

應用遺傳規劃法建立櫻花鉤吻鮭魚群數量 對環境之反應函數

Application of Genetic Programming to Establish the Environmental Response Function for Taiwan Salmon

國立台灣大學生物環境系統工程學系
副教授

童慶斌

Ching-Pin Tung

國立台灣大學生物環境系統工程學系
碩士班研究生

李宗祐

Tsung-Yu Lee

摘 要

本研究結合了資料分類與遺傳規劃法建立櫻花鉤吻鮭魚群數量與河川溫度及流量的關係，櫻花鉤吻鮭是分佈在最低緯度之陸封型鮭魚，因此對其保育有極重要的生態意義。櫻花鉤吻鮭對於河川溫度非常敏感，且從未在超過水溫 18°C 的水域發現其蹤跡，七家灣溪是櫻花鉤吻鮭的主要生活棲地，但近年來已經有部分的河段水溫已經超過 17°C；除此之外，洪水發生將魚群沖刷至下游因而導致死亡亦是魚群減少的主要因素。近年來的研究顯示，氣候變遷可能會使七家灣溪水溫升高並有流量增加的情形，因此建立河川溫度及流量與櫻花鉤吻鮭族群數量的關係成爲了解環境因子改變對於魚群數量影響的重要依據。本研究在此提供一個較新的方法來模擬魚群數量，利用資料分類將魚群數量依溫度及流量的屬性劃分爲四類，再利用遺傳規劃法優選每個分類下溫度與流量跟魚群數量的關係，結果顯示資料分類結合遺傳規劃的模式可合理用來模擬櫻花鉤吻鮭的數量。

關鍵詞：生態系統，生態模擬，啓發式演算法。

ABSTRACT

The genetic programming with data classification is proposed to identify the functions to project the effects of stream temperature and flood on Taiwan Salmon. Taiwan Salmon is the land locked species, which can be found in the lowest latitude in the world. Thus, it is an important indicator for Taiwan's ecological conservation. Taiwan

Salmon is very sensitive to stream temperature, and can not exist with stream temperature higher than 18°C. The ChiChaWan Creek is the important habitat of Taiwan Salmon. Stream temperature more than 17°C has been observed in some stream sections in recent years. Besides, it has also been observed that population of Taiwan Salmon is reduced significantly due to flood flushing them to downstream and causes death. Recent research suggests that climate change may result in higher stream temperature and increasing flood in Taiwan Salmon's habitat. Thus, it is very necessary to develop functions to describe the effects of stream temperature and flood on Taiwan Salmon's population. A new approach is proposed to identify the functions by using genetic programming with data classification. Data classification is applied to divide fish population into different categories according to magnitude of flood and temperature, and then genetic programming is utilized to develop functions for each category, respectively. The results indicate that the proposed approach can successfully develop functions to provide reasonable predictions of Taiwan Salmon's population based on given temperature and possible flood.

Keywords: Ecosystem, Ecological modeling, Heuristic algorithm.

一、前言

櫻花鉤吻鮭（學名 *Oncorhynchus masou formosanus*）是台灣本島之自然文化資產。這種鮭鱒魚類原屬於太平洋洄游性魚類，具有溯河洄游的生活形態，可能在冰河時期遭遇板塊運動和河川變遷，而被環境隔絕在大甲溪與高山溪流中，成為陸封型的魚種。這種陸封型的鮭鱒魚類，也只曾經在日本、韓國及大陸東北地區發現過，而台灣是全球鮭魚分佈區域緯度最低的地區。

櫻花鉤吻鮭在日據時代，廣布於大甲溪上游的合歡溪、南湖溪、司界蘭溪、七家灣溪、武陵溪及有勝溪等支流，但是到了民國四十年至七十年代，就僅在司界蘭溪、七家灣溪及武陵溪發現有鮭魚的蹤跡，進入八十年代，依據林曜松教授及汪靜明教授的調查報告指出，目前櫻花鉤吻鮭分佈僅限於雪霸國家公園武陵地區的武陵溪下游、七家灣溪流及其上游桃山西溪（汪靜明，1994）。造成棲地縮減的原因，除了人為造成的棲地型態改變及水質惡化外，水溫上升亦可能是造成棲地縮減的主要因素。大甲溪上游水域分佈圖如圖 1 所示。

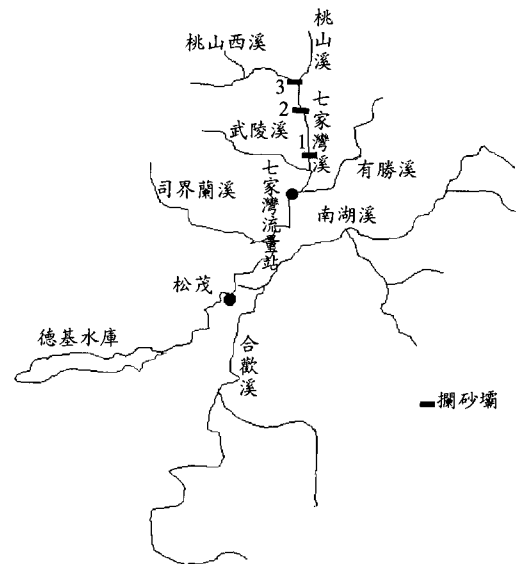


圖 1 大甲溪上游水域分佈圖

河川溫度一直被認為是影響水生生態最重要的因素之一，溫度的變化除影響水中魚類及其他水棲生物的成長、存活率外，魚類的遷徙、繁殖、孵化等活動，甚至魚類的競爭能力、河川的自淨能力、污染傳輸及河川內部的化學反應也都深受河川溫度的影響。其中以魚類的活動最直接

也最容易觀察，有些魚類會在特定溫度範圍的水域裡活動，以櫻花鉤吻鮭為例，其生活水域水溫介於 9~17°C 之間，到了繁殖、孵化期間水溫更需降低至 12°C 以下（曾晴賢，1999）。由於大部分的水生動物對水溫的要求很高，所以河川的水溫在生態的研究中，一直扮演著重要的角色。一般而言，溪流水溫由上游往下游遞增，根據調查報告顯示，七家灣溪在民國 74 年至 86 年間，其溪流平均溫度之 12°C 等值線往上游退縮約 1.56 公里（楊正雄，1997），嚴重影響河川生物之棲地分布，因此探討水溫之變動情形，成了延續魚類生命相當重要的課題。

除了河川溫度對魚群所造成的影響之外，洪水亦是一個影響魚群分佈的重要因素，部分研究顯示（曾晴賢，2001）每年因颱風引起的洪水常常會使得櫻花鉤吻鮭的魚群數量大量減少，洪水發生時會夾帶大量的泥砂導致水質劣化，部分魚群因無法承受如此惡劣的環境而死亡，部分魚群則被洪水沖至下游河段，而下游河段通常不為適合櫻花鉤吻鮭生長的棲地環境，再加上因為台灣的地形之故河川多短且急，所以必須在上游地區建築大量的攔砂壩以延長水庫的壽命，而攔砂壩往往成為魚群遷移的障礙，以櫻花鉤吻鮭為例，由於攔砂壩的建築，使得櫻花鉤吻鮭活動範圍被限制在相鄰壩體之間，造成近親交配，除此之外，被洪水沖至下游的魚群也將喪失重返上游棲地的機會，因為下游環境往往高於其適合之生長溫度，使得其暴露在不適當之棲地環境而招致死亡。

國外學者很早就開始觀察鮭鱒魚，因為鮭鱒魚為他們重要的魚獲來源之一，只是當時的焦點均擺在各項環境因子對魚隻體型大小之影響，直到近幾年，由於鮭鱒魚的數量有逐漸減少之趨勢，於是開始紛紛投入對魚群數量的研究，評估分析各項環境因子對於魚群數量的影響，如 Chen et al. (2000) 利用前一時期產卵的魚群數量及海面平均水溫結合模糊分類及遺傳演算法決定線性迴歸函數之係數，進而模擬每年迴游之鮭魚數量；Bradford et al. (2000) 提出近年來土地利用、氣候變遷及人為過渡補抓行為對於鮭魚密度的

負面影響；Shaw and Richardson (2001) 提出河川泥砂含量對鱒魚的影響；Cattaneo et al. (2002) 利用統計分析方法提出大流量為影響鱒魚數量的最主要因素。上述的參考文獻中除了 Chen et al. (2000) 所建立的模式可用來模擬並預測魚群數量外，大部分的研究都是將魚群數量對單一環境因子作單變數的線性迴歸分析，只可以看出單一環境因子對魚群數量的影響趨勢，而無法利用所建立的模式預測魚群數量。

生態系統中所存在的環境因素眾多且關係複雜，魚群的數量亦受到環境因子的複雜影響，如環境因子如何影響其數量？影響的程度又如何？若不同環境同時作用又會造成什麼影響？都是模擬魚群數量所必須考慮到的課題，本研究結合資料分類及遺傳規劃法所發展之模式即能綜合考慮上述因子，與線性迴歸的差別是，本研究並非將所有的影響因子與反應變數直接作迴歸分析，而是先透過資料分類將環境因子與魚群數量的關係分為四類，分別是發生在「高溫度、高流量」「高溫度、低流量」、「低溫度、高流量」及「低溫度、低流量」的魚群數量，再透過迴歸分析推求各類分組下魚群數量與流量及溫度間的關係。但環境因子對魚群數量的影響關係複雜，若單純用線性迴歸分析將很難完全描述環境因子與魚群數量的關係，而遺傳規劃法就提供了這樣的機制，在不限定方程式為線性的情形下，利用優選的方式產生最適合來描寫環境因子與魚群數量之函數，比線性迴歸分析更具有彈性，經由優選的方式更能凸顯環境因子在不同分類下對魚群數量之影響。

在接下來的章節裡，本研究先透過環境因子分析，分析櫻花鉤吻鮭數量受到現地哪些環境因子的影響，再利用遺傳規劃法建立魚群數量與環境因子間的關係。

二、環境因子分析

本研究地區七家灣溪流位於台灣中部地區，又稱武陵，位於北緯 24°20' 與 24°25' 及東經 121°10' 與 121°20' 之間，流域範圍位於雪霸國家公園的東南隅。本研究探討的重點為櫻花鉤吻鮭

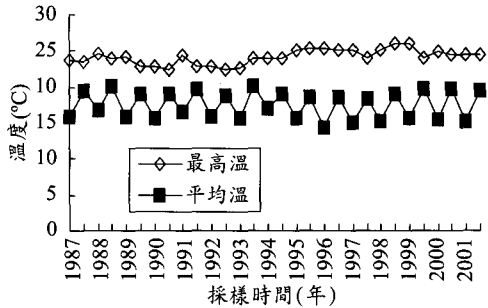


圖 2 採樣前最高氣溫與平均氣溫之關係圖

數量受環境因子影響的情形，而七家灣溪一號壩到三號壩間其主要的分佈區之一，所以本研究主要針對此河段進行魚群數量之模擬。

在建立魚群數量模式前，必須先瞭解哪些環境因子會對魚群數量造成影響，參考國內外學者的研究，發現很多因素都會影響到魚群的數量，如水溫（楊正雄，1997）、泥砂量(Shaw and Richardson, 2001)及大流量(Cattaneo et al., 2002)等因素均會對魚群的數量產生影響，為了使模式的運算更有效率且有意義，本研究根據七家灣溪流域之量測資料事先針對幾個可能造成櫻花鉤吻鮭數量減少的环境因素作評估，挑選出較具代表性的環境因素且因子間應僅量相互獨立。其中日氣溫資料取自松茂氣象站，日流量及日輸砂量資料取自七家灣流量站，櫻花鉤吻鮭數量調查資料則根據雪霸國家公園管理處提供之「櫻花鉤吻鮭族群監測與生態調查」報告書中取得(曾晴賢，1998；1999；2000；2001)，分析期間介於1987~2001年。

在水溫方面，水溫已被證實是影響櫻花鉤吻鮭生長的重要因素，其生活水域的水溫介於9~17°C之間，到了繁殖、孵化期間水溫更需降低至12°C以下(曾晴賢，1999)，所以水溫為模式中不可缺的環境因子。但由於資料取得的困難，本研究以大甲溪上游松茂站的氣溫資料取代水溫，因為從過去的文獻得知，水溫與氣溫相關(Mohseni, 1999)，以氣溫模擬魚群數量和直接以水溫模擬魚群數量，只會反應在變數前的係數上。櫻花鉤吻鮭魚群數量在每年夏、秋兩季各調查一次，研究中是以前次調查與此次調查間每日

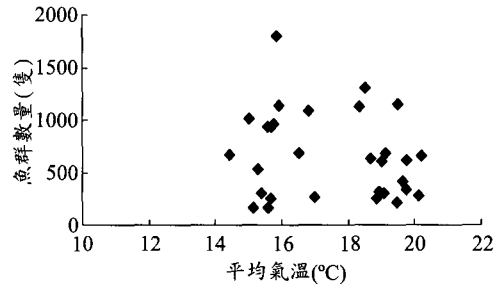


圖 3a 魚群數量與平均溫度之關係圖

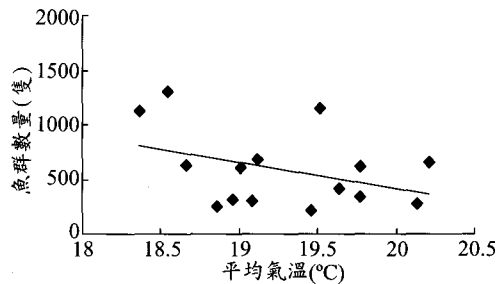


圖 3b 平均溫度大於 18°C 時，魚群數量與平均溫度之關係圖

平均氣溫的平均值作為影響此次魚群數量的環境因子，原因為兩次調查間平均氣溫隨季節的變化較為劇烈，較能反應不同採樣魚群數量點間溫度的差異，不像最高溫幾乎維持在一定值（如圖 2 所示）。而魚群數量與平均氣溫的關係以圖 3a 來表示，圖中虛線表示整體資料的趨勢線，從整體資料來看，魚群數量似乎與平均氣溫的相關性甚低，但若將魚群資料以平均氣溫 18°C 為界分為兩組資料，則可以發現，當平均氣溫大於 18°C 時，魚群數量會隨著溫度的升高而有明顯的遞減趨勢，如圖 3b 所示，這樣結果也顯示當溫度較低時魚群數量較不受影響，但溫度較高時則不然。

在流量方面，是以前次調查與此次調查間每日流量中的最大值作為影響此次魚群數量的環境因子，魚群數量與採樣前發生之最大流量關係如圖 4 所示，最大流量與魚群數量的關係略呈現三角形的分佈，三角形的分佈反映了兩個特性，第一、流量越大則魚群數量越少；第二、流量越小則魚群數量的變化範圍越大，第一個特性反映

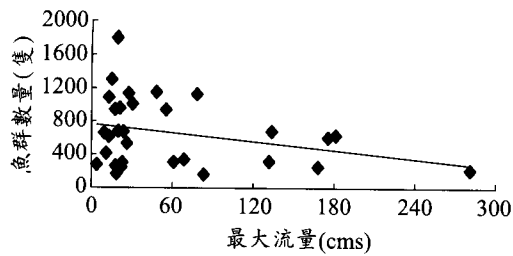


圖 4 魚群數量與採樣前發生之最大流量之關係圖

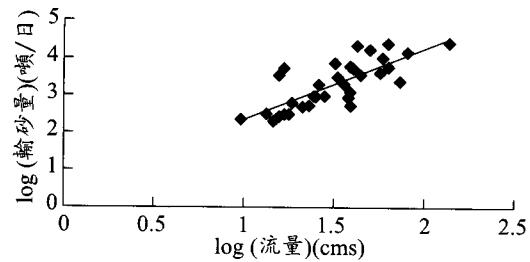


圖 7 輸砂量與平均流量之關係圖

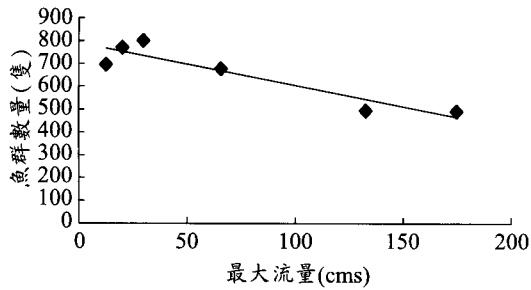


圖 5 依超越機率分組之各組相對應魚群數量的平均數相對於依超越機率分組之各組平均流量

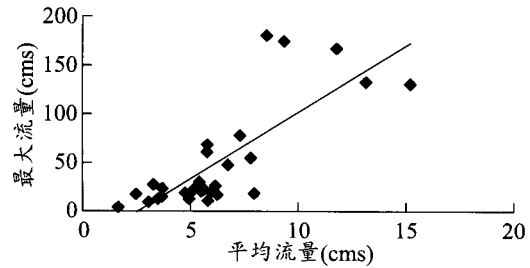


圖 8 採樣前最大流量與平均流量之關係圖

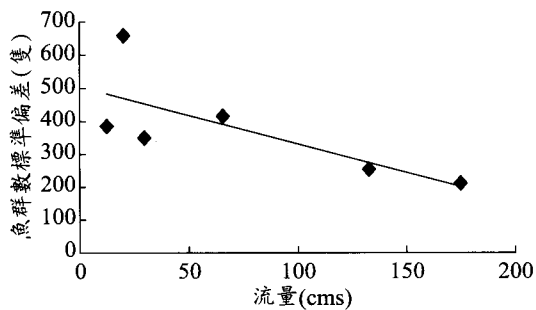


圖 6 依超越機率分組之各組相對應魚群數量的標準偏差相對於依超越機率分組之各組平均流量

在平均魚群數量上，而第二個特性則會反映在魚群的標準偏差上，於是本研究依上述之動機作了以下分析。

以歷年採樣前發生之最大流量計算其超越機率，依據超越機將歷年發生之最大流量分為六組，分別計算在超越機率 $<10\%$ 、 $10\% \sim 20\%$ 、 $20\% \sim 30\%$ 、 $30\% \sim 50\%$ 、 $50\% \sim 70\%$ 、 $70\% \sim 100\%$ 組內之魚群平均數及標準偏差，其結果如圖 5、圖 6 所示。這樣的結果顯示，當流量足夠大時流

量完全影響了魚群的數量，而當流量不足夠大時則可能有其他的環境因素使得魚群數量產生變動。

在河川輸砂量方面，雖然輸砂量亦會影響魚群數量，但研究中發現輸砂量(S)與平均流量(Q)之對數值成線性相關，如圖 7 所示，而平均流量又與最大流量成線性相關，如圖 8 所示，這表示模式在建立時輸砂量對魚群數量的影響可以反應在變數-最大流量的係數上，所以模式中並未將輸砂量納為影響魚群之環境因子。

根據上述資料分析的結果，本研究最後採用採樣前平均氣溫及發生之最大流量視為影響魚群數量之主要環境因素。而從分析的過程可以發現，似乎魚群對於不同程度大小的環境因子有不同的反應情形，所以本研究希望藉由資料分類的方式，探討在不同大小環境因子下，魚群數量與環境因子間的關係。

三、研究方法

本研究是希望透過遺傳規劃法優選得到櫻花鉤吻鮭魚群數量(Y)與採樣前平均氣溫(X_1)及發生之最大流量(X_2)間之函數關係式，若直接將

蒐集到的原始資料直接代入遺傳規劃法模式中，只能優選出一個函數巨觀的描寫環境因子與魚群數量的關係，爲了深入了解不同程度大小環境因子對於魚群數量的影響，所以本研究事先將原始資料進行分類後再帶入遺傳規劃模式中優選，藉此觀察在不同環境條件下對魚群數量的影響爲何。

3.1 資料分類

資料分類的依據是在各個自變數(X)中找出一門檻值(k_x)，若變數 X 超過門檻值 k_x ，則研究中將它歸納至「大」的分類下，否則將它歸納至「小」的分類下，所以一個變數可以分爲兩個類別；以此類推，若研究中有 n 個變數，且每個變數均可分爲「大」「小」兩類，則自變數可歸納出的分類規則就包含了 2^n 個類別。

舉例來說，若研究中只有兩個變數，分別爲 X_1 及 X_2 ，則自變數可歸納出的分類規則就包含了 $2^2=4$ 個類別，其分別爲：

Category 1: if $x_{1i} > k_{x1}$ and $x_{2i} > k_{x2}$,

then $(x_{1i}, x_{2i}, y_i) \in \text{Category 1}$

Category 2: if $x_{1i} > k_{x1}$ and $x_{2i} \leq k_{x2}$,

then $(x_{1i}, x_{2i}, y_i) \in \text{Category 2}$

Category 3: if $x_{1i} \leq k_{x1}$ and $x_{2i} > k_{x2}$,

then $(x_{1i}, x_{2i}, y_i) \in \text{Category 3}$

Category 4: if $x_{1i} \leq k_{x1}$ and $x_{2i} \leq k_{x2}$,

then $(x_{1i}, x_{2i}, y_i) \in \text{Category 4}$

本研究中的自變數爲兩個，分別爲採樣前平均氣溫(X_1)及採樣前發生之最大流量(X_2)，所以依上述資料分類規則所分類出來的四組資料分別表示在「高溫高流量」、「高溫低流量」、「低溫高流量」及「低溫低流量」四種環境條件下所調查之櫻花鉤吻鮭魚群數量，如此分類的目的是因爲大部分的研究都指出大流量及高氣溫均是影響魚群數量的重要因素，若將原始資料未經分類直接進行優選，將無法透視在不同環境條件下魚群與環境因子間的關係，透過分類將可以清楚的看到在不同環境條件下魚群數量分佈的情形，根據原始資料分佈的情形(如圖 3、圖 4 所

示)，本研究選取的各變數之門檻值分別爲 18°C (k_{x1})及 60cms (k_{x2})。而各分類下魚群數量與環境因子間的關係將透過遺傳規劃法來建立。此門檻值的設定目前是以現地魚群資料與環境因子分佈的情形(圖 3a 及圖 4)來做決定，因爲我們發現在平均溫度超過 18°C 及最大流量超過 60cms 時，魚群數量均會有明顯遞減的趨勢，未來則可考慮利用優選的方式決定或以敏感度分析來判定門檻值選擇的優劣。

3.2 遺傳規劃法

生物在自然環境中如何演變讓自己適應環境，靠的是世代與世代之間的基因演變，而遺傳規劃法就提供了這樣的機制，將問題的解以類似染色體的結構作編碼，染色體之間藉由互相交配逐漸衍生出較好的子代。本節將介紹遺傳規劃法主要操作元件，包括編碼、產生初始解、適合度、複製、交配、挑選、突變、與菁英策略等。

編碼(The structures undergoing adaptation)

在遺傳規劃法中，其編碼方式是具有階層的，且每條染色體的大小、形狀會作動態的改變，也是就每代間染色體的大小不會維持在一定的大小及形狀。其編碼的方式是分別從函數集合(F)及端點集合(T)中反覆選取原件(基因)來建構一串完整的染色體。而函數集合中的每個函數都有特定可以向下延伸串連基因的個數。若函數集合爲 $F=\{f_1, f_2, f_3, \dots, f_n\}$ ，端點集合 $T=\{t_1, t_2, t_3, \dots, t_n\}$ ，則每個函數可向下延伸串連的基因個數爲 $z(f_i)$ ， $i=1\sim n$ 。在端點集合中則可能爲變數或是常數，而函數集合裡的函數可能有下列幾類：

- (1) 算數符號 (i.e., +, -, ×, ÷)
- (2) 數學符號 (i.e., sin, cos, exp, log)
- (3) Boolean operations (i.e., AND, OR, NOT)
- (4) 條件敘述 (i.e., If-Then-Else)
- (5) 迭代運算符號 (i.e., Do-Until)

舉例來說，現在假設有一函數集合 $F=\{+, \times, \sin, \exp\}$ ，且有一端點集合 $T=\{X_1, X_2, X_3\}$ ，可以將這兩個集合結合起來成爲一個新的集合 $C=\{+, \times, \sin, \exp, X_1, X_2, X_3\}$ ，則可將端點集合中每個變數的 $z(t_i)$ 視爲 0，因爲端點集合中的變

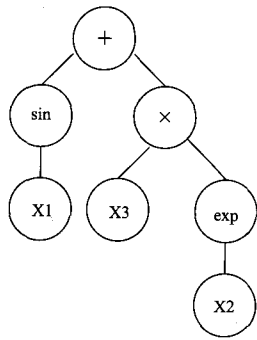


圖 9 $\sin X1 + X3 \times \exp(X2)$ 函數在遺傳規劃法的編碼方式

數一旦被選中就無法繼續向下延伸串連更多的基因。所以 C 集中每個原件分別的 $z(c_i)$ 可以視為 2,2,1,1,0,0,0, 也就是分別表示他們可以向下延伸串連基因的個數。

假設我們現在想要創造一個「 $\sin X1 + X3 \times \exp(X2)$ 」的函數式，這個函數式在遺傳規劃法裡的編碼方式如圖 9 所示，遺傳規劃法裡利用樹枝狀的結構來建構出所有可能的函數式，其中我們把樹枝結構狀的最頂層的節點稱為「根」(+)，樹枝狀結構裡面的節點稱為「枝」(sin, ×, exp)，而樹枝狀結構最底層的節點稱為「葉」(X1, X2, X3)。若將圖 9 所代表的函數式予以文字化，則可以表示為“(sin X1) (× X3 (exp X2))”，這是一般在程式設計者在建構遺傳規劃法時，其表達樹狀結構型編碼的轉換方式。由遺傳規劃法的編碼方式看來，其編碼方式跟傳統的遺傳演算法比較起來有很大的不同，在遺傳規劃法裡處理的是具有階層的染色體，而在傳統的遺傳演算法裡則是處理固定長度單一維度的染色體。

初始解(Initial solution)

在遺傳規劃法裡，所謂的初始解，就是一群具有上述樹枝狀結構的解。其初始解的決定是隨機去挑選樹枝狀結構的「根」、「枝」及「葉」，其產生的步驟如下。

第一步要決定的是樹枝狀結構的「根」，所以我們先隨機在函數集合 F 中找一個函數，我們把「根」限制在 F 集合中是因為希望在初始解都能擁有階層的樹枝狀結構，若們從端點集合 T 或



圖 10 樹狀結構「根」的決定

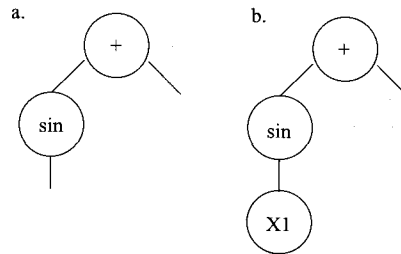


圖 11 樹狀結構的延伸

聯合集合 C 中去找到了變數來作為「根」，則此初始解就會退化成一個值而已。如圖 10 所示，其表示的為我們在 F 集合中找到了「+」作為此樹狀結構的「根」，而因為「+」有向下延伸串連兩個基因的機會(即 $z(“+”)=2$)，所以隨著「+」的選取亦提供了在此基因之後必須連接兩個基因的資訊。

每當樹狀結構中的一個節點選取自 F 集合，則就會有 $z(f)$ 條線從該節點延伸下去，而每條線的末端就可從聯合集合 C 中隨機選取一個函數或是變數。假如線的末端從 C 集中選取的為一個函數，則重複上述的步驟。舉例來說，如圖 11a 所示，由「根」所延伸出來的第一條線從 C 集中選取了函數 sin，則因為 $z(“sin”)=1$ ，所以從節點「sin」底下尚可延伸連接另外的節點。假如線的末端從 C 集中選取的為一個變數，則終止該節點的繼續向下發展。舉例來說，如圖 11b 所示，由「sin」節點所延伸的一條線從 C 集中選取到了變數 X1，則因為 $z(“X1”)=0$ ，所以從節點「X1」不會有任何線延伸出來，這個部分的發展就停終止。同樣的步驟一直「由上往下，由左往右」重複的進行，直到每條從節點延伸出來的線選取到 T 集中的原件為止。

適合度(Fitness)

適合度在遺傳演算法與遺傳規劃法中扮演的角色相同，適合度提供的是一個讓解往優的方向去的驅動力。在自然界中，生物的適合度被用

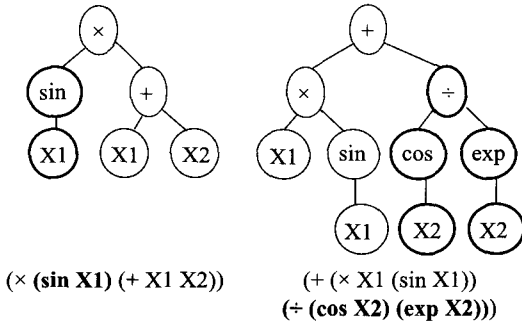


圖 12 選取預交配之親代

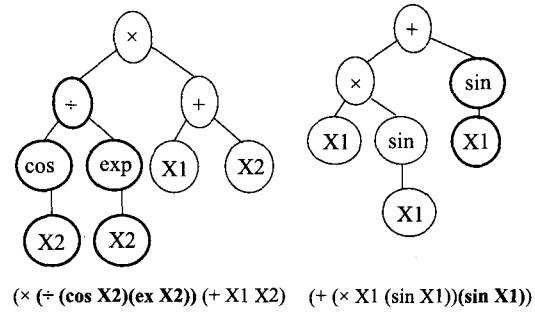


圖 13 交配產生之子代

來表示他可以存活到繁殖年紀跟真正繁殖下一代的機率，所以這樣的機率往往被轉換用來表示其後代的數量；在遺傳規劃法裡，我們利用適合度的量測來決定解是否發生交配或繁殖去改變解的結構，以符合問題的需求。適合度量測的方法和自然界相同，那就是適合度越高則存在且存活的機率越高。

複製(Reproduction)

適合度是判斷親代優劣的依據，其目的是希望藉由挑選適合度較高的親代透過繁殖、交配的過程使得往後的世代能逐漸演化出更好的子代。而繁殖就是遺傳演算法中繁衍下一代最基本的演算步驟，也是達爾文適者生存說最具體的表現方式。繁殖屬於無性生殖，也就是說繁殖的過程中只需要一個親代，且經過繁殖過後亦只會產生一個子代。繁殖的過程包含兩個步驟，第一、根據適合度的優劣選取一個親代，且適合度越高的親代越有機會被挑選到；第二、將此親代不作任何的改變，完全複製成一個新的子代。

選取(Selection)

選取親代進行複製的方法很多，其中最常用的方法即為親代被挑選的機率與其適合度好壞成比例關係的方法，假設 $f(s_i(t))$ 代表親代 s_i 在第 t 世代之適合度，則親代 s_i 被挑選複製成爲子代的機率爲：

$$\frac{f(s_i(t))}{\sum_{i=1}^M f(s_i(t))}$$

在電腦運算的過程中，大部分的時間都是花在親代適合度的計算上，而複製是減少電腦計算

時間的一種方式，因為複製的過程是將親代不經由改變完完全全複製到下一代，所以在進行下一代的適合度運算時，就不需要去計算這些已知親代的適合度。舉例來說，若下一代的子代中有 10% 是經由複製所產生，則電腦運算時間就會減少約 10%，對於需要較多運算時間的問題而言，這是一個最直觀節省時間的方法。

交配 (Crossover)

交配屬於有性生殖，需要兩個親代藉由交換彼此的染色體以產生新的子代，交配是增加遺傳規劃法中解的變異之主要方法。交配需要兩個親代，每個親代選取的方式和選取預進行複製親代的方式相同。再從這兩個被選取準備進行交配的親代，隨機在每個親代中選取一個節點，即交配點準備交配，要注意的是這兩個被選取的親代通常不具有同樣的大小、形式，且每個親代的交配點也不同。兩個親代中交配點的以下的樹狀結構，就像準備被截斷的染色體一樣，將被互相交換而完成交配的程序。交配點以下的樹狀結構其實亦是一個完整的樹狀結構，且至少包含一個端點節點(terminal)。

舉例來說，我們選取了兩個如圖 12 的親代來準備進行交配，並分別選取了兩個親代的「sin」及「÷」節點作爲交配點，準備進行交換的染色體以黑色粗線的圖形表示，從圖中可以發現，交配點以下的樹狀結構其實都可以獨立運算，就像是一個層數較少的親代，這也是爲什麼兩親代進行交配後，樹狀結構仍可以作運算的原因。交配的結果如圖 13 所示，將兩親代部分的染色體交換將會得到兩個與親代結構不同之子代。而在文

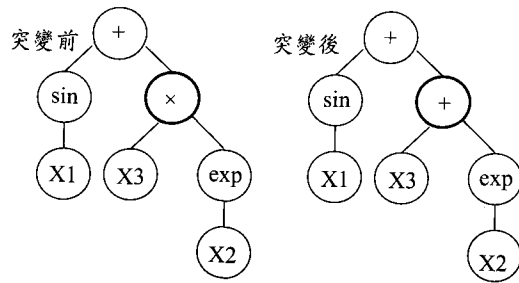


圖 14 點突變示意圖

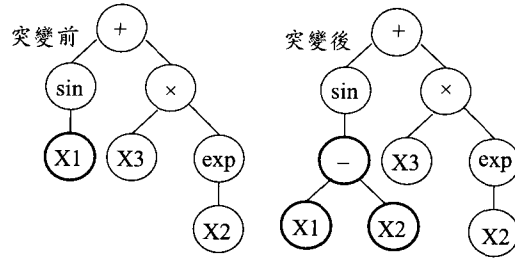


圖 15 面突變示意圖

字化編碼的交配方式上也相當容易，只要將相對於樹狀結構圖上交配點以下的染色體（即黑色粗線字體）直接交換即可。

因為被選取的親代並不會因此被排除在親代族群外，所以會有選取到兩個相同親代進行交配的機會，儘管如此，因為兩個親代的交配點為隨機選取，所以只要是交配點位置的選擇不同就會產生出兩個形式不同的子代。這對傳統的遺傳演算法而言是一種突破，因為在固定長度的帶狀編碼方式，通常是一個交配點應用在兩個親代上，所以交配出來的子代會和親代完全一樣。

突變 (Mutation)

所謂的突變，就是在染色體的結構上作隨機的改變。在遺傳演算法中突變有其必要性，因為解族群會有收斂的趨勢而落入區域最佳解中，突變正提供了一個跳脫區域最佳解的機制，藉由隨機選取改變染色體中的基因，將可以大大改變原有解的性質。突變也因此可以增加解的多樣性。與遺傳演算法相較，在遺傳規劃法中，突變功能不像在遺傳演算法中那麼重要，因為與遺傳演算法固定長度的編碼方式不同，遺傳規劃的編碼方式具有動態改變其形狀及大小的特性，經由交配的過程就可以增加解的變異，擴大搜尋的範圍，達到部分突變的效果，所以遺傳規劃法的解不易發生收斂的情形，但巧妙的突變往往可以增加求解的效率，所以儘管扮演的重要性不若在遺傳演算中，但仍廣泛被應用在求解演算的過程中。

發生突變時只需有一個樹狀結構參與，其發生的時機分為兩種，第一、親代發生突變，根據適合度的優劣選取一個親代進行突變成為子代；第二、子代發生突變，在兩親代交配所產生

的子代中，根據突變機率決定子代是否發生突變。不管是哪種發生的時機，突變都是在預備突變的樹狀結構中隨機選取一突變點準備進行突變，而突變的形式又分為兩種，一種為點的突變，也就是突變的過程中只改變染色體中的一個基因，若突變點為函數則突變為另一函數，若突變點為端點則突變為另一變數或常數；一種為面的突變，突變點下的樹狀結構都會在突變發生後產生改變，突變點隨機從函數集合及端點集合之聯合集中選擇一個突變後的節點，若為函數則繼續向下產生新的「枝」、「葉」，若為端點則完成突變。

舉例來說，如圖 14 所示，假設節點「×」為隨機選擇的突變點，則根據點突變的原則，因為「×」為函數，所以再任意從函數集中選取相異的函數取代，即完成點突變的過程；若為面的突變，如圖 15 所示，假設端點「X1」為隨機選取的突變點，則根據面突變的原則，從函數及端點集合的聯合集中選取任一節點，假設「-」符號被選擇，則此樹狀結構必須根據編碼的原則繼續往下作延伸，直到樹狀結構的末端為端點為止。

菁英策略 (Elitist)

菁英策略是複製之一種特例情形，是將一世代中最好的一個親代直接複製至子代族群中，除了可以在演算的過程中記錄每一世代的最佳解外，還能深入調整演算的行為，使得演算的過程更有效率。當然，依照使用者的需求，不一定要選取最好的一個親代，而可以選擇最好的若干個（如 5 個）親代執行菁英策略，直接複製到子代中。

表 1 遺傳規劃法各項參數設定

目標	找到一個足以描寫該分類下 (X_{1i}, X_{2i}, y_i) 間之關係式。
端點集合	X_1, X_2 (自變數)
函數集合	$+, -, \times, \div$
Raw fitness	即迴歸與實際的誤差值，是採用 Root mean square error
Parameters	
親代數量	500 個 (每個世代都維持 500 個親代)
初始解決定方式	初始解的大小、形狀均為隨機產生
運算世代數量	200 個世代
繁殖百分比	10% (親代中有 10% 是以繁殖的方式直接複製到子代中)
菁英策略	保留每個世代中最佳的解至下個世代
突變時機	親代交配後子代判斷是否發生突變
突變形式	點突變
突變機率	0.05 (交配過後的每 100 個子代中約有 5 個發生突變)
最大層數設定	6 層 (樹狀結構最多發展到 6 層，最多有 63 個節點)

註：Root mean square error (RMSE) = $\sqrt{\frac{\sum_{i=1}^n (y_{i,pred} - y_{i,obs})^2}{n}}$ ，其中 $y_{i,pred}$ 為樹狀結構在第 i 組輸入 (X_{1i}, X_{2i}) 時的預測值， $y_{i,obs}$ 為第 i 組資料的觀測值， n 為資料組數。

3.3 分析流程

本研究應用遺傳規劃法(Koza, 1992)進行優選時，是將資料分類後的各組分別以遺傳規劃法進行優選，也就是共要優選四次，每組資料進行優選的參數相同，各組各項參數設定如表 1 所示。

研究中的演算過程可以圖 16 來表示，在將欲優選的資料放入模式後，模式先依照分類的規則將資料分類，本研究共將資料分為四類，然後開始利用遺傳演算法將第一類資料進行優選，完成第一類資料的優選後再接續第二類資料的優選，如此重複進行演算，直到所有類別均已優選完畢。

四、分析結果

本研究模擬的魚群數量為 1987~2001 年，每年夏、秋兩季採樣兩次，介於七家灣溪一號壩至三號壩間之魚群數量，並將採樣前發生之日最大流量(cms)及平均氣溫(°C)作為影響魚群數量的兩環境因子。流量資料取得來自七家灣流量站，而氣溫則來自松茂站。模擬的結果分為四個類別來討論，並分別探討在不同類別下櫻花鉤吻鮭魚

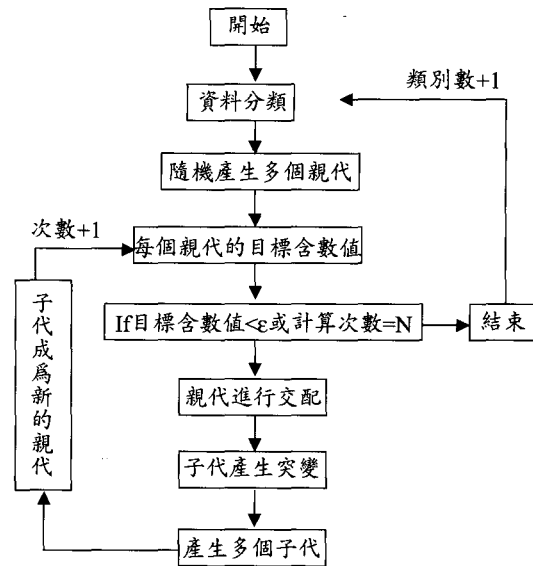


圖 16 模式演算流程圖

群數量與流量及溫度變化之關係。在「高溫度高流量」類別中，模式模擬的結果如圖 17 所示，在該分類中模式模擬魚群數量與實測魚群數量之誤差(RMSE)為 167 隻，且模擬魚群數量與實測魚群數量之相關係數(R)為 0.80。在「高溫度

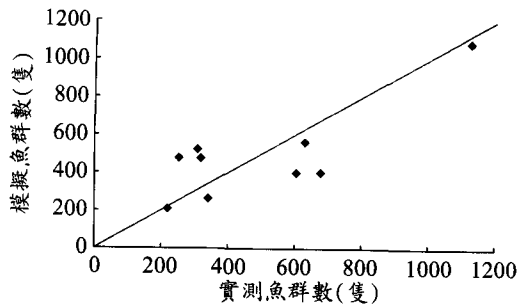


圖 17 「高溫度高流量」分類下，模擬魚群數與實測時群數之 XY 散佈圖

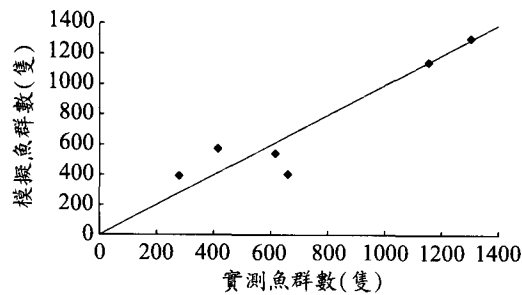


圖 18 「高溫度低流量」分類下，模擬魚群數與實測魚群數之 XY 散佈圖

低流量」類別中，模式模擬的結果如圖 18 所示，在該分類中模式預測魚群數量與實測魚群數量之誤差為 135 隻，模擬魚群數量與實測魚群數量之相關係數為 0.93。

其中圖 17、圖 18 中的斜直線表示百分之百正確模擬的直線，而從圖中可以看出，不管在那個分類下所模擬的魚群數均落在此斜直線的附近，這表示利用遺傳規劃法所產生的魚群數量與溫度及流量的函數關係可以描寫發生在「高溫度高流量」和「高溫度低流量」的魚群數量。

在「低溫度低流量」的分類中，模式模擬的結果如圖 19 所示，在該分類中模式預測魚群數量與實測魚群數量之誤差為 432 隻，模擬魚群數量與實測魚群數量之相關係數為 0.1。此分類下模式模擬的結果似乎不盡理想，原因是當環境條件屬於低溫度低流量時，魚群數量並未受到溫度及流量的明顯影響，圖 20 及圖 21 分別表示在「低溫度低流量」分類下魚群數量分佈與溫度及流量的關係，從圖 20 及圖 21 可以看出，在此分類條

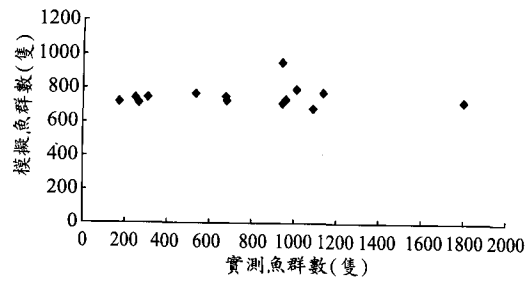


圖 19 「低溫度低流量」分類下，模擬魚群數與實測魚群數之 XY 散佈圖

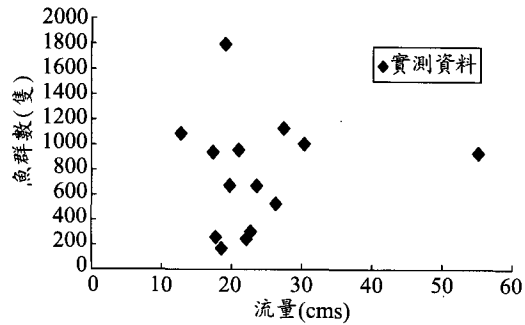


圖 20 「低溫度低流量」分類下，魚群數量與流量關係之 XY 散佈圖

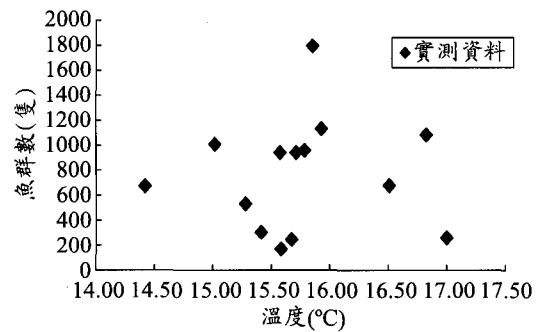


圖 21 「低溫度低流量」分類下，魚群數量與氣溫關係之 XY 散佈圖

件下的魚群數量並未與溫度及流量有明顯的函數關係，所以遺傳規劃法在進行優選時，僅優選出一條平均曲線來模擬該分類下的魚群數。

在「低溫度高流量」分類下，由於在台灣地區高流量的發生集中在豐水期，也就是 5~10 月，在這段時間內正值春、夏兩季，平均溫度均高於模式中設定之門檻值(18°C)，所以高流量發

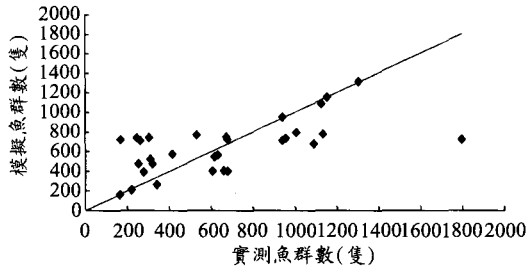


圖 22 所有資料整合下，模擬魚群數量與實測魚群數量之 XY 散佈圖

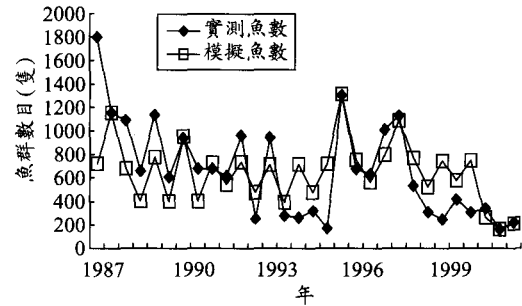


圖 23 模擬魚群數與實測魚群數隨時間變化之關係圖

生的時期溫度均達高溫度的標準；同樣地，低溫度大多發生在台灣地區的枯水期，以就是 11~4 月，此時氣溫低，流量亦低，所以低溫度及高流量幾乎不同時發生，在這樣的情形下，幾乎沒有資料點被分類到該類，所以無法進行優選。

若將所有分類過後的資料整合來看，如圖 22 及圖 23 所示，圖 22 表示在所有資料整合下，模擬魚群數量與實測魚群數量之 XY 散佈圖，圖 23 則表示，在每個時間點上模擬魚群數量與實測魚群數的關係，整體來看其誤差值為 315 隻，相關係數為 0.62，但如上述所言，所有的模擬誤差幾乎來自於「低溫度低流量」分類下的情形，若將該分類排除在外，則模式模擬的誤差及相關係數分別為 155 隻和 0.89。各分類下模式模擬的情形整理如表 2 所示。

五、模式應用

在合理建構出一套櫻花鉤吻鮭數量與環境因子間關係的模式後，本研究建立的 GP 模式將

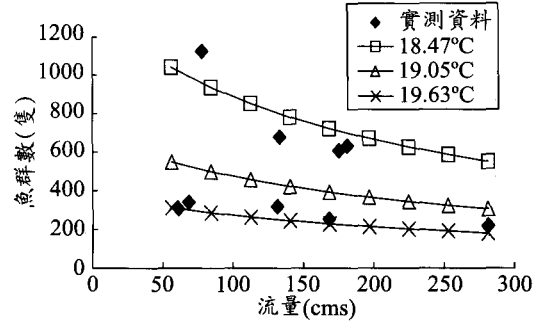


圖 24 「高溫度高流量」分類下，溫度及流量變化對魚群數量的影響(1)

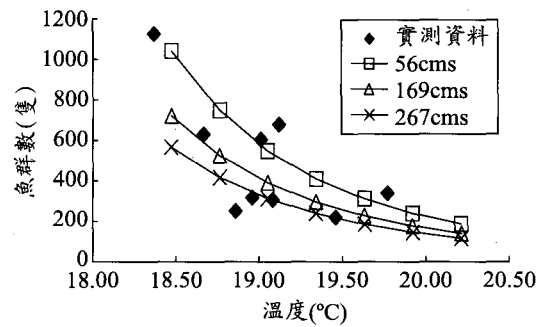


圖 25 「高溫度高流量」分類下，溫度及流量變化對魚群數量的影響(2)

表 2 各分類下遺傳規劃法模擬魚群數量之結果

類別	溫度	流量	RMSE	R
Category 1	高	高	167	0.80
Category 2	高	低	135	0.93
Category 3	低	高	Null	Null
Category 4	低	低	432	0.1
Total			315	0.62
Category 1 & 2	高	高/低	155	0.89

可應用於預測環境變遷對櫻花鉤吻鮭族群的影響，在本章節先透過敏感度分析的動作來了解不同分類下溫度及流量的變化會對魚群數量造成怎樣的影響。

在「高溫度高流量」的分類下，氣溫及流量變化對魚群數量的影響如圖 24 及圖 25 所示，兩圖中黑點分別表示實測魚群數量與流量及氣溫的關係，而實線則表示 GP 模式在此分類下模擬的結果，每一條實線分別是在固定的氣溫或流量

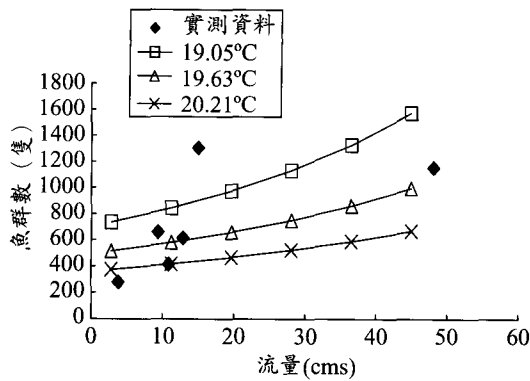


圖 26 「高溫度低流量」分類下，溫度及流量變化對魚群數量的影響(1)

下模擬的結果。從圖 24 的結果中我們發現，在固定的氣溫下流量越大魚群數量有減少的趨勢，且在固定的流量下氣溫越高魚群數量亦有少的趨勢；從圖 25 的結果中我們也發現，在固定的流量下氣溫越高魚群的數量越少，且在固定的氣溫下流量越大魚群數量有減少的趨勢。這樣的模擬結果與實際觀測的情形相當吻合，也顯示在高溫高流量的情形下的確對櫻花鉤吻鮭的生存造成一定程度的威脅。

在「高溫度低流量」的分類下，氣溫及流量變化對魚群數量的影響如圖 26 及圖 27 所示，從圖 26 的結果中可以看出，在固定的氣溫下流量增加魚群數量有增加的趨勢，而在固定的流量下氣溫增加魚群數量有減少的趨勢；從圖 27 的結果中亦可以得到相同的結論。在此分類環境條件下，固定溫度下魚群數量隨流量增加而遞增的原因是，因為在此分類下的流量並未對魚群的生存造成衝擊，但氣溫則不然，固定氣溫下流量越高相對而言水溫就越低，因為較大的流量使得水溫的反應趨於平緩，所以說在氣候變遷的環境條件下，過低的流量及高溫將會是另一個威脅魚群生存的重要環境因素，這是以往從未見人提起的。

近年來相關的研究顯示，未來氣候變遷的發生於櫻花鉤吻鮭活動的棲地將會有氣溫升高、豐水期流量增高、枯水期流量減少的趨勢（童慶斌與李宗祐，2001），也是就是說未來將會趨於「高溫度高流量」及「高溫度低流量」的環境條

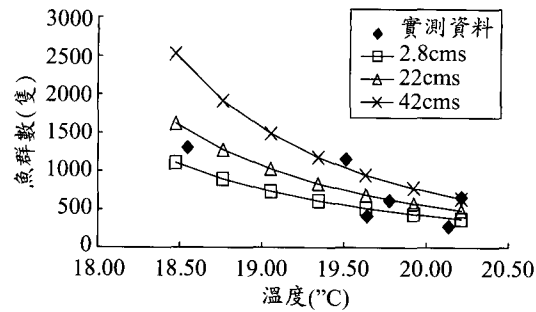


圖 27 「高溫度低流量」分類下，溫度及流量變化對魚群數量的影響(2)

件，從上述的分析結果看來，氣候變遷下的環境條件將會對櫻花鉤吻鮭造成極大的衝擊，當然本研究預測之魚群數量仍具有不確定性，不過已可合理提供環境變遷下魚群數量可能發生之趨勢，從預測的結果來看讓人不得不憂心在未來氣候變遷下櫻花鉤吻鮭所面臨之威脅。

六、結論與建議

從上述的研究方法及結果來看，利用遺傳規劃法所建立的模式的確提供了一個較新的方法來模擬魚群數量，儘管模擬上仍有些許誤差，但就像之前所提到的，魚群數量受到環境中多個因素複雜的影響，本研究中僅以兩個因子來模擬魚群數量實在是略顯不足，未來將進一步瞭解哪些環境會影響魚群的數量，並利用遺傳規劃法建構出魚群數量與環境因子間的關係。研究中分類環境因子大小門檻值的設定是以魚群資料對應不同環境因子分佈的情形來決定，未來則可以利用敏感度分析判斷門檻值選擇的優劣或是直接以優選的方式來決定。在合理建構出一套櫻花鉤吻鮭數量與環境因子間關係的模式後，將可以應用於評估環境變遷下魚群數量改變的情形，透過模式的建立，我們亦可以瞭解在何種環境條件下不適於櫻花鉤吻鮭的生長，再透過適當的工程方法改善其棲地環境，使得櫻花鉤吻鮭得以永續地悠遊在七家灣溪的上游。

參考文獻

1. Bradford, M. J., and Irvine, J. R., "Land use,

- fishing, climate change, and the decline of Thompson River, British Columbia, coho salmon,” *Can. J. Fish. Aquat. Sci.* 57: 13-16, 2000.
2. Cattaneo, F., Lamouroux, N., Breil, P., and Capra, H., “The influence of hydrological and biotic process on brown trout (*Salmo trutta*) population dynamics,” *Can. J. Fish. Aquat. Sci.* 59: 12-22, 2002.
 3. Chen D. G., Hargreaves N. B., Ware D. M., and Liu Y., “A fuzzy logic model with genetic algorithm for analyzing fish stock-recruitment relationships,” *Can. J. Fish. Aquat. Sci.* 57: 1878-1887, 2000.
 4. Koza, John R., *Genetic Programming: On the Programming of Computers by Means of Natural Selection*, Cambridge, Mass.: MIT Press., 1992.
 5. Mohseni O., Stefan H. G., “Stream temperature/air temperature relationship: a physical interpretation,” *Journal of Hydrology* 218: 128-141, 1999.
 6. Shaw, E. Al and Richardson, John S., “Direct and indirect effects of sediment pulse duration on stream invertebrate assemblages and rainbow trout growth and survival,” *Can. J. Fish. Aquat. Sci.* 58: 2213-2221, 2001.
 7. 汪靜明，子遺的國寶-台灣櫻花鉤吻鮭專輯，內政部營建署雪霸國家公園管理處，1994。
 8. 曾晴賢，櫻花鉤吻鮭族群監測與生態調查(一)，內政部營建署雪霸國家公園管理處，1998。
 9. 曾晴賢，櫻花鉤吻鮭族群監測與生態調查(二)，內政部營建署雪霸國家公園管理處，1999。
 10. 曾晴賢，櫻花鉤吻鮭族群監測與生態調查(三)，內政部營建署雪霸國家公園管理處，2000。
 11. 曾晴賢，櫻花鉤吻鮭族群監測與生態調查(四)，內政部營建署雪霸國家公園管理處，2001。
 12. 童慶斌，李宗祐，氣候變遷對七家灣溪流量之影響評估，*農業工程學報*，第 47 卷第 1 期，65~74 頁，2001。
 13. 楊正雄，1997。水溫對櫻花鉤吻鮭族群之影響。國立清華大學生命科學研究所碩士論文，1997。

收稿日期：民國 92 年 5 月 2 日
修正日期：民國 92 年 6 月 17 日
接受日期：民國 92 年 6 月 24 日