

行政院國家科學委員會專題研究計畫 成果報告

以基因演算法建立高性能混凝土材料行為模式 研究成果報告(精簡版)

計畫類別：個別型
計畫編號：NSC 96-2221-E-216-032-
執行期間：96年08月01日至97年07月31日
執行單位：中華大學資訊管理學系

計畫主持人：葉怡成

計畫參與人員：碩士-兼任助理人員：張萬鈞
碩士-兼任助理人員：蔡宜真
碩士-兼任助理人員：侯宏孺

處理方式：本計畫可公開查詢

中華民國 97年06月03日

行政院國家科學委員會補助專題研究計畫期中進度報告

以基因演算法建立高性能混凝土材料行為模式(I)

計畫類別： 個別型計畫 整合型計畫

計畫編號：NSC-96-2221-E-216-032-

執行期間：96年8月1日至97年7月31日

計畫主持人：葉怡成

共同主持人：

計畫參與人員：彭建華、張萬鈞、蔡宜真、侯宏孺

成果報告類型(依經費核定清單規定繳交)： 精簡報告 完整報告

本成果報告包括以下應繳交之附件：

- 赴國外出差或研習心得報告一份
- 赴大陸地區出差或研習心得報告一份
- 出席國際學術會議心得報告及發表之論文各一份
- 國際合作研究計畫國外研究報告書一份

處理方式：除產學合作研究計畫、提升產業技術及人才培育研究計畫、列管計畫及下列情形者外，得立即公開查詢

涉及專利或其他智慧財產權， 一年 二年後可公開查詢

執行單位：中華大學

中 華 民 國 97 年 5 月 30 日

以基因演算法建立高性能混凝土材料行為模式

計劃編號：NSC-96-2221-E-216-032-

執行期限：96/08/01~97/07/31

主持人：葉怡成 中華大學土木工程學系 教授

摘要

迴歸分析可以建立外顯的公式來預測高性能混凝土強度，但其產生之公式精確性不良；倒傳遞網路能夠建立相當精確的模型，但無法產生一個外顯的公式；而演化運算樹能產生外顯的公式，其預測的準確度介於迴歸分析與倒傳遞網路之間。此外，雖然演化運算樹能產生公式，但公式經常過於複雜，以致無法解釋其中的意義。因此，本研究發展出一個後向刪除法，來簡化演算運算樹公式的複雜性，這個公式嘗試用「中位常數」替代樹型枝葉端的「變數」，以簡化公式的複雜性，並維持公式的準確度。為了驗證本法的優點，本研究以 404 筆實驗數據，比較逐步迴歸、逐步倒傳遞網路、修剪的演化運算樹三者的模型準確度與複雜度。實證結果顯示，在修剪的演化運算樹可以產生一個簡化但準確的高性能混凝土強度預測公式。

關鍵字：倒傳遞網路、演化運算樹、高性能混凝土、向後式修剪技術、中位數。

Modeling 28 days Strength of High-Performance Concrete Using Genetic Operation Trees with Backward Pruning Techniques

I-Cheng Yeh¹ Chien-Hua Peng² Li-Chuan Lien³

¹ Professor, Chung Hun University

² Research Assistant, Chung Hun University

³ Research Assistant, National Taiwan University of Science and Technology

ABSTRACT

Regression analysis (RA) can build an explicit formula to predict the strength of High-Performance Concrete (HPC); however, the accuracy of the formula is poor. Back-Propagation Networks (BPN) can build rather accurate model to predict the strength of HPC; however, it can not produce an explicit formula. On the other hand, Genetic Operation Trees (GOT) can produce an explicit formula to predict the strength of HPC, and the accuracy of the formula is more accurate than the RA and less accurate than the BPN. Although the GOT can produce an explicit formula; however, their formula are usually too complex to explain the meaning of the formula. Therefore, this study developed a Backward Pruning Technique (BPT) to simplify the complexity of the formula produced by GOT. This technique tries to replace the variables in the leaves of the tree-styled formula with median constant to decrease the complexity of the formula, at the same time, tries to remain the accuracy of the formula. The results showed that, after pruning, the strength formula of HPC became simple enough to explain the meaning in them as well as not to sacrifice their accuracy.

Keywords: back-propagation networks, genetic operation trees, high-performance concrete, backward pruning technique, median constant.

一、前言

高性能混凝土[1]之主要材料組成除了傳統混凝土之水、水泥、粗骨材及細骨材基本構材外，並添加了許多填充膠結料，例如爐石、飛灰，及其它化學摻料，例如強塑劑，除了提升材料主要性質之高強度與工作度外，同時兼顧耐久性、安全性、工作性、經濟性及生態性之需求[2]。傳統建立預測模型的關係式是以統計的方法，例如多變數線性迴歸分析(Linear Regression Analysis, LRA)等，雖可產生預測之公式，唯其準確度較低，因此將其應用在複雜的高性能混凝土非線性系統模型並不可行。

在過去，已有相當多的研究應用倒傳遞網路[3](Back-Propagation Network, BPN)於材料的領域，雖然文獻提出許多建構混凝土強度模型的方法[4-12]，並能準確的預測混凝土之強度行爲，但無法提出合理且讓人理解之材料行爲模型公式。爲解決產生明確公式的問題，許多研究者提出了人工智慧中演化式計算的方法[4-6, 11, 12]。遺傳演算法(Genetic Algorithms, GA)是人工智慧中演化式計算的一門分支，其依據自然演化及適者生存的觀念衍生而成[13, 14]，模擬遺傳的選擇、交配、突變等機制。其強韌性以及平行處理能力，能夠在各種不同的環境裡讓效率與精確率達到平衡[15]，並且適用於解答空間大、複雜、非線性的問題。

從過去文獻[12]的預測準確度及模型可解釋的能力可知，遺傳演算法結合運算樹(Genetic Algorithms of Operation Tree)的演化運算樹(Genetic Operation Tree, GOT)方法確實是一個可以產生自組織公式的方法，其產生之公式準確度低於倒傳遞網路，卻遠比迴歸分析(Regression Analysis, RA)及其它已經發表的文獻[4-6]好。雖然演化運算樹與迴歸分析同樣能產生模型公式，但卻有公式過度複雜的問題，使得雖然有公式但確難以解讀其意義。

爲解決公式過度複雜的問題，本研究提出一個運算樹修剪技術(Pruning Technique,

PT)，即將運算樹的各個節點的「變數」改用樣本的變數中位數(Median)之「常數」替代，以精簡運算樹的結構而簡化公式。本研究將以大量的實驗數據驗證此一方法確實可以產生較簡化、可理解，但不失預測準確度的預測公式。

二、運算樹、遺傳演算法及運算樹的修剪

1. 運算樹

運算樹是一種表達數學公式的樹狀資料結構，圖 1 是一個運算樹模型，其中 X_1 由運算子(+、-、 \times 、 \div 、ln 或 exp 等)構成， $X_2\sim X_7$ 由運算元(變數或常數)或運算子構成， $X_8\sim X_{15}$ 只能由運算元構成[11, 12]。當 $X_1\sim X_{15}$ 設定特定的運算子及運算元後，運算樹可以表達一個特定的數學公式。

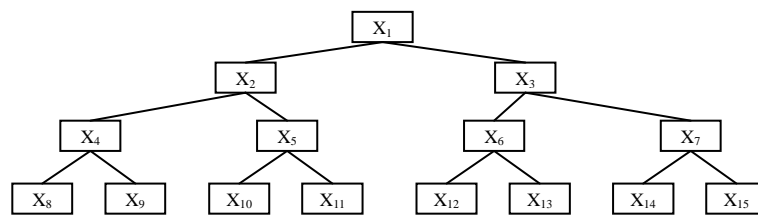


圖 1 運算樹模型示意圖

例如圖 2 是一個運算樹模型實例，其模型共使用了+、-、 \times 、 \div 及 ln 的運算子，而運算

元變數產生了 X_1 , X_2 , 0.5, 5 及 60，而其運算樹公式為 $Y = (X_1 \times 0.5 - \frac{5}{X_2}) + \ln(60)$ 。

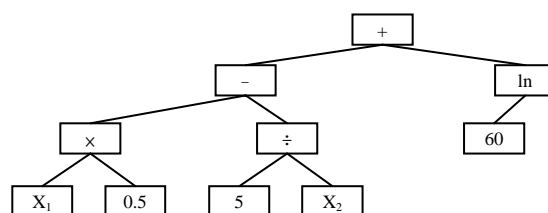


圖 2 運算樹模型實例示意圖

運算樹可用在數據的迴歸問題上，其方法是在運算樹的節點設定適當的運算子及運算元，使此樹狀構對應的數學公式能最配適數據，即能使將數據的自變數代入公式下產生的因變數預測值，與數據的因變數實際值之間的誤差平方和最小化。利用運算樹的結構可以突破傳統迴歸分析的公式結構必須預設，迴歸分析只能作調整迴歸係數的工作的瓶頸。但運算樹是一個離散資料結構，故無法像迴歸分析一樣以微積分中的極值定理導出解迴歸係數的聯立方程式，即用運算樹來迴歸數據是一個離散最佳化問題。因此，本研究採用具有解離散最佳化問題能力的遺傳演算法，來建構最配適數據的運算樹。

2. 遺傳演算法

遺傳演算法是近年來極具發展潛力的最佳化方法。它以獨特的搜尋方式，可跳離局部最佳值，並趨近整體最佳值。此方法最早於1975年由Holland所發展，並出版Adaptation in Natural and Artificial System一書。遺傳演算法的概念源自於達爾文「物競天擇，適者生存」的想法，即仿效自然界生物生存競爭機制，較能適應環境的個體有較高的存活機會，這些個體透過兩兩交配的方式產生後代個體，而後代藉由組合前代個體的基因而擁有近似但不同於前代個體的特性，此外基因突變也產生一些變異。新的一代也在生存競爭的機制下產生下一代，如此一代接一代，個體將朝向最能適應環境的方向演化。而遺傳演算法以「最佳化問題」為「演化問題」；以「解答」為「個體」；「解答的最佳化」過程為「個體的演化」過程，使解答能逼近最佳解。

遺傳演算法的要點如下[12]：

- 染色體的編碼與解碼(encoding and decoding)：在遺傳演算法中，一個解答由許多個染色體構成，將解答以染色體表達稱為編碼；反之，將染色體表達成解答稱為解碼。例如，將一個數學公式以運算樹來表達可視為一種編碼法；反之，將一個運算樹表

達成數學公式可視為一種解碼法。

- 適應度函數(fitness function)與適應度(fitness)：每一個生物個體對於大自然環境都有不同的適應性，適者生存，不適者淘汰。適應函數就如同大自然環境，用來衡量每一個個體(解答)適應演化問題(最佳化問題)的程度。將個體的染色體解碼成一個解答後，代入適應函數即可求得個體的適應度(fitness)。演化的目的即在最大化適應度。
- 選擇(selection)：遺傳演算法的選擇機制是模擬自然界適者生存的現象，適應度越高的個體，存活機率較高；反之，存活機率較低。因此，適應度較高的個體所擁有的後代可能較多，如果某一個個體的適應度高於其他個體，就有機會藉由世代交替，使其染色體逐漸成為族群內的主流。
- 交配(crossover)：將兩個個體(稱之為父個體及母個體)做訊息的互換，以產生兩個新的子個體。
- 突變(mutation)：將一個個體做訊息的變異以產生一個新的子個體。
- 結束規則：遺傳演算法的演算過程為一個迴圈，因此必須有決定演算過程終止之規則。常見的結束規則有四種：
 - (1) 當連續無法找到更佳解答的演化世代數達到預設的次數，例如30個世代。
 - (2) 當演化的世代數達到預設的最大世代數，例如200個世代。
 - (3) 當染色體適應值已滿足預設的目標值。
 - (4) 當族群內各染色體的同質性已達到預設的水準。例如，族群內最好和最差的染色體，其適應值的差異已達到預定的範圍。
- 精英策略：在演化過程中，將適應度最大的個體保留至下一次演化。由於遺傳演算法的選擇機制是適應度越高的個體，存活機率較高；反之，存活機率較低。但並不

保證適應度最高的個體一定能存活。此策略可確保最終世代的最佳個體是整個演化過程所產生的所有個體中的最佳個體。一個完整的遺傳演算法程序可參考圖 3。

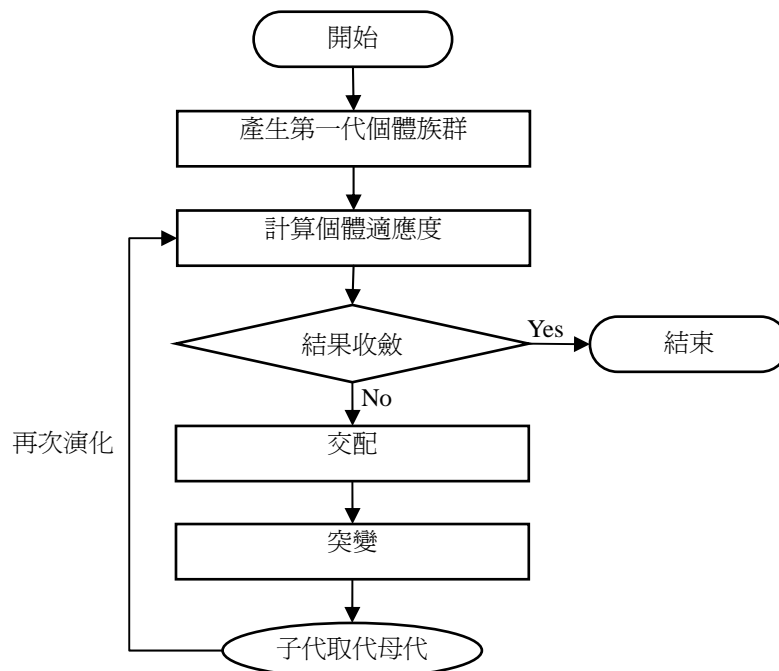


圖 3 遺傳演算法流程圖

3. 運算樹的修剪

遺傳演算法產生的運算樹的末端節點為「變數」者，可考慮改用樣本變數中位數之「常數」來替代，即修剪運算樹，以簡化公式。因此可令末端節點修剪與否為二元狀態：維持「變數」狀態、改用樣本變數中位數之「常數」狀態，即不修剪與修剪二種狀態。本研究提出二種修剪方式：

- (1) 窮盡式修剪法：當運算樹有 N 個「變數」的末端節點時，二元狀態有 2^N 個組合，即運算樹有 2^N 個修剪方式。當 N 不是很大，例如10以下時，可以進行窮盡式的搜尋。
- (2) 後向式修剪法：當 N 很大時，窮盡式的搜尋不可行，此時可用類似逐步迴歸中的後向刪除法，即先嘗試修剪一個最不會降低運算樹預測準確度的末端節點；接著在修剪該末端節點下，再嘗試修剪第二個最會降低運算樹預測準確度的末端節點；如此

下去直到修剪所有「變數」末端節點為止。詳細的步驟如下：

步驟一：分別在修剪一個節點下，以 N-1 個節點建立模式(共有 N 個模式)，選擇其中誤差最小的模式，並命名其修剪的節點為 X1。

步驟二：在修剪 X1 節點下，分別再修剪一個節點下，以 N-2 個節點建立模式(共有 N-1 個模式)，選擇其中誤差最小的模式，並命名其修剪的節點為 X2。

步驟三：在修剪 X1 及 X2 節點下，分別再修剪一個節點下，以 N-3 個節點建立模式(共有 N-2 個模式)，選擇其中誤差最小的模式，並命名其丟棄的節點為 X3。

步驟四：仿照步驟三，不斷減少節點數目，直到減少更多的節點，會明顯劣化預測模型為止。

三、模型建構

1. 實驗數據收集

HPC 強度模型的輸入的變數共有 11 個自變數，分別為：水泥用量(C)、飛灰用量(FL)、爐石用量(SL)、水用量(W)、強塑劑用量(SP)、粗骨材用量(CA)、細骨材用量(FA)、水灰比(W/C)、水膠比(W/B)、水固比(W/S)及骨膠比(TA/B)，因變數為極限抗壓強度(f'_c , psi)。本研究從收集 404 筆實驗數據之 28 天強度，共 404 筆資料，並以隨機的方式取樣 300 筆為訓練範例，104 筆做為驗證的測試範例之高性能混凝土配比資料[8]，以評估所產生的模型是否具備普遍性，其強度資料的值域及型態如表 1 所示。

表 1 強度資料中各變數的值域及型態

變數名稱	值域	單位	型態	
C	71.00~896.00	Kg/m ³	連續	自變數
FL	0.00~200.10	Kg/m ³	連續	
SL	0.00~359.40	Kg/m ³	連續	
W	118.00~314.00	Kg/m ³	連續	
SP	0.00~32.20	Kg/m ³	連續	
CA	595.00~1820.00	Kg/m ³	連續	
FA	486.00~1300.00	Kg/m ³	連續	
W/C	0.24~2.73	比例	連續	
W/B	0.24~0.90	比例	連續	
W/S	0.04~0.13	比例	連續	
TA/B	2.18~9.85	比例	連續	
f_c	1238.00~17691.00	psi	連續	因變數

其中 $W/C=(W+SP)/(C)$ 、 $W/B=(W+SP)/(C+FL+SL)$ 、 $W/S=(W+SP)/(C+FL+SL+CA+FA)$ 、 $TA/B=(CA+FA)/(C+FL+SL)$ 。

2. 運算樹規則及基因編碼方式

本研究為產生自組織公式，採用了運算樹的公式表達方式。圖 4 為一個五層的運算樹示意圖，而運算樹的運算子及運算元的編碼方式如表 2 及表 3 所示。其中：

- 第一層的節點(X_1)限使用運算子，因此值域為 1~6 的整數；
- 第二、三、四層節點($X_2\sim X_{15}$)可使用運算子或運算元，因此值域為 1~18 的整數，其中當編碼為 18 時，代表使用常數 K，而 K 的值域限定為-100~100 的連續數值；
- 第五層的節點($X_{16}\sim X_{31}$)限使用運算元，因此值域為 7~18 的整數。

此外，此樹狀結構遵守下列規則：

- 當節點使用 ln 運算子時，則此運算子只對「左」節點運算。
- 當節點使用運算元(變數或常數)時，則該節點下方的節點將被忽略。

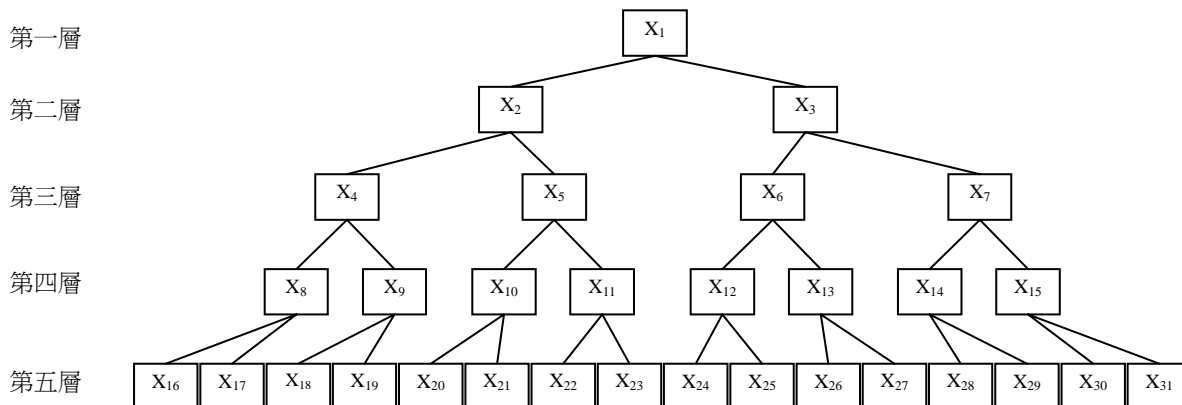


圖 4 五層的運算樹示意圖

表 2 運算子之編碼

編碼	1	2	3	4	5	6
運算子	+	-	×	÷	x ^y	ln

表 3 運算元(變數或常數)之編碼

編碼	7	8	9	10	11	12	13	14	15	16	17	18
運算元	C	FL	SL	W	SP	CA	FA	W/C	W/B	W/S	TA/B	K

3. 運算樹預測值之迴歸修正

運算樹的預測值有時會有偏斜現象，即預測值與實際值間存在平移及旋轉關係。為修正運算樹的偏斜現象，本研究以下列單變數迴歸公式(1)改正之：

$$y = \alpha + \beta \cdot f \dots\dots\dots(1)$$

其中

y = 以迴歸公式改進後之運算樹預測強度值

f = 運算樹預測強度度值

$$\alpha = \bar{y} - \beta \cdot \bar{f} \dots\dots\dots(2)$$

$$\beta = \frac{\sum_{i=1}^n (f_i - \bar{f}) \times (y_i - \bar{y})}{\sum_{i=1}^n (f_i - \bar{f})^2} \dots\dots\dots(3)$$

其中

- \bar{y} = 所有資料的實際強度平均值
- \bar{f} = 所有資料的運算樹預測強度平均值
- y_i = 第 i 筆資料的實際強度值
- f_i = 第 i 筆資料的運算樹預測強度值

4. 適應度函數

本研究以訓練範例的誤差均方根(Root of Mean Square, RMS)最小化作為適應度函數，並以測試範例的誤差均方根評估產生的預測公式是否具有普遍性。RMS 公式如下：

$$RMS = \sqrt{\frac{\sum_{i=1}^n (\hat{y}_i - y_i)^2}{n}} \dots\dots\dots(4)$$

其中 \hat{y}_i 為第 i 筆資料的預測強度值； y_i 為第 i 筆資料的實際強度值； n 為訓練範例或測試範例之總筆數。

5. 演算法參數設定

文獻[16]建議的遺傳演算法參數設定為：(一)交配後產生之個體數 10~200 個；(二)交配率設定為 0.4~0.99；(三)突變率設定為 0.0001~0.1。本研究所採用之遺傳演算法參數設定為：(一)交配後產生之個體數設定為 100 個；(二)交配率設定為 0.9；(三)突變率設定為 0.001。此外，使用精英策略強迫保留目前為止所搜尋到的最佳個體，並設定收斂條件為連續 1000 個世代不再進步則停止程式。

四、結果

1. 修剪前的運算樹

遺傳演算法的解答具有隨機性，因此本研究以不同的亂數種子執行三次，得到如圖 5~7 的演化運算樹，運算樹的相應預測公式如下。

$$y = 434.51 + 64.48 \times \frac{\ln(C \times W / S)}{W / S \times W / B} \dots\dots\dots(5)$$

$$y = 1.94 - 84635.21 \times \frac{\ln(C)}{\left(\frac{TA/B}{\ln(C)} - W - SP \right) \times W / B} \dots\dots\dots(6)$$

$$y = 263.401 - 871.06 \times \frac{\ln(W / C \times W / S)}{W / B} \dots\dots\dots(7)$$

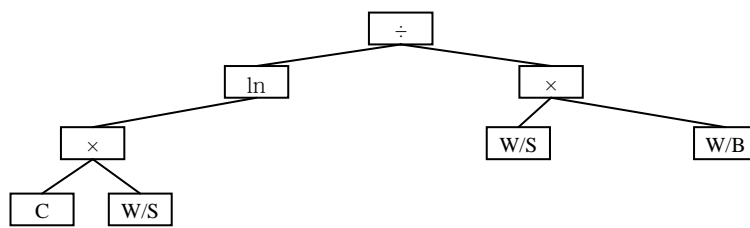


圖 5 第一組 GOT 產生之運算樹

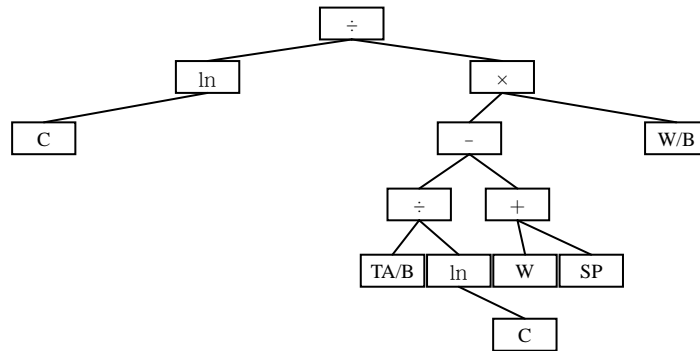


圖 6 第二組 GOT 產生之運算樹

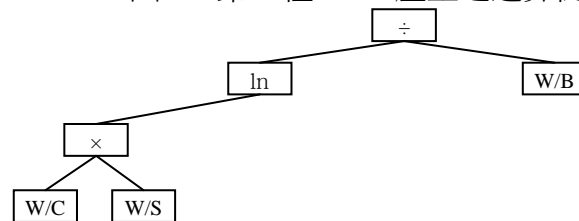


圖 7 第三組 GOT 產生之運算樹

其訓練範例及測試範例之 RMS 如表 4 所示，訓練範例及測試範例之散佈圖如圖 8~圖 13 所示。

表 4 GOT 修剪前的統計數據

組別	訓練 RMS	測試 RMS
1	1320	1561
2	1348	1466
3	1320	1539

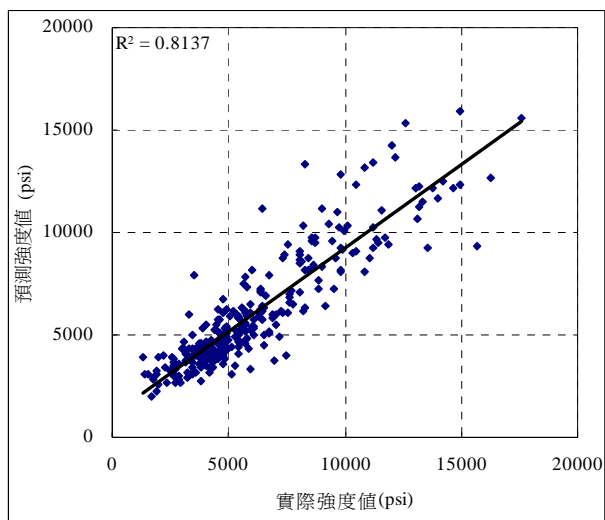


圖 8 第一組 GOT 訓練範例強度散佈圖

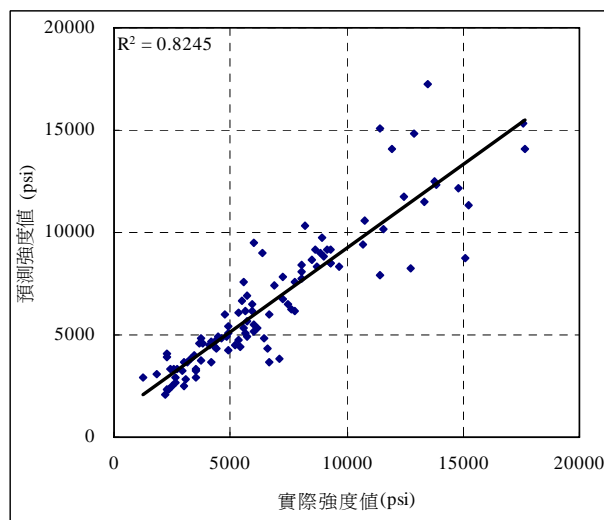


圖 9 第一組 GOT 測試範例強度散佈圖

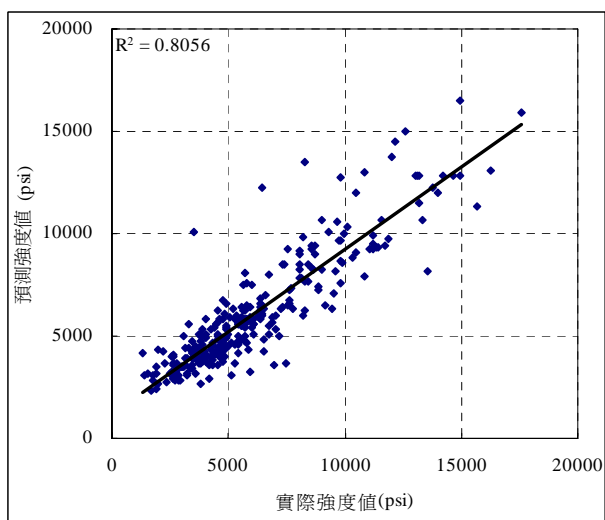


圖 10 第二組 GOT 訓練範例強度散佈圖

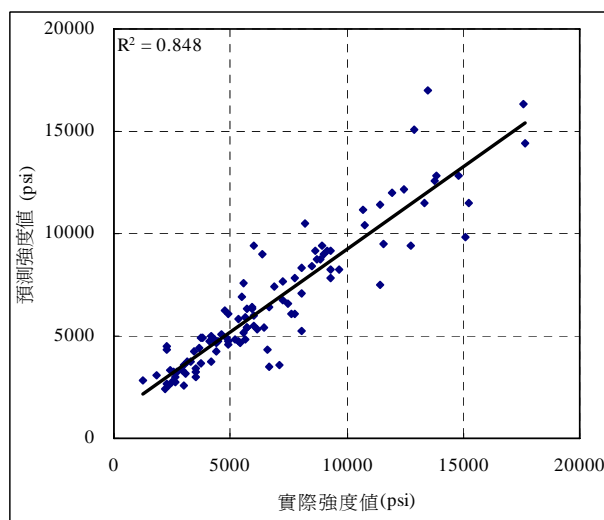


圖 11 第二組 GOT 測試範例強度散佈圖

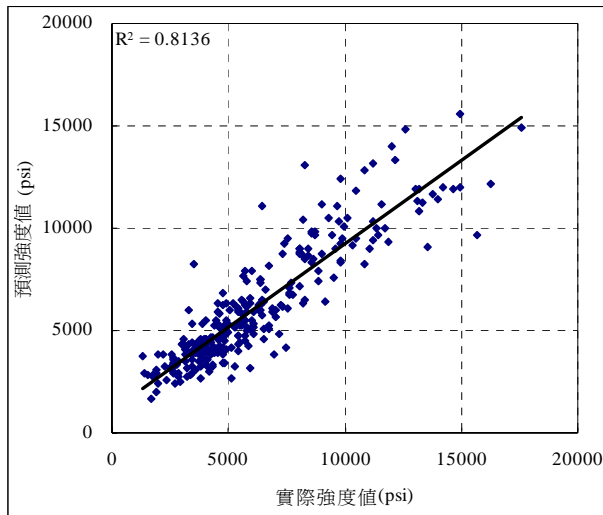


圖 12 第三組 GOT 訓練範例強度散佈圖

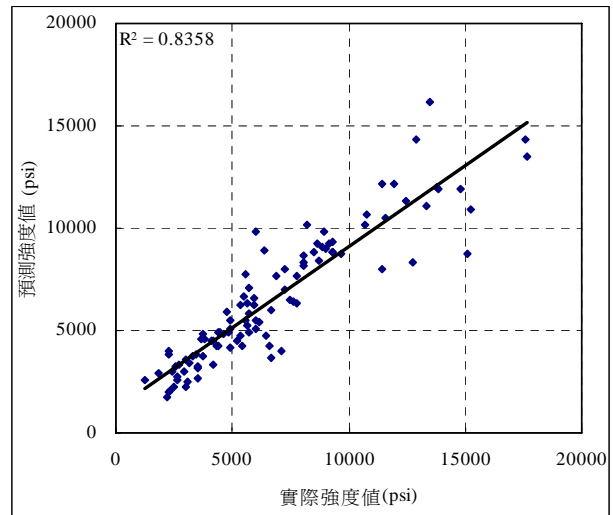


圖 13 第三組 GOT 測試範例強度散佈圖

2. 修剪後的運算樹

各組運算樹採用後向式修剪法的詳細修剪過程及統計數據的變化如圖 14~圖 16 及表 5~7 所示。為了探討節點數的影響，本研究在修剪運算樹的過程中，並不依照「不斷減少節點數目，直到減少更多的節點，會明顯劣化預測模型為止。」的終止條件而停止修剪，而是修剪到只剩一個節點為止。上述三個運算樹經修剪後的最簡公式均相同：

$$y = -3313.95 + \frac{4131.42}{W/B} \dots\dots\dots(8)$$

由上述三組公式可知，採用本研究所提出的向後式修剪法進行演化運算樹的修剪，其最後存留的公式型態皆相同(僅常數合併因小數點取位的不同而有些誤差)，顯示演化運算樹所產生的公式具備穩定性，且皆認為預測 28 天高性能混凝土強度的最重要的變數為水膠比(W/B)，並認為水膠比與強度成反比，這樣的結果與實務上符合。

此外，表 5 步驟 3、表 6 步驟 5、表 6 步驟 5 等三個公式如下：

$$y = -865.83 + 942.84 \times \frac{\ln(C) - 2.43}{W/B} \dots\dots\dots(9)$$

$$y = -1888.57 + 618.84 \times \frac{\ln(C)}{W/B} \dots\dots\dots(10)$$

$$y = -272.60 - 965.86 \times \frac{\ln(W/C) - 2.43}{W/B} \dots\dots\dots(11)$$

公式(9)與(10)顯示，在水膠比相同下，水泥用量越大強度越高。在公式(11)中，因水灰比常小於 1.0，故 $\ln(W/C)$ 幾乎不可能大於 2.43，故 $\ln(W/C)-2.43$ 幾乎可以肯定小於 0，因此(11)式與(9)式、(10)式一樣，都顯示強度與水膠比成反比。將(11)式改寫成

$$y = -272.60 + (-965.86 \cdot \ln(W/C) + 2347) \times \frac{1}{W/B} \dots\dots\dots(11)$$

可知水灰比越小，因其 $\ln(W/C)$ 為負值，故強度越大。

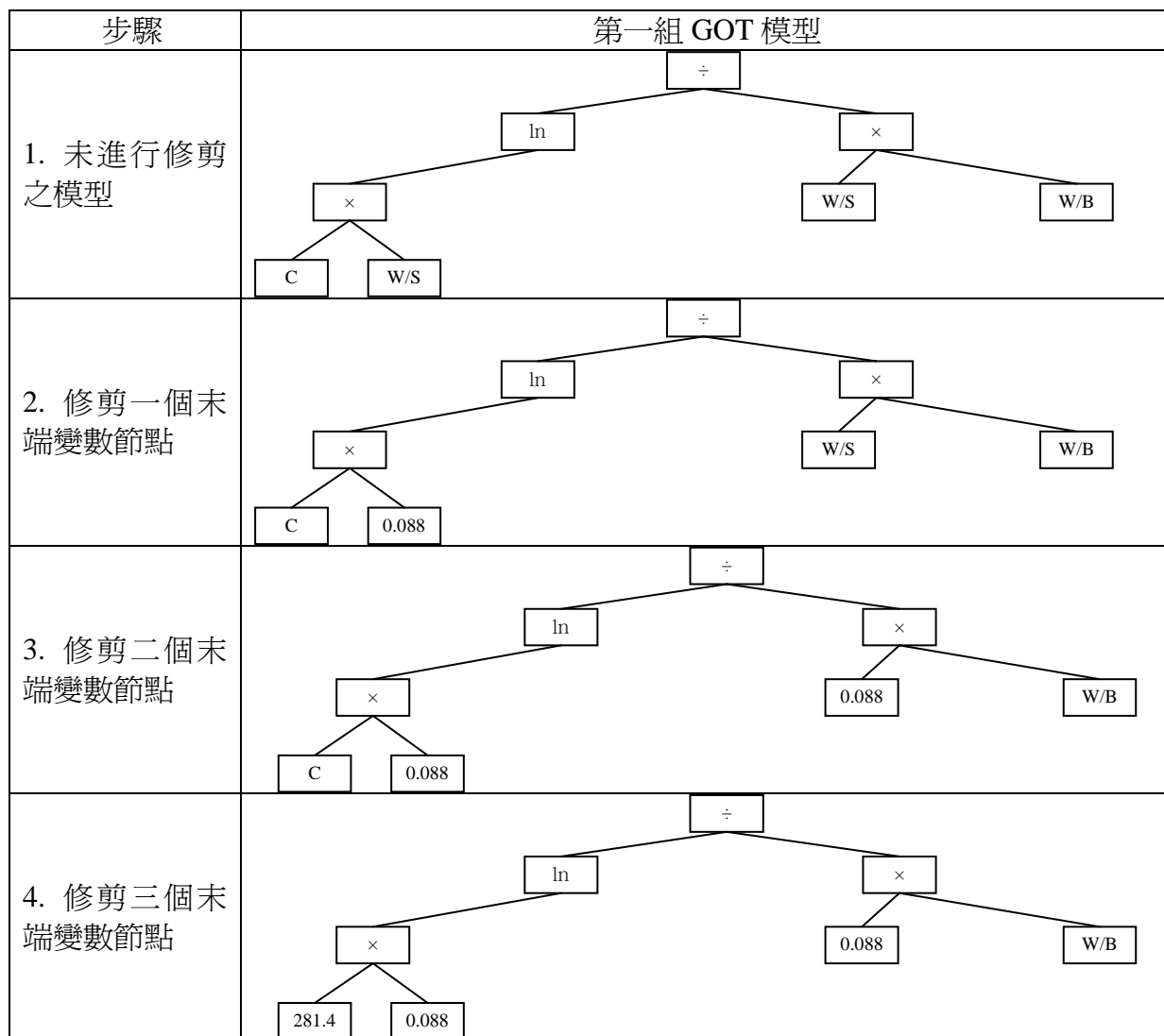


圖 14 第一組 GOT 的修剪過程

表 5 第一組 GOT 修剪過程的統計數據

步驟	公式	訓練 R^2	測試 R^2	訓練 RMS	測試 RMS
1	$y = 434.51 + 64.48 \times \frac{\ln(C \times W/S)}{W/S \times W/B}$	0.81	0.82	1320	1567
2	$y = 958.13 + 57.19 \times \frac{\ln(C) - 2.43}{W/S \times W/B}$	0.81	0.81	1345	1623
3	$y = -865.83 + 942.84 \times \frac{\ln(C) - 2.43}{W/B}$	0.79	0.77	1412	1848
4	$y = -3313.95 + \frac{4131.42}{W/B}$	0.76	0.76	1482	1880

步驟	公式二
1. 未進行修剪之模型	
2. 修剪一個末端變數節點	
3. 修剪二個末端變數節點	
4. 修剪三個末端變數節點	
5. 修剪四個末端變數節點	

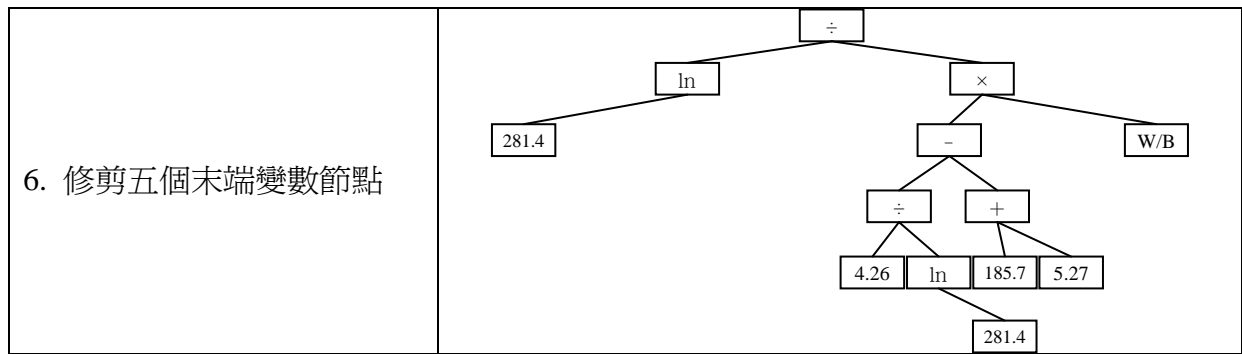


圖 15 第二組 GOT 的修剪過程(續)

表 6 第二組 GOT 修剪過程的統計數據

步驟	公式	訓練 R^2	測試 R^2	訓練 RMS	測試 RMS
1	$y = 1.94 - 84635.21 \times \frac{\ln(C)}{\left(\frac{TA}{B} / \ln(C) - W - SP\right) \times W / B}$	0.81	0.85	1348	1466
2	$y = 11.28 - 84494.88 \times \frac{\ln(C)}{\left(\frac{4.26}{\ln(C)} - W - SP\right) \times W / B}$	0.81	0.85	1347	1467
3	$y = 14.83 - 84440.97 \times \frac{\ln(C)}{(0.755 - W - SP) \times W / B}$	0.81	0.85	1347	1467
4	$y = -1963.75 - 118962.63 \times \frac{\ln(C)}{(0.755 - 185.7 - SP) \times W / B}$	0.80	0.81	1363	1706
5	$y = -1888.57 + 618.84 \times \frac{\ln(C)}{W / B}$	0.79	0.78	1389	1815
6	$y = -3313.95 + \frac{4135.76}{W / B}$	0.76	0.76	1483	1880

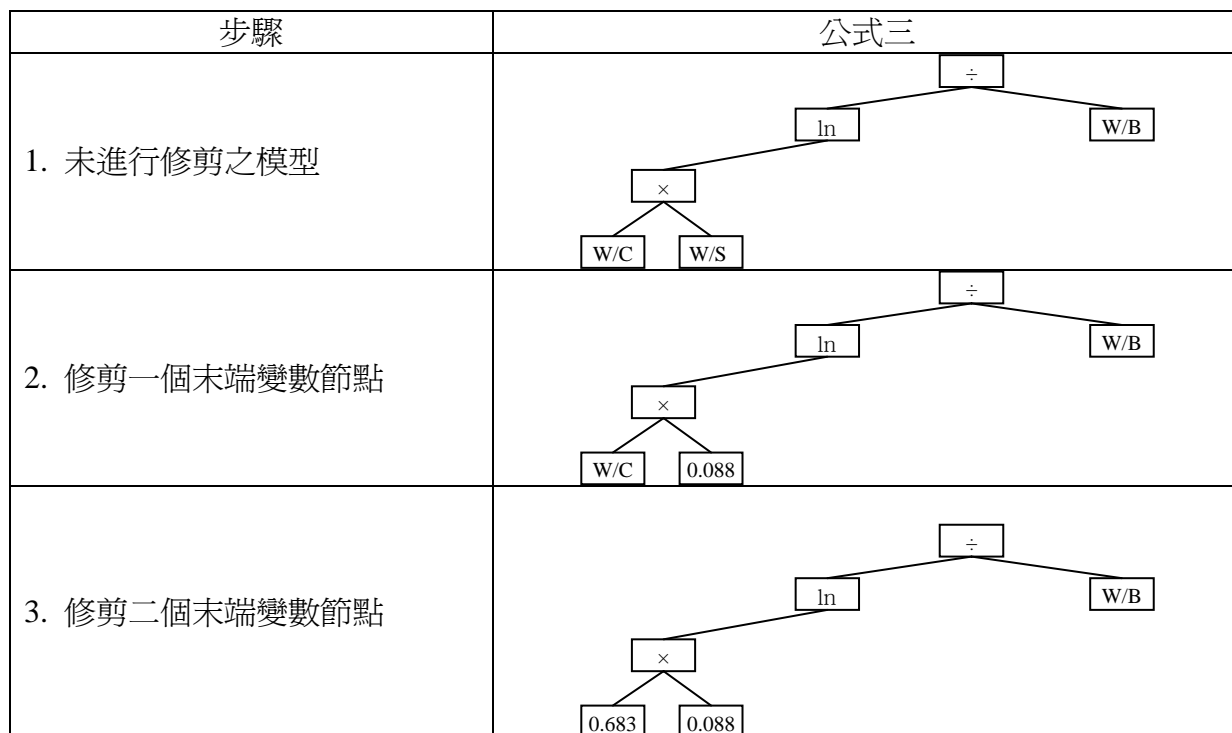


圖 16 第三組 GOT 的修剪過程

表 7 第三組 GOT 修剪過程的統計數據

步驟	公式	訓練 R^2	測試 R^2	訓練 RMS	測試 RMS
1	$y = 263.40 - 871.06 \times \frac{\ln(W/C \times W/S)}{W/B}$	0.81	0.84	1320	1538
2	$y = -272.60 - 965.86 \times \frac{\ln(W/C) - 2.43}{W/B}$	0.81	0.81	1339	1663
3	$y = -3313.95 + \frac{4134.29}{W/B}$	0.76	0.76	1483	1880

3. 迴歸分析

爲了與 GOT 作比較，本研究採用後向刪除法的逐步迴歸分析，其步驟如下：

步驟一：用所有 N 個自變數建立一迴歸分析模式。分別在丟棄一個自變數下，以 N-1

個自變數建立迴歸分析模式，選擇其中誤差最小的模式，並命名其丟棄的自變數爲 X1。

步驟二：在丟棄 X1 下，分別再丟棄一個自變數下，以 N-2 個自變數建立迴歸分析模式，

選擇其中誤差最小的模式，並命名其丟棄的自變數爲 X2。

步驟三：在丟棄 X1 及 X2 下，分別再丟棄一個自變數下，以 N-3 個自變數建立自變數

模式，選擇其中誤差最小的模式，並命名其丟棄的自變數爲 X3。

步驟四：仿照步驟三，不斷減少自變數的數目，直到減少更多的自變數，會明顯劣化預

測模型爲止。

以逐步迴歸之後向刪除法建立的迴歸公式及其訓練範例及測試範例之 RMS 如表 8 所示。爲了探討變數數目的影響，本研究在逐步迴歸分析的建構過程中，並不依照上述終止條件而停止建模，而是繼續丟棄自變數到只剩一個爲止。由表 8 的結果發現：

1. 模型的訓練範例評估：在刪減變數過程中，訓練範例的準確度在刪除前六個變數，而只剩水泥用量(C)、飛灰用量(FL)、爐石用量(SL)、水用量(W)等四個變數時，只有微幅變低。這四個變數正好是水膠比(W/B)的組成，顯示若變數的組成不能反應水膠比，則模型無法準確預測強度，此一結果與演化運算樹相同。接著依序刪除的是水飛灰用量(FL)、爐石用量(SL)、水用量(W)，最後只剩水泥用量(C)，在此過程模型的準確度迅速降低，這顯示此四個變數對強度最爲重要，且從其順序可知，這三種膠結料對強度的貢獻勿由小而大依序是飛灰用量(FL)、爐石用量(SL)、水泥用量(C)。

2. 模型的測試範例評估：在刪減變數過程中，測試範例的準確度在刪前六個變數時，反而有遞增的趨勢，這顯示使用 11 個變數的模型有過度配適的現象，因此簡化的模型其「普遍化」預測能力反而獲得提升。

表 8 逐步迴歸之後向刪除法

步驟	公式	訓練 R^2	測試 R^2	訓練 RMS	測試 RMS	丟棄 變數
1	$f'_c = -36390.90 + 41.46C + 35.00FL + 36.46SL - 341.27W - 378.06SP$ $+ 23.99CA + 24.39FA + 0AGE - 340.44W / C + 7539.82W / B$ $+ 561452.70W / S - 993.69TA / B$	0.82	0.68	1313	2241	None
2	$f'_c = -35514.50 + 41.21C + 33.71FL + 35.21SL - 335.87W - 370.94SP$ $+ 23.39CA + 23.70FA + 5889.38W / B + 555362.9W / S$ $- 880.69TA / B$	0.82	0.69	1314	2210	W/C
3	$f'_c = -36644.20 + 41.09C + 33.20FL + 34.99SL - 333.89W - 369.50SP$ $+ 22.84CA + 23.31FA + 579143.2W / S - 341.44TA / B$	0.82	0.68	1315	2233	W/B
4	$f'_c = 11417.09 + 21.45C + 12.28FL + 14.69SL - 113.23W - 143.78SP$ $+ 0.99CA + 91990.31W / S - 214.787TA / B$	0.80	0.81	1374	1682	FA
5	$f'_c = 13127.86 + 20.44C + 11.37FL + 13.59SL - 104.64W - 140.13SP$ $+ 72477.77W / S - 264.72TA / B$	0.80	0.81	1377	1711	CA
6	$f'_c = 10648.46 + 23.86C + 14.96FL + 17.14SL - 114.35W - 148.43SP$ $+ 91558.32W / S$	0.80	0.81	1380	1710	TA/B
7	$f'_c = 10386.55 + 21.79C + 14.41FL + 16.38SL - 66.52W - 101.53SP$	0.78	0.81	1420	1680	W/S
8	$f'_c = 9274.43 + 20.78C + 8.82FL + 14.32SL - 59.42W$	0.77	0.76	1483	1873	SP
9	$f'_c = 10871.31 + 18.50C + 11.75SL - 61.11W$	0.74	0.77	1559	1833	FL
10	$f'_c = 12238.92 + 14.83C - 57.89W$	0.65	0.73	1809	2005	SL
11	$f'_c = 1254.752 + 16.30C$	0.45	0.47	2258	2819	W

4. 神經網路

本研究以 11-11-1(11 輸入、11 節點、1 輸出)的倒傳遞網路來做為準確度的比較，其中網路的設定值如下所示：

- 初始權值範圍= ±0.3
- 學習速率下限值=0.1
- 學習速率=1.0
- 慣性因數初始值=0.5
- 學習速率衰減率=0.99
- 慣性因數衰減率=0.99

- 慣性因數下限值=0.1

- 學習循環=1000

爲了與 GOT 及逐步迴歸作比較，本研究採用類似逐步迴歸中的後向刪除法，提出「逐步神經網路」其步驟如下：

步驟一：用所有 N 個自變數建立一類神經網路模式。分別在丟棄一個自變數下，以 N-1 個自變數建立類神經網路模式，選擇其中誤差最小的模式，並命名其丟棄的自變數爲 X1。

步驟二：在丟棄 X1 下，分別再丟棄一個自變數下，以 N-2 個自變數建立類神經網路模式，選擇其中誤差最小的模式，並命名其丟棄的自變數爲 X2。

步驟三：在丟棄 X1 及 X2 下，分別再丟棄一個自變數下，以 N-3 個自變數建立類神經網路模式，選擇其中誤差最小的模式，並命名其丟棄的自變數爲 X3。

步驟四：仿照步驟三，不斷減少自變數的數目，直到減少更多的自變數，會明顯劣化預測模型爲止。

以逐步神經網路建立的強度模型其訓練範例及測試範例之 RMS 如表 9 所示。爲了探討變數數目的影響，本研究在逐步神經網路的建構過程中，並不依照上述終止條件而停止建模，而是繼續丟棄自變數到只剩一個爲止。

由表 9 的結果可知，逐步神經網路刪減變數的順序與準確度變化的歷程與逐步迴歸分析十分接近，即最後三個被刪除的變數依序仍是水飛灰用量(FL)、爐石用量(SL)、水用量(W)，最後只剩水泥用量(C)；在刪減變數過程中，測試範例的準確度在刪前六個變數時，反而有遞增的趨勢，這顯示使用 11 個變數的模型有過度配適的現象，因此簡化的模型其「普遍化」預測能力反而獲得提升。

表 9 逐步神經網路

步驟	訓練 R^2	測試 R^2	訓練 RMS	測試 RMS	丟棄變數
1	0.85	0.88	1173	1280	None
2	0.87	0.83	1097	1554	W/C
3	0.87	0.87	1096	1378	W/B
4	0.87	0.87	1105	1380	CA
5	0.87	0.88	1117	1307	FA
6	0.84	0.87	1133	1329	TA/B
7	0.85	0.87	1194	1352	W/S
8	0.81	0.79	1331	1704	SP
9	0.77	0.82	1458	1617	FL
10	0.69	0.76	1713	1851	SL
11	0.5	0.5	2176	2654	W

5. 各方法的比較

表 10 及表 11 與圖 17 至圖 20 為上述方法的在模型準確度及複雜度的比較。模型的準確度指的是對於問題的預測能力，亦即 RMS 越小則表示準確度越優；模型的複雜度由變數數目或節點數目來表示，數目越大模型越複雜，其可理解性越低。由表及圖可知：

- 複雜度：在節點或變數數目大於五以上，準確度由高而低是逐步神經網路、GOT、逐步迴歸；但在節點或變數數目小於三以下時，逐步神經網路、逐步迴歸的準確度開始劇降，而GOT的準確度並沒有發生劇降的情況，準確度由高而低變成GOT、逐步神經網路、逐步迴歸。
- 複雜度：上述三組GOT產生的運算樹模型所使用的變數為3~6個，其訓練範例 R^2 平均為0.81及RMS平均為1329。而逐步迴歸需要使用9個變數、神經網路需要使用4個變數以上方能達到相同的準確度。顯示在相同的預測準確度下，模型複雜度由低而高是逐步神經網路、GOT、逐步迴歸。

因此，本研究提出的修剪法是一個可以簡化 GOT 產生的 28 天高性能混凝土強度公式，但簡化過程不會導致準確度劇降的方法。

表 10 各種方法的 R^2 比較

變數或節點數	第一組 GOT		第二組 GOT		第三組 GOT		逐步迴歸		逐步神經網路	
	訓練範例	測試範例	訓練範例	測試範例	訓練範例	測試範例	訓練範例	測試範例	訓練範例	測試範例
11	NA	NA	NA	NA	NA	NA	0.82	0.68	0.85	0.88
10	NA	NA	NA	NA	NA	NA	0.82	0.69	0.87	0.83
9	NA	NA	NA	NA	NA	NA	0.82	0.68	0.87	0.87
8	NA	NA	NA	NA	NA	NA	0.80	0.81	0.87	0.87
7	NA	NA	NA	NA	NA	NA	0.80	0.81	0.87	0.88
6	NA	NA	0.81	0.85	NA	NA	0.80	0.81	0.84	0.87
5	NA	NA	0.81	0.85	NA	NA	0.78	0.81	0.85	0.87
4	0.81	0.82	0.81	0.85	NA	NA	0.77	0.76	0.81	0.79
3	0.81	0.81	0.80	0.81	0.81	0.84	0.74	0.77	0.77	0.82
2	0.79	0.77	0.79	0.78	0.81	0.81	0.65	0.73	0.69	0.76
1	0.76	0.76	0.76	0.76	0.76	0.76	0.45	0.47	0.5	0.5

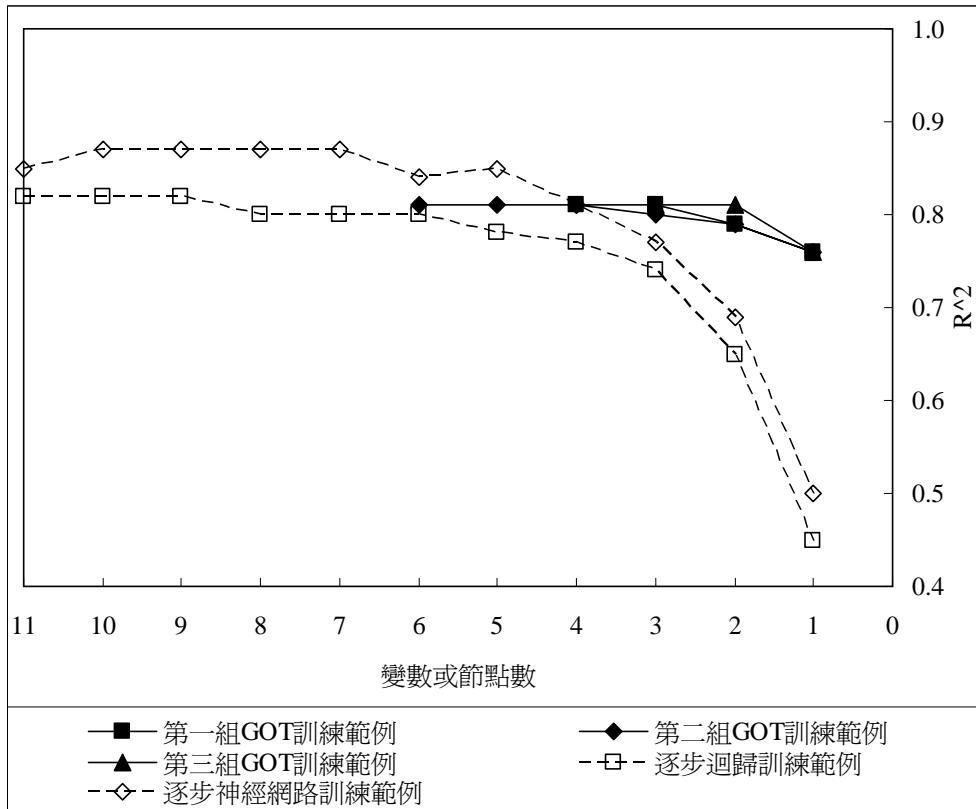


圖 17 各種方法的訓練範例 R^2 比較折線圖

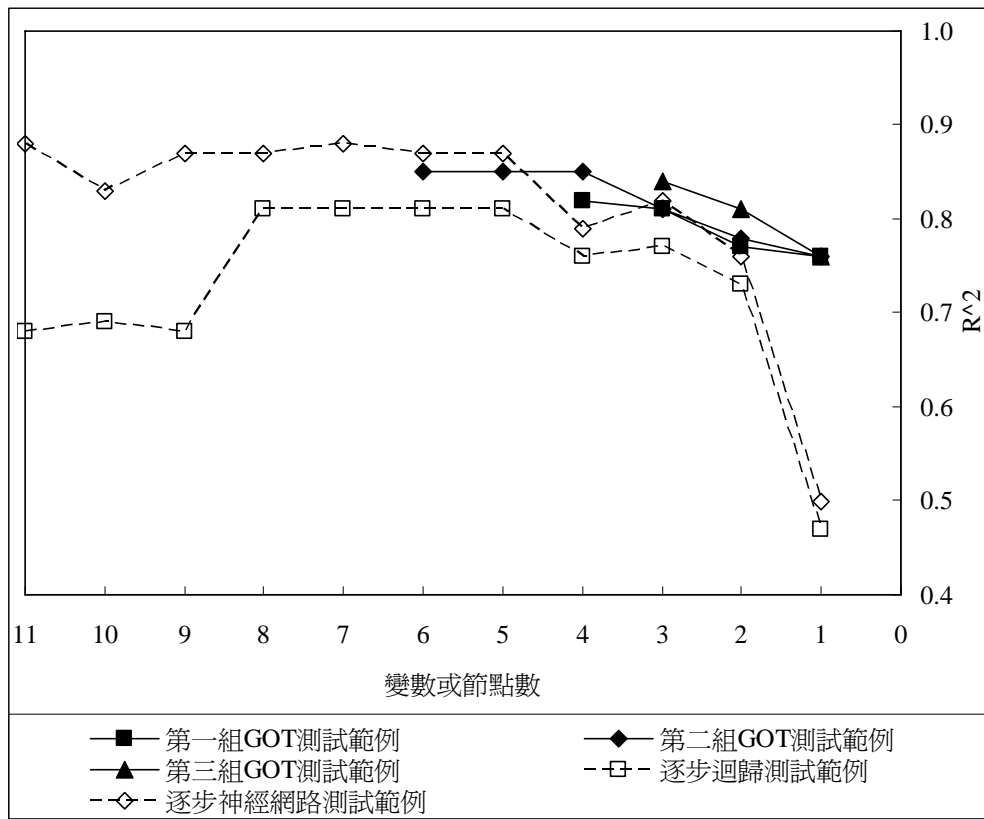


圖 18 各種方法的測試範例 R^2 比較折線圖

表 11 各種方法的 RMS 比較

變數或節點數	第一組 GOT		第二組 GOT		第三組 GOT		逐步迴歸		逐步神經網路	
	訓練範例	測試範例	訓練範例	測試範例	訓練範例	測試範例	訓練範例	測試範例	訓練範例	測試範例
11	NA	NA	NA	NA	NA	NA	1313	2241	1173	1280
10	NA	NA	NA	NA	NA	NA	1314	2210	1097	1554
9	NA	NA	NA	NA	NA	NA	1315	2233	1096	1378
8	NA	NA	NA	NA	NA	NA	1374	1682	1105	1380
7	NA	NA	NA	NA	NA	NA	1377	1711	1117	1307
6	NA	NA	1348	1466	NA	NA	1380	1710	1133	1329
5	NA	NA	1347	1467	NA	NA	1420	1680	1194	1352
4	1320	1567	1347	1467	NA	NA	1483	1873	1331	1704
3	1345	1623	1363	1706	1320	1538	1559	1833	1458	1617
2	1412	1848	1389	1815	1339	1663	1809	2005	1713	1851
1	1482	1880	1483	1880	1483	1880	2258	2819	2176	2654

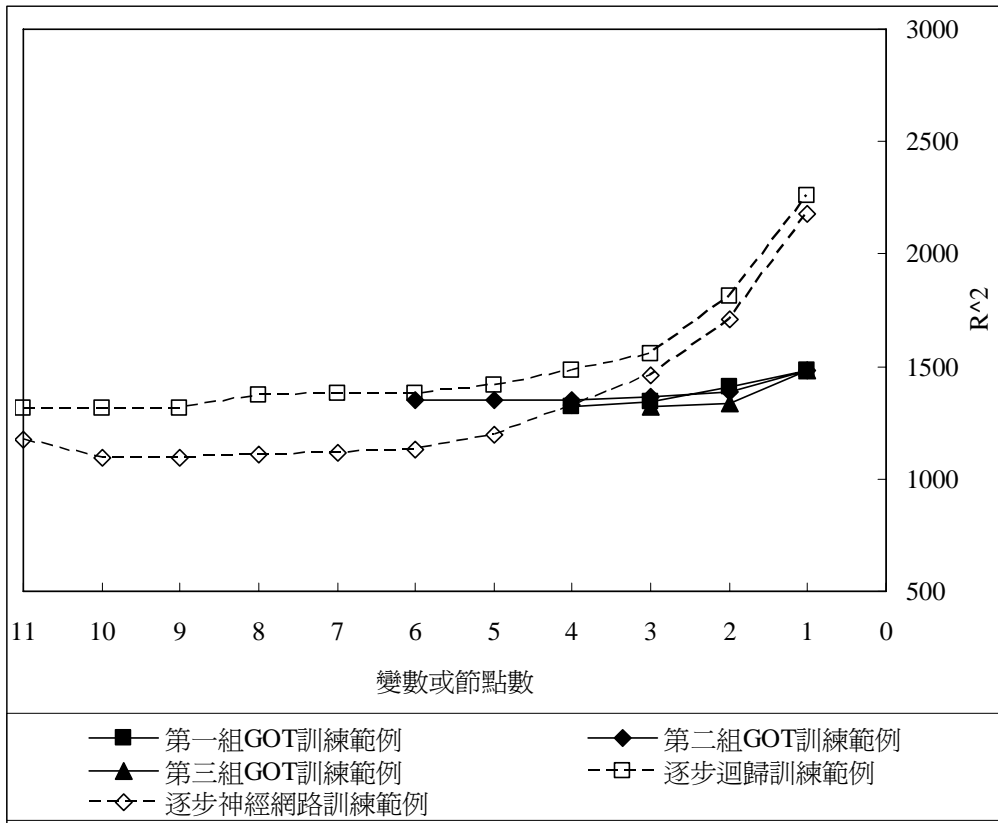


圖 19 各種方法的訓練範例 RMS 比較折線圖

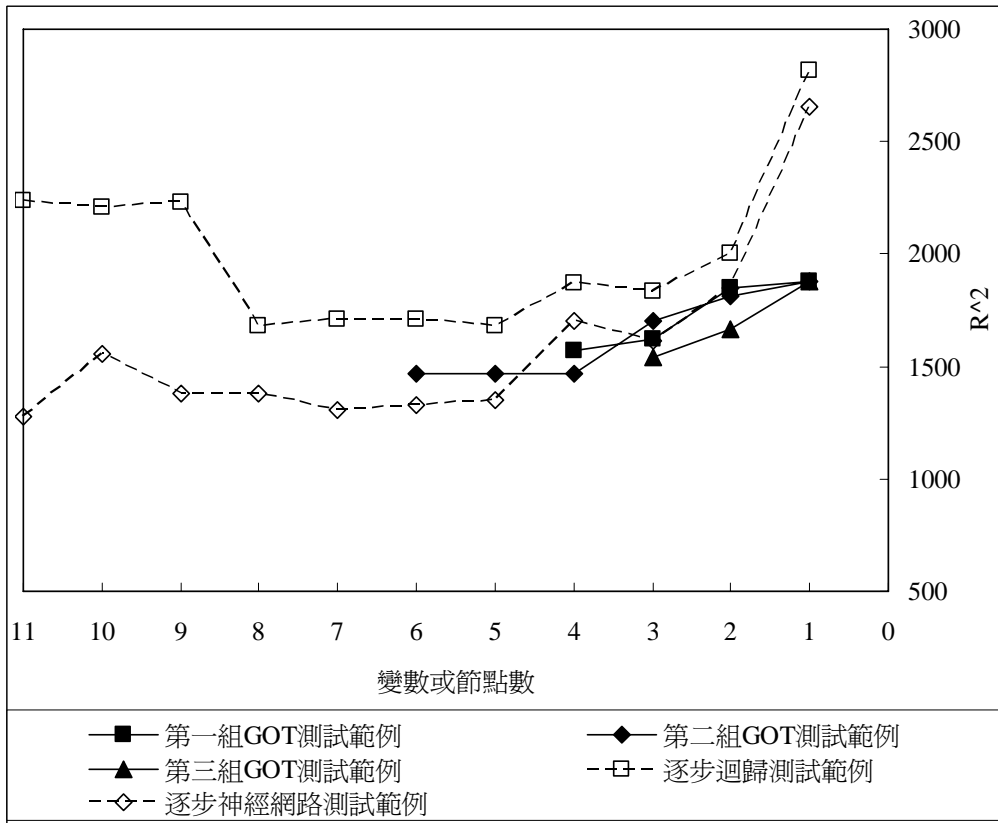


圖 20 各種方法的測試範例 RMS 比較折線圖

五、結論及建議

本文研究結論歸納如下：

1. 採用本研究所提出的向後式修剪法進行演化運算樹的修剪，其最後存留的公式型態皆相同，顯示其產生的公式具備穩定性，且皆認為預測 28 天高性能混凝土強度的最重要的變數為水膠比(W/B)，並認為水膠比與強度成反比，這樣的結果與實務上符合。
2. 逐步神經網路刪減變數的順序與準確度變化的歷程與逐步迴歸分析十分接近，即最後三個被刪除的變數依序都是水飛灰用量(FL)、爐石用量(SL)、水用量(W)，最後只剩水泥用量(C)；在刪減變數過程中，測試範例的準確度在刪前六個變數時，反而有遞增的趨勢，這顯示使用 11 個變數的模型有過度配適的現象，因此簡化的模型其「普遍化」預測能力反而獲得提升。
3. 在模型準確度的比較方面，在節點或變數數目大於五以上，準確度由高而低依序是逐步神經網路、GOT、逐步迴歸；但在節點或變數數目小於三以下時，逐步神經網路、逐步迴歸的準確度開始劇降，而 GOT 的準確度並沒有發生劇降的情況，準確度由高而低變成 GOT、逐步神經網路、逐步迴歸。
4. 在模型複雜度的比較方面，GOT 產生的運算樹模型所使用的變數為 3~6 個，其訓練範例 R^2 平均為 0.81 及 RMS 平均為 1329。而逐步迴歸需要使用 9 個變數、神經網路需要使用 4 個變數以上方能達到相同的準確度。顯示在相同的預測準確度下，模型複雜度由小而大依序是逐步神經網路、GOT、逐步迴歸。

綜合以上四點，可以證明本研究提出的修剪法是一個可以產生簡化，但能保持高準確度的 HPC 強度公式的方法。

未來研究方向包括：

- (1) 不同齡期的強度模型：由於本研究僅針對 28 天齡期之高性能混凝土進行強度模型公式的研究及修剪，未來可將本研究延伸至齡期小於 28 天或大於 28 天的情況。
- (2) 不同水膠比分類的強度模型：傳統 ACI 的經驗公式是以水膠比為基礎來進行高性能混凝土的強度計算，未來可延伸本研究至不同水膠比的混凝土，例如將樣本分成高、中、低水膠比，分別產生強度預測公式，以探討不同水膠比下強度預測公式的變化。

參考文獻

1. 黃兆龍，混凝土性質與行為，詹氏書局 (1999)。
2. 葉基棟、彭耀南、黃兆龍、高健章、林建弘等，『高強度混凝土設計與施工準則初步研究』，中國土木水利工程學會，建研所專題研究成果報告 A09~22，第 8~28 頁 (2000)。
3. 葉怡成，應用類神經網路，儒林圖書有限公司 (2004)
4. Chen, L., "A study of applying macroevolutionary genetic programming to concrete strength estimation," ASCE, Journal of Computing in Civil Engineering, Vol.17, No.4, pp.290-294 (2003).
5. 葉怡成，免疫演算法於高性能混凝土配比設計多目標最佳化之研究(1/2)，行政院國家科學委員會專題研究計畫成果報告，NSC 92-2211-E-216-015 (2004)。
6. Chen, L., Tasi, C. S., and Chen, H. M., "A study of applying grammar evolution to concrete strength estimation," Chung Hua Journal of Science and Engineering Vol.2, No.2, pp.55-62 (2004).
7. 賴鴻成，「倒傳遞網路在混凝土材料強度之應用」，碩士論文，交通大學，新竹市 (1993)。
8. 彭鈞哲，「以倒傳遞網路在高性能混凝土抗壓強度之應用」，碩士論文，中華大學，

新竹市 (1995)。

9. Hamid-Zadeh, N., Jamali, A., Nariman-Zadeh, N., and Akbarzadeh, H., "Prediction of concrete compressive strength using evolved polynomial neural networks," WSEAS Transactions on Systems Vol.6, No.4, pp. 802-807 (2007).
10. Ahmet O., Murat P., Erdogan O., Erdogan K., Naci C., and Bhatti M. A., "Predicting the compressive strength and slump of high strength concrete using neural network," Construction and Building Materials Vol.20, pp.769-775 (2006).
11. 連立川、葉怡成、張皓博、謝明勳，「以遺傳演算法及運算樹作高性能混凝土工作度建模」，2005 營建技術暨管理研討會，斗六市 (2005)。
12. 連立川、葉怡成、鄭明淵，「以遺傳演算法及運算樹作高性能混凝土強度建模」，技術學刊，第二十一卷，第一期，第 41-54 頁 (2006)。
13. Davis, L., Handbook of Genetic Algorithms, Van Nostrand Reinhold, NY. (1991).
14. Holland, J. H., Adaptation in Natural and Artificial System, University of Michigan Press, Ann Arbor. (1975).
15. Goldberg, D. E., Genetic Algorithms in Search Optimization and Machine Learning, Addison-Wesley Publishing Company. (1989).
16. 雷英杰、張善文、李續武、周創明，MATLAB 遺傳算法工具箱及應用，西安電子科技大學出版社 (2005)。

研究成果自評

本計畫是一個三年期計畫「以基因演算法建立高性能混凝土材料行為模式」(96-2221-E-216-032-)的第二年。申請人過去曾在 89~91, 92~93, 94~95 學年, 分別完成一個三年期與二個二年期等三個應用人工智慧於混凝土科技的國科會專題研究計畫。其研究成果一共發表了 18 篇期刊論文。包括下列 9 篇國外 SCI/EI 期刊論文：

1. I-Cheng Yeh, "Design of High Performance Concrete Mixture Using Neural Networks," *J. of Computing in Civil Engineering*, ASCE, Vol. 13, No. 1, pp. 36-42 (1999). **(SCI, EI)**
2. I-Cheng Yeh, "Modeling of strength of high performance concrete using artificial neural networks," *Cement and Concrete Research*, Vol. 28, No. 12, pp. 1797-1808 (1998). **(EI)**
3. I-Cheng Yeh, "Modeling Concrete Strength with Augment-Neuron Networks," *J. of Materials in Civil Engineering*, ASCE, Vol. 10, No. 4, pp. 263-268 (1998). **(EI)**
4. Yeh, I-Cheng (2006), "Exploring concrete slump model using artificial neural networks," *J. of Computing in Civil Engineering*, ASCE, Vol.20, No.3, pp.217-221. **(EI, SCI) (SCI 2007 Impact Factor 0.975)**
5. Yeh, I-Cheng (2006), "Analysis of strength of concrete using design of experiments and neural networks," *Journal of Materials in Civil Engineering*, ASCE, Vol.18, No.4, pp.597-604. **(SCI, EI) (SCI 2007 Impact Factor 0.380)**
6. Yeh, I-Cheng (2006), "Generalization of strength versus water-cementitious ratio relationship to age," *Cement and Concrete Research*, Vol.36, No.10, 1865-1873. **(SCI, EI) (SCI 2007 Impact Factor 1.185)**
7. Yeh, I-Cheng (2006), "Computer-aided design for optimum concrete mixture," *Cement and Concrete Composites*, Vol.29, No.3, 193-202 (2007). **(SCI, EI) (SCI 2007 Impact Factor 0.791)**
8. Yeh, I-Cheng, "Modeling slump flow of concrete using second-order regressions and artificial neural networks," *Cement and Concrete Composites*, Vol.29, 474-480 (2007). **(SCI, EI) (SCI 2007 Impact Factor 0.791)**
9. Yeh, I-Cheng, "Prediction of workability of concrete using design of experiments for mixtures," *COMPUTERS AND CONCRETE* (已接受)(2007) **(SCI, EI)**