

# 短期列車旅運需求之類神經網路預測模式 建構與評估<sup>1</sup>

## ARTIFICIAL NEURAL NETWORKS FOR SHORT-TERM RAILWAY PASSENGER DEMAND FORECASTING

蔡宗憲 Tzung-Hsien Tsai<sup>2</sup>

李治綱 Chi-Kang Lee<sup>3</sup>

魏健宏 Chien-Hung Wei<sup>4</sup>

(94 年 9 月 13 日收稿，94 年 12 月 6 日第一次修改，95 年 3 月 8 日  
第二次修改，95 年 11 月 30 日定稿)

### 摘要

短期列車旅運需求預測模式可以提供軌道運輸營運者短期旅運資訊，有益於短期營運規劃之設計。本研究以類神經網路為基礎，建構短期列車旅運需求預測模式，探討三個模式構建課題：模式輸入變數設計對預測績效之影響、類神經網路模式與其他方法之績效比較、綜合模式 (combined model) 對預測績效之影響。我們蒐集臺鐵實際售票紀錄來進行模式構建以及驗證，有以下發現：其一，不適當的變數選用會導致類神經網路預測績效惡化。其二，在預測績效表現上，類神經網路優於隨機模式、去季節化隨機模式以及移動

- 
1. 感謝國科會專題研究計畫之財務補助，並感謝臺灣鐵路管理局提供之票務資料。
  2. 國立成功大學交通管理科學系博士 (E-mail: thtsai@mail.tcm.ncku.edu.tw)。
  3. 南台科技大學行銷與流通管理系教授 (聯絡地址：71005 臺南縣永康市南台街一號南台科技大學行銷與流通管理系；電話 06-2533131 轉 8108；E-mail: leeck@mail.stut.edu.tw)。
  4. 國立成功大學交通管理科學系教授 (聯絡地址：701 臺南市大學路 1 號成功大學交通管理科學系，電話 06-2757575 轉 53233，E-mail: louiswei@mail.ncku.edu.tw)。

平均模式，但與指數平滑法相近。其三，綜合模式績效優於個別預測模式，惟必須挑選較佳的個別模式來建立綜合模式。

**關鍵詞：**短期旅運量預測；類神經網路；預測模式比較；綜合模式

## ABSTRACT

*Short-term railway passenger demand forecasting can offer essential information to benefit short-term operational planning. This study constructed short-term forecasting models for railway passenger demand and discusses three modeling issues: the effects of input design on forecasting performance, validity of artificial neural networks and validity of combined models. We collected data from Taiwan Railway Administration for model construction and validation. Three findings were obtained. First, inappropriate design or use of input variables may result in unsatisfactory forecasting performance. Second, Artificial Neural Networks outperform random walk model, deseasonalized random walk model and moving average model, but have similar performance to exponential smoothing model. Third, combined models outperform individual models. However, candidates should be carefully selected for combining.*

**Key Words:** Short-term forecasting, Artificial neural networks, Model comparison, Combined model

## 一、緒 論

### 1.1 短期列車需求預測之價值

短期列車旅運需求預測可以提供鐵路營運者短期旅運需求資訊，以擬定適當的短期輸運計畫來達成營運目標。進一步從營收管理系統的角度來分析，短期預測為一重要前端輸入因子，諸如列車艙位配置以及票價因素，若能在準確的需求預測資訊下作出正確決策，對於營收會有顯著提升。Kimes<sup>[1]</sup>指出航空業、旅館業以及租車業在實施營收管理後，可以顯著提升 3% 至 5% 的營收，由此可顯示短期規劃的效益。Lee<sup>[2]</sup>在其模擬實驗中則得到若能提升營收管理系統高需求航班 10% 之預測準確性，則大約能提升航空公司整體 0.5% 至 3% 營收的結論，由此顯示預測系統在營收管理的重要性。因此如能建立一個準確之預測系統，對於後續營收管理系統之建置將有正面助益。其他短期營運規劃諸如列車排班以及相對應的行車人員排輪班，也都可由短期預測結果來作適當調整。

本研究旨在建立軌道運輸短期列車旅運需求預測模式，並以臺鐵長途市場每日旅運量為實證分析對象。短期預測問題對於臺鐵之所以重要，主要是因為臺鐵目前對於各班次所服務起迄點市場的座位分配，採取無彈性的上限額度配置方式。額度配置在一年內只有數次整體額度上的分配調整，並不會隨著星期特性不同而有所變動。此一機制將無法有效針

對每日旅運需求作出有效之動態回應。因此一旦長途旅客購票需求達售票上限，臺鐵則拒絕旅客之購票要求。此時長途旅客可能轉移至其他運具、購買鄰近班次或是購買無座票。倘若此時該班車上若干區段仍有空位，可以組合成一個長途旅客所要求起迄點座位，則臺鐵將因不適當決策導致有營收損失的風險。因此若能根據需求狀況動態調整每班車所服務起迄點市場的購票上限額度，不僅可以減少空位現象提高臺鐵營收，也可以滿足長途旅客要求，提升顧客滿意度。除此之外，也能根據需求尖離峰特性來擬定不同的長途市場座位銷售策略。為了達到此等目標，短期需求預測就很重要。

## 1.2 研究範圍

目前臺鐵服務西部幹線每日南下北上班次約 110 班次，每個班次所服務之起迄點隨車種以及服務屬性有所不同，除少數特定班次外，大多會超過 15 個甚至更多，因此針對整個系統所需要預測之模式數量就很龐大。在此本研究先從單一班次、單一個起迄點的需求預測模式進行探討，因為當此模式建立之後，其他不同起迄點及班次的預測模式比較容易進行擴展。本研究選擇 1008 班次自強號，高雄至北部配票區（包含松山、臺北、板橋、桃園、中壢）的每日旅運需求作為研究對象，來進行預測模式構建。選擇自強號是因為其占臺鐵整體營收約 30%~40%，是最受歡迎的車種；選擇高雄至北部配票區這個起迄點市場，是因為其包含臺灣前兩大都市，長途城際旅運需求相較其他起迄點市場高；選擇 1008 班次是因為其早上 8:00 由高雄出發，抵達終點松山為 12:42，時間上對於旅客較有吸引力。這個需求模式所預測之結果為 1008 班次高雄至北部配票區未來每日的搭乘人數（旅運量），屬於單期預測的範疇（one-step ahead forecast）。

本研究在此也假設研究資料未來的特性會延續過去發生之特性，以趨勢分析的概念來進行預測模式構建。因為針對預測方法與預測課題間的關係，我們可以隨著資料特性之轉變來選擇不同的預測方法（Armstrong<sup>[3]</sup>）。例如若研究課題為一新的數列，亦即無歷史資料可供參考，則可以採用市場調查法，或是尋找類似性質的標的來進行分析。若預期未來資料趨勢型態將產生顯著變化，則應進一步在預測模式中考量專家意見，或從歷史資料中找尋類似效果並反映在模式之中。若預期未來資料型態將持續過去資料型態的特性，則本研究所採用之趨勢分析法較為適宜。換言之，針對研究資料我們可以根據其特性，採用不同的預測方法，在此我們先從趨勢分析的角度切入。

## 1.3 本研究之課題

針對預測模式構建與評析之相關議題，本研究將進行以下三個課題的探討：其一，我們選用類神經網路作為研究方法，並探討該如何針對資料特性設計適合的變數，讓網路能夠有效學習到歷史旅運量的趨勢；此一課題之重要性，在於主觀認定的輸入變數型態，並不一定能夠讓網路學習到資料特性，而必須透過適當的變數設計。至於為何選用類神經網路作為研究方法，將於本文第二節進行討論。其二，本研究所發展的模式是否具有效力，

必須與其他預測方法相比較並觀察其是否能在預測績效 (out-of-sample forecasting) 上獲得較好的結果；此一課題之重要性，在於除了表達驗證效力的要求外，更可以觀察資料特性、預測方法與預測績效三者之間的關係。其三，由於在第二個研究課題上本研究會構建數個不同類型的預測模式，因此本研究第三個研究課題，是參酌不同模式預測結果進行結合，以建立綜合模式 (combined model)，並探討綜合模式是否會比單一模式績效結果更好；此一課題之重要性，主要在於其具有顯著提升模式預測績效的潛力。

針對此三個課題，本研究進行下列的討論：其一，我們以逐步篩選 (stepwise selection) 的方式嘗試幾種不同的變數設定方式，觀察每一解釋變數對類神經網路模式績效的影響。其二，我們與隨機模式 (random walk)、去季節化隨機模式 (deseasonalized random walk)、星期特性移動平均模式以及 Holt-Winters 指數平滑模式比較預測績效，以驗證類神經網路模式效力。其三，我們將按所有模式最後之績效表現建立綜合模式，並觀察績效與個別預測模式的差異。

在章節安排方面，第二節將簡述文獻上對於研究方法以及模式輸入變數選擇的討論。第三節將透過敘述統計及圖形繪製來觀察資料特性，這些特性將為後續模式輸入變數選擇的參考依據。第四節將概略介紹所使用的研究方法以及績效指標。第五節以蒐集而來的資料進行實證研究分析，並檢核各個研究課題之預測績效。第六節則綜合本研究所得的結論。

## 二、文獻回顧

方法選擇與輸入變數的使用，為構建預測模式重要的兩個決策因子，本節針對研究方法以及輸入變數型態進行文獻上之回顧，並進而決定本研究在此兩重要因子上的選擇。

### 2.1 預測方法

針對趨勢分析預測方法，常見的幾種方法有 (去季節化) 隨機模式、移動平均法、指數平滑法、時間序列模式、無母數迴歸、模糊預測、灰色預測以及類神經網路等。Grubb 與 Mason<sup>[4]</sup>修正 Holt-Winters 指數平滑法來預測英國航空市場的旅運量，獲得良好的績效結果。Williams<sup>[5]</sup>等人利用 SARIMA 時間數列構建高速公路交通量預測模式，發現該模式預測績效優於無母數迴歸、指數平滑模式以及類神經網路模式。Williams 與 Hoel<sup>[6]</sup>更進一步比較 SARIMA 時間數列與隨機模式、指數平滑模式及其變形於高速公路旅運量預測模式上之表現，結論也發現 SARIMA 在預測績效上占有優勢。Smith<sup>[7]</sup>等人以無母數迴歸理論為基礎，發展多種啟發式的權重系統來改善預測績效，並與 SARIMA 模式比較，雖然啟發式權重系統能改進傳統無母數迴歸方法之預測績效，然最終卻無法優於 SARIMA 模式。Clark<sup>[8]</sup>則將無母數迴歸的概念由單變量擴展至多變量，以找尋車流、速度及占有率之間的關係來預測快速道路的交通狀況，發現該方法在車流以及占有率的預測上較隨機

模式佳。許巧鶯與溫裕弘<sup>[9]</sup>使用灰色預測的概念建構航空旅運量預測模式，結果發現灰色預測模式優於 ARIMA 以及乘幂迴歸模式。Sfetsos<sup>[10]</sup>使用整合模糊理論與類神經網路模式進行風速預測，並與數種線性及非線性結構比較，發現該整合模式績效最佳。而有關類神經網路於短期交通量之預測也有很多研究成果，諸如 Smith 與 Demetsky<sup>[11]</sup>、Hill<sup>[12]</sup> 等人、Dharia 與 Adeli<sup>[13]</sup>、Ishak<sup>[14]</sup>、魏健宏等人<sup>[15]</sup>都發現類神經網路於短期交通量預測上具有潛力。

文獻上尚無法根據資料特性來保證最佳之預測模式 (Prudêncio 與 Ludermit<sup>[16]</sup>)，相關這方面預測模式的選擇仍受到廣泛討論，例如知名的 M-Competition 嘗試建立一個比較平台供數種不同模式結構加以比較 (Makridakis<sup>[17]</sup> 等人；Fields 等人<sup>[18]</sup>；Makridakis 與 Hibon<sup>[19]</sup>)，這一類研究雖能透過大規模比較獲得一般性之結論，然卻無法獲得有效且一致的證據，顯示資料特性與模式選擇之間的關係。有些研究建立模式選擇的模型來選擇適合資料特性的模式以進行後續預測工作 (Shah<sup>[20]</sup>；Prudêncio 與 Ludermit<sup>[16]</sup>)，這類研究可以透過分類模式有效地顯示資料特性與模式選擇的關係，然其缺點為在真正進行預測工作之前，必須大規模地針對數列資料嘗試各種不同的預測模式，以獲得「最佳模式」結果來建立分類模式，因此在模式構建成本上較為昂貴。除此之外，更多的實證研究必須被建立以用來支持此類模式。目前常用來驗證個別研究所發展模式效力的方法，為必須與其他預測方法在預測績效上作一比較 (Adya 與 Collopy<sup>[21]</sup>)。

## 2.2 輸入變數之選擇

在輸入變數方面，文獻上進行趨勢分析通常由時間數列 (time series) 的角度切入，利用歷史資料當成解釋變數來構建預測模式 (Smith 與 Demetsky<sup>[11]</sup>；Hill 等人<sup>[12]</sup>；Ledoux<sup>[22]</sup>；Faraway 與 Chatfield<sup>[23]</sup>；Smith 等人<sup>[7]</sup>)。然由於歷史資料的使用可能造成過多的輸入變數，因此可利用不同的方法來進行變數篩選。例如 Lachtermacher 與 Fuller<sup>[24]</sup>觀察 Box-Jenkins 模式的最終結果來選擇延滯變數的期數；Zhang<sup>[25]</sup>計算自我相關係數 (autocorrelation coefficient) 及偏自我相關係數 (partial autocorrelation coefficient)，作為選擇延滯變數的參考依據；Ishak 與 Alecsandru<sup>[26]</sup>利用主成份分析的方法來去除不重要之變數；Vlahogianni 等人<sup>[27]</sup>更進一步在一個模式之中考量不同延滯期數對績效的實際影響，使用遺傳演繹法來進行變數篩選；Piramuthu 等人<sup>[28]</sup>則以決策樹建立可能之變數組合，以進行後續模式構建。另外一個較少被討論到的角度為使用時間特性變數 (temporal variables)來構建預測模式，這些變數之選擇，主要來自於先驗知識 (domain knowledge) 以及對資料特性之觀察。Nam 與 Schaefer<sup>[29]</sup>使用 11 個虛擬變數來表示月份效果及一個時間連續變數來表示月份趨勢，來預測航空公司每月旅客總量。Sun 等人<sup>[30]</sup>則更細部地同時預測每日各航班的旅運量，其使用時間特性虛擬變數如日期、星期等特性，以及航班屬性如起迄點、機型、停站數，以及目前預約狀況來進行預測模式之構建。

## 2.3 本研究之選擇

本研究根據 Vlahogianni 等人<sup>[27]</sup>的歸納結論，選擇以類神經網路作為模式發展基礎，因其發現類神經網路模式，在短期高速公路旅運量的預測上優於其他傳統方法。雖然該結論是在不同的個別研究上來加以比較，其基準並不相同 (Fildes 與 Lusk<sup>[31]</sup>)，然該歸納結論仍提供一個可行的嘗試方向。而本研究的研究對象與高速公路旅運量具有下述相似特性：其一，短期高速公路旅運量也為短期預測的範疇，資料觀測以分鐘為單位，所以資料的變異相較於總體資料來得大，和本研究以日為單位的資料波動特性較為接近；其二，高速公路旅運量也具有以日為長度的季節特性，這和先驗知識上鐵路旅運量以週為長度之季節特性相似；其三，短期高速公路旅運量趨勢並不明顯，這和鐵路旅運量的特性也類似。值得注意的是，高速公路旅運量之空間關係，較鐵路運輸各起迄對間的空間關係來的明顯且重要，然此一現象並不會影響類神經網路於研究課題上的潛力，因為不考慮空間關係實為考慮空間關係的一特例。

在輸入變數類型方面，本研究將先從時間特性變數的角度切入來進行模式構建。選用時間特性變數來構建預測模式，除了可以減輕對大量歷史資料變數做篩選的動作外，還具有兩個優點：其一為時間特性變數可以事先得知，因此不需有其他變數蒐集的動作。值得注意的是社經變數，如所得、車輛持有，以及列車屬性，如票價、服務水準，並不會對短期列車旅運需求提供解釋的效力。其二為時間特性變數較容易作多期預測，因為其不需要用到歷史資料的資訊。換言之，在未來每一日時間屬性都能事先掌握的情形下，要預測未來某一特定期數的旅運量是比較直觀的，並不需要以預測值再去進行未來多期預測的動作。

## 三、鐵路資料特性分析

### 3.1 旅運量資料分析

本研究蒐集 1999 年與 2000 年每日 (時間  $t$ ) 由高雄 (起點  $i$ ) 至北部配票區 (迄點  $j$ ) 搭乘臺鐵 1008 班次 (班次  $s$ ) 的旅運需求紀錄，以  $y_{ij}^s(t)$  表示。其中 1999 年資料被用來進行模式構建，2000 年資料被用來進行模式驗證 (validation)。必須注意的是，用來驗證的資料必須沒有在模式構建中被使用過，如此才具有驗證效力。

時間趨勢分布圖 (time plot) 常被用來觀察數值分布的特性，以作為後續預測模式構建的基礎。一般觀察時間數列圖形分析的三大重點為離群值 (outlier) 是否存在、時間上的特性諸如趨勢 (trend)、季節特性 (seasonality) 以及循環特性 (cycle) 為何、資料分布上是否產生結構變化 (Chatfield<sup>[32]</sup>)。下圖 1 為原始資料在 1999 與 2000 年相對應星期特性的旅運量分布，相關特性分述如下。

### 3.1.1 離群值分布

就離群值的角度來觀察，在本研究中我們將國定假日 (statutory festivals) 視為當然之離群值，此外也將一些極端分布的數值視為離群值。首先國定假日之所以被視為離群值，主要是因為國定假日歷史資料的旅運型態，並無法有效預測未來國定假日的旅運型態，主要之原因有二：其一，國定假日旅運量分布特性受到假期長度的影響而有所不同。其二，國定假日旅運型態更顯著地受到政府所提出交通輸運計畫所影響，致使型態上有所不同。圖 2 中顯示研究期間春節假期以及清明假期的旅運量分布圖，可以明顯看出型態上的差異。由此可知國定假日旅運量實屬模式無法控制，在本研究中我們將國定假日旅運量視為離群值並進行修正，將研究焦點集中在日常一般時間特性的旅運量數值。而從圖 1 中也可清楚發現幾個異常分布的數值，我們以虛線圈圍在圖 1 中註記以舉例表示。現階段臺鐵對於這些異常數值發生之原因並無紀錄可循，因此無法給予有效的解釋變數設定。而該如何有系統地定義出資料中的異常數值並進行修正，我們將於之後介紹。

有關國定假日以及極端值之數值則值得更進一步進行探討，例如針對國定假日，由於歷史資料所能提供之資訊有限，因此透過主觀經驗來修正模式預測值，預期可以獲得較好的預測績效值；針對離群值，則應透過蒐集可能影響因素，待預期未來類似事件可能發生時，再將歷史資訊透過模式或是人為介入之方式修正模式預測值。因此針對國定假日以及離群值，更適合的模式結構值得被進一步討論，惟此已超過本研究之範圍。

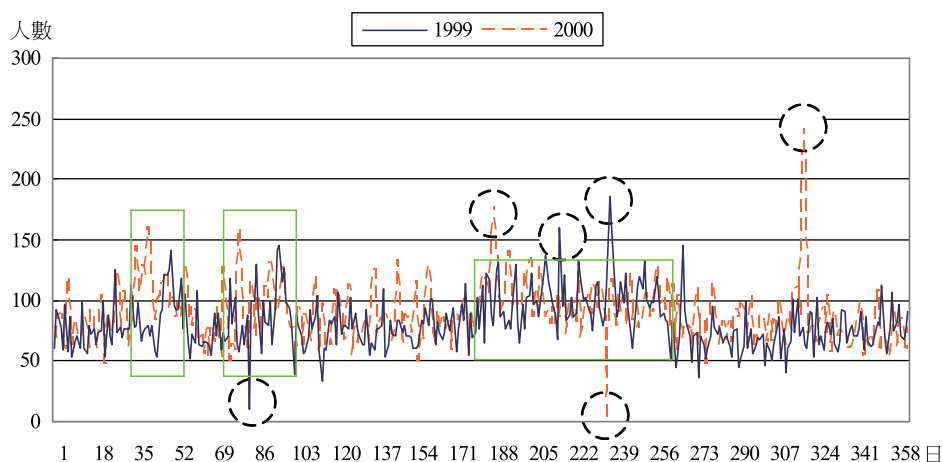


圖 1 研究期間原始旅運量分布圖

### 3.1.2 時間特性

就時間特性的角度來觀察，我們分別對趨勢、季節性以及循環的存在進行判斷。其一，由圖 1 分布可以發現旅運量呈現一種向平均數迴歸的力量 (Armstrong<sup>[3]</sup>)，而並無持續向上 (grow) 或向下 (decay) 之現象。由此可知在研究期間，旅運量的趨勢並不明顯。其二，

就先驗知識而言，旅運量存在著以週為長度的季節特性，例如對於研究資料而言，工作日型態和週末型態就有所不同。圖 3 顯示 1999 年自我相關係數的分布圖，該指標可以計算不同旅運量延滯期數間的相關性。由圖中可以看出延滯 7、14、21 期的相關係數很高(圖中畫圈之處)，顯示出週間樣式 (pattern) 具有顯著之關係，亦即以週為長度的季節特性。此外，延滯第 1 期與第 2 期的相關係數很高，顯示出資料的週內樣式也有很顯著之關係。值得注意的是，在研究期間臺灣實施隔週休二日的政策。其三，就先驗知識來看，循環之特性並不存於資料之中。循環和季節性同樣是重複出現類似的週期型態效果，兩者唯一不同為循環的週期長度並不定，因此很難像季節性一樣進行觀察，而循環最常出現在經濟資料之中。根據 Armstrong<sup>[3]</sup> 的看法，當資料或先驗知識中沒有強烈證據支持循環特性時，應該將該因素捨去，因此本研究將不在後續模式構建時放入循環因素的影響。除了上述三個時間上的特性之外，我們也可以觀察到影響資料分布的特性和資料平均值是呈現相依 (dependent) 的關係，亦即時間特性因子的影響並非一個定性關係。圖 4 顯示 1999 年 1 月份 (一般期間) 以及 8 月份 (暑假期間) 旅運量按星期特性劃分之平均值，可以發現兩者之間的差異並非圍繞在一固定常數。因此按照資料分解 (data decomposition) 的觀點，影響旅運量分布的數個特性之間，應是存在一乘方的關係 (multiplicative)，而非加方的關係。

### 3.1.3 結構性變化

就結構方面的角度來觀察，我們可以發現以下兩個特性：其一，根據臺鐵紀錄，在研究資料的期間，並無重大之政策會使得旅運量數值產生顯著的變化 (level change)，因此並無外在因素的影響導致資料分布結構產生不連續的情形 (discontinuity)。其二，這兩年的資料都存在兩個短暫尖峰，以及一個持續尖峰，如圖 1 中矩形所圍成區域。暫時尖峰分別表示寒假 (包含春節假期) 及清明假期，而持續之尖峰則表示暑假期間。由於本研究將國定假日視為當然離群值，因此我們將不討論春節假期以及清明假期的影響。而寒假期間縱使不考慮春節假期的影響，其旅運量預期也比一般日要來得高。換言之，寒暑假期的發生對於旅運量會產生平均值顯著的變化，而當寒暑假期一結束時，旅運量波動就又回復到原先平均值的幅度。此一現象也顯示資料具有不平穩 (non-stationary) 的特性，事實上此一現象也可由圖 3 自我相關係數下降緩慢的現象來加以定斷。就時間數列 ARIMA 模式的角度來看，不平穩之資料必須先經過差分來平穩資料，以便進行模式設定以及估計 (Chatfield<sup>[1]</sup>)。然對類神經網路而言平穩化資料並非一前提要件，因此我們保留使用原始資料，並透過找尋適當解釋變數來進行後續預測模式的構建。

### 3.2 離群值偵測與修正

文獻上對於離群值之認定可以分成兩個方向，第一個方向為在模式構建之前利用偵測的方式將離群值篩選出來，之後再利用剩餘之資料來構建模式。第二個方向為利用不受離群值影響之方法結構來進行模式構建，之後再利用模式的輸出結果來判定離群值的位置。



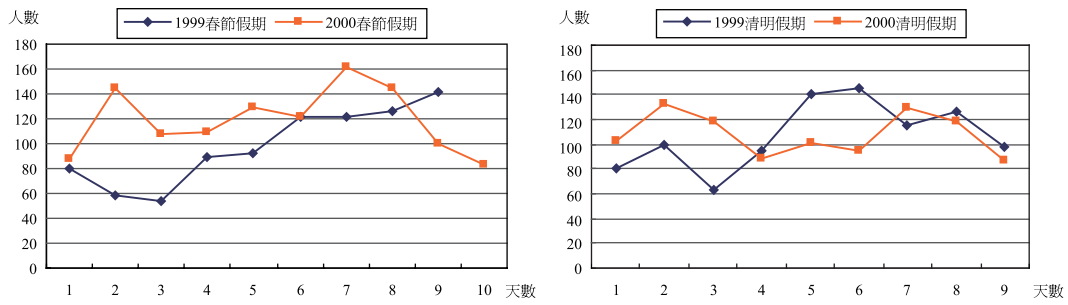


圖 2 (a) 研究期間春節假期旅運量分布

(b) 研究期間清明假期旅運量分布

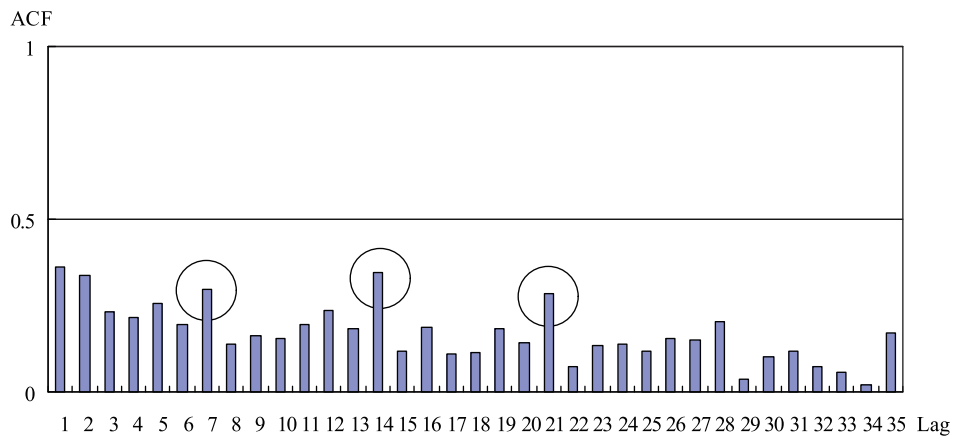


圖 3 1999 年旅運量自我相關係數分布圖

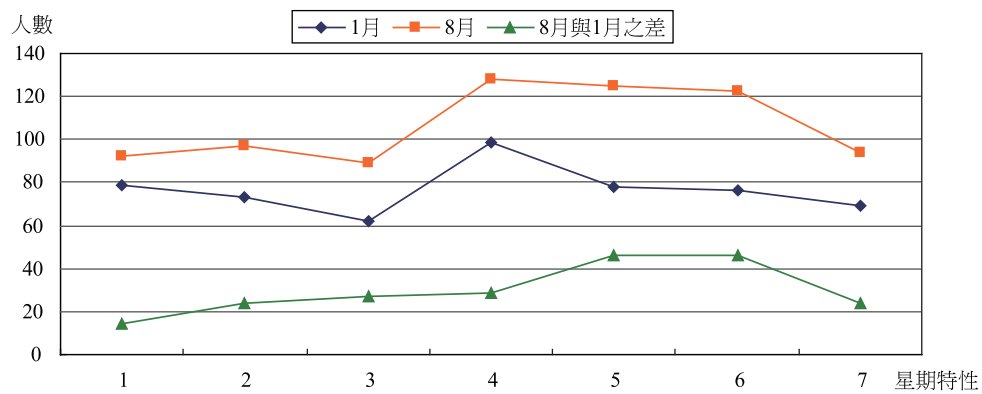


圖 4 1999 年 1 月份與 8 月份旅運量平均值分布

相關的離群值討論可以參考 Rousseeuw 與 Leroy<sup>[33]</sup>、Fildes 等人<sup>[18]</sup>、Liu 等人<sup>[34]</sup>、Kumar<sup>[35]</sup> 之研究。本研究採用第一個方向，亦即利用偵測的方式將離群值篩選出來，首先針對旅運量分布特性，建立一套離群值偵測以及篩選的機制，以避免離群值對於預測績效的扭曲。由於國定假日可以由行政院人事行政局的公告得知，因此可以根據這些資料先就國定假日進行標記。此外，觀察旅運量次數分布圖，如圖 5 所示，我們可以發現大體上兩年的分布形狀皆近似常態分布，此一步驟可更進一步確立旅運量極端值並非常態。由於離群值大部分在數列資料中都以極高或是極低的現象存在，因此我們直觀的採用平均數以及兩倍標準差所構成的範圍來進行離群值篩選（不含國定假日影響），因為在常態分配中，此一範圍可以涵蓋約 95% 的資料區間。而旅運數值超過此一範圍的樣本點被視為非常態之現象，以離群值視之來進行後續的資料處理。這裡值得注意的是，我們更進一步按照寒暑假的發生期間將一年分成五個區間，這是因為如果以一年為長度去進行偵測時，有可能會將寒暑假旅運量較高的現象視為離群值，因此劃分為年初至寒假前、寒假期間、寒假後至暑假前、暑假期間以及暑假至年底。各區間每一日的旅運量若有超過各區間上下限者則加以標記，圖 6 顯示 1999 年離群值判斷的界線。最後將所有被標記的日期按照星期特性（day-of-week），以各區間其他相同星期特性旅運量的平均值加以取代。後續資料特性以及模式構建就以修正後的旅運數值來進行分析。

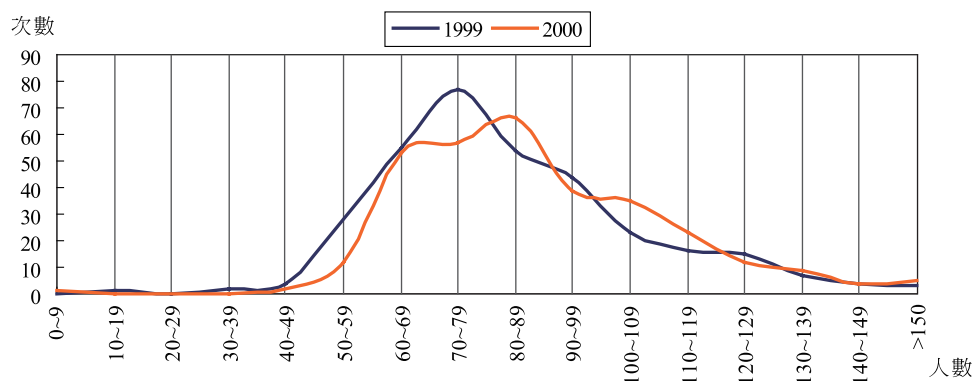


圖 5 研究期間旅運量次數分布圖

此一小節討論了研究資料的分布特性。由於資料上的限制，使得現階段無法對國定假日以及極端值的現象進行有效之處理，因此本研究根據資料特性，發展一個簡單的偵測及修正流程來除去這些影響，將焦點先著重在一般日旅運特性。在一般日旅運需求分布下，我們分析了數個資料可能存在的時間特性，就臺鐵資料而言，這些時間特性對其他班次預期同樣有效，惟重要性可能會略為不同，例如對某些班次而言寒暑假因素可能並不顯著。總而言之，本文接下來所討論的模式構建流程，對於其他鐵路班次預期同樣有效，並不會因為研究班次改變而使得本研究之結論產生重大變化。

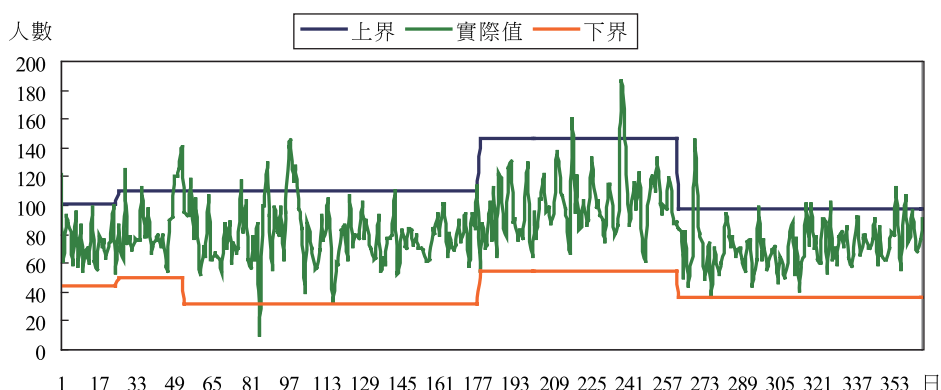


圖 6 1999 年離群值偵測區間

## 四、預測方法

此一小節介紹類神經網路以及四個比較模式的基本概念，此四種模式分別為隨機模式、去季節化隨機模式、星期特性模式以及指數平滑法之 Holt-Winters 模式，分述如後。此外，我們也論述綜合模式的基本概念以及本研究所使用的績效評估指標。

### 4.1 多層次前向式類神經網路 (multi-layer feed-forward neural network, MLFN)

此節論述本研究所採用之神經網路模式以及相關的模式設定。類神經網路主要起源於模仿人腦運作狀態，利用簡單神經元的連結與互動來處理複雜問題 (Haykin<sup>[36]</sup>)。適用於預測課題的網路結構屬於監督式類神經網路 (supervised network)，文獻上此類網路結構常見的有反傳遞網路 (Dharia 與 Adeli<sup>[13]</sup>)、半徑式神經網路 (Haykin<sup>[36]</sup>) 以及本研究擬使用之多層次前向式類神經網路 (MLFN)。多層次前向式類神經網路在文獻上廣泛被使用在各種課題上並獲致良好績效結果，大部分預測文獻都是以此網路結構為基礎去進行模式應用與發展。其他高階網路結構，例如 autoregressive neural network (Dougherty 與 Cobbett<sup>[37]</sup>；Tsai 等人<sup>[38]</sup>)、time-delay neural network (Lingras 與 Mountford<sup>[39]</sup>)、Elman/Jordan neural network (Ishak 等人<sup>[26]</sup>)、diagonal neural network (Ku 與 Lee<sup>[40]</sup>)，也都是以此網路結構為基礎模式來構建預測模式。圖 7 為一個典型的多層次前向式類神經網路架構圖，第一層為輸入層，主要在於接受輸入變數；第二層為隱藏層，主要進行資料的特徵擷取；第三層為輸出層，主要在於產生網路預測值。透過網路的學習功能，可以將輸入變數以及輸出變數之間的關係作一有效對應 (mapping)，用以提供後續預測之目的。下式(1)為 MLFN 的基礎數學架構，其中  $\hat{y}_{ij}^s(t+1)$  為未來一期旅運量預測值、 $y_{ij}^s(t+1)$  為未來一期旅運量實際值、 $w_{hl}(t)$  為第  $t$  期隱藏層與輸出層之權重連結 (本研究輸出值只有一個所以  $l=1$ )、 $w_{kh}(t)$  為第  $t$  期輸入層與隱藏層之權重連結、 $x_{kij}^s(t)$  為第  $t$  期第  $k$  個輸入變數、 $f(\cdot)$  為 Sigmoid 轉換函

數、 $\theta_h(t)$ 與 $\theta_l(t)$ 分別為隱藏層與輸出層第  $t$  期門檻值。網路首先由輸入層接收外部資訊  $x_{kij}^s(t)$ ，透過非線性轉換函數轉變成隱藏層之輸出  $f(\sum_k w_{kh}(t)x_{kij}^s(t) - \theta_h(t))$ ，之後再將資訊傳送至輸出層，透過非線性轉換而得到網路輸出值  $\hat{y}_{ij}^s(t+1)$ 。網路訓練目的主要在於求得最佳  $w_{kh}(t)$ 、 $w_{hl}(t)$ 、 $\theta_h(t)$ 與 $\theta_l(t)$ ，一般使用的目標函數為極小化預測誤差之平方  $(y_{ij}^s(t+1) - \hat{y}_{ij}^s(t+1))^2$ ，利用最陡坡降法之概念求取。

$$\hat{y}_{ij}^s(t+1) = f(\sum_h w_{hl}(t)f(\sum_k w_{kh}(t)x_{kij}^s(t) - \theta_h(t)) - \theta_l(t)) \quad (1)$$

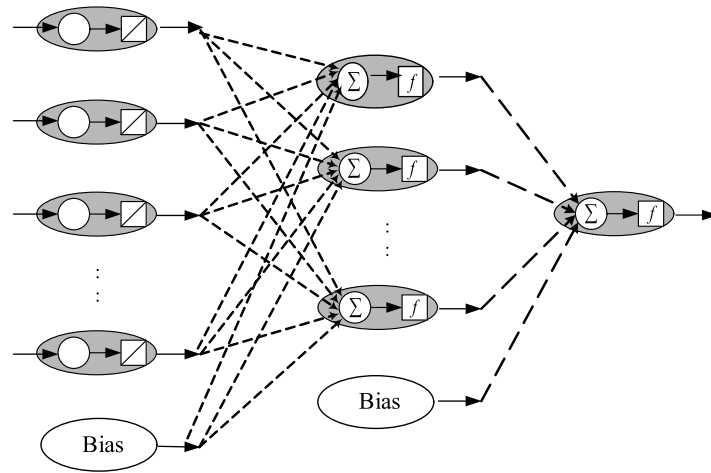


圖 7 多層次前向式類神經網路架構

網路進行訓練時，必須先對其結構參數進行設定，本研究參酌類似主題之文獻內容採取以下之設定：其一，隱藏層層數一層，而隱藏層神經元數目為輸入神經元與輸出神經元之算術平均數，此為一經驗值，目前文獻上對於最佳之隱藏層神經元個數仍無一定論。其二，隱藏層以及輸出層之非線性轉換函數採用 Sigmoid 函數，如式 (2) 所示。其三，網路權重求取之演算法採用倒傳遞網路演算法 (back-propagation learning algorithm)。其四，網路學習停止法採用 early-stop 訓練法則來避免過度學習 (overfitting/overtraining) 的情形。由於採用此法則，因此必須將模式構建的訓練資料（亦即 1999 年資料）再分成兩個子資料集。本研究從原始訓練資料中隨機抽取總數的 1/6，以決定學習停止時機，剩餘資料則用來訓練網路權重值。之後再以測試資料集（亦即 2000 年的資料）來驗證模式預測的績效。

$$f(x) = \frac{1}{1 + \exp^{-x}} \quad (2)$$

## 4.2 統計模式

除了發展以類神經網路為基礎之預測模式外，本研究也應用四個常見統計模式來構建預測模式作為比較之基礎，以辨別類神經網路模式是否具有較佳的預測績效。四個統計模式分別為隨機模式、去季節化隨機模式、移動平均模式以及指數平滑模式。此四種統計模式以及類神經網路模式簡單的特性比較如下表 1 所示，主要從模式複雜度、模式潛力以及參數多寡來作一概略比較，以下更進一步分別簡介這四種模式之概念。

表 1 不同預測方法之特性比較

| 預測方法     | 特 性                                                                                                                      |
|----------|--------------------------------------------------------------------------------------------------------------------------|
| 隨機模式     | <ol style="list-style-type: none"> <li>1. 假設簡單且易於了解與使用</li> <li>2. 可表現純隨機現象之資料</li> <li>3. 無需參數</li> </ol>               |
| 去季節化隨機模式 | <ol style="list-style-type: none"> <li>1. 假設簡單且易於了解與使用</li> <li>2. 可處理季節因素之影響</li> <li>3. 無需參數</li> </ol>                |
| 移動平均模式   | <ol style="list-style-type: none"> <li>1. 假設簡單且易於了解與使用</li> <li>2. 利用平均消除資料部分噪音</li> <li>3. 無需參數</li> </ol>              |
| 指數平滑法    | <ol style="list-style-type: none"> <li>1. 複雜度增加但可進一步考慮資料近期以及遠期影響</li> <li>2. 有效抓住季節 (或趨勢)</li> <li>3. 模式參數個數少</li> </ol> |
| 類神經網路    | <ol style="list-style-type: none"> <li>1. 模式較上述四種方法複雜</li> <li>2. 透過設計可以表現資料特性</li> <li>3. 模式參數個數多</li> </ol>            |

### 4.2.1 隨機模式 (random walk, RW)

隨機模式適用於資料甚為混亂的情形，其假設資料下一步走向為目前位置再加上一個隨機誤差項，就像一個喝醉的人走路一般。下一期預測值為本期之實際值，其預測式如下式 (3)，其中  $y_{ij}^s(t)$  為第  $t$  期的實際旅運量， $\hat{y}_{ij}^s(t+1)$  為第  $t+1$  期預測值。隨機模式常被用來處理不平穩 (non-stationary) 資料，常見於經濟及股市資料，常被用來作為模式比較基礎 (Makridakis 等人<sup>[41]</sup>；Williams 與 Hoel<sup>[6]</sup>)。

$$\hat{y}_{ij}^s(t+1) = y_{ij}^s(t) \quad (3)$$

#### 4.2.2 去季節化隨機模式 (deseasonalized random walk, DRW)

去季節化隨機模式與隨機模式基本精神相同，惟在模式中須先將資料作去季節化處理，之後進行隨機模式預測過程，最後再考慮季節化效果的影響以形成最終預測值。去季節化隨機模式的優點，是可以避免季節化因素與隨機因素產生交互作用而影響預測績效，一般也可用來作為模式比較基礎 (Makridakis 等人<sup>[41]</sup>)。為了方便說明，在此先將原先定義之  $y_{ij}^s(t)$ ，根據其星期特性  $k$  以及用來計算季節性指標所用的週期個數  $m$  改寫成  $y_{ij}^s(m, k)$ ，以下式 (4) 至式 (6) 為模式架構，其中， $y_{ij}^s(m, k)$  為去季節化後之旅運數值、 $S_{ij}^s(k)$  為星期特性  $k$  的季節指標。首先將原始資料透過季節性指標先進行去季節化之程序，如式 (4) 所示，先不考慮季節性因素對旅運量之影響。其次這些去季節化的資料可以再以一個隨機模式來加以反應，如式 (5) 所示。最後再利用季節性指標修正上一步驟所獲得之預測值，如式 (6) 所示。至於季節指標之求取，本研究採用 ratio-to-moving-average 的概念 (DeLurgio<sup>[42]</sup>)，其概念及演算流程如圖 8 所示，該算法主要為資料分解的概念，由一乘方關係式開始。而在資料分析時，我們發現研究資料影響因素之間呈現乘方關係，如 3.1.2 節所述，符合該算法之前提假設。在本研究中，由於星期特性為週一至週日，因此  $k$  最多可至 7，而用來計算的週期個數共有 52 週，因此  $m$  最多可至 52。

$$y_{ij}^s(m, k) = \frac{y_{ij}^s(m, k)}{S_{ij}^s(k)} \quad (4)$$

$$\hat{y}_{ij}^s(m, k+1) = y_{ij}^s(m, k) \quad (5)$$

$$\hat{y}_{ij}^s(m, k+1) = \hat{y}_{ij}^s(m, k+1) S_{ij}^s(k+1) \quad (6)$$

#### 4.2.3 星期特性移動平均模式 (moving average, MV)

移動平均法利用平均之概念來消除隨機噪音，此一方法假設未來的星期特性變動和過去一樣，利用模式構建階段資料求算每一個星期特性的平均值，用以當成預測未來每一個星期特性的預測值，概念如下式 (7)。其中， $y_{ij}^s(t|t \in k)$  為第  $t$  期的旅運量且其星期特性為  $k$ ， $n_k$  為星期特性  $k$  在 1999 年的個數， $\bar{y}_{ij}^s(t|t \in k)$  為星期特性  $k$  的平均值，也為未來星期特性  $k$  之預測值。此一方法雖然可以抓住簡單的歷史趨勢，但當資料之波動具有較複雜型態時，預測績效會受到影響。

$$\bar{y}_{ij}^s(t|t \in k) = \frac{\sum_{t=1}^{n_k} y_{ij}^s(t|t \in k)}{n_k}, \quad k \in \{1, 2, 3, 4, 5, 6, 7\} \quad (7)$$

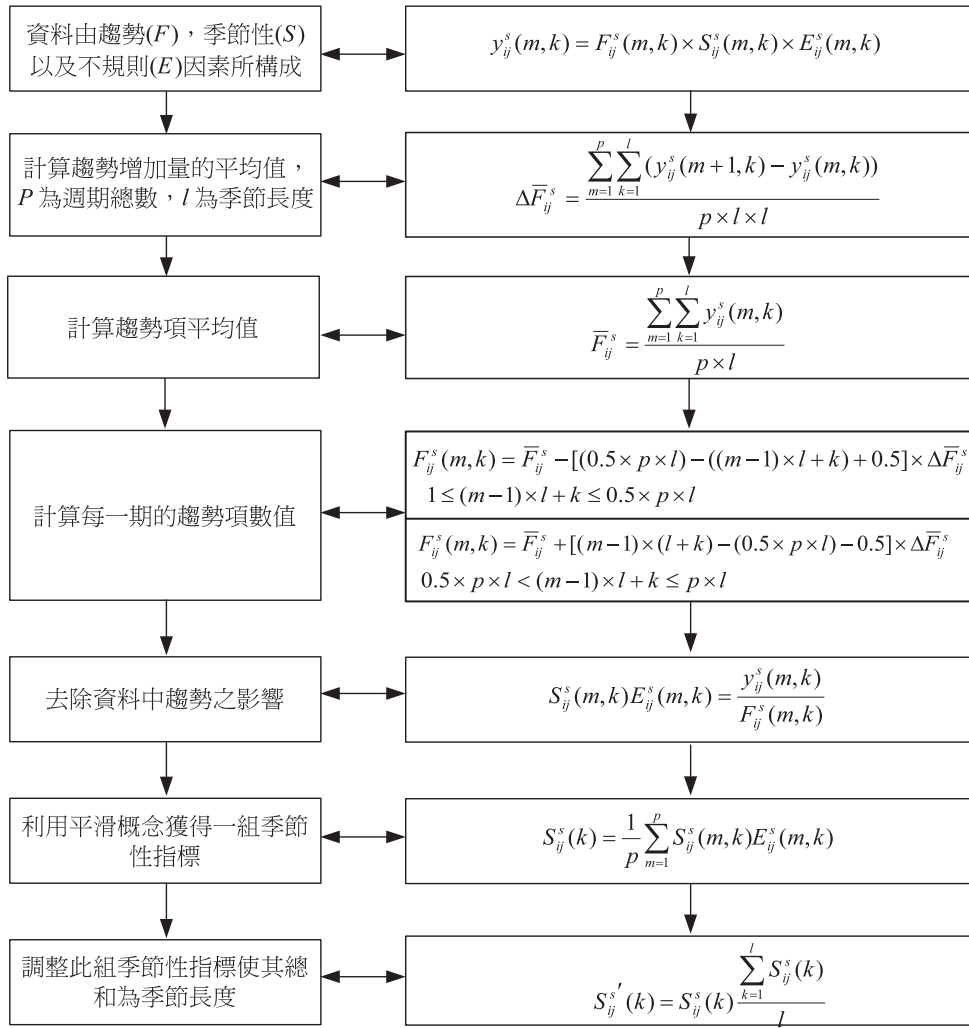


圖 8 ratio-to-moving average 概念以及演算流程

#### 4.2.4 Holt-Winters 模式 (HW)

指數平滑法更進一步將移動平均法視每一期歷史資料的貢獻皆為相同的假設，擴展為隨時間滾動具有遞遠遞減的指數關係。除此之外，指數平滑法也是由資料分解的概念所組合而成的模式，因此可依據資料中是否具有趨勢以及季節性，更進一步將模式分成三大類。第一類為資料中不含線性趨勢以及季節性影響，例如單一指數平滑法 (single exponential smoothing method)。第二類為資料中含有線性之趨勢但不含季節性影響，例如 Holt's 模式。第三類為資料同時含有線性趨勢以及季節性影響，例如本研究所採用的 Holt-Winters 模式。由於本研究使用之資料具有以週為長度之季節性影響，因此以

Holt-Winters 模式較為適合用來發展預測模式。Holt-Winters 模式假設資料由三個基礎因子所組成，分別為基礎量 ( $L_{ij}^s(t)$ )、線性趨勢 ( $T_{ij}^s(t)$ ) 以及季節性因素 ( $S_{ij}^s(t)$ ) (Makridakis 等人<sup>[41]</sup>)。為進一步了解此一演算流程，可以先從式 (11) 看出該方法之結構：輸出值  $y_{ij}^s(t+1)$  主要由基礎量  $L_{ij}^s(t)$  先加上一當期趨勢項  $T_{ij}^s(t)$  後，再乘以相對應的季節性指標  $S_{ij}^s(t-l+1)$ 。其中基礎量  $L_{ij}^s(t)$  可以解釋為數列資料平均值加上截至  $t-1$  期的趨勢影響總合。首先模式先進行去季節化之程序，如式 (8) 所示，由於去季節化後資料中還殘餘基礎量以及當期趨勢之影響，因此透過一個平滑參數  $\alpha$  並利用前一期基礎量以及趨勢之影響，來進行平滑之動作。之後利用相鄰兩期去季節化後之資料，可以去除基礎量之影響，而分離出線性趨勢，如式 (9) 所示。此時由於資料中只殘存線性趨勢，因此透過一個平滑參數  $\beta$  以及上一期線性趨勢量來進行平滑之動作。季節指標之求取則透過式 (10) 來計算，將原始資料除以基礎量即可獲得。之後再透過預測式，如式 (11)，將此三個影響因子結合在一起以提供預測值，其中  $l$  為季節長度。在 Holt-Winters 模式中，有三個係數需要求取，亦即  $\alpha$ 、 $\beta$  以及  $\gamma$ ，這三個係數反應近期資料的重要性，當數值愈大表示近期資料影響愈大。至於這三個係數之求取，為避免主觀選取之缺點，本研究使用 Hooke-Jeeves 演算法來進行搜尋 (Bazarra 等人<sup>[43]</sup>)。

$$L_{ij}^s(t) = \alpha \frac{y_{ij}^s(t)}{S_{ij}^s(t-l)} + (1-\alpha)(L_{ij}^s(t-1) + T_{ij}^s(t-1)) \quad (8)$$

$$T_{ij}^s(t) = \beta(L_{ij}^s(t) - L_{ij}^s(t-1)) + (1-\beta)T_{ij}^s(t-1) \quad (9)$$

$$S_{ij}^s(t) = \gamma \frac{y_{ij}^s(t)}{L_{ij}^s(t)} + (1-\gamma)S_{ij}^s(t-l) \quad (10)$$

$$\hat{y}_{ij}^s(t+1) = (L_{ij}^s(t) + T_{ij}^s(t))S_{ij}^s(t-l+1) \quad (11)$$

### 4.3 綜合模式

綜合模式 (combined model) 的概念，建立在其認為不同資料或是方法，都能獲得不同程度有用的資訊，因此透過整合這些有用資訊，將可以進一步有效提升預測績效。綜合模式在預測文獻上所指為針對不同模式結構之輸出，採取更進一步結合之方式，常見的為加權平均法。由於在大規模的比較實驗中發現，很難有一種預測方法可以在眾多方法中保持一致的優勢，隨著資料特性的改變，預測方法之選擇也必須跟著改變，綜合模式則成為一種很好的方式來擷取不同模式的優點，並且具有一般性 (Makridakis 等人<sup>[17]</sup>; Makridakis 與 Hibon<sup>[19]</sup>)。Armstrong<sup>[3]</sup> 在其文章中對於結合的方式有一探討，其認為針對不同方法採取不同權重來進行結合時，必須有充分的先驗知識為基礎，不然簡單的平均法則已足夠。



在 Hibon 與 Evgeniou<sup>[44]</sup> 的研究中採用簡單之平均法來進行模式的結合，以減低選用不適當預測方法時所造成之風險，也顯現出簡單平均法是結合不同種類型模式一種非常有效之方法。

本研究除了構建以類神經網路為基礎之預測模式外，也進一步構建四個統計模式，以供預測模式績效之比較。除了可以探討不同模式間競爭之結果外，也可以觀察模式結合之結果是否能更進一步提升預測績效，以及進行模式結合時是否需要其他條件。

#### 4.4 績效指標

有關績效指標之使用，文獻上也有相關探討 (Fildes<sup>[45]</sup>)。針對單一個預測模式，預測誤差最常被用來作為績效評估之準則。由此所衍伸出來的包括誤差平方 (mean square error, MSE)、誤差平方根 (root mean square error, RMSE)、絕對誤差百分比 (mean absolute percentage error, MAPE) 等，也有考慮模式使用係數個數所衍伸的 Akaike's Information Criterion 或 Bayesian Information Criterion 評估準則 (Chatfield<sup>[32]</sup>)。適當的評估準則應該是判定模式預測能力，而非模式配適能力。本研究將觀察模式 MSE 以及 MAPE 值在測試資料集的預測能力。針對預測問題，此二指標特性略有不同。MSE 單純計算模式預測誤差值，而 MAPE 則可以更進一步反應相同預測誤差值在尖離峰區間所代表的不同意義，例如同樣是預測誤差 10 人，在尖峰區間或許是可以容忍，但在離峰區間可能變成無法容忍。MSE 以及 MAPE 公式如式 (12) 及式 (13) 所示，其中  $N$  為測試樣本數， $y_{ij}^s(t)$  為第  $t$  期之實際旅運量，而  $\hat{y}_{ij}^s(t)$  為第  $t$  期預測值。為了比較兩個不同預測模式間差異，以提供模式架構選擇之參考依據，本研究採用成對  $t$  檢定來加以辨別兩個模式 MSE 與 MAPE 值，是否具有統計上的顯著差異 (Hill 等人<sup>[12]</sup>)。

$$MSE = \frac{1}{N} \sum_{t=1}^N (y_{ij}^s(t) - \hat{y}_{ij}^s(t))^2 \quad (12)$$

$$MAPE = \frac{1}{N} \sum_{t=1}^N \frac{|y_{ij}^s(t) - \hat{y}_{ij}^s(t)|}{y_{ij}^s(t)} \times 100\% \quad (13)$$

## 五、實證分析

此節我們將以資料分析所得結果為基礎建立預測模式，探討輸入變數設計、MLFN 網路模式的效力以及綜合模式的效力。

### 5.1 輸入變數設計

輸入變數設計主要是以資料分析所獲得的特性為基礎，設計數種不同輸入方式來測試

MLFN 預測績效，並觀察網路最後模式架構是否與資料分析的結果一致。

### 5.1.1 設計情境

針對資料分析所得之特性、趨勢、季節性、寒暑假尖峰效果為重要的測試標的，其中更以季節性以及寒暑假尖峰因子在資料分析時的影響最為明顯。我們針對各個輸入變數，設計幾個不同輸入型態來解釋旅運量變化。其一，星期特性變數被用來表示以週為長度的季節特性。假設個別星期特性都被視為具有不同的影響效果，可以有離散型（設計 1）以及連續型（設計 2）兩種輸入方式。若假設寒暑假期間季節性影響和平日並不相同，則可以按寒暑假期間不同，設定個別星期特性變數（設計 3 及設計 4）。其二，假期尖峰虛擬變數被用來加強寒暑假效果，針對寒假以及暑假，我們可以將其對旅運量所帶來的效果假設為相同或者是不同，而有兩種不同之設計（設計 5 及設計 6）。其三，週在月份中的順序變數被使用來表示不同週的影響可能不同，這是因為受到隔週休二日的政策所影響，在每月份第二、第四週為週休二日，其餘週末則禮拜六早上仍需上班。而在變數設定方面也可以有離散型或是連續型（設計 7 及設計 8）。其四，從圖 1 來看，雖然資料呈現向平均數迴歸的情形，沒有明顯向上或向下的長期趨勢，但在此我們仍以一個連續型的時間變數利用模式，來測試長期趨勢是否存在（設計 9），其中測試檔輸入變數之數值為訓練檔之延續，以持續該趨勢效果。其五，我們更進一步假設月份具有一個慣性 (momentum) 的影響力量，例如今年 1 月份的旅運需求會影響明年 1 月份的旅運需求。為此我們採用月份順序變數來表示資料，在變數設定方面也可以有離散型（設計 10）或是連續型（設計 11）。其六，我們將趨勢的影響改以月份總量的角度來觀察，這是因為以短期的角度來觀察長期趨勢並不容易，若以較為總體之變數或許可以觀察出趨勢變化。因此我們以月份為單位來測試是否有月份趨勢存在，可用一個連續變數來表示（設計 12），其中測試檔輸入變數之數值為訓練檔之延續，以持續該趨勢效果。下表 2 為上述變數所形成之變數表，各欄位意義簡述如下：設計編號主要用來區別每一種不同型態的變數設計，以便後續章節解說時使用。情境特性主要為資料特性之描述，而輸入變數則為該情境特性之命名。由於每一種輸入變數又可以有不同之輸入型態，所以輸入型態假設之欄位主要論述假設的意義。數值設計欄位主要說明在每種不同輸入型態假設下變數的輸入情形，例如我們可以用 1 個連續變數 (1,2,3,4,5,6,7) 來表示星期特性變數（設計 1），我們也可以用 3 個 0/1 虛擬變數來表示星期特性（設計 2）。最後一個欄位則為每一種設計所用到的數值範圍。此外，趨勢變數在訓練及測試檔之設定具有連續之概念。

為了觀察模式最後整體績效，我們採用逐步選取 (stepwise selection) 的概念來選擇輸入變數，觀察模式預測績效以及模式輸出值，以辨別這些模式是否能透過輸入變數的設計，成功擷取到旅運需求特性。訓練資料為 1999 年全年資料，而測試資料為 2000 年全年資料。

表 2 輸入變數列表

| 設計編號 | 情境特性           | 輸入變數   | 輸入型態假設                 | 數值設計   | 數值範圍                     |
|------|----------------|--------|------------------------|--------|--------------------------|
| 設計1  | 以週為長度<br>季節特性  | 星期特性變數 | 連續型輸入可以表示該特性           | 1個連續變數 | 1,...,7                  |
| 設計2  |                |        | 離散型輸入可以表示該特性           | 3個虛擬變數 | 0,1                      |
| 設計3  |                |        | 寒暑假星期特性與一般日不同 (離散型)    | 6個虛擬變數 | 0,1                      |
| 設計4  |                |        | 寒假、暑假、一般日星期特性皆不同 (離散型) | 9個虛擬變數 | 0,1                      |
| 設計5  | 寒暑假尖峰<br>時段    | 尖峰虛擬變數 | 寒暑假效果一樣                | 1個虛擬變數 | 0,1                      |
| 設計6  |                |        | 寒假與暑假效果不一樣             | 2個虛擬變數 | 0,1                      |
| 設計7  | 隔週休二日<br>特性    | 週次順序變數 | 連續型輸入可以表示該特性           | 1個連續變數 | 1,...,6                  |
| 設計8  |                |        | 離散型輸入可以表示該特性           | 3個虛擬變數 | 0,1                      |
| 設計9  | 持續成長或<br>下降特性  | 日趨勢變數  | 連續型態                   | 1個連續變數 | 訓練1,...,365<br>測試366,... |
| 設計10 | 月份慣性           | 月份順序變數 | 連續型輸入可以表示該特性           | 1個連續變數 | 1,2,...,12               |
| 設計11 |                |        | 離散型輸入可以表示該特性           | 3個虛擬變數 | 0,1                      |
| 設計12 | 月份總量呈<br>現成長趨勢 | 月份趨勢變數 | 連續型態                   | 1個連續變數 | 訓練1,...,12<br>測試13,...   |

### 5.1.2 變數篩選結果

表 3 為模式最終測試結果，其中有幾個重要發現。其一，針對星期特性變數的連續型（設計 1）以及離散型（設計 2）輸入，雖然兩者 MSE 與 MAPE 值績效差異不大，但從模式輸出值我們發現離散型輸入方式，才能有效表示出星期特性波動上的差異，如圖 9 所示。其次，進一步觀察寒暑假與一般日的星期特性分布是否一樣，可以發現若將寒假以及暑假效果視為不同時（設計 4），對寒假時期的預測明顯不好。而將寒暑假效果視為相同（設計 3），比較能提供準確的預測值。

其二，若模式加入一個虛擬變數來表示寒暑假尖峰特性（設計 3+5），則發現該變數並無改善模式預測績效；若將寒假以及暑假視為兩種不同尖峰效果（設計 3+6），則可以發現模式績效會略為改善。為了進一步確認寒暑假尖峰變數與寒暑假內星期特性變數的效果是否重疊，亦即上述寒暑假與一般日星期特性差異，其實是因為寒暑假尖峰關係而並非星期特性真的和平常日有所不同，我們測試了結合離散型星期變數（設計 2）與尖峰寒暑假效果（設計 5）模式。結果發現此一模式結果略差於先前模式（設計 3+6）的績效，此一結果顯示寒暑假內星期特性與一般日不同除了由寒暑假尖峰因子所引起之外，也包含些許星期特性不同的成分在內。值得注意的是，設計 3+6 在 2000 年寒假期間的預測值產生了低估

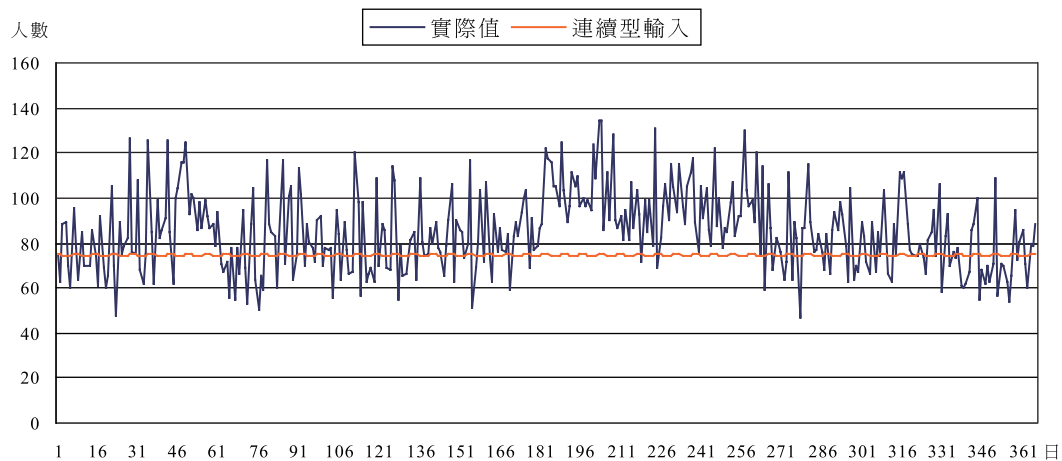


圖 9(a) 星期特性連續型輸入之預測結果

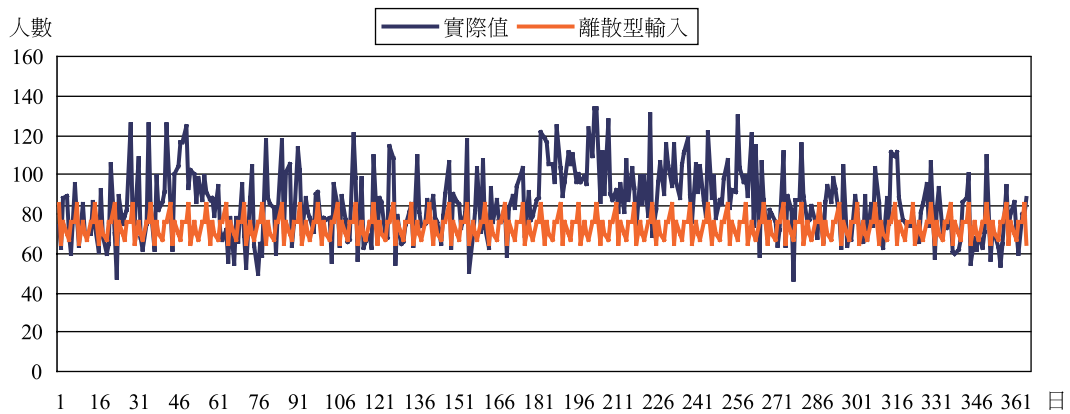


圖 9(b) 星期特性離散型輸入之預測結果

的現象，顯示出 1999 年的寒假尖峰型態，並無法有效預測 2000 年的寒假尖峰形態。下圖 10(a)擷取出 1999 年以及 2000 年實際寒假旅運數值來加以比較，可以發現 1999 年的尖峰型態，並不如 2000 年寒假明顯。為了解決此一現象，本研究將 1999 年寒假時期視為一般日，不再給予一尖峰因子，而改以 1999 年暑假旅運型態去訓練並預測 2000 年寒假以及暑假的旅運型態，給與模式設計編號 3+6'。預測績效結果顯示設計 3+6' 可以有效改善模式預測績效，圖 10(b) 也顯示出此種設計可以解決先前設計 3+6 在寒假期間預測低估的現象。

其三，若模式中再加入一個連續變數（設計 3+6'+7）或者三個虛擬變數（設計 3+6'+8）來表示隔週休二日之政策時，可以發現以虛擬變數之設定方式較能夠反應此一特性，模式 MSE 值獲得改善，但 MAPE 值則略微上升。總而來說，隔週休二日的政策在先驗知

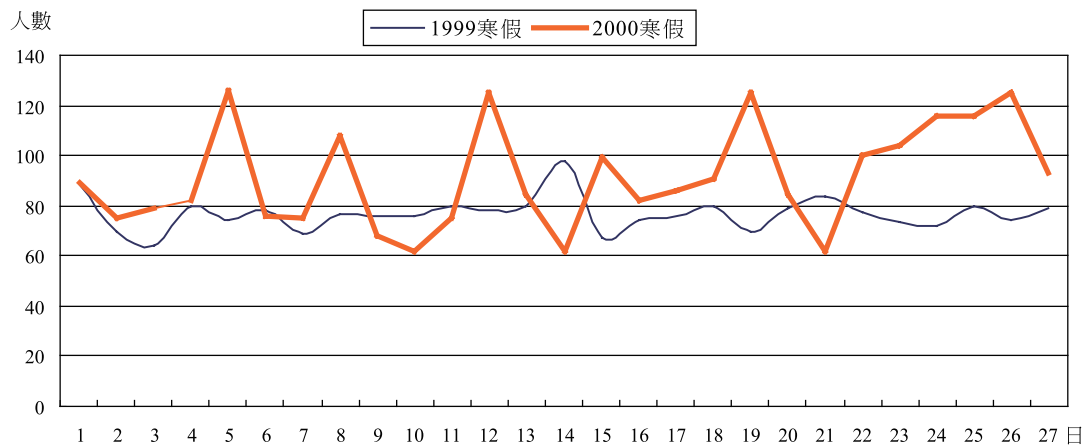


圖 10(a) 1999 年與 2000 年寒假旅運量分布

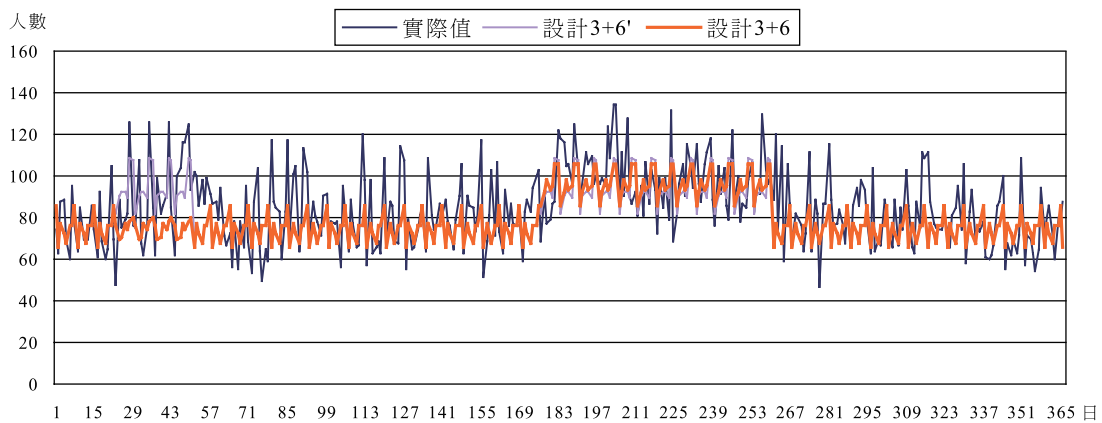


圖 10(b) 修正 1999 年寒假影響前後之差異

識上對於旅運行為會產生變化，例如休閒旅次在週休二日的星期會有所增加，而且模式 MSE 值也獲得改善，因此本研究選擇保留該變數。

其四，若模式中加入日趨勢變數（設計  $3+6'+8+9$ ），則發現會使模式績效惡化。由圖 1 的分析中可以概略發現，資料中並不存有日趨勢之現象，若強制加諸於模式之中，反而會導致反效果。因此對多層次前向式類神經網路來說，不適當的變數使用反而有可能使模式績效變差。雖然理論上時間數列資料大都具有趨勢特性，但對於研究資料而言，長期趨勢之特性並不明顯。

其五，若模式加入月份慣性變數，可以發現不論是以連續變數（設計  $3+6'+8+10$ ）或是虛擬變數（設計  $3+6'+8+11$ ）來表示該特性，都會得到反效果。此一結果顯示，月份慣性現象並不存在。進一步思考該變數意義與目前模式所使用的資料型態，可以發現月份

表 3 模式輸入變數之篩選

| 輸入變數   | 設計編號         | MSE           | MAPE (%)     |
|--------|--------------|---------------|--------------|
| 星期特性變數 | 設計 1         | 423.67        | 17.70        |
|        | 設計 2         | 424.16        | 17.13        |
|        | 設計 3         | <b>277.70</b> | <b>14.54</b> |
|        | 設計 4         | 299.20        | 14.73        |
| 尖峰虛擬變數 | 設計 3+5       | 283.61        | 14.58        |
|        | 設計 3+6       | 271.18        | 14.33        |
|        | 設計 2+5       | 274.33        | 14.53        |
|        | 設計 3+6'      | <b>251.50</b> | <b>14.34</b> |
| 周次順序變數 | 設計 3+6'+7    | 269.41        | 14.76        |
|        | 設計 3+6'+8    | <b>239.55</b> | <b>14.43</b> |
| 日趨勢變數  | 設計 3+6'+8+9  | 336.60        | 17.00        |
| 月份慣性變數 | 設計 3+6'+8+10 | 313.81        | 15.83        |
|        | 設計 3+6'+8+11 | 371.96        | 18.26        |
| 月份趨勢變數 | 設計 3+6'+8+12 | 329.34        | 16.74        |

慣性特性應會在資料具有月份重現性事件 (recurrent events) 時較有作用。而如同上一部分之測試結果，多餘變數將會使預測績效變差。相同的現象也發生在模式加入月份趨勢項(設計 3+6'+8+12) 之情形，此一結果顯示月份趨勢的假設也並不成立。

從變數的設計以及篩選過程中可以得到以下幾點結論：其一，從被選入的輸入變數型態上來看，以離散型輸入方式較連續型輸入方式來得好，MLFN 可以由離散型輸入得到比較正確的資訊。此點說明人們認知的現象，網路並不一定能認知到，必須透過適當的變數設計來達成。其二，變數中以星期特性變數、尖峰虛擬變數以及週次順序變數最為重要，對模式預測績效具有正面助益。其三，寒暑假和一般日除了有尖離峰差別外，在星期特性上也有所不同。其四，當資料特性出現變化時，必須適當地調整模式設定，如 1999 年寒假無法反應 2000 年寒假型態，則改以 1999 年暑假取代，因此人為的監控以及調整有其必要。其五，若模式中加入多餘解釋變數，則對於 MLFN 預測績效反而會有傷害。

## 5.2 模式預測績效之比較

為了比較本研究所發展模式之效力，我們將 MLFN 與其他四個預測方法比較。而在此之前，我們先考慮以動態更新模式參數估計值的概念來定期更新模式。在上述 MLFN 模式中，我們使用一年期資料來構建模式，並以下一年期資料來測試模式。此一方法雖然可行，但卻忽略了應用當時間演進時新進可得之資料來抓住較新的旅運趨勢。因此更進一

步之作法是隨著資料的可得，動態更新模式訓練樣本以及參數估計值，以便將新的旅運趨勢資訊提供給模式。如此一來，模式便會變成一滾動之狀態 (rolling state)，讓模式能反應近期的旅運趨勢。本研究應用移動時窗學習法 (moving window learning method) 定期以一個月為區間，更新模式訓練樣本以及參數估計值。

其他四個預測方法，我們也以相同的概念進行模式的更新，以獲得預測值。另一方面，由於倒傳遞演算法以及 Hooke-Jeeves 演算法的一個缺點為皆有機會收斂至局部最佳解 (local minima)，為了減輕此一現象對網路績效的影響，本研究嘗試 100 個不同的隨機起始點，並以較佳的 50 個績效結果平均值，來加以比較不同模式間的差異。圖 11、圖 12 顯示出五個模式在 12 個測試子集中 MAPE、MSE 的績效值分布。

從圖 11 與圖 12 中，若個別從每一個月份績效值來進行比較，我們發現很難有一致之結論。表 4 列出各個模式 MSE 與 MAPE 值的平均績效表現，由表中可以發現 MLFN 的平均績效表現優於 RW、DRW 以及 MV，但卻略遜於 HW 模式。若進一步應用成對 t 檢定對這五個模式加以進行比較，則可以得到表 5 與表 6 的結果。表 5 與表 6 中每一欄位分別為兩個模式績效差異的檢定結果，虛無假設為兩種模式並無績效上之差異，樣本數皆為 12 筆。表 5 及表 6 中除了提供 t 值之外，也提供多種不同信心水準下的結論，信心水準愈高，表示因為樣本隨機性所導致錯誤決策的機會愈低，從表中可以發現幾個現象：其一，HW 優於 MV，表示資料具有某種波動型態，且簡單之移動平均法並無法有效抓住此一型態。其二，HW 及 MV 皆優於 DRW，表示本研究使用之資料並非只存在隨機與季節之效果，其他可能之影響因素應該進一步被考慮。其三，HW、MV、DRW 皆優於 RW，除了可有上述第二點結論外，也顯示去季節化處理的效益。其四，MLFN 優於 RW 及 DRW，但和 MV 以及 HW 並無顯著之差異，此一結果值得進一步討論。首先 MLFN 優於 RW 以

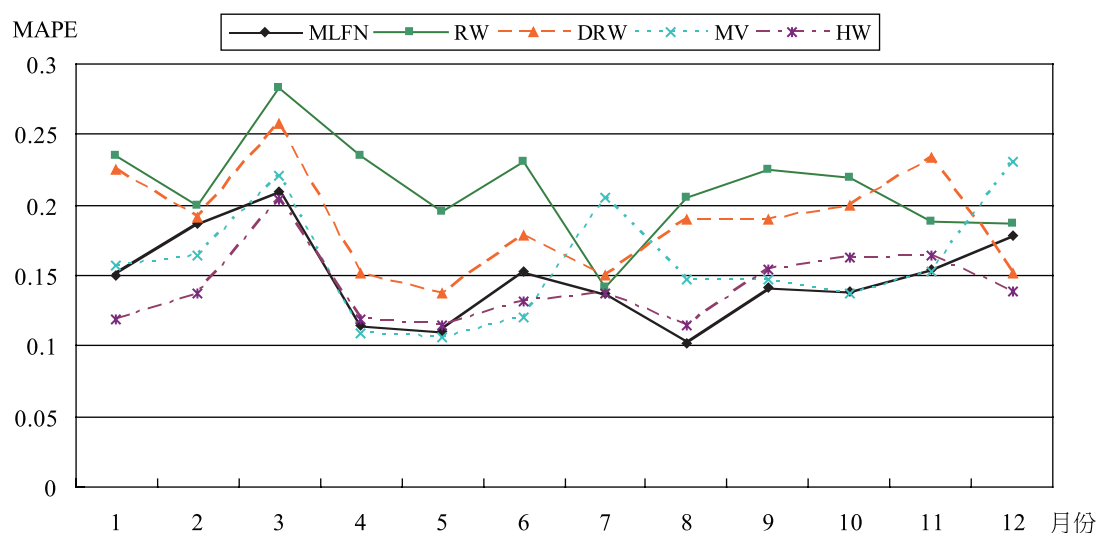


圖 11 MLFN 與四個統計模式之 MAPE 值比較

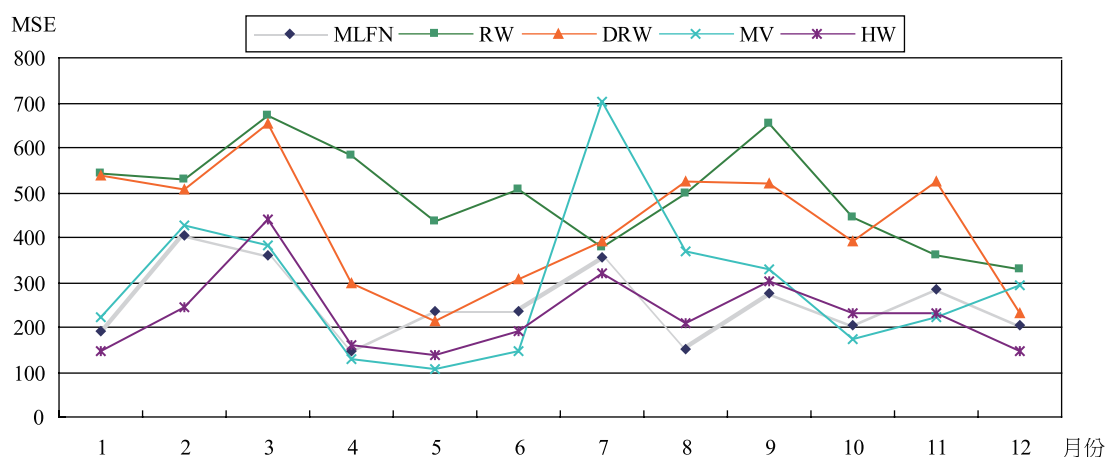


圖 12 MLFN 與四個統計模式之 MSE 比較

表 4 各模式 MAPE 與 MSE 的平均績效表現

|      | MAPE | MSE    |
|------|------|--------|
| MLFN | 0.15 | 254.59 |
| RW   | 0.21 | 493.95 |
| DRW  | 0.19 | 424.29 |
| MV   | 0.16 | 292.49 |
| HW   | 0.14 | 230.35 |

表 5 模式 MAPE 值成對 t 檢定之結果 (t 值)

|      | MLFN   | RW    | DRW    | MV      | HW |
|------|--------|-------|--------|---------|----|
| MLFN |        |       |        |         |    |
| RW   | -5.65* |       |        |         |    |
| DRW  | -4.05* | 2.52* |        |         |    |
| MV   | -1.15  | 3.27* | 2.12** |         |    |
| HW   | 0.97   | 7.14* | 5.88*  | 1.59*** |    |

註：\* 表 95% 信心水準；\*\* 表 90% 信心水準；\*\*\* 表 80% 信心水準。

及 DRW，表示資料的確具有某種型態，單純的隨機效果或再加上考量去季節化處理，都無法得到滿意的預測績效。其次 MLFN 和 MV 以及 HW 之間無顯著差異，可以透過圖 13 至圖 15 來進一步觀察預測值與實際值之間的分布。由圖 13 與圖 14 的比較，我們可以發現 MLFN 較 MV 在尖峰時期，亦即寒暑假區間，可以得到較好的預測型態，如兩圖中矩



表 6 模式 MSE 值成對 t 檢定之結果 (t 值)

|      | MLFN   | RW     | DRW   | MV      | HW |
|------|--------|--------|-------|---------|----|
| MLFN |        |        |       |         |    |
| RW   | -6.29* |        |       |         |    |
| DRW  | -4.47* | 1.93** |       |         |    |
| MV   | -0.99  | 3.38*  | 2.63* |         |    |
| HW   | 1.22   | 8.56*  | 6.39* | 1.60*** |    |

註：\* 表 95% 信心水準；\*\* 表 90% 信心水準；\*\*\* 表 80% 信心水準。

形所圍住之區域。MV 之機制顯然無法有效學習到此一尖峰特性，因為尖峰效果已在計算平均值時被消除了。因此就預測值與實際值分布的情形來看，MLFN 比 MV 之表現要來的好。由圖 13 以及圖 15 的比較，可以發現在 2 月份後期的預測中，HW 有高估而 MLFN 卻有低估的現象，如兩圖中的橢圓型所圍住之區域。這是由於 HW 會考量近期資料之重要性，並以平滑係數之大小加以反應，因此 HW 持續寒假時期的趨勢，在非寒假期間產生高估的預測值。在 MLFN 中將每一期資料都一視同仁，在尖峰變數的作用下，於寒假結束後，將預測值調回正常時間特性型態，由於寒假效果似乎有殘留跡象（但不如寒假期間效果大），因此產生低估的預測值。根據本段的平均績效結果、檢定結果以及模式輸出結果，本研究認為 MLFN 以及 HW，是五個模式中較為適合短期需求預測的候選模式。

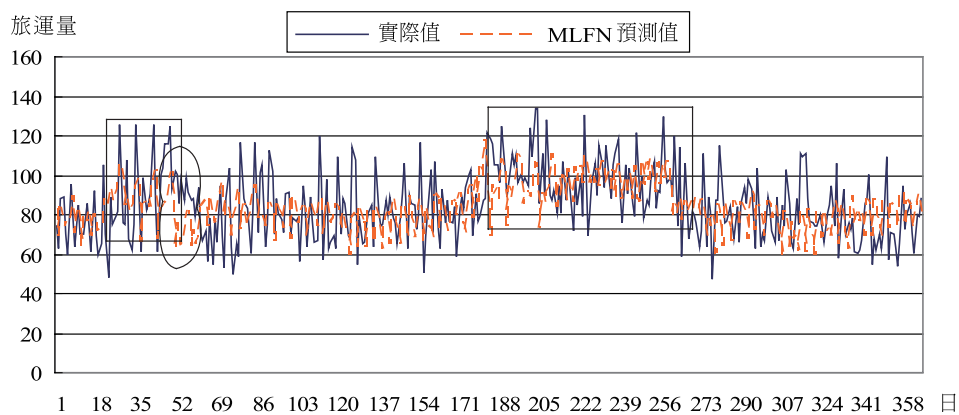


圖 13 MLFN 預測值與實際值分布

### 5.3 綜合模式 (combined model)

在上述實證分析中，我們使用了五種性質不同的模式，最終發現針對研究資料 MLFN 與 HW 的預測績效相近，並且優於 RW、DRW 以及 MV。此一小節將驗證綜合模式的概

念，是否可以有效提升預測績效。直觀上針對個別模式的績效結果，我們認為只需針對 MLFN 以及 HW 進行結合即可，因為此兩者比較能掌握資料的型態，並且獲得較佳的預測績效。至於若將 RW、DRW 以及 MV 也一併考慮進入綜合模式，可能反而會使預測績效變差。為了驗證此段論述，我們可以測試所有可能的組合型態，產生 26 個組合模式 ( $C_2^5 + C_3^5 + C_4^5 + C_5^5 = 26$ )，不過其意義上並不大。在此我們將 MLFN 與 HW 先進行組合，以建立基本的組合模式，並觀察預測績效是否可以比個別模式好。若此一基礎組合模式能顯著改善預測績效，之後再以此基礎的組合模式與 MV、DRW 以及 RW 進行結合並觀察績效改進的情形，若績效能獲得改進，則進一步考慮其他可能組合，否則停止結合的程序。

表 7 先列出 MLFN 與 HW 結合時的平均績效結果，之後再觀察與其他三個模式結合，是否能更進一步提升平均的績效表現。結果發現 MLFN 與 HW 的結合，的確可以有效提升平均預測績效，但更進一步與其他三個模式進行結合時，則發現平均績效結果並不能獲

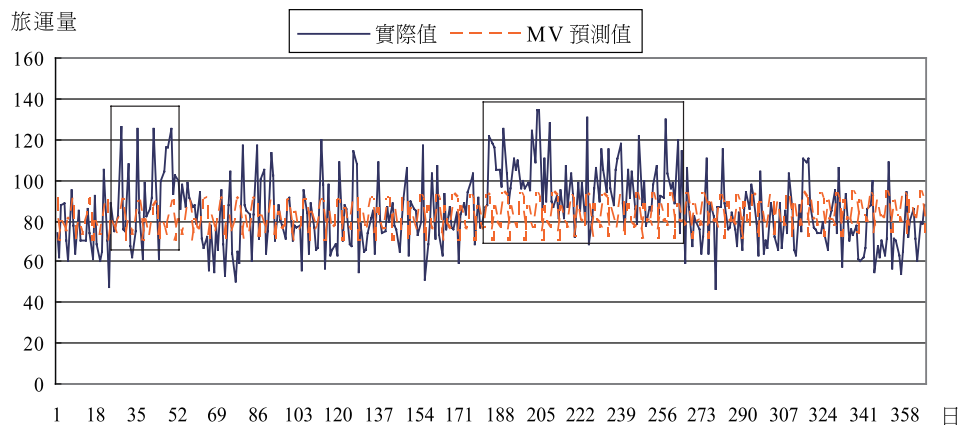


圖 14 MV 預測值與實際值分布

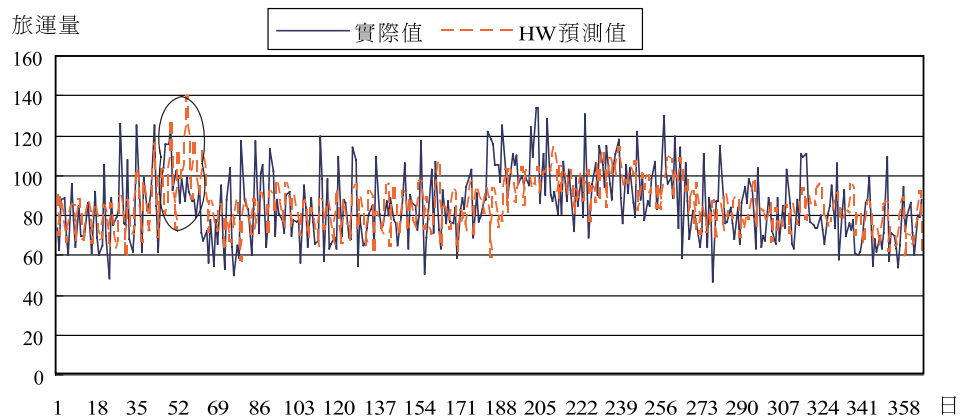


圖 15 HW 預測值與實際值分布

得更進一步的改善。表 8、表 9 及表 10 為模式結合的檢定結果，其中每一欄位分別為兩個模式績效差異的檢定結果，虛無假設為兩種模式並無績效上之差異，樣本數為 12 筆。由表 8 與表 9 之結果可以顯示，基礎綜合模式比 MLFN 或 HW 的個別績效表現，無論在 MAPE 或是 MSE 上，都獲得顯著改善。若進一步將基礎綜合模式與 MV 進行結合，表 10 顯示兩個不同結果：其一，若基礎綜合模式再加入 RW 或者 DRW，都會導致績效結果顯著惡化。其二，若基礎綜合模式再加入 MV，則並無法進一步提升預測績效。由此小節的測試結果來看，綜合模式的觀念的確有用，但是必須從個別模式中謹慎選取候選模式來進行結合，而在個別模式測試中表現較好的模式，則是較佳的選擇。

#### 5.4 模式健全性討論

在資料分析部分，我們對所選取的研究對象（1008 班次）進行特徵觀察之動作。在修正離群值以及國定假日後，發現研究資料具有以週為長度的週期特性、受隔週休二日影

表 7 綜合模式的平均預測績效

|                 | MAPE | MSE    |
|-----------------|------|--------|
| MLFN + HW       | 0.13 | 209.7  |
| MLFN + HW + RW  | 0.16 | 234.52 |
| MLFN + HW + DRW | 0.14 | 231.89 |
| MLFN + HW + MV  | 0.14 | 216.38 |

表 8 基礎綜合模式與個別模式 MAPE 值成對 t 檢定之結果

|           | MLFN  | HW    | MLFN + HW |
|-----------|-------|-------|-----------|
| MLFN      |       |       |           |
| HW        | 0.97  |       |           |
| MLFN + HW | 2.30* | 2.39* |           |

註：\*表 95% 信心水準。

表 9 基礎綜合模式與個別模式 MSE 值成對 t 檢定之結果

|           | MLFN  | HW    | MLFN + HW |
|-----------|-------|-------|-----------|
| MLFN      |       |       |           |
| HW        | 1.22  |       |           |
| MLFN + HW | 2.80* | 2.90* |           |

註：\*表 95% 信心水準。

表 10 基礎綜合模式再加入單一個別模式預測績效成對 t 檢定之結果

|                 | MLFN + HW |        |
|-----------------|-----------|--------|
|                 | MAPE      | MSE    |
| MLFN + HW + RW  | -2.46*    | -2.11* |
| MLFN + HW + DRW | -1.61***  | -2.42* |
| MLFN + HW + MV  | -0.42     | -0.83  |

註：\*表 95% 信心水準；\*\*\*表 80% 信心水準。

響、寒暑假尖峰影響的顯著特性。針對這些特性，本研究在類神經網路的結構下，嘗試不同的變數輸入方式，觀察何種輸入方式對於預測績效最有幫助，並與四種統計模式加以比較。然上述特性觀察只限於 1008 班次，並不能包含所有可能存在於鐵路資料內的特性，例如某些班次可能具有明顯的日趨勢形態，亦即資料特性可能並不齊一(non-homogenous)。在資料特性齊一的情形之下，一個模式就能有效表達所有可能存在之特性。但在資料特性不齊一的情形之下，針對個別數列發展適合之預測模式的方式則較為恰當，而此也為本研究之切入角度。雖然本研究所發展之模式需要經過修改，才能適用於其他鐵路數列資料，然當其他鐵路數列資料也具有前述三項觀察之資料特性時，相同的變數設計方式以及模式構建經驗，也可以被加以應用。至於模式的績效比較以及模式選擇等課題，在本研究所得之結論，對於具有相同資料屬性的數列，預期也會得到類似結論。但對於資料屬性不同之數列，在模式績效比較以及模式選擇課題上，則值得更進一步進行檢驗，惟不同資料屬性下之模式選擇課題，則已超出本研究之範圍。

針對本研究所使用的其他四類統計模式，其健全性之說明如下：首先此四類模式在構建時的樣本使用方式以及模式更新方式，與類神經網路是相同的。在此條件下，唯一不同的是不同模式方法的機制，是否能有效表現研究數列的資料特性。隨機模式假設資料是完全隨機的，因此並無將上述觀察的資料特性考慮在模式之中。去季節化隨機模式則考量了以週為長度之季節特性，利用一組季節指標將該因子之影響消除後，假設剩餘資料分布為隨機現象。移動平均模式的概念則也能適當表達以週為長度之季節特性，然其他可能存在之因素，則由於平均之使用，被混入以週為長度的特性因子之中。Holt-Winters 模式也考慮了以週為長度的週期因子，更重要的是該模式能考慮不同樣本點的重要性（亦即近期資料較遠期資料重要）。此一特性讓 Holt-Winters 模式能在短時間內反應寒暑假之特性，因此獲得與類神經網路類似之績效。同樣的，由於資料特性並不齊一，現階段的模式架構可能並不適合所有的鐵路數列資料，而適當的模式修改，才能更進一步表達不同的資料特性。

## 六、結論

本研究以類神經網路為基礎來構建短期列車旅運需求預測模式，主要學術貢獻在於我

們詳細分析旅運需求特性，並根據這些特性來構建一具資料特性導向的類神經網路模式。在修正資料離群值以及國定假日之後，研究資料主要呈現以週為長度之季節特性、受隔週休二日政策影響、寒暑假尖峰時期具不平穩性、無明顯日趨勢的特性。在此資料特性下，我們以類神經網路構建預測模式，並探討三個研究課題，得到以下發現：其一，我們發現星期特性、寒暑假尖峰特性以及週次特性最為重要，而這些變數以離散型輸入設定能讓網路獲得較佳績效結果。此外，這幾個重要變數也符合資料分析時所獲得的資料特性。其二，若與其他統計方法加以比較，則發現類神經網路績效值優於隨機模式、去季節化隨機模式以及星期特性模式，但和指數平滑法之 Holt-Winters 模式績效結果相近。其三，綜合模式績效結果優於個別模式之績效結果，惟候選模式應以在個別模式中表現較佳的模式為考慮對象。

其他值得進一步探討之議題，建議如下：針對鐵路資料進行較多數列資料之比較，並建立資料特性與模式使用的關係；使用或發展網路模式，以擷取出資料中所隱含的其他資訊，來提升預測績效；以歷史資料為輸入變數的角度，探討模式結構、輸入變數篩選以及預測績效；探討不同模式結構對於單期預測以及多期預測的績效表現。

## 參考文獻

1. Kimes, S. E., "Perceived Fairness of Yield Management", *Cornell Hotel and Restaurant Administration Quarterly*, Vol. 43, No. 1, 2002, pp. 21-30.
2. Lee, A. O., "Airline Reservations Forecasting: Probabilistic and Statistical Models of the Booking Process", Ph. D. dissertation, MIT, USA, 1990.
3. Armstrong, J. S., *Principles of Forecasting*, Kluwer Academic Publishers, Boston, US, 2001.
4. Grubb, H. and Mason, A., "Long Lead-time Forecasting of UK Air Passengers by Holt-Winters Methods with Damped Trend", *International Journal of Forecasting*, Vol. 17, 2001, pp. 71-82.
5. Williams, B. M., Durvasula, P. K., and Brown, D. E., "Urban Freeway Traffic Flow Prediction: Application of Seasonal Autoregressive Integrated Moving Average and Exponential Smoothing Models", *Transportation Research Record 1644*, 1998, pp. 132-141.
6. Williams, B. N. and Hoel, L. A., "Modeling and Forecasting Vehicular Traffic Flow as a Seasonal ARIMA Process: Theoretical Basis and Empirical Results", *Journal of Transportation Engineering*, Vol. 129, 2003, pp. 664-672.
7. Smith, B. L., Williams, B. M., and Oswald, R. K., "Comparison of Parametric and Nonparametric Models for Traffic Flow Forecasting", *Transportation Research Part C*, Vol. 10, 2002, pp. 303-321.
8. Clark, S., "Traffic Prediction Using Multivariate Nonparametric Regression", *Journal of Transportation Engineering*, Vol. 129, No. 2, 2003, pp. 161-168.
9. 許巧鶯、溫裕弘，「臺灣地區國際航空客運量之預測—灰色預測模式之應用」，**運輸計劃**

季刊，第二十六卷，第三期，民國八十六年，頁 525-556。

10. Sfetsos, A., "A Comparison of Various Forecasting Techniques Applied to Mean Hourly Wind Speed Time Series", *Renewable Energy*, Vol. 21, 2000, pp. 23-35.
11. Smith, B. L. and Demetsky, M. J., "Short-term Traffic Flow Prediction: Neural Network Approach", *Transportation Research Record 1453*, 1994, pp. 98-104.
12. Hill, T., O'Connor, M., and Remus, W., "Neural Network Models for Time Series Forecasts", *Management Science*, Vol. 42, No. 7, 1996, pp. 1082-1091.
13. Dharia, A. and Adeli, H., "Neural Network Model for Rapid Forecasting of Freeway Link Travel Time", *Engineering Applications of Artificial Intelligence*, Vol. 16, 2003, pp. 607-613.
14. Ishak, S. and Alecsandru, C., "Optimizing Traffic Prediction Performance of Neural Networks under Various Topological, Input, and Traffic Conditions Setting", *Journal of Transportation Engineering*, Vol. 130, No. 4, 2004, pp. 452-465.
15. 魏健宏、林士傑、李穎，「高速公路客運車輛旅行時間預測之實證評析」，*運輸計劃季刊*，第三十二卷，第四期，民國九十二年，頁 651-680。
16. Prudêncio, R. B. C. and Ludermir, T. B., "Meta-learning Approaches to Selecting Time Series Models", *Neurocomputing*, Vol. 61, 2004, pp. 121-137.
17. Makridakis, S., Andersen, A., Carbone, R., Fildes, R., Hibon, M., Lewandowski, R., Newton, J., Parzen, E., and Winkler, R., "The Accuracy of Extrapolation (Time Series) Methods: Results of a Forecasting Competition", *Journal of Forecasting*, Vol. 1, 1982, pp. 111-153.
18. Fildes, R., Hibon, M., Makridakis, S., and Meade, N., "Generalising about Univariate Methods: Further Empirical Evidence", *International Journal of Forecasting*, Vol. 14, 1998, pp. 339-358.
19. Makridakis, S. and Hibon, M., "The M3-Competition: Results, Conclusions and Implications", *International Journal of Forecasting*, Vol. 16, 2000, pp. 451-476.
20. Shah, C., "Model Selection in Univariate Time Series Forecasting Using Discriminant Analysis", *International Journal of Forecasting*, Vol. 13, 1997, pp. 489-500.
21. Adya, M. and Collopy, F., "How Effective Are Neural Networks at Forecasting and Prediction? A Review and Evaluation", *Journal of Forecasting*, Vol. 17, 1998, pp. 481-495.
22. Ledoux, C., "An Urban Traffic Flow Model Integrating Neural Networks", *Transportation Research Part C*, Vol. 5, No. 5, 1997, pp. 287-300.
23. Faraway, J. and Chatfield, C., "Time Series Forecasting with Neural Networks: A Comparative Study Using the Airline Data", *Applied Statistics*, Vol. 47, No. 2, 1998, pp. 231-250.
24. Lachtermacher, G. and Fuller, J. D., "Backpropagation in Time-series Forecasting", *Journal of Forecasting*, Vol. 14, 1995, pp. 381-393.
25. Zhang, H. M., "Recursive Prediction of Traffic Conditions with Neural Network Models", *Journal of Transportation Engineering*, Vol. 126, No. 6, 2000, pp. 472-481.
26. Ishak, S., Kotha, P., and Alecsandru, C., "Optimization of Dynamic Neural Networks Performance for Short-term Traffic Prediction", *Transportation Research Record 1836*, 2003, pp. 45-56.

27. Vlahogianni, E. I., Golias, J. C., and Karlaftis, M. G., "Short-term Traffic Forecasting: Overview of Objectives and Methods", *Transport Reviews*, Vol. 24, No. 5, 2004, pp. 533-557.
28. Piramuthu, S., Ragavan, H., and Shaw, M. J., "Using Feature Construction to Improve the Performance of Neural Networks", *Management Science*, Vol. 44, No. 3, 1998, pp. 416-429.
29. Nam, K. and Schaefer, T., "Forecasting International Airline Passenger Traffic Using Neural Networks", *Logistics and Transportation Review*, Vol. 31, No. 3, 1995, pp. 239-251.
30. Sun, X., Brauner, E., and Hormby, S., "A Large Scale Neural Network for Airline Forecasting in Revenue Management", *Operations Research in the Airline Industry*, Kluwer Academic Publishers, Boston, US, 1997.
31. Fildes, R. and Lusk, E. J., "The Choice of a Forecasting Model", *OMEGA*, Vol. 12, No. 5, 1984, pp. 427-435.
32. Chatfield, C., *Time Series Forecasting*, Chapman & Hall/CRC, USA, 2000.
33. Rousseeuw, P. J. and Leroy, A. M., *Robust Regression & Outlier Detection*, John Wiley & Sons, USA, 1987.
34. Liu, L. M., Bhattacharyya, S., Sclove, S. L., and Chen, R., "Data Mining on Time Series: An Illustration Using Fast-food Restaurant Franchise Data", *Computation Statistics & Data Analysis*, Vol. 37, 2001, pp. 455-476.
35. Kumar, U. A., "Comparison of Neural Network and Regression Analysis: A New Insight", *Expert Systems with Application*, Vol. 29, 2005, pp. 424-430.
36. Haykin, S., *Neural Networks: A Comprehensive Foundation*, Prentice hall, USA, 1999.
37. Dougherty, M. S. and Cobbett, M. R., "Short-term Inter-urban Traffic Forecasts Using Neural Networks", *International Journal of Forecasting*, Vol. 13, No. 1, 1997, pp. 21-31.
38. Tsai, T. H., Lee, C. K., and Wei, C. H., "An Artificial Neural Networks Approach to Forecast Short-term Railway Passenger Demand", *Journal of the East Asia Society for Transportation Studies*, Vol. 4, 2003, pp. 221-235.
39. Lingras, P. and Mountford, P., "Time Delay Neural Networks Designed Using Genetic Algorithms for Short Term Inter-city Traffic Forecasting", IEA/AIE 2001, LNAI 2070, 2001, pp. 290-299.
40. Ku, C. C. and Lee, K. Y., "Diagonal Recurrent Neural Networks for Dynamic Systems Control", *IEEE Transactions on Neural Networks*, Vol. 6, No. 1, 1995, pp. 144-156.
41. Makridakis, S., Wheelwright, S. C., and Hyndman, R. J., *Forecasting: Methods and Applications*, Third Edition, John Wiley & Sons, Inc., USA, 1998.
42. DeLurgio, S. A., *Forecasting Principles and Applications*, McGraw-Hill, 1998.
43. Bazaraa, M. S., Sherali, H. D., and Shetty, C. M., *Nonlinear Programming: Theory and Algorithms*, Wiley, New York, 1993.
44. Hibon, M. and Evgeniou, T., "To Combine or Not to Combine: Selecting among Forecasts and Their Combinations", *International Journal of Forecasting*, Vol. 21, 2005, pp. 15-24.
45. Fildes, R., "The Evaluation of Extrapolative Forecasting Methods", *International Journal of Forecasting*, Vol. 8, 1992, pp. 81-98.

