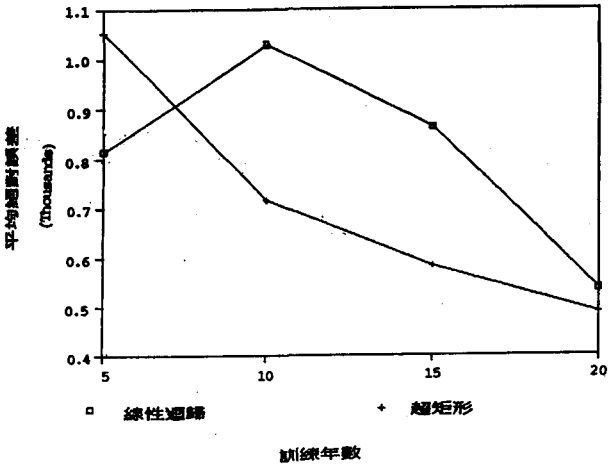

圖6　超矩形模式與線性迴歸在平均絕對誤差的比較

六、討論與結論

1.「巢狀超矩形學習模式」是一種以歷史性的範例爲基礎的學習方式，適合應用於分類或預測，其基本假設爲事件的相似性，即過去的事件將來有重覆發生的情況，如果不符合此條件，則模式演算結果將不理想。

2.「距離計量」公式可動態調整各項參數值，以求得最佳的結果，不同於一般分羣（Cluster）原理中固定不變的純幾何意義，但如何精確的修正參數需要靠經驗法則，並無嚴謹的論證，此部份爲模式建立的一大難題。

3.本模式最主要的優點在於利用極少的資源（每個超矩形只需記憶對角線上的兩個點），即可達到令人滿意分類預測，且將過去發生的所有有用事件都予以考慮，將來亦隨着樣本事件的增加而使結果愈來愈正確。

4.本研究提出一個假設範例與三個應用實例，前者爲一數學函數，完全符合模式的基本假設，所以可獲至相當完美的預測結果。後者第一個爲推估河川流量，屬於分類的一環，亦有成功的表現。但第二個雨量補遺的例子，因爲資料本身缺乏一致性，資料結構常有自相矛盾之處，所以預測結果未達預期之理想。而第三個年平均流量的補遺實例，結果比傳統的線性複迴歸分析所得的平均絕對誤差較小，展現了優良的補遺能力。由以上應用之結果顯示巢狀超矩形學習模式對水資源系統之預測與分類提供另一種極具效用的替代方式，值得進一步的探討與推展。

七、誌　　謝

本研究承蒙臺灣大學農工系易任教授之支持與多方鼓勵及中山大學應用數學系官大智教授提供許多寶貴意見，謹此併致謝忱，另評審委員提出數項改善意見，使本文更臻完善，於此亦表感激之意。

參 考 文 獻

1.王如意、易任，「應用水文學（上册）」茂昌圖書有限公司，1985。

2.張斐章、徐國麟，「利用模糊集理論推估河川流量之研究」中國農業工程學報第三十六卷第四期，1990。

3.Aha, D. and Kibler, D., "Noise-Tolerant Instance-Based Learning Algorithms" Proceedings of the International Joint Conference on Artificial Intelligenec, Mogan Kaufmann Publishers, 1989.

4.Everitt, Brian, "Cluster Analysis" Gower Publishing Co. Ltd., Hampshire, England, 1980.

5.Fisher, R. A., "The use of Multiple Measurements in Taxonomic Problems." Annals of Eugenics 7(1), 1936.

6.Salas, J. D., J. W. Delleur, V. Yevjevich, and W. L. Lane,"Applied Modeling of Hydrologic Time Series." Water

Resource Publications, 1980.
7.Medin, D. and Schaffer, M., "Context Theory of Classification Learning." Psychological Review, 85(3), 207—238, 1978.
8.Michalski, R., Carbonell, J., and Mitchell, T., "Machine Learning" Tioga Publishing Co., 1983.
9.Michalski, R., Mozetic, I., Hong, J., and Lavrac, N., "The Multi-Purpose Incremental Learning System AQ15 and Its Testing Application to Three Medical Domains." Proceedings of AAAI-86, Philadelphia, Pennsylvania, 1041—1045, 1986.
10.Salzberg, Steven, "Heuristic for Inductive Learning". Proceedings of IJCAI 85, Los Angeles, California, 603—610, 1985.
11.Salzberg, Steven: "Pinpointing Good Hypothesis with Heuristics" In Artificial Intelligence and Statistics." W. Gale (ed.), Addison-Wesley, 133—159, 1986.
12.Salzberg, Steven: "Exemplar-Based Learning: Theory and Implementation." Technical Report TR-10-88, Center for Research in Computing Technology, Harvard University, 1988.

收稿日期：民國81年8月13日
修正日期：民國81年8月24日
接受日期：民國81年9月9日