大葉學報　第十一卷　第二期　民國九十一年

Journal of Da-Yeh University, Vol. 11, No. 2, pp. 115-127 (2002)

應用灰色理論於時間序列轉折點之分析與預測

羅傑瀛　林彥宏　王正賢
大葉大學工業工程學系
彰化縣大村鄉山腳路 112 號

摘　要

灰色預測模型之特點為只需要少量的歷史資料即可達到預測的效果。然而，單純之灰色預測模型 GM(1,1) 雖然可以對未來變動之趨勢有大致之預測成效；但是，在趨勢曲線之轉折處往往無法達到更精確的預測效果。因此，本研究即針對此現象提出改良之方法。首先，可利用對所有造成趨勢轉折之因素進行分析，並轉換成量化的因素影響權重值，同時使用此數值將實際數據間的波動予以減緩。同時，在經過 GM(1,1) 進行預測後，對於預期在下一時間點產生影響之因素加以考量，而求得下一時間點之預測值。此外，由於預測之誤差是不可避免，因此本研究也利用自迴歸模型對於誤差之變化予以模式化，進而能將下一期之預測結果進行修正調整。透過實際之數據資料進行驗證，改良後之 GM(1,1) 預測方法與誤差修正方法之結合確能有效提升原始 GM(1,1) 模型之預測精準度，同時，實例中也同時說明此方法較一般傳統之方法有更好之預測績效。

關鍵詞：灰色預測，影響權重，自迴歸模型，台灣股價指數

An Analysis and Prediction of Turning Points in A Time Series Based on the Grey Theory

CHIEH-YING LO, YEN-HUNG LIN and JEAN-SHYAN WANG
Department of Industrial Engineering, Da-Yeh University
112 Shan-Jiau Rd., Da-Tsuen, Changhua, Taiwan

ABSTRACT

The advantage of the Grey Model is that it can obtain a good forecasting effect as long as a little historical data are provided. The simple Grey Model GM(1,1) can forecast changeable trends; however, it may produce serious errors at the turning points in a curve. Hence, the aim of this research is to discover a solution to improve the condition just descrived. First, the factors which cause the turning points can be analyzed and digitalized as influential weights. In addition, the usage of these factors can also make the historical data smoother. Through a forecast from GM(1,1), these factors still need to be added to the GM(1,1) Model forecasting values at the next time point so that the forecasting values can be produced there. Meanwhile, deviation in any forecast cannot be avoided; hence; this research also uses the Auto-regression Model to forecast deviation and modify it from the forecasting values. Actual data can be used to verify the research. The accuracy of the

GM(1,1) Model is indeed increased by a combination of an improved GM(1,1) Model and deviation modification. The examples in this research are verified to be, indeed, better than the traditional forecasting models.

Key Words**:** grey model, influential weight, auto-regression model

一、前言

在過去中，許多有關於預測的數學理論模型紛紛被提出，例如：各種迴歸分析、時間序列分析等傳統的方法，乃至於近年來廣泛被提及與應用的模糊預測理論或灰色預測理論等等。各種模型之提出，其目的皆是為了能更廣泛地運用在各領域上，且能達到更準確的預測效果。在各傳統方法中，需要收集長期的、大量的歷史資料，以建立適當之預測模型，但在現今凡事求新求快的時代中，由於其及時反應之能力不佳，所以對短期且及時之預測問題並無法提供可靠的預測成效。也因此，灰色理論之提出即是針對此一需求而產生。

灰色理論是由鄧聚龍教授所提出 [18]，理論中包含了各種不同之應用領域，其中灰色預測為極為重要之一部分。其中，GM(1,1) 即為此理論中最為廣泛應用之預測模型。在過去之研究中，對於經濟、財務、工程等領域皆有理想之成效。相較於傳統之預測模型需要大量的歷史資料才可建立預測模型，GM(1,1) 不需要大量資料即可快速的建立模型，根據其模型之定義，最少只需要四筆的數據即可建立模型並進行預測，而這也是此模型之最大特點。雖然其模型簡單而且只需要少筆的歷史資訊，但其預測精準度在不同的情況下會產生不同的效果，例如，用來預測波動起伏劇烈之數據時，雖然預測值和實際值有著相似的趨勢，但是在趨勢曲線轉折點會產生高估或低估的現象，也就是說，持續遞增或遞減後出現之反轉現象，其在反轉點之下一期的預測結果必定出現與實際值反向的趨勢，而如此之特性亦造成了在應用上之成效並不盡理想。因此，本研究即針對此一現象進行分析並提供改善之方法。

首先，對於所有造成趨勢轉折之因素進行蒐集與分析，並利用灰關聯分析與模糊階層分析法進行模式化之處理，其中，灰關聯分析用以處理可形成序列型態之量化型因素；而模糊階層分析法則用於處理非量化型之因素。透過兩種方法之應用，可將相關語意概念之因素轉換成量化的因素影響權重值，同時使用此數值對原始之實際數據進行修正，以減弱數據間之波動程度。同時，在經由灰色預測之模型而求得之預測結果後，也可利用在下一時間點預期會產生影響之因素的影響權重值來修正預測之結果，以期望能反應這些因素對未來所產生之影響效果。此外，由於預測之誤差仍舊是無法避免的且誤差之出現為隨機型態，因此本研究也利用自迴歸模型對於誤差序列之變化予以模式化，而誤差序列的產生是根據過去幾期預測值與實際值之差距而建立。透過模式化之模型，因而能對未來之誤差進行預估，進而將下一期之預測結果進行修正調整。

最後，本研究將以台灣股價加權指數為例驗證上述之方法；同時，預測之結果將分別與傳統之預測方法，例如迴歸分析與時間序列分析，和原始單純之灰預測模型進行比較，透過模型間之比較，以瞭解本研究中之方法對於波動型數據之預測成效。

二、灰色預測理論與模型

灰色理論中最重要的應用領域即為灰色預測之提出，灰色之意函即代表了事件之不確定性或資訊之不完全性。在不確定或不完全的情境下，試圖對未知的未來進行推估或判斷，即為灰色預測之目的。

在灰色預測理論中，最廣泛使用之預測模型為 GM(1,1)，其括弧中符號之意義為一階單變數 [1, 6, 11, 12, 18]。今假設 $x^{(0)}$ 為原始序列

$$x^{(0)} = (x^{(0)}(1), x^{(0)}(2), ..., x^{(0)}(n)), \quad n \geq 4 ,$$

令 $x^{(1)}$ 為 $x^{(0)}$ 之轉換序列，且

$$x^{(1)} = (x^{(1)}(1), x^{(1)}(2), ..., x^{(1)}(n)) 。$$

若滿足

$$x^{(1)}(k) = \sum_{m=1}^{k} x^{(0)}(m) \tag{1}$$

則上述轉換稱 $x^{(1)}$ 為 $x^{(0)}$ 之累加生成（accumulated generating operation，簡稱 AGO）序列。累加生成之目的為

藉著改變原始序列之層次以發現變動之規律，使得序列能更易於建構適當之模型，相關之研究可參考 [11]。則灰色預測模型可由上述轉換序列建構，並可表達成以下之方程式：

$$x^{(0)}(k)+aZ^{(1)}(k)=b \qquad (2)$$

$$Z^{(1)}(k)=0.5x^{(1)}(k)+0.5x^{(1)}(k-1) \qquad (3)$$

其中 $Z^{(1)}(k)$ 稱為白化背景值；a、b 為未定之常數。依據灰導數之定義 [11, 12]，則公式（2）可近似於下列公式：

$$\frac{dx^{(1)}}{dt}+ax^{(0)}=b \qquad (4)$$

因此可依照一般微分方程式求解方法對上式求解，可得方程解為

$$\hat{x}^{(1)}(k+1)=\left(x^{(0)}(1)-\frac{b}{a}\right)e^{-ak}+\frac{b}{a} \qquad (5)$$

此解又被稱為 GM(1,1) 之白化響應式。由於此響應式是由累加生成序列所得，因此需將其還原成原始之序列，此步驟稱為逆累加生成（Inverse AGO），則可得最終之預測值為

$$\hat{x}^{(0)}(k+1)=\hat{x}^{(1)}(k+1)-\hat{x}^{(1)}(k) \qquad (6)$$

在公式(5)中尚有兩未定參數需決定，其分別可依照下列公式求出：

$$a=\frac{\sum_{k=2}^{n}z^{(1)}(k)\sum_{k=2}^{n}x^{(0)}(k)-(n-1)\sum_{k=2}^{n}z^{(1)}(k)x^{(0)}(k)}{(n-1)\sum_{k=2}^{n}\left[z^{(1)}(k)\right]^2-\left[\sum_{k=2}^{n}z^{(1)}(k)\right]^2} \qquad (7)$$

$$b=\frac{\sum_{k=2}^{n}[z^{(1)}(k)]^2\sum_{k=2}^{n}x^{(0)}(k)-\sum_{k=2}^{n}z^{(1)}(k)\sum_{k=2}^{n}z^{(1)}(k)x^{(0)}(k)}{(n-1)\sum_{k=2}^{n}[z^{(1)}(k)]^2-[\sum_{k=2}^{n}z^{(1)}(k)]^2} \qquad (8)$$

三、灰關聯分析法與模糊階層分析法

本節將分別使用灰關聯分析法（grey relational analysis, GRA）與模糊階層分析法（fuzzy analytical hierarchy process, Fuzzy AHP）對量化因素與非量化因素進行分析，以了解各相關因素對主要變數之影響程度。同時，並對兩種分析法下之不同測度間之一致性問題進行說明。

（一）灰關聯分析法

1. 一般性灰關聯度

灰關聯分析法 [10] 為灰色理論中另一個重要的分析方法。其目的為探究兩序列間之關聯程度，並以灰關聯度（grey relational grade, GRG）表達。

若今有 m 個序列，且各序列各有 n 個元素，如

$$X=\{x_0,x_1,\cdots,x_{m-1}\},$$

$$x_i=[x_i(1),x_i(2),\cdots,x_i(n)]$$

當有一序列 x_0 為參考序列，其他序列為比較序列時，則灰關聯係數可定義為：

$$\gamma(x_0(k),x_i(k))=\frac{\Delta_{\min.}+\varsigma\,\Delta_{\max.}}{\Delta_{0i}(k)+\varsigma\,\Delta_{\max.}} \qquad (9)$$

$$i=1,\cdots,m-1 \quad k=1,\cdots,n$$

$$\Delta_{\min.}=\min_{\forall i}.\min_{\forall k}.\left|x_0(k)-x_i(k)\right| \qquad (10)$$

$$\Delta_{\max.}=\max_{\forall i}.\max_{\forall k}.\left|x_0(k)-x_i(k)\right| \qquad (11)$$

其中

x_0 為參考序列；

x_i 為一特定之比較序列；

Δ_{0i}：x_0 和 x_i 之間第 k 個元素差的絕對值；

$\Delta_{\min.}$：各比較序列與參考序列間各元素差之最小的最小值；

$\Delta_{\max.}$：各比較序列與參考序列間各元素差之最大的最大值；

ζ：為辨識係數並介於 0 與 1 之間。

而辨識係數（ζ）的功能主要是作參考物和待測物之間的對比，其大小可根據實際的需要作適當之調整，由實際的數學證明得知其只會改變相對數值的大小，不會影響灰關聯度的排序。故一般而言，ζ 值均取 0.5 附近。

當求得灰關聯係數之後，一般取灰關聯係數之平均值為灰關聯度：

$$\gamma(x_0,x_i)=\frac{1}{n}\sum_{k=1}^{n}\gamma(x_0(k),x_i(k)) \tag{12}$$

但在實際系統上, 各個因子對系統的重要程度並不一定見得完全相同，所以我們正視各個因子的權重不相等的實際情形，將延伸上式的灰關聯度的定義為：

$$\gamma(x_0,x_i)=\sum_{k=1}^{n}\beta_k\gamma(x_0(k),x_i(k)) \tag{13}$$

其中 β_k 表示因子 k 的常態化權重，而且 $\sum_{k=1}^{n}\beta_k=1$。

如同前文中所述, 灰關聯度為一排序用之測度, 因此其為定性化之數值;但在本研究中希望能得到定量化之數值以代表相關因素之影響權重, 換句話說, 定性化之灰關聯度必須加以轉換成為定量化之灰關聯度以符合本研究中之應用。在文獻 [18] 中討論了灰關聯度之定量化問題，經由文獻中之證明得知，定量化之灰關聯度可由下式表達：

$$\gamma_{0i}=\gamma(x_0,x_i)=\frac{\Delta_{\min.}+\Delta_{\max.}}{\overline{\Delta}_{0i}(k)+\Delta_{\max.}} \tag{14}$$

$$\overline{\Delta}_{0i}(k)=\frac{1}{n}\sum_{k=1}^{n}[\Delta_{0i}(k)] \tag{15}$$

因此本研究將以定量化之灰關聯度為灰關分析之基礎, 如此可將灰關聯度視為各相關因素對主要變數之影響權重。

2. 利用比率值生成序列之相似性灰關聯度

對於兩序列間之關係而言, 可粗略區分為相近性與相似性 [15]。相似性灰關聯度是不同於以序列間接近性測量方式為主的灰關聯分析, 其為一種對序列間的相似性質的另一種動態測量的形式。在上述之一般性之灰關聯度中, 其測度可由兩序列間元素差而得之；換言之, 一般性灰關聯度為兩序列間相近性之概念。因此, 本研究將針對另一序列間相似性之概念提出以比率值為基礎之相似性灰關聯度。

本研究利用了徐守德、黃玉娟、余明芳之論文 [5] 中所運用之方法提出了另一種序列生成方式，稱為比率法生成，如下所列

$$\mathrm{R}_i={P_{i+1}}/{P_i} \tag{16}$$

其中 R 代表比率；i 代表第 i 筆數據；P 則代表各序列的值。

則生成序列可表示成

$$\tilde{x}_i(k)=\left[\frac{x_i(k)}{x_i(k-1)}\right], \tag{17}$$

$i=1,\cdots,m-1$，　$k=1,2,\cdots,n$

生成後每一項因素的比率序列, 再利用灰關聯之公式求出灰關聯度。如此生成的方法可以隨著序列遞增或遞減的特性, 求出不同的增減比率, 進而找出序列間的相似程度。也因為一般性灰關聯本身是根據序列間接近性質的數學模式, 所以利用此方式可找出序列間增減比率的接近關聯度, 如此即為趨勢之相關程度。

3. 最佳化灰關聯度

最佳化灰關聯度之觀念來自 Chang [16], 其觀念是一種同時結合相近性與相似性灰關聯度, 並以兩者之線性組合表示。最佳化灰關聯度是一種獨立測量分析序列間關係的方法, 其被用於測量序列間之非線性關係, 並經由最大熵值法以求得相近性與相似性灰關聯度的最佳組合。兩種不同觀念之灰關聯度的最佳化線性組合如下所示：

$$\Gamma_{0i}=\alpha A_{0i}+\beta S_{0i} \tag{18}$$

其中

Γ_{0i} ： A_{0i} 和 S_{0i} 兩種不同灰關聯度之最佳化灰關聯度；

A_{0i} ：接近性概念之灰關聯度；

S_{0i} ：相似性概念之灰關聯度。

由公式（18）得知, 兩種灰關聯度之發生機率可由下式表示：

$$p_1=\frac{\alpha A_{0i}}{\alpha A_{0i}+\beta S_{0i}} \tag{19}$$

$$p_2=\frac{\beta S_{0i}}{\alpha A_{0i}+\beta S_{0i}} \tag{20}$$

如果 P_1 與 P_2 為熵值方程式, 則可利用最大熵值法解出參數,

$$\begin{aligned}&\textbf{MAX.}\ H=H_1+H_2=-p_1\ln p_1-p_2\ln p_2\\&\quad\textbf{s.t.}\ \ p_1+p_2=1\\&\qquad\quad \alpha+\beta=1\end{aligned} \tag{21}$$

因此，最佳化關聯度 Γ_{0i} 為

$$\Gamma_{0i}=\frac{2A_{0i}S_{0i}}{A_{0i}+S_{0i}} \tag{22}$$

因此可利用此結合公式之結果，針對想要結合之灰關聯特性，予以結合，以更客觀的評估出序列與序列間的關聯程度。

（二）**模糊階層分析法**

層級分析法是美國作業研究家Satty於70年代所提出的理論 [9]，是一種定性與定量的分析相結合的多目標規劃分析方法。但層級分析法在準則評價上的分析觀點上是以一個數值來代表，但如果評估者認為其評價值應該界於兩個評估使度之間時 AHP 便無法解決諸如此類據模糊性的問題，因此便有模糊階層分析法的誕生。

模糊 AHP 法是近幾年所衍生的方法，在我們所學過的數學裡面，對概念必須給出一個明確的定義，但是在生活中，我們對於一些事物並不能以明確的字眼或名稱說明一件事或物。例如：你覺得那個人是高或矮，並不能絕對性的說明，但可以絕對的說出是比誰高或矮，藉由這一些的描述你才能知道他到底多高或不高。所以有很多事件，它都是邊界不清楚的、模糊的、很難用數學來描述的。

在階層分析法的架構下，最重要的工具即為問卷的設計，而在問卷勾選的內容上，在傳統中大部分都是以評分或數據化的值來做勾選，而在本研究中將以比較程度上的差異來做勾選，並利用口語化的比較詞與模糊數學中的模糊數加以變化設計。而尺度的多寡也會影響問卷的價值。一般而言問卷的尺度以"五尺度"為問卷進行中，受測者比較容易且喜歡作答的問卷方式，因此本研究之評估尺度與語意概念可如表 1 中所示。然而在融合模糊理論的觀念下 [7]，本組運用語意變數轉換成模糊數的方法，讓受測者可以主觀、清楚、明確、快速的勾選出其專業認知上的評估值，而勾選出的值，經由模糊語意轉換成客觀的三角模糊數，如表 2 所示；接著再進行 AHP 架構下的矩陣運算，以求得權重值。

在論文 [4, 20] 中提到有關模糊 AHP 方法之權重計算方式，本研究經整理後列出下列數個步驟：

步驟一、群體整合

此步驟根據模糊數的運算定義，進行專家所填寫的問卷轉成模糊數的整合。

表 1. 評估尺度語意說明

評估尺度	定義	說明
1	同等重要	兩種比較方案貢獻程度同等重要
3	稍重要	經驗與判斷稍微傾向喜好某一方案
5	頗重要	經驗與判斷強烈傾向喜好某一方案
7	極重要	經驗與判斷非常強烈傾向喜好某方案
9	絕對重要	有足夠的證據肯定絕對喜好某一方案
2、4、6、8	相鄰尺度中間值	須要折衷值時

表 2. 語意措詞與模糊數之轉換

語意措辭	模糊數
極大	(7,9,9)
頗大	(5,7,9)
相等	(3,5,7)
頗小	(1,3,5)
極小	(1,1,3)

$$\tilde{M}_{ij}=\left(\frac{1}{n}\right)\otimes\left(\tilde{m}_{ij}{}^{1}\oplus\tilde{m}_{ij}{}^{2}\oplus\tilde{m}_{ij}{}^{3}\oplus\cdots\oplus\tilde{m}_{ij}{}^{n}\right) \tag{23}$$

其中

$\tilde{M}_{ij}$：整合後的三角模糊數；

n：專家人數；

$\tilde{m}_{ij}{}^{n}$：專家 n 對第 i 個評估要素與第 j 個評估要素之比較值。

步驟二、建立模糊正倒值矩陣

此步驟根據之前模糊數的運算定義，進行 AHP 架構下，以模糊數表示專家兩兩要素間相對重要程度之看法形成左右對稱的比較正倒值矩陣。

$$\tilde{M}=\left[\tilde{R}_{ij}\right]_{N\times N} \tag{24}$$

其中

$\tilde{M}$：模糊正倒值矩陣

$\tilde{R}_{ij}$：模糊數,

$\tilde{R}_{ij}=1/\tilde{R}_{ij}$，$\forall ij=1,2,3,\cdots,n$。

步驟三、計算各個因素之模糊權重

$$\tilde{Z}_i = \frac{1}{n}(\tilde{a}_{i1} \oplus \tilde{a}_{i2} \oplus \cdots \oplus \tilde{a}_{in}), \tag{25}$$

$$\tilde{w}_i = \tilde{Z}_i \otimes \left(\tilde{Z}_1 \oplus \tilde{Z}_2 \oplus \cdots \oplus \tilde{Z}_n\right)^{-1} \tag{26}$$

$\forall i = 1,2,3,\cdots,n$

$\tilde{w}_i$ ：為各個因素的模糊權重值

步驟四、將模糊數予以反模糊化

$$\tilde{w}_i = \left(L_{ij}, M_{ij}, U_{ij}\right)$$

$$DF_{ij} = \left[\frac{\left(U_{ij} - L_{ij}\right) + \left(M_{ij} - L_{ij}\right)}{3}\right] + L_{ij}, \tag{27}$$

$\forall ij = 1, 2, 3, \cdots, n$

DF_{ij} ：為反模糊化後得到的值。

（三）測度之一致性

在上述方法中，量化影響因素之權重是以灰關聯度表達；而非量化影響因素的權重是由模糊 AHP 方式所求得，針對此兩種不同方法下所產生之權重，勢必考量其一致性問題。為求得一致性之權重，在方法的使用上特別考量此一需求而有所設計。

首先，在問卷設計上，將所有量化型因素一併列入考量；換言之，所有量化因素將可分別自兩種方法架構下得到不同之權重。因此，可以利用同一因素兩種權重的關係，將其他非量化性因素的模糊 AHP 值轉變成灰關聯度權重，使其所有的影響因素皆可在同一基準點上做其他計算評比的使用。而衡量影響因素的方式採灰關聯權重，是因為灰關聯權重是根據數據序列與數據序列間實際特性而衡量，此方式下之權重為一客觀性較強之數值；反之，模糊 AHP 權重牽涉到較多主觀意念上的判斷，因此，將後者轉換成為前者將可加強各因素的客觀性。

為求得迴歸係數，可利用已知的量化性影響因素的模糊 AHP 權重值作為迴歸方程式的 X 值，再以其灰關聯值為 Y 值，找出兩種權重間的線性關係，再以此迴歸方程式為轉換式，將其他非量化因素的模糊 AHP 權重值轉換成相對應的灰關聯值，如此可得到所有因素的一致性權重。

轉換方式採用簡單之線性轉換，即建構線性迴歸方程式。線性轉換方程式可表達如下所示：

$$Y = aX + b \tag{28}$$

其中

Y： 灰關聯值；

X： 模糊 AHP 權重；

a, b： 線性迴歸方程式之係數。

四、因素之影響程度與數據調整

因為灰色理論的預測方式若用在劇烈波動的數據時，會產生預測上的缺失，尤其單純使用 GM(1,1) 的方法時發現 GM(1,1) 的預測值會在趨勢轉折處產生高估或低估的現象，這種現象即是造成大量誤差影響預測效果的主因。針對此一現象，本研究提出以影響因素調整原始數據之方式使數據波動幅度減小，如此一來便可使灰預測之模型變得更準確。

（一）預測前調整方法

根據欲消除影響因子影響程度之數據，找出每筆數據其伴隨發生之所有可能影響因素，對其做評分，評分之目的是因為利用前述方法所得每項因素之灰關聯度權重為固定值，但因素在每一筆數據發生之影響程度各有些許不同，例如影響性是遞增或者是遞減，或者影響性是不固定的增減其影響效能；也就是說，相同因素在不同數據發生時就可能產生不同的影響效果，所以給予評分以判斷因素所影響程度之大小，再與各因素的灰關聯度權重做結合以消除影響程度。調整數學模式如下：

$$T_{pure}(i) = T_i - \frac{\sum \pm A_{ij} \times W_{ij}}{\sum A_{ij} \times W_{ij}} \times T_i \times G \tag{29}$$

其中

$\pm$： 當因素為正影響取正符號，若為負影響則取負號；

$A_{i,j}$： 第 i 筆數據之第 j 項因素最佳化灰關聯值（有選取之因素才列入計算）；

$W_{i,j}$： 第 i 筆數據第 j 項因素之評分權重（有選取之因素才列入給予評分）；

T_i： 第 i 筆之數據；

$T_{pure}(i)$： 將數據消除影響因子影響程度之值；

G： 控制數據調整範圍之參數。

假若當有 m 個影響因素發生，並令所有權重和為 1，如

$$W_1 + W_2 + \cdots + W_m = 1 \tag{30}$$

又"±"之出現是因為因素可能影響有正面或反面，若為正面的影響因素則取"＋"；若為負面的影響因素即取"－"。再者，控制數據調整範圍之參數可依實際之需求做不同之設定，以達到減緩波動之目的。

此方式之目的是為了使原始數據間之大量波動的現象減緩，換言之即為平滑原始數據，如此一來便可消弱數據本身激烈波動的峰態，迎合 GM(1,1) 預測時所產生之弱點，增加利用少筆數據進行預測之效果。

（二）預測後調整方法

當預測值產生後，為反應未來預測值可能出現的波動特性，因此，本研究在此進行另一次的數據調整。因為消除影響因子影響程度之方式是以所欲剔除之該筆數據與其前一筆數據增減量為基準量，而用因素權重控制其大小來作影響因素的消除，但求出 GM(1,1) 預測值之後，因為實際值仍未發生，所以勢必無法找出欲預測之該筆數據與其前一筆數據的增減量，是故利用預測值前一筆數值乘上控制數據最大增減變動範圍之參數，再利用預測值當時之發生之影響因子發生所伴隨之影響因素權重比，放大或縮小預測當日之影響因素所影響之點數，其該扣除或增加之正負值亦由預測當日之影響因素決定，再與預測當日相加，即為還原影響因素後的預測值。其數學表示式如下：

$$\tilde{T}_{i+1} = \overline{T}_{i+1} + \frac{\sum \pm W_{i+1,j} \times A_{i+1,j}}{\sum A_{i+1,j} \times W_{i+1},j} \times T_i \times \mathrm{G} \tag{31}$$

其中

- $\pm$：當因素為正影響取正符號，若為負影響則取負號；
- $A_{i,j}$：第 i 筆數據第 j 項因素之灰關聯最佳化值（有選取之因素才列入計算）；
- $W_{i,j}$：第 i 筆數據第 j 項因素之評分權重（有選取之因素才列入給予評分）；
- T_i：第 i 筆之數據；
- $\overline{T}_{i+1}$：第 i+1 筆之經過消除影響因子之影響後的數據經過 GM(1,1) 預測後之預測值；
- $\tilde{T}_{i+1}$：第 i+1 筆之還原後預測值；
- G：控制數據最大增減之參數。

在此方法中，為了克服數據在未來時間點上之未知性，所造成計算上之困難，因此利用已發生的數值來估計未發生之數值所應有之狀態，這也是許多預測方法中所經常使用的準則之一。也就是說，透過對鄰近數據的瞭解，即可將隱含的資訊找出，以利於對未來之狀態進行推估。

五、應用自迴歸模型之誤差修正方法

在預測領域中，另一項重要的工作即是對預測結果進行合理的誤差修正。也由於誤差產生的隨機性，因此可利用 Box 與 Jenkins 教授在 1970 年所發表的時間序列中之自迴歸模型來進行處理 [2, 3]。是故利用時間序列模型來預測改良數據後之 GM(1,1) 預測值與實際值間的誤差，來對其作誤差的修正。本研究中收集了所有時間點上在預測值與實際值間的絕對誤差值來建立自迴歸預測模型，即可得到下一時間點上的誤差預測值，然後再使用它去從 GM(1,1) 的預測值中扣除，就可以得到與實際值更相似的數值，與波動趨勢曲線更接近的趨勢。

要建立一合理且適當的自迴歸模型，首先可用自相關函數（ACF）和偏自相關函數（PACF）來判斷誤差所形成之預測模型為何。

（一）自相關函數的定義與求法

在隨機變數 Z_t 和 Z_{t+k} 間的自相關函數 (ρ_k) 如，

$$\rho_k = \frac{Cov(Z_t, Z_{t+k})}{\sqrt{Var(Z_t)}\sqrt{Var(Z_{t+k})}} = \frac{r_k}{\sqrt{r_0^2}} = \frac{r_k}{r_0} \tag{32}$$

其中

$$r_k = \frac{c_k}{c_0}, k = 1,2,3,\cdots \tag{33}$$

$$r_0 = E\left(Z_t - \overline{Z_t}\right)^2 = \sigma_Z{}^2 \tag{34}$$

$$c_k = \frac{1}{N}\sum_{t=1}^{N-k}\left(Z_t - \overline{Z}\right)\left(Z_{t+k} - \overline{Z}\right) \tag{35}$$

$$c_0 = \frac{1}{N}\sum_{t=1}^{N}\left(Z_t - \overline{Z}\right)^2 \tag{36}$$

（二）偏自相關函數（PACF）的定義與求法

在時間差 k，當在 Z_t 和 Z_{t+k} 擺脫所有變數 $Z_{t+1}, Z_{t+2}, \ldots, Z_{t+k-1}$ 的影響後，再去考慮 Z_t 和 Z_{t+k} 間的自相關函數，稱 Z_t

和 Z_{t+k} 間的偏自相關函數 (ρ_{kk}),

$$\rho_{kk}=\frac{Cov\left(Z_t,Z_{t+k}\left|Z_{t+1},Z_{t+2},\cdots,Z_{t+k-1}\right.\right)}{\sigma\left(Z_t\right)\sigma\left(Z_{t+k}\right)} \tag{37}$$

利用 Cramer’s rule，得到

$$\rho_{kk}=\frac{\begin{vmatrix}1 & \rho_1 & \rho_2 & \cdots & \rho_{k-2} & \rho_1\\ \rho_1 & 1 & \rho_1 & \cdots & \rho_{k-3} & \rho_2\\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots\\ \rho_{k-1} & \rho_{k-2} & \rho_{k-3} & \cdots & \rho_1 & \rho_k\end{vmatrix}}{\begin{vmatrix}1 & \rho_1 & \rho_2 & \cdots & \rho_{k-1}\\ \rho_1 & 1 & \rho_1 & \cdots & \rho_{k-2}\\ \vdots & \vdots & \vdots & \ddots & \vdots\\ \rho_{k-1} & \rho_{k-2} & \rho_{k-3} & \cdots & 1\end{vmatrix}} \tag{38}$$

（三）建立自迴歸模型

欲找出該使用多少個時間點來建立此預測模型，自相關函數與偏自相關函數是重要的工具，需依照以下規則，

1. 自相關函數的分布會像指數分布一樣趨近於零。
2. 偏自相關函數在時間相隔 k 後皆為零。

經由自相關函數和偏自相關函數的求得後，判斷出模型之架構。自迴歸模型之基本模型如

$$Z_t=\mu+\phi_1 Z_{t-1}+\phi_2 Z_{t-2}+\cdots+\phi_n Z_{t-n} \tag{39}$$

其中 Z_t, Z_{t-1}, …, Z_{t-k} 為在每一時間點的隨機誤差。

當決定模型後，其模型之各項係數亦要被確定，所以利用求迴歸係數的方式；首先，先建立一誤差值經由自迴規模型排列之數據矩陣如下：

$$\mathbf{D}=\begin{bmatrix}1 & x_{11} & x_{12} & \cdots & x_{1k} & y_1\\ 1 & x_{21} & x_{22} & \cdots & x_{2k} & y_2\\ 1 & x_{31} & x_{32} & \cdots & x_{3k} & y_3\\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots\\ 1 & x_{n1} & x_{n2} & \cdots & x_{nk} & y_k\end{bmatrix} \tag{40}$$

再根據下列步驟計算係數

$$\mathbf{A}=\mathbf{D}'\mathbf{D} \tag{41}$$

假如 **A** 是一個 $(n+1)\times(n+1)$ 的矩陣，經過 n 次的掃出後，可以在 **A** 矩陣的最後一行得到所有係數，如

$$\mathbf{A}_{..(n)}=\begin{bmatrix}1 & 0 & 0 & \cdots & 0 & \hat{\mu}\\ 0 & 1 & 0 & \cdots & 0 & \hat{\phi}_1\\ 0 & 0 & 1 & \cdots & 0 & \hat{\phi}_2\\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots\\ 0 & 0 & 0 & \cdots & 0 & S_r\end{bmatrix} \tag{42}$$

在建立誤差預測模型後將所預測出誤差項從改良數據後之 GM(1,1) 預測值中剔除，即可得到一經由修正改善後的預測序列。

六、實例驗證與討論

本研究以台灣股價指數之預測驗證上述所提之方法，數據收集時間範圍為 2001/2/1 至 2002/1/31，而蒐集影響因素之時間範圍為 2001/10/2 至 2002/1/31，用作模型驗證時之計算依據。採用此實例之優點為資料蒐集容易，且股價之變動容易受外在因素的影響，容此更可瞭解透過因素校正後之預測效果。

首先在因素蒐集的部份，包括了容易影響股價變動的量化型因素與非量化型因素，量化型因素將直接以最佳化灰關聯度來做處理；而非量化型因素將由模糊 AHP 方式處理，同時，所有的量化型因素與非量化型因素都將列入模糊 AHP 分析法所需之問卷上，在問卷中有關量化型因素的資訊是用以當成兩種測度間轉換之基礎。最後經由專家填寫、回收、整理後求得本研究所需的數據值。

因素蒐集的過程中，從網路上各大網站所提供的股市訊息以及過去專家們所評估的因素，再加上分析師所提供的操盤意見，大致上影響國內股市的層面可分成政治面、籌碼面、基本面、技術面四大部分。本研究將之分成 21 項的影響因素，並歸納如表 3 所示，本研究找出之 21 項因素為概括性的局部影響，因素並非代表影響股市整體漲跌的控制項目，本研究亦想盡可能找出所有控制因素，但實際上非常困難，故以大部分代替，但也在給予評分權重上予以降低所佔據之比例，使其所找出之影響當比數據之因素控制量不會產生百分之百的控制，況且本研究找出此 21 項因素目的只為了消除投資人在投資時，產生之不理性的心理因素導致而成的漲跌量，所以因素只是為了使原始數據更平滑以便預測模型進行預測的工具，若將股市比喻成海，如此一來影響海浪波動的因素有風、溫度、洋流、深淺…等，此 21 項因素就好比風，比較為大眾所接受之影響因素，風大海浪大，反之則小，故先消除風之影響，即可減少多數高低起伏的大浪；

表 3. 影響股價指數之量化與非量化因素

量化因素	非量化因素	
美股那斯達克	經濟景氣燈	FED 調降利率
美股道瓊指數	台灣內政	國內突發事件
日本日經指數	台股技術面	國際情勢
恆生指數	外匯存底	國安基金
股市成交量	失業率	CPI 指數
	三大法人	台灣外交
	央行調降利率	台幣美元的波動
	經濟成長率	各類股營運狀況

故此 21 項因素並非絕對的因素，用以做輔助之目的亦使數據之轉折峰減少，如此便可將預測模型套用至此消除影響因素之數據上作預測，再予以還原產生有意義之預測值。至於盤前之影響因素用作減緩數據波動之用途；盤中因素勢必無法在預測時考慮到，但亦可作為當日調整還原數據之依據；盤後之因素即用作還原預測值的參考。

經過回收問卷的分析計算出模糊 AHP 各因素與台股間的三角模糊數值。經由反模糊化求得模糊 AHP 之權重植，如表 4。同時，針對量化型因素，以灰關聯度方法計算後可得到各量化型因素之灰關聯權重值，表 5 整理了以在各種灰關聯度觀念下所產生之結果，而本研究將以最佳化灰關聯度為參考之依據。最後利用測度一致化的方式，將 21 項因素皆化成相同衡量方式之最佳化灰關聯度值，作為評判與台灣股價指數間的關聯程度與衡量依據，其結果如表 6。

其次，在公式（29）與（31）中之合理調整範圍 G 值尚需決定。本研究以過去一整年的歷史資料共 243 筆資料，針對波動型態做初步分析，分析項目有最大、最小漲跌幅、平均漲跌幅等，其中圖 1、圖 2 中顯示過去一年資料的波動型態以及漲跌幅度情形，再以最末之 79 筆數據進行模型驗證。同時，經過 Excel 計算統計區間次數如表 7，並將數據整理得知相關分析項目如表 8。在分析討論所有數值後，本研究採用平均單日最大漲跌幅度為 G 值，其原因如下：（1）

表 4. 各因素之模糊 AHP 權重值

因素名稱	模糊 AHP 權重	因素名稱	模糊 AHP 權重
國內突發事件	0.0996	各類股營運狀況	0.0495
國際情勢	0.0963	股市成交量	0.0476
三大法人	0.0882	經濟成長率	0.0443
央行調降利率	0.0844	美股道瓊指數	0.0440
美股那斯達克	0.0825	國安基金	0.0380
台灣內政	0.0795	經濟景氣燈	0.0323
台幣美元的波動	0.0614	日本日經指數	0.0262
台股技術面	0.0595	CPI 指數	0.0203
失業率	0.0566	外匯存底	0.0132
FED 調降利率	0.0564	恆生指數	0.0093
台灣外交	0.0530		

表 5. 各種灰關聯度之結果與排序

一般性灰關聯度	比率值生成之灰關聯度	最佳化灰關聯度
恆生指數	日本日經指數	美股那斯達克
0.781530	0.980102802	0.947504216
美股那斯達克	美股那斯達克	美股道瓊指數
0.764653	0.980096932	0.920673686
美股道瓊指數	美股道瓊指數	股市成交量
0.731401	0.980086308	0.867829529
日本日經指數	恆生指數	日本日經指數
0.7214696	0.97964302	0.859649822
股市成交量	股市成交量	恆生指數
0.6895581	0.781141975	0.843374016

表 6. 各因素對台股之影響權重

因素名稱	灰關聯度值	因素名稱	灰關聯度值
國內突發事件	0.9518	各類股營運狀況	0.8815
國際情勢	0.9469	股市成交量	0.8786
三大法人	0.9358	經濟成長率	0.8778
央行調降利率	0.9317	美股道瓊指數	0.8724
美股那斯達克	0.9307	國安基金	0.8689
台灣內政	0.9237	經濟景氣燈	0.8621
台幣美元的波動	0.8999	日本日經指數	0.854
台股技術面	0.8976	CPI 指數	0.8471
失業率	0.8967	外匯存底	0.8381
FED 調降利率	0.8927	恆生指數	0.8332
台灣外交	0.8883		

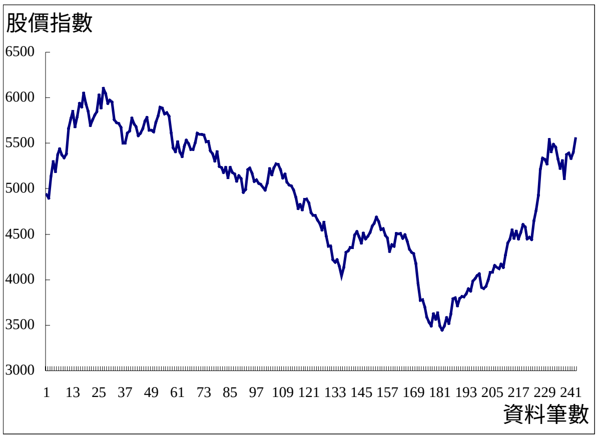

圖 1. 歷史股價指數之波動趨勢

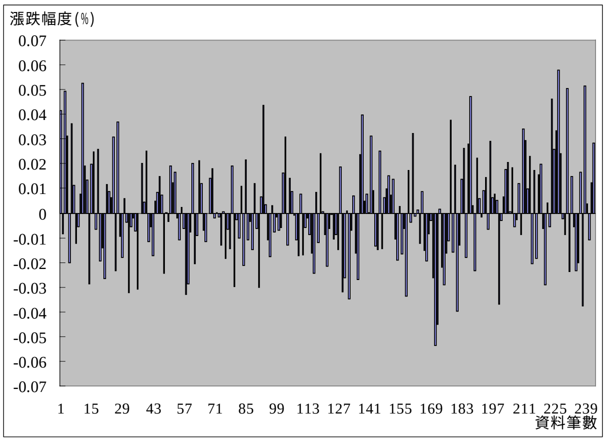

圖 2. 歷史股價之漲跌幅度

表 7. 歷史資料之漲跌幅次數統計

組距	次數	組距	次數
-0.07 ~ -0.06	0	0.00 ~ 0.01	43
-0.06 ~ -0.05	1	0.01 ~ 0.02	33
-0.05 ~ -0.04	1	0.02 ~ 0.03	20
-0.04 ~ -0.03	9	0.03 ~ 0.04	11
-0.03 ~ -0.02	23	0.04 ~ 0.05	5
-0.02 ~ -0.01	41	0.05 ~ 0.06	4
-0.01 ~ 0.00	52	0.06 ~ 0.07	0

表 8. 期望漲跌與單日最大漲跌

期望值		單日最大值	
漲幅期望值	1.76%	日最大漲幅	5.77%
跌幅期望值	-1.47%	日最大跌幅	-5.37%
平均漲跌幅期望值	1.61%	平均日最大漲跌幅	5.57%

根據台灣股市交易規章，最大漲跌幅度為 7%，而過去一年歷史數據指出平均最大漲跌幅為 5.57%，顯示研究中所探討的因素並不會影響股市漲跌到達 7%；(2) 各種期望漲跌幅僅只有正負百分之一點多，其區間範圍過窄，所以宜向上調整。綜合以上兩點之考量，因此以平均單日最大漲跌幅為調整範圍之參數。

最後可由蒐集之數據資料與每日所發生之影響因素產生隔日之預測值，其中，除 GM(1,1) 為主要預測模型外，本研究同時利用傳統的非線性迴歸與時間序列法作為比較的對象 [2, 3, 8, 14]，特別的是，為符合灰色預測之特性，所有預測模型皆採用滾動建模方式，也就是說，僅以四筆數據建模預測，再反覆的以新的數據取代舊的數據並建立新的模型，但在時間序列預測模型之建立必需靠 20 筆資料以上

之客觀數據才可建立其預測模型, 所以可說滾動建模用以預測上實在不容易, 預測之結果如圖 3 至圖 8。由圖 3 中可知, 透過因素校正後, 對於原始數據之明顯轉折處能有效的平滑, 如此將可增加灰色預測模型之穩定性。

最後將各預測之結果轉換成數值之資料以比較各方法之差異, 評估之準則為均方差（mean square error, MSE）與平均絕對差（mean absolute error, MAE）[16], 比較之結果如表 9 所示。由表中可知, 經由因素校正後之 GM(1,1) 在預測表現上有顯著的提昇, 同時, 透過自迴歸模型之誤差修正方法, 將可使預測誤差更進一步的縮小。

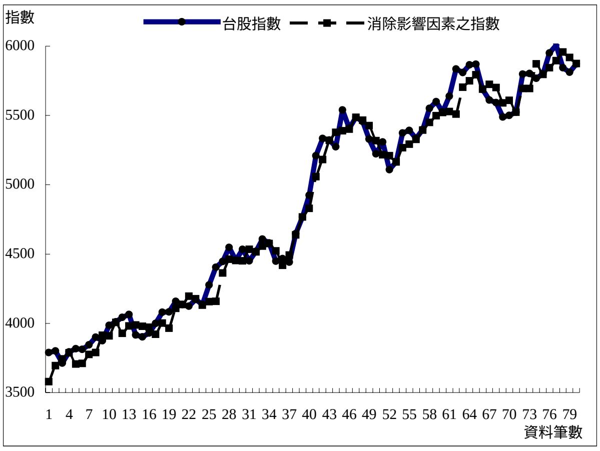

圖 3. 因素校正後無 GM(1,1) 預測之指數值與原始指數值比較

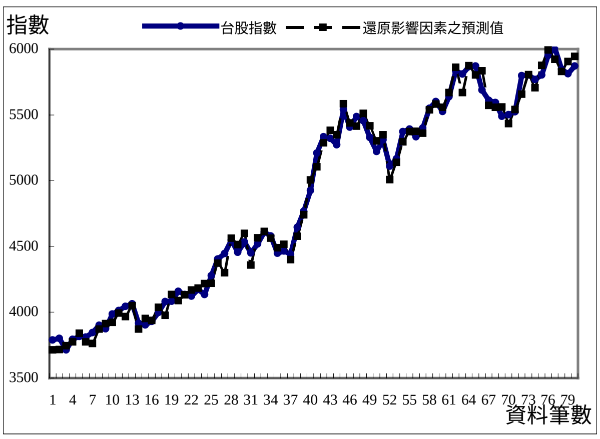

圖 4. 因素校正後無 GM(1,1) 預測之指數值與原始指數值比較

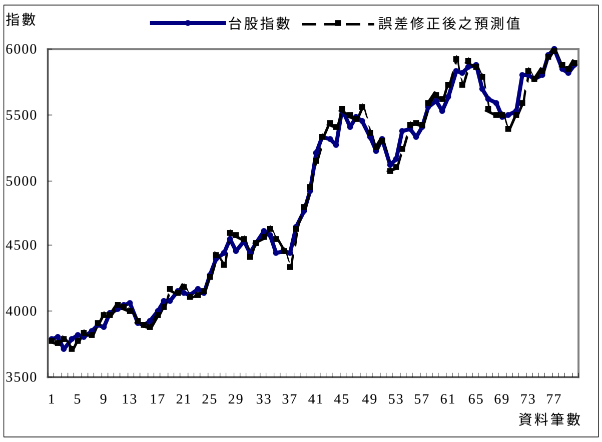

圖 5. 因素校正後經 GM(1,1)預測＋誤差修正後之預測結果

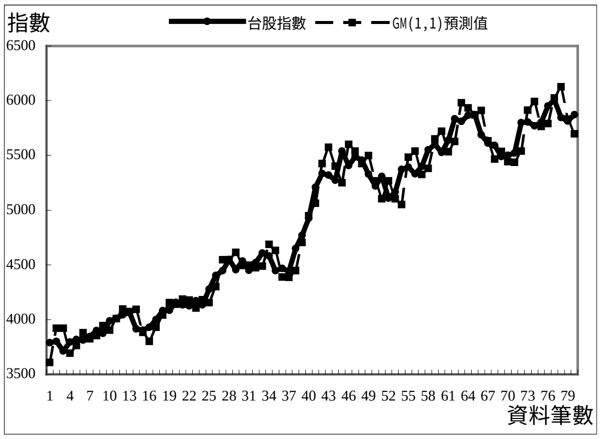

圖 6. 原始 GM(1,1) 之預測結果

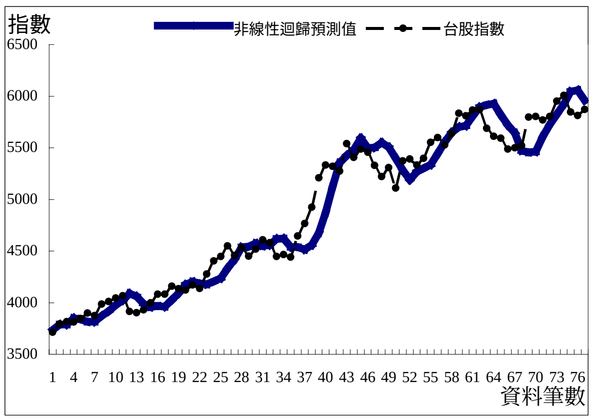

圖 7. 非線性迴歸模型之預測結果

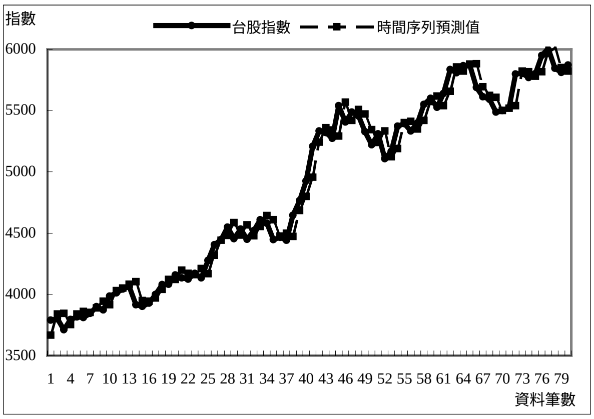

圖 8. 時間序列模型之預測結果

表 9. 各預測模型之誤差比較

預測模型	MAE	MSE
GM(1,1)+因素校正+誤差修正	48.74	1059.13
GM(1,1)+因素校正	53.20	3955.82
GM(1,1)	106.58	17223.48
時間序列	78.39	9970.14
非線性迴歸模型	134.0	29317.3

七、結論

在許多預測模型及方法極力追求與實際值之誤差最小化之狀況下，模型精簡及參考之歷史資料越少，越能突顯此模型的優勢，灰色理論即是在此種觀念下所創造出之理論。GM(1,1) 模型有著只需少筆資料（本研究僅以四筆資料來建立預測模型），即能進行預測的優點，但其對於一序列中之轉折點之預測效果卻不盡理想。所以本研究即針對此一問題進行修正，期望能提昇 GM(1,1) 之預測績效。

對於序列之波動特性，本研究從造成數據波動之原因切入，透過對影響數據波動之相關因素的觀察而進行分析，分析之方法為結合灰關聯分析法與模糊階層分析法，其中前者用以處理量化型因素；而後者用以處理非量化型因素。經由分析的過程，並對原始數據進行調整，使其轉折點處能獲得平滑之效果。而實例驗證中，圖 4 旨在說明利用因素修改方式修改後的數據使用 GM(1,1) 預測模型，已有非常接近之預測效果，雖有稍許過頭的現象，但並無像圖 6 中原始 GM(1,1) 模型造成許多轉折點上反向預測，而產生大量的誤差；圖 5 旨在說明利用時間序列預測方法所作之誤差修正，確實亦可增加預測數據與實際值得接近性雖然一開始建模使用初期仍會有些微不穩定現象，有些是因為誤差預測時所產生的不預期修正現象，但經過一段期間之滾動建模後，漸漸出現穩定現象，亦說明了此方法能有效的改善並減低傳統灰色預測模型之缺點。再者，透過自迴歸模型之誤差修正方法，能使誤差進一步的降低。而圖 6 在說明利用純 GM(1,1) 預測模型所做預測時所產生在數據轉折處，所產生的高估與低估之示意圖；圖 7、圖 8 亦說明了非線性回歸模型與時間序列模型在預測時的不利處。

綜合上述，灰色預測方法能克服傳統方法所需大量歷史資料之缺點，從圖 4 中已可顯示出本研究之研究方法已經改善許多，甚至許多與測點已經非常精準，而圖 5 之目的只為了得到更佳的預測值。因此，若能對灰色模型本身之缺點加以改良，即能在預測準確性上有理想之成果。

參考文獻

1. 刑治宇、蔡勇斌（民 89），GM(1,1) 預估值之簡化模式探討，第五屆灰色系統理論與應用研討會，頁 271-275。
2. 林茂文（民 81），時間數列分析與預測，華泰書局，台北。
3. 吳柏林（民 84），時間數列分析導論，華泰書局，台北。
4. 郭怡華、巫沛倉（民 90），修正模糊層級分析法之應用－以基金經紀人之選股策略為例，頁 1-18，中國工業工程學會九十年會暨學術研討會。
5. 徐守德、黃玉娟、余明芳（民 90），台股指數期貨日內交易型態之研究－摩根台股期貨與台灣指數期貨之比較，管理評論，20(2)，頁 31-53。
6. 袁嘉祖（民 80），灰色系統理論及其應用，科學出版社，台北。
7. 區亦勤、張先迪編著（民 80），模糊數學原理及應用，儒林圖書有限公司，台北。
8. 陳耀茂（民 83），迴歸分析導論，全華科技圖書股份有限公司，台北。
9. 陳耀茂（民 87），階層構造分析法入門，東海大學圖書館，台中。
10. 張偉哲、溫坤禮、張廷政（民 89），灰關聯模型方法與應用，高立圖書有限公司，台北。
11. 鄧聚龍（民 89），灰色系統理論與應用，高立圖書有限公司，台北。
12. 鄧聚龍、郭洪、溫坤禮、張廷政、張偉哲（民 88），灰

預測模型與應用，高立圖書有限公司，台北。

13. 劉應興（民 87），非線性迴歸與相關分析應用線性迴歸模型－補篇，華泰書局，台北。
14. Chang, B. R. (2001) Alternative view of grey relational analysis. *The Journal of Grey System*, 13(1), 31-40.
15. Chang, B. R. (2001) An optimal grey relational measurement. *Proceeding of International Joint Conference on Neural Networks*, 1609-1614.
16. Diebold, F. X. (2001) *Elements of Forecasting*, 2nd Ed., 291-294, South-Western, Cincinnati, Ohio.
17. Deng, J. L. (1989) Introduction to grey system theory. *The Journal of Grey System*, 1(1), 1-24.
18. Wu, J. H. and C. B. Chen (1999) An alternative form for grey relational grades. *The Journal of Grey System*, 11(1), 7-12.
19. Zhu, K. J., Y. Jing and D. Y. Chang (1999) A discussion on extent analysis method and applications of fuzzy AHP. *European Journal of Operational Research*, 116, 450-456.

收件：**91.07.10**　修正：**91.09.25**　接受：**91.10.17**