

## 應用基因模糊邏輯控制建構事件自動偵測系統

邱裕鈞<sup>1</sup> 艾嘉銘<sup>2</sup> 陳怡霖<sup>3</sup>

### 摘要

高速公路在國內南北城際運輸中扮演相當重要的角色，一旦發生行車擁塞或交通事故，即可能造成嚴重的車流延滯，甚至發生追撞事故。因此，有必要建立一套有效的事件自動偵測系統，以便精確、快速地提供事件警報資訊給相關管理單位，俾便進行及時之處理。此外，為避免以往依據事件模擬資料所建構之事件自動偵測模式在實際應用時所可能發生之落差，本文乃以實際事件資料及偵測器測得交通資料為基礎，利用 GFLC 建構一高速公路之事件自動偵測模式。本模式係以上下游偵測器所測得之流量、速度和佔有率之差距作為狀態變數，控制變數則為是否有事件發生。以實際高速公路事件資料加以驗證後，發現本模式雖可達 100% 之偵測率及 53.33 秒之平均偵知時間，但誤報率卻高達 36.36%，偵測績效不甚良好。究其原因可能有三：狀態變數之選擇不當、事件登錄資料錯誤，以及偵測資料有誤等。未來應蒐集更精確之事件資料及嘗試不同狀態變數組合，以利事件自動偵測模式之構建。

### 壹、前言

高速公路在國內南北城際運輸中扮演相當重要的角色，一旦發生行車擁塞或交通事故，即可能造成嚴重的車流延滯，甚至發生追撞事故。因此，建立一套有效的自動事件偵測系統，能精確且快速地提供事件警報資訊給相關管理單位，以便進行及時之處置，實為一重要課題。為提升高速公路網運轉效率，節省運輸能源消耗並減少整體社會成本的支出，有必要發展一套迅速且精確的事件偵測系統以求於第一時間即加以通報與處理。基此，已有相當多之事件偵測的演算法發表，例如 Ai-deek 等人[1]提出之加州演算法、Dia 和 Rose[2]提出之類神經網路法、Adeli 和 Karim[3]提出之模糊類神經網路、周義華等人[4]提出之模糊類聚分析，以及王秀帆[5]提出之灰預測演算法等。惟由於事件及其對應之偵測交通資料蒐集不易，許多研究大多利用車流模擬軟體，產生事件發生及交通變化狀況等資料，來進行模式之訓練與驗證，或為國外高速公路之車流資料。因此，各演算法雖均能達到一定水準之事件偵測準確率與誤報率，但能否實際應用於國內高速公路之事件偵測，仍有待考驗。因此，最好能利用我國高速公路之實際事件資料，進行模式之

<sup>1</sup>逢甲大學交通工程與管理學系助理教授。

<sup>2</sup>逢甲大學交通工程與管理學系講師。

<sup>3</sup>逢甲大學交通工程與管理學系研究生。

構建與驗證，以確保模式之可用性。因此，本文乃蒐集國內高速公路之事件與其對應之交通資料，建構事件偵測模式。

此外，基因模糊邏輯控制 (genetic fuzzy logic controller, GFLC) 於相當多領域之應用績效，已獲驗證 (例如，Chen 等人[6]、Xiong 和 Litz[7]，以及邱裕鈞和藍武王[8]等)，但尚未有應用於事件偵測之相關研究發表。基此，本文乃嘗試利用 GFLC 建構一套事件自動偵測系統，除可透過模糊理論處理資訊之不確定性因素外，尚能利用遺傳演算法 (genetic algorithms, GAs) 自動學習邏輯規則與隸屬函數參數之校估，利用反覆尋優程序來建構 GFLC 事件偵測演算法，期使事件偵測績效達到最佳，以供相關單位在規劃高速公路事件偵測系統時之參考。本文章節安排如后：第二節說明以 GAs 建構 FLC 之模式，第三節說明實例研究，以驗證 GFLC 之精確度與適用性，最後提出結論與建議。

## 貳、模式建構

### 一、GFLC 模式

本文採用邱裕鈞和藍武王[8]所發展之反覆演化 GFLC 模式為基礎，進行事件偵測模式之構建。以下簡介該模式之編解碼與運作方式。

#### (一) 邏輯規則庫之編解碼

該模式在建構邏輯規則庫的部份主要是根據 Thrift[9] 所提之編解碼方式來操作。因為此一方法可使染色體長度最為精簡，有效減少尋優過程中花費之時間及電腦記憶體的空間需求。假設模式設定之狀態變數為 2 個，控制變數為 1 個，各具有五等級語意值，其染色體之長度為  $5 \times 5 = 25$  個基因，每個基因值為 0~5，分別以 0 代表該條規則未入選，以 1~5 代表該規則獲選外，亦代表其控制變數之語意值，其編解碼方式如圖一所示。

假設一條染色體為 0000020030000000000004005，代表共有 4 條邏輯規則入選，此 4 條規則分別為：

Rule 1: IF  $x_1 = NS$  AND  $x_2 = NL$  THEN  $y = NS$

Rule 2: IF  $x_1 = NS$  AND  $x_2 = PS$  THEN  $y = ZE$

Rule 3: IF  $x_1 = PL$  AND  $x_2 = NS$  THEN  $y = PS$

Rule 4: IF  $x_1 = PL$  AND  $x_2 = PL$  THEN  $y = PL$

#### (二) 隸屬函數之編解碼

就模式隸屬函數建立的部分，一般以三角模糊數為例，其第  $k$  等級之隸屬參數之左底角、頂點及右底角座標之參數值分別為  $ck_l$ 、 $ck_c$  及  $ck_r$ ，假設一個具有五等級語意值之變數，則此變數所待校估之參數為 15 個。此外，由於各等級模糊數之參數間尚有一定程度相對大小之關係須先加以釐清。就同一等級模糊數而言：應符合  $ck_r \geq ck_c \geq ck_l$ ；而就不同等級模糊數間則應符合  $ck_r \geq ck-1_r$ 、 $ck_c \geq ck-1_c$  以及  $ck_l \geq ck-1_l$ 。如直接以基因演算法校估各參數，勢必要將這些限制條件皆納入模式中，對於模式尋優之績效將會有嚴重的影響。故邱裕鈞與藍武王[8]乃建立一套隸屬函數之編解碼方式進行參數校估：

假設隸屬函數形式：第1與第5等級之隸屬函數均為左傾及右傾之直角三角形，即 $c_{1c} = c_{1l} = c_{\min}$ 及 $c_{5c} = c_{5r} = c_{\max}$ ，其中， $c_{\min}$ 及 $c_{\max}$ 分別為該變數之最小值與最大值，為外生變數；第2~4等級之隸屬函數均為等腰三角形，即 $c_{kc} = (c_{kr} + c_{kl})/2$ ， $k=2, 3, 4$ ，因此，所須校估之參數僅8個。參數相對大小：

$$c_{\max} \geq c_4^r \geq \frac{c_5^l}{c_3^r} \geq \frac{c_4^l}{c_2^r} \geq \frac{c_3^l}{c_1^r} \geq c_2^l \geq c_{\min}$$

，其中， $c_{5l}$ 與  $c_{3r}$ 、 $c_{4l}$ 與  $c_{2r}$ 及 $c_{3l}$ 與  $c_{1r}$ 之

相對大小未定。故設計9個位置變數： $r_1 \sim r_9$ ，用以代表其相對大小位置。每個位置變數 $r_i$ 係以4個實數基因代表之（如圖二），以使求解參數之精確度達小數點後兩位，其最大值為99.99，最小為0。因此，以2個狀態變數與1個控制變數，且各具有五等級語意值為例，其染色體長度共計 $3 \times 9 \times 4 = 108$ 個基因。最後，再計算參數值：各位置變數經校估後，再依公式（1）~（8）以及等腰三角形之特性，分別計算各參數值。

$$c_2^l = c_{\min} + r_1 \times sf \tag{1}$$

$$c_1^r = c_2^l + r_2 \times sf \tag{2}$$

$$c_3^l = c_2^l + r_3 \times sf \tag{3}$$

$$c_2^r = \max\{c_1^r, c_3^l\} + r_4 \times sf \tag{4}$$

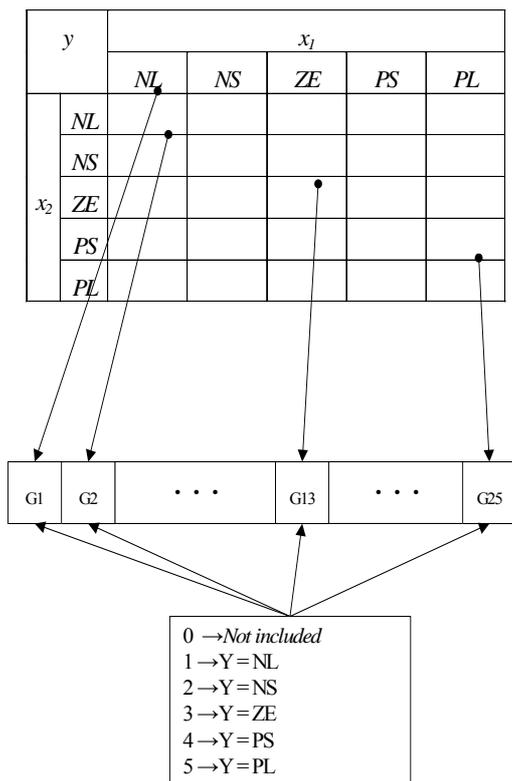
$$c_4^l = \max\{c_1^r, c_3^l\} + r_5 \times sf \tag{5}$$

$$c_3^r = \max\{c_2^r, c_4^l\} + r_6 \times sf \tag{6}$$

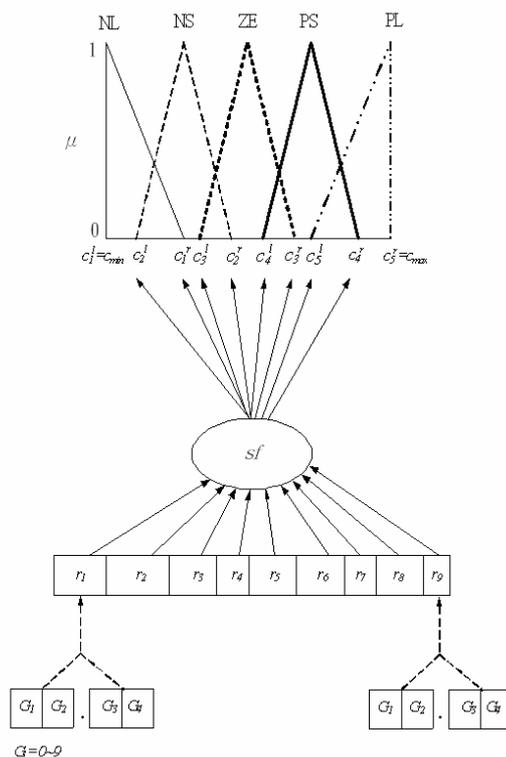
$$c_5^l = \max\{c_2^r, c_4^l\} + r_7 \times sf \tag{7}$$

$$c_4^r = \max\{c_3^r, c_5^l\} + r_8 \times sf \tag{8}$$

其中， $sf = \frac{(c_{\max} - c_{\min})}{\sum_{i=1}^9 r_i}$ 。



圖一 邏輯規則庫編解碼方式示意圖



圖二 隸屬函數之編解碼方式示意圖

(三)遺傳演算法則之運作

針對遺傳運算法則之交配與突變方式，邱裕鈞與藍武王[8]除採用 Max-min-arithmetical 交配方式，同時納入雙點交配法(two-point crossover)如圖三所示，第一個交配點左半部基因與第二個交配點右半部之基因不做變動，僅是染色體中間部分做基因交換之動作。根據以上兩種交配運算機制，每次進行交配運作時，將由2個母代染色體產生8個子代染色體，再依據染色體適合度值挑選出最優的2個染色體，置回族群中，其他6個則予以刪除。

1.Max-min-arithmetical 交配演化方式：

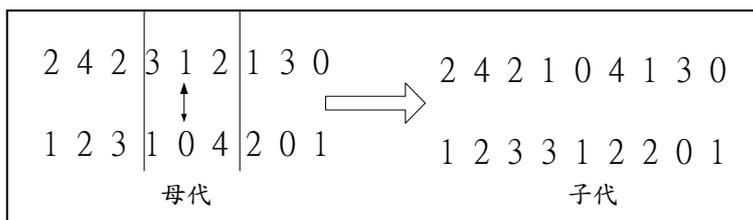
$$G_1^{t+1} = aG_w^t + (1-a)G_v^t \tag{9}$$

$$G_2^{t+1} = aG_v^t + (1-a)G_w^t \tag{10}$$

$$G_3^{t+1} \text{ with } g_{3k}^{t+1} = \min\{g_{wk}^t, g_{vk}^t\} \tag{11}$$

$$G_4^{t+1} \text{ with } g_{4k}^{t+1} = \max\{g_{wk}^t, g_{vk}^t\} \tag{12}$$

其中， $G_{it+1}(i=1\sim 4)$ 代表產生之子代染色體， $g_{ikt}$ 代表該染色體之第 $k$ 個基因。  $G_w^t$ 及  $G_v^t$ 為獲選之母代染色體， $g_{wkt}$ 及  $g_{vkt}$ 分別代表其染色體之第 $k$ 個基因。  $a$ 為一設定之參數， $t$ 代表演化之世代數。



圖三 雙點交配方式示意圖

## 2. Non-uniform 突變演化方式：

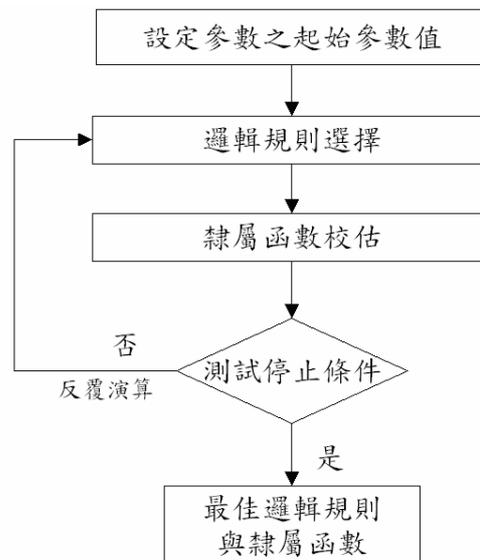
Non-uniform 突變方式，不同於 SGAs 之基因突變僅作 0 或 1 之變動，其納入模擬退火(simulated annealing)觀念，於尋優初期之突變率較大，隨著演化世代數之增加，突變率逐漸縮小，以便尋優結果達到收斂。其演化方式如下：

$$g_k = \begin{cases} g_k + \Delta(t, g_k^u - g_k) & \text{if } b = 0 \\ g_k - \Delta(t, g_k - g_k^l) & \text{if } b = 1 \end{cases} \quad (13)$$

其中， $g_k$  為產生子代染色體之第  $k$  個基因。 $g_k^u$  及  $g_k^l$  分別為該基因之上、下限值。 $b$  為一隨機數值(random number)，其值為  $\{0,1\}$ 。 $\Delta(t, y)$  為一機率函數，給定其參數後，將輸出一數值，位於  $[0, y]$ 。其機率函數型式為： $\Delta(t, y) = y(1 - r^{(1-t/T)^h})$ 。其中， $r$  為一隨機數值，其值位於  $[0,1]$ 。 $T$  為最大演化世代數。 $h$  為一設定之參數。

### (四) GFLC 反覆演化機制

該模式利用一反覆演化之方式來進行模糊邏輯規則之挑選與隸屬函數參數之校估。模式在起始演化時先自行設定各變數之隸屬函數參數，經由運算後得到最適之邏輯規則，以此規則組成下求解最佳之隸屬函數參數，如此反覆運算直至達到演化尋優條件為止，亦即可使隸屬函數參數與邏輯規則相互配合達到最佳之組合。其演化流程如下圖四所示。



圖四 GFLC 反覆演化機制

## 二、 GFLC 事件偵測模式

### (一) 變數選擇

有關事件偵測之輸入變數，大致包括兩大類，一類是偵測器之偵測數值，包括事件上、下游偵測器或同一偵測器前、後時段所測得交通資訊。另一類則是上、下游偵測器或同一偵測器前、後時段所測得交通資訊之差距。至於一般偵測所得之交通資訊，則包括流量、速度、佔有率等三種。因此，可採用之變數甚多。惟 GFLC 模式中，由於潛在邏輯規則及隸屬函數校估參數之數量會隨著狀態變數之增加而大幅成長。故一般在應用上，狀態變數之個數大

多限制在三個以下。在此一限制下，本模式無法如類神經網路模式一樣，同時考量許多輸入變數。因此，變數之選擇極為重要。參考相關文獻及考量所能蒐集資料，本文將以上、下游偵測器所測得之流量（ $Q_u$ 、 $Q_d$ ；u表上游、d表下游）、速度（ $V_u$ 、 $V_d$ ）、佔有率（ $K_u$ 、 $K_d$ ）等三項交通資訊之差值，即 $\Delta Q (= Q_u - Q_d)$ 、 $\Delta V (= V_u - V_d)$ 及 $\Delta K (= K_u - K_d)$ 作為狀態變數。控制變數則設為是否有事件。假設各變數之語意等級均為5個，且均為三角形之模糊數。在邏輯規則之選擇部份，共計有 $5 \times 5 \times 5 = 125$ 條規則可供選擇，每一個基因代表一條規則，每一個染色體共有125個基因。在隸屬函數之校估部份，設計9個位置變數下，共計有 $4 \times 9 = 36$ 個參數可供校估，每四個基因代表一個參數，每一個染色體共有 $4 \times 9 \times 4 = 144$ 個基因。

### (二)適合度值

本文以GFLC建構之事件偵測系統之主要目標是在使訓練樣本值與預測值間之誤差達最小化。以下為總誤差平方和之計算方式：

$$E = \frac{1}{2} \sum_{i=1}^N [t_i - y_i]^2 \quad (14)$$

其中，E為總誤差平方和。N為訓練樣本總數。ti為第i個訓練樣本之事件發生與否模式預測值。yi為第i個訓練樣本之事件發生與否值。由於GAs的適合度值以值愈大為愈佳之原則，因此，適合度值訂為 $f=1/E$ 。

### (三)績效評估

一般在評估事件偵測模式之優劣，大多採用三個指標：偵測率（detection rate, DR）、誤報率（false alarm rare, FAR）以及平均偵知時間（mean time to detection, MTD）。說明如下：

1. 偵測率之定義為：在所有發生的事件中，演算法所偵測出事件數佔實際發生事件數的百分比。其公式如下所示：

$$\text{偵測率} = \frac{\text{演算法偵測之事件總數}}{\text{所有實際發生之事件總數}} \times 100\%$$

而偵測率的認定係設定為在同一事件的訊息，只要事件未結束，這過程中無論產生多少筆事件訊息，皆視為同一事件之訊息。

2. 誤報率之定義為：使用無事件資料進行離線測試時，產生錯誤訊息次數佔總訊息判斷次數的比率。其公式如下所示：

$$\text{誤報率} = \frac{\text{演算法誤報為有事件之訊息總數}}{\text{無事件之訊息總數}} \times 100\%$$

而誤報率的認定係設定為在同一無事件的資料內，只要產生一筆事件訊息，皆視為誤報。

3. 平均偵知時間之定義為：事件實際發生時間與被偵測出時間的平均延遲時間。

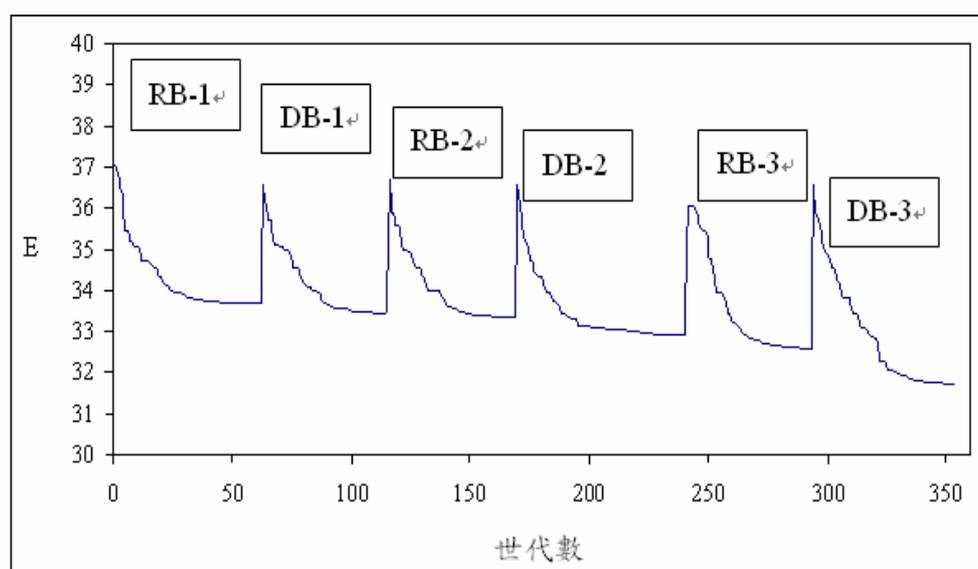
## 參、實例應用

### 一、資料說明與參數設定

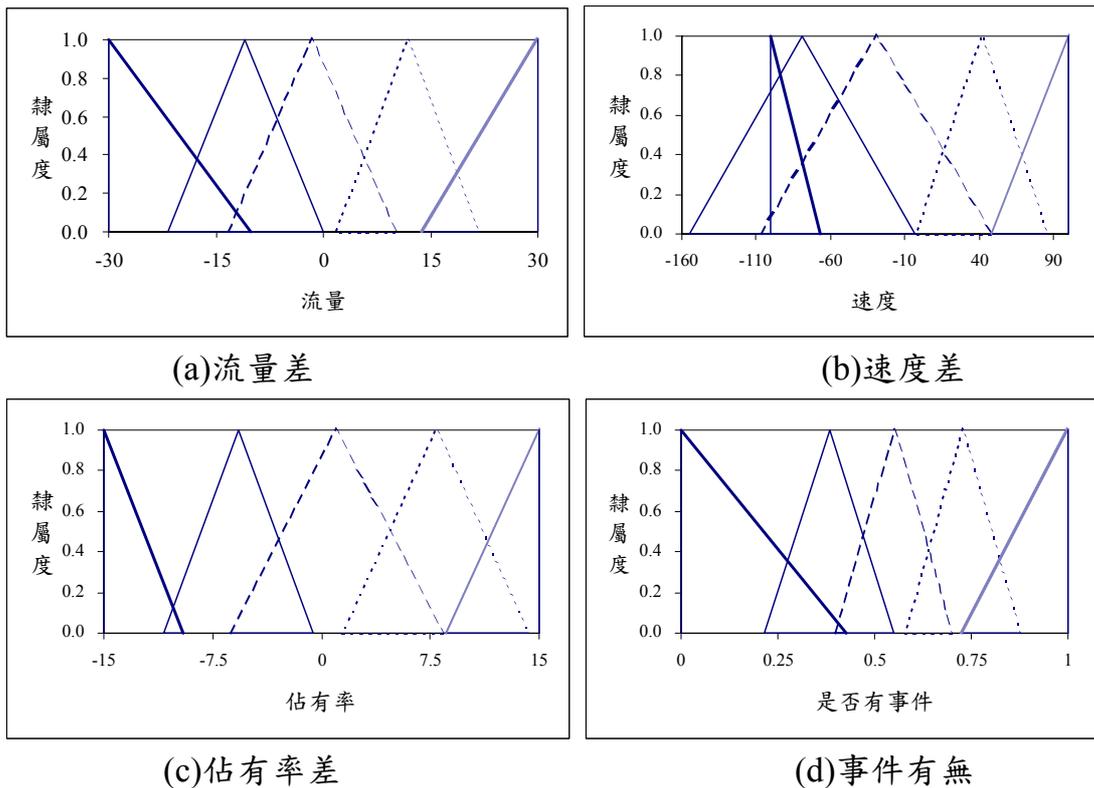
本文以中山高北上路段民國 93 年 11 月份之事件為例，進行各事件發生地點、時間及前後 5 分鐘內之上、下游偵測器每 20 秒記錄之交通資訊，包括流量、速度及佔有率等三項資料。由於各車道均埋設有偵測器，故將同一地點各車道所測得之交通資料予以彙整（流量予以加總，速度及佔有率則加以平均）成路段交通資料。過濾偵測器資料明顯有誤之事件樣本後，共挑選 11 個有事件及 11 個無事件（每個均有 30 筆交通資料）作為模式之訓練與驗證資料。本模式之參數設定：族群數=200，交配率=0.9，突變率=0.01，成熟率=80%， $h=0.5$ ， $\epsilon=0.01$ ，採用面積法來解模糊化。根據資料分佈情形，將三個狀態變數及一個控制變數之隸屬函數參數分別予以設定：流量差之最小值為-30pcu/hr，最大值為 30pcu/hr；速度差之最小值為-100km/hr，最大值為 100km/hr；佔有率差之最小值為-15，最大值為 15；事件發生與否設定其最小值為 0，最大值為 1。

### 二、結果分析

本模式之學習演化結果，如圖五所示。其中 RB-# 表示選擇邏輯規則程序之第#次之尋優過程；DB-#表示選擇隸屬函數校估程序之第#次之尋優過程。RB-1 和 DB-1 為第一次尋優結果；RB-2 和 DB-2 為第二次尋優結果；RB-3 和 DB-3 為第三次尋優結果，透過三次反覆運算，得到之最佳邏輯規則及隸屬函數共經過 353 個世代之尋優過程。其事件發生與否之總誤差平方和，由最初的 37.05 達至最佳之 31.74。各變數之最佳隸屬函數型式，如圖六所示。



圖五 模式之演化尋優過程



圖六 各變數隸屬函數之校估結果

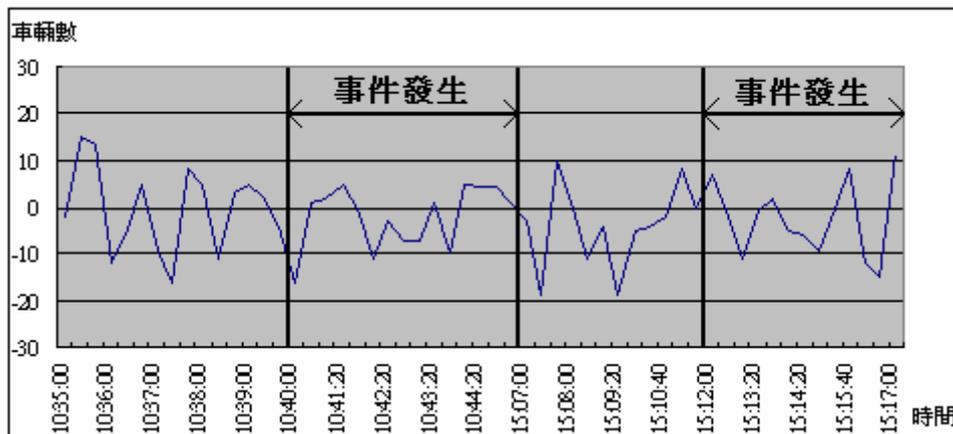
由於模式輸出之結果為一 $[0,1]$ 間之連續變數，為能明確預測是否有事件發生。本文以 0.5 作為判斷門檻，輸出值大於 0.5 則預測有事件發生，小於 0.5 則預測無事件發生。本模式之運作績效如下：偵測率 100%、誤報率 36.36%、平均偵知時間則為 53.33 秒。此一績效顯較相關研究為差，有必要進一步加以檢討。

#### 肆、討論

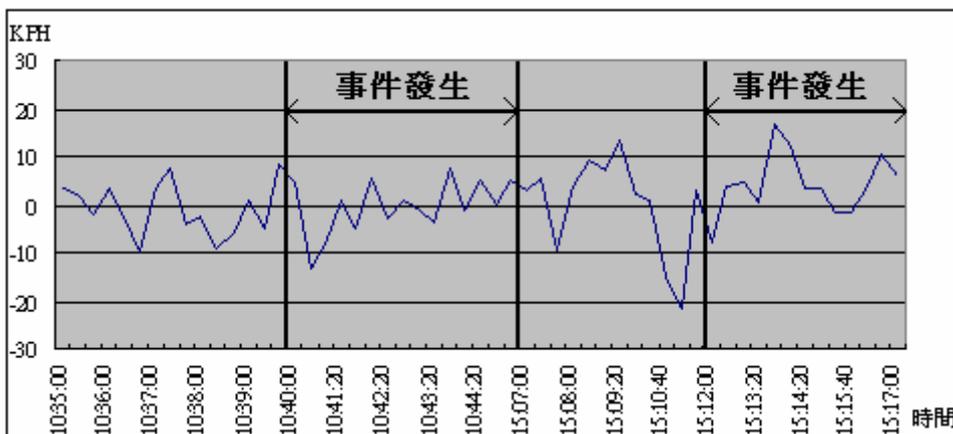
由實例應用之績效分析結果顯示，本模式之績效仍有待改善。進一步分析其原因可歸納有三：狀態變數之選擇不當、事件登錄資料錯誤，以及偵測資料有誤等。其中，在狀態變數選擇上，如前述，GFLC 在應用上囿於狀態變數個數之限制，無法像類神經網路模式同時考量十多個變數，因此，有必要嘗試不同輸入變數組合，以達到最佳偵測效果。另外，為能同時考量三個以上變數，也可考慮將數個變數整合成一變數，以有效減少變數個數需求，又能考慮必要之輸入變數。至於事件記錄或偵測資料有誤，也會影響本模式之訓練與偵測績效。為檢驗資料之正確性，茲挑選兩件事件之偵測資料加以繪圖分析，結果如圖七及圖八所示。圖七為上、下游偵測器所測得之交通資訊差距，圖八則為上、下游偵測器所測得之交通資訊。由此兩圖觀之，事件發生前後之交通資訊並未發現有顯著變化，此亦可能為導致本模式績效不佳的另一原因。而此一原因可能是事件之登錄時間與實際發生時間不同或偵測資料準確率較低所致。其中，又以前者之可能性更高，因為本模式係利用每 20 秒一筆之交通資訊進行訓練，而每一事件僅蒐集前後各 5 分鐘資料。因此，如果事件登錄時間，與實際發生時間差距達 5 分鐘以上，則極可能無法找到其對應之交通資訊。因此，蒐集更為正確之事件資料及其所對應之交通資料實為一重要課題。

## 伍、結論與建議

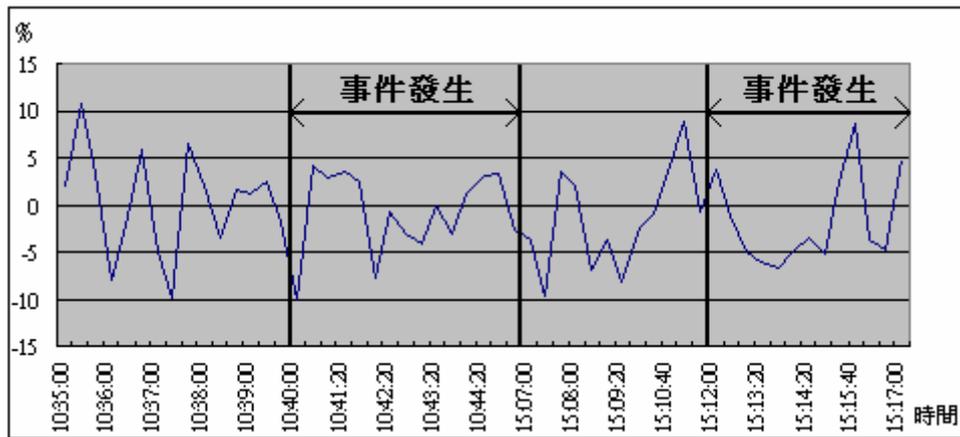
為避免以往依據車流模擬產生事件資料所建構之事件自動偵測模式在實際應用時，可能發生落差之問題，本研究旨在以實際事件資料及偵測器測得交通資料為基礎，利用 GFLC 建構一高速公路之事件偵測模式。本模式以上下游偵測器所測得之流量、速度和佔有率之差距作為狀態變數，控制變數則為是否有事件發生。以實際高速公路事件資料加以驗證後，發現本模式雖可達 100% 之偵測率及平均偵知時間為 53.33 秒，但誤報率卻高達 36.36%，偵測績效不甚良好。究其原因可能有三：狀態變數之選擇不當、事件登錄資料錯誤，以及偵測資料有誤等。而針對所蒐集資料加以檢視後，確發現事件發生前後，未出現顯著之交通狀況變化情形。因此，建議後續研究可考慮嘗試不同狀態變數組合，以達到最佳偵測效果。而且為能同時考量三個以上變數，也可考慮將數個變數整合成一變數，以有效減少變數個數需求。另外，如何蒐集更為正確的事件資料及其所對應之交通資料，亦值得加以研析，以供未來事件偵測模式訓練之基礎。最後，由於無論是那一種事件偵測模式，均需奠基於正確之交通資訊，方能成功。惟目前所使用的車輛偵測器，無論是環路線圈、紅外線、微波或影像辨識，均存有相當程度之誤差，在無法克服此一技術問題前，或可考慮建立一套交通偵測資料之調校模式，以產生更可信、精確的交通資訊，以利事件自動偵測系統之落實應用。



(a) 流量差

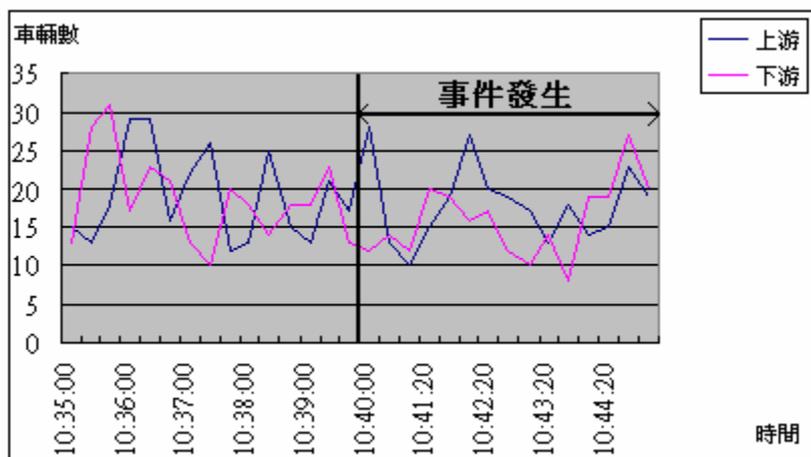


(b) 速度差

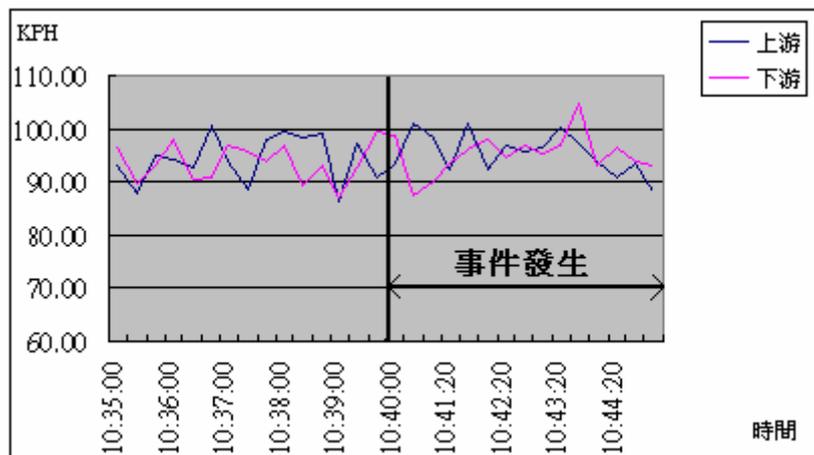


(c)佔有率差

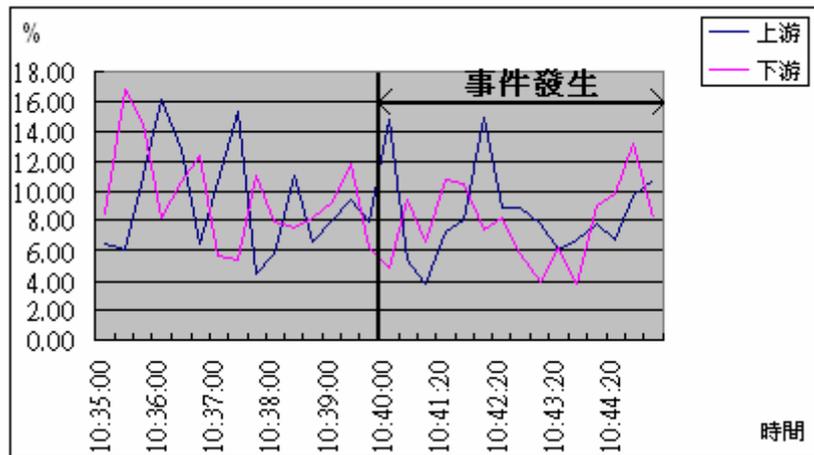
圖七 上、下游偵測器所測得之交通資訊差距



(a)流量



(b)速度



(c)佔有率

圖八 上、下游偵測器所測得之交通資訊

### 參考文獻

1. Ai-Deek, H.M., Ishak, S.S. and Khan, A.A., "Impact of freeway geometric and incident characteristics on incident detection," Journal of Transportation Engineering, Vol.122, No.6, pp.440-446, 1996.
2. Dia, H. and Rose, G., "Development and evaluation of neural network freeway incident detection models using field data," Transportation Research Part C , Vol.5 , No.5 , pp.313-331, 1997.
3. Adeli, H. and Karim, A., "Fuzzy-wavelet RBFNN model for freeway incident detection," Journal of Transportation Engineering, Vol.126, No.6, pp.464-471, 2000.
4. 周義華、許鉅秉、李啟仲,「高速公路事故屬性即時自動鑑別之方法研究」, 運輸計劃季刊, 第二十九卷第三期, 第 499~528 頁, 民國 89 年。
5. 王秀帆,「應用灰色預測於高速公路事件自動偵測之研究」, 國立中央大學土木工程學所碩士論文, 民國 92 年。
6. Shieh, C.S., "Genetic fuzzy control for time-varying delayed uncertain systems with a robust stability," Applied Mathematics and Computation, 131, pp.39-58, 2002.
7. Xiong, N. and Litz, L., "Reduction of fuzzy control rules by mans of premise learning-method and case study," Fuzzy sets and systems, 132, pp.217-231, 2002.
8. 邱裕鈞、藍武王,「應用遺傳演算法建構適應性模糊邏輯控制系統—以跟車行為為例」, 中華民國運輸學會第 16 屆學術論文研討會, 第 515-526 頁, 民國 90 年。
9. Thrift, P., "Fuzzy logic synthesis with genetic algorithms," Proceeding of the Fourth International Conference on Genetic Algorithms, pp.509-513, 1991.

