國家發展研究　第九卷第二期
2010 年 6 月　　頁 1-26

歐美地區對台灣之旅遊需求預測*

邱鳳臨**、陳奕均***

收稿日期：2010 年 1 月 15 日　　接受日期：2010 年 3 月 3 日

* 作者感謝兩位匿名審查人的評論和建議。方嘉閩和高美菁的打字、資料蒐集和登錄協助，一併致謝。
** 臺灣大學國家發展研究所副教授。電話：3366-3326。MAIL: flchu@ntu.edu.tw。
*** 臺灣大學國家發展研究所碩士。

臺灣大學學術期刊資料庫

摘　要

回顧歷年來之來台旅客，雖皆以亞洲旅客居多，但有鑒於美、加及歐洲地區亦是來台旅遊之新興市場，同時也爲我國在觀光客倍增計畫中努力拓展來台旅遊之目標市場，故本篇研究將以美國、加拿大和歐洲爲研究對象，採用交通部觀光局所公佈的 1971 年 1 月至 2008 年 12 月之旅客入境人數年、季和月資料，分別對此三個地區建立簡單線性迴歸模型（linear trend）、Holt-Winters 趨勢的季節性加法預測模型、自我迴歸自我迴歸移動平均模型（autoregressive autoregressive moving average）、季節性與非季節性自我迴歸整合移動平均模型（seasonal-nonseasonal autoregressive integrated moving average）和部分整合自我迴歸移動平均模型（fractionally integrated autoregressive moving average）等五種模型，進行樣本內的預測，並使用平均絕對百分比誤差和均方根誤差統計量檢驗不同預測模型之預測能力。SARS（嚴重急性呼吸系統綜合症）爆發於 2003 年 3 月初，如果因爲這一事件產生一個結構突破，在假設整個期間是一個常數參數結構下得到的預測結果將受到懷疑。因此本文最後將作一項帶有結構破壞的單根檢驗之時間序列分析。

關鍵詞：旅遊、單根檢驗、預測、平均絕對百分比誤差

臺灣大學學術期刊資料庫

壹、研究動機與目的

世界觀光組織（World Tourism Organization；WTO）發佈最新資料顯示，2008 年全球仍有許多動亂不安的事件，然而在金融海嘯的影響下，全年全球國際觀光遊客人次仍超乎預期，達到 9 億 2,400 萬人次，較 2007 年成長 2%； 亞太地區 2008 年到訪旅客人次約爲 1 億 8,800 萬人次，與 2007 年相較成長 1.6%。台灣位於亞太地區的樞紐，具有先天上的地理優勢，更應掌握此全球主流旅遊觀光市場的移轉時機。根據世界旅遊及觀光委員會（World Travel & Tourism Council；WTTC 2009）的報告，2008 年全球旅遊產業提供約 2 億 2,500 萬個相關工作機會，占總就業的 9.3%；國際旅遊收入達到 9,440 億元，與 2007 年相較成長 1.8%。綜合上述，觀光產業不但能增加外匯，也能提供就業機會；而台灣在外部競爭模式改變，內部經濟結構尚未完全轉型、失業率居高不下的情況下，全力拓展觀光已是刻不容緩。

爲了掌握時機將台灣打造成最具有競爭力的國家，我國政府亦將發展觀光旅遊訂爲「挑戰 2008：國家發展重點計畫」之一項重要項目，並在 2002 年 8 月啓動「觀光客倍增計劃」，希望將台灣打造成爲「觀光之島」。另外行政院也在 2009 年通過一項觀光拔尖領航方案，增加旅客來台觀光的計畫，在該計畫下 300 億新台幣預算將應用於發展台灣成爲東亞主要旅遊中心，每年增加吸引 160 萬外國旅客，並將 2012 年目標人次訂爲 550 萬。因此，在這些目標之中，對未來之來台觀光旅遊人數的需求預測就顯得極爲重要，而如何建立一套系統化的研究方法以客觀推估未來旅客人次就構成了本研究的動機之一。其次，觀察歷年來之來台旅客，雖皆以亞洲爲主要來台旅遊的市場，然而在發展國際觀光之時，任何地區皆有可能是來台旅遊之潛在市場，而美國、加拿大和歐洲地區除了具有發展潛力外，亦是我國觀

臺灣大學學術期刊資料庫

光局努力拓展來台旅遊之目標市場。另外，針對這些地區之來台旅遊人數的需求預測也將有助於觀光資源之配置與投資，因此，掌握不同市場之旅遊人次脈動以供相關部門做爲決策依據就構成了本研究的動機之二。本研究將以美國、加拿大和歐洲爲研究對象，採用交通部觀光局所公佈的 1971 年 1 月至 2008 年 12 月之旅客入境人數年、季和月資料，分別對此三個地區建立五種時間序列模型，進行樣本內的預測，並使用平均絕對百分比誤差和均方根誤差統計量檢驗不同預測模型之預測能力，以選取美、加和歐洲各自之最適預測模型。如此方能彌補國內關於來台旅遊需求預測研究文獻之不足，並期望預測結果能提供相關單位和業者做爲觀光政策或觀光行銷策略制定和實施的參考。

貳、文獻探討

國外旅遊需求預測的實證研究文獻爲數相當龐大，約略可分爲採用迴歸分析的和純時間數列分析。在應用迴歸方面，Leob (1982) 探究實質所得、匯率和相對價格對旅遊輸出的影響。Uysal and Crompton (1985) 評述質化方法和量化方法並建議應用兩種方法的組合會產生較準確的預測。Malenberg and Van Soest (1996) 應用橫切面資科分析荷蘭家庭的年旅遊支出。Kulendran and King (1997) 比較用於預測美國、日本、英國和紐西蘭到澳洲的季旅遊流量的預測模型，結論是誤差修正模型（error correction model）表現不如時間序列模型。Law and Au (1999) 是第一位應用類神經網絡預測日本到香港的旅客人數。Song and Witt (2006) 應用向量自動迴歸模型預測到澳門的旅客人數。Athanasopoulos and Hyndman (2008) 應用帶有外生變數的變動狀態空間模型（innovation state space model）預測澳洲國內旅遊需求。

在時間序列分析架構下，Geurts and Ibrahim (1975) 應用夏威夷

臺灣大學學術期刊資料庫

旅遊資料比較 Box-Jenkins 方法和指數平滑預測法。Choy (1984) 應用時間序列迴歸預測旅客抵達亞太地區來檢驗預測技術的準確度。Martin and Witt (1989) 評估英、美、法和德國的境外旅遊預測表現，其結論是簡算法 I 產生最準確的一年期預測。Chan (1993) 應用若干個預測模型探討國際旅客抵達新加坡的數字，其發現波動性時間序列非線性迴歸模型（sine wave nonlinear regression model）比簡算法 I、簡算法 II、簡單線性和 ARIMA (2,1,2) 模型較優。Lim and McAleer (2002) 在模型選擇和執行預測之前，透過單根和季節單根檢定，分析國際旅遊時間序列是否穩定，繼而應用 Box-Jenkins ARIMA 模型估計香港、馬來西亞和新加坡抵達澳洲的旅客人數。Gil-Alina (2005) 應用季節性部份整合方法於抵達紐西蘭旅客次數的預測上；該文發現季節性模型在預測表現上勝過非季節性模型。Chen and Wang (2007) 應用支援向量迴歸方法（support vector regression）於中國的季旅遊需求上，實證研究顯示支援向量迴歸超越類神經網路和 ARIMA 的表現。Wang, Song, Witt and Wu (2007) 探討旅遊預測的組合是否能產生較準確的預測。Chu (2009) 應用 ARMA 爲基礎的模型於亞太地區九個國家的旅客人次的預測；實證研究顯示 ARMA 爲基礎的模型將 MAPE 值保持在 10 以下。

回顧國內過去針對來台旅客爲研究主題之文獻，大多是對總體來台旅客人數進行研究，對各別來台客源市場的旅遊需求預測之探討則大多侷限在日本，而過去雖有對美國和歐洲來台旅客進行需求預測的研究（如：時巧煒（1994）以簡算法、時間趨勢模型、指數平滑法、轉移函數、計量經濟模型、自我相關整合移動平均模型建立和評估美國和歐洲來台旅客需求預測模型，發現兩個地區皆以時間趨勢模型的預測能力最好；江麗文（1995）以計量經濟模型建立美國來台旅客需求模型；施瑞賢（2003）則替美國和歐洲以觀光爲目之來台觀光者建構自我相關整合移動平均模型），但卻沒有利用各種不同的時間序列

臺灣大學學術期刊資料庫

模型（尤其是自我迴歸自我迴歸移動平均模型和部分整合自我迴歸移動平均模型）建立預測模式並加以評估其預測能力之研究。除此以外，關於加拿大地區，國內則尚未有學者針對這個市場之來台旅遊需求進行預測。故本研究以美國、加拿大和歐洲地區之來台旅客爲研究對象，採用簡單線性迴歸模型（Linear Trend）、Holt-Winters 趨勢的季節性加法預測模型、自我迴歸自我迴歸移動平均模型（autoregressive autoregressive moving average, ARARMA 或 ARAR）、季節性與非季節性自我迴歸整合移動平均模型（seasonal-nonseasonal autoregressive integrated moving average, ARIMA 或 SARIMA）和部分整合自我迴歸移動平均模型（fractionally integrated autoregressive moving average, ARFIMA）等五種不同之時間序列模型，針對這三個來台市場建構此五種預測模型，並就各模型加以評估其預測能力，以選取美、加和歐洲各自之最適預測模型。文章結構安排如下：第一節爲研究動機與目的，第二節爲文獻探討，第三節爲模型設計，第四節爲資料來源與模型預測能力分析，第五節爲單根檢定，第六節爲討論與建議。

參、模型設計

一、簡單線性迴歸模型（Simple Linear Regression Model）

簡單線性迴規模型認爲序列 $\{F_t, t=1,2,\ldots,n\}$ 在每一期都會以固定的絕對數量成長，此模型以時間 t 做爲模型中的自變數，再將對應之各期之入境人數作爲因變數 F_t；並且假定實際的入境人數變化可以利用線性迴歸方程式表示入境人數發展的趨勢。因此，使用最小平方法（Least Squares Method）可求得一個迴歸方程式：

$$F_t = \alpha + \beta t + \varepsilon_t \tag{1}$$

臺灣大學學術期刊資料庫

其中，F_t 表示入境人數，在本研究中則表示在第 t 期入境台灣人數之觀察值。t 表示時間變量；α、β 爲迴歸式中待估計的參數。其次，因爲此模型不包含季節調整的功能，所以在進行模型參數估計之前，本研究將先就原始序列 $\{F_t, t=1,2,\ldots,n\}$ 進行乘積的移動平均之季節調整以得到 12 月（四季）之季節指數和一個新變數（YSA），再利用此一新變數（YSA）代替原迴歸式的應變量 F_t 建立迴歸模型。並且，在以此迴歸模型進行預測時，需將所得到之相對應的各個對應值分別乘以進行季節調整後所得到的季節指數，才能得到最後的預測值。如果殘差項 ε_t 有自相關，有幾種估計方法可以產生比最小平方法更有效（more efficient）的估計值（Cochrane and Orcutt, 1949；Hildreth and Lu, 1960；Durbin, 1960；Prais and Winsten, 1954；Magnus, 1978）。因爲這些估計方法結果沒有太大的差異，本文應用 Cochrane-Orcutt 重複法修正自相關問題。

二、自我迴歸自我迴歸移動平均模型（ARARMA 或 ARAR）

Parzen (1982) 建議一個 Box-Jenkins 方法的延續模型，稱其爲 ARARMA 或 ARAR，這是一項二階段評估程序；第一階段先將資料用於配適一個長期滯後自動迴歸模型（long-lag autoregressive model）：

$$\widetilde{F}_t = F_t - \alpha F_{t-s} + \varepsilon_t \tag{2}$$

其中 s 爲季節期數（用在月資料時，s=12），之後將第一階段估計產生的殘差項配適一個 ARMA 模型。這裡所用的 ARAR 運算法是以配適 AR 子集代替 ARMA 配適步驟，進行配適時趨向於配適比較高階的 AR 過程。例如：對著名的航空公司資料（airline data），Box-Jenkins 建議一個帶有 4 個估計參數的 ARIMA$(0,1,1)\times(0,1,1)_{12}$ 模型，Parzen 則應用 $\widetilde{F}_t = F_t - 1.0248F_{t-12}$ 作爲長期記憶轉換爲短期記憶的式子，隨後是帶有 13 個參數的 AR 濾波（filter）之短期記憶系列。詳細模型

臺灣大學學術期刊資料庫

可以參考 Parzen (1982)。

三、季節性與非季節性自我迴歸整合移動平均模型（ARIMA and SARIMA）

ARIMA 模型是由 Box-Jenkins (1976) 所設立的，此模型結構基本上包含了自我相關、移動平均及差分三個部份。其中差分過程主要是在將非穩定過程轉換成穩定的序列，而執行差分的次數則需視差分多少次始能將此非穩定過程轉換成穩定過程而定。其基本模式爲：

$$\phi_p(B)(1-B)^d F_t = \theta_q(B)\varepsilon_t \tag{3}$$

其中 p 表示自我相關階數，q 表示移動平均階數，d 則表示差分次數。除此以外，$\phi(B)$ 爲自我迴歸運算符號，$\theta(B)$ 爲移動平均運算符號，而 B 爲後移運算符號；而此模式穩定的條件爲 $\phi(B)=0$ 之根必須落在單位圓之外，且 $\theta(B)=0$ 之根必須落在單位圓之外。 同時 ARIMA 模型需要經過模型鑑定（model identification）、參數估計（parameter estimation）、診斷查驗（diagnostic checking）等三個步驟方能確立。（SARIMA）爲 ARIMA 模型進一步演化之預測模型，當時間序列具有週期性的規則變化時，需以季節性差分運算方式將季節因素消除，再配適 ARIMA 模式即成爲季節性與非季節性自我整合移動平均模型。模式可記爲 ARIMA$(p,d,q)(P,D,Q)_s$，其中 P 爲季節自我相關階數，Q 是季節移動平均階數，D 則爲季節差分階數， s 是季節週期長度。其基本模式如下：

$$\phi_p(B)\Phi_P(B^s)\nabla^d\nabla_s^D F_t = \theta_q(B)\Theta_Q(B^s)\varepsilon_t \tag{4}$$

其中 $\Phi_p(B^S)$ 爲季節性的 P 階自我相關運算符號，$\Theta_Q(B^S)$ 爲季節性的 Q 階移動平均運算符號，$\nabla^d\nabla_s^D$ 爲差分及週期 s 之季節差分的運算符號，故此種模型又被稱爲相乘季節型之自我相關整合移動平均模

臺灣大學學術期刊資料庫

型 （Multiplicative Models）。ARIMA 和 SARIMA 的預測表現在各類的時間序列裡皆甚穩定且非常準確，往往很難擊敗（Granger and Engle, 1987）。詳細模型可以參考 Box-Jenkins (1970)。

四、Holt-Winters 趨勢的季節性加法預測模型

此模型是簡單指數平滑法的一般化版本（Holt 1957; Winters 1961）。時間序列$\{F_t\}$是以下列形式存在：

$$F_t = T_t + S_t + W_t \tag{5}$$

其中T_t 、 S_t和W_t代表趨勢、季節和非規則因素。以G_t作爲時間點 t 的趨勢斜率。爲了預測未來值，趨勢、季節因子和趨勢斜率必須應用觀察值估計。設$\hat{T}_t$ 、 $\hat{S}_t$和$\hat{G}_t$爲T_t 、 S_t和G_t的估計值，在 Holt-Winters 方法中，這些估計值是透過下列三個平滑方程式獲得:

$$\hat{T}_t = \lambda_1(F_t - \hat{S}_{t-p}) + (1-\lambda_1)(\hat{T}_{t-1} + \hat{G}_{t-1}), \tag{6}$$

$$\hat{G}_t = \lambda_2(\hat{T}_t - \hat{T}_{t-1}) + (1-\lambda_2)\hat{G}_{t-1}, \tag{7}$$

$$\hat{S}_t = \lambda_3(F_t - \hat{T}_1) + (1-\lambda_3)\hat{S}_{t-p}, \tag{8}$$

其中λ_1 、 λ_2和λ_3是介於 0 與 1 之間的平滑參數。然後再把三個平滑結果用一個預測公式結合起來，進行預測。其模型爲：

$$\hat{F}_{t+h} = \hat{T}_n + h\hat{G}_n + \hat{S}_{n-p+j(h)} \tag{9}$$

(9) 式中 $j(h)$ 是整數使得$1 \le j(h) \le p$ ，p 代表季節期的正整數。Holt-Winters 方法的預測表現在若干實証研究皆有討論（Makridakis et al., 1984），其預測表現在各類的時間序列裡皆甚穩定且合理地準確。詳細模型可以參考 Holt (1957) 和 Winters (1961)。

臺灣大學學術期刊資料庫

五、部分整合自我迴歸移動平均模型（ARFIMA）

時間序列有一重要的特性就是記憶性（memory），此記憶性是指時間序列中本期的觀察值會受前一期的觀察值影響所顯現出的特性。Box and Jenkins (1970) 針對非定態序列提供一個解決模型，即整合自我迴歸移動平均（ARIMA）：

$$\phi_p(B)(1-B)^d F_t = \theta_q(B)\varepsilon_t \tag{3'}$$

其 d 是整合級數（integrated order）且限定爲整數。Granger and Joyeux (1980) 和 Hosking (1981) 分別獨立地提出部分整合自我迴歸移動平均（Fractional Integrated Autoregressive Moving Average, ARFIMA (p, d, q)）模型，其中 d 可爲任意實數，可用來解決具有高持續特性的時間序列模型。當 d<0.5 時，(3') 式爲一恆定過程，且可化爲無窮期移動平均多項式。而若 d>−0.5，該式爲一可逆過程，且可化爲無窮期自我迴歸多項式。因此當部分差分參數介於 0.5 及 −0.5 之間時，該模型爲恆定且可逆，其中，若 0<d<0.5，則稱時間序列有緩長記憶；若 −0.5<d<0，則稱過程具有中等記憶（intermediate memory）（Chung and Baillie, 1993）。Hosking (1981) 解釋所謂長期記憶（long memory）現象，即爲伴隨時間的增加自我相關係數遞減的速度會比自我迴歸移動平均模式來的緩慢，此即爲 ARFIMA 模式之特性。而 ARFIMA 模式現已被廣泛地用來分析存在長期記憶現象的時間序列資料。

其譜密度函數（spectral density）爲

$$f(\lambda) = \frac{\sigma^2}{2\pi}\frac{\left|\theta(e^{-i\lambda})\right|^2}{\left|\phi(e^{-i\lambda})\right|^2}\left|1-e^{-i\lambda}\right|^{-2d} \tag{10}$$

方程式 (10) 右側的第一項是 ARMA 模式的譜密度函數，第二項是長

臺灣大學學術期刊資料庫

期記憶的譜密度函數。估計 ARFIMA 的方法可以歸類爲時域（time domain）和頻域（frequency domain）；前者包括確切最大慨似估計法（exact maximum likelihood）和準最大慨似估計法（quasi maximum likelihood），這類估計方法可以參考 Granger and Joyeux (1980)、Sowell (1992) 和 Beran (1994) 等 。在頻域下包括 Whittle 估計方法和半參數估計方法（semi-parametric methods），這類方法可以參考 Fox and Taqqu (1986)、Dahlhaus (1989)、Giraitis and Surgailis (1990) 和 Robinson (1995) 等。應用最大概似法估計需要承受龐大的運算負擔，而對一個對數概似函數的 whittle 近似值 L_w 取最大化相對比較容易，其定義如下：

$$-2\ln(L_w) = n\ln(2\pi) + 2n\ln\sigma + \sigma^{-2}\sum_j \frac{I_n(\omega_j)}{g(\omega_j)} + \sum_j \ln g(\omega_j) \quad (11)$$

其中 I_n 爲序列的週期圖（periodogram），$\sigma^2 g/(2\pi)(=f)$ 爲 ARFIMA 的譜密度函數，$\sum_j$ 爲非零博立葉頻率 $\omega_j = 2\pi j/n \in (-\pi,\pi]$ 的總和。詳細模型可以參考 Chung and Baillie (1993) 和 Baillie (1996)。

肆、資料來源與模型預測能力分析

本研究所採用的資料爲依據交通部觀光局所公佈之「觀光月報」中的「來台旅客居住地分析」統計表，並以此表所統計之居住地爲美國、加拿大及歐洲地區的來台旅客入境人數年、季和月資料進行分析。資料範圍爲 1971 年 1 月至 2008 年 12 月。其次，爲了測試預測之準確性，本研究針對美國、加拿大及歐洲皆僅以 1971 年 1 月至 2006 年 12 月來台旅客人數年、季和月資料進行此五種時間序列模型之模型分析，並分別對美、加及歐洲就這五種不同的預測模型預測其於 2007 年 1 月至 2008 年 12 月之來台旅客預測值，共計 2 筆（8 筆季和 24 筆

臺灣大學學術期刊資料庫

月資料）預測值，然後再與 3 個地區各自之樣本內的實際值比較，如此方能進一步進行模型有效性之相關分析，以選取最適之預測模型。

經由模型預測出來的數值，與未來實際發生所得的數值不可能完全一致，這其間必定有偏差存在，這些偏差稱之爲預測誤差。而一般衡量模型預測能力之優劣就是利用預測誤差的大小來判斷。將這些預測值所產生之預測誤差加以比較，即可從中選取各來台市場之最適預測模型。 由於旅遊需求預測的相關文獻中大多以平均絕對百分比誤差（MAPE)和均方根誤差（RMSE）作爲衡量預測能力的指標（Makridakis et al., 1984），故在本研究中，將取 MAPE 和 RMSE 作爲判斷何種時間序列模型爲最佳預測模型的衡量工具。其計算方程式如下：

$$\text{MAPE} = \frac{1}{K}\sum_{t=N+1}^{N+K}\left|\frac{\hat{F}_t - F_t}{F_t}\right| \times 100 \tag{12}$$

$$\text{RMSE} = \sqrt{\frac{1}{K}\sum_{t=N+1}^{N+K}(\hat{F}_t - F_t)^2} \tag{13}$$

其中，F_t 代表實際值，$\hat{F}_t$ 代表預測值，K 代表預測的次數。由 MAPE 和RMSE 所表示的意義來看，若是 MAPE 和RMSE 統計量之值越小，表示利用該模型進行預測後所產生的預測誤差將越小。

根據上一節所敘述之 MAPE 和 RMSE 公式加以計算，可以得到之前所分析之五個不同的預測模型在美國、加拿大和歐洲等三個地區，其各自不同模型下的 MAPE 和 RMSE 統計量之值(表 1 至表 6)。其中，預測表現最佳的模型所計算的 MAPE 和 RMSE 值以粗體字顯示。 由表 1 和表 2 可以發現，在年預測上，從粗體字可以證實非季節性自我相關整合移動平均模型（ARIMA）的預測表現最佳（年資料沒有季節問題，因此這裡使用 ARIMA 而不用 SARIMA），亦即其所產生之 MAPE 和 RMSE 最小，並且其 MAPE 值小於 10，表示此

臺灣大學學術期刊資料庫

模型產生高度準確（highly accurate）的預測（Lewis, 1982）。其實，就美國、加拿大和歐洲而言，ARIMA 的 MAPE 值皆小於 2，顯示模型預測出來的數值，與實際發生所得的數值偏差非常小。在歐洲旅客的預測表現上，ARFIMA 和 Linear Trend 緊跟在 ARIMA 之後；而另兩種預測模型 ARAR 和 Holt-Winters 的表現比較不出色。Linear Trend 和 Holt-Winters 在預測加拿大旅客人次的表現僅次於 ARIMA；另外 ARFIMA 預測所產生的 MAPE 值亦維持在 2 上下。至於美國，除了 ARIMA 表現亮麗外，ARFIMA 和 ARAR 的預測表現亦在可接受的範圍以內。至於從 RMSE 指標觀察，除了歐洲的 2 年期預測輸給 ARFIMA 外，ARIMA 仍然保持最好的預測模型地位，這點可以從粗體數字確認。

表 1　美、加和歐洲地區各模型之 MAPE 比較表（年資料）

國家／模型	1 年期	2 年期
歐　洲		
Linear Trend	1.82	2.82
Holt-Winters	3.61	5.47
ARAR	5.53	3.29
ARIMA	**0.96**	**1.77**
ARFIMA	1.25	1.86
加拿大		
Linear Trend	1.97	2.14
Holt-Winters	1.30	1.83
ARAR	4.18	10.00
ARIMA	**0.38**	**0.54**
ARFIMA	2.24	2.05
美　國		
Linear Trend	4.91	8.91
Holt-Winters	2.91	6.31
ARAR	3.66	4.94
ARIMA	**1.06**	**1.41**
ARFIMA	1.52	4.42

臺灣大學學術期刊資料庫

表 2　美、加和歐洲地區各模型之 RMSE 比較表（年資料）

國家 / 模型	1 年期	2 年期
歐　洲		
Linear Trend	3,387	5,923
Holt-Winters	6,735	11,455
ARAR	10,328	7,456
ARIMA	**1,784**	3,890
ARFIMA	2,332	**3,886**
加拿大		
Linear Trend	1,131	1,268
Holt-Winters	742	1,137
ARAR	2,393	6,950
ARIMA	**219**	**332**
ARFIMA	1,281	1,205
美　國		
Linear Trend	19,557	38,027
Holt-Winters	11,563	28,005
ARAR	14,559	19,911
ARIMA	**3,994**	**5,682**
ARFIMA	6,051	20,503

從表 3 和表 4 明顯可以看出，在季旅遊人次的預測上，幾乎所有的 MAPE 值仍然維持在 10 以下，即此五種模型仍然有能力產生高度準確的預測。SARIMA 的表現已經不具領導地位，繼而興起的是 ARFIMA 預測模型。從粗體數字觀察之，ARFIMA 在預測歐洲的 2 和 4 季、美國的 6 和 8 季的來台旅客人次上表現最好；另外在預測加拿大的 2、4、6 和 8 季的來台旅客人次上表現也超越其它四種預測模型。ARAR 和 Holt-Winters 是表現僅次於 ARFIMA 的兩種預測模型。同樣的結論亦可以從觀察 RMSE 指標而獲得。最後，月旅客人次的預測結果顯示在表 5 和表 6。從 MAPE 指標觀察，ARFIMA 和 SARIMA 在預測能力上平分秋色，ARAR 則緊追在後，接下來是 Holt-Winters 和 Linear Trend。RMSE 指標則顯示 ARFIMA 比 SARIMA 有更好的表現。

臺灣大學學術期刊資料庫

表 3　美、加和歐洲地區各模型之 MAPE 比較表（季資料）

	MAPE 值			
國家 / 模型	2 季	4 季	6 季	8 季
歐　洲				
Linear Trend	5.85	8.54	9.07	10.90
Holt-Winters	3.14	2.70	3.86	4.51
ARAR	3.01	2.37	**1.83**	**2.05**
SARIMA	3.69	4.28	6.18	6.38
ARFIMA	**2.68**	**1.89**	3.11	3.23
加拿大				
Linear Trend	9.58	8.18	7.69	7.91
Holt-Winters	4.33	4.19	4.59	5.41
ARAR	6.77	6.66	8.07	9.08
SARIMA	4.61	3.64	4.28	4.58
ARFIMA	**2.97**	**2.49**	**3.21**	**3.47**
美　國				
Linear Trend	4.55	6.11	6.54	9.54
Holt-Winters	**3.84**	**3.53**	3.51	5.97
ARAR	9.98	5.51	5.33	9.18
SARIMA	5.23	4.57	4.63	7.02
ARFIMA	4.65	4.07	**3.29**	**5.28**

表 4　美、加和歐洲地區各模型之 RMSE 比較表（季資料）

	RMSE 值			
國家/模型	2 季	4 季	6 季	8 季
歐　洲				
Linear Trend	13,192	4,839	4,453	6,446
Holt-Winters	1,747	1,550	2,231	2,534
ARAR	1,808	1,429	**1,187**	**1,267**
SARIMA	1,743	2,425	3,587	3,561
ARFIMA	**1,429**	**1,132**	2,196	2,105
加拿大				
Linear Trend	1,428	1,342	1,268	1,388
Holt-Winters	599	609	686	845
ARAR	940	962	1,235	1,414
SARIMA	636	551	653	708
ARFIMA	**532**	**435**	**599**	**625**
美　國				
Linear Trend	4,785	6,444	7,521	11,039
Holt-Winters	**3,739**	**3,603**	**3,562**	7,143
ARAR	11,859	8,422	7,442	11,850
SARIMA	5,276	4,729	4,799	8,008
ARFIMA	4,913	4,269	3,633	**6,275**

臺灣大學學術期刊資料庫

表 5　美、加和歐洲地區各模型之 MAPE 比較表（月資料）

國家 / 模型	MAPE 值			
	6 個月	12 個月	18 個月	24 個月
歐　洲				
Linear Trend	9.72	9.77	9.79	16.78
Holt-Winters	5.37	4.68	6.09	5.78
ARAR	5.58	4.40	4.63	**4.38**
SARIMA	3.62	4.36	5.11	5.16
ARFIMA	**3.03**	**3.81**	**4.38**	4.49
加拿大				
Linear Trend	13.32	13.41	13.07	12.99
Holt-Winters	13.40	10.38	10.49	10.09
ARAR	**8.35**	8.36	8.38	8.60
SARIMA	9.19	8.10	**7.56**	7.26
ARFIMA	9.17	**7.93**	7.78	**6.61**
美　國				
Linear Trend	9.82	11.10	10.67	12.78
Holt-Winters	12.60	16.09	16.31	19.73
ARAR	**6.88**	5.88	7.72	9.17
SARIMA	7.74	**5.52**	**5.41**	**6.26**
ARFIMA	9.58	6.01	6.89	8.15

表 6　美、加和歐洲地區各模型之 RMSE 比較表（月資料）

國家 / 模型	RMSE 值			
	6 個月	12 個月	18 個月	24 個月
歐　洲				
Linear Trend	2,527	3,130	3,214	3,492
Holt-Winters	1,283	1,107	1,458	1,342
ARAR	1,146	975	1,009	937
SARIMA	734	905	1,056	1,043
ARFIMA	**621**	**809**	**912**	**926**
加拿大				
Linear Trend	668	761	693	814
Holt-Winters	630	527	547	544
ARAR	**493**	**483**	497	498
SARIMA	516	497	473	447
ARFIMA	525	489	**464**	**409**
美　國				
Linear Trend	3,550	4,018	3,987	4,762
Holt-Winters	5,001	6,096	6,142	7,250
ARAR	**2,777**	2,329	2,885	3,327
SARIMA	3,008	**2,295**	**2,169**	**2,412**
ARFIMA	3,618	2,666	2,958	3,254

臺灣大學學術期刊資料庫

從表1至表6可以看出,年旅客人次預測比季旅客人次預測準確，而季旅客預測又比月旅客人次預測準確，這是因爲一年裡的季（月）預測誤差之間有相互抵消的效應（就如同產業預測比個別公司預測更準確）。其次，表 1 至表 6 也呈現一種現象，即出現在第 2 年期（第 8 季和第 24 個月）的預測能力評估指標（MAPE 或 RMSE）皆稍微偏高，這是因爲 2008 年全球經歷了金融海嘯，某種程度上影響到美、加和歐洲，對台灣旅遊需求，這項衝擊尤其在第 8 季和第 24 個月更爲明顯。另外，儘管 Linear Trend 在季和月旅客預測上不如其它四種預測模型,但是在年旅客人次預測上,與其它四種模型的差異並不大。概括地說，經過 MAPE 和 RMSE 的分析後，在這兩項預測能力指標的架構下，在年預測上，以 ARIMA 的表現最好。至於季和月預測，則以 ARFIMA 的表現最佳。

伍、單根檢定

SARS（嚴重急性呼吸系統綜合症）爆發於 2003 年 3 月初，如果因爲這一事件產生一個結構破壞（structural break），在假設整個期間是一個常數參數結構下得到的預測結果將受到懷疑。因此本文最後將作一項帶有結構破壞的單根檢定之時間序列分析。帶有結構破壞的一項單根檢定的計量程序，牽涉到將樣本拆分成兩部份，然後應用 Dickey-Fuller 檢定於每一部分樣本（Dickey and Fuller, 1979）。這種程序導致每一個迴歸結果自由度遞減的難題。因此寧可用整組樣本做單獨一次檢定。然而,運用傳統的單根檢定如ADF或Phillips and Perron 的 Z 檢定常產生檢定力較低的問題（Phillips and Perron, 1988）。而且，當出現結構破壞時，Dickey-Fuller 和 Phillips-Perron 檢定偏向於不拒絕單根。本文將應用 Perron (1989) 所提出的帶有結構破壞的單根檢定。在此架構下，可以檢定單根過程中水平（level）的一次上升

臺灣大學學術期刊資料庫

（下降）之虛無假設對著趨勢平穩過程中截距（intercept）的一次改變之對立假設（對美國而言，其對立假設是趨勢平穩過程中截距和斜率的一次改變）。其迴歸模型如下：

加拿大和歐洲：$f_t = a_0 + \alpha_1 D_P + \alpha_2 D_L + a_2 t + a_1 f_{t-1} + \sum_{i=1}^{k} \beta_i \Delta f_{t-i} + \varepsilon_t$　(14)

美國：$f_t = a_0 + \alpha_1 D_P + \alpha_2 D_L + \alpha_3 D_T + a_2 t + a_1 f_{t-1} + \sum_{i=1}^{k} \beta_i \Delta f_{t-i} + \varepsilon_t$　(15)

其中，D_P(2003)=1，否則爲 0

$$D_L = D_T = 0 \text{，若 } t < 2003 \text{ 且 } D_L = 1 \text{、} \quad D_T = t \text{ 若 } t \geq 2003$$

在方程式 (14) 和 (15) 裡，D_P 代表一個跳動虛擬變量（pulse dummy variable）D_L 代表一個水平虛擬變量（level dummy variable），D_T 代表一個斜率虛擬變量（slope dummy variable）。f_t 表示旅客人次的對數值，滯後期（k 值）是應用 β_k 係數的 t 檢定來決定。SARS 事件爆發以後，美國旅遊序列的斜率有所改變，因此，另外估計帶斜率虛擬變量的方程式 (15)。對 (14) 式而言，虛無假設是 $a_0 = 0$ 、 $a_2 = 0$ 、 $a_1 = 1$ 和 $\alpha_1 = 0$，對立假設是 $a_0 \neq 0$ 、 $a_2 \neq 0$ 、 $a_1 < 1$ 和 $\alpha_1 \neq 0$。對 (15) 式而言，虛無假設是 $a_0 = 0$ 、 $a_2 = 0$ 、 $a_1 = 1$ 和 $\alpha_3 = 0$ ，對立假設是 $a_0 \neq 0$ 、 $a_2 \neq 0$ 、 $a_1 < 1$ 、 $\alpha_1 \neq 0$ 和 $\alpha_3 \neq 0$ 。迴歸估計結果如表 7 和 8：

表 7　帶有結構破壞的單根檢驗結果

國　家	λ	T	k	a_0	α_1	α_2	a_2	a_1
加拿大	0.87	38	2	3.39	-0.332	0.084	0.017	0.632
				(3.82)	(-4.26)	(1.83)	(2.82)	(-6.22)***
歐　洲	0.87	38	0	2.85	-0.304	0.003	0.012	0.733
				(2.56)	(-3.44)	(0.53)	(1.74)	(-6.71)***

說明：(1) 適當的 t 值表示在的括號內；***、** 和 * 分別是 1%、2.5% 和 5% 的顯著水準。
(2) λ 值代表發生結構破壞之前的樣本相對於總樣本之比率。
(3) 檢定 a_1=1 的 Perron t 統計量臨界值是：λ=0.80 分別是 -4.33（1% 的顯著水準）、-3.99（2.5% 的顯著水準）和 -3.75（5% 的顯著水準），λ=0.90 分別是 -4.27（1% 的顯著水準）、-3.97（2.5% 的顯著水準）和 -3.69（5% 的顯著水準），本文樣本是 38，λ=0.87。

臺灣大學學術期刊資料庫

表 8　帶有結構破壞的單根檢驗結果

國　家	λ	T	k	a_0	α_1	α_2	α_3	a_2	a_1
美　國	0.87	38	0	5.41	-0.469	2.02	-0.057	0.019	0.533
				(3.53)	(-4.27)	(2.40)	(-2.43)	(3.44)	(-4.00)**

說明：(1) 適當的 t 值表示在的括號內；***、** 和 * 分別是 1%、2.5% 和 5% 的顯著水準。
(2) λ 值代表發生結構破壞之前的樣本相對於總樣本之比率。
(3) 檢定 a_1=1 的 Perron t 統計量臨界值是：λ=0.80 分別是 -4.70（1% 的顯著水準）、-4.31（2.5% 的顯著水準）和 -4.04（5% 的顯著水準），λ=0.90 分別是 -4.41（1% 的顯著水準）、-4.10（2.5% 的顯著水準）和 -3.80（5% 的顯著水準），本文樣本是 38，λ=0.87。

表 7 和 8 的最後一欄明顯地顯示單根假設的支持度是不足的；所有 a_1 估計值皆與 1 顯著不同。對所有國家，我們可以在 5% 或較好的水平下拒絕單根。我們可以在 1% 的水平下拒絕加拿大和歐洲的旅遊序列；也可以在 5% 的水平下拒絕美國的旅遊序列。表 7 顯示，加拿大和歐洲的常數項（ a_0 ）、破壞後常數虛擬變量（ α_1 ）和趨勢（ a_2 ）的估計係數都顯著的。表 8 顯示，美國的常數項（ a_0 ）、破壞後常數虛擬變量（ α_1 ）、破壞後斜率虛擬變量（ α_3 ）和趨勢（ a_2 ）的估計係數都顯著的。綜合上述，從各項點估計值結果觀察，似乎比較傾向不拒絕數列帶有結構破壞的趨勢平穩的對立假設。資料似乎支持美、加和歐洲的旅遊序列是趨勢平穩的，儘管有 SARS 事件促使一次的結構破壞。以上檢定結果也可以支持 2007～2008 年之間的資料結構與特性與建立模型的資料之間是沒有顯著差異的。如此，本文前面所作預測的結果是可靠的。

陸、討論與建議

本文以美國、加拿大和歐洲三個市場之入境旅客人次作爲研究對象，採用五種時間序列模型進行分析。首先，對三個來台市場建立不

臺灣大學學術期刊資料庫

同的預測模型，並進行樣本內的預測。然後，以 MAPE 和 RMSE 統計量來分析三個地區在不同模型下的預測誤差，藉此選擇美國、加拿大和歐洲地區各自之最佳預測模型。爲了預測工作達到可靠性（reliability），本文使用年，季和月資料執行預測。在經過這些項目的分析後，發現三個來台市場在年資料方面 ARIMA 模型的預測能力最佳； 在季和月資料方面，則 ARFIMA 模型的預測能力最好。因此，建議相關單位及旅遊業者從事旅遊預測時可以應用此兩種模式。

臺灣在 2003 年爆發 SARS 事件，使實證所用的資料可能產生結構破壞，爲了預測結果不受到懷疑，本文也做一項帶有結構破壞的單根檢定之時間序列分析；檢定結果顯示不拒絕數列帶有結構破壞的趨勢穩定的對立假設，也因此確認本文所作預測的結果是可靠的。另外，本文研究結果也證實年預測比季預測和月預測更準確，因此建議間接應用年預測去推算季和月預測，這個過程需要借用穩定可靠的旅遊相關資訊，這個工作也是我們將來努力的方向。

臺灣大學學術期刊資料庫

參考文獻

中華民國交通部觀光局（1971～2008）。《台灣地區觀光統計月報》。台北：交通部觀光局。

中華民國交通部觀光局（2001）。《交通部觀光局成立三十週年特刊》。台北：交通部觀光局。

中華民國交通部觀光局（2001）。《觀光政策白皮書》。台北：交通部觀光局。

中華民國交通部觀光局（2002）。《觀光客倍增計畫》。台北：交通部觀光局。

中華民國行政院（2009）。《六大新興產業發展規劃—觀光拔尖領航方案》。台北：交通部觀光局。

江麗文（1995）。《來華旅客需求計量經濟模式之研究》。台北：中國文化大學觀光事業研究所碩士論文。

施瑞賢（2003）。《來華觀光旅客人數需求預測之研究》。台中：朝陽科技大學休閒事業管理研究所碩士論文。

時巧煒（1994）。《來華觀光旅客需求預測模式建立之研究》。台北：政治大學統計學研究所碩士論文。

Athanasopoulos, G. & Hyndman, R., 2008. Modeling and forecasting Australia domestic tourism. *Tourism Management,* 29, 19-31.

Baillie, R. T., 1996. Long memory processes and fractional integration in econometrics, *Journal of Econometrics*, 73, 5-59.

Beran, J., 1994. On a class of *M*-estimators for Gaussian long-memory models. *Biometrika*, 81(4), 755-766.

Box, G. E. P., & Jenkins, G. M., 1970. *Time series Analysis: forecasting and control*. San Francisco: Holden Day.

臺灣大學學術期刊資料庫

Chen, K. & Wang, C., 2007. Support vector regression with genetic algorithms in forecasting tourism demand. *Tourism Management*, 28, 515-526.

Chan, Y. M., 1993. Forecasting tourism: a sine wave time series regression approach. *Journal of Travel Research*, 32, 58-60.

Choy, D. L., 1984. Forecasting tourism revisited. *Tourism Management*, 5, 171-176.

Chu, F. L., 2009. Forecasting tourism demand with ARMA-based methods. *Tourism Management*, 30, 740-751.

Chung, C. F. & Baillie, R. T., 1993. Small sample bias in conditional sum-of-squares estimators of fractionally integrated ARMA methods. *Empirical Economics*, 18, 791-806.

Cochrane, D. & Orcutt, G. H., 1949. Application of least squares regressions to relationships containing autocorrelated error terms. *Journal of American Statistical Association*, 44, 32-61.

Crouch, I. G., 1994. The study of international tourism demand: A survey of practice. *Journal of Travel Research*, 32, 41-55.

Crouch, G. I., Schultz, L., & Valerio, P., 1992. Marketing international tourism to Australia: A regression analysis. *Tourism Management*, 13, 196-208.

Dahlhaus, R., 1989. Efficient parameter estimation for self-similar processes. *Annals of Statistics 17,* 1749-66.

Dickey, D. & Fuller, W., 1979. Distribution of the estimates for autoregressive time series with a unit root. *Journal of American Statistical Association*, 74, 427-431.

Durbin, J., 1960. Estimation of parameters in time series regression models. *Journal of Royal Statistical Society*, Series B, 22, 139-153.

臺灣大學學術期刊資料庫

Fox, R., & Taqqu, M. S., 1986. Large sample properties of parameter estimates for strongly dependent stationary Gaussian time series. *Annals of Statistics* 14, 517-532.

Geurts, M. & Ibrahim I. B., 1975. Comparing the Box-Jenkins approach with the exponentially smoothed forecasting model: Application to Hawaii tourists. *Journal of Marketing Research*, 12, 182-188.

Gil-Alana, L. A., 2005. Modeling international monthly arrivals using seasonal univariate long-memory processes. *Tourism Management*, 26, 867-878.

Giraitis, L. & Surgailis, D., 1990. A central limit theorem for quadratic forms in strongly dependent linear variables and its application to asymptotical normality of Whittle's estimate. *Probability Theory Related Fields*, 86(1), 87-104.

Granger, C. W. J., Joyeux, R., 1980. An introduction to long-memory time series models and fractional differencing. *Journal of Time Series Analysis*, 1, 15-39.

Granger, C, W. J. & Engle, R. F., 1987. Econometric forecasting: A brief survey of current and future techniques. In K. Land & S. Schneider (eds.), *Forecasting in the Social and Natural Sciences*, pp. 117-140. Dordrecht: D. Reidel Publishing.

Hildreth, C. & Lu, J. Y., 1960. Demand relations with autocorrelated disturbances. *Technical Bulletin* 276, Michigan State University Agricultural Station.

Holt, C. C., 1957. *Forecasting seasonals and trends by exponentially weighted moving averages*, Pittsburgh PA: Carnegie Institute of Technology.

Hosking, J. R. M., 1981. Fractional differencing. *Biometrika*, 68, 165-176.

臺灣大學學術期刊資料庫

Kulendran, N. & King M. L., 1997. Forecasting international quarterly tourist flows using error-correction and time-series models. *International Journal of Forecasting*, 13, 319-327.

Leob, P. D., 1982. International travel to the United States: An econometric evaluation. *Annals of Tourism Research*, 9, 7-20.

Lewis, C. D., 1982. *Industrial and business forecasting methods.* London: Butterworths.

Lim, C., & McAleer, M., 2002. Time series forecasts of international travel demand for Australia. *Tourism Management*, 23, 389-396.

Magnus, J. R., 1978. Maximum likelihood estimation of the GLS model with unknown parameters in the disturbance covariance matrix. *Journal of Econometrics*, 7, 305-306.

Makridakis, S., Anderson, A., Carbone, R., Fildes, R., Hibon, M., Lewandowske, R., Newton, J., Parzen, E. (Eds.), 1984. *The accuracy of major forecasting procedure.* New York: John Wiley.

Martin, C. A. & Witt, S. F., 1989. Forecasting tourism demand: A comparison of the accuracy of several quantitative methods. *International Journal of Forecasting*, 5, 7-19.

Malenberg, B., & Van Soest, A., 1996. Parametric and semiparametric modeling of vacation expenditures. *Journal of Applied econometrics*, 11, 59-76.

Parzen, E., 1982. ARARMA models for time series analysis and forecasting. *Journal of Forecasting*, 1, 67-82.

Perron, P., 1989. The great crash, the oil price shock and the unit root hypothesis. *Econometrica*, 57, 1361-1401.

Phillips, P. & Perron, P., 1988. Testing for a unit root in time series regression. *Biometrica*, 75, 335-346.

臺灣大學學術期刊資料庫

Prais, S. J. & Winsten, C. B., 1954. Trend estimators and serial correlation. *Cowles commission Discussion Paper No. 383*, Chicago.

Robinson, P. M., 1995. Gaussian semiparametric estimation of long range dependence. Ann. Statist., 23(5): 1630-1661.

Song, H. & Witt, S. F., 2006. Forecasting international tourist flows to Macau. *Tourism Management*, 27, 214-224.

Sowell, F., 1992. Maximum likelihood estimation of stationary univariate fractionally integrated time series models. *Journal of Econometrics*, 53, 165-188.

Uysal, M. & Crompton, J., 1985. An overview of approaches used to forecast tourism demand. *Journal of Travel Research*, 23, 7-15.

Winters, P. R., 1960. Forecasting Sales by Exponentially Weighted Moving Averages, *Management Science*, 6: 324-342.

Wong, K. F., Song, H., Witt, S. F. & Wu, D., 2007. Tourism forecasting: To combine or not to combine. *Tourism Management,* 28(4), 1068-1078.

World Tourism Organization website: *http://www.world-tourism.org*

World Travel & Tourism Council website: http://www.wttc.org

臺灣大學學術期刊資料庫

Forecasting Tourism Demand for Taiwan from the US, Canada and Europe

Fong-Lin Chu & Yi-Chun Chen

Abstract

This paper examines five time-series forecasting models and fits them to the datasets of yearly, quarterly and monthly visitor arrivals in Taiwan from the United States, Canada and Europe. We then use the resulting estimated models to make 'out-of-sample' forecasts over the next 2 years (8 quarters and 24 months). We find that forecasts generated using the autoregressive integrated moving average and fractionally integrated autoregressive moving average models provide more accurate forecasts than those from other models when judged on the basis of mean absolute percentage errors and root mean square errors criteria. If there has been a structure break because of SARS outbreak in early march 2003, the result obtained by assuming a constant parameter structure during the entire period will be suspect. An analysis of the time series, including unit root tests with structural break is carried out as well.

Keywords: Tourism, forecast, unit roots test, mean absolute percentage errors.

臺灣大學學術期刊資料庫