

以多變數迴歸與類神經網路模式模擬水庫水質

Application of Multivariate Linear Regression and Neural Network Models for Simulating Water Quality in Reservoir

國家災害防救科技中心
助理研究員

國立聯合大學土木與防災工程學系
教授
暨國家實驗研究院台灣颱風洪水研究中心
副主任

陳偉柏

Wei-Bo Chen

柳文成*

Wen-Cheng Liu

摘 要

水質為水庫操作與管理重要考量的因子，而卡爾森優養指標考量的水質變數包括溶氧、總磷、葉綠素 a 及透明度為決定水庫水質優養程度的指標。本研究以類神經網路模式(包括輻射基底類神經網路及調適性模糊類神經網路)與多變數線性迴歸模式以預測台灣中部德基水庫之溶氧、總磷、葉綠素 a 及透明度。類神經模式與多變數線性迴歸模式之輸入水質變數則由線性迴歸後之相關係數決定。為評估水質預測之優劣，以統計誤差包括平均絕對誤差、均方根誤差及相關係數判定。水質預測結果顯示，調適性模糊類神經網路模式優於多變數線性迴歸模式及輻射基底類神經網路模式，調適性模糊類神經網路模式可以較準確的預測水庫之溶氧、總磷、葉綠素 a 及透明度；因此，調適性模糊類神經網路模式可被應用於水庫水質管理之較佳工具。

關鍵詞：多變數線性迴歸，輻射基底類神經網路，調適性模糊類神經網路，優養，水質，德基水庫。

ABSTRACT

The water quality is one of the key factors in operation and management of reservoirs. The Carlson index of dissolved oxygen (DO), total phosphorus (TP), chlorophyll a (Chl a), and secchi disk depth (SD) are commonly used as the indicators for

*通訊作者，國立聯合大學土木與防災工程學系教授暨國家實驗研究院台灣颱風洪水研究中心副主任，36063 苗栗市南勢里聯大二號，wcliu@nuu.edu.tw

determining the status of eutrophication in the reservoirs. In this study, various artificial neural network models (i.e. radial basis function neural network, RBFN and adaptive neuro-fuzzy inference system approach, ANFIS) and multilinear regression (MLR) model were developed to predict DO, TP, Chl *a*, and SD in the Techu Reservoir of Central Taiwan. The input variables of the neural network and MLR models are determined using the linear regression. The performances using RBFN, ANFIS, and MLR models were evaluated based upon the statistical errors including mean absolute error, root mean square error, and correlation coefficient computed from the measured and model simulated DO, TP, Chl *a*, and SD values. The results indicate that ANFIS model's performances are superior to those of MLR and RBFN models. Study results show that the neural network using ANFIS model is suitable to predict water quality variables with reasonable accuracy, suggesting that the ANFIS model can be used as a valuable tool for reservoir management in Taiwan.

Keywords: Multivariate linear regression, Radial basis function neural network, Adaptive network-based fuzzy inference system, Eutrophication, Water quality, Techu Reservoir.

一、前言

1.1 研究背景與目的

隨著工商業的發達，時代的進步，水，在人類的生活中日趨重要。在台灣冬季的東北季風與夏季的西南氣流以及颱風為降雨的主要來源，年降雨量約為世界平均值的 2.6 倍，但因台灣的地形山勢、空間與時間分布及人口眾多的關係，每人可得到的水資源卻只有世界平均值的 1/7，為世界水資源貧乏區。當乾旱一發生時，仍然有缺水的危機，所以，政府為了解決這個重要的民生問題，開發興建了許多大大小小的水庫。但由於水庫集水區上游的管理不當及暴雨沖刷，導致大量的營養鹽流入水庫，使得水庫藻類滋生，水質產生惡化，即為水庫優養化。所以控制營養鹽(氮與磷)流入水庫，才是提升水庫水質的根本方法。

以數值模式預測具富營養鹽水體的水質狀況是極具困難性，因為水體中複雜的物理、化學及生物作用。質量平衡為基礎的湖泊或水庫優養模擬起源於 40 多年前(Chen, 1970; Di Toro *et al.*, 1971; Thomann *et al.*, 1975)，為瞭解優養化的過程及電腦計算能力的進步，多維的湖泊與水庫的

水理與水質模式已陸續被研發並應用於預測與管理水質的問題(Thomann and Mueller, 1987; Chapra, 1997, Martin and McCutcheon, 1999)。在這些研究中顯示，水理與水質模式是非常複雜，而且需要大量的實測資料進行模式的校驗證。近年來，非線性系統的類神經網路模式已廣泛應用於表面水體的水質模擬與預測(Dogan *et al.*, 2009; Singh *et al.*, 2009; Khataee *et al.*, 2010; Chen *et al.*, 2010; Jiang *et al.*, 2013; Najah *et al.*, 2013; Wen *et al.*, 2013; Xiang *et al.*, 2013; Abyaneh, 2014; Ay ad Kisi, 2014; Kim *et al.*, 2014; Najah *et al.*, 2014)。

本研究目的應用多變數線性迴歸與類神經網路模式(幅狀基底類神經網路及調適性模糊類神經網路)模擬台灣中部德基水庫之水質參數，包括：溶氧、透明度、總磷及葉綠素 *a*，此四種水質變數為判斷水庫優養狀況的重要指標，另以誤差統計以分析模擬預測之水質與實測水質之誤差值，以提出建議適用於德基水庫水質模擬預測之非線性類神經網路模式。

1.2 文獻回顧

國內近年來有前人對於水體水質預測有相

關之研究；謝明翰(2002)利用倒傳遞類神經網路對水庫水質進行時間上連續的預測；王英義(2005)建立類神經網路水質(總磷與葉綠素 *a*)模式；張麒顯(2008)以類神經網路模式與衛星影像於石門水庫水質之推估；梁惠儀等(2009)結合類神經網路與遺傳演算於翡翠水庫最佳控制優養；吳冠儀(2013)利用 CE-QUAL-W2 水理水質模式模擬水庫內營養鹽及藻類生長情形，最後利用類神經網路取代複雜的 CE-QUAL-W2 水庫模式，較簡便地推估水庫營養鹽產量及藻類生長情形。

國外有關於類神經網路模式應用於湖泊或水庫的模擬預測，則有 Soyupak *et al.* (2003)應用類神經網路模式計算土耳其 Keban Dam 水庫、Kuzgun Dam 水庫及 Doganci Dam 水庫於穩態的時間與空間的溶氧濃度；Kuo *et al.* (2006)結合類神經與基因演算法應用於翡翠水庫集水區營養鹽控制的預測對於水質的影響；Kuo *et al.* (2007)利用倒傳遞類神經網路模式預測影響德基水庫的四種重要水質參數，為溶氧、總磷、葉綠素 *a* 及透明度；Rankovic *et al.* (2010)應用前向類神經網路模式預測賽爾維亞 Gruza 水庫的溶氧濃度；Soltani *et al.* (2010)結合水質模擬模式與基因演算法用以決定不同水庫出口的最佳操作策略，其中水質模擬模式為調適性模糊類神經網路 (ANFIS)；Akkoyunlu *et al.* (2011)應用倒傳遞類神經網路模式、幅射基底類神經網路及多變數線性迴歸於預測土耳其 Lake Iznik 的溶氧濃度；Rankovic *et al.* (2012)再以調適性模糊類神經網路預測賽爾維亞 Gruza 水庫的溶氧濃度；Chen and Liu (2014)以多變數線性迴歸、倒傳遞類神經網路模式以及調適性模糊類神經網路模式預測翡翠水庫的溶氧濃度。其中，Kuo *et al.* (2007)僅應用單一類神經網路模式進行德基水庫水質預測，缺少利用不同類神經網路模式間之水質預測能力比較；Chen and Liu (2014)雖然使用三種模式預測翡翠水庫溶氧，並比較三種模式間之溶氧預測能力，然而只針對單一水質項目進行比較，缺少模式預測不同水質變數之結果。本研究不僅利用不同的類神經網路模式預測水同的水質項目

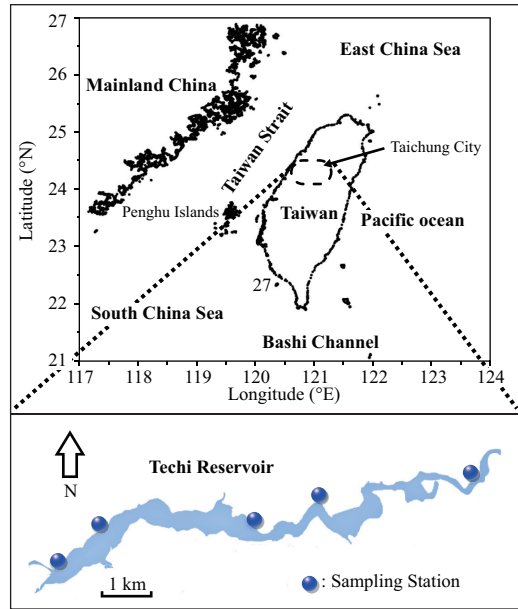


圖 1 德基水庫之位置圖及水質採樣點

(溶氧、透明度、總磷及葉綠素 *a*)，並比較其預測能力。

由以上的文獻資料可知，類神經網路模式應用於水庫水質參數的模擬預測，並可做為水庫水質管理之預測工具，但至於何種類神經網路較適用於水庫水質參數的模擬預測，因為各水庫的水質特性不同，則需要予以詳細研究才能確定。

二、研究區域概述及資料蒐集分析

2.1 研究區域概述

德基水庫又名達見水庫，位於德基溪之下游屬於台中市和平區(見圖 1 所示)，為台灣的第四大水庫，水庫最大的表面積為 4.54 km²，初始設計之蓄水體積約為 232 × 10⁶ m³，年入流量為 1.2 × 10⁹ m³，約為 5 倍大的水庫體積，但四分之三流量來自於雨季。德基水庫集水區為 592 km²，集水區的高度變化從標高 3,884 公尺至 1,408 公尺(正常水位)，高度變化超過 2,400 公尺，主流坡度大部分超過 50%，平均坡度通常超過 30%。

德基水庫壩型為雙曲線薄型混凝土拱壩，壩身高 180 公尺，為台灣最高壩，並為世界十大高

表 1 德基水庫水質變數之統計資料

變數	最小值	最大值	平均值	標準偏差	變異數
透明度(m)	0.05	8.40	3.32	1.76	0.53
水溫(°C)	10.90	25.20	17.22	3.09	0.18
pH	6.80	9.87	8.16	0.46	0.06
電導度(μmho/cm 25°C)	157.00	366.00	250.31	31.96	0.13
溶氧(mg/L)	0.30	12.60	7.51	1.79	0.24
濁度(NTU)	0.10	550.00	7.02	26.18	3.73
懸浮固體物(mg/L)	0.10	702.00	8.29	31.23	3.77
總硬度(mg/L)	19.00	167.00	119.95	16.45	0.14
總鹼度(mg/L)	58.70	123.07	86.45	12.07	0.14
葉綠素 <i>a</i> (μg/L)	0.10	126.10	5.47	11.46	2.10
總磷(μg/L)	2.00	266.00	18.21	24.81	1.36
硝酸氮(mg/L)	0.01	0.95	0.07	0.09	1.23
生化需氧量(mg/L)	0.50	92.40	6.13	6.62	1.08

註：變異數=標準偏差/平均值

壩之一，水庫容量超過 2 億立方公尺，為中部地區最大者。德基水庫於民國 58 年開國際標發包，為國內之創舉，於民國 63 年完成。德基電廠是台灣電力公司大甲溪電力開發的骨幹，共裝有三部發電機，其規模僅次於青山電廠，設有台灣最高電梯，高 222 公尺。本水庫排洪設施除壩頂溢洪道外，另設有排洪隧道，洪水經 650 公尺隧道排洩於青山水庫下游。

德基水庫水源除大甲溪外，並自志樂溪越域引水。水庫直接功能為發電，年平均發電量 3.6 億度，所調蓄水量經放流青山等電廠逐步發電後，於石岡壩引取供給水與灌溉使用；本水庫並具防洪與觀光功能，為多標的水庫(經濟部水利署網站)。

2.2 水庫水資料蒐集與分析

本研究蒐集德基水庫 1994 年至 2013 年每季水量測一次，5 監測站(見圖 1)，共 20 年的水質監測資料，包括透明度、水溫、pH 值、電導度、溶氧、濁度、懸浮固體濃度、總硬度、總鹼度、葉綠素 *a*、總磷、硝酸氮、生化需氧量等 13 種物理與化學性的水質，經過統計分析後，獲得 13 種水質監測項目之最大值、最小值、平均值、標準偏差及變異數如表 1 所示，顯示這 13 種水質

監測的變化極大，以葉綠素 *a* 與溶氧濃度之變化分別可介於 0.1~126.1 μg/L 及 0.3~12.6 mg/L。

Carlson (1977)建議於判斷水庫水質優劣的指標，稱為卡爾森指標，此指標是以透明度、溶氧、葉綠素 *a* 及總磷等水質變數做為綜合性的判斷，台灣水庫即採用卡爾森指標來判斷水庫水質是否為優養、中養或貧養狀況。圖 2 為德基水庫優養指標隨時間的變化，圖中顯示水庫大部分時間處於中養(Mesotrophic)與貧養(Oligotrophic)狀況，2013 年水質則為中養至優養(Eutrophic)，表示 2013 年水質有變壞的趨勢。

三、模式理論與方法

3.1 多變數線性迴歸模式

多變數線性迴歸用以模擬相依變數與一個或多個獨立變數間的線性關係，相依變數為被預測值，而獨立變數為預測值(Predictor)。多變數線性迴歸模式可以用觀測值與被預測值間之最小平方來建立(Weisberg, 1985)。該方程式可寫成：

$$y = b_0 + b_1x_1 + b_2x_2 + \dots + b_ix_i \dots\dots\dots (1)$$

式中， x_i 為第 i 項獨立變數； b_0 為迴歸常數； b_i 為第 i 項獨立變數的迴歸係數。

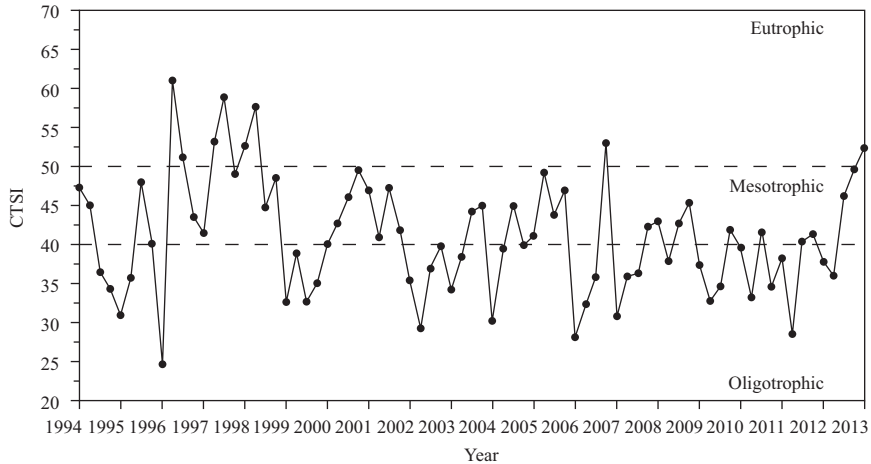


圖 2 德基水庫之優養指標

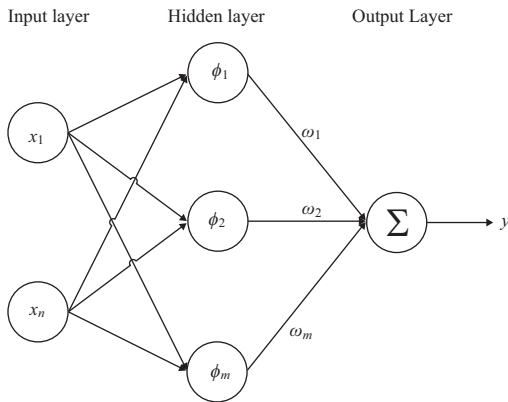


圖 3 幅射基底類神經網路模式路架構圖

3.2 幅狀基底函數類神經網路模式

幅狀基底函數類神經網路為以函數逼近 (curve fitting) 的方式在輸入層、隱藏層以及輸出層間進行線性映射與組合，如圖 3 所示；輸入層是輸入訓練範本進入網路之介面層，經非線性活化函數轉換至隱藏層，最後由輸出層將隱藏層之輸出作線性組合以得到輸出值；通常需建構高維度之隱藏層架構來獲得更加精確之預估，選取中心點並無特定之方法，因此一般採用複合式學習法，在模式前階段以非監督式學習法獲得中心點，模式後階段則以監督式學習法調整連結權重 (Dibike *et al.*, 1999; 張斐章與張麗秋, 2005)。

針對單一輸出的幅狀基底函數類神經網路

模式於 m 個隱藏層可以表示為：

$$y = \sum_{i=1}^m \omega_i \phi_i(x) \quad \dots\dots\dots (2)$$

式中 x 為輸入向量； ω_i 為連接隱藏層與輸出層間的權重； $\phi_i(x)$ 為經過幅射基底函數轉換後於隱藏層的神經元輸出值。本研究以高斯函數作為幅射基底函數，如下式所示：

$$\phi_i(x) = \exp\left(-\frac{\|x - c_i\|^2}{2\sigma^2}\right) \dots\dots\dots (3)$$

式中 c_i 為隱藏層中第 i 個神經元之中心向量； σ 定義為 $\sigma = d_{\max} / \sqrt{nc}$ ； d_{\max} 為於 c_i 中的最大距離； nc 為中心向量之個數； $\|x - c_i\|$ 為 x 與 c_i 間的歐基里德距離 (Euclidean distance)。決定 nc 所採用的方法為正交最小平方方法 (Chen *et al.*, 1991; Ham and Kostanic, 2001; Kecman, 2001)。

3.3 調適性模糊類神經網路模式

類神經網路雖然具有極佳之自我組織與推理能力，卻無法處理定性知識與邏輯推論。因此，Jang (1993) 以模糊推論系統 (Fuzzy Inference System, FIS) 為網路模式基礎，並結合調適性網路自我組織特性，提出調適性模糊類神經網路模式。Yang *et al.* (2000) 提出模糊控制器系統，係以 IF-THEN 規則補充了對問題的不確定性與學習

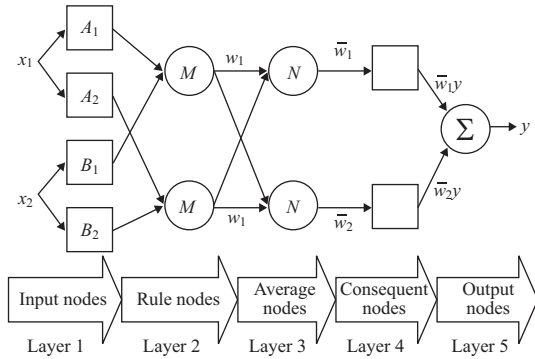


圖 4 調適性模糊類神經網路模式架構圖

過程中之認知能力，與類神經網路演算法結合後，使模式具有更佳之處理能力、自我學習與組織能力，進而強化調適性模糊類神經網路模式。其架構除輸入層與輸出層外，多了模糊化之規則層、正規化層與解模糊化之結論推論層，如圖 4 所示。

調適性模糊類神經網路係採用監督式學習的前饋式類神經網路模式之學習架構，於學習訓練之過程中，有許多參數可作為調整，在輸入層進入模糊集合之前可設定每一種輸入變數之集合數目，進而排列組合出所有可能之規則，但過多增加規則庫之數目則會使非線性最佳化問題越加複雜，造成收斂速度減慢且收斂性差，並且增加運算時間。

第一層：將輸入層之神經元分別映射至模糊集合中，而模糊集合內則個別有相對應之隸屬函數，通常採用鐘型函數，如下表示：

$$Q_{1,i} = \mu_{A_i}(x_1), \quad i=1,2 \dots\dots\dots(4)$$

$$Q_{1,i} = \mu_{B_{i-2}}(x_2), \quad i=3,4 \dots\dots\dots(5)$$

式中 x_1 及 x_2 為輸入至神經元 i ； A_i 及 B_i 為下標符號； μ_{A_i} 及 $\mu_{B_{i-2}}$ 為隸屬函數。

本研究採用鐘型函數為隸屬函數：

$$\mu_{A_i} = \frac{1}{1 + \left| \frac{x_1 - c_i}{a_i} \right|^{2b}} \mu_{B_{i-2}} = \frac{1}{1 + \left| \frac{x_2 - c_i}{a_i} \right|^{2b}} \dots\dots\dots(6)$$

式中 a_i ， b_i 及 c_i 為參數。

第二層：將各輸入變數間的模糊集合進行排

列組合之配對後，進行模糊邏輯運算得到權重，其計算式如下：

$$O_{2,i} = \mu_{A_i}(x_1) \cdot \mu_{B_i}(x_2) = w_i, \quad i=1,2 \dots\dots\dots(7)$$

第三層：將第二層各節點所得的結果進行正規化，使其輸出值可介於 0 到 1 之間，其計算式如下：

$$O_{3,i} = \bar{w}_i = \frac{w_i}{w_1 + w_2}, \quad i=1,2 \dots\dots\dots(8)$$

第四層：將第三層獲得之正規化結果與 Sugeno 模糊模式相乘，其計算式如下：

$$O_{4,i} = \bar{w}_i y_i = \bar{w}_i (p_i x + q_i y + r_i) \quad \text{for } i=1,2 \dots\dots(9)$$

式中 p_i ， q_i ， r_i 為線性參數或稱結論部分之參數。

第五層：輸出層總和上層之輸出結果並計算出整體之輸出值，如式(10)所示。

$$O_{5,1} = \sum_{i=1}^2 \bar{w}_i y_i = \frac{\sum_{i=1}^2 w_i y_i}{\sum_{i=1}^2 w_i} \dots\dots\dots(10)$$

3.4 統計誤差指數

為評估多變數線性迴歸模式與類神經網路模式，三種統計誤差分析用以比較模式預測結果與觀測值，包括平均絕對誤差(Mean absolute error, MAE)、均方根誤差(Root mean square error, RMSE)以及相關係數(Correlation coefficient, R)。其計算公式分別說明如下：

$$MAE = \frac{1}{N} \sum_N |(C_p)_N - (C_m)_N| \dots\dots\dots(11)$$

$$RMSE = \sqrt{\frac{1}{N} \sum_N [(C_p)_N - (C_m)_N]^2}, \dots\dots\dots(12)$$

$$R = \frac{\sum_N [(C_p)_N - \bar{C}_p][(C_m)_N - \bar{C}_m]}{\sqrt{\sum_N [(C_p)_N - \bar{C}_p]^2 \sum_N [(C_m)_N - \bar{C}_m]^2}} \dots\dots(13)$$

式中 N 為資料總筆數； C_p 為預測的水質變數； C_m 為實測之水質變數； \bar{C}_p 與 \bar{C}_m 分別表示預測與實測水質變數之平均值。

表 2 不同水質變數與總磷之相關係數

水質變數	相關係數(R)
透明度(SD)	0.2906
水溫(WT)	0.0519
pH	0.1109
電導度(EC)	0.2248
濁度(TB)	0.3646
溶氧(DO)	0.0023
懸浮固體物(SS)	0.4033
葉綠素 <i>a</i> (Chl- <i>a</i>)	0.2534
生化需氧量(BOD)	0.1539
硝酸氮(NO ₃ -N)	0.5685
總硬度(TH)	0.1875
總鹼度(TA)	0.2264

表 4 不同水質變數與溶氧之相關係數

水質變數	相關係數(R)
總磷(TP)	0.0023
水溫(WT)	0.2377
pH	0.4595
電導度(EC)	0.1679
濁度(TB)	0.0019
透明度(SD)	0.2831
懸浮固體物(SS)	0.0054
葉綠素 <i>a</i> (Chl- <i>a</i>)	0.2895
生化需氧量(BOD)	0.0201
硝酸氮(NO ₃ -N)	0.0522
總硬度(TH)	0.2010
總鹼度(TA)	0.2014

表 3 不同水質變數與透明度之相關係數

水質變數	相關係數(R)
總磷(TP)	0.2906
水溫(WT)	0.2184
pH	0.2145
電導度(EC)	0.1566
濁度(TB)	0.2369
溶氧(DO)	0.2831
懸浮固體物(SS)	0.2104
葉綠素 <i>a</i> (Chl- <i>a</i>)	0.2013
生化需氧量(BOD)	0.0452
硝酸氮(NO ₃ -N)	0.1556
總硬度(TH)	0.1132
總鹼度(TA)	0.1366

表 5 不同水質變數與葉綠素 *a* 之相關係數

水質變數	相關係數(R)
總磷(TP)	0.2534
水溫(WT)	0.2114
pH	0.2755
電導度(EC)	0.2440
濁度(TB)	0.0258
透明度(SD)	0.2013
懸浮固體物(SS)	0.0032
溶氧(DO)	0.2895
生化需氧量(BOD)	0.2042
硝酸氮(NO ₃ -N)	0.1346
總硬度(TH)	0.2299
總鹼度(TA)	0.2859

四、結果與討論

4.1 選擇輸入之水質參數

選擇適當的輸入因子為以多變數線性迴歸模式、幅狀基底函數類神經網路模式及調適性模糊類神經網路模式預測水庫水質變數重要的一環；Akkoyunlu *et al.* (2011)與 Rankovic *et al.* (2012)已使用不同的水質變數做為類神經網路模式之輸入，以進行水庫水質之模擬預測。Chang *et al.* (2014)應用非線性因子選取工具 Gamma Test (GT)，選擇適當的相依變數為輸入，建立 ANFIS (GT)用以評估臺灣北部磺港溪之砷濃度；Chang *et al.* (2015)同樣以 GT 為工具，選擇主要因子為

類神經網路模式之輸入，模擬大漢溪氨氮(NH₃-N)的濃度。

為瞭解輸入水質因子對於輸出結果的相對重要性與貢獻度，以及簡化選擇輸入因子的程序，本研究先採用線性迴歸法，探討不同的相依變數與水質變數(溶氧、總磷、葉綠素 *a*、透明度)。從行政院環境保護署蒐集 1994 年至 2013 年的德基水庫水質資料共計 715 筆，做為線性迴歸之依據，表 2 至表 5 分別為個別相依變數與總磷、透明度、溶氧以及葉綠素 *a* 之相關係數值(R)。為使三種模式對四種水質變數之預測結果可互相比較，且避免過度擬合(overfitting)與過度訓練(overtraining)狀況發生，本研究選用與欲預測之

表 6 多變數線性迴歸模式預測不同水質變之誤差分析

水質變數	訓練階段			測試階段		
	MAE	RMSE	R	MAE	RMSE	R
DO	1.07 mg/L	1.51 mg/L	0.52	1.16 mg/L	1.55 mg/L	0.53
TP	8.51 µg/L	15.23 µg/L	0.74	9.56 µg/L	20.43 µg/L	0.71
Chl- <i>a</i>	5.15 µg/L	10.96 µg/L	0.48	4.74 µg/L	8.01 µg/L	0.43
SD	1.28 m	1.54 m	0.51	1.22 m	1.49 m	0.45

水質變數 R 值前 6 高之個別相依變數為模式輸入。表 2 顯示影響總磷的最重要水質因子為透明度(SD)、濁度(TB)、懸浮固體物(SS)、葉綠素 *a* (Chl-*a*)、硝酸氮(NO₃-N)及總鹼度(TA)；表 3 可知影響透明度之重要水質因子為總磷(TP)、水溫(WT)、pH、濁度(TB)、溶氧(DO)及懸浮固體物(SS)；由表 4 可得影響溶氧的重要水質因子為水溫(WT)、pH、透明度(SD)、葉綠素 *a*、總硬度(TH)及總鹼度(TA)；表 5 顯示影響葉綠素 *a* 之重要水質因子有總磷(TP)、pH、電導度(EC)、溶氧(DO)、總硬度(TH)及總鹼度(TA)。

4.2 多變數線性迴歸模式

多變數線性迴歸模式用以預測溶氧(DO)、總磷(TP)、葉綠素 *a* (Chl-*a*)及透明度(SD)，由 4.1 節的輸入水質變數之選擇，可建立多變數線性迴歸模式見式(14)至(17)所示。

$$DO = -0.014 \times WT + 1.5207 \times pH - 0.1692 \times SD + 0.0199 \times Chl - a - 0.0035 \times TH - 0.0099 \times TA - 2.9313 \dots\dots\dots (14)$$

$$TP = -1.6142 \times SD - 0.4637 \times TB + 0.6821 \times SS + 0.2948 \times Chl - a + 143.8025 \times NO_3 - N - 0.0762 \times TA + 15.8915 \dots\dots\dots (15)$$

$$Chl - a = 0.1085 \times TP + 5.2206 \times pH - 0.0221 \times EC + 1.0238 \times DO + 0.0179 \times TH - 0.1377 \times TA - 31.5261 \dots\dots\dots (16)$$

$$SD = -0.0215 \times TP - 0.0891 \times WT - 0.2783 \times pH - 0.0702 \times TB - 0.2035 \times DO + 0.0525 \times SS + 9.0964 \dots\dots\dots (17)$$

式中 EC 為電導度；NO₃-N 為硝酸氮；pH 為 pH 值；SS 為懸浮固體物；TA 為總鹼度；TB 為濁度；TH 為總硬度；WT 為水溫度。

方程式(14)是用以預測溶氧，於預測過程中分為訓練階段與測試階段；訓練階段共有資料 500 筆(佔總資料筆數之 70%)，測試階段共有資料 215 筆(佔總資料筆數之 30%)。表 6 為以多變數線性迴歸模式預測不同水質變數於訓練階段與測試階段之統計誤差分析結果，表中顯示於訓練階段預測溶氧之 MAE、RMSE 及 R 值分別為 1.07 mg/L、1.51 mg/L 及 0.52；於測試階段預測溶氧之 MAE、RMSE 及 R 值分別為 1.16 mg/L、1.55 mg/L 及 0.53。對於總磷之預測而言，於訓練階段預測總磷之 MAE、RMSE 及 R 值分別為 8.51 µg/L、15.23 µg/L 及 0.74；於測試階段預測總磷之 MAE、RMSE 及 R 值分別為 9.56 µg/L、20.43 µg/L 及 0.71。對於葉綠素 *a* 之預測而言，於訓練階段預測葉綠素 *a* 之 MAE、RMSE 及 R 值分別為 5.15 µg/L、10.96 µg/L 及 0.48；於測試階段預測葉綠素 *a* 之 MAE、RMSE 及 R 值分別為 4.74 µg/L、8.01 µg/L 及 0.43。對於透明度之預測而言，於訓練階段預測透明度之 MAE、RMSE 及 R 值分別為 1.28 m、1.54 m 及 0.51；於測試階段預測透明度之 MAE、RMSE 及 R 值分別為 1.22 m、1.49 m 及 0.45。

Akkoyunlu *et al.* (2011)應用類神經網路模式與多變數線性迴歸模式預測土耳其 IZnik 湖泊的

表 7 幅狀基底函數類神經網路模式預測不同水質變之誤差分析

水質變數	訓練階段			測試階段		
	MAE	RMSE	R	MAE	RMSE	R
DO	0.84 mg/L	1.17 mg/L	0.74	0.89 mg/L	1.26 mg/L	0.75
TP	6.93 μ g/L	11.25 μ g/L	0.87	8.83 μ g/L	18.44 μ g/L	0.76
Chl- <i>a</i>	3.44 μ g/L	5.57 μ g/L	0.87	3.50 μ g/L	5.75 μ g/L	0.88
SD	0.63 m	0.84 m	0.88	0.71 m	0.91 m	0.84

表 8 類神經網路模式之相關輸入變數

類神經網路模式種類	相關變數			
	輸入變數數量	輸出變數數量	規則數	準確率
RBFN	6	1	-	0.85
ANFIS	6	1	2	-

-：表示無須輸入此變數

溶氧，發現以多變數線性迴歸模式預測溶氧的精確度不高。本研究以多變數線性迴歸模式預測德基水庫的溶氧、總磷、葉綠素 *a* 及透明度，預測的準確度均偏低。

4.3 幅狀基底函數類神經網路模式

由於多變數線性迴歸模式無法精確預測水庫水質，類神經網路模式嘗試應用於水庫水質之預測。首先採用幅狀基底函數類神經網路模式進行水庫溶氧、總磷、葉綠素 *a* 及透明度之預測，此模式於訓練與測試階段所採用的資料筆數與前節多變數線性迴歸模式相同。

表 7 為以幅狀基底函數類神經網路模式預測不同水質變數於訓練階段與測試階段之統計誤差分析結果，表中顯示於訓練階段預測溶氧之 MAE、RMSE 及 R 值分別為 0.84 mg/L、1.17 mg/L 及 0.74；於測試階段預測溶氧之 MAE、RMSE 及 R 值分別為 0.89 mg/L、1.26 mg/L 及 0.75。對於總磷之預測而言，於訓練階段預測總磷之 MAE、RMSE 及 R 值分別為 6.69 μ g/L、11.25 μ g/L 及 0.87；於測試階段預測總磷之 MAE、RMSE 及 R 值分別為 8.83 μ g/L、18.44 μ g/L 及 0.76。對於葉綠素 *a* 之預測而言，於訓練階段預測葉綠素 *a* 之 MAE、RMSE 及 R 值分別為 3.44 μ g/L、5.57 μ g/L 及 0.87；於測試階段預測葉綠素 *a* 之 MAE、

RMSE 及 R 值分別為 3.50 μ g/L、5.75 μ g/L 及 0.88。對於透明度之預測而言，於訓練階段預測透明度之 MAE、RMSE 及 R 值分別為 0.63 m、0.84 m 及 0.88；於測試階段預測透明度之 MAE、RMSE 及 R 值分別為 0.71 m、0.91 m 及 0.84。比較表 6 與表 7 可以發現，幅狀基底函數類神經網路模式所預測之不同水質變數均優於多變數線性迴歸模式之預測結果。幅狀基底函數類神經網路模式所需之相關變數如表 8 所示。

4.4 調適性模糊類神經網路模式

另一類神經網路模式：調適性模糊類神經網路模式應用預測水庫水質變數，此模式於訓練與測試階段所採用的資料筆數與前節幅狀基底函數類神經網路模式相同。

圖 5 為預測溶氧與實測溶氧之散佈圖，訓練階段預測溶氧之 MAE、RMSE 及 R 值分別為 0.68 mg/L、0.97 mg/L 及 0.84；於測試階段預測溶氧之 MAE、RMSE 及 R 值分別為 0.52 mg/L、0.74 mg/L 及 0.91 (見表 9)。圖 6 為預測總磷與實測總磷之散佈圖，於訓練階段預測總磷之 MAE、RMSE 及 R 值分別為 6.57 μ g/L、13.85 μ g/L 及 0.86；於測試階段預測總磷之 MAE、RMSE 及 R 值分別為 4.75 μ g/L、7.06 μ g/L 及 0.91(見表 9)。

表 9 調適性模糊類神經網路模式預測不同水質變之誤差分析

水質變數	訓練階段			測試階段		
	MAE	RMSE	R	MAE	RMSE	R
DO	0.68 mg/L	0.97 mg/L	0.84	0.51 mg/L	0.74 mg/L	0.91
TP	6.57 $\mu\text{g/L}$	13.85 $\mu\text{g/L}$	0.86	4.75 $\mu\text{g/L}$	7.06 $\mu\text{g/L}$	0.91
Chl- <i>a</i>	2.97 $\mu\text{g/L}$	5.11 $\mu\text{g/L}$	0.88	3.04 $\mu\text{g/L}$	5.49 $\mu\text{g/L}$	0.91
SD	0.76 m	1.02 m	0.81	0.65 m	0.91 m	0.86

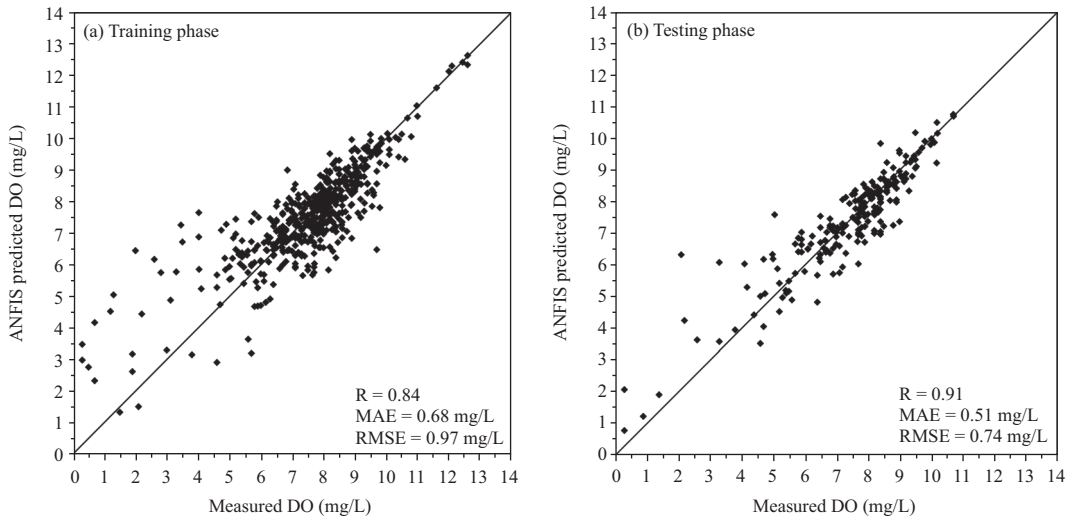


圖 5 調適性模糊類神經網路模式預測溶氧與實測溶氧之散佈圖(a)訓練階段；(b)測試階段

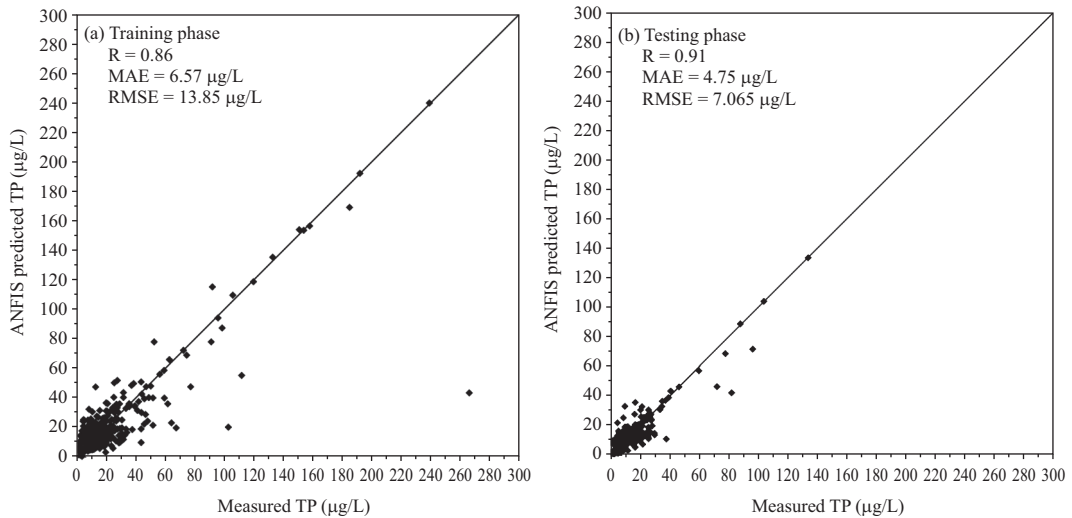


圖 6 調適性模糊類神經網路模式預測總磷與實測總磷之散佈圖(a)訓練階段；(b)測試階段

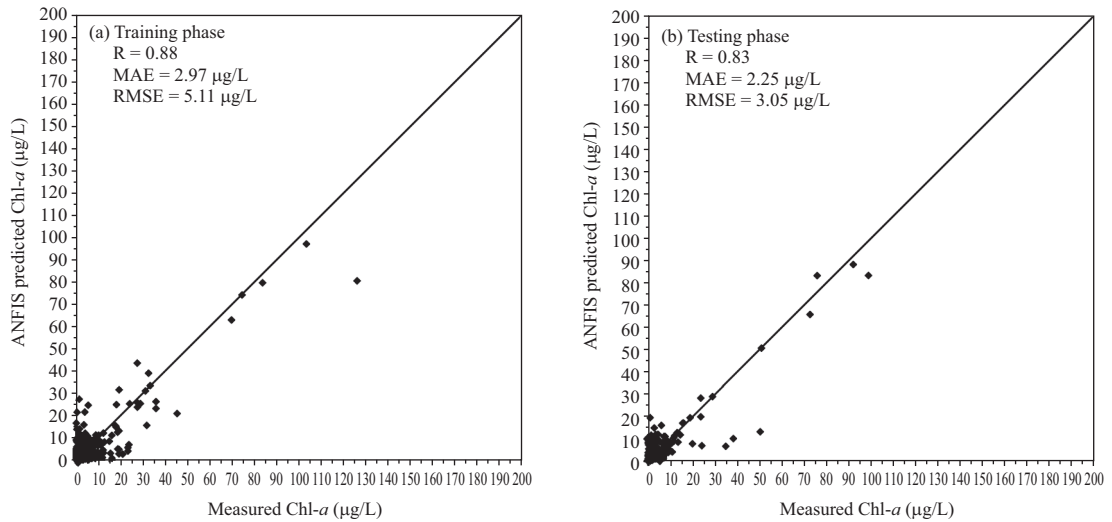


圖 7 調適性模糊類神經網路模式預測葉綠素 a 與實測葉綠素 a 之散佈圖(a)訓練階段；(b)測試階段

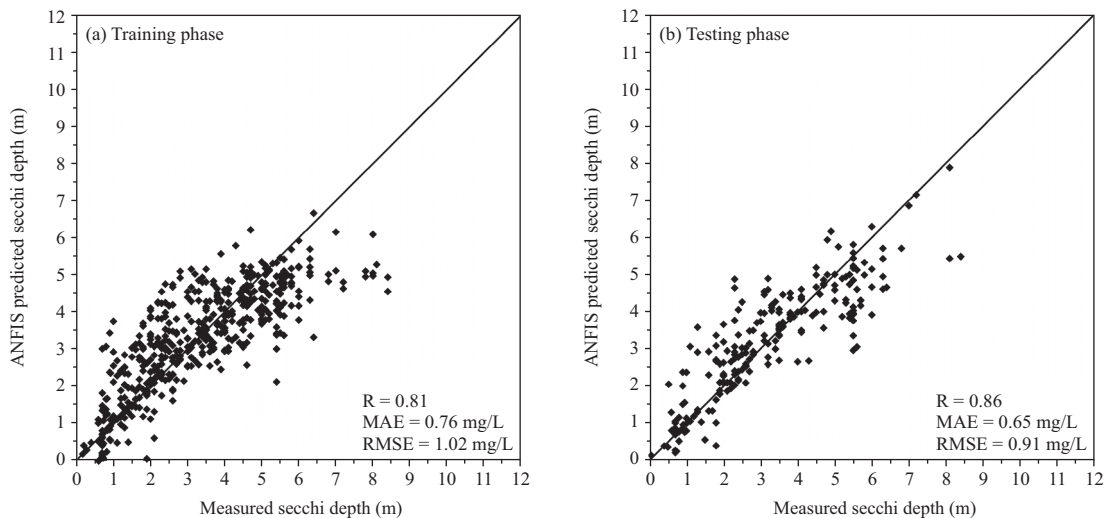


圖 8 調適性模糊類神經網路模式預測透明度與實測透明度之散佈圖(a)訓練階段；(b)測試階段

圖 7 為預測葉綠素 a 與實測葉綠素 a 之散佈圖，於訓練階段預測葉綠素 a 之 MAE、RMSE 及 R 值分別為 $2.97 \mu\text{g/L}$ 、 $5.11 \mu\text{g/L}$ 及 0.88 ；於測試階段預測葉綠素 a 之 MAE、RMSE 及 R 值分別為 $3.04 \mu\text{g/L}$ 、 $5.49 \mu\text{g/L}$ 及 0.91 (見表 9)。圖 8 為預測透明度與實測透明度之散佈圖，於訓練階段預測透明度之 MAE、RMSE 及 R 值分別為 0.76 m 、 1.02 m 及 0.81 ；於測試階段預測透明度之 MAE、RMSE 及 R 值分別為 0.65 m 、 0.91 m

及 0.86 (見表 9)。調適性模糊類神經網路模式預測透明度於訓練階段之 MAE、RMSE 及 R 值略差於幅狀基底函數類神經網路模式預測結果，但於測試階段則調適性模糊類神經網路模式優於幅狀基底函數類神經網路模式預測透明度。其他水質變數:溶氧、總磷及葉綠素 a ，調適性模糊類神經網路模式預測結果均優於幅狀基底函數類神經網路模式及多變數線性迴歸模式。由此可知，因為水庫水質具有非線性的特性，因此非線

性之類神經網路模式於預測德基水庫水質變數優於線性之迴歸模式。Rankovic *et al.* (2012)的研究結果同樣顯示：以非線性之類神經網路模式描述複雜的非線性輸入與輸出變數間之關係，可以得到有效率且精確的結果。就本研究所採用的兩種類神經網路模式，整體而言，調適性模糊類神經網路模式預測結果仍優於幅狀基底函數類神經網路模式之預測結果，故可建議調適性模糊類神經網路模式應用於德基水庫水質之模擬預測。調適性模糊類神經網路模式所需之相關變數詳見表 8。

4.5 討論

整體而言，本研究結果顯示，類神經網路模式包括幅狀基底函數類神經網路模式與調適性模糊類神經網路模式，在水庫水質預測的效能表現上均較多變數線性迴歸模式為佳。然而，多變數線性迴歸模式在水質模擬與預測仍可視為一個簡易且快速的工具，舉例來說，Chenini and Khemiri (2009)使用多變數線性迴歸模式評估突尼西亞中部地區 Maknassy 流域的地下水水質；Thoe and Lee (2014)採用多變數線性迴歸模式進行香港海灘的每日水質預測。

本研究之誤差分析結果顯示，在調適性模糊類神經網路模式對德基水庫水質預測中，除葉綠素 *a* (Chl-*a*)外，其餘水質變數之預測均呈現測試階段優於訓練階段的情況(見表 9)，即測試階段誤差小於訓練階段。Rankovic *et al.* (2010)使用類神經網路進行單一變數模擬水庫溶氧時亦有類似情況發生；Pourtousi *et al.* (2015)使用計算流體動力與調適性模糊類神經網路模式預測氣泡塔 (bubble column)的液體流速時也呈現測試階段略優於訓練階段的情形。推斷可能原因有二，一為測試階段所使用的資料數量遠少於訓練階段，因此可能產生大誤差的機率較低；再則為測試所使用的輸入資料與預測變數間，有較好的關聯性。此外，本研究結果也顯示，調適性模糊類神經網路模式在水質變數預測的精度表現較一般類神經網路模式為佳，此結果不但與 Chen and Liu (2014)於翡翠水庫溶氧預測的研究結果一致

(調適性模糊類神經網路模式之溶氧預測結果優於倒傳遞類神經網路模式)，亦與 Chau (2010)比較調適性模糊類神經網路模式與倒傳遞類神經網路模式於中國漫灣水庫流量模擬之效能結果一致，即調適性模糊類神經網路模式之流量預測結果優於倒傳遞類神經網路模式。

五、結 論

本研究以多變數線性迴歸模式與類神經網路模式(幅狀基底函數類神經網路模式及調適性模糊類神經網路模式)預測德基水庫之水質包含溶氧、總磷、葉綠素 *a* 以及透明度，因為此四種水質變數為影響水庫優養的重要指標，經水質模擬預測後，可獲致以下的結論分述如下：

1. 對於輸入模式水質變數之選取採用單變數線性迴歸法，求取輸出水質變數(溶氧、總磷、葉綠素 *a* 以及透明度)與其他個別水質變數之相關係數，結果得知影響總磷相關係數較高的水質因子為透明度(SD)、濁度(TB)、懸浮固體物(SS)、葉綠素 *a* (Chl-*a*)、硝酸氮(NO₃-N)及總鹼度(TA)；影響透明度之重要水質因子為總磷(TP)、水溫(WT)、pH、濁度(TB)、溶氧(DO)及懸浮固體物(SS)；影響溶氧的重要水質因子為水溫(WT)、pH、透明度(SD)、葉綠素 *a*、總硬度(TH)及總鹼度(TA)；影響葉綠素 *a* 之重要水質因子有水溫(WT)、pH、電導度(EC)、溶氧(DO)、總硬度(TH)及總鹼度(TA)。
2. 以多變數線性迴歸建立的四種預測不同水質的模式，結果顯示於訓練階段的相關係數值介於 0.48~0.74；於測試階段相關係數值介於 0.45~0.71。且各項誤差值(平均絕對誤差及均方根誤差)均有偏高的現象，表示多變數線性迴歸模式並無法準確預測溶氧、總磷、葉綠素 *a* 及透明度。
3. 以幅狀基底函數類神經網路模式預測不同的水質變數，可知此類神經網路模式所預測之水質變數，於訓練階段的相關係數值介於 0.74~0.88；於測試階段相關係數值介於 0.75~0.88。預測各項水質變數之結果相較於以多變數線性迴歸模式為佳。

4. 以調適性模糊類神經網路模式預測不同水質變數，得知於訓練階段的相關係數值介於0.81~0.88；於測試階段相關係數值介於0.86~0.91。整體而言，此類神經網路模式相較於幅狀基底函數類神經網路模式所預測的水質為佳，顯示調適性模糊類神經網路模式最適用於德基水庫之水質預測。

謝 誌

本研究承蒙行政院科技部研究計畫(編號：MOST 103-2625-M-239-001)的經費補助及行政院環境保護署提供之實測資料得以完成，謹致謝忱。

參考文獻

1. 王英義：「整合類神經網路與遺傳演算法於水庫水質管理之研究」，國立臺灣大學土木工程學研究所博士論文，2005。
2. 吳冠儀：「應用類神經網路與衛星影像於水庫水質推估-以石門水庫為例」，國立中興大學環境工程學系所碩士論文，2013。
3. 張麒顯：「應用類神經網路於水庫之水質預測分析」，國立中興大學土木工程學系所碩士論文，2008。
4. 張斐章、張麗秋：類神經網路，臺灣東華書局股份有限公司，2005。
5. 梁惠儀、陳莉、王泰盛、蔡祐竹：「結合類神經網路與遺傳演算法於翡翠水庫最佳控制優養化之研究」，農業工程學報，第55卷第3期，pp. 42-52，2009。
6. 經濟部水利署網站，<http://www.wra.gov.tw/>。
7. 謝明翰：「應用倒傳遞類神經網路於水庫水質之分析—以德基與曾文水庫為例」，國立臺灣大學土木工程學研究所碩士論文，2009。
8. Abyaneh, H. Z., "Evaluation of multivariate linear regression and artificial neural networks in prediction of water quality parameters," *Journal of Environmental Health Science and Engineering*, 12(4), 1-8, 2014.
9. Akkoyunlu, A., Altun, H., and Cigizoglu, H. K., "Depth-integrated estimation of dissolved oxygen in a lake," *Journal of Environmental Engineering*, 137(10), 961-967, 2011.
10. Ay, M., and Kisi, O., "Modelling of chemical oxygen demand by using ANNs, ANFIS and k-means clustering techniques," *Journal of Hydrology*, 511, 279-289, 2014.
11. Carlson, R. E., "Trophic state index for lake," *Limnology and Oceanography*, 22(2), 361-369, 1977.
12. Chang, F. J., Chung, C. H., Chen, P. A., and Liu, C. W., Coynel, A., Vachaud, G., "Assessment of arsenic concentration in stream water using neuro fuzzy networks with factor analysis," *Science of Total Environment*, 494-495, 202-210, 2014.
13. Chang, F. J., Tsai, Y. H., Chen, P. A., and Vachaud, G., "Modeling water quality in an urban river using hydrological factors-data driven approaches," *Journal of Environmental Management*, 151, 87-96, 2015.
14. Chapra, S. C., Surface Water-quality Modeling, McGraw-Hill, NY, p. 844, 1997.
15. Chen, C. W., "Concepts and utilities of ecological model," *Journal of Sanitary Engineering, ASCE*, 96(SA5), 1085-1097, 1970.
16. Chau, K. W., "Modelling for coastal hydraulics and engineering," *Spon Press*, 2010.
17. Chen, D., Lu, J., and Shen, Y., "Artificial neural network modelling of concentrations of nitrogen, phosphorus and dissolved oxygen in a non-point source polluted river in Zhejiang Province, southern China," *Hydrological Processes*, 24(3), 290-299, 2010.
18. Chen, W. B., and Liu, W. C., "Artificial neural network modeling of dissolved oxygen in reservoir," *Environmental Monitoring and Assessment*, 186(2), 1203-1217, 2014.
19. Chen S., Cowan C. F. N., Grant P. M., "Orthogonal least squares learning algorithm for

- radial basis function network,” *IEEE Transaction on Neural Networks*, 2(2), 302-309, 1991.
20. Chenini, I., and Khemiri, S., “Evaluation of ground water quality using multiple linear regression and structural equation modeling,” *International Journal of Environmental Science & Technology*, 6 (3), 509-519, 2009.
 21. Dibike, Y. B., Solomatine, D. P., Abbott, M. B., “On the encapsulation of numerical-hydraulic models in artificial neural network,” *Journal of Hydraulic Research*, 37(2), 147-161, 1999.
 22. Di Toro, D. M., O’Connor, D. J., Thomann, R. V., “A dynamic model of phytoplankton population in the Sacramento-San Joaquin Delta,” *Advances in Chemistry, American Chemistry Society*, 106, 131-180, 1971.
 23. Dogan, E., Sengorur, B., and Koklu, R., “Modeling biological oxygen demand of the Melen River in Turkey using an artificial neural network technique,” *Journal of Environmental Management*, 90(2), 1229-1235, 2009.
 24. Ham, F. M., Kostanic I., Principles of Neurocomputing for Science and Engineering, McGraw-Hill, New York, NY, 2001.
 25. Jang, J. S. “ANFIS: adaptive-network-based fuzzy inference system,” *IEEE Transactions on Systems, Man, and Cybernetics*, 23, 665-685, 1993.
 26. Jiang, Y., Nan, Z., and Yang, S., “Risk assessment of water quality using Monte Carlo simulation and artificial neural network method,” *Journal of Environmental Management*, 122, 130-136, 2013.
 27. Kecman V., *Learning and Soft Computing: Support Vector Machine, Neural Networks, and Fuzzy Logic Models*, Cambridge, MA, MIT Press, 2001.
 28. Khataee, A. R., Zarel, M., and Pourhassan, M., “Bioremediation of malachite green from contaminated water by three microalgae: neural network modeling,” *Clean*, 38(1), 96-103, 2010.
 29. Kim, Y., Shin, H. S., and Plummer, J. D., “A wavelet-based autoregressive fuzzy model for forecasting algae blooms,” *Environmental Modelling and Software*, 62, 1-10, 2014.
 30. Kuo, J. T., Hsieh, M. H., Lung, W. S., and She, N., “Using artificial network for reservoir eutrophication prediction,” *Ecological Modelling*, 200(1-2), 171-177, 2007.
 31. Kuo, J. T., Wang, Y. Y., and Lung, W. S., “A hybrid neural-genetic algorithm for reservoir water quality management,” *Water Research*, 40(7), 1367-1376, 2006.
 32. Martin, J. L., and McCutcheon, S. C., Hydrodynamics and Transport for Water Quality Modeling, CRC Press, Boca Raton, Florida, USA, P. 794, 1999.
 33. Najah, A., El-Shafie, A., Karim, O. A., and El-Shafie. A. H., “Performance of ANFIS versus MLP-NN dissolved oxygen prediction models in water quality monitoring,” *Environmental Science and Pollution Research*, 21(3), 1658-1670, 2014.
 34. Najah, A., El-Shafie, A., Karim, O. A., and El-Shafie. A. H., “Application of artificial neural networks for water quality prediction,” *Neural Computing and Applications*, 22(1), S187-S201, 2013.
 35. Pourtousi, M., Sahu, J. N., Ganesan, P., Shamsirband, S., and Redzwan, G., “A combination of computational fluid dynamics (CFD) and adaptive neuro-fuzzy system (ANFIS) for prediction of the bubble column hydrodynamics,” *Powder Technology*, 274, 66-481, 2015.
 36. Rankovic, V., Radulovic, J., Radojevic, I., Ostojic, A., and Comic, L., “Neural network modeling of dissolved oxygen in the Gruza reservoir, Serbia,” *Ecological Modelling*, 221(8), 1239-1244, 2010.
 37. Rankovic, V., Radulovic, J., Radojevic, I.,

- Ostojic, A., Comic, L., "Prediction of dissolved oxygen in reservoirs using adaptive network-based fuzzy inference system," *Journal of Hydroinformatics*, 14(1), 167-179, 2012.
38. Singh, K. P., Basant, A., Malik, A., and Jain, G., "Artificial neural network modeling of the river water quality-a case study," *Ecological Modelling*, 220(6), 888-895, 2009.
39. Soltani, F., Kerachian, R., and Shirangi, E., "Developing operating rules for reservoirs considering the water quality issues: application of ANFIS-based surrogate models," *Expert Systems with Applications*, 37(9), 6639-6645, 2010.
40. Soyupak, S., Karaer, F., Gurbuz, H., Kivrak, E., Senturk, E., and Yazici, A., "A neural network-based approach for calculating dissolved oxygen profiles in reservoirs," *Neural Computing & Applications*, 12(3-4), 166-172, 2003.
41. Thoe, W., and Lee, J., "Daily forecasting of Hong Kong beach water quality by multiple linear regression models," *Journal of Environmental Engineering*, 140 (2), 04013007, 2014.
42. Thomann, R. V., and Mueller, J. A., Principles of Surface Water Quality Modeling and Control, Harper & Row, NY, P. 644, 1987.
43. Thomann, R. V., Di Toro, D. M., Winfield, R. P., and O'Connor, D. J., "Mathematical modeling of phytoplankton in Lake Ontario, 1: model development and verification," EPA-600/3-75-005, USEPA Ecological Research Series, Corvallis, Oregon, 1975.
44. Weisberg, S., Applied Linear Regression (2nd ed.), John Wiley, New York, 324pp, 1985.
45. Wen, X., Fang, J., Diao, M., and Zhang, C., "Artificial neural network modeling of dissolved oxygen in the Heihe River, Northwestern China," *Environmental Monitoring and Assessment*, 185(5), 4361-4371, 2013.
46. Xiang, X., Xu, X., and Tao, J., "Modelling chlorophyll-*a* in Bohai Bay based on hybrid soft computing approach," *Journal of Hydroinformatics*, 15(4), 1099-1108, 2013.
47. Yang, Y., Xu, X., and Zhang, W., "Design neural networks based fuzzy logic," *Fuzzy Sets and System*, 114(2), 325-328, 2000.

收稿日期：民國 104 年 4 月 7 日

修正日期：民國 104 年 6 月 1 日

接受日期：民國 104 年 7 月 30 日