

# 記憶性 $k$ -Factor GARMA 模式之研究

## Studies on $k$ -Factor GARMA Long-Memory Models

淡江大學水資源政策及管理研究中心主任

淡江大學水資源及環境工程研究所研究助理

虞國興

何琮裕

Gwo-Hsing Yu

Tsung-Yu He

### 摘 要

以往在長記憶水文時序之分析上，常利用差分、除勢等轉換方法。近年來，研究顯示  $k$ -factor GARMA 模式在長記憶時間序列上有著不錯的預測結果。本研究主要探討此一模式之適用性，並將其應用於臺灣河川月流量資料上，與傳統除勢模式做一比較。

本研究於  $k$ -factor GARMA 模式之參數推估上，採用 Hosking(1984)所提出之近似最大概似法；至於模式之判定上本研究則採用 AIC、BIC 兩種判斷準則。另外，為瞭解模式預測能力之優劣，本研究分別採用 MSE、MAPE 及 UI 三種預測指標。研究結果顯示，於合成資料之參數推估上，近似最大概似法受到樣本數的影響較大，在大樣本時可得較高之準確性；而於模式判定方面，受到模式參數與樣本數之影響，整體而言 BIC 之偵測能力明顯優於 AIC。於實測資料分析上，就統計特性保存能力方面， $k$ -factor GARMA 模式於偏態係數之保存能力優於除勢模式；除勢模式則於平均值及變異數之表現上較佳，另於預測能力表現上， $k$ -factor GARMA 模式於低流量部份之預測能力明顯優於除勢模式。

整體而言， $k$ -factor GARMA 模式於預測能力上優於傳統除勢模式，且可改善除勢模式無法掌握之偏態保存能力，加上參數精簡原則之考慮， $k$ -factor GARMA 模式於臺灣月流量資料之分析上具有較佳之適用性。

**關鍵詞：**長記憶性過程，GARMA 模式，近似最大概似法。

### ABSTRACT

In the past, difference and detrended methods were usually used in analyzing long-memory hydrological time series. Recently, a  $k$ -factor GARMA model was proposed to model long-memory time series, and its forecasting results are better than conventional methods. The major purpose of this study is to examine the suitability of this model, and

then apply it to monthly riverflow data of Taiwan.

In this study, approximate maximum likelihood method proposed by Hosking(1984) is used to estimate the parameters of  $k$ -factor GARMA model, and AIC and BIC are used for model identification. Besides, MSE, MAPE and UI are used as criteria for comparison of forecasting. The accuracy of approximate maximum likelihood method is affected by sample size. The larger the sample size, the better parameter estimation accuracy. The results of model identification are affected by parameter and sample size. The BIC criterion has better results in model identification than AIC criterion. On the other hand, the performance of  $k$ -factor GARMA model is better than detrended model in preserving skewness when real data are analyzed. To the opposite, detrended model fits better to the mean and variance than  $k$ -factor GARMA model. The forecasting results show that  $k$ -factor GARMA model is better than detrended model, especially the low streamflow .

In conclusion,  $k$ -factor GARMA model has better forecasting ability than detrended model, and improves preserving skewness. In general, the  $k$ -factor GARMA model is a reasonable model for the monthly riverflow forecasting of Taiwan.

**Keywords:** Long-Memory process, Gegenbauer ARMA models, Approximate maximum likelihood method.

## 一、前言

水文時序多為非定常性(Non-stationary)及非常態性(Non-Gaussian)歷程，在缺乏直接分析的理論下，往往須藉用轉換方法，將其轉換為定常性序列。以往在長記憶水文時序之分析上，常利用差分、除勢法(Detrended Method)等轉換方式，虞及莊(1989)曾針對上述長記憶水文時間序列模式加以探討。近年來， $k$ -factor GARMA (Gegenbauer ARMA)模式的提出在週期性資料上有著不錯的預測結果(Woodward 等人, 1998)。因此本研究針對  $k$ -factor GARMA 模式之參數推估、模式判定及合成資料統計特性保存能力等進行研究，並探討其於臺灣河川月流量資料之適用性。本研究主要探討下列幾項主題：

- (1)採用合成資料，探討  $k$ -factor GARMA 模式參數之推估。
- (2)採用合成資料，探討  $k$ -factor GARMA 模式偵測之準確性。
- (2)採用實測資料，探討  $k$ -factor GARMA 模式所產生之合成資料其統計特性保存效果如

何，並與除勢模式進行比較。

- (4)採用實測資料，進行  $k$ -factor GARMA 模式及除勢模式之預測能力比較。

## 二、理論基礎

### 2-1 長記憶過程(Long-Memory Process)

序列之記憶性，一般可用序列資料間所有稽延長度之自相關函數絕對值總和加以區分：

- (1)無記憶性過程：

若序列資料間所有稽延長度之自相關函數絕對值總和趨近於零，則為一無記憶性過程，例如白噪音(White Noise)過程。

- (2)短記憶性過程：

若序列資料間所有稽延長度之自相關函數絕對值總和趨近於某定數，則為一短記憶性過程，例如定常性的 ARMA 序列。

- (3)長記憶性過程：

若序列資料間所有稽延長度之自相關函數絕對值總和趨近於無窮大，則為一長記憶性過程，例如一般具有趨勢或週期之序列。

## 2-2 除勢模式

傳統上，在分析非定常性序列時，通常將非定常性時序轉換為定常性時序直接加以分析，若此時序具有週期，則利用除勢法將資料轉換為定常性序列加以分析，除勢模式建立步驟如下：

- (1)繪製資料圖，決定週期。
- (2)求出資料 $\{Y_{v,h}\}$ 之週期平均值 $\bar{Y}_h$ 及週期標準偏差 $S_h$ ，並將資料予以標準化。

$$Z_{v,h} = \frac{Y_{v,h} - \bar{Y}_h}{S_h}$$

其中 $v$ 為年份， $v=1,2,\dots,M$ (年)

$h$ 為年份之時距(Time Interval within the Year)例如月( $h=1,2,\dots,12$ )、旬( $h=1,2,\dots,36$ )、日( $h=1,2,\dots,365$ )等。

- (3)轉換後資料 $Z_{v,h}$ 採用最大似法推求參數，並以AIC、BIC選取最適ARMA模式。
- (4)利用上述所得之模式進行預測得 $Z'_{v,h}$ ，再利用先前求得之 $\bar{Y}_h$ 及 $S_h$ 反轉換為 $Y'_{v,h}$ 。

$$Y'_{v,h} = \bar{Y}_h + Z'_{v,h} \cdot S_h$$

## 2-3 $k$ -factor GARMA 模式

根據GARMA模式(Gray等人,1989)所延伸之 $k$ -factor GARMA模式之一般形式為：

$$\phi(B) \prod_{i=1}^k (1 - 2u_i B + B^2)^{\lambda_i} (Y_t - \mu) = \theta(B) a_t \dots (1)$$

其中： $u_i$ 指出長記憶性質之頻率， $u_i = \cos(2\pi f_{0i})$ 。

$\lambda_i$ 為自相關函數遞減程度， $\lambda$ 愈大，遞減速率愈緩慢。

$$(1 - 2uB + B^2)^{-\lambda} = \sum_{l=0}^{\infty} C_l^{(\lambda)}(u) B^l \dots (2)$$

$$C_l^{(\lambda)}(u) = \sum_{k=0}^{[l/2]} \frac{(-1)^k (2u)^{l-2k} \Gamma(\lambda - k + 1)}{k!(l-2k)! \Gamma(\lambda)}, \lambda > 0 \quad (3)$$

$C_l^{(\lambda)}(u)$ 為Gegenbauer函數(Magnus等人,1966)。

由式(1)可得知， $k$ -factor GARMA模式為長

記憶Gegenbauer多項式與短記憶ARMA模式所結合而成，其頻譜為：

$$P(f) = \sigma_a^2 \frac{|\theta(e^{2\pi i f})|^2}{|\phi(e^{2\pi i f})|^2} \left| \prod_{j=1}^k \left\{ -2u_j e^{-2\pi i f} + e^{-4\pi i f} \right\}^{\lambda_j} \right|^2$$

$$= \sigma_a^2 \frac{|\theta(e^{2\pi i f})|^2}{|\phi(e^{2\pi i f})|^2} \prod_{j=1}^k \left[ 4 \left\{ \cos(2\pi f) - u_j \right\}^2 \right]^{\lambda_j},$$

$$0 \leq f \leq 0.5 \quad \dots (4)$$

$k$ -factor GARMA模式之定常性與長記憶性條件為：

- (1)若 $\phi(B) = 0$ 之所有根皆落於單位圓外，且當 $|u_i| < 1$ ， $\lambda_i < \frac{1}{2}$ ，或 $|u_i| = 1$ ， $\lambda_i < \frac{1}{4}$ 時，式(1)中 $Y_t$ 為一定常性序列。
- (2)若序列滿足上述條件，且 $\lambda_i > 0$ ，則式(1)中 $Y_t$ 為一定常性長記憶序列。

Gray(1994)更指出GARMA模式在上述條件之下，模式之自相關函數滿足：

$$\rho_k \sim k^{2\lambda-1} \cos(2\pi k f_0), \text{ 當 } k \rightarrow \infty$$

其中 $f_0$ 為Gegenbauer頻率。

因此GARMA模式適用於描述自相關函數遞減緩慢且具週期性之資料型態，而 $k$ -factor GARMA模式則可以用來分析具有多個週期之資料，在使用上更具彈性。

## 2-4 $k$ -factor GARMA 模式之合成資料繁衍方法

Ramsey(1974)指出，任何期望值為零之定常性線性過程 $Z_t = \sum_{m=0}^{\infty} \psi_m B^m a_t$ ，其條件期望值與變異數分別為：

$$m_t = E(Z_t | Z_0, Z_1, \dots, Z_{t-1}) = \sum_{j=1}^t \phi_{tj} Z_{t-j} \dots (5)$$

$$v_t = \text{Var}(Z_t | Z_0, Z_1, \dots, Z_{t-1}) = \gamma_0 \prod_{j=1}^t (1 - \phi_{jj}^2) \dots (6)$$

$\phi_{ij}$ 及 $\phi_{tj}$ 可以Durbin-Levinson遞迴式求得：

$$\phi_{tt} = N_t / D_t$$

$$\phi_{ij} = \phi_{i-1,j} - \phi_{ii}\phi_{i-1,t-j}, \quad j = 1, 2, \dots, t-1$$

其中

$$N_0 = 0, \quad D_0 = 1$$

$$N_t = \rho_t - \sum_{j=1}^{t-1} \phi_{t-1,j} \rho_{t-j}$$

$$D_t = D_{t-1} - N_{t-1}^2 / D_{t-1}$$

Ford (1991)針對式中自相關函數之推求，採用頻譜之傅立葉反轉換(Inverse Fourier Transformation)，先求出模式之自相關變異數：

$$\gamma_k = 2 \int_0^{0.5} P(f) \cos(2\pi k f) df \dots\dots\dots(7)$$

因此，根據 Ford 所建議之方法產生  $k$ -factor GARMA 模式之合成資料，可以依據下列步驟：

- (1) 求得理論模式之頻譜並進行傅立葉反轉換，藉以推求自相關函數。
- (2) 從  $N(0, v_0)$  分布產生序列之起始值，其中  $v_0$  為序列之變異數。
- (3) 利用 Durbin-Levinson 遞迴式計算  $\phi_{ij}$  及  $\phi_{ii}$ ，進一步求出條件期望值與變異數。
- (4) 從  $N(m_t, v_t)$  繁衍合成資料。

詳細流程如圖 1 所示。

### 2-5 $k$ -factor GARMA 模式之模式判定與參數推估

根據 Hosking (1984)所建議之近似最大概似法(Approximate Maximum Likelihood Method)，可將  $k$ -factor GARMA 模式改寫為：

$$\phi(B)Z_t = \theta(B)a_t$$

其中，

$$Z_t = \prod_{i=1}^k (1 - 2u_i B + B^2)^{\lambda_i} (Y_t - \mu) \dots\dots\dots(8)$$

根據式(8)，改變參數  $u$  及  $\lambda$ ，可以求得在候選模式 ARMA(p,q)下所對應之  $Z_t$  並進行最大概似推估，接著找出在不同  $Z_t$  下，其最大概似函數值中之最大值，此時之  $u$ 、 $\lambda$  與所推得之  $\phi$ 、 $\theta$ 、 $\sigma_a^2$  即為所求模式之參數。至於  $\phi(B)$  與  $\theta(B)$  之階數則是利用判斷準則，從不同之 ARMA 候選模式中找出最適模式。

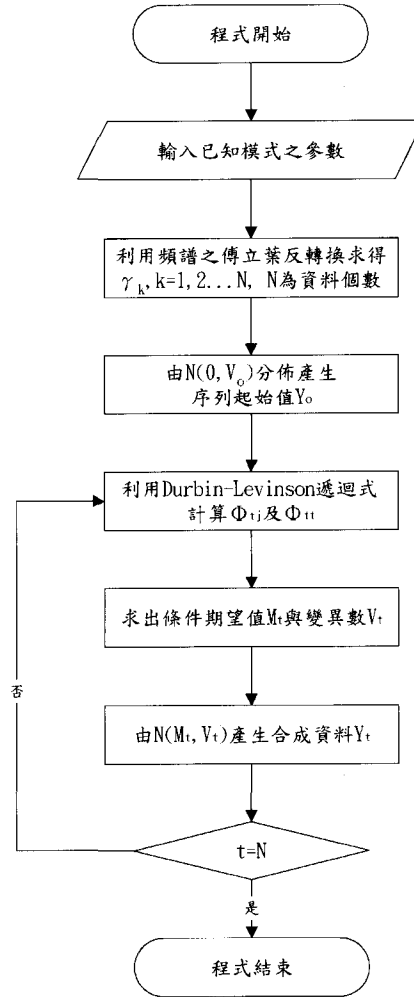


圖 1  $k$ -factor GARMA 模式之資料繁衍流程圖

上述方法整理如下列步驟：

- (1) 從頻譜中找出  $\{\lambda_1, u_1, \dots, \lambda_k, u_k\}$  之可能範圍。
- (2) 利用高階 AR 模式進行向後預測 (Backcast)，得到  $Y_{-1}, Y_{-2}, \dots, Y_{-M+1}$ ，其中  $M$  為一足夠大之整數。
- (3) 針對  $\{\lambda_1, u_1, \dots, \lambda_k, u_k\}$  不同之組合，利用下式將原序列轉換成 ARMA 序列，若序列具有趨勢，則需另行分析。

$$Z_t = \prod_{i=1}^k \left\{ \sum_{l=0}^{t+M-1} C_l^{(-\lambda_i)} (u_i) B^l \right\} (Y_t - \mu) \dots\dots\dots(9)$$

- (4)計算  $Z_t$  之最大概似函數值。
  - (5)在所有最大概似函數值中，具有最大值之  $\{\lambda_1, u_1, \dots, \lambda_k, u_k\}$  組合，即為近似最大概似推估值。
  - (6)計算上述結果之 AIC (或 BIC)值。
  - (7)重覆步驟 3 至 6，找出不同  $p$ 、 $q$  下之候選模式，而具有最小 AIC (或 BIC)值之模式即為最適模式。
- 上述計算流程如圖 2 所示。

### 2-6 $k$ -factor GARMA 模式之預測

引用 Box 及 Jenkins 之  $\pi$  權重 ( $\pi$ -weights) 預測方式，可將  $k$ -factor GARMA 模式之  $\pi$  權重展開為：

$$\sum_{j=0}^{\infty} \pi_j B^j = \theta^{-1}(B) \phi(B) \prod_{i=1}^k (1 - 2u_i B + B^2)^{\lambda_i}$$

$$= \theta^{-1}(B) \phi(B) \prod_{i=1}^k \left\{ \sum_{l=0}^{\infty} C_l^{(-\lambda_i)} (u_i) B^l \right\} \dots \dots \dots (10)$$

因此  $Y_{t_0+h}$  的預測值可以利用下式求出：

$$\hat{Y}_{t_0}(h) = - \sum_{j=1}^{t_0-1+h} \pi_j Y_{t_0+h-j} \dots \dots \dots (11)$$

其中，當  $h > 0$  時， $Y_{t_0}(h)$  為預測值；當  $h \leq 0$  時， $Y_{t_0}(h)$  為觀測值。

### 2-7 其他相關準則

#### 2-7-1 模式判定準則

直覺上模式階數愈高，其擬合情況會愈好，但是必須注意到過度擬合 (Overfitting) 之影響，為了估計模式之品質並避免模式參數過度擬合情形，本研究採用下列準則來判斷最適模式為何：

$$AIC(M) = n \ln \hat{\sigma}_a^2 + 2M \dots \dots \dots (12)$$

$$BIC(M) = n \ln \hat{\sigma}_a^2 + M \ln n \dots \dots \dots (13)$$

其中， $M$  為模式之參數個數。

最適模式之選取以 AIC 或 BIC 最小值決定。(虞及劉，1995)

#### 2-7-2 模式預測指標

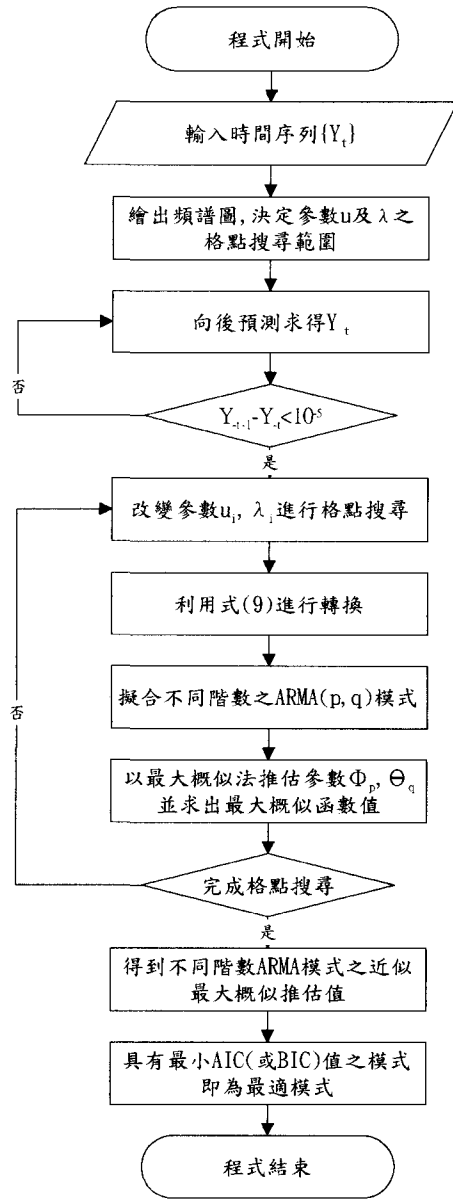


圖 2  $k$ -factor GARMA 模式之模式判定流程圖

本研究對於預測能力採用三種指標來比較其優劣，分別為 MSE、MAPE 及 UI，其敘述如下：

$$MSE = \frac{1}{K} \sum_{t=T+1}^{T+K} \{ \hat{Y}_t(1) - Y_{t+1} \}^2 \dots \dots \dots (14)$$

表 1 實測月流量資料之統計特性

站名	流域名稱	起訖年	資料長度(月)	平均值(C.M.S)	變異數	偏態係數
玉峰	淡水河	1958-1997	480	17.67	380.78	2.85
稜角	淡水河	1958-1997	480	7.11	48.15	2.82
福山	淡水河	1958-1997	480	17.71	242.51	2.74
秀巒	淡水河	1958-1997	480	6.03	58.41	4.23
高義	淡水河	1958-1997	480	30.26	1200.95	3.40
三峡	淡水河	1958-1997	480	8.81	60.97	2.33
打鹿坑	後龍溪	1958-1997	480	13.36	370.66	2.92
桶頭	濁水溪	1958-1997	480	18.97	742.59	2.34
北港	北港溪	1958-1997	480	24.45	931.48	2.19
延平	卑南溪	1958-1997	480	33.88	1730.86	2.52
台東大橋	卑南溪	1958-1997	480	90.50	14017.75	2.39
山腳	南澳溪	1958-1997	480	5.70	44.63	4.58
蘭陽大橋	蘭陽溪	1958-1997	480	57.32	3719.34	2.62

$$MAPE = \frac{1}{K} \sum_{t=T+1}^{T+K} \frac{|\hat{Y}_t(1) - Y_{t+1}|}{Y_{t+1}} \times 100\% \dots\dots(15)$$

$$UI = \left[ \frac{\sum_{t=T+1}^{T+K} \{\hat{Y}_t(1) - Y_{t+1}\}^2}{\sum_{t=T+1}^{T+K} \{Y_{t+1}\}^2} \right]^{1/2} \dots\dots\dots(16)$$

其中， $Y_t(t)$  為時間  $t$  之前一步預測值，且 MSE 大小可以瞭解預測值與真值之離散程度；從 MAPE 大小可以瞭解預測值與真值之相對誤差百分比；而 UI 則是將預測值與真值之誤差予以無因次化。上述指標愈小，預測能力愈佳。

### 三、本研究所採用之資料

#### 3-1 合成資料

本研究共使用七種 GARMA 模式產生合成資料，其樣本數分別為 250 及 500。由 IMSL 產生常態分布之隨機亂數，每一種模式之合成資料皆產生 100 組。其中長記憶項特徵方程式根之絕對值為 1。

#### 3-2 實測資料

本研究所使用之臺灣河川月流量資料，考慮到資料完整性及記錄長度，故選取擁有最近 40 年(1958-1997)流量記錄之測站，共計 13 站，其

基本統計特性如表 1 所示。其中，山腳站之月流量資料在 1958 年 7 月及 8 月有所遺漏，補遺方式採取該站每年同一月份之平均月流量代表該流量資料。

### 四、結果與討論

本研究之結果分為合成資料與實測資料兩部份討論。合成資料部份探討  $k$ -factor GARMA 模式之資料繁衍、參數推估與模式判定；實測資料方面則對  $k$ -factor GARMA 模式於臺灣河川月流量資料之適用性及預測能力進行討論，其結果分述如下：

#### 4-1 合成資料

##### 4-1-1 $k$ -factor GARMA 模式之資料繁衍

為瞭解本研究所採用之資料繁衍方式其穩定性如何，故繁衍 100 組樣本數為 5000 之資料加以分析，其相關性結構如圖 3 所示，由圖中可以得知其自相關函數(ACF)及偏自相關函數(PACF)皆與理論值相差不大。研究中並取前 500 與後 500 個資料進行比較，發現模式結構並不會隨資料擷取位置不同而有所改變。而在樣本個數方面，長記憶性模式需要較多之資訊，故樣本數愈大時，所求得之 ACF 與 PACF 愈接近理論值。因此本研究所採用之資料繁衍方式在模式結構

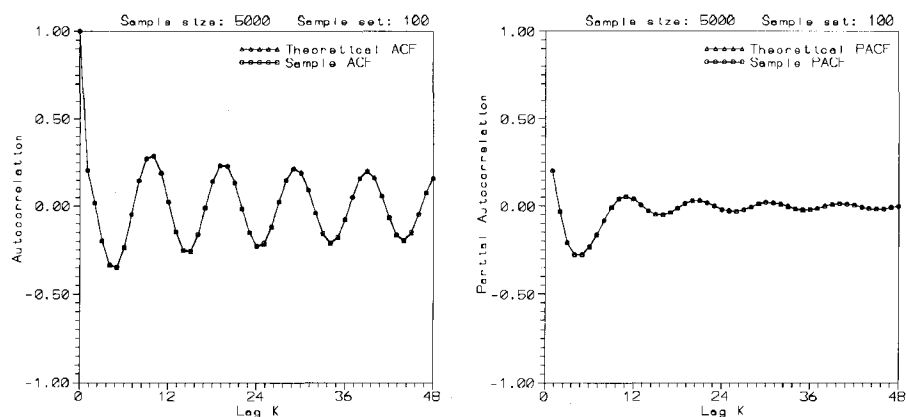


圖 3 模式 3 之自相關函數與偏自相關函數圖

表 2 合成資料參數推估結果 (樣本數 500)

模式	模式參數(真值)								模式參數(推估值)								
	$\phi_1$	$\phi_2$	$\theta_1$	$\theta_2$	$u_1$	$u_2$	$\lambda_1$	$\lambda_2$	$\phi_1$	$\phi_2$	$\theta_1$	$\theta_2$	$u_1$	$u_2$	$\lambda_1$	$\lambda_2$	$\sigma_u^2$
1-factor GARMA(0,0)	--	--	--	--	0.80	--	0.40	--	--	--	--	--	0.80 (0.000)	--	0.40 (0.001)	--	1.01 (0.004)
1-factor GARMA(1,0)	0.70	--	--	--	0.80	--	0.40	--	0.68 (0.002)	--	--	--	0.81 (0.000)	--	0.41 (0.001)	--	1.01 (0.004)
1-factor GARMA(0,1)	--	--	0.70	--	0.80	--	0.40	--	--	--	0.71 (0.001)	--	0.80 (0.000)	--	0.41 (0.001)	--	1.00 (0.006)
1-factor GARMA(2,0)	1.20	-0.50	--	--	0.80	--	0.40	--	1.19 (0.012)	-0.49 (0.011)	--	--	0.80 (0.000)	--	0.40 (0.005)	--	1.00 (0.004)
1-factor GARMA(0,2)	--	--	0.30	0.40	0.80	--	0.40	--	--	--	0.31 (0.003)	0.40 (0.002)	0.80 (0.000)	--	0.41 (0.001)	--	1.01 (0.004)
1-factor GARMA(1,1)	0.80	--	-0.60	--	0.80	--	0.40	--	0.77 (0.004)	--	-0.60 (0.002)	--	0.81 (0.001)	--	0.42 (0.001)	--	1.02 (0.005)
2-factor GARMA(1,0)	0.30	--	--	--	0.80	0.60	0.40	0.30	0.28 (0.006)	--	--	--	0.79 (0.003)	0.62 (0.004)	0.40 (0.003)	0.30 (0.002)	1.00 (0.004)

註：模式參數(推估值)欄中為 100 組合成資料參數推估值之平均值，(·)內為 100 組合成資料參數推估值之 MSE。

保有能力上尚有不錯之表現，且此一方式中起始值之影響並不顯著，因此適用於長記憶性之資料繁衍。

#### 4-1-2 k-factor GARMA 模式之參數推估

本研究所使用之近似最大似法在參數  $u$  及  $\lambda$  之推估上，採用格點搜尋法(Grid Search Method)，當格點間距較小時，所推得之參數較為精準，但使用之 CPU 計算時間亦相對增加。本研究採取 0.01 為其搜尋間隔。

為探討參數推估之精確度，本研究採用已知  $k$ -factor GARMA 模式所繁衍之合成資料進行參

數推估，並計算 100 組之平均值及 MSE，其結果如表 2 所示。就平均值而言，樣本數為 500 時之結果均較樣本數 250 時之結果更接近理論模式之參數，而就參數推估之 MSE 來看，樣本數 500 時之推估誤差明顯小於樣本數 250 時之結果，因此在大樣本時參數推估之精確度將會較小樣本之參數推估結果為佳。

整體結果顯示在小樣本之情況下，參數推估之結果較差，但是當樣本個數增加時，參數推估之精確度亦隨之增加，故本研究採用之近似最大似法在大樣本之參數推估上，可以得到不錯

表 3 合成資料各模式判斷準則選取正確階數之組數 (樣本數 500)

模式	模式選取準則	
	AIC	BIC
1-factor GARMA(0,0)	74	99
1-factor GARMA(1,0)	80	98
1-factor GARMA(0,1)	71	97
1-factor GARMA(2,0)	91	99
1-factor GARMA(0,2)	73	98
1-factor GARMA(1,1)	80	98
2-factor GARMA(1,0)	60	67

之結果。

#### 4-1-3 $k$ -factor GARMA 模式之模式判定

在模式判定方面，序列經轉換為短記憶性後所擬合之最佳 ARMA 模式，本研究採用 AIC 及 BIC 兩種判斷準則加以偵測，其結果如表 3 所示。考慮之擬合模式依序為 ARMA(0,0)、ARMA(1,0)...至 ARMA(2,2)共計 9 種模式。其結果顯示當樣本大小為 250 時，除 2-factor GARMA(1,0) 模式之 AIC 選取正確組數高於 BIC 外，其餘選取結果皆為 BIC 較佳，AIC 大約能選取到 70% 左右之正確組數，而 BIC 方面則可以選取到 90% 以上之正確組數。而當樣本大小增為 500 時，各準則之選取結果皆隨之增加，此一現象表示準則之選取結果受到資料長度之影響很大。另外，針對 2-factor GARMA(1,0) 模式， $\phi_1 = 0.3$  時，即使樣本數增至 500 時仍無法有效選取正確模式，BIC 能選取之正確組數僅有 67 組，此為參數不夠顯著，造成 BIC 容易低估之緣故。整體而言，BIC 所得之結果明顯較 AIC 更適於  $k$ -factor GARMA 模式之判定，故本研究之實測資料部份將採用 BIC 為模式判定之準則。另本研究並未對參數顯著程度於模式判定之影響進行整體探討，不過由上述結果可知傳統判定準則仍適用於  $k$ -factor GARMA 模式之判定。

#### 4-2 實測資料

為探討長記憶性  $k$ -factor GARMA 模式於實

測資料上之適用性，本研究採用臺灣 13 站月流量資料進行分析，主要項目包括模式之適合度、統計特性保存能力及模式預測能力等，其結果並與傳統除勢模式加以比較。其中統計特性包括實測資料之平均值、變異數及偏態係數；預測為取前 35 年之資料推估模式，剩餘 5 年資料採向前一步預測(One-Step-Ahead Forecasting)，進行模式預測能力之探討。本研究為滿足 GARMA 模式之定常性條件，在分析  $k$ -factor GARMA 模式時，將先對實測資料進行對數轉換，使其接近定常性。

#### 4-2-1 $k$ -factor GARMA 模式之模式判定

$k$ -factor GARMA 模式之判定，首先遭遇之困難為  $k$  值之決定，其餘之參數則較無爭議。由實測資料之頻譜圖如圖 4 所示，較為顯著之頻率僅有一個(相對應之週期為 12)，故本研究之  $k$ -factor GARMA 模式判定一律採用  $k = 2$ ，用以描述一主要頻率及一次要頻率，其模式判定結果如表 4 所示。就模式適合度而言，在  $k$ -factor GARMA 模式部份，桶頭、北港、山腳、蘭陽大橋未通過殘差之 Q 檢定( $\alpha = 0.05$ )，代表此四站在資料轉換後之 ARMA 模式判定上仍有偏差，可能為轉換方式無法有效地將序列轉為定常性序列之緣故；而針對殘差週期性檢定部份，各站模式皆通過 C-P 檢定，因此在長記憶項之判定上並無重大問題。最後在模式之參數個數方面， $k$ -factor GARMA 模式所需之參數(7~9 個)明顯較除勢模式(26~27 個)為少，其模式較為精簡。

#### 4-2-2 $k$ -factor GARMA 模式之統計特性保存能力

各站模式之基本統計特性保存能力如表 5 所示，在  $k$ -factor GARMA 模式之資料偏態係數保存能力方面，除打鹿坑、桶頭、北港、台東大橋四站之偏態係數過高外，其餘各站均較除勢模式為佳，而除勢模式之偏態係數普遍偏低，整體而言，原始資料之偏態係數大約在 2.19~4.58 間，而除勢模式與  $k$ -factor GARMA 模式資料偏態之範圍則分別為 1.17~2.21 與 2.35~5.43，相較之下後者較接近原始資料。另就資料平均值及變異數而言，除三峽及台東大橋兩站以外，除勢模式均較  $k$ -factor GARMA 模式為佳，此乃除勢模



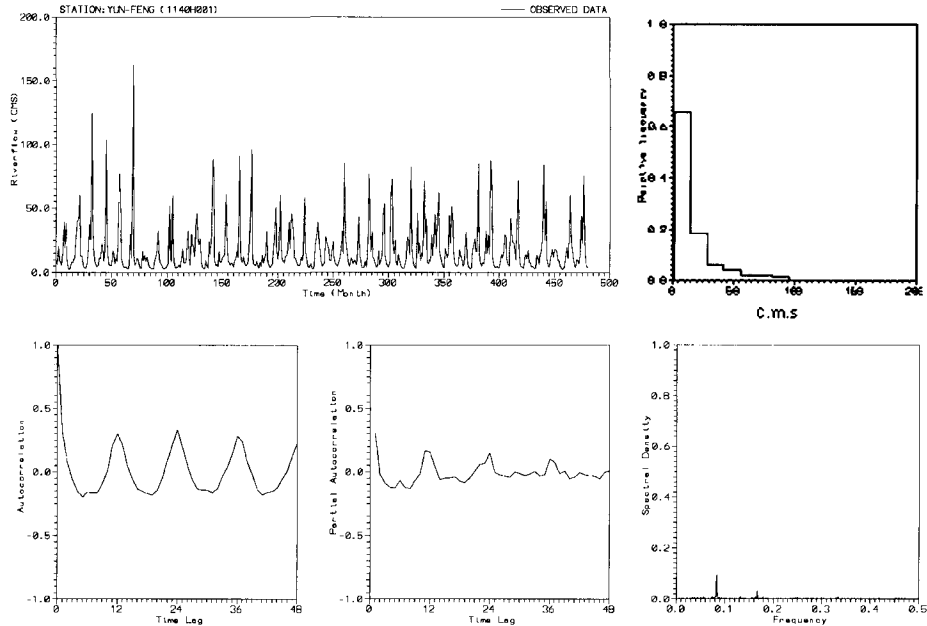


圖 4 玉峰站之月流量資料圖 (1958.01~1997.12)

式在分析時掌握了資料之週期平均值及變異數之緣故。

#### 4-2-3 $k$ -factor GARMA 模式之預測能力

本研究採用之臺灣月流量資料，受到季節性降雨如梅雨、冬雨與颱風等影響，資料之變異相當大，故豐水期、枯水期間之差異相當明顯，因此在水資源之利用上，合理預估未來之水文時序以提供水資源系統規劃上之參考，為目前水文研究之重要課題。

長記憶性  $k$ -factor GARMA 模式於預測上之表現如表 6 所示，結果顯示在 13 個測站當中，玉峰、福山、秀巒、高義、北港、延平、台東大橋及山腳等八站無論在何種預測指標上， $k$ -factor GARMA 模式之預測結果均優於除勢模式；若針對預測指標 MAPE 而言， $k$ -factor GARMA 模式之結果皆優於除勢模式，此為除勢模式將過去高流量之訊息過份引入，以致於逢枯水年時其豐水期之流量較易高估，而  $k$ -factor GARMA 模式則無此現象。若將預測結果依照所對應之實測資料大小依序排列，如圖 5 所示，圖中折線代表實測

資料與預測結果之比較，而垂直線則表示各時間位置上之預測誤差，由圖中得知  $k$ -factor GARMA 模式在低流量部份之預測結果較除勢模式為佳，而於高流量之部份兩種模式普遍無法準確預估。

根據上述結果， $k$ -factor GARMA 模式於預測能力上優於除勢模式，且可改善傳統除勢模式較無法掌握之偏態保存能力，雖於平均值及變異數上之保存能力較差，此可能為資料於轉換時無法滿足定常性條件，或轉換後資料分布仍為非常態所致，造成模式選取上有所偏差，若找出一良好之轉換方式，應可改善其結果。故整體而言， $k$ -factor GARMA 模式在臺灣河川月流量資料之分析上應具有其適用性。

## 五、結論與建議

### 5-1 結論

本研究分別針對長記憶性合成資料及臺灣河川月流量資料進行分析，其結果可歸納如下：

(1)  $k$ -factor GARMA 模式之選取以 BIC 準則為

表 4 各流量站之模式適合度比較

測站	模式	模式參數		模式適合度	
		型態	個數	Q-test	CP-test
玉峰	GARMA	$(1-1.74B+B^2)^{0.25}(1-1.00B+B^2)^{0.11}(\text{Ln}Y_t-2.487) = a_t$	7	Pass	Pass
	除勢模式	$(1-0.358B)Z_t = a_t$	26	Pass	Pass
稜角	GARMA	$(1-1.74B+B^2)^{0.31}(1+0.96B+B^2)^{0.09}(\text{Ln}Y_t-1.672) = a_t$	7	Pass	Pass
	除勢模式	$(1-0.361B)Z_t = a_t$	26	Pass	Pass
福山	GARMA	$(1-1.74B+B^2)^{0.36}(1+B^2)^{0.12}(\text{Ln}Y_t-2.621) = (1-0.321B)a_t$	8	Pass	Pass
	除勢模式	$Z_t = (1+0.295B)a_t$	26	Pass	Pass
秀巒	GARMA	$(1-1.74B+B^2)^{0.25}(1-1.00B+B^2)^{0.11}(\text{Ln}Y_t-1.332) = a_t$	7	Pass	Pass
	除勢模式	$(1-0.335B)Z_t = a_t$	26	Pass	Pass
高義	GARMA	$(1-1.74B+B^2)^{0.35}(1-1.00B+B^2)^{0.14}(\text{Ln}Y_t-3.042) = (1-0.309B)a_t$	8	Pass	Pass
	除勢模式	$(1-0.347B)Z_t = a_t$	26	Pass	Pass
三峽	GARMA	$(1-1.74B+B^2)^{0.19}(1+B^2)^{0.07}(\text{Ln}Y_t-1.866) = a_t$	7	Pass	Pass
	除勢模式	$(1-0.221B)Z_t = a_t$	26	Pass	Pass
打鹿坑	GARMA	$(1+0.301B)(1-1.74B+B^2)^{0.44}(1+1.00B+B^2)^{0.13}(\text{Ln}Y_t-1.838) = a_t$	8	Pass	Pass
	除勢模式	$(1-0.298B)Z_t = a_t$	26	Pass	Pass
桶頭	GARMA	$(1-1.74B+B^2)^{0.49}(1+1.00B+B^2)^{0.08}(\text{Ln}Y_t-2.033) = (1-0.392B)a_t$	8	No Pass	Pass
	除勢模式	$Z_t = (1+0.215B+0.227B^2)a_t$	27	Pass	Pass
北港	GARMA	$(1-1.74B+B^2)^{0.38}(1-1.00B+B^2)^{0.38}(\text{Ln}Y_t-2.479) = (1-0.765B+0.500B^2)a_t$	9	No Pass	Pass
	除勢模式	$(1-0.239B)Z_t = a_t$	26	Pass	Pass
延平	GARMA	$(1-1.74B+B^2)^{0.46}(1+0.54B+B^2)^{0.03}(\text{Ln}Y_t-3.064) = (1-0.360B)a_t$	8	Pass	Pass
	除勢模式	$(1-0.414B)Z_t = a_t$	26	Pass	Pass
台東大橋	GARMA	$(1-1.74B+B^2)^{0.39}(1+1.84B+B^2)^{0.07}(\text{Ln}Y_t-3.788) = a_t$	7	Pass	Pass
	除勢模式	$(1-0.429B)Z_t = a_t$	26	Pass	Pass
山腳	GARMA	$(1+0.310B)(1-1.74B+B^2)^{0.38}(1+B^2)^{0.19}(\text{Ln}Y_t-1.414) = a_t$	8	No Pass	Pass
	除勢模式	$(1-0.315B-0.031B^2)Z_t = a_t$	27	No Pass	Pass
蘭陽大橋	GARMA	$(1-1.74B+B^2)^{0.29}(1+B^2)^{0.11}(\text{Ln}Y_t-3.650) = a_t$	7	No Pass	Pass
	除勢模式	$(1-0.409B)Z_t = a_t$	26	Pass	Pass

註：  $Z_{v,h} = (Y_{v,h} - \bar{Y}_h) / S_h$

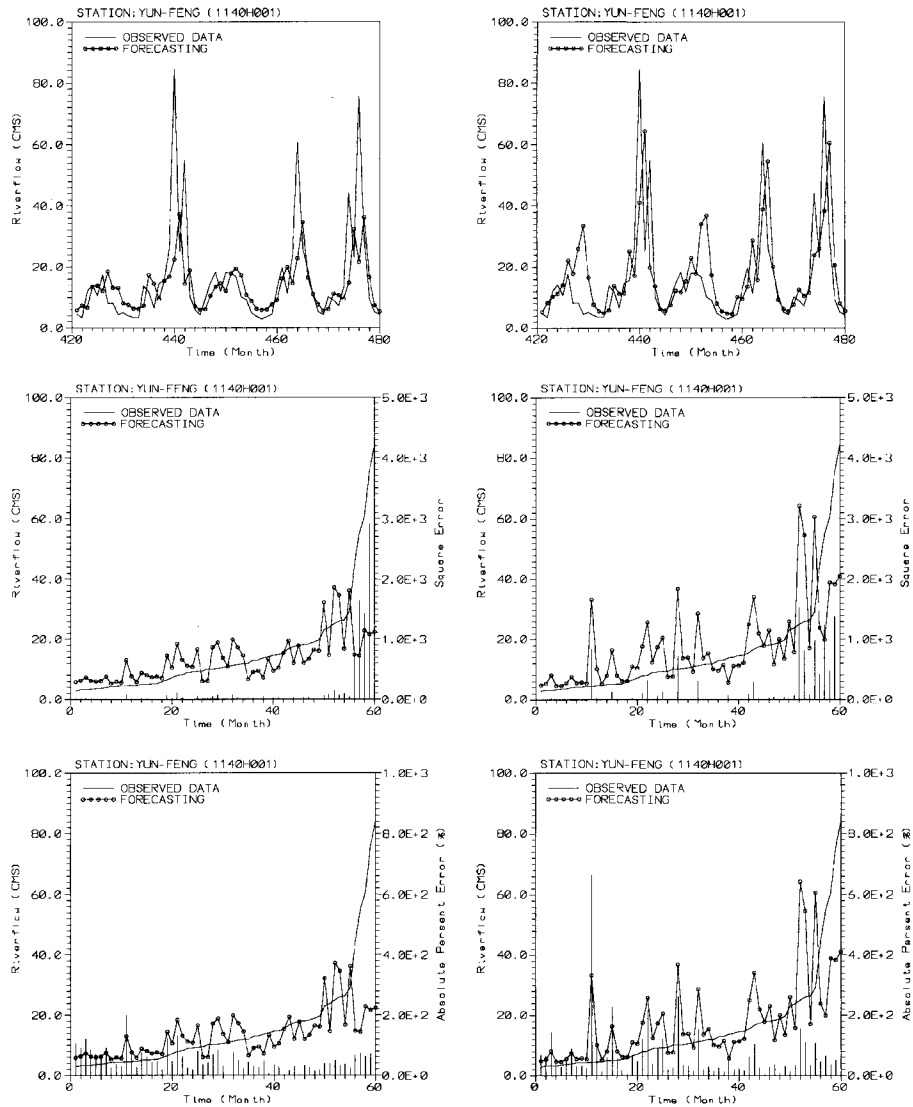


圖 5 玉峰站月流量資料預測結果， $k$ -factor GARMA 模式 (左)、除勢模式 (右) (中、下圖為依實測資料由小至大排列之預測結果及預測誤差)

佳。

- (2) 在月流量資料之模式適合度方面，本研究探討之  $k$ -factor GARMA 模式可以有效地將長記憶性水文時序轉換為短記憶性，並減少模式所需之參數個數。
- (3) 在統計特性保存能力方面，除勢模式在平

均值與變異數之保存能力上表現較佳，但是就偏態係數之保存能力而言， $k$ -factor GARMA 模式則優於除勢模式。

- (4) 在預測能力方面，整體而言以  $k$ -factor GARMA 模式之結果較佳，且於低流量部份之預測結果亦優於除勢模式。

表 5 各流量站統計特性保存能力比較

測站	模式	平均值	變異數	偏態係數
玉峰	原始序列	17.67	380.78	2.85
	GARMA	16.17	227.36	3.05
	除勢模式	17.50	372.57	1.60
稜角	原始序列	7.11	48.15	2.82
	GARMA	6.78	31.03	2.62
	除勢模式	7.06	47.17	1.60
福山	原始序列	17.71	242.51	2.74
	GARMA	16.88	149.59	2.35
	除勢模式	17.68	242.91	1.66
秀巒	原始序列	6.03	58.41	4.23
	GARMA	5.70	40.32	3.59
	除勢模式	5.96	57.18	1.70
高義	原始序列	30.26	1200.95	3.40
	GARMA	28.24	708.47	3.03
	除勢模式	29.97	1171.33	1.82
三峽	原始序列	8.81	60.97	2.33
	GARMA	8.81	69.89	3.16
	除勢模式	8.75	60.16	1.17
打鹿坑	原始序列	13.36	370.66	2.92
	GARMA	13.97	576.38	5.03
	除勢模式	13.22	363.27	1.51
桶頭	原始序列	18.97	742.59	2.34
	GARMA	16.99	776.67	4.42
	除勢模式	18.88	737.56	1.74
北港	原始序列	24.45	931.48	2.19
	GARMA	24.03	1632.46	5.43
	除勢模式	24.28	918.50	1.51
延平	原始序列	33.88	1730.86	2.52
	GARMA	35.25	2248.02	3.92
	除勢模式	33.63	1707.30	1.33
台東大橋	原始序列	90.50	14017.75	2.39
	GARMA	90.49	25277.48	5.35
	除勢模式	89.86	13868.87	1.18
山腳	原始序列	5.70	44.63	4.58
	GARMA	5.33	20.33	2.66
	除勢模式	5.67	43.12	2.21
蘭陽大橋	原始序列	57.32	3719.34	2.62
	GARMA	54.10	2767.41	3.19
	除勢模式	57.02	3694.27	1.72

表 6 各流量站模式預測能力比較

測站	模式	MSE	MAPE	UI
玉峰	GARMA	203.9	51.4	0.63
	除勢模式	205.0	70.2	0.63
稜角	GARMA	37.1	42.5	0.61
	除勢模式	36.0	58.8	0.60
福山	GARMA	104.2	41.0	0.52
	除勢模式	114.7	55.4	0.54
秀巒	GARMA	103.3	59.4	0.82
	除勢模式	106.8	87.1	0.83
高義	GARMA	610.0	57.3	0.62
	除勢模式	680.0	78.9	0.65
三峽	GARMA	46.3	66.8	0.60
	除勢模式	38.8	88.0	0.55
打鹿坑	GARMA	256.2	61.8	0.70
	除勢模式	208.8	80.9	0.63
桶頭	GARMA	217.6	56.8	0.55
	除勢模式	191.8	68.7	0.51
北港	GARMA	190.9	37.0	0.55
	除勢模式	218.7	60.1	0.59
延平	GARMA	383.7	113.9	0.66
	除勢模式	566.1	121.7	0.80
台東大橋	GARMA	6011.9	110.6	0.70
	除勢模式	6714.7	141.6	0.74
山腳	GARMA	4.8	45.7	0.43
	除勢模式	13.0	68.0	0.71
蘭陽大橋	GARMA	1627.6	48.8	0.55
	除勢模式	1542.4	67.3	0.54

## 5-2 建議

- (1)  $k$  值之決定，可配合週期檢定及頻譜分析加以研究。
- (2) 本研究採用之實測資料為臺灣河川月流量資料，亦可考慮使用臺灣河川旬流量資料加以分析。

## 參考文獻

1. 虞國興、莊明德，「長記憶水文時序之分析—新方法」，臺灣水利季刊，第三十七卷，第三期，第 23 頁-32 頁，1989。

2. 虞國興，莊明德，「長記憶水文時間序列模式之探討」，臺灣水利季刊，第三十七卷，第四期，第 37 頁-49 頁，1989。
3. 虞國興、劉治均，「時間序列模式選取準則之比較」，農業工程學報，第四十一卷，第一期，第 15 頁-26 頁，1995。
4. Akaike, H. (1973), Information theory and an extension of the maximum likelihood principle, *Proc. 2<sup>nd</sup> International Symposium on Information Theory*, (Eds. B. N. Petrov and F. Csaki), 267-281, Akademiai Kiado, Budapest.
5. Akaike, H. (1974a), A new book at the statistical model identification, *IEEE Transactions on Automatic Control*, **AC-19**, 716-723.
6. Box, G. E. P., Jenkins, G. M. and Reinsel, G. C. (1994), *Time Series Analysis: Forecasting and Control*, Printice Hall, New Jersey.
7. Chung, C. F. (1996), "A Generalized Fractionally Integrated Autoregressive Moving-Average Process", *Journal of Time Series Analysis*, Vol. 17, No. 2, Pages 111-140.
8. Ford, C. R. (1991), The Gegenbauer and Gegenbauer Autoregressive Moving Average Long-Memory Time Series Models. Ph.D. Dissertation, Southern Methodist University.
9. Gray, H. L., Zhang, N. F. and Woodward, W. A. (1989), "On Generalized Fractional Process", *Journal of Time Series Analysis*, Vol. 10, Pages 233-257.
10. Hosking, J. R. M. (1984), "Modeling Persistence In Hydrological Time Series Using Fractional Differencing", *Water Resources Research*, Vol.20, No. 12, Pages 1898-1908.
11. Magnus, W., Oberhettinger, F. and Soni, R. P. (1966), *Formulas and Theorems for the Special Functions of Mathematical Physics*, Berlin: Springer.
12. Ramsey, F. L. (1974), Characterization of the partial autocorrelation function. *Ann. Stat.* 2, 1296-1301.
13. Shibata, R. (1976), Selection of the order of an autoregressive model by Akaike's information criterion, *Biometrika*, **63**, 117-126.
14. Wei, W. W. S. (1994), *Time Series Analysis: Univariate and Multivariate Methods*, Addison Wesley, Canada.
15. Woodward, W. A., Cheng, Q. C. and Gray, H. L. (1998), "A  $k$ -Factor GARMA Long-Memory Model", *Journal of Time Series Analysis*, Vol. 19, No. 4, Pages 485-504.

收稿日期：民國 89 年 9 月 21 日

接受日期：民國 89 年 10 月 30 日