

以複迴歸分析法探討水質因子與優養化全模式之研究

陳鴻烈⁽¹⁾ 蔡大偉⁽²⁾

摘要

本研究以德基水庫水質變化情形為對象，利用複迴歸分析方法探討控制優養化之動力因子，以期更了解該區控制優養化變化的機制。研究共分為三部分，第一部分是優養化重要水質因子進行複迴歸分析，探討其對優養化變化的解釋能力。第二部分是藉由多元共線性來檢定該迴歸模式之共線性程度，並討論共線性是否會影響迴歸式之解釋能力。第三部分則是運用殘差分析來檢定觀測值所構成的時間數列是否有自我相關情形，以及離群值存在。研究顯示，複迴歸模式之解釋能力可達 0.9116 的高解釋力，而各個水質因子中以懸浮固體、COD 及葉綠素 a 有較嚴重的共線性。另外，由殘差分析結果可知，觀測值時間數列中並無自我相關的現象發生，但存有 5 個可能的離群值。離群值經過修正後，即可得到最後的優養化全模式，而研究顯示，此模式不但可充分解釋優養化之變化情形，亦可作為未來水庫管理的參考依據。

(**關鍵詞**：優養化、複迴歸分析、水質因子、德基水庫)

An Investigation of Water Quality Factors and Full-Eutrophication Model with Multiple Regression Analysis

Paris Honglay Chen⁽¹⁾, David D-W. Tsai⁽²⁾

Professor⁽¹⁾ and Graduate Students⁽²⁾, Department of Soil and Water Conservation,
National Chung Hsing University, Taichung, Taiwan, R.O.C

ABSTRACT

This study was to investigate the kinetic factors of eutrophication in Te-Chi reservoir with multiple regression analysis to learn more about the mechanism of eutrophication. The study was divided into three parts. The first part was to investigate the explanation ability of important water quality factors of eutrophication by the multiple regression method. The second part was to test the degree of collinearity for the regression model by the multicollinearity method and to discuss whether the collinearity affected the explanation ability of regression model or not. Finally, the residual

(1) 國立中興大學水土保持學系教授

(2) 國立中興大學水土保持學系碩士班研究生

analysis was used to inspect whether there were self-relative conditions and outliers or not in the time series constructed from the observation values. The results showed that the the multiple regression model had a high explanation ability indicated by R^2 value of 0.9116. Water quality factors of suspended solids, COD, and chlorophyll-a showed obvious collinearity. Additionally the results of residual analysis indicated that the autoregression didn't exist in the time series but five possible outliers were observed. After calibrating the outlets, we could get the full eutrophication model. The results showed that this model could well explain the variations of eutrophication for use in reservoirs management.

(**Key words:** Eutrophication, Multiple regression analysis, Water quality factor, Te-Chi reservoir)

前 言

近年來台灣地區遭受 921 集集大地震的影響，造成地質鬆軟，以致於每遭受豪大雨侵襲時，坡地安全問題就令人擔憂，同時亦突顯坡地嚴重開發的問題。其中，坡地農業的大量開發，更是污染了台灣大部分的河川，致使絕大多數的水庫都面臨了優養化的問題，因而造成飲水品質的下降，甚至有危害人體健康之虞。

本研究以德基水庫優養化為研究對象(經, 1983~1999)，該水庫位於大甲溪上游，遭受梨山地區坡地農業開發之污染已久，致使優養化問題趨於嚴重 (Antikainen, 2005)。本研究主要目的即是探討造成優養化情形之背後動力因子，此領域相關研究中使用複迴歸及相關統計分析是常見的方法之一 (Arheimer, 2000)，可利用複迴歸分析模型來嘗試找出對優養化管理最重要的因子。研究結果可對日後發展優養化管理模式有所助益，譬如 BMP 管理策略的運用 (Hiscock, 2003)、時間數列分析預測 (陳, 1997b, 1998, 1999, 2004, 2005) 等，冀望能藉此降低管理成本並提升經濟效益，使優養化情形可獲得有效的控制。

理論分析

一、複迴歸分析 (multiple regression)

當線性迴歸模型有 1 個以上的自變數時，即為複迴歸分析，其一般式可表示如下 (林, 2003a)：

$$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \dots + \varepsilon_i$$
$$= \beta_0 + \sum_{j=1}^n \beta_j x_{ji} + \varepsilon_i$$

x_{ji} ：第 i 組變數， x_{ji} 為獨立且已知的觀測值

y_i ：對應於 x_{ji} 的反應變數，又稱為因變數

β_0 、 β_1 、 β_2 、...、 β_n ：線性關係係數

ε_i ：散佈於迴歸線附近的誤差。一般而言，我們通常假設 ε_i 為相互獨立且平均數為 0，變異數 σ^2 的常態分配

複迴歸模型之迴歸係數亦可依最小平方方法求得 (陳, 1997a)，例如：有 3 個自變數時，與因變數可構成一個四維迴歸平面，稱之為超平面 (hyperplane)，此時需要 4 個正規方程式來求解 (潘, 2003)。另外，當變數為定性函數而非定量函數 (譬如做生態環境棲地適宜性指數之研究) 時，則需使用特殊的迴歸方法，像是 logistic regression (林, 2004)。一般兩個以上之迴歸式之計算極為繁複，因此

通常需藉由電腦計算方可求得。

一般複迴歸模式對變異度的解釋能力可用決定係數 (R^2) 來表示,但如果需要比較多個迴歸模式,且模式間之觀測值或參數個數不同時,則需要使用調整後之決定係數值 ($Adj.R^2$) 來做調整。至於校正模式間自由度與觀測值數目不同的問題,為了使比較時能有相同的基準,計算時是以自由度改變對模式無法解釋的部分進行修正,其修正式如下:

$$Adj.R^2 = 1 - (1 - R^2) \frac{N - 1}{N - k}$$

N : 觀測值個數

k : 模式參數個數

二、多元共線性檢定

當迴歸模型之自變數間有高度相關時,會導致共線性的現象發生,而降低模型的解釋能力。因此,必須檢驗迴歸模型之共線性程度,以利於進行進一步的校正,來克服迴歸係數標準誤及預測值變異數膨脹的缺點。共線性之檢驗法可用以下幾個方法來檢定(陳, 2003):

1. 具有複迴歸模型決定係數 (R^2) 大,但迴歸係數多數未達顯著的情形。

2. 自變數間的相關性檢定

將各個自變數之相關係數列出,形成自變數相關矩陣,便可進行皮爾森相關分析(張, 2004)。若自變數間有高度相關情形發生,則表示可能存在共線性的問題。

3. 容忍度 (tolerance) 與變異數波動因素 (variance inflation factor, VIF)

容忍度等於 1 減去決定係數值 (林, 2002), 因此, 其算法可表示為

$$\text{容忍度} = 1 - R_i^2$$

R_i^2 : 以其他自變數預測第 i 個自變數所得之決定係數

決定係數值愈大,表示愈有共線性的問題;亦即容忍度數值越小,表示共線性問題愈嚴重。另外,變異數波動因素 (VIF) 則為容忍度之倒數,即

$$VIF = \frac{1}{1 - R_i^2}$$

R_i^2 : 以其他自變數預測第 i 個自變數所得之決定係數

變異數波動因素的判別標準是:如果數值大於 10,則代表該變項與其他預測變項有共線性的問題(陳, 2003)。

4. 特徵值 (eigenvalue)

分別計算各個自變數之特徵值,若數值接近 0 的話,則代表此 X 矩陣可能是特異矩陣 (singular matrix),也就是有線性相依的情形發生。

5. 條件指數 (condition index, CI)

條件指數為根據特徵值所計算的指數,其計算方法如下:

$$CI = \sqrt{\frac{\lambda_{\max}}{\lambda_i}}$$

λ_{\max} : 最大特徵值

λ_i : 第 i 項之特徵值

條件指數判別標準為:當其數值介於 30~100 之間時,為中度共線性;若大於 100 以上,則是高度共線性(陳, 2003)。

6. 變異數比例 (variance proportions)

根據矩陣 X 所求之特徵向量 (eigen vector) 來判定,在任一列上有任兩個變項以

上之係數非常接近1者，或是數值均大於0.5以上者，則表示這些變項間可能有共線性的情形 (Katz, 2005)。

三、殘差分析

進行殘差分析可對複迴歸模式做適合性的診斷 (林, 2003b) 與誤差變方的一致性檢測 (葉, 2001)，包括可檢定出觀測值中是否有特殊的觀察體，如極端值或是具有高度影響力的數值，並且可以檢定時間數列中是否有自我相關的現象，其檢驗法列舉如下 (陳, 2003)：

1. 自我相關性檢測 — Durbin-Watson D 檢定法

時間序列之迴歸分析的基本假設之一為殘差間無自我相關性，即 Durbin-Watson D 檢定法計算之 DW 值會在 2 左右，其計算方法表示如下 (汪, 2001)：

$$DW = \frac{\sum_{t=2}^n (e_t - e_{t-1})^2}{\sum_{t=1}^n e_t^2}$$

e ：殘差值

2. 離群值判定分析

離群值的判定可由兩方面來進行，一是根據模式殘差值的計算，判斷離群值的可能位置，一般經過標準化後，殘差值大於 2 或 3 個標準差者，即有可能為離群值。另一個方法是找尋對模式最有影響力的觀測值，經由逐一刪除觀測值後對模式影響的計算方法，因此，尋找最有影響力之觀測值，亦為離群值可能之所在。

(1) 殘差值分析：

A) 標準化殘差值 — 某個觀察值之殘差等於觀察值減去預測值，計算出殘差後，再使用迴歸標準誤加以標準化，可以下式表

示：

$$\frac{e_i}{S} = \frac{e_i}{\sqrt{MSe}}$$

e_i ：殘差值。

S ：迴歸標準誤

當計算結果之絕對值大於 2 或 3 個標準差時，就表示該觀測值可能是極端值。

B) 經 t 標準化殘差值 — 經過 t 分配後的標準化殘差值，其計算方式如下：

$$\text{Stud. Residual} = \frac{e_{(i)}}{\sqrt{MSe_{(i)} \times (1 - h_{ii})}}$$

e_i ：殘差值。

$MSe_{(i)}$ ：第 i 個觀測值之迴歸標準誤

h_{ii} ：第 i 個觀測值之槓桿值

此一指標之判斷標準，亦為當計算結果之絕對值大於 2 或 3 個標準差時，則代表該觀測值可能是極端值。

C) Rstudent 檢定 — 可用來測驗觀測值中是否有離群值存在，其算法與 t 標準化殘差值類似，不同之處在於計算標準差時為刪除第 i 個觀測值後的估計值，其計算式為：

$$Rstudent = \frac{e_{(i)}}{\hat{\sigma}_{(i)} \sqrt{1 - h_{(i)}}}$$

$e_{(i)}$ ：殘差值

$\hat{\sigma}_{(i)}$ ：第 i 個觀測值之標準差估計值

h_i ：第 i 個觀測值之槓桿值

一般而言 Rstudent 檢定結果若絕對值大於 2 個標準差，則表示該觀測值可能是離群值。

(2) 刪除觀測值殘差分析法：

A) Cook's D 距離值 — 是刪除第 i 個觀察值後所得之迴歸係數改變值，其值大約呈 $F(1-\alpha; k+1, N-k-1)$ 分配。當其值大於 0.5 時，則表示可能是具有影響力之觀測值 (Turner-Fairbank HRC, 2000)。此檢定法亦可同時檢定自變數或應變數中具有影響力的觀測值，其計算式可表示為：

$$D_i = \frac{e_{(i)}^2}{p \times MSE_{(i)}} \left[\frac{h_{ii}}{(1-h_{ii})^2} \right]$$

D_i : Cook's D 值

p : 參數個數

$MSE_{(i)}$: 第 i 個觀測值之迴歸標準誤

h_{ii} : 第 i 個觀測值之槓桿值

B) Hat Diag H — 計算 Hat 矩陣的槓桿值 (leverage value)，即矩陣中對角線元素，可求得每個觀測值對模型的影響力。因此，可藉此找出是否有離群值存在，其計算方法如下：

$$h_{ii} = X_i' (X'X)^{-1} X_i$$

h_{ii} : 第 i 個觀測值之槓桿值

X : 自變數矩陣

此一指標判斷門檻值為 $2(k+1)/n$ ， k 為參數個數， n 為觀測值個數，若大於該門檻值，即可能為離群值所在處 (陳，2003)。

C) 共變數率 (covariance ratio) — 其定義為「刪除第 i 個觀測值後對迴歸係數精確性的改變」，因此，其值若大於 1，則精確性增加；若小於 1，則精確性減少。如果共變數率與 1 的差值 (即 $|\text{COVRATIO}_i - 1|$) 大於 $3(k+1)/N-k-1$ (k 為參數個數， N 為觀測值個數)，代表該觀測值可能具有高度影響力 (陳，2003)。其計算式可以下式來表示：

$$\text{COVRATIO}_i = \frac{S_{(i)}^2 |(XX)_{(i)}^{-1}|}{S^2 |(XX)^{-1}|}$$

X : 自變數矩陣

S^2 : 觀測值誤差變方

D) 刪除後預測差異 (DFFITs) — 其計算方法為計算以完整迴歸方程式預測第 i 個觀測值，與用刪除第 i 個觀測值所得之迴歸方程式預測第 i 個觀測值之差距，其計算式為：

$$\text{DFFITs} = \frac{\hat{Y} - \hat{Y}_{(i)}}{\sqrt{MSE_{(i)} \times h_{ii}}}$$

\hat{Y} : 刪除第 i 個觀測值所得迴歸方程式預測第 i 個觀測值

$\hat{Y}_{(i)}$: 完整迴歸方程式預測第 i 個觀測值

$MSE_{(i)}$: 第 i 個觀測值之迴歸標準誤

h_{ii} : 第 i 個觀測值之槓桿值

此一指標之標準為 $2\sqrt{(k+1)/N}$ ， k 為參數個數， N 為觀測值個數，當檢定值超過此標準時，該觀測值就有可能是具有影響力之觀測值 (陳，2003)。

E) 刪除後標準化差異 (DFBETA) — 可以檢定觀測值是否具有高度影響力，其計算方式為算出刪除第 i 個觀測值後之迴歸係數改變值，計算方法如下：

$$\text{DFBETA} = \frac{b - b_{(i)}}{\sqrt{MSE_{(i)} \times \text{diag}(XX)^{-1}}}$$

b : 完整迴歸方程式之迴歸係數

$b_{(i)}$: 第 i 個觀測值之迴歸係數

$MSE_{(i)}$: 第 i 個觀測值之迴歸標準誤

X : 自變數矩陣

當刪除後標準化差異值大於 $2/\sqrt{n}$ (n 為觀測值個數) 時, 則代表該觀測值可能具有高度影響力。而當研究觀測值為小樣本時, 則可用 1 作為門檻值 (Gillen, 2005)。

結果達顯著之重要優養水質因子進行複迴歸分析, 得到優養複迴歸模型。然複迴歸模式所選用的參數必須觀測值一致才可進行分析, 而受到缺失值的影響, 只有 23 個時間點有完整的資料, 因此, 最後只能利用 23 個觀測值來做分析。

接著, 進一步藉由 SAS 統計軟體實施共線性與殘差分析, 探討該模式之合理性後, 再經過離群值的檢討修正, 即得到最後的優養化全模式 (整個研究流程如圖 1 所示)。

研究方法

本研究先採用 1983~1999 之德基水庫優養化指標資料進行單迴歸分析, 再利用分析

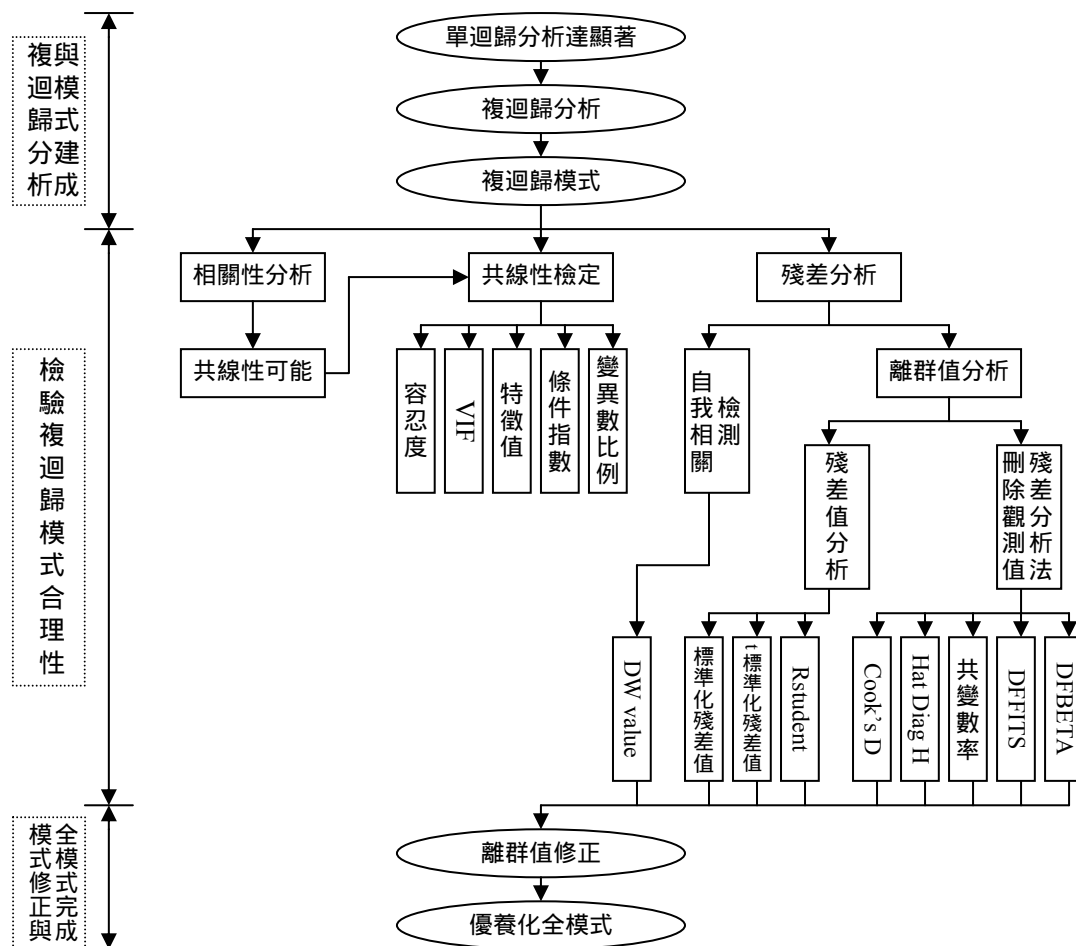


圖 1. 研究流程圖

Figure 1. Overview of research method.

結果與討論

一、複迴歸分析

1. 獨立變數之選擇

複迴歸分析所需要的獨立變數，是根據單迴歸分析後所篩選出來達顯著水準的相關因子來進行分析，可以更進一步闡釋各個相關因子的相互關係及變異程度。因此，本研究用來探討優養化相關因子之複迴歸分析共選擇水溫、溶氧、懸浮固體、濁度、氨氮、硝酸鹽、總磷、磷酸鹽、鈉、化學需氧量、葉綠素 a 及透明度等 12 個因子，這些因子的決定係數 (R^2) 如表 1 所示。

表 1. 複迴歸選用自變數之決定係數表

Table 1. R^2 of independent variables in multiple regression.

排序	自變數	R^2	P-value
1	總磷	0.7997	<0.0001
2	懸浮固體	0.5400	<0.0001
3	透明度	0.3956	<0.0001
4	葉綠素a	0.3818	<0.0001
5	COD	0.3259	<0.0001
6	濁度	0.2293	<0.0001
7	水溫	0.1324	0.0034
8	鈉	0.1292	0.0194
9	溶氧	0.1290	0.0038
10	硝酸鹽	0.1055	0.0094
11	氨氮	0.1009	0.0142
12	磷酸鹽	0.0961	0.0321

資料來源：陳，2005。

2. 複迴歸分析結果

以優養化指標 TSI 為應變數，而以單迴歸分析篩選後之 12 個相關因子作為獨立變數，進行複迴歸分析，其變異分析結果如表 2 所示。由表中數據可知，複迴歸所選用因子正確，其 F 值為 8.60，迴歸結果達 99.91% 顯著，決定係數亦達 0.9116 之高度解釋能

力。此外，12 個相關因子與應變數 TSI 間的參數估計如表 3 所示，而此一迴歸模型方程式可表示如下：

$$y = 49.13032 - 0.09661x_1 - 0.30594x_2 + 0.89422x_3 - 0.74840x_4 - 4.75770x_5 - 2.63486x_6 + 0.22938x_7 - 0.07382x_8 - 0.07329x_9 + 0.27991x_{10} - 0.00530x_{11} - 0.06516x_{12}$$

迴歸式中 $y = \text{TSI}$ ， $x_1 = \text{水溫}$ ， $x_2 = \text{溶氧}$ ， $x_3 = \text{懸浮固體}$ ， $x_4 = \text{濁度}$ ， $x_5 = \text{氨氮}$ ， $x_6 = \text{硝酸鹽}$ ， $x_7 = \text{總磷}$ ， $x_8 = \text{磷酸鹽}$ ， $x_9 = \text{COD}$ ， $x_{10} = \text{鈉}$ ， $x_{11} = \text{葉綠素 a}$ ， $x_{12} = \text{透明度}$ 。

由表 3 第 3 欄可發現許多變數的迴歸係數呈現負值，表示這些因子的變化與優養情形成反比。然而，除了透明度以外，其他變數的負值並不符合一般優養化的概念，因此，我們依變數的性質分類說明如下：

(1) 水溫

水溫係數為負值，代表水溫越高，優養程度反而下降。此一現象反映出這個區域的藻類族群有可能是屬於 psychrophilic 生物，它們最適合生存的溫度約為 20–25 左右。然而台灣夏季高溫，但因過高的水溫反而不利於這些藻類的生長，故造成溫度與優養化成反比的狀況。不過，因負值係數 (-0.09661) 不大，代表雖有此特殊狀況發生，但程度並不嚴重，仍有待進一步深入研究確認。

(2) 溶氧、濁度、COD、葉綠素 a

溶氧係數為負值，是因氧氣為藻類行光合作用後之廢物，當水溫高時，水中溶氧易達飽和，在飽和狀態下，溶氧可能會抑制藻類的生長，進而降低優養程度。

濁度分為有機與無機態，有機態之濁度即代表藻類族群，若族群量過高，會因超過環境承载力而自動下降。另外，當濁度過高時，不論為有機或無機態均會遮蔽陽光，進

表 2. 複迴歸分析之變異分析表

Table 2. ANOVA table of multiple regression analysis.

Source	DF	SS	MS	F value	Pr > F	R-Square
Model	12	3453.27607	287.77301	8.60	0.0009	0.9116
Error	10	334.74476	33.47448			
Corrected Total	22	3788.02083				

表 3. 迴歸式之參數估計表

Table 3. Parameter estimation of multiple regression equation.

Variable	DF	Parameter Estimation	Standard Error	t value	Pr > t	Standardized Estimation
截距	1	49.13032	23.86713	2.06	0.0666	0.00000
水溫	1	-0.09661	0.77791	-0.12	0.9036	-0.03119
溶氧	1	-0.30594	1.51376	-0.20	0.8439	-0.03975
懸浮固體	1	0.89422	0.64141	1.39	0.1935	0.61453
濁度	1	-0.74840	0.74842	-1.00	0.3409	-0.25528
氨氮	1	-4.75770	18.07072	-0.26	0.7977	-0.04967
硝酸鹽	1	-2.63486	9.61464	-0.27	0.7896	-0.05148
總磷	1	0.22938	0.04648	4.94	0.0006	0.91851
磷酸鹽	1	-0.07382	0.04862	-1.52	0.1599	-0.21800
COD	1	-0.07329	0.32639	-0.22	0.8268	-0.09113
鈉	1	0.27991	0.82934	0.34	0.7427	0.05102
葉綠素a	1	-0.00530	0.04374	-0.12	0.9059	-0.06198
透明度	1	-0.06516	3.11406	-0.02	0.9837	-0.00451

而妨礙藻類生長，此乃濁度係數呈現負值的原因。

COD 代表水中有機碳含量，因此也包括藻類族群，故 COD 過高，同樣代表水中藻類族群呈飽和現象。而藻類族群量過高，會因超過環境承載力而自動下降，此乃造成 COD 係數出現負值的原因。

葉綠素 a 乃直接量測藻類族群量的指標之一，所以葉綠素 a 係數為負值，與優養指數呈反比的情形，同樣代表藻類族群達飽和現象。

以上 4 個指標所代表的含意，均指出此區域優養程度嚴重，藻類族群生長已達飽和，甚至因超過環境承載力，而開始出現衰退的現象。

(3) 氨氮、硝酸鹽

藻類主要組成物質之一為蛋白質，故量測水中總氮含量，將可做為藻類族群量的指標。氨氮與硝酸鹽為總氮中無機氮的部分，因此，這兩個指標與優養化出現反比狀態，代表此區藻類的生長可能已達飽和，而有嚴重優養化情形發生。

(4) 磷酸鹽

磷為一般淡水優養的營養控制因子，磷酸鹽是代表營養鹽磷可被藻類吸收利用的狀態，只要水中磷含量增加，藻類即開始快速繁殖。但在此水體中，磷酸鹽與優養化卻呈反比現象，代表水中藻類出現生長過飽和的現象，且因族群密度過高而導致藻類生長不增反減的趨勢。此外，因藻類族群已達飽和，故此時磷酸鹽可能已成為非主要控制因子。

綜合以上 8 個變數對優養現象的解釋，均指出此水體已出現藻類族群生長過飽和的情形，亦即優養化現象已相當嚴重。

另外，表 3 中第 6 欄參數估計中只有總磷達 95% 顯著水準，表示此迴歸模型可能有共線性問題存在。至於表中第 7 欄之標準化迴歸係數，可以用來表示各變項之相對重要性，因此可看出總磷之預測作用最大，其次為懸浮固體含量。

經過複迴歸分析後，我們更能確定優養化的動力機制可以水溫、溶氧、懸浮固體、濁度、氨氮、硝酸鹽、總磷、磷酸鹽、鈉、化學需氧量、葉綠素 a 及透明度等 12 個因子來解釋。而這些因子中又以總磷與懸浮固體影響最為顯著，所以日後進行相關研究時可以這些因子為重點，以節省人力物力的投入。

二、相關性檢定

將所有的自變數之相關係數依矩陣型式排列 (表 4)，相關係數若達 0.6 以上，可視為高度相關 (楊，2004)。

1. 顯著水準 1%

經過相關顯著性統計檢定後，可得到顯著機率 p 值，藉由分析結果發現，相關係數 0.6 約為 1% 的顯著水準，而達顯著水準者可視為有顯著相關性。因此，由表 4 可知，若以建構 TSI 優養指標之 3 個變數為探討對象，可分為以下 3 組相關：

(1) 總磷

總磷與氨氮、懸浮固體有中度相關，代表水體中氮、磷循環的影響，亦即水中氮、磷含量的多寡與優養化情形有密切關聯 (Nijboer, 2004)。磷是一般淡水水體藻類生長之營養鹽控制因子，而氨氮與懸浮固體可用來代表藻類族群的濃度指標，故營養鹽磷與藻類濃度指標氨氮、懸浮固體間有顯著相關。

(2) 葉綠素 a

葉綠素 a 與懸浮固體、濁度、COD、磷酸鹽呈顯著相關。其中，懸浮固體、濁度、COD 均與藻類濃度相關，而磷更是一般淡水中藻類生長之營養鹽限制因子，因此，這 4 個因子均與藻類族群指標葉綠素 a 有顯著相關，意味著這些因子均為量測優養化的重要因素。

(3) 透明度

透明度與水溫、溶氧有負顯著相關，表示當水溫與溶氧上升時，可能會致使水中污染物質增加 (Cheng, 2003)，造成透明度下降。

進行複迴歸分析時，自變數之相關性應越低越好，才能完整的表現出自變數與應變數之關聯性 (俞，1990)，因此，在自變數相關檢定中，有些自變數具有高度相關性，代表此迴歸結果可能會產生共線性的問題。

2. 顯著水準 5%

由於一般統計分析均採 5% 為統計顯著水準，因此，本研究乃針對達 5% 顯著者 (表 5) 依同樣方式探討自變數間之相關性。

(1) 總磷

與總磷達顯著相關者有懸浮固體、濁度、氨氮、磷酸鹽、COD、葉綠素 a 與透明度，若和達顯著水準 1% 之相關變數進行比較，發現多了濁度、磷酸鹽、COD、葉綠素

表 4. 兩自變數間之相關矩陣 (達 1% 顯著水準)

Table 4. Correlation matrix between two independent variables. (reach 1% significance level)

Pearson Correlation Coefficients, N=19												
Prob. > r under H ₀ : R _{h0} = 0												
	水溫	溶氧	懸浮固體	濁度	氨氮	硝酸鹽	總磷	磷酸鹽	COD	鈉	葉綠素a	透明度
水溫	1	0.33443	0.51552	0.23984	0.50954	-0.25679	0.39430	0.18304	0.52679	0.19034	0.44145	-0.61396
P-Value	-	0.1188	0.0118	0.2703	0.0130	0.2369	0.0626	0.4032	0.0098	0.3843	0.0350	0.0018
溶氧	0.33443	1	0.22402	0.08488	0.12262	0.41161	0.32571	0.05465	0.17170	-0.13557	0.09469	-0.68963
P-Value	0.1188	-	0.3041	0.7002	0.5773	0.0510	0.1294	0.8044	0.4334	0.5374	0.6674	0.0003
懸浮固體	0.51552	0.22402	1	0.83642	0.41110	-0.20642	0.56858	0.56463	0.89324	0.07501	0.91294	-0.41434
P-Value	0.0118	0.3041	-	<.0001	0.0513	0.3447	0.0046	0.0050	<.0001	0.7338	<.0001	0.0493
濁度	0.23984	0.08488	0.83642	1	0.28067	-0.18945	0.41789	0.33944	0.73398	0.19102	0.78125	-0.32025
P-Value	0.2703	0.7002	<.0001	-	0.1946	0.3866	0.0472	0.1131	<.0001	0.3826	<.0001	0.1363
氨氮	0.50954	0.12262	0.41110	0.28067	1	-0.09423	0.66001	0.18031	0.46998	-0.06907	0.44310	-0.28229
P-Value	0.0130	0.5773	0.0513	0.1946	-	0.6689	0.0006	0.4103	0.0236	0.7542	0.0342	0.1919
硝酸鹽	-0.25679	0.41161	-0.20642	-0.18945	-0.09423	1	-0.14711	-0.25712	-0.36188	0.17154	-0.38844	-0.17324
P-Value	0.2369	0.0510	0.3447	0.3866	0.6689	-	0.5030	0.2363	0.0897	0.4338	0.0670	0.4292
總磷	0.39430	0.32571	0.56858	0.41789	0.66001	-0.14711	1	0.47440	0.51494	-0.08203	0.49371	-0.41443
P-Value	0.0626	0.1294	0.0046	0.0472	0.0006	0.5030	-	0.0222	0.0119	0.7098	0.0167	0.0493
磷酸鹽	0.18304	0.05465	0.56463	0.33944	0.18031	-0.25712	0.47440	1	0.47187	-0.11601	0.55207	-0.07709
P-Value	0.4032	0.8044	0.0050	0.1131	0.4103	0.2363	0.0222	-	0.0230	0.5981	0.0063	0.7266
COD	0.52679	0.17170	0.89324	0.73398	0.46998	-0.36188	0.51494	0.47187	1	-0.12060	0.95137	-0.26523
P-Value	0.0098	0.4334	<.0001	<.0001	0.0236	0.0897	0.0119	0.0230	-	0.5836	<.0001	0.2213
鈉	0.19034	-0.13557	0.07501	0.19102	-0.06907	0.17154	-0.08203	-0.11601	-0.12060	1	-0.03395	-0.11777
P-Value	0.3843	0.5374	0.7338	0.3826	0.7542	0.4338	0.7098	0.5981	0.5836	-	0.8778	0.5925
葉綠素a	0.44145	0.09469	0.91294	0.78125	0.44310	-0.38844	0.49371	0.55207	0.95137	-0.03395	1	-0.17978
P-Value	0.0350	0.6674	<.0001	<.0001	0.0342	0.0670	0.0167	0.0063	<.0001	0.8778	-	0.4118
透明度	-0.61396	-0.68963	0.41434	-0.32025	-0.28229	-0.17324	-0.41443	-0.07709	-0.26523	-0.11777	-0.17978	1
P-Value	0.0018	0.0003	0.0493	0.1363	0.1919	0.4292	0.0493	0.7266	0.2213	0.5925	0.4118	-

註：1. 以粗體字標記者為達統計水準 1% 顯著相關者。

2. 以斜線框標記者為經共線性檢定而有可能相關者。

a 與透明度等 5 個變數。其中，濁度在有機態部分包括藻類族群；磷酸鹽是總磷的一部分；COD 為量測有機碳含量，同樣包含藻類；另外，葉綠素 a 與透明度已被証實與總磷同為優養化重要變數，因此，這 7 個變數達顯著相關為合理現象。

(2) 葉綠素 a

葉綠素 a 在 5% 顯著水準下較 1% 時多了水溫、總磷與氨氮 3 個變數達顯著相關。其中，水溫的變化會影響藻類族群的生長速度，總磷為淡水中藻類生長的控制因子，而氨氮則為總氮中無機氮的一部份，均可在某種程度下代表優養化的現象，因此，呈顯著相關性亦為合理現象。

表 5. 兩自變數間之相關矩陣 (達 5% 顯著水準)

Table 5. Correlation matrix between two independent variables. (reach 5% significance level)

Pearson Correlation Coefficients, N=19												
Prob. > r under H ₀ : Rh ₀ = 0												
	水溫	溶氧	懸浮固體	濁度	氨氮	硝酸鹽	總磷	磷酸鹽	COD	鈉	葉綠素a	透明度
水溫	1	0.33443	0.51552	0.23984	0.50954	-0.25679	0.39430	0.18304	0.52679	0.19034	0.44145	-0.61396
P-Value	-	0.1188	0.0118	0.2703	0.0130	0.2369	0.0626	0.4032	0.0098	0.3843	0.0350	0.0018
溶氧	0.33443	1	0.22402	0.08488	0.12262	0.41161	0.32571	0.05465	0.17170	-0.13557	0.09469	-0.68963
P-Value	0.1188	-	0.3041	0.7002	0.5773	0.0510	0.1294	0.8044	0.4334	0.5374	0.6674	0.0003
懸浮固體	0.51552	0.22402	1	0.83642	0.41110	-0.20642	0.56858	0.56463	0.89324	0.07501	0.91294	-0.41434
P-Value	0.0118	0.3041	-	<.0001	0.0513	0.3447	0.0046	0.0050	<.0001	0.7338	<.0001	0.0493
濁度	0.23984	0.08488	0.83642	1	0.28067	-0.18945	0.41789	0.33944	0.73398	0.19102	0.78125	-0.32025
P-Value	0.2703	0.7002	<.0001	-	0.1946	0.3866	0.0472	0.1131	<.0001	0.3826	<.0001	0.1363
氨氮	0.50954	0.12262	0.41110	0.28067	1	-0.09423	0.66001	0.18031	0.46998	-0.06907	0.44310	-0.28229
P-Value	0.0130	0.5773	0.0513	0.1946	-	0.6689	0.0006	0.4103	0.0236	0.7542	0.0342	0.1919
硝酸鹽	-0.25679	0.41161	-0.20642	-0.18945	-0.09423	1	-0.14711	-0.25712	-0.36188	0.17154	-0.38844	-0.17324
P-Value	0.2369	0.0510	0.3447	0.3866	0.6689	-	0.5030	0.2363	0.0897	0.4338	0.0670	0.4292
總磷	0.39430	0.32571	0.56858	0.41789	0.66001	-0.14711	1	0.47440	0.51494	-0.08203	0.49371	-0.41443
P-Value	0.0626	0.1294	0.0046	0.0472	0.0006	0.5030	-	0.0222	0.0119	0.7098	0.0167	0.0493
磷酸鹽	0.18304	0.05465	0.56463	0.33944	0.18031	-0.25712	0.47440	1	0.47187	-0.11601	0.55207	-0.07709
P-Value	0.4032	0.8044	0.0050	0.1131	0.4103	0.2363	0.0222	-	0.0230	0.5981	0.0063	0.7266
COD	0.52679	0.17170	0.89324	0.73398	0.46998	-0.36188	0.51494	0.47187	1	-0.12060	0.95137	-0.26523
P-Value	0.0098	0.4334	<.0001	<.0001	0.0236	0.0897	0.0119	0.0230	-	0.5836	<.0001	0.2213
鈉	0.19034	-0.13557	0.07501	0.19102	-0.06907	0.17154	-0.08203	-0.11601	-0.12060	1	-0.03395	-0.11777
P-Value	0.3843	0.5374	0.7338	0.3826	0.7542	0.4338	0.7098	0.5981	0.5836	-	0.8778	0.5925
葉綠素a	0.44145	0.09469	0.91294	0.78125	0.44310	-0.38844	0.49371	0.55207	0.95137	-0.03395	1	-0.17978
P-Value	0.0350	0.6674	<.0001	<.0001	0.0342	0.0670	0.0167	0.0063	<.0001	0.8778	-	0.4118
透明度	-0.61396	-0.68963	0.41434	-0.32025	-0.28229	-0.17324	-0.41443	-0.07709	-0.26523	-0.11777	-0.17978	1
P-Value	0.0018	0.0003	0.0493	0.1363	0.1919	0.4292	0.0493	0.7266	0.2213	0.5925	0.4118	-

註：以粗體字標記者為達統計水準 5% 顯著相關者。

(3) 透明度

在 5% 顯著水準下，透明度除了水溫與溶氧外，亦和懸浮固體及總磷達顯著相關。其中，懸浮固體含量會直接影響透明度，而總磷所代表的藻類生長潛力亦會影響透明度的量測，因此，它們之間呈現顯著相關的結果亦為合理。

由顯著水準 1% 與 5% 的顯著相關性檢定結果可知，以 5% 為顯著水準時，達顯著

相關的變數均有其合理性，因此，為了減少因誤判而導致的誤差，本研究仍建議以 5% 為顯著水準，來探討變數間的相關性，然後再針對顯著相關的部分作進一步的分析。

三、多元共線性檢定

經過複迴歸參數顯著水準檢定與兩自變數相關分析後，均發現模式中可能有共線性的問題存在，因此，進一步使用統計方法來確定其共線性，並檢測出是哪幾個自變數共

線性最嚴重。本研究共使用 5 種檢定方法，分別是容忍度、變異數波動因素、特徵值、條件指數與變異數比例，分別討論如後。

1. 容忍度與變異數波動因素

容忍度 (tolerance) 與變異數波動因素 (variance inflation, VIF) 兩種共線性問題檢定法之檢定結果如表 6 所示，由表中可知，容忍度以葉綠素 a 之值最小，表示此一指標與其他變項之共線性問題最為嚴重。另外，在變異數波動因素的欄位中，懸浮固體 COD 與葉綠素 a 之數值均大於 10，表示這三個水質指標有較嚴重的共線性問題。由於兩種檢定法互為倒數，因此檢定結果有相似的情形。

表 6. 容忍度與變異數波動因素之共線性檢定表

Table 6. Colinearity tests of tolerance and VIF.

Variable	DF	Tolerance	VIF
截距	1	-	0
水溫	1	0.14013	7.13623
溶氧	1	0.22849	4.37649
懸浮固體	1	0.04548	21.98720
濁度	1	0.13560	7.37477
氨氮	1	0.24827	4.02789
硝酸鹽	1	0.25044	3.99290
總磷	1	0.25513	3.91959
磷酸鹽	1	0.42865	2.33290
COD	1	0.05366	18.63539
鈉	1	0.38674	2.58571
葉綠素a	1	0.03378	29.59916
透明度	1	0.18986	5.26709

註：以粗體標記者為大於該檢定門檻值者。

懸浮固體的量測對象為水體中的懸浮微粒，若優養化問題嚴重時，藻類繁殖旺盛，因此水中富含有機微粒，此時懸浮固體的測量就會與葉綠素 a 的量測重疊，均為測出水中藻類族群總量分布。至於 COD，即為量測

生物無法分解的部分，包括藻類的細胞壁，因此，3 測量結果均有重疊的部分，所以這 3 個參數存有共線性的問題。

2. 特徵值

特徵值 (eigenvalue) 計算結果如表 7 第 2 欄所示，第 1 欄則為特徵值由大到小之排序。由於自變數加截距共有 13 個，所以特徵值總和為 13，而排序 12 與 13 兩個最小特徵值均小於 0.01，並接近 0，因此可判斷出模式中可能存有共線性的問題。

3. 條件指數

條件指數 (condition index, CI) 計算結果如表 7 之第 3 欄，其計算方式是依據第 2 欄之特徵值 (eigenvalue) 來計算，結果發現數值大於 30 的有兩個，屬於中度共線性範圍 (30~100)，表示模式存在共線性的問題。

4. 變異數比例

變異數比例 (variance proportions) 即為特徵向量的計算，其結果如表 7 的第 4 欄。在特徵向量的表格中，可發現各列中沒有係數同時比較大而接近 1 者，表示此檢定法在此並無法明顯看出各變項共線性的情形。但若以係數 0.5 作為標準時，則可發現水溫、溶氧、氨氮與硝酸鹽較有可能有共線性的問題存在。

由於藻類主要組成為蛋白質，因此量測在水中的總氮含量即可推測優養化情形，而氨氮與硝酸鹽即為水中氮含量的指標。另外，氮含量亦與水溫、溶氧量有關，因此 4 個因子間存有共線性十分合理。而在表 4 中，我們亦以「斜線框」框出前述 4 因子的相關性，由表中可知此 4 因子的相關性均不屬於中高度相關的範疇，故較懸浮固體 COD 與葉綠素 a 三者相關性低，因此，共線性仍以懸浮固體、COD 與葉綠素 a 三者較強。

經過以上 5 種共線性檢測方法檢驗後，

表 7. 特徵值、條件指數、變異數比例之共線性檢定表

Table 7. Colinearity tests of eigenvalue, condition index, and variance proportions.

Number	Eigenvalue	Condition Index	Variance Proportions					
			Intercept	水溫	溶氧	懸浮固體	濁度	氨氮
1	10.09437	1.00000	0.0000228	0.0000543	0.0000603	0.0001430	0.0003891	0.0003987
2	1.38496	2.69973	0.0000953	0.0001023	0.0002042	0.0008216	0.0008315	0.0001400
3	0.50619	4.46561	0.0000540	0.0000054	0.0000953	0.0012100	0.0103400	0.0000179
4	0.34102	5.44065	0.0000001	0.0000803	0.0002213	0.0004880	0.0080600	0.0256000
5	0.27654	6.04167	0.0002075	0.0000017	0.0000518	0.0031583	0.0013500	0.0007405
6	0.14303	8.40094	0.0000942	0.0000029	0.0061700	0.0053800	0.0139000	0.0415600
7	0.10557	9.77840	0.0000017	0.0094000	0.0010100	0.0008035	0.1749000	0.0122600
8	0.06122	12.84031	0.0011500	0.0061500	0.0068400	0.0000197	0.0238300	0.3606700
9	0.04690	14.67089	0.0032900	0.0219500	0.0027200	0.0054400	0.2032500	0.0455300
10	0.01904	23.02362	0.0001853	0.0103700	0.0425600	0.5640600	0.0581800	0.0001815
11	0.01586	25.23026	0.0002798	0.0010600	0.0447300	0.0112000	0.0444900	0.0000031
12	0.00343	54.26813	0.0032000	0.5232900	0.6631800	0.4101100	0.1168400	0.5103600
13	0.00187	73.54101	0.9914200	0.4275200	0.2321500	0.0000099	0.3436400	0.0025500

Number	Variance Proportions						
	硝酸鹽	總磷	磷酸鹽	COD	鈉	葉綠素a	透明度
1	0.0001957	0.0007347	0.0010800	0.0001842	0.0007694	0.0001281	0.0001836
2	0.0019500	0.0012200	0.0686900	0.0015900	0.0065600	0.0018700	0.0015600
3	0.0007073	0.0030100	0.4216700	0.0038400	0.0014900	0.0024500	0.0027200
4	0.0000088	0.0872100	0.0656000	0.0002597	0.1090600	0.0005389	0.0018900
5	0.0000655	0.0641600	0.0055900	0.0055600	0.1333700	0.0034400	0.0316400
6	0.0500700	0.0199900	0.0062300	0.0000665	0.1121800	0.0036000	0.0309400
7	0.0001203	0.1321500	0.0128100	0.0164700	0.0286300	0.0022700	0.0333700
8	0.0579600	0.2330500	0.0835800	0.0106200	0.0048100	0.0006052	0.0025900
9	0.1457900	0.1216200	0.0789000	0.0182000	0.0579000	0.0239000	0.0144600
10	0.0047000	0.0002418	0.1148200	0.0816700	0.0825200	0.0621700	0.0433600
11	0.0190500	0.0000718	0.1278400	0.6749400	0.0166400	0.4423200	0.0044800
12	0.7039100	0.3343900	0.0023500	0.1489800	0.3954300	0.4288300	0.0521500
13	0.0154700	0.0021700	0.0108600	0.0376300	0.0506400	0.0278700	0.7806300

註：以粗體標記者為大於該檢定之門檻值者。

可發現模式所使用的參數中，以懸浮固體、COD 與葉綠素 a 這 3 個參數 (水中有機物含量指標) 之共線性問題最為嚴重。為了降低共線性的影響，提升模式對變異度之解釋能力，未來必須採用去除模式共線性的方法來最佳化模式。

四、殘差分析

在殘差分析中，可依目的分為兩部分來

進行，第一部分為自我相關性檢測 — 用來檢驗殘差間是否有自我相關的情形，亦即測定殘差間是否獨立，這部分本研究使用 Durbin-Watson D 檢驗法來檢定。第二部分為離群值的判定分析 — 使用一些統計方法來檢測觀測值中可能的離群值，然後再針對可能的離群值做進一步的探討。離群值判定研究中共使用標準化殘差值、t 標準化殘差值、Rstudent、Cook's D 距離值、Hat Diag H、共

變數率、刪除後預測差異與刪除後標準化差異等共 8 種方法來檢定。

是直接由模式殘差值來加以判斷，若標準化後之殘差值大於 2 或 3 個標準差時，該觀測值即有可能為離群值；第二類方法為找出最

1. 自我相關性檢測

經過 Durbin-Watson D 檢驗法之檢定結果如表 8 所示，由表中可知 DW 值 (= 1.968) 接近於 2，且誤差間之一次相關性等於 -0.005，為低相關性，代表殘差間彼此獨立，表示此時間數列中無殘差自我相關情形。

表 8. Durbin-Watson D 檢驗法之檢定結果

Table 8. Results of Durbin-Watson D test.

Durbin-Watson D	1.968
Number of Observations	23
1st Order Autocorrelation	-0.005

2. 離群值判定分析

離群值的判定方法可分為兩類，第一類

表 9. 殘差分析表

Table 9. Residual analysis.

Obs.	Dep. Var. TSI	Predicted Value	Std. Error Mean Prediction	Redidual	Std. Error Residual	Standardized Residual	Student Residual	Rstudent
1	76.0790	79.5632	4.7246	-3.4842	3.3400	-0.8932	-1.0430	-1.0485
2	41.2470	45.8149	4.1379	-4.5679	4.0440	-1.1710	-1.1300	-1.1473
3	34.4920	44.7870	2.8651	-10.2950	5.0260	-2.6393	-2.0480	-2.5503
4	47.1060	46.3984	3.4129	0.7076	4.6720	0.1814	0.1510	0.1438
5	79.8120	83.3164	4.7464	-3.5044	3.3090	-0.8984	-1.0590	-1.0664
6	71.2600	71.1331	3.7441	0.1269	4.4110	0.0325	0.0288	0.0273
7	71.3060	67.5370	4.7545	3.7690	3.2970	0.9662	1.1430	1.1632
8	37.4970	44.6768	3.2206	-7.1798	4.8060	-1.8406	-1.4940	-1.6078
9	58.4420	52.3183	2.7552	6.1237	5.0880	1.5699	1.2040	1.2348
10	66.4720	67.1935	4.4469	-0.7215	3.7010	-0.1850	-0.1950	-0.1853
11	66.9120	66.7136	4.3375	0.1984	3.8290	0.0508	0.0518	0.0492
12	77.2440	77.8119	5.7175	-0.5679	0.8850	-0.1456	-0.6410	-0.6214
13	53.6680	51.4859	5.0792	2.1821	2.7710	0.5594	0.7880	0.7715
14	58.0470	55.2946	4.8085	2.7524	3.2180	0.7056	0.8550	0.8430
15	62.7840	62.3449	5.7573	0.4391	0.5720	0.1126	0.7670	0.7500
16	72.7790	73.1358	3.0456	-0.3568	4.9190	-0.0915	-0.0725	-0.0688
17	48.9400	50.8348	4.0479	-1.8948	4.1340	-0.4857	-0.4580	-0.4395
18	52.5400	47.0556	3.5331	5.4844	4.5820	1.4059	1.1970	1.2269
19	63.2460	62.5718	5.4295	0.6742	1.9990	0.1729	0.3370	0.3218
20	73.2660	71.1685	4.6902	2.0975	3.3880	0.5377	0.6190	0.5990
21	55.1410	53.2115	2.8052	1.9295	5.0600	0.4946	0.3810	0.3644
22	55.6900	51.0054	4.7731	4.6846	3.2700	1.2010	1.4330	1.5246
23	45.8290	44.4258	5.0134	1.4032	2.8880	0.3597	0.4860	0.4665

註：以粗體標記者為大於該檢定之門檻值者。

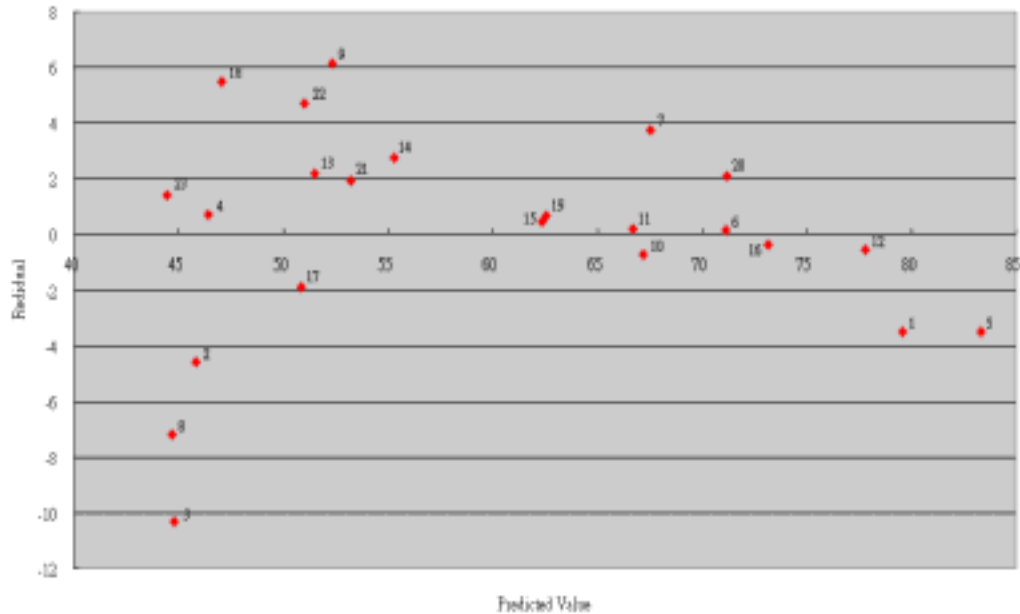


圖 2. 殘差分布圖

Figure 2. Distribution of residual.

具影響力之觀測值，再逐一刪除這些觀測值，以找出刪除後對模式的影響。

(1) 殘差值分析：

殘差值之計算結果為表 9 的第 5 欄，分布情形則如圖 2 所示。由圖中可看出模型預測的趨勢大部分有偏高的情形，顯示此迴歸模型的預測結果會有偏大的傾向。本研究採用的殘差值標準化方法共有 3 種，包括標準化殘差值、t 標準化殘差值與 Rstudent，將一一分述如下：

A) 標準化殘差值 — 將殘差標準化後，結果為表 9 中的第 7 欄，其分布則如圖 3 所示。圖 3 與圖 2 相似，均以預測值為 x 軸，殘差值計算為 y 軸，且在每個座標點均標示觀測值的序號，因此，可從圖 3 明顯看出大部分的數值並未大於本研究所採用的 2 個標準差門檻，僅第 3 個觀測

值超過這個標準，為可能極端值之位置。

B) t 標準化殘差值 — t 標準化殘差值亦可用來檢定極端值的存在與否，其計算結果如表 9 的第 8 欄。其分布情形如圖 4，由圖中可看出亦只有第 3 個觀測值之絕對值大於 2 個標準差，表示該組觀測值有可能是極端值。

C) Rstudent — Rstudent 檢定值為表 9 之第 9 欄，由計算結果同樣可看出只有第 3 個觀測值之檢定值大於 2，為可能的離群值，其分布情形則如圖 5 所示。

經過殘差值的檢定後，得到第 3 個觀測值為可能的離群值之一致結果，然而必須進一步針對此觀測值計算其對模式解釋能力之影響來加以判斷。

(2) 刪除觀測值殘差分析法：

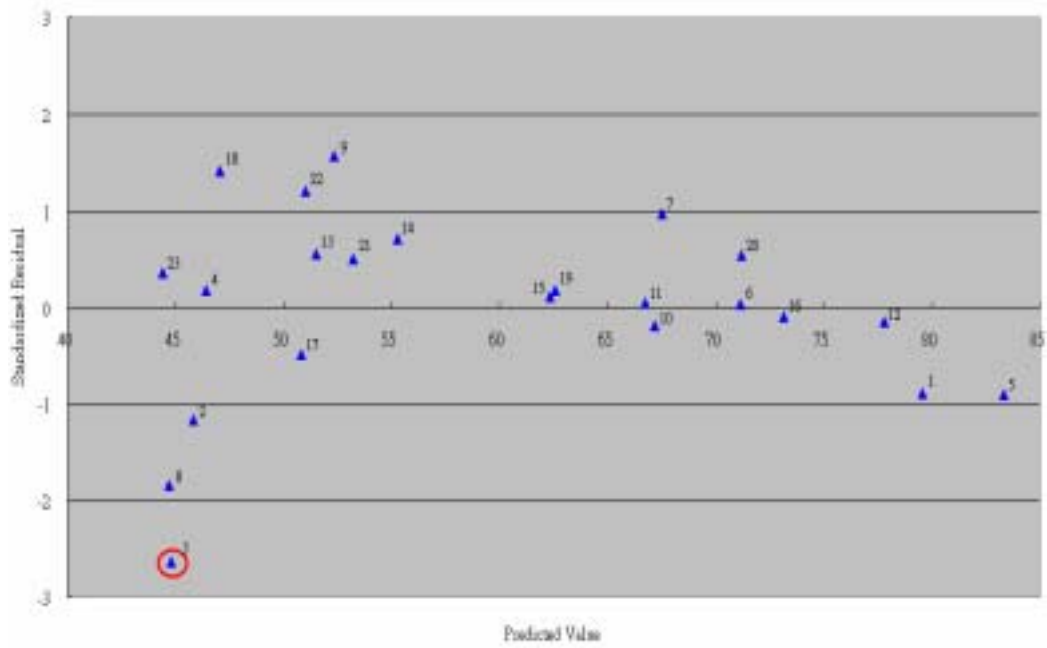


圖 3. 標準化殘差分布圖
 Figure 3. Distribution of standardized residual.

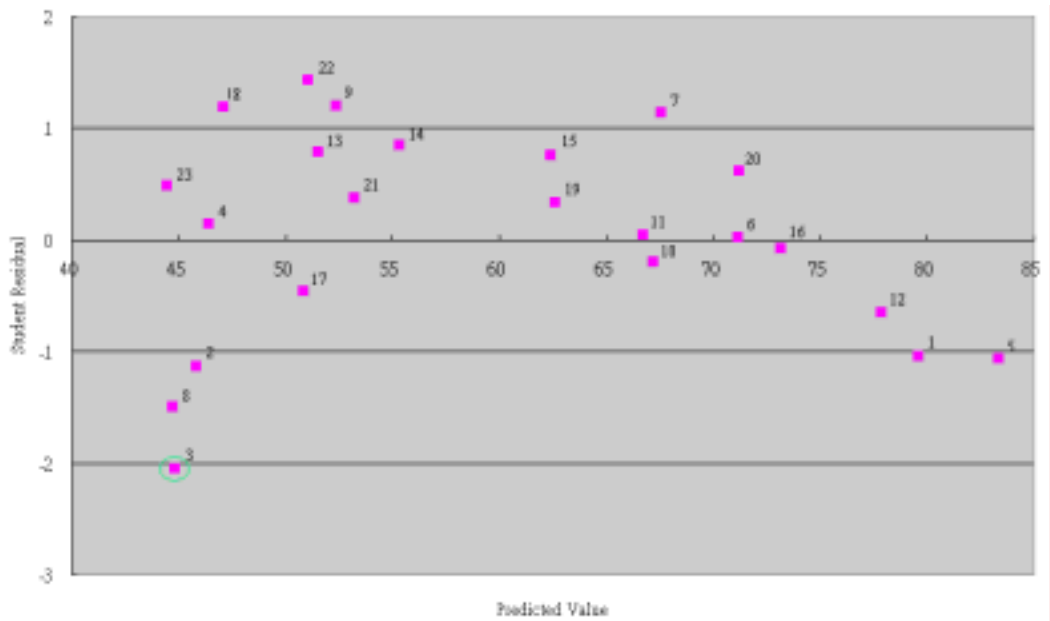


圖 4. t 標準化殘差分布圖
 Figure 4. Distribution of t standardized residual.

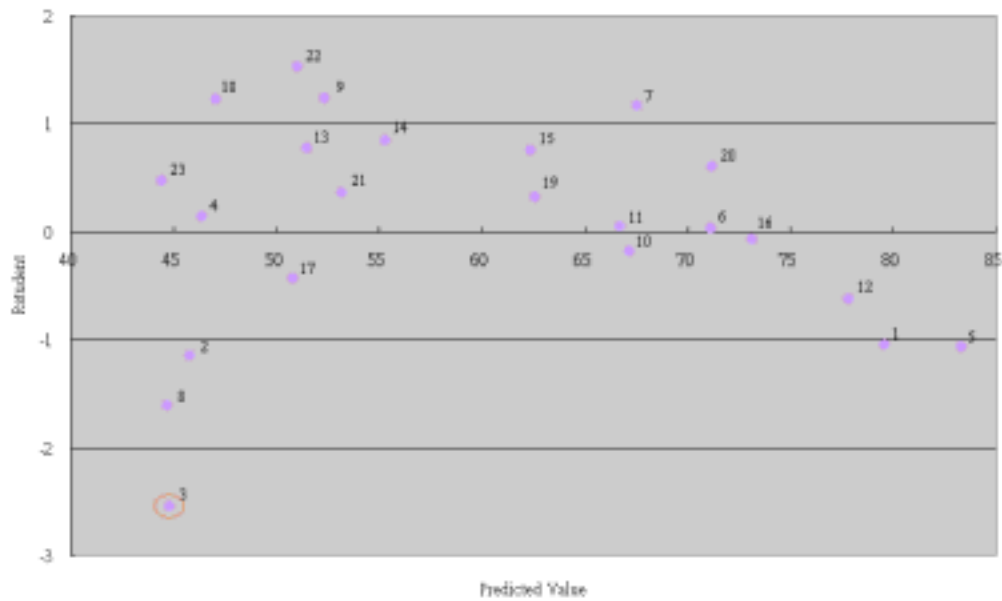


圖 5. Rstudent 殘差分布圖

Figure 5. Distribution of Rstudent residual.

表 10. 具影響力之觀測值的檢定結果

Table 10. Test results of influential observed values.

Obs.	Cook's D	Hat Diag H	Cov. Ratio	DFFITs
1	0.168	0.6668	2.6397	-1.4832
2	0.103	0.5115	1.3655	-1.1740
3	0.105	0.2452	0.0044	-1.4537
4	0.001	0.3480	5.8561	0.1051
5	0.178	0.6730	2.5612	-1.5299
6	0.000	0.4188	6.7613	0.0232
7	0.209	0.6753	1.9617	1.6774
8	0.077	0.3099	0.2140	-1.0773
9	0.033	0.2268	0.6651	0.6687
10	0.004	0.5908	9.1489	-0.2226
11	0.000	0.5620	8.9513	0.0557
12	1.319	0.9766	97.2794	-4.0122

Obs	Cook's D	Hat Diag H	Cov. Ratio	DFFITs
13	0.160	0.7707	7.4625	1.4144
14	0.126	0.6907	4.7367	1.2598
15	4.577	0.9902	182.7149	7.5433
16	0.000	0.2771	5.4050	-0.0426
17	0.015	0.4895	5.8474	-0.4303
18	0.066	0.3729	0.8402	0.9461
19	0.065	0.8806	28.4050	0.8741
20	0.057	0.6572	6.9035	0.8292
21	0.003	0.2351	4.2514	0.2020
22	0.336	0.6806	0.6215	2.2255
23	0.055	0.7508	11.5744	0.8098

註：以粗體標記者為大於該檢定之門檻值者

對模式最具影響力之觀測值，亦可能為離群值的位置。本研究共採用 5 種刪除檢定法來計算觀測值的影響力，分別是 Cook's D 距離值、Hat Diag H、共變數率、刪除後預測

差異與刪除後標準化差異。

A) Cook's D 距離值 — 計算結果為表 10 第 2 欄，發現第 12 與 15 個觀測值大於其檢定

門檻值 0.5，代表此兩者可能是離群值

- B) Hat Diag H — 計算結果為表 10 的第 3 欄。由於模式共採用 12 個參數，使用 23 個觀測值，因此其鑑定門檻值計算後為 1.130。由欄中數據可發現並無觀測值大於門檻值的情形，表示此檢測法無法找出觀測值中的離群值。
- C) 共變數率 (covariance ratio) — 經共變數率檢定之計算結果如表 10 之第 4 欄，可明顯看出第 12 與 15 個觀測值的檢定值最大，且遠大於其檢定門檻值 3.9，因此最有可能為離群值。
- D) 刪除後預測差異 (DFFITs) — DFFITS 值之計算結果為表 10 中的第 5 欄，發現第 12 與 15 個觀測值仍是最大檢定值，且超過檢定標準 1.504，表示這兩個觀測值為離群值的可能性最大。
- E) 刪除後標準化差異 (DFBETA) — 在 DFBETA 檢定下，其鑑定結果如表 11。因本研究共有 23 個觀測值，故樣本數屬於小樣本，所以門檻值為 1。由表中可看出第 7、12、15 與 22 個觀測值達門檻值標準，均大於 1，而其中又以第 12 與 15 個觀測值有最大改變值，與前面的 Cook's D 距離值、共變數率、刪除後預測差異等分析法有一致的結果。

因此，經過上述 8 種殘差分析檢定後，最後結果整理如表 12。發現本研究所採用的 23 組觀測值中，第 3、7、12、15 與 22 個觀測值的殘差值與影響力相對較大，表示這 5 個觀測值最有可能為離群值所在。

為了檢驗經殘差分析後所得到第 3、7、12、15 與 22 個觀測值為離群值之結論是否正確，本研究將此 5 個可能為離群值之觀測值與其可能組合，共計 31 種組合，分別由模

式剔除後重新進行複迴歸分析，再與原來使用全部觀測值的複迴歸模式進行比較，由 R^2 的變化觀察其模式解釋能力的改變程度。其計算結果經排序後如表 13 所示，由表中第 5 欄之迴歸模式 F 統計顯著水準值可發現，所有模式中均達 1% 以上的統計顯著水準，表示所有模式均為合理模式。

此外，在 R^2 值的變化上，就刪除單一觀測值方面，發現除了刪除第 12 個觀測值後，模式 R^2 值有下降的情形外，其他均是刪除後有所提升，而其中又以刪除第 3 個觀測值後上升最多。因此，5 個可能的離群值中，以第 3 個觀測值對模式解釋能力之影響最強，而第 12 個觀測值為離群值之可能性最低。另外，在其可能的刪除組合中，以刪除第 3、7、15 與 22 個觀測值後，模式之 R^2 值最高，可達 0.9664。

由於刪除觀測值後所形成的模式觀測值個數與原來複迴歸模式不同，因此使用調整後的 R^2 值 (Adj. R^2) 校正總觀測值個數後再進行觀察會較為準確，其計算排序如表 14。結果發現在刪除單一觀測值檢定方面，刪除第 3、7 與 8 個觀測值後解釋能力增加；刪除第 12、15 則降低。而 5 個可能的離群值中，仍以第 3 個觀測值之影響力最強，亦即最有可能為離群值。另外，在各種可能的離群值組合中，仍以刪除第 3、7、15、22 個觀測值之複迴歸模式解釋能力最佳，調整後 R^2 值亦最高，可達 0.8993，相較於原模式的 0.8056 有所提升。因此，更可確定 5 個可能的離群值中，可能性最低的為第 12 個觀測值。

再者，經過離群值修正後，可得到本研究所採用之優養化全模式，亦即刪除原本觀測值中第 3、7、15 與 22 個觀測值的複迴歸模式，其變方分析結果為表 15。由於最後使用 19 個觀測值，所以總自由度變為 18，其 F 檢定值為 14.40，達 99.81% 的顯著水準； R^2

表 11. 刪除後標準化差異值之計算結果

Table 11. Results of DFBETA values.

Obs.	DFBETA Values						
	Intercept	水溫	溶氧	懸浮固體	濁度	氨氮	硝酸鹽
1	-0.1445	-0.1519	0.3908	0.2220	-0.6463	-0.0782	-0.0313
2	-0.9922	0.5475	0.4948	0.3220	0.4199	0.1882	-0.0843
3	0.5430	-0.0752	-0.7962	-0.3335	-0.2232	-0.3599	0.7096
4	0.0179	-0.0351	0.0189	0.0156	-0.0142	0.0804	-0.0295
5	0.3322	0.0717	-0.3714	0.3391	-0.1987	-0.5374	-0.0493
6	-0.0010	-0.0086	0.0104	0.0142	-0.0058	0.0036	-0.0131
7	-0.1788	0.3411	-0.4164	-0.2131	0.0035	-0.8729	0.5724
8	-0.3007	-0.3215	0.6537	0.5849	-0.1442	0.7261	-0.6381
9	0.3751	-0.3693	0.0512	-0.0656	-0.1659	-0.0166	-0.1954
10	-0.0774	0.1708	-0.0726	-0.1701	0.1370	-0.0927	0.1233
11	-0.0026	0.0137	-0.0016	0.0040	-0.0035	-0.0220	0.0121
12	0.0076	-0.2206	0.3446	0.0795	-0.2598	0.2850	-0.3530
13	-0.2334	-0.0316	0.3101	-0.2651	0.1390	0.1226	-0.0416
14	-0.0589	0.2624	-0.3341	0.2468	0.1615	0.0781	0.6119
15	-0.5951	0.1617	-0.1451	-0.0769	0.0292	-0.0434	0.1434
16	-0.0023	-0.0039	0.0102	0.0089	0.0002	0.0163	-0.0163
17	0.1919	-0.1768	-0.0133	-0.1578	-0.0512	-0.0357	-0.0413
18	0.2830	-0.1698	-0.2973	-0.3283	0.0175	0.3304	0.1456
19	0.0615	0.2113	-0.2248	-0.0623	-0.2888	-0.1358	0.1262
20	0.0959	0.0114	0.0973	-0.0191	-0.2446	0.0247	-0.2372
21	-0.0032	-0.0235	0.0132	0.0890	-0.0260	0.0411	-0.0728
22	-0.1588	0.7407	-0.0005	-0.2859	1.3203	-0.4057	-0.4652
23	-0.4167	0.1161	0.4366	-0.1976	0.2733	-0.0453	0.0631

Obs.	DFBETA Values					
	總磷	磷酸鹽	COD	鈉	葉綠素a	透明度
1	-0.2211	0.2834	-0.1085	0.3331	0.0791	0.1669
2	0.0192	-0.0415	-0.1554	-0.1828	-0.3477	0.9697
3	0.5694	-0.2253	0.0093	-0.1461	0.5238	-0.7662
4	-0.0726	0.0308	0.0041	0.0161	-0.0193	-0.0169
5	-0.3339	-0.1441	-0.0607	-0.1477	0.0104	-0.1178
6	-0.0009	-0.0031	0.0105	0.0067	-0.0169	0.0057
7	1.2838	-0.4655	-0.0214	-0.2250	0.1660	0.5203
8	-0.4173	-0.0003	0.1114	0.3821	-0.4889	0.2832
9	0.0008	-0.0559	0.0239	0.1234	0.1010	-0.2962
10	0.0814	0.0665	-0.0466	-0.0979	0.0824	0.0254
11	0.0126	-0.0299	-0.0382	-0.0211	0.0337	-0.0048
12	-0.0798	-2.8070	-0.0600	0.1756	0.1663	0.0300
13	-0.3833	-0.0451	-0.0856	0.0815	0.4453	0.2486
14	-0.1822	0.0184	-0.3679	-0.5634	0.0218	-0.0739
15	0.3368	0.1107	-0.0798	4.5202	0.0497	0.1546
16	-0.0178	0.0121	-0.0151	0.0050	-0.0013	0.0059
17	0.0349	0.0815	0.2113	0.1703	-0.0041	-0.2673
18	-0.2195	0.2869	0.3320	0.0987	-0.0331	-0.2069
19	-0.0153	-0.0302	0.2569	-0.0048	-0.0156	-0.1486
20	0.1577	-0.1173	-0.4325	-0.0118	0.3662	-0.1293
21	-0.0369	-0.0226	0.0552	0.0139	-0.1041	0.0774
22	-0.0667	0.3634	-0.2646	-0.6370	-0.3794	0.0598
23	-0.1371	0.2461	0.2337	0.0879	-0.1764	0.2607

註：以粗體標記者為大於該檢定之門檻值者。

表 12. 殘差分析判定離群值之檢定結果

Table 12. To adjudge outliers by residual analysis.

殘差檢定法	可能為離群值之觀測值
標準化殘差值	3
t 標準化殘差值	3
Rstudent	3
Cook's D	12、15

殘差檢定法	可能為離群值之觀測值
Hat Diag H	無
Cov. Ratio	12、15
DFFITS	12、15
DFBETA	7、12、15、22

表 13. 刪除可能的離群值後與原模式之比較
(以 R² 進行排序)

Table 13. Comparisons between original and adjusted model after deleting possible outliers.
(order by R²)

排序	被刪除的觀測值	R ²	Adj. R ²	p-value
1	3, 7, 15, 22	0.9664	0.8993	0.0019
2	3, 7, 12, 15, 22	0.9641	0.8778	0.0076
3	3, 7, 15	0.9608	0.8935	0.0009
4	7, 15, 22	0.9581	0.8862	0.0011
5	3, 7, 12, 15	0.9563	0.8688	0.0040
6	7, 12, 15, 22	0.9549	0.8647	0.0043
7	3, 7, 22	0.9499	0.8641	0.0020
8	3, 15, 22	0.9498	0.8637	0.0020
9	3, 7	0.9487	0.8717	0.0007
10	3, 12, 15, 22	0.9461	0.8383	0.0071
11	3, 7, 12, 22	0.9444	0.8332	0.0078
12	15, 22	0.9441	0.8602	0.0010
13	3, 22	0.9433	0.8583	0.0010
14	3, 7, 12	0.9431	0.8455	0.0030
15	12, 15, 22	0.9400	0.8372	0.0036
16	7, 15	0.9398	0.8494	0.0013
17	3, 15	0.9395	0.8488	0.0013
18	3	0.9379	0.8552	0.0005
19	3, 12, 22	0.9379	0.8315	0.0040
20	7, 12, 15	0.9348	0.8230	0.0047
21	3, 12, 15	0.9337	0.8199	0.0049
22	3, 12	0.9321	0.8302	0.0020
23	7, 22	0.9312	0.8280	0.0021
24	22	0.9295	0.8354	0.0009
25	7, 12, 22	0.9241	0.7939	0.0076
26	12, 22	0.9229	0.8073	0.0032
27	7	0.9201	0.8136	0.0015
28	15	0.9166	0.8054	0.0018
29	7, 12	0.9150	0.7875	0.0046
30	12, 15	0.9118	0.7795	0.0053
31	原複迴歸模式	0.9116	0.8056	0.0009
32	12	0.9073	0.7836	0.0027

表 14. 刪除可能的離群值後與原模式之比較
(以調整後的 R² 進行排序)

Table 14. Comparisons between original and adjusted model after deleting possible outliers.
(order by adj. R²)

排序	被刪除的觀測值	R ²	Adj. R ²	p-value
1	3, 7, 15, 22	0.9664	0.8993	0.0019
2	3, 7, 15	0.9608	0.8935	0.0009
3	7, 15, 22	0.9581	0.8862	0.0011
4	3, 7, 12, 15, 22	0.9641	0.8778	0.0076
5	3, 7	0.9487	0.8717	0.0007
6	3, 7, 12, 15	0.9563	0.8688	0.0040
7	7, 12, 15, 22	0.9549	0.8647	0.0043
8	3, 7, 22	0.9499	0.8641	0.0020
9	3, 15, 22	0.9498	0.8637	0.0020
10	15, 22	0.9441	0.8602	0.0010
11	3, 22	0.9433	0.8583	0.0010
12	3	0.9379	0.8552	0.0005
13	7, 15	0.9398	0.8494	0.0013
14	3, 15	0.9395	0.8488	0.0013
15	3, 7, 12	0.9431	0.8455	0.0030
16	3, 12, 15, 22	0.9461	0.8383	0.0071
17	12, 15, 22	0.9400	0.8372	0.0036
18	22	0.9295	0.8354	0.0009
19	3, 7, 12, 22	0.9444	0.8332	0.0078
20	3, 12, 22	0.9379	0.8315	0.0040
21	3, 12	0.9321	0.8302	0.0020
22	7, 22	0.9312	0.8280	0.0021
23	7, 12, 15	0.9348	0.8230	0.0047
24	3, 12, 15	0.9337	0.8199	0.0049
25	7	0.9201	0.8136	0.0015
26	12, 22	0.9229	0.8073	0.0032
27	原複迴歸模式	0.9116	0.8056	0.0009
28	15	0.9166	0.8054	0.0018
29	7, 12, 22	0.9241	0.7939	0.0076
30	7, 12	0.9150	0.7875	0.0046
31	12	0.9073	0.7836	0.0027
32	12, 15	0.9118	0.7795	0.0053

表 15. 優養化全模式之變異分析

Table 15. ANOVA of the full model of eutrophication.

Source	DF	SS	MS	F value	Pr > F	R ²	Adj. R ²
Model	12	2885.931	240.4942	14.40	0.0019	0.9664	0.8993
Error	6	100.2112	16.70187				
Corrected Total	18	2986.142					

表 16. 優養化全模式之參數估計表

Table 16. Parameter estimation of the full model of eutrophication.

Variable	DF	Parameter Estimates	Standard Error	t value	Pr > t
截距	1	77.69014	26.22922	2.961969	0.025219
水溫	1	-0.98623	0.703218	-1.40246	0.210344
溶氧	1	1.609134	1.251053	1.286223	0.245766
懸浮固體	1	1.392002	0.477315	2.916319	0.026759
濁度	1	-1.27132	0.788522	-1.61228	0.158027
氨氮	1	24.92924	17.06291	1.461019	0.194315
硝酸鹽	1	-13.6577	9.100533	-1.50076	0.184091
總磷	1	0.108723	0.061176	1.777227	0.125859
磷酸鹽	1	-0.05942	0.040555	-1.4652	0.193214
COD	1	0.035232	0.239473	0.147123	0.887854
鈉	1	-6.51942	4.218651	-1.54538	0.173209
葉綠素a	1	-0.02211	0.034421	-0.6422	0.544487
透明度	1	-1.88521	2.811001	-0.67065	0.527398

表 17. 原模式與全模式之比較

Table 17. Comparisons between original and full model.

模 式	R ²	Adj. R ²
原模式	0.9116	0.8056
全模式	0.9664	0.8993

值原為 0.9664，調整後亦可達 0.8993 的高解釋能力。全模式之參數估計結果如表 16 所示，由表中 t 檢定結果發現，僅懸浮固體參數檢定值達 95% 的顯著水準，表示模式仍有

共線性問題存在。而其中最接近標準者為總磷，表示此模式仍以總磷與懸浮固體參數之影響最大，亦即為最重要的參數。

最後，再將原來的複迴歸模式與優養化全模式兩者互相比較 (表 17)，可知原模式的 R² 值為 0.9116，全模式則提升為 0.9664；而在調整後的 R² 值比較方面，全模式為 0.8993，亦較原模式 (= 0.8056) 高。由此可知，經離群值修正後，模式對變異度的解釋能力有所提升，因此，最後即採此複迴歸式作為優養化全模式，日後將可藉由此全模式為基礎，作進一步的相關分析。

結論與建議

經過複迴歸分析後，證明優養化現象可考慮以水溫、溶氧、懸浮固體、濁度、氨氮、硝酸鹽、總磷、磷酸鹽、化學需氧量、鈉、葉綠素 a 及透明度等 12 個因子來表現其動力機制。產生的迴歸模型解釋變異度可達 0.9116，其中又以總磷因子最為重要，證明淡水水域優養化現象是以水中磷的含量為控制因子 (Lau, 2002)，而磷的主要來源一般為農業開發 (Smith, 2005)，因此，推論與梨山地區的農業開發息息相關。

經相關性檢定與共線性檢定後發現，懸浮固體、COD 與葉綠素 a 可能有較嚴重的共線性，因此會影響整個迴歸模型的解釋度，使模式預測能力下降。至於殘差分析檢定，結果顯示：(1) 此時間數列並無自我相關情形，不需要再做自我相關校正。(2) 第 3、7、8、15 與 22 個觀測值可能為極端值，或是影響力過大的觀測值，因此，必須藉由個別刪除後對模式解釋能力的改變程度來判別該觀測值之重要程度，結果發現以刪除第 3、7、15 與 22 個觀測值後之複迴歸模式表現最佳，解釋能力 R^2 值可達 0.9664，而調整後的 R^2 值亦達 0.8993，均較原模式為佳，因此最後本研究以此模式作為優養化全模式。

由於研究發現使用水溫、溶氧、懸浮固體、濁度、氨氮、硝酸鹽、總磷、磷酸鹽、化學需氧量、鈉、葉綠素 a 及透明度等 12 個因子可高度解釋優養化的發生機制，因此，日後在進行相關模式研究時，可考慮將這些因子納入，以提升模式預測的準確性。至於模式中仍存在有共線性的問題，則冀望以後能進行更深入的統計研究，如逐步迴歸分析與主成分分析等，來去除共線性，提升模式解釋之變異度。

參考文獻

- 1.汪三華 (2001)，「台股期貨與現貨之價格及報酬率長短期關聯性探討」，碩士論文，中國文化大學。
- 2.林志融 (2004)，「棲地適宜性分析應用於生態廊道規劃之研究 --- 以山羌及有勝溪流流域為例」，碩士論文，東華大學。
- 3.林師模、陳范欽 (2003a)，「多變量分析~管理上的應用」，雙葉書廊有限公司，台北市。
- 4.林羣晏 (2003b)，「土壤之物理性質與力學參數相關性分析 --- 台灣地區山坡地土壤」，碩士論文，中華大學。
- 5.林慧姿、張嘉容、黃春松、廖萃邁 (2002)，「統計學」，新科技書局，台北縣。
- 6.俞其海 (1990)，「實用統計學」，昭人出版社，台中市。
- 7.張秉元 (2004)，「花蓮地區棲地組成與黃嘴角鴉數量之關係」，碩士論文，東華大學。
- 8.陳正昌、程炳林、陳新豐、劉子鍵 (2003)，「多變量分析方法：統計軟體應用」，五南圖書出版股份有限公司，台北市。
- 9.陳登源、楊錦洲、林茂文、蔡豐清 (1997a)，「管理數學」，國立空中大學發行，臺北縣。
- 10.陳鴻烈、梁家柱、鄭慧玲、王久泰 (1999)，「利用相加性季節變動模式之德基水庫優養化時間數列分析」，水土保持學報，第三十一卷，第三期，第 139~144 頁。
- 11.陳鴻烈、梁家柱、羅惠芬、鄭慧玲 (2004)，「水庫優養化時間序列模式分析比較研究」，水土保持學報，第三十六卷，第二期，第 169~178 頁。
- 12.陳鴻烈、鄭慧玲 (1997b)，「台中縣政府大甲溪及德基水庫水質監測研究方案」，台中縣環境保護局委託研究計畫，pp.94~117。

- 13.陳鴻烈、鄭慧玲 (1998), 「水庫優養化之時間數列分析研究」, 水土保持學報, 第三十卷, 第四期, 第 331~337 頁。
- 14.陳鴻烈、蔡大偉 (2005), 「不同模式之預測能力研究」, 水土保持學報, 第三十七卷, 第二期。
- 15.葉怡巖 (2001), 「水庫水質之變異對於淨水廠中三鹵甲烷生成之影響」, 碩士論文, 逢甲大學。
- 16.楊錫賢、簡淑美、丁月琴、謝連德 (2004), 「以迴歸分析解析油品參數對柴油引擎排放懸浮微粒之影響研究」, 2004 年台灣環境資源永續發展研討會論文集。
- 17.經濟部水利署德基水庫集水區管理委員會 (1983~1999), 「德基水庫集水區整體治理計畫水質監測與管理模式研究」。
- 18.潘南飛編譯 (2003), 「工程統計」, 全威圖書有限公司, 台北縣。
- 19.Antikainen, R., R. Lemola, J. I. Nousiainen, L. Sokka, M. Esala, P. Huhtanen and S. Rekolainen (2005), "Stocks and Flows of Nitrogen and Phosphorus in the Finnish Food Production and Consumption System", *Agriculture, Ecosystems & Environment*, Vol.107, pp.287~305.
- 20.Arheimer, B. and R. Lidén (2000), "Nitrogen and Phosphorus Concentrations from Agricultural Catchments — Influence of Spatial and Temporal Variables", *Journal of Hydrology*, Vol.227, pp.140-159.
- 21.Cheng, W. P. and F.-H. Chi (2003), "Influence of Eutrophication on the Coagulation Efficiency in Reservoir Water", *Chemosphere*, Vol.53, pp.773~778.
- 22.Gillen, D. (2005), Lecture 6, *Statistical Methods II Lecture Notes*, Department of Statistics, University of California, Irvine.
- 23.Hiscock, J. G., C. S. Thourot and J. Zhang (2003), "Phosphorus Budget — Land Use Relationships for the Northern Lake Okeechobee Watershed, Florida", *Ecological Engineering*, Vol.21, pp.63-74.
- 24.Katz, A., M. V. Hau and J. Mahoney (2005), *Explaining the Great Reversal in Spanish America: Fuzzy-Set Methods Versus Statistical Methods*, Department of Sociology, Brown University.
- 25.Lau, S. S. S. and S. N. Lane (2002), "Biological and Chemical Factors Influencing Shallow Lake Eutrophication: A Long-Term Study", *Science of the Total Environment*, Vol.288, pp.167~181.
- 26.Nijboer, R. C. and P. F. M. Verdonchot (2004), "Variable Selection for Modelling Effects of Eutrophication on Stream and River Ecosystems", *Ecological Modelling*, Vol.177, pp.17~39.
- 27.Smith, R. V., C. Jordan and J. A. Annett (2005), "A Phosphorus Budget for Northern Ireland: Inputs to Inland and Coastal Waters", *Journal of Hydrology*, Vol.304, pp.193-202.
- 28.Turner-Fairbank HRC (2000), *Speed Prediction for Two-Lane Rural Highways*, Research, Development, and Technology, Turner-Fairbank Highway Research Center, Publication No.99-171, VA.

95 年 07 月 08 日 收稿

95 年 08 月 16 日 修改

95 年 08 月 22 日 接受

水土保持學報 38(3) : 243-266 (2006)

Journal of Soil and Water Conservation, 38(3) : 243-266 (2006)