

應用類神經網路推估區域地下水中砷污染之研究

Applying ANNs for Estimating the Regional Arsenic Pollution in Groundwater

國立台灣大學生物環境系統工程學系

博士班研究生

高力山

Li-shan Kao

國立台灣大學生物環境系統工程學系

教授

張斐章*

Fi-john Chang

摘要

本研究中應用類神經網路建立地下水砷濃度推估模式，以解決高度非線性砷污染傳輸問題，及提高砷濃度推估之準確性。本研究以雲林縣沿海地區為研究區域，採用類神經網路建構地下水中砷推估模式，模式分為單一水井水質及區域水井水質之類神經網路模式，單一水井水質類神經網路模式主要針對單一監測井藉由砷與其他水質因子相關性，建立類神經網路推估模式，區域水井水質類神經網路模式則是應用全區資料，建立適用研究區域範圍之砷濃度推估模式，研究中除探討輸入因子及網路架構對模式誤差之影響外，並針對地下水水質模式較易遭遇到資料過少問題，提出交叉驗證法及修正型目標函數加以改善模式推估誤差，其中單一水井水質類神經網路模式之平均誤差(rmse)為 65.7ug/l，區域水井水質類神經網路模式之平均誤差(rmse)為 112ug/l，大部分監測井推估結果屬可接受範圍，僅#7 監測井誤差較大，但對於本區砷變動較大且複雜地區，而監測資料有限下，藉由類神經網路達到可接受誤差，本模式成功解決過去傳統模式不易推估區域地下水中砷污染問題。最後，本研究將模式推估結果結合地理資訊系統(GIS)展示雲林縣沿海地區地下水中砷污染分布情形，可作為日後政府管理地下水之參考依據。

關鍵詞：砷、地下水、人工智慧、類神經網路、修正型目標函數

*通訊作者，臺灣大學生物環境系統工程所教授，106 臺北市大安區羅斯福路 4 段 1 號，電話(02)3366-3452，changfj@ntu.edu.tw

ABSTRACT

The groundwater extracted by some regional farmers leads to a lower level of the groundwater and a release of the poisonous substance in the groundwater. That affects the health of local residents, even the inhabitants who do not ingest the local agriculture and aquaculture products. The aim of the study is to build the arsenic water quality model by adopting the artificial neural networks (ANNs). Taking the YUN-LIN County as an example, the ANNs were constructed and assessed. The models are divided into two parts: (1) single well models and (2) regional well models. In the process of constructing the models, the optimal input factors and structures of the ANN models were discussed in this study. At the same time, we applied the cross validation method and the modified objective function to solving the data scarce problems of the monitoring wells. The results produced by the single well models and the regional well models were compared and demonstrated their applicability. Finally, the results obtained by the ANN models were integrated with GIS to display the distribution of the arsenic concentration at the coastal area in the YUN-LIN County. The results can offer a good reference to government decision-makers for the management of the groundwater and installation of monitoring wells.

Keywords : Arsenic, Groundwater, Artificial Intelligence (A.I.), Artificial Neural Networks (ANNs), Modify-Objective Function

一、前言

砷普遍存在於自然界中，臺灣含高砷濃度之地下水區域，主要分布於嘉南平原、濁水溪沖積扇南翼以及蘭陽平原，台灣地區於 1960 年代出現大規模之烏腳病疫情，可歸因於當地居民長期飲用高砷含量之地下水，過去研究報告亦指出，地下水中無機砷與烏腳病具高度相關性，且對人體有致癌高危險風險，因此，了解地下水水質特性將有助於地下水使用之安全與管理，而雲林縣除為砷污染較嚴重區域外，並具有長期地下水水位及水質監測資料，故本研究選擇雲林縣為研究區域，其位置如圖 1。

傳統地下水水質模式建立，必須蒐集完整地層參數、地下水水位及初始污染物濃度，建立模式前須蒐集龐大地層及水文資料，而在經費及預算有限下，相關地層參數資料僅限於局部地區，而研究者必須針對全區各項參數進行推估及驗證，推估參數與實際地層狀況往往產生很大差異，使得模式推估誤差情形無法有效改善，並且傳統地下水水質模式難以整合不同類型資料，無法適切應用於雲林縣沿海地區地下水受多重影響之複雜地區，故本研究嘗試採用類神經網路整合地下水中不同水質因子，建立地下水中砷濃度推估模式。

本研究藉由分析歷史資料及相關研究選用與地下水中砷含量最相關因子，作為模式輸入因子，建立單一水井水質與區域水井水質類神經網路，以推估砷濃度在地下水中變化，在建立模式過程中除須評估最佳模式架構外，並須考慮水質資料過少之特性，加入交叉驗證法及修正型目標函數，再運用多種方法避免建構模式過程中過度訓練之問題，以求達到可接受模式推估結果，最後，本研究將模式推估結果應用地理資訊系統(GIS)展現歷年砷濃度在雲林縣沿海地區變化情形，本推估模式藉由地下水中其他水質因子，推估地下水中砷污染情形，可作為地下水污染防治及地下水使用之參考。



圖 1 雲林縣位置圖

本研究採用類神經網路作為推估地下水中砷濃度之方法，過去類神經網路在水文系統或地下水方面應用相當廣泛，且均獲得滿意結果，包括陳韋佑及陳昶憲(1993)以定率倒傳遞類神經網路為主架構，經評選訂定網路架構之後，建構多組符合學習精度要求之之倒傳遞類神經網路模式，再依驗證資料之輸入資料結構特性，以比對之方式由前述模式中尋求合適之推估模式，並將模式應用於烏溪流域洪流量之預報上，其驗證結果良好，證明群策倒傳遞類神經網路可提供決策者在流量預測時，更具可信度之預測模式;郭益銘(1999)應用多變量統計分析與類神經網路分析雲林沿海地區地下水質變化；Gumrah et. al. (2000)考量石油抽取後大量廢鹽水注入地下水層造成地下水污染，採用氯離子與地下水水位之過去資料作為訓練，架構類神經網路，有效預測鹽水擴散情形，並可預測地下水水位變化情形，可有效控制地下水污染與使用；Martos(2002)等人藉由 SOM 分析方法將多種地下水水質離子濃度變化歸納在卡式座標中網格，此方法將西班牙南部複雜地下水水質狀況，簡化為讓人較容易了解座標中，地下水水質在卡式座標中分布位置，即表示該水質之污染情形及特性;Coppola et. al. (2003)考慮狀態變數、氣候及抽水條件，架構類神經網路預測多層地下水中地下水位變化，並將類神經模式與物理模式之數值方法相比較，證實此法可精確預測複雜之地下水變化；Mohammad(2010)比較傳統補遺方法與類神經網路補遺集水區中流量資料，發現正常比例法(the normal ratio method ,NRM)及相關係數法(the correlation method ,CM)推估誤差較類神經網路為高，雖其差異性不大，但是最重要類神經網路可解決異質性資料整

合，具高度資料補遺之價值，這些類神經網路相關應用證明類神經網路具有強大非線性及複雜問題之處理能力。

另在砷污染相關研究上，過去研究顯示嘉南沿海地區發生嚴重的烏腳病事件，其主因可能為飲用水中含有較高砷濃度引起(曾文賓,1976);王榮德等人(1983)應用統計分析出嘉南沿海地區的烏腳病是由於飲用當地深井地下水引起，並在地下水抽出時檢測出高濃度的砷;而雲林地區沿海地區歷年地下水水質檢驗成果(成功大學水工試驗所,1991~1998)發現，雲林四湖鄉沿海區域砷濃度皆有偏高現象。

在地下水水質相關研究方法方面，Madhumita Das(2010)應用統計多變數分析技術(multivariate techniques)評估都市地區飲用地下水質，發現PH值、鐵離子及硝酸根離子為影響地下水水質最重要三項因子，可作為地下水處理之指標;廖少威(1991)利用主成份分析法探討台灣灌溉水質受工業、都市、畜牧廢水污染之變異性。

二、理論說明

2.1 類神經網路

類神經網路主要的基本概念為模仿人類腦神經系統，因此在邏輯的組成單元與架構上皆與生物神經網路類似，其中最重要的三項功能為學習、回想及歸納推演。而類神經網路經過了長時間的發展，已有許多不同的網路模型被提出，常見的網路架構有倒傳遞類神經網路、輻狀基底函數網路及自組織特徵圖網路等，因其具有高度訊息處理與非線性映射的能力，可解決物理模式難以描述與建構之非線性關係，因此被大量地應用於水文領域上。

在國內張斐章等人(1998)即採用反傳遞模糊類神經網路應用於河川流量推估。張和孫(1997)應用倒傳遞類神經網路於降雨及逕流過程之研究。近年來，Chang et. al. (2001, 2002, 2006)更採用不同類神經網路架構於不同水文量之推估與研究，均獲致可靠水文量預測與推估成果，如2001年應用模糊倒傳遞類神經網路對河川流量推估，2002年應用即時回饋類神經網路預測河川流量，研究中並對各種網路架構之優劣進行探討，提供後續研究之參考依據。

在各種類神經網路中，以倒傳遞類神經網路具有學習精度高、回想速度快及

能處理複雜的樣本識別與高度非線性的函數合成問題，因此為應用範圍廣泛之網路，本研究在評估各種網路型式後，仍以倒傳遞類神經網路為較適合地下水中砷濃度推估水質模式之網路型式，其架構為一多層前饋式網路，可分為輸入、隱藏及輸出三層(如圖 2 所示)。其使用監督式學習法，藉由網路輸出值與目標輸出值間之誤差來修正模式，因此網路學習的過程即為使誤差函數 E 達到最小值。

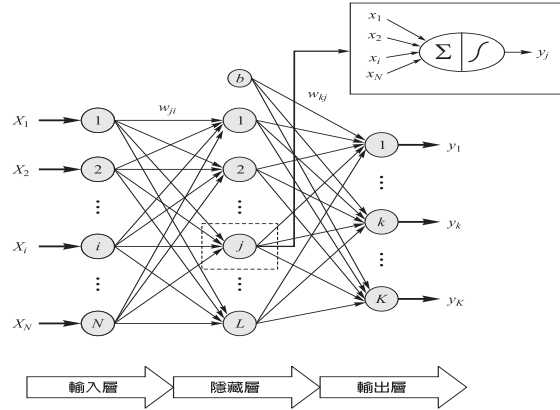


圖 2 倒傳遞類神經網路架構圖

在倒傳遞類神經網路中，第 n 層第 j 個神經元的輸入值為第 $n - 1$ 層神經元輸出值的非線性函數：

$$y_j^n = f(net_j^n) \quad (1)$$

式中 y_j^n 為第 n 層的輸出值，對第一層來說為輸入值， f 為活化函數， net_j^n 為第 $n - 1$ 層輸出值的權重累加值，可表示為 (2) 式：

$$net_j^n = \sum_i w_{ji}^n y_i^{n-1} + b_j^n \quad (2)$$

式中 w_{ji}^n 為第 n 層第 j 個神經元與第 $n - 1$ 層第 i 個神經元的連結權重， b_j^n 為第 n 層第 j 個神經元的偏權值。由於倒傳遞類神經網路屬監督式學習，其學習的目的在降低網路輸出值與目標輸出值間的差距，誤差函數 E 定義為

$$E = (1/2) \sum_k (d_k - y_k)^2 \quad (3)$$

式中 d_k 為第 k 個神經元的目標輸出值， y_k 為輸出層第 k 個神經元的網路輸出值。因此，網路的學習過程，即為使誤差函數 E 最小化的過程，在此方面前人研究提供不同演算法搜尋 E 的最佳解。

在完成倒傳遞演算法流程後，尚有許多網路架構問題須決定，例如如何決定輸入層個數，選用何種因子作為輸入層、選用何種活化函數、隱藏層個數及選擇何種演算法等均會影響網路輸出誤差。

在活化函數選用上，過去研究較常使用對數雙彎曲活化函數、正切彎曲活化函數、線性轉移活化函數，使用者可根據問題須要選用較適當活化函數。至於隱藏層神經元個數的常用方法有兩種：其一是使用網路修剪法 (pruning algorithm) 由 Abrahart et al. (1998) 提出，另一種方法則是網路增長法 (constructive algorithm) 是在 1997 年由 Kwok 和 Yeung 提出。另一重要問題即是演算法選用，在過去較常使用為梯度下降法與動量的梯度下降法，但是這兩種演算法對於實際問題求解太過耗時，近年來提出一些較高效率之演算法，第一類為啟發式演算法，第二類為標準的數值最佳化，其中啟發式演算法又可分為 1. 可變學習速率 (Variable Learning Rate) 及 2. 有彈性的 (Resilient) 倒傳遞演算法，在此要決定何種演算法需考慮問題的複雜度、訓練數據多寡、網路中權重值及誤差目標值大小，因此，各種演算法須配合整體網路架構、資料數量及資料特性等，以找尋最佳化之倒傳遞網路模式。

2.2 交叉驗證法

交叉驗證法 (cross validation) 常應用資料過少時校驗網路推估值之良莠，其在類神經網路應用上，為每筆資料交替使用在訓練時期及驗證時期，由於本區監測井可取得水質資料有限，若由人為選擇訓練與驗證資料，在架構類神經網路時，將顯著影響模式表現優劣，為消除人為效應，本研究採用交叉驗證法，其計算步驟說明如下。

- (1) 假設每座監測井具有 N 筆 (假設 50 筆) 水質資料，包括砷濃度及其他水質因子。
- (2) 將每座監測井 N 筆水質，切割為 M 段 (假設 5 段)，每段中資料筆數為 N/M ($50/5=10$ 筆)。
- (3) 在訓練階段，選用 M 段中 M-1 段作為訓練段，在本研究中為選擇 4 段 (編號 11~50 資料，共 40 筆) 作為訓練資料，在訓練過程中，藉由目標函數調整網路中權重，最後計算網路之估計值與實際值之差，計算訓練誤差。
- (4) 驗證過程中，選用訓練不使用之 1 段 (1~10 筆資料) 作為驗證之用，並根據訓練過程中得到網路權重，計算驗證過程中，網路輸出之估計值與實際值之差，即

為驗證誤差，此為完成第 1 段交叉驗證，接著回到步驟(3)，重新選擇第 2 段(編號 11~20 資料)作為驗證資料，其餘資料第 1 段(編號 1~10 資料)及第 3~5 段(編號 21~50 資料)作為訓練資料，直至 1~5 段資料均作為驗證資料，並且其他段資料均作為訓練資料後，即可稱完成本監測井之交叉驗證工作，可得到 M 組(5 段)之誤差。

- (5) 在本模式中將資料分為 M 段(假設 5 段)，故在分段訓練及驗證過程中可得到 M 組訓練及驗證之誤差值，將 M 組誤差值平均後，可得到平均訓練及驗證之誤差值，本誤差可客觀表示本網路架構之誤差值。
- (6) 完成第 1 號監測井之交叉驗證工作後，接續第 2 號監測井之交叉驗證工作，直至所有監測井水質模式均完成交叉驗證工作，即可獲得交叉驗證結果。

整體交叉驗證過程，資料充份混合及重新排序，可得到整體網路較穩定之參數值，比較交叉驗證時，訓練與驗證之誤差，即可了解模式之穩定性及可靠性。

2.3 修正型目標函數

在監測井水質資料過少，網路結構過於龐大，常發生過度訓練情形，即網路在訓練過程誤差極小，但於驗證階段卻產生極大誤差，其產生誤差原因為在訓練過程中，過多網路中參數調整至符合訓練資料，卻未能符合資料特性，故在後續驗證時，模式將產生極大誤差。

由於本研究採用資料特性具有高度複雜性但資料卻過少，極容易產生上述過度訓練情形，為減少過度訓練情形，過去相關研究提出三類主要改善方法(羅華強，2005)，分別為降低隱藏層數目、修改目標函數及提早停止模式，本研究嘗試採用修改目標函數改善過度訓練情形，在修改目標函數上，一般倒傳遞類神經網路採用式(4)作為目標函數及評估模式之表現，本研究考量採用修正型目標函數式(5)，而式(5)為結合式(6)及式(7)，並採用性能比(γ)分別給與兩數值不同權重，且兩數值權重相加總合必須為 1，式(4)中 N 為資料筆數， t_i 為目標輸出值， a_i 為網路輸出值， e_i 為目標輸出值與網路輸出值之差值；式(5)中 $msereg$ 為本研究採用之修正型目標函數， γ 為性能比；式(6)為計算平均誤差值平方，式(7)為計算平均網路權重平方，其中 n 為網路權重個數， w_j 為網路中連結權重。使用修正型目標函數將造成網路有較小的權重值，並且強制網路輸出數值變化幅度減小，有效減少過度訓練之機會及網路不穩定情形，如設定性能比較大，對改善過度訓練效果

較差，模式穩定度不佳，但對於極端值預測卻較準確，相反地，如性能比較小，較能減少過度訓練情形，但卻需降低極端值之推估精確度，因此，在採用修正型目標函數時，須兼顧精確度與減少過度訓練發生，選取較適合之性能比。

$$rmse = \sqrt{\frac{1}{N} \sum_{i=1}^n (e_i)^2} = \sqrt{\frac{1}{N} \sum_{i=1}^n (t_i - a_i)^2} \quad (4)$$

$$msereg = \gamma mse + (1 - \gamma) msw \quad (5)$$

$$mse = \frac{1}{N} \sum_{i=1}^N (e_i)^2 = \frac{1}{N} \sum_{i=1}^N (t_i - a_i)^2 \quad (6)$$

$$msw = \frac{1}{n} \sum_{j=1}^n w_j^2 \quad (7)$$

三、結果與討論

3.1 研究區域及材料

本研究以臺灣西南沿海之濁水溪下游沖積扇(雲林縣)為例。雲林縣位在台灣的中部，在嘉南平原最北端，北邊沿著濁水溪和彰化縣接壤，東西最寬的地方有 50 公里，南北最長的地方有 38 公里，全縣面積總計 1,290 平方公里，其中十分之九為平原，十分之一為山地，人口約 73 萬 5 千人。

根據水利署及中央地質調查所於雲林縣設置地下水位觀測站址之水文地質鑽探資料，顯示本地區自地表至 300 公尺深度間，共包含 4 個含水層和 4 個阻水層，由上而下命名為含水層一、阻水層一、含水層二、阻水層二、含水層三、阻水層三、含水層四及阻水層四。

經濟部水利署鑒於地下水污染對於居民健康有重大影響，於 1992 年起執行台灣地區地下水觀測站網整體計畫(1992~2008)，在台灣地下水使用較頻繁且污染情形較嚴重地區設置地下水水位及水質監測水井，其中在雲林離島基礎工業區及工業區附近，其水質監測井分佈於台 17、雲 129 及雲 3 公路；於口湖、四湖、台西與麥寮等四個濱海鄉鎮，本研究共蒐集 22 座水質監測井(編號#1 ~ #22)資料，

如圖 3，本研究所蒐集資料時間長度為 1992 至 2005 年間調查的地下水質資料，每站約具 10~50 次調查資料，每座監測井共 50 筆地下水水質相關資料，水質資料包括砷、溫度、pH、電導度、溶氧、鹼度、TDS、氯鹽、硫酸鹽、游離氨氮、硝酸鹽氮、亞硝酸鹽、COD、總有機碳、BOD、鎘、鉻、銅、鋅、鎳、鈉、鉀、鐵、鉛、鎂、鈣、硬度、鈉吸著率，共 28 項水質項目，因本研究 22 座監測井中僅 6 座監測井具備完整降雨量、地下水水位及 17 項水質資料，其平均值及標準偏差如表 1，故本研究以此 6 座地下水監測井之 17 項水質資料作為主要研究對象，並分析此 6 座監測井之降雨量、地下水水位及 17 水質項目與砷濃度相關係數，如表 2。

由表 1 中了解本區地下水中平均砷濃度最高達 450ug/l，標準偏差達 315ug/l，顯示本區確實為砷高度污染區，且由於砷變化差異也非常大，加以各項水質變化差異也相當大，故建立推估模式確有其困難度。

表 2 顯示各監測水井與地下水中砷的相關係數，可知本區南北距離雖僅 40 公里，但由於本區地形地質條件差異、地下水流向及地下水位升降等因素，造成本區 6 座監測井影響地下水中砷之水質因素差異很大，如南部#3 以氯鹽及硫酸鹽為主要因子，中部#6 以月平均地下水位及總有機碳為主要因子，北部#17 以鈣及硬度為主要因子，全區以 pH 值及鈣為影響地下水中砷之最主要因子，pH 值與砷濃度具正相關性，鈣離子與砷濃度為負相關，其相關性係數為 0.34，這些影響砷濃度重要因子可作為模式之輸入因子推估砷濃度變化。

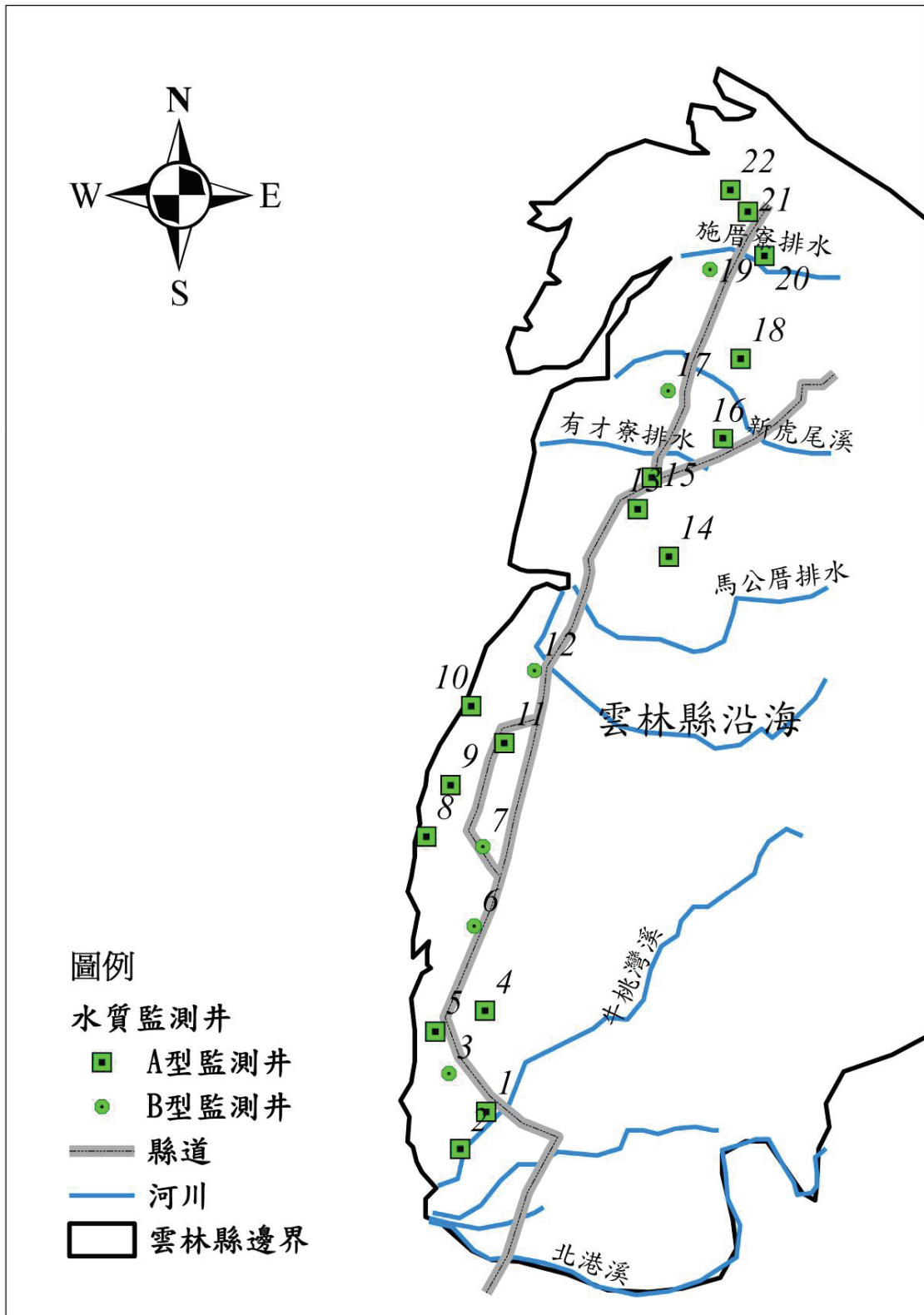


圖 3 雲林縣沿海地區水質站分布圖

說明:單一水井水質模式僅推估 B 型監測井之砷濃度

區域水井水質模式推估 A 型及 B 型監測井之砷濃度

表 1 雲林沿海地區觀測井水質資料統計表

監測井編號		#3		#6		#7		#12		#17		#19	
項目	單位	平均值	標準偏差	平均值	標準偏差	平均值	標準偏差	平均值	標準偏差	平均值	標準偏差	平均值	標準偏差
月平均降雨量	mm	305.2 ± 298.8	298.8	305.2 ± 298.8	298.8	305.2 ± 298.8	298.8	245.1 ± 262.5	262.5	245.1 ± 262.5	262.5	245.1 ± 262.5	262.5
月平均地下水水位	m	-2.2 ± 0.4	0.4	-2.2 ± 1.1	1.1	-5.4 ± 2.8	2.8	-2.4 ± 0.7	0.7	-0.2 ± 0.4	0.4	0.5 ± 0.2	0.2
砷	ug/L	75.9 ± 67.6	67.6	177.0 ± 109.5	109.5	450.4 ± 314.3	314.3	43.7 ± 30.7	30.7	39.5 ± 47.6	47.6	38.1 ± 30.7	30.7
溫度	C	25.8 ± 1.0	1.0	25.7 ± 0.9	0.9	25.8 ± 1.0	1.0	25.9 ± 1.3	1.3	26.1 ± 1.4	1.4	26.1 ± 1.3	1.3
pH		7.7 ± 0.4	0.4	7.9 ± 0.4	0.4	7.9 ± 0.2	0.2	7.7 ± 0.5	0.5	7.6 ± 0.4	0.4	7.6 ± 0.3	0.3
電導度		23382.7 ± 18221.0	18221.0	16509.4 ± 8317.4	8317.4	2209.0 ± 912.2	912.2	21794.6 ± 13885.9	13885.9	1408.4 ± 970.8	970.8	17295.1 ± 7677.2	7677.2
溶氧	uS/cm	1040.3 ± 7346.7	7346.7	419.0 ± 2922.6	2922.6	51.0 ± 349.1	349.1	1.3 ± 1.0	1.0	1.3 ± 1.0	1.0	1.3 ± 1.0	1.0
鹼度	ug/L	356.2 ± 137.3	137.3	560.2 ± 138.5	138.5	504.4 ± 84.8	84.8	384.1 ± 123.9	123.9	315.0 ± 55.7	55.7	504.2 ± 143.0	143.0
TDS	ug/L	15821.6 ± 10285.3	10285.3	10963.3 ± 4148.6	4148.6	1432.4 ± 626.2	626.2	13994.3 ± 7302.1	7302.1	885.9 ± 637.7	637.7	11790.0 ± 5425.5	5425.5
氯鹽	ug/L	6850.7 ± 4651.8	4651.8	4478.7 ± 1819.3	1819.3	391.9 ± 215.6	215.6	5944.9 ± 3418.7	3418.7	233.6 ± 236.3	236.3	4937.4 ± 2154.7	2154.7
硫酸鹽	ug/L	690.9 ± 784.1	784.1	512.0 ± 270.8	270.8	102.9 ± 82.8	82.8	960.0 ± 588.1	588.1	64.5 ± 57.4	57.4	515.6 ± 467.0	467.0
游離氨氮	ug/L	11.5 ± 10.6	10.6	19.0 ± 13.5	13.5	8.2 ± 3.5	3.5	4.4 ± 3.3	3.3	2.3 ± 2.5	2.5	10.6 ± 7.9	7.9
COD		42.5 ± 84.9	84.9	45.7 ± 103.9	103.9	17.4 ± 21.3	21.3	44.4 ± 103.6	103.6	11.6 ± 21.6	21.6	129.7 ± 287.5	287.5
總有機碳	ug/L	1.9 ± 1.1	1.1	4.5 ± 3.0	3.0	4.3 ± 2.2	2.2	1.5 ± 0.7	0.7	2.4 ± 1.8	1.8	2.2 ± 1.3	1.3
鈉	ug/L	3771.5 ± 2603.6	2603.6	2707.6 ± 998.5	998.5	293.5 ± 90.8	90.8	3297.1 ± 1806.0	1806.0	179.2 ± 121.7	121.7	2756.1 ± 1132.7	1132.7
鉀	ug/L	201.1 ± 105.6	105.6	145.2 ± 46.1	46.1	38.7 ± 15.5	15.5	133.4 ± 57.6	57.6	17.0 ± 11.4	11.4	142.4 ± 91.9	91.9
鎂	ug/L	598.8 ± 954.6	954.6	254.7 ± 205.5	205.5	73.2 ± 30.3	30.3	427.5 ± 296.4	296.4	32.2 ± 22.6	22.6	323.8 ± 156.3	156.3
鈣	ug/L	216.0 ± 133.1	133.1	74.9 ± 50.1	50.1	59.0 ± 19.7	19.7	281.8 ± 174.3	174.3	88.0 ± 46.7	46.7	150.4 ± 75.8	75.8
硬度	ug/L	2462.4 ± 1610.2	1610.2	1137.4 ± 901.2	901.2	643.1 ± 621.2	621.2	2405.8 ± 1595.7	1595.7	363.5 ± 197.5	197.5	1706.4 ± 788.3	788.3

表 2 雲林沿海地區觀測井水質與砷濃度相關係數表

監測井編號	#3	#6	#7	#12	#17	#19	全區
當月平均降雨量	0.06	0.22	-0.17	-0.14	0.14	0.15	0.02
當月平均水位	-0.14	<u>0.65</u>	-0.26	0.01	0.09	<u>0.38</u>	-0.33
溫度	0.31	-0.17	0.02	0.07	0.06	-0.10	-0.06
pH	0.49	-0.26	0.46	<u>0.40</u>	0.48	<u>-0.20</u>	<u>0.34</u>
電導度	-0.60	0.23	-0.61	-0.08	-0.64	-0.06	-0.29
溶氧	-0.14	-0.01	-0.09	0.16	0.05	-0.04	-0.03
鹼度	0.49	-0.24	-0.53	<u>0.32</u>	-0.33	0.04	0.18
TDS	-0.61	0.43	-0.58	0.06	-0.60	0.02	-0.30
氯鹽	<u>-0.63</u>	0.36	<u>-0.66</u>	-0.14	-0.66	-0.11	-0.32
硫酸鹽	<u>-0.63</u>	0.41	-0.54	-0.16	-0.67	-0.10	-0.27
游離氨氮	0.36	0.45	-0.30	-0.04	0.68	0.07	0.13
COD	0.11	-0.22	0.08	0.17	0.12	-0.17	-0.09
總有機碳	0.10	<u>-0.59</u>	-0.50	-0.11	0.15	-0.03	0.13
鈉	-0.62	0.42	-0.59	-0.13	-0.51	-0.01	-0.30
鉀	-0.45	0.45	-0.58	0.05	-0.62	-0.13	-0.25
鎂	-0.23	0.02	<u>-0.69</u>	-0.21	-0.54	0.11	-0.19
鈣	-0.54	0.18	-0.60	-0.24	<u>-0.70</u>	0.11	<u>-0.34</u>
硬度	-0.56	0.12	-0.32	-0.23	<u>-0.71</u>	0.12	-0.29

3.2 單一水井水質類神經網路模式

砷在地下水中受到多種水質作用，為有效掌握砷在地下水中變化，須建立地下水水質模式，但一般傳統地下水水質傳輸模式，須具備完整地層中水力傳導係數、儲水係數及地下水水位高程，且須有污染物初始濃度及邊界條件等，始能採用傳統有限元素法或有限差分法加以模擬及推估，但由於砷不僅受地下水水流方向而擴散，另受到地層中物理及化學作用而改變其濃度，由表 2 中可了解不同地層之地下水砷濃度具不同影響因子，故本研究藉由其它水質因子建立砷濃度推估模式，由本模式可推估砷濃度擴散情形，有效降低當地居民使用受污染地下水之機率。

考慮地下水中砷具有顯著地域性，不同地層中監測水井應具有不同類神經網路之架構，包括輸入因子、隱藏層個數、修正目標函數係數等，因此，本研究從 28 座監測水井中選擇其中水質資料較完整之 6 座監測水井，在輸入因子選擇上，由於 14 年中每座監測水井僅有約 50 筆資料可供訓練及驗證模式，且每座監測水井具有 18 項與砷有關水質項目，但在資料不足條件下，過多輸入因子，將會造成更大誤差，故本研究僅從表 2 中挑選相關性最高之水質因子作為輸入項目，另因每座監測水井約有 50 筆，對於建構模式略顯不足，且採用人為選擇訓練及驗證資料，容易造成模式誤差受人為選擇影響而產生偏差，故在建立單一水井水質類神經網路模式同時採用交叉驗證法及修正型目標函數，且須儘量減少隱藏層個數，避免過度訓練情形，經採用試誤法優選單一水井模式，最佳之隱藏層個數及修正型目標函數係數優選結果，如表 3，各監測水井之砷濃度推估模式架構圖，如圖 4。

由表 3 中 6 座監測水井模式訓練時期之平均地下水中砷濃度為 65.7ug/l，並將 6 座監測井之目標砷濃度及推估砷濃度，繪成如圖 5，6 座監測井中 5 座推估誤差介於 27~73ug/l，仍屬可合理範圍，惟有 #7 井位於高砷濃度地區，其砷濃度已超過模式可預測範圍，因此，推估誤差較大，大體而言，單一水井水質類神經網路模式推估誤差在容許誤差範圍內，且證明類神經網路在缺乏監測水質資料下仍可藉由少數水質因子可靠推估地下水中砷濃度變化，另在訓練及驗證過程中誤差數值(rmse)相當接近，展現加入修正型目標函數及交叉驗證法可減少模式過度訓練之效果。

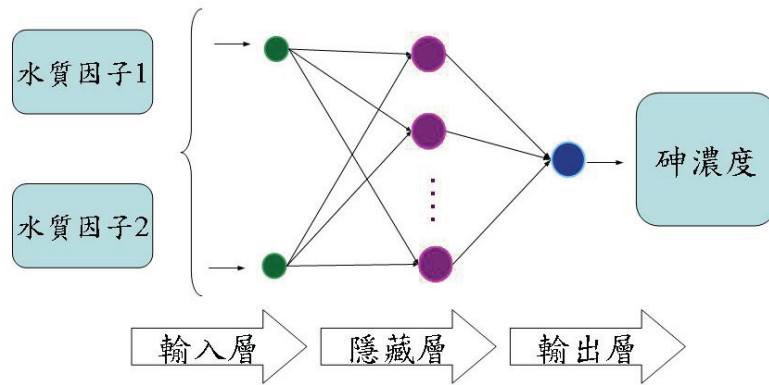


圖 4 單一水井水質類神經網路模式網路架構圖

表 3 雲林縣沿海地區單一水井水質類神經網路模式之架構及誤差表

區域	水井編號	輸入因子	隱藏層個數	γ	訓練時期		驗證時期	
					rmse (ug/l)	誤差百分比	rmse (ug/l)	誤差百分比
南部	#3	氯鹽及硫酸鹽	4	0.8	46.6	41.1%	41.5	47.0%
中部	#6	地下水水位及總有機碳	2	0.5	73.2	31.0%	66.9	32.5%
中部	#7	氯鹽及鎂	3	0.3	190.8	30.8%	183.8	36.1%
中部	#12	PH 及鹼度	5	0.8	27.3	47.4%	26.7	49.9%
北部	#17	鈣及硬度	2	0.5	27.8	42.8%	23.5	61.1%
北部	#19	地下水水位及 PH	3	0.1	28.9	49.3%	25.5	45.0%
平均值					65.7	40.4%	61.3	45.3%

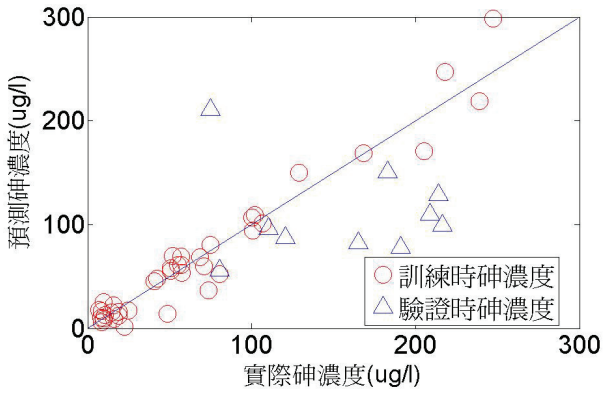


圖 5(a) #3 監測井訓練及驗證圖

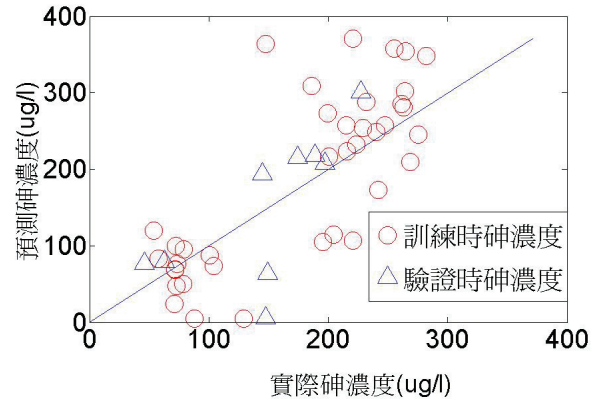


圖 5(b) #6 監測井訓練及驗證圖

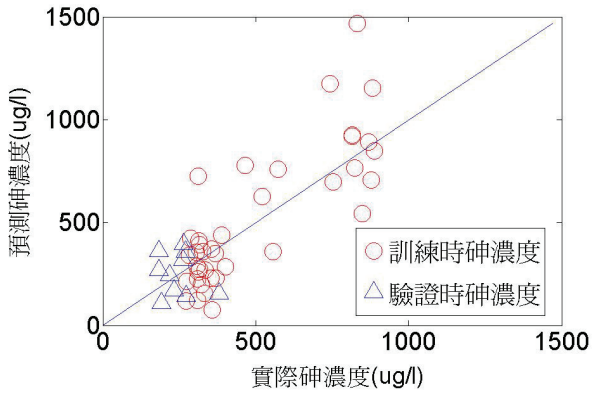


圖 5(c) #7 監測井訓練及驗證圖

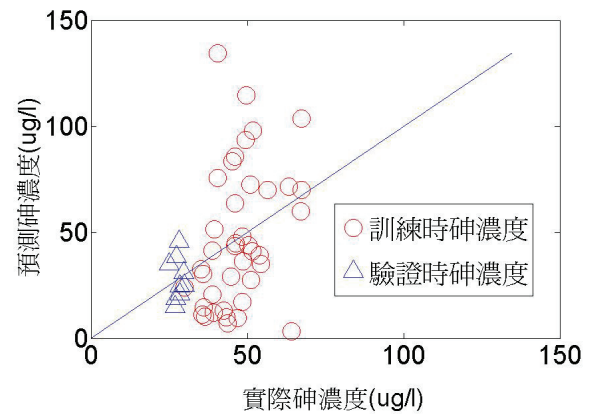


圖 5(d) #12 監測井訓練及驗證圖

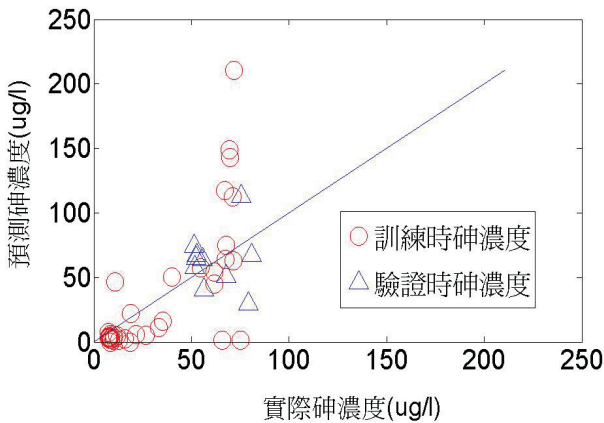


圖 5(e) #17 監測井訓練及驗證圖

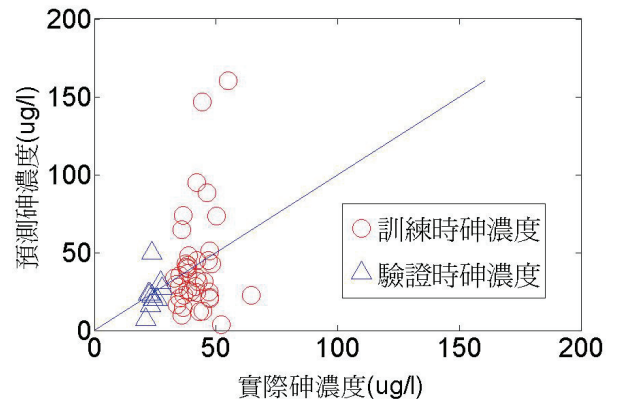


圖 5(f) #19 監測井訓練及驗證圖

3.3 區域水井水質類神經網路模式

為推估雲林縣沿海地區砷濃度變化情形，須建立可適用於研究區域之區域水井水質類神經網路模式，應用 6 座淺水井中較完整 18 項因子，包括月平均降雨量、月平均水位、地下水之溫度、pH、電導度、溶氧、鹼度、TDS、

氯鹽、硫酸鹽、游離氨氮、COD、總有機碳、鈉、鉀、鎂、鈣、硬度、砷為本研究主要探討因素。

為探討輸入層個數對類神經網路推估誤差之影響，分別由 Ph 值、電導度、鹼度、TDS、氯鹽、鈉、鉀、鎂、鈣、硬度增加網路中輸入層節點個數由 1 個至 8 個，輸入因子 1 個時，採用 ph 值作為輸入項；輸入因子為 2 個時，採用 Ph 值、電導度作為輸入項，依此類推；至輸入因子為 8 個時，即採用 Ph 值、電導度、鹼度、TDS、氯鹽、鈉、鉀、鎂、鈣、硬度 8 個作為輸入項，並相對應增加隱藏層個數，如圖 6，圖 6 可見不論隱藏層個數為何，在訓練階段類神經網路推估誤差雖隨輸入因子增加而減少，但在驗證階段類神經網路推估誤差卻隨輸入因子增加而增加，可知在本案例中增加輸入因子個數對於改善模式誤差並不明顯，在過多輸入項個數下，卻無法相對增加資料數量下，反而會造成網路中過多參數無法確定，而增加推估誤差，因此，本模式僅採用 2 個因子作為類神經網路之輸入因子。

在網路之輸入因子選用上，根據過去研究受到砷污染地下水嚴重之孟加拉地區，與地下水中砷濃度最相關為地下水之 ph 值、去鐵微生物、有機物含量及競爭性離子，其中競爭性離子特別以硫酸鹽為重要因子，與本研究較符合為 ph 值與砷相關係數達 0.34，如表 4，但表 4 中顯示雲林沿海地區以 ph 值及鈣離子與砷濃度相關性最高(0.34)，由於鐵離子在本研究中無完整資料，無法與孟加拉地區相對照，總有機碳則不與砷明顯相關。

考量過去研究資料及相關係數表，與砷最相關因子應為 ph 值及鈣離子，另在因子間相關係數上，發現較具代表性之水質因子除 ph 值外，應以電導度與其它水質因子最相關，在表 4 中 18 項水質因子中有 8 項水質因子與電導度相關係數高於 0.5，且電導度與硫酸鹽之相關係數達 0.86，更重要是電導度與各重金屬均呈現高度相關，可見電導度可替代硫酸鹽或鈣離子成為推估砷濃度另一重要因子。

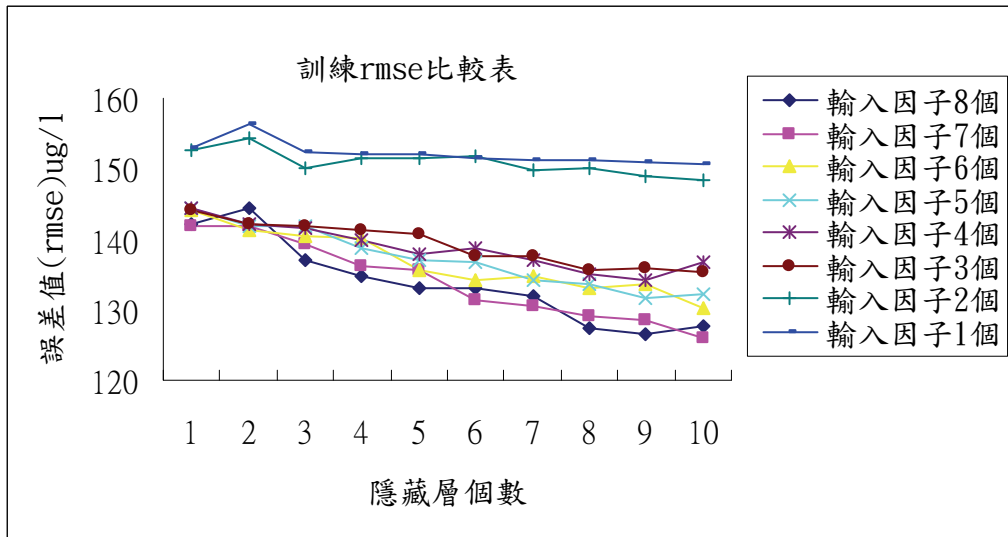


圖 6(a) 訓練階段輸入因子個數對類神經網路推估誤差之影響

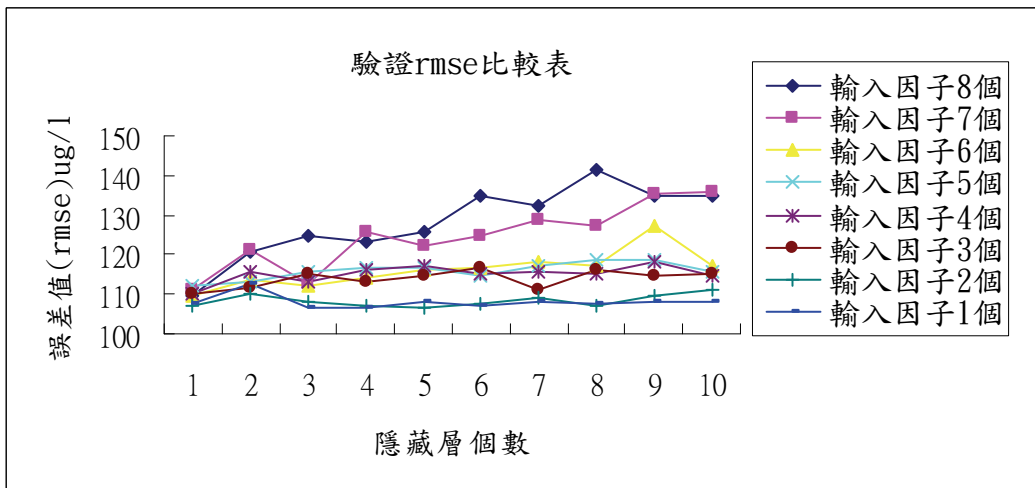


圖 6(b) 驗證階段輸入因子個數對類神經網路推估誤差之影響

表 4 雲林縣沿海地區地下水中水質相關係數表

	砷	溫度	pH	電導度	溶氧	鹼度	TDS	氯鹽	硫酸鹽	游離氨 氮	COD	總有機 碳	鈉	鉀	鎂	鈣	硬度	
砷	1.00																	
溫度	-0.06	1.00																
pH	0.34	-0.03	1.00															
電導度	-0.29	-0.09	-0.48	1.00														
溶氧	-0.03	0.01	-0.09	-0.08	1.00													
鹼度	0.18	0.05	0.44	-0.26	-0.13	1.00												
TDS	-0.30	-0.09	-0.43	0.92	0.09	-0.21	1.00											
氯鹽	-0.32	-0.06	-0.48	0.95	0.12	-0.28	0.94	1.00										
硫酸鹽	-0.27	-0.08	-0.48	0.86	0.15	-0.34	0.83	0.89	1.00									
游離氨氮	0.13	0.00	0.02	0.11	0.08	0.29	0.19	0.16	-0.03	1.00								
COD	-0.09	0.12	0.02	0.10	-0.02	-0.01	0.12	0.16	0.19	0.02	1.00							
總有機碳	0.13	-0.12	0.31	-0.31	-0.02	0.44	-0.31	-0.34	-0.28	0.04	-0.13	1.00						
鈉	-0.30	-0.11	-0.45	0.95	0.08	-0.24	0.93	0.97	0.86	0.17	0.09	-0.31	1.00					
鉀	-0.25	-0.07	-0.35	0.80	0.15	-0.15	0.82	0.85	0.70	0.26	0.24	-0.28	0.86	1.00				
鎂	-0.19	0.01	-0.32	0.52	0.09	-0.17	0.52	0.56	0.44	0.19	0.00	-0.21	0.53	0.48	1.00			
鈣	-0.34	0.01	-0.53	0.72	0.10	-0.39	0.70	0.76	0.73	-0.04	0.05	-0.39	0.72	0.62	0.53	1.00		
硬度	-0.29	-0.04	-0.52	0.81	0.15	-0.30	0.79	0.86	0.77	0.12	0.02	-0.35	0.83	0.74	0.61	0.87	1.00	

經上述討論，區域水井水質類神經網路模式，以 2 個輸入層節點數為主要架構，分別採用 pH 值及電導度(EC)作為輸入層，地下水中砷濃度作為輸出層之推估項目，在資料上使用全區 6 座監測水井之資料，共 700 筆資料，作為網路訓練及驗證使用，在訓練及驗證過程中仿照單一水井模式，使用修正型目標函數及交叉驗證方法，並以試誤法決定網路中最佳隱藏層個數，經試誤法獲得最佳模式，修正型目標函數之係數(γ)為 0.2，在演算法上搭配 Levenberg-Marquardt 演算法，轉移函數則採用線性轉換函數，其架構如圖 7。

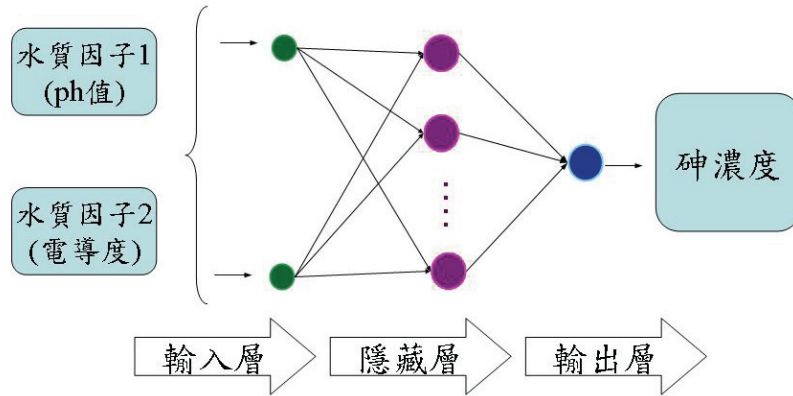


圖 7 區域水井水質類神經網路模式架構圖

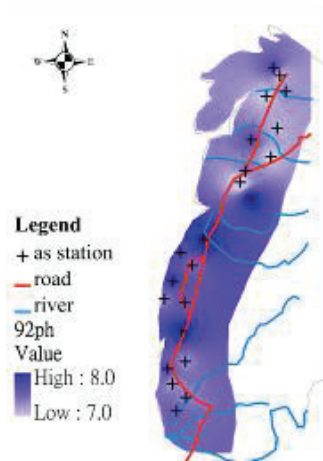
本研究應用 6 座監測井之 1992 年至 2005 年水質資料進行區域水井水質類神經網路模式訓練及驗證工作，並推估 1992、1994、1996 及 1998 年 22 座監測井之砷濃度，其結果如表 5，另為了解砷在地下水中空間分布情形，本研究應用地理資訊系統(Geographic Information system, 簡稱 GIS)具強大空間資料處理能力之軟體，將類神經網路推估 22 座監測井之地下水中實際水質分佈情形與推估砷濃度，藉由 GIS 提供空間推估模式轉換為雲林縣沿海地區砷污染分佈圖，如圖 8 及圖 9，圖 8 中分別顯示 1992 年~1998 年雲林縣沿海地區地下水中實際 pH 值、電導度(EC)及砷濃度，圖 9 為採 1992 年~1998 年雲林縣沿海地區地下水中實際 pH 值、電導度(EC)，應用區域水井水質類神經網路模式推估地下水中砷分布情形，圖 9 與實際砷濃度(圖 8)相比較可知模式之差誤值分布情形，及顯示本模式推估結果符合實際砷濃度趨勢變化，砷濃度仍以中部最高，在探討 pH 值、電導度(EC)及砷空間分布上，pH 值越高，砷濃度相對較高，在此分佈圖明顯可見，說明地下水水質趨向鹼性環境，較可能促進砷之生成；另外，電導度(EC)較高，則砷濃度反而較低，電導度(EC)代表水中重金屬濃度升高(如鈉、鉀、鎂、鈣及硬度)，也代表水質不佳，在重金屬濃度升高，水質不佳情形下，砷濃度反降低，本現象雖不如 pH 值變化明顯，但仍是值得本研究區域探討之課題。

探討模式誤差時，最佳推估誤差為 81ug/l，最差推估誤差為 129.9ug/l，但是對於雲林縣中部由於存在超高濃度#7 監測井，本監測井位於雲林縣中部，其地質、地下水水質及水文等不確定因素較高是造成模式推估誤差過大之主因，因

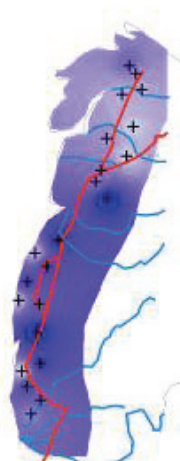
此，未來沿海地區之監測井中以雲林縣中部監測井較不適用模式推估其水質及砷濃度且應採實際監測方式取得水質資料，其餘雲林縣北區及南區監測井可利用模式建立 pH 值及電導度(EC)與砷濃度關係得到良好推估結果。

表 5 區域水井水質類神經網路模式之訓練、驗證及推估結果表

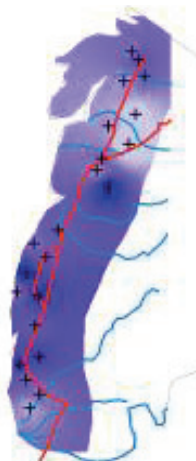
		rmse (ug/l)	誤差百分比
模式訓練		76.7	64.0%
模式驗證		3.8	23.7%
1992 年 5 月	推估	118.5	64.6%
1994 年 4 月	推估	129.9	73.0%
1996 年 4 月	推估	121.6	87.2%
1998 年 6 月	推估	81.0	90.6%



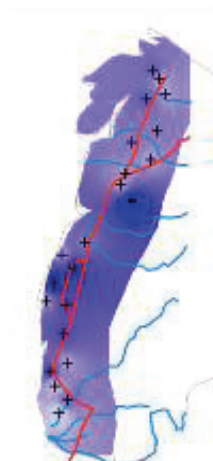
1992 年地下水之 pH 值



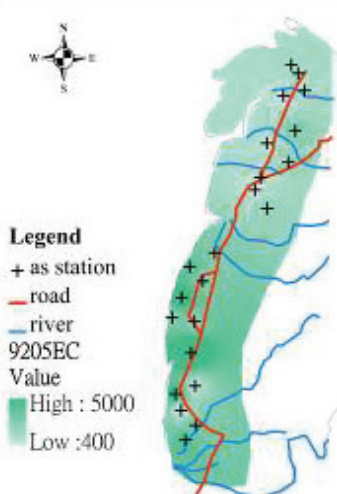
1994 年



1996 年



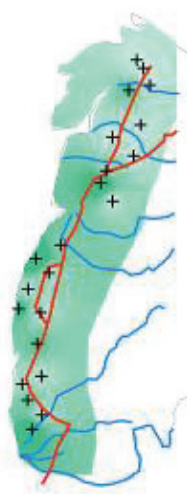
1998 年



1992 年地下水之電導度 (EC)



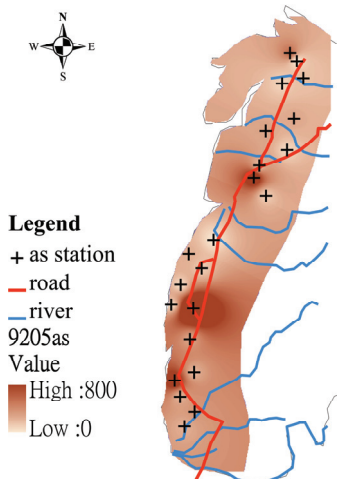
1994 年



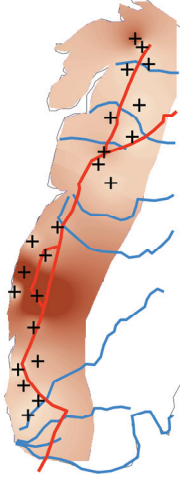
1996 年



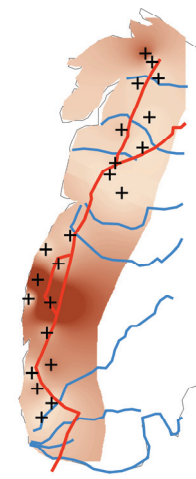
1998 年



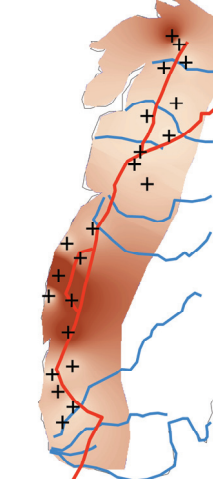
1992 年地下水之砷



1994 年

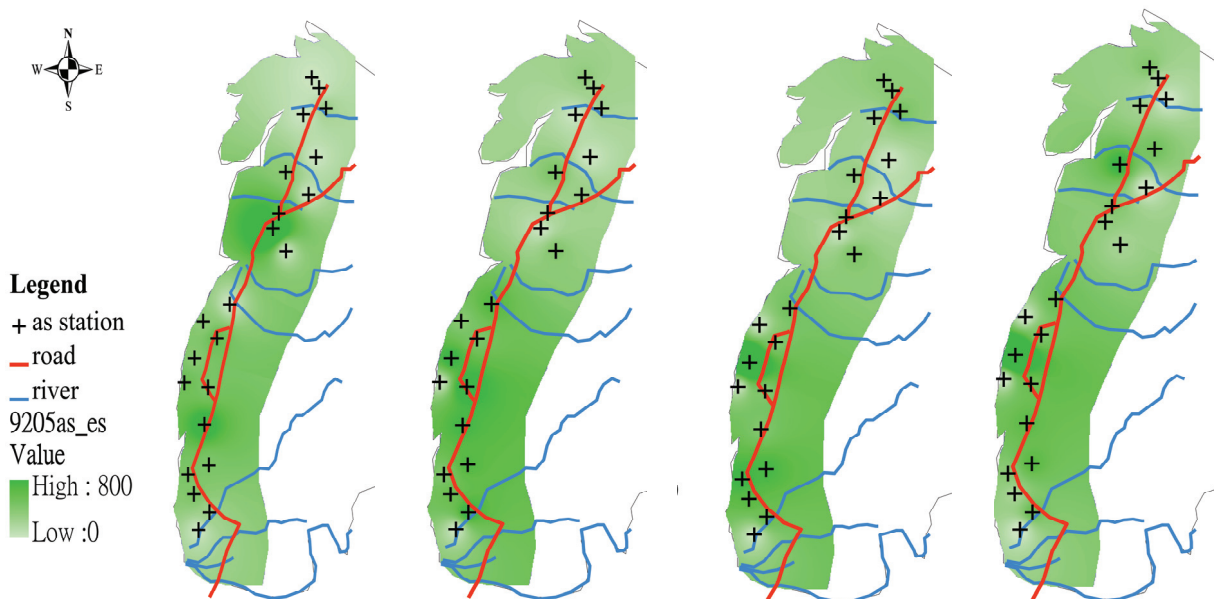


1996 年



1998 年

圖 8 1992~1998 年雲林縣地海區沿海地區水質分布情形圖



1992 年推估地下水之砷

1994 年

1996 年

1998 年

圖 9 1992~1998 年雲林縣地海區沿海地區類神經網路推估地下水中砷濃度分布圖

四、結論

本研究主要目的建立雲林縣沿海地區地下水中砷之水質模式，過去傳統模式僅能應用有限元素法或有限差分法，應用切割網格方法推估區域砷濃度，但是如本研究分析，砷在地下水具有複雜物理及化學作用主導，加以地下水升降影響，一般傳統水質模式或統計模式很難解決整合異質性資料模式之推估問題，本研究中應用類神經網路解決不同水質資料整合問題，在架構模式中可發現雲林縣沿海地區，砷濃度受到多重水質因子影響，將雲林縣沿海地區監測水井分為北中南三區，所得到模式輸入因子皆不同，可知影響砷濃度之因子變異性相當大，因此，在建立單一水井水質類神經網路模式之平均誤差大多介於可以接受範圍，但是建立區域水井水質類神經網路模式得到誤差值(rmse)較高，可了解建立本區全區模式之困難度，雖然如此，本研究區域大部分地區仍可採用本模式進行地下水中砷濃度推估工作，除雲林縣中部#7監測井附近，由於本區地下水水質環境較為獨立且特殊，造成模式推估結果不佳，建議本區可直接設置長期監測井方式獲得地下水中污染資料，總體而言，本模式推估之成果可協助地下水管理者及使用者有效掌控地下水中污染擴散情形，針對受污染較嚴重地區進行地下水使用管制及地下水污染防治工作，減少居民使用受砷污染地下水之風險。

謝誌

本研究承蒙國科會專案(NSC 95-2313-B-002-051-MY3)經費之補助，特此致

謝。

五、參考文獻

1. Abrahart, R. J., See, L. and Kneale, P. E., “New Tools for Neurohydrologists: Using Network Pruning and Model Breeding Algorithms to Discover Optimum Inputs and Architectures”, Proceedings of the 3rd International Conference on Geocomputation, University of Bristol, pp.17-19 September, 1998.
2. Chang, F. J. and Chen, Y. C., “A counterpropagation fuzzy-neural network modeling approach to real-time streamflow prediction”, Journal of Hydrology, Vol. 245, pp. 153-164, 2001.
3. Chang, F. J., Chang, L. C. and Huang, H. L., “Real-time recurrent neural network for stream-flow forecasting”, Hydrological Processes, Vol. 16, No. 13, pp. 2577-2588, 2002.
4. Chang, F. J. and Chang, Ya-Ting, “Adaptive neuro-fuzzy inference system for prediction of water level in reservoir”, Advances in Water Resources, Vol. 29, pp.1-10, 2006.
5. Chang, F. J., Chang, L. C. and Wang, Y. S., “Enforced Self-Organizing Map Neural Networks for River Flood Forecasting”, Hydrological Processes, Vol. 21, pp.741-749,2006.
6. Coppola, E., Szidarosky F., Poulton, M. and Charles, E., “Artificial Neural Network Approach for Predicting Transient Water Levels in a Multilayered Groundwater System under Variable State, Pumping, and Climate Conditions”, Journal of Hydrologic Engineering, Vol. 8, No. 6, pp. 348-360, 2003.
7. Gumrah, F., Oz, B., Guler, B. and Evin, S., “The Application of Artificial Neural Networks for the Prediction of Water Quality of Polluted Aquifer”, Water, Air and Soil Pollution, Vol. 119, pp. 275-294, 2000.
8. Kwok, T. Y. and Yeung, D. Y., “Constructive algorithms for structure learning in feedforward neural networks for regression problems”, IEEE

- Transaction on Neural Networks, Vol. 8, pp. 630-645, 1997.
9. Liu, C. W., Jang, C. S., and Liao, C. M., "Evaluation of arsenic contamination potential using indicator kriging in the Yun-Lin aquifer (Taiwan) ", Science of the Total Environment ,Vol.321(1-3), pp.173-188, 2004.
 10. Liu, C.W., Wang, S.W., Jang, C.S. and Lin, K.H., "Occurrence of Arsenic in Ground Water in the Choushui River Alluvial Fan, Taiwan", Journal of Environmental Quality, Vol.35, pp. 68-75, 2006.
 11. Lin, Y. B., Lin, Y. P., Liu, C. W. and Tan, Y. C., "Mapping of spatial multi-scale sources of arsenic variation in groundwater on ChiaNan floodplain of Taiwan", Science of the Total Environment, Vol.370, pp. 168-181, 2006.
 12. Martos, F. S., Aguilera, P. A., Garrido-Frenich, A., Torres, J. A. and Pulido-Bosch, A., "Assessment of Groundwater Quality by Means of Self-Organizing Map: Application in a Semiarid Area", Environment Management, Vol. 30, No.5, pp.716-726, 2002.
 13. Madhumita D. A., Kumar M. M., Muduli S. D.," Evaluation of drinking quality of groundwater through multivariate techniques in urban area " Environ Monit Assess Vol. 166, pp.149-157,2010.
 14. Mohammad T. D., Alireza M., Jamshid P. ,and Miguel R.-R.," Application of ANN and ANFIS models for reconstructing missing flow data", Vol. 166, pp.421-434,2010.
 15. Nolan, B. T., Arthur, L. Baehr and Leon, J. Kauffman, "Spatial Variability of Groundwater Recharge and its Effect on Shallow Groundwater Quality in Southern New Jersey", Vadose Zone Journal, Vol. 2, issue 4, pp. 677-691, 2003.
 16. Ruiz, F., Gomis, V. and Blasco, P., "Application of factor analysis to the hydrogeochemical study of a coastal aquifer", Journal of Hydrology, Vol. 119, pp. 169-177, 1990.

17. 王榮德、胡賦強、江錦華、吳新英，「飲用水改善前後，烏腳病發生率之研究」，烏腳病之研究報告，第 18 輯，台灣省烏腳病防治中心，第 21-25 頁，1983。
18. 郭益銘，「應用多變量統計分析與類神經網路分析雲林沿海地區地下水水質變化」，國立台灣大學農業工程研究所碩士論文，1999。
19. 張斐章、胡湘帆、黃源義，「反傳遞模糊類神經網路於流量推估之應用」，農業工程學報，第四十四卷，第二期，第 26-38 頁，1998。
20. 張斐章、孫建平，「類神經網路及其應用於降雨-逕流過程之研究」，中國農業工程學報，第四十三卷，第一期，第 34-49 頁，1997。
21. 張斐章、張麗秋：類神經網路，東華書局，2005。
22. 陳韋佑，「群策倒傳遞類神經網路之建構與應用」，逢甲大學土木及水利工程研碩士論文，1993。
23. 曾文賓，「烏腳病之診斷、治療與預防」，台灣省政府烏腳病防治中心，第一輯，第 1-25 頁，1976。
24. 廖少威，「以主成份分析法探討台灣灌溉水質受工業、都市、畜牧廢水污染之變異分析」，國立台灣大學農業工程研究所碩士論文，1991。
25. 工業技術研究院能源與資源研究所，「臺灣地區地層下陷之監測、調查及分析(2/4)」，經濟部水利署，2002。
26. 工業技術研究院能源與資源研究所，「台灣地區地層下陷之監測、調查及分析(3/4)」，經濟部水利署，2003。
27. 羅華強：類神經網路-MATLAB 的應用，高立圖書有限公司，2005。