時空數列模式在訂位保留比率控制之應用

吳　柏　林*

摘　要

近年來我國經濟發展快速，台北已經成爲亞洲地區重要的貿易中心和觀光據點。頻繁的商務往來以及爲數頗多的觀光客，使得台北市區的觀光飯店和國內航線的班機出現一位難求的情況。若將多數房間或機位供旅客預訂，很容易發生顧客臨時退房或退票的狀況，而造成公司的損失。如果能夠事先預測出顧客人數，公司就能提出適當比例的客房和機位供旅客預訂，使公司獲得較大收益。

本研究將針對飯店和航空公司，利用過去住宿人數與搭機人數的資料，以時空數列的STARMA模式的方法，建立一個模式，找出應將多少比例的房間或機位提供給顧客做預定保留，以不免不必要的損失。

關鍵字：Space-time ARMA模式，到達率，保留比率，預測。

壹、前　言

單變量時間數列模式的發展，將動態資料分析的技術向前推進一大步。很多各種短期的預測均可利用單變量ARIMA：轉換函數；多變量VARMA，來求得預測模式。其主要原因

*作者爲本校應用數學所教授

在於：傳統所使用的迴歸模型只能考慮獨立樣本的資料，而ARIMA時間數列模型將樣本間的自相關性列入考慮，使得一些與前期有密切關係的資料可以得到較佳的配適（fitting）。

但是，這只是考慮到過去某一特定點或在空間中某一區域……等的時間序列，卻忽略了其他對等空間資料的影響，及其空間之交互影響權數問題。因爲在實務上，許多資料是以等時距的方式蒐集，這些資料發生的位置也是一個重要的因素。因此時空數列模式（spatial time series）之探討，對具有空間相關特性之資料相當重要。單變量ARIMA模型主要概念在於來自資料本身自相關（autocorrelation）與偏自相關（partial autocorrelation）的關係。時空數列除了時間因素之外，更强調鄰近地區同一類觀察值的影響。

時空數列模式分析乃同時處理N個區域的時間數列問題。對來自於N 個區域相關的資料的問題，將此N個區域彼此的空間關係列入模式中考慮。時空數列最大的特色，就是以加權矩陣（weighting matrix）描述各個位置（site）間的空間關係。在不增加變數的情況下，利用現有的資訊，建構更合適的模式，以提高預測準確度。

時空數列模式描述地區與地區之間的時間和空間動態關係，若觀察值在鄰近不同區域之間呈現出某種系統性的相關，此種現象稱爲「空間相關」（spatial correlation）。由於時間數列僅分析變數之間自迴歸及時間相關的關係，並沒有考慮到這些資料蒐集的位置對研究結果產生的影響。時空數列模式即是企圖解釋這（weighting matrix ）來描述各個位置（site）間的空間關係。

時空數列模式在物理和環境科學的領域應用較多，但是社會、經濟方面的問題也有一些學者使用時空數列分析，如：考慮雨量對河水流量的影響、研究波士頓地區犯罪率的情況（Deutsch and Pfeifer, 1980a）；探討旅遊人數變動的趨勢（Stoffer, 1986）；在法國對傳染病的調查（Flahault, 1988）；探討旅遊人數變動的趨勢（Pfeifer and Bodily, 1990）；研究波士頓地區犯罪率的情況（Deutsch and Pfeifer, 1980b）；台北地區空氣污染指標分析（吳柏林與廖敏治，1993）及台灣地區失業率的時空數列分析（吳柏林與陳雅玫，1994）等。

加權矩陣在時空數列模式分析中扮演相當重要的角色。它可視爲建立模式者事前對於各個地區變數之間的關係程度，這也是時空數列模式不同於多變量時間數列（Multivariate time series）的地方。這些地區變數之間的關係程度，包括地緣的交互影響與時間的落差。

貳、時空數列模型分析

2.1 時空數列模式STARMA模式的理論與架構

假設空間中有N個固定位置，Z_t為在時間t時，在N個定位上分別得到的觀察值所組成的（N×1）向量

$$Z_t=\begin{pmatrix} Z_1(t) \\ Z_2(t) \\ \vdots \\ Z_N(t) \end{pmatrix}, t=1,2,\ldots,T$$

則Z_t的STARMA模式可表示為：

$$Z_t=\sum_{k=1}^{\rho}\sum_{\ell=0}^{\lambda_K}\phi_{k\ell}W^{(\ell)}Z_{t-k}-\sum_{k=1}^{q}\sum_{\ell=0}^{m_K}\theta_{k\ell}W^{(\ell)}a_{t-k}+a_t \qquad (2.1)$$

記做STARMA（P_{λ_1}，λ_2，…λ_p，q_{m_1,m_2}，…m_p）model

其中：

p ：自我迴歸（AR）階數，

q ：移動平均（MA）階數，

λ_k ：第K個自我迴歸項的空間階數，

m_k ：第K個移動平均項的空間階數，

$\phi_{k\ell}$：時間落後K期、空間落後ℓ期的自我迴歸項係數，

$\theta_{k\ell}$：時間落後K期、空間落後ℓ期的移動平均項係數，

$W^{(1)}$：空間階數為l的加權矩陣，通式為

$$W^{(\ell)}=\begin{pmatrix} W_{11}^{(\ell)} & W_{12}^{(\ell)} & \cdots & W_{1N}^{(\ell)} \\ W_{21}^{(\ell)} & W_{22}^{(\ell)} & \cdots & W_{2N}^{(\ell)} \\ \vdots & \vdots & \ddots & \\ W_{N1}^{(\ell)} & W_{N2}^{(\ell)} & \cdots & W_{NN}^{(\ell)} \end{pmatrix}$$

且$W^{(0)}=I_N$，

a_t　：時間t時的隨機誤差項，服從多變量常態分配，且具有下列性質：

$$E(a_t\, a'_{t+s}) = \begin{cases} G & ,s=0 \\ O & ,s\neq 0' \end{cases}$$

$$E(a_t\, a'_{t+s}) = O，\ \forall\ s>0$$

本文為了簡化問題與計算方便，假設$G=\sigma^2 I_N$。

2.2 空間階數與加權矩陣

時空數列模式架構過程中，首先決定每個位置與其相鄰的空間階數（spatila order）vo加權矩陣（weighting matrix）。空間階數和加權矩陣的決定視實際需要由模式建立者依照資料的空間關係與研究主題的性質來決定。

(a)決定空間階數

Besag（1974）提出，如何在規則的空間系統（regularly spaced system）中訂定空間階數。即在規則的空間系統中，對特定位置與其鄰近位置，決定一度空間的線性系統及二度空間的格狀系統空間階數的例子。在空間階數決定後，對於具有相同空間階數者，依據等值加權法（Equal Scaled Weighting）定義加權矩陣，以下將會有詳細的說明。至於不規則的空間系統，可利用上述的定義方式再稍作修改，即對特定位置，只考慮其與鄰近位置的距離，對於距離相近者給予相同之空間階數。

此外，在分析非空間系統的問題時，〝位置”與〝空間關係”應另行給予特殊且適當的定義。例如：〝位置”可代表不同的物品，各個物品間相互替代與互補的情況則視為〝空間關係”。此時空間階數通常都定義為1。

(b)定義加權矩陣

決定空間階數之後，接著要定義加權矩陣，一般使用的方法有下列幾種：

1. 規則的空間系統：以等值加權法決定權數

在規則的空間系統中，Besag（1974）對於空間階數相等者給予相同之權數，稱爲等值加權（Equal Scaled Weighting）。其基本原則爲：與一特定位置有相同距離者，定義相同之空間階數；而空間階數相等者給予相同之權數，其表示方式如下

$$W_{ij}^{(\ell)}=\begin{cases}\dfrac{1}{n_i^{(\ell)}} & \text{若位置j相對於位置i的空間階數爲l}\\ 0 & \text{其它}\end{cases}$$

其中$W_{ij}^{(\ell)}$爲$W^{(\ell)}$中第（i, j）個元素

$n_i^{(\ell)}$爲相對於位置i，其空間階數爲ℓ的位置個數

因爲等值加權的緣故，$W^{(\ell)}$中不全爲0的列滿足

$$\sum_{j=1}^{N} W_{ij}^{(\ell)}=1$$

且加權矩陣W不對稱。

2. 不規則的空間系統：以距離倒數的函數作爲權數

此處「距離」的定義，除了一般所使用的兩個區域中心點的直線距離之外，其它較常用的定義方式還包括：兩區域相連的長度，兩區域的交通網路密度……等，由研究者依其需要自行選定。若以研究兩國之間的空間關係爲例，共同邊界的長度，兩國間聯絡道路的數目及天然障礙（河流、高山……）都可列入考慮，見Cliff and Ord（1981）。

事實上，絕大多數的空間系統都是不規則的，基於距離愈近者影響愈大之原則，以兩位置間的距離倒數的函數作爲權數，距離愈近者其權數愈大。若以dij表示位置i與位置j之間的距離，較常用的函數有下列數種：

$$\text{(a) } W_{ij}^{(\ell)}=\frac{1}{d_{ij}}$$

$$\text{(b) } W_{ij}^{(\ell)}=\frac{1/d_{ij}}{\sum_{j} 1/d_{ij}}$$

(c) $W_{ij}^{(l)}=\dfrac{1/d_{ij}}{(d_{ij}+c)^{\alpha}}$，C>0，$\alpha>0$

這樣的定義方式，可能會造成 $\sum_{j=1}^{N} W^{(l)} \neq 1$，但W對稱。

3. 抽象空間系統：由資料的屬性之關係決定權數

對於一些特殊情況，由於研究內容並非眞實的空間系統，其空間關係無法以距離表示。例如研究不同品牌的同類商品在市場上的替代性，我們可以透過問卷調查的方式了解消費者的偏好，以代替眞實空間系統的空間關係。將各品牌在消費者心目中的地位轉換成適當的權數，據以決定加權矩陣。如此可擴大時空數列模式的應用範圍，使其延伸到社會和經濟的領域，而不再侷限於物理和環境科學。

例如：

以$W_{ij}^{(l)}$代表消費者買不到第i種品牌的商品時，願意以第j種品牌替代的比例。因爲對商品偏好的程度不同，這樣的定義方式可能會出現$W_{ij}^{(l)} \neq W_{ij}^{(l)}$，$i \neq j$的情況，即W不對稱。因爲$W_{ij}^{(l)}$爲機率值，$\sum_{j=1}^{N} W_{ij}^{(l)}=1$。

$$\begin{cases} W_{ii}^{(l)}=0 \\ \sum_{j=1}^{n} W_{ij}^{(l)}=1 \end{cases}$$

由於擴散有方向性的緣故，如同非空間系統一樣，$W_{ij}^{(l)} \neq W_{ji}^{(l)}$，$i \neq j$，是被允許的。

其它如：流行病的傳染、海洋中魚群的移動、公路網交通流量的分析，亦可用擴散原理來定義加權矩陣。

最後，値得一提的是除了規則的空間系統之外，其餘三種類型對於空間階數的選取並沒有一定的規則，而是由研究者依其對問題的了解程度自行選取適當的階數。一般來說，爲了避免將眞正的空間關係複雜化，空間階數最大取到2已經足夠。同時，只有在考慮機率的情況下，才會限制 $\sum_{j=1}^{n} W_{ij}^{(l)}=1$。通常 $W_{ij}^{(l)}$只要能適當表示出位置i與位置j之間，互相影響的程度即可，因此也允許 $W_{ij}^{(l)} \neq W_{ji}^{(l)}$ 的情形出現。

參、建立模式

時間數列模式的建構過程，大都採用Rox & Jenkins（1976）所提出的三階段程序建立。而時空數列模式的建構過程所使用的方法，在認定過程中相當複雜。Deutsch and Pfeifer（1980a）曾提出模式階數認定方法，推廣Box & Jenkins的模式的建構概念。本研究則對Deutsch and Pfeifer所提出模式階數認定方法進一步修正與改良。

3.1 模式階數認定（Order Identification）

當我們決定使用STARMA模式配適資料時，首先遭遇到的問題，便是模式中的參數（$p, q, \lambda_1, \cdots, \lambda_p, m_1, \cdots, m_q$）應如何決定。第一個步驟—確認模式，我們利用樣本的時空自我編相關函數STACF（space-time autocorrelation function）及時空自我編相關函數STPACF（space- time partial autocorrelation function），來探討這個問題。

(a)計算STACF

設$\Gamma(s) = E[Z(t)Z(t+s)]$，時空共變異函數（Space-time covariance function）$\gamma_{lk}(s)$可表示爲

$$\gamma_{lk}(s) = E\left\{\frac{[W^{(l)}Z(t)]'[W^{(k)}Z(t+s)]}{N}\right\} \tag{3.1}$$

$$= tr\left\{\frac{W^{(k)1}W^{(l)}\Gamma(s)}{N}\right\}$$

我們可將STACF定義爲

$$\rho_{lk}(s) = \frac{\gamma_{lk}(s)}{[\gamma_{ll}(0)\gamma_{kk}(0)]^{1/2}} \tag{3.2}$$

（3.2）式爲時間落後S期時，空間階數ℓ與k的自我相關函數。

通常我們以

$$\hat{\Gamma}(s)=\sum_{t=1}^{T-s}\frac{z(t)z(t+s)'}{T-s} \tag{3.3}$$

估計$\Gamma(s)$，因此（3.1）式中$\gamma_{\ell k}(s)$的估計式爲

$$\hat{\gamma}_{\ell k}(s)=\frac{\sum_{i=1}^{N}\sum_{t=1}^{T-s}[\sum_{j=1}^{N}W_{ij}^{(\ell)}Z_j(t)][\sum_{j=1}^{N}W_{ij}^{(k)}Z_j(t+s)]}{N(T-s)} \tag{3.4}$$

則樣本的時空自我相關函數（sample STACF）定義爲

$$\hat{\rho}_{\ell k}(s)=\frac{\hat{\gamma}_{\ell k}(s)}{[\hat{\gamma}_{\ell\ell}(0)\hat{\gamma}_{kk}(0)]^{1/2}} \tag{3.5}$$

(b)計算STPACF

ATPACF是由Yule-Walker方程式來決定，其一般式如下：

$$\gamma_{ho}(s)=\sum_{j=1}^{\rho}\sum_{l=0}^{\lambda_p}\phi_{j\ell}\gamma_{h\ell}(s-j) \tag{3.6}$$

其中$s=1,2\cdots\cdots k$；$h=0,1\cdots\cdots\lambda$，構成一組聯立方程式，以矩陣方式表示成

$$\begin{pmatrix}\gamma(1)\\ \gamma(2)\\ \vdots\\ \gamma(k)\end{pmatrix}\begin{pmatrix}\Gamma(0) & \Gamma(-1) & \cdots & \Gamma(1-k)\\ \Gamma(1) & \Gamma(0) & \cdots & \Gamma(2-k)\\ \vdots & \vdots & & \vdots\\ \Gamma(k-1) & \Gamma(k-2) & \cdots & \Gamma(0)\end{pmatrix}\begin{pmatrix}\phi_1\\ \phi_1\\ \vdots\\ \phi_k\end{pmatrix}$$

其中

$$\gamma(k)=\begin{pmatrix}\gamma_{00}(k)\\ \gamma_{10}(k)\\ \vdots\\ \gamma_{\lambda 0}(k)\end{pmatrix},\qquad \phi_k=\begin{pmatrix}\phi_{k0}\\ \phi_{k1}\\ \vdots\\ \phi_{k\lambda}\end{pmatrix} \tag{3.7}$$

$$\Gamma(k)=\begin{pmatrix}\gamma_{00}(k) & \gamma_{01}(k) & \cdots & \gamma_{0\lambda}(k)\\ \gamma_{10}(k) & \gamma_{11}(k) & \cdots & \gamma_{1\lambda}(k)\\ \vdots & \vdots & & \vdots\\ \gamma_{\lambda 0}(k) & \gamma_{\lambda 1}(k) & \cdots & \gamma_{\lambda\lambda}(k)\end{pmatrix}$$

將（3.7）式中的ϕ_{sl},$s=1,2\cdots\cdots,k$；$l=0,1,\cdots\cdots,\lambda$解出，即爲樣本的時空自我偏相關函數（sample STPACF）。

由（3.5）式及（3.7）式將樣本的STACF和STPACF求出後，即可參考表3.1決定時間及空間階數，找出一個暫時適合的模式，然後進行第二個步驟—估計參數。

最後要特別强調的是，當我們在計算STPACF時，算式中的λ爲事先所決定空間階數的最大值。Deutsch與Pfeifer（1980a）建議，在一般情況下λ值取到2階已經足夠，若討論的空間系統較大時，可將λ增加到3階或4階。因爲λ值太大，會將眞正的空間關係過度複雜化，同時也容易造成計算上的困擾。

表3.1　時空數列模式STACF及STPACF理論值的特徵

模　式	STACF的理論值	STPACF的理論值
白干擾	不顯著	不顯著
STAR（$p_{\lambda_1,\cdots,\lambda_p}$）	漸漸消失	時間落差P期後消失 空間落差λ_p期後消失
STARMA($q_{m_1,\cdots,m_q}$)	時間落q期後消失 空間落差m_q期後消失	漸漸消失
STMA($p_{\lambda_1,\cdots,\lambda_p}$，$q_{m_1,\cdots,m_q}$)	漸漸消失	漸漸消失

3.2 估計參數（Parameters Estmation）

經由確認模式的步驟找到初步的模式後，接下來必須估計此模式中的參數，使殘值達到最小。通常我們使用的方法有最大概似估計法或條件最小平方法：

(1)最大概似估計法（Maximum Likelihood Estimate）：

假設at爲一多變量常態分配，平均數爲0，變異數共變異數矩陣爲 σ 2INT。但是at爲未知隨機誤差項，必須用已知的Zt代入

$$a_t = Z_t - \sum_{k=1}^{p} \sum_{l=0}^{\lambda_k} \phi_{k\ell} W^{(\ell)} Z_{t-k} + \sum_{k=1}^{q} \sum_{\ell=0}^{m_k} \theta_{k\ell} W^{(\ell)} a_{t-k}, \quad t = 1 \cdots, N$$

以求出a_t的估計值。當N足夠大時，條件M.L.E.與眞實M.L.E.相當接近，並且在計算上相當節省時間。

(2)條件最小平方法（Conditional Least Square Estimate）：

因爲MLE與LSE均爲將誤差平方和最小化所求得，在at爲常態分配的假設之下，兩者應相等。現以STAR（2_{11}）爲例，說明如何以條件最小平方法求出參數估計值。

STAR（2_{11}）模式：

$$Z_t = \phi_{10} Z_{t-1} + \phi_{11} W^{(1)} Z_{t-1} + \phi_{20} Z_{t-2} + \phi_{21} W^{(1)} Z_{t-2} + a_t$$

$$t = 3, 4 \cdots, T$$

將上式寫成一般線性迴歸模式Z＝XΦ＋a，即

$$\begin{pmatrix} Z_1 \\ Z_2 \\ Z_3 \\ \vdots \\ Z_T \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ Z_1 & W^{(1)}Z_1 & 0 & 0 \\ Z_2 & W^{(1)}Z_2 & Z_1 & W^{(1)}Z_1 \\ \vdots & \vdots & \vdots & \vdots \\ Z_{T-1} & W^{(1)}Z_{T-1} & Z_{T-2} & W^{(1)}Z_{T-2} \end{pmatrix} \begin{pmatrix} \phi_{10} \\ \phi_{11} \\ \phi_{20} \\ \phi_{21} \end{pmatrix} + \begin{pmatrix} a_1 \\ a_2 \\ a_3 \\ \vdots \\ a_T \end{pmatrix}$$

誤差項可以矩陣表示爲

$$a = Z - X\Phi$$

令Q表示誤差平方和，則

$$Q = a'a = (Z - X\Phi)'(Z - X\Phi)$$
$$= Z'Z - 2\Phi'X'Z + \Phi'X'X\Phi$$

欲使誤差平方和最小，利用最小平方法

$$\frac{\partial Q}{\partial \Phi} = -X'Z' + X'X\Phi = 0$$

$$\rightarrow \quad (X'X)\hat{\Phi} = X'Z \tag{3.8}$$
$$\rightarrow \quad \hat{\Phi} = (X'X) - 1X'Z$$

即可將參數估計出來。

由一般的線性迴歸理論，

$$\frac{(\Phi - \hat{\Phi})X'X(\Phi - \hat{\Phi})}{\dfrac{KS_*(\hat{\Phi})}{TN - K}} \sim F_{K, TN-K}$$

其中K是Φ的維度，S*(Φ)爲殘差平方和，

$\hat{\Phi}$爲Φ的最小平方估計向量。

雖然線性迴歸中，如：自變數X矩陣不固定，而由隨機且可重複抽取之樣本決定；系統變數之間應互相獨立……等假設，時空數列模式並不滿足。但已有學者證明（Mann and

Wald, 1943），類似STAR這樣的線性模式，一般使用線性迴歸中的最小平方法估計參數，即可求出參數之近似值。

值得注意的是，時空數列模式中僅STAR模式爲線性，STMA及STARMA 皆爲非線性模式，故我們通常以梯度法（Gradient Method）估計非線性模式參數。

3.3 模式診斷及檢驗（Models Diagnostic Checking）

當初步的模式被選取且參數估計出來之後，我們必須診斷這個模式是否適合解釋這組資料。可能造成模式不適合的原因有兩個：第一、此模式可能對這組資料的解釋不夠充分，使殘差未能達到白干擾。第二、模式過度複雜，模式中存在不必要的參數。

(a)檢查殘差值是否爲白干擾（white moise）通常觀察殘差的STAC F及STPACF爲最有效的方法，若兩者均不顯著則殘差爲白干擾。即此模式已將資料充分解釋。當我們判定殘差值非白干擾時，此時殘差值可能爲另一個STARMA模式，我們必須求出此一模式以便修正初步的模式，使其殘差值達到白干擾爲止。

(b)檢定參數是否顯著不爲零：當估計的參數被檢定爲不顯著時，必須將它從模式中去除，得到一個簡化的新模式，並重新估計參數及檢驗模式。我們定檢定方法如下（比較Deutsch and Pfeifer, 1980a）：

Ho：$\phi_{ij}=0$ vs. H_1：$\phi_{ij}\neq 0$

Φ_{ij}爲Φ中的一個元素。其檢定統計量爲

$$F=\frac{(TN-K)[S_*(\hat{\Phi}^*)-S_*(\hat{\Phi})]}{S^*(\hat{\Phi})} \tag{3.9}$$

其中$\hat{\Phi}$爲H_1成立時，Φ的最小平方估計向量。

$\hat{\Phi}^*$爲H_0成立時，Φ的最小平方估計向量。

在虛無假設H_0成立的條件下，$F\sim F_{1,TN-K}$。假設顯著水準爲α，若$F<F_{1,TN-K;\alpha}$，則不否定H_0，即ϕ_{ij}應從模式中移去。對Φ中每一個元素ϕ_{ij}使用上述方法檢定，以剔除不

必要的參數。

我們不斷將新的候選模式，重覆使用確認、估計及診斷檢驗等三階段過程，直到模式通過檢驗的程序；即參數經檢定顯著不為零，殘差值達到白干擾；就得到最適合的模式。

肆、實例分析

本節我們分別以航空公司客艙機位及觀光飯店客房數為例，利用過去住宿人數與搭機人數的資料，以時空數列的STARMA模式，建立模式過程作一實證分析與探討。由於航空公司客艙機位及觀光飯店客房數之時間數列資料收集不易，且部份資料亦屬於商業機密，因此本節僅以符合實際狀況作一模擬說明與分析。

例 1.考慮3家航空公司，其搭機率的時空數列$X_{i,t}$為：

$X_{i,t}$＝（實際搭機數－預訂搭機數）／總座位，i＝1,2,3

我們應用STAR(1_1)模式，模擬此3家航空公司各135筆資料，其中3家航空公司之市場佔有率決定矩陣權數為：

$$W^{(1)}=\begin{array}{c} \\ A \\ B \\ C \end{array}\begin{array}{c}\begin{array}{ccc} A & B & C \end{array} \\ \begin{pmatrix} 0 & 0.3 & 0.2 \\ 0.3 & 0 & 0.2 \\ 0.3 & 0.3 & 0 \end{pmatrix}\end{array}$$

針對此3家航空公司之前120筆資料（我們留最後之15筆資料以作為比較預測結果依據），按第3節陳述之模式建構方式，選取一最佳的時空數列模式：

$$\text{STAR}(1_1):X_t=0.81X_{t-1}-0.53X_{t-1}W^{(1)}+\varepsilon_t$$

這裡 $X_t=\begin{pmatrix} X_{1,t} \\ X_{2,t} \\ X_{3,t} \end{pmatrix}$ ，t＝1,2,……120；，$\varepsilon_t N(0,0.5)$。

爲了比較此預測結果，我們對此3家航空公司之前120筆資料分別建構一最佳之ARIMA（p，q）模式如下：

變數	最佳ARIMA模式
X_1	$X_{1,t}=-0.0033+0.74X_{1,t}-1-0.37X_{1,t-2}+\varepsilon_{1,t}$
X_2	$X_{2,t}=-0.0041+0.61X_{2,t}-2+\varepsilon_{2,t}$
X_3	$X_{3,t}=-0.0027+0.41X_{3,t}-1-0.73X_{3,t-2}+\varepsilon_{3,t}$

由此3家航空公司之資料，選取一最佳的時空數列模式與最佳之ARI MA（p，q）模式的預測結果列於表4.1（其中SSE表示預測誤差平方和）。

表4.1　航空公司資料之預測結果比較

A公司的預測值

	時空數列		時間數列
期　數	預測值	眞實值	預測值
121	-0.0946	-0.0010	-0.0340
122	-0.0816	-0.0490	-0.0136
123	-0.0702	-0.0820	-0.0193
124	-0.0605	-0.0930	-0.0059
125	-0.0521	-0.0310	-0.0060
126	-0.0448	-0.0860	-0.0099
127	-0.0386	-0.0020	-0.0084
128	-0.0331	-0.0310	-0.0058
129	-0.0285	-0.0040	-0.0045
130	-0.0244	-0.0850	-0.0044
131	-0.0210	-0.1820	-0.0049
132	-0.0180	-0.1270	-0.0053
133	-0.0154	-0.0220	-0.0054
134	-0.0132	-0.0040	-0.0053
SSE	0.90		0.94

B公司的預測值

	時空數列	時間數列	
期　數	預測值	眞實值	預測值
121	-0.1354	-0.0860	-0.1038
122	-0.1132	-0.1540	-0.0672
123	-0.0943	-0.1400	-0.0450
124	-0.0789	-0.1570	-0.0315
125	-0.0661	-0.0960	-0.0233
126	-0.0554	-0.1720	-0.0183
127	-0.0466	-0.0550	-0.0152
128	-0.0393	-0.1200	-0.0134
129	-0.0331	-0.0510	-0.0123
130	-0.0280	-0.0250	-0.0116
131	-0.0237	-0.1250	-0.0112
132	-0.0200	-0.0920	-0.0109
133	-0.0170	-0.0560	-0.0107
134	-0.0144	-0.0580	-0.0107
SSE	0.0009		0.003

C公司的預測值

	時空數列	時間數列	
期　數	預測值	眞實值	預測值
121	-0.0679	0.0510	0.0043
122	-0.0629	-0.0040	0.0514
123	-0.0563	-0.0900	0.0150
124	-0.0504	-0.0780	-0.0339
125	-0.0448	-0.0480	-0.0274
126	-0.0397	-0.0170	0.0108
127	-0.0350	0.0030	0.0216
128	-0.0308	-0.0330	-0.0017
129	-0.0269	0.0010	-0.0191
130	-0.0235	0.1060	-0.0092
131	-0.0205	0.1610	0.0075
132	-0.0179	-0.0090	0.0070
133	-0.0155	-0.1270	-0.0052
134	-0.0135	-0.0370	-0.0099
SSE	0.071		0.139

由表4.1我們可以看出時空數列模式預測的誤差平方和均遠小於最佳之ARIMA（p，q）模式的預測結果。雖然在這裡航空公司搭機率之資料乃由特定的時空數列模式預測與之比較，對後者可能較不公平。但是本研究的另一個目的乃說明一事實：當時間數列其走勢具有時空相關時（此處之空間關係乃可傳遞的市場佔有率）時空數列模式預測之結果應比傳統ARIMA模式來的精確。

例 2.考慮6家觀光飯店，其進住率時空數列Yi,t為：

$Y_{i,t}$ =（實際進住房間數—預訂房間數）／總房間數，i＝1,2……6

我們應用STAR（3m）模式，模擬此6家觀光飯店各135筆資料，我們以各家飯店間之車程距離倒數決定矩陣權數如下：

車程距離	A飯店	B飯店	C飯店	D飯店	E飯店	F飯店
A飯店	0					
B飯店	5	0				
C飯店	5	5	0			
D飯店	14	9	14	0		
E飯店	10	6	12	3	0	
F飯店	4	2	3	8	6	0

$$\text{加權矩陣 } W^{(1)} = \begin{pmatrix} 0 & & & & & \\ 0.200 & 0 & & \text{Symmetric} & & \\ 0.200 & 0.200 & 0 & & & \\ 0.083 & 0.111 & 0.071 & 0 & & \\ 0.100 & 0.167 & 0.083 & 0.333 & 0 & \\ 0.250 & 0.500 & 0.333 & 0.125 & 0.167 & 0 \end{pmatrix}$$

$$Y_t = \begin{pmatrix} Y_{1,t} \\ Y_{2,t} \\ Y_{3,t} \\ Y_{4,t} \\ Y_{5,t} \\ Y_{6,t} \end{pmatrix}, \quad t=1,2,\cdots 135$$

針對此6家觀光飯店之前120筆資料（我們留最後之15筆資料以作爲比較預測結果依據），按第3節陳述之模式建構方法，選取一最佳的時空數列模式：

STAR(3m)：

$$
\begin{aligned}
Y_t = {} & 0.23Y_{t-1} + 0.30Y_{t-1}W^{(1)} - 0.06Y_{t-2} \\
& + 0.31Y_{t-3}W^{(1)} - 0.1Y_{t-2} - 0.32Y_{t-3}W^{(1)} + \varepsilon_t
\end{aligned}
$$

爲了比較此預測結果，我們對此6家觀光飯店之前120筆資料分別建構一最佳之ARIMA（p，q）模式如下：

變數	最佳ARIMA模式
Y_1	$Y_{1,t} = -0.0009 + 0.57Y_{1,t-1} - 0.27Y_{t-2} + \varepsilon_{1,t}$
Y_2	$Y_{2,t} = -0.0018 + 0.37Y_{2,t-1} + \varepsilon_{2,t}$
Y_3	$Y_{3,t} = -0.0010 + 0.29Y_{3,t-1} - 0.70Y_{3,t-2} + \varepsilon_{3,t}$
Y_4	$Y_{4,t} = -0.0005 + 0.89Y_{4,t-1} - 0.54Y_{4,t-2} + \varepsilon_{4,t}$
Y_5	$Y_{5,t} = -0.0015 + 0.17Y_{5,t-1} - 0.62Y_{5,t-2} + \varepsilon_{5,t}$
Y_6	$Y_{6,t} = -0.0004 + 0.45Y_{6,t-1} + \varepsilon_{6,t}$

表4.2 觀光飯店資料之預測結果比較

	A公司的預測值		
	時空數列		時間數列
期　數	預測值	眞實值	預測值
121	0.0027	0.0177	0.0223
122	0.0152	0.0639	-0.0013
123	-0.0143	0.0491	-0.0075
124	-0.0049	-0.0615	-0.0048
125	-0.0142	-0.0285	-0.0016
126	-0.0000	-0.0001	-0.0004
127	-0.0008	0.0462	-0.0007
128	-0.0069	0.1190	-0.0011
129	-0.0027	0.0652	-0.0013
130	-0.0032	0.0977	-0.0013
131	-0.0014	0.1126	-0.0012
132	-0.0011	0.0520	-0.0012
133	-0.0024	-0.0380	-0.0012
134	-0.0006	-0.0521	-0.0012
SSE	0.053		0.062

	B公司的預測值		
	時空數列		時間數列
期　數	預測值	眞實值	預測值
121	0.0041	0.0260	0.0128
122	0.0149	0.0928	0.0029
123	-0.0212	0.0592	-0.0006
124	-0.0057	-0.0266	-0.0020
125	-0.0185	0.0253	-0.0025
126	0.0002	0.0069	-0.0027
127	-0.0013	0.0509	-0.0028
128	0.0093	0.1137	-0.0028
129	0.0036	0.0638	-0.0028
130	0.0045	0.1420	-0.0028
131	-0.0018	0.1430	-0.0028
132	-0.0014	0.1038	-0.0028
133	-0.0033	0.0377	-0.0028
134	-0.0009	0.0273	-0.0028
SSE	0.084		0.056

	C公司的預測值		
	時空數列		時間數列
期　數	預測值	眞實值	預測值
121	0.0220	0.0521	-0.0335
122	0.0233	0.0352	-0.0501
123	-0.0171	-0.0116	-0.0079
124	-0.0081	-0.1063	-0.0363
125	-0.0164	-0.0056	-0.0040
126	0.0002	0.0774	-0.0251
127	-0.0005	0.0677	-0.0110
128	0.0076	0.0574	-0.0134
129	0.0029	-0.0271	-0.0106
130	0.0035	0.0340	-0.0072
131	-0.0015	0.0963	-0.0104
132	-0.0011	0.0091	-0.0010
133	-0.0026	-0.1114	-0.0066
134	-0.0007	-0.0586	-0.0002
SSE	0.023		0.048

	D公司的預測值		
	時空數列		時間數列
期　數	預測值	眞實值	預測值
121	-0.0080	0.0100	0.0557
122	0.0104	0.0549	0.0190
123	-0.0146	0.0588	-0.0133
124	-0.0040	-0.0527	-0.0227
125	-0.0108	-0.0587	-0.0136
126	0.0014	-0.0217	-0.0005
127	-0.0001	0.0504	0.0062
128	0.0054	0.1475	0.0054
129	0.0017	0.1106	0.0009
130	0.0023	0.1081	-0.0026
131	-0.0011	0.1091	-0.0033
132	-0.0007	0.0481	-0.0021
133	-0.0018	-0.0580	-0.0006
134	-0.0005	-0.0951	-0.0000
SSE	0.058		0.085

表4. E飯店的預測值

期數	時空數列預測值	眞實值	時間數列預測值
121	0.0084	-0.0166	-0.0402
122	0.0085	0.1220	0.0481
123	-0.0227	0.0027	-0.0343
124	-0.0060	0.0072	0.0338
125	-0.0114	0.0201	-0.0283
126	0.0022	0.0098	0.0240
127	-0.0002	0.0546	-0.0230
128	0.0060	0.0981	0.0171
129	0.0019	0.0413	-0.0185
130	0.0027	0.1510	0.0121
131	-0.0012	0.0960	-0.0150
132	-0.0009	0.1097	0.0084
133	-0.0021	0.1204	-0.0121
134	-0.0006	0.0579	-0.0056
SSE	0.047		0.073

表4. F飯店的預測值

期數	時空數列預測值	眞實值	時間數列預測值
121	0.0020	0.0237	0.0221
122	0.0263	0.0673	0.0095
123	-0.0207	0.0611	0.0039
124	-0.0084	-0.0495	0.0013
125	-0.0217	-0.0397	0.0001
126	0.0003	-0.0122	-0.0003
127	-0.0011	0.0427	-0.0005
128	0.0103	0.1291	-0.0006
129	0.0039	0.0935	-0.0007
130	0.0049	0.1184	-0.0007
131	-0.0020	0.1338	-0.0007
132	-0.0016	0.0755	-0.0007
133	-0.0036	-0.0265	-0.0007
134	-0.0010	-0.0627	-0.0007
SSE	0.075		0.082

由表4.2我們可以看出時空數列模式各期的預測誤差大都小於最佳之ARIMA（p，q）模式的預測結果。雖然在這裡觀光飯店之進住率資料乃由特定的時空數列模式預測與之比較，對後者可能較不公平。但是本研究的另一個目的乃說明一事實：當時間數列其走勢具有時空的地緣相關時，時空數列模式預測之結果應比傳統ARIMA模式來的精確。

伍、結　論

在實務上，許多資料不僅以等時距的方式蒐集，這些資料所發生的位置也是一個重要的因素，這時只能分析某一特定地點的單變量時間數列就無法處理。而多變量時間數列只可討論多個變量間的關係，對於空間關係無法處理，因此時空數列的發展有其必要。單變量ARIMA 模型主要概念在於時間較接近的觀察值，彼此間的交互影響較大。而時空數列則強調鄰近位置觀察值的影響。

本研究將針對飯店和航空公司，利用過去住宿人數與搭機人數的資料，以時空數列的

STARMA模式與狀態空間方程式的方法，建立一個模式，找出應將多少比例的房間與機位提供給顧客預訂之用，避免不必要的損失。

參 考 文 獻

吳柏林，廖敏治（1983）。大台北地區空氣污染指標之時空數列分析。中國統計學報。31卷2期，139-167。

吳柏林，陳雅玫（1994）。台灣地區失業率的時空數列分析。人力資源學報。31-34。

Ali, M. M. （1979）. Analysis of stationary spatial temporal processes : estimation and prediction, Biometrika, 66, 513-518.

Baczkowski, A. J. and Mardia, K. V. （1989）. A Test of spatialindependence, Journal of Aplied Statistics, 16, 235-351.

Bennett, R. J. （1979）. Spatial time series, Pion Limited, London.

Bennett, R. J. （1984）. Advances in the analysis of spatial time series, Spatial Statistics and Models,66, 235-251.

Besag, J. （1974）. Spatial interaction and the statistical analysis of lattice systems, Journal of the Royal Statistical Society, B, 36, 192-225.

Box, G. E. P. and Jenkins, G. M. （1976）. Time series analysis forecasting and control, 2nd ed. ,Holden-Day, San-Francisco.

Cliff, A. D. and Ord, J. K. （1981）. Spatial processes : models and applications, Pion Limited, London.

Deutsch, S. J. and Pfeifer, P. E. （1980a）. A three-stage iterative procedure for space-time modeling, Technometrics, 22, 35-47.

Deutsch, S. J. and Pfeifer, P. E. （1980b）. Identification and interpretation of first order space-time ARMA models, Technometrics, 22, 397-408.

Deutsch, S. J. and Pfeifer, P. E. （1981）. Space-Time ARMA modeling with contemporaneously correlated innovations, Technometrics, 23, 401-409.

Flahault, A. et al. （1988）. Modeling the 1985 influenza epidemic in Frence, Statistics in Medicine, 7, 1147-1155.

Griffith, D. A., Haining, R. P. and Bennett, R. J. （1985）. Estimating missing values in space-time data series, Time Series Analysis : Theory and Practice 6, 273-296.

Haslett, J. and Raftery, A. E. （1989）. Space-Time modelling with long-memory dependence : assessing Ireland's wind power resource, Applied Statistics, 38, 1-50.

Mann, H. B. and Wald, A. （1943）. On the statistical treament of linear stochastic difference equations, Econometrika, 11, 173-270.

Pfeifer, P. E. and Bodily, S. E. （1990）. A Test of space-time ARMA modelling and forcasting of hotel data, Journal of Forecasting, 9, 155-272.

Raftery, A. E., Haslett, J. and McColl, E. （1982）. Wind power : A space-time process, Time Series Analysis : Theory and Practice 2, 191-202.

Rogers, A. （1974）. Statistical analysis of spatial dispersion : the quadrat method, Pion Limited, London.

Stoffer, D. S. （1986）. Estimation and identification of space-time ARMAX models in the presence of missing data, Journal of the American Statistical Association, 81, 762-772.

Tjostheim, D. （1978）. Statistical spatial series modelling , Adv. Appl. Prob., 10, 130-154.

Wei William W. S. （1990）. Time series analysis : univariate and multivariate method, Addison-Wesley.

Zeger. S. L. （1985）. Exploring an ozone spatial time series in the frequency domain, Journal of the American Statistical Associations, 80, 323-331.