



混合複數類神經模糊與自動回歸差分平均移動方法之智慧型時間序列預測模型

李俊賢* 江泰緯

國立中央大學資訊管理學系

摘要

本研究提出一個複數模糊類神經系統 (CNFS)，其結合複數模糊集合、類神經模糊系統以及差分自回歸移動平均模型 (ARIMA)，形成 CNFS-ARIMA 模型並應用於時間序列預測之研究。為了更新 CNFS-ARIMA 模型之參數，本研究提出一複合式進化式學習演算法，其結合粒子群最佳化演算法與遞迴最小平方估計之學習方法；其中，粒子群最佳化演算法用來調整系統的前鑑部參數，而遞迴最小平方估計用以更新系統之後鑑部參數。為了測試本研究所提出之方法的效能，使用兩個標竿時間序列資料集作為實驗範例。在實驗中，將比較並分析本研究所提出之 CNFS (複數型) 以及其傳統的 NFS (實數型) 之效能差異，並與文獻所提出之方法進行比較。由實驗結果可證實本研究所提出之系統方法可以獲得良好的效能。

關鍵詞：複數模糊集合、複數模糊類神經系統、粒子群最佳化演算法、遞迴最小平方估計法、時間序列預測

Intelligent Time Series Forecasting Model Combining Complex Neuro-Fuzzy Computing Model and ARIMA Method

Chunshien Li Tai-Wei Chiang

Department of Information Management, National Central University, Taiwan

Abstract

A complex neuro-fuzzy system, using complex fuzzy sets (CFSs), neuro-fuzzy theory, and autoregressive integrated moving average (ARIMA) model, is proposed to the problem of time series forecasting. The proposed computing system is denoted as CNFS-ARIMA. To update the free parameters of the proposed CNFS-ARIMA, a novel hybrid learning

* 通訊作者 電子信箱：jamesli@mgt.ncu.edu.tw

本研究由國科會計畫 NSC99-2221-E-008-088 補助。



method is devised, combining both the particle swarm optimization (PSO) algorithm and the recursive least squares estimator (RLSE) algorithm. The PSO is used to adjust the premise parameters of the proposed predictor, and the RLSE is used to update the consequent parameters. To test the proposed approach, two benchmark time series datasets are used. The experimental results by the proposed approach are compared with those by its neuro-fuzzy counterpart and by other approaches in literatures. The experimental results have illustrated the merits of CFSs in the proposed approach with excellent performance for the two examples of time series forecasting. Through performance comparison, the experimental results indicate that the proposed approach outperforms the compared approaches.

Keywords: Complex fuzzy set (CFS), complex neuro-fuzzy system (CNFS), particle swarm optimization (PSO), recursive least square estimator (RLSE), time series forecasting.

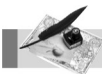
1. 緒論

時間序列是指在連續時空下由某一種動態過程所產生的一連串資料點，通常經過等時距觀察某一種變數或某些變數而得。在物理系統之動態過程，物理變數可能是溫度、氣壓、濕度、速度、壓力與濃度等。在財務和經濟系統的動態過程，財經變數可能是匯率、股價、商品銷售量、國內生產總值、失業率、進出口量、物價指數、銀行利率與經濟成長率等。時間序列可以分為連續型的（continuous）與離散型的（discrete）時間序列。以時間的維度觀察，若時間序列在時間的維度的觀察是連續型的，則所得到的連續型的時間序列，反之則為離散型的時間序列。以機率的角度，時間序列可分為確定型的（deterministic）或隨機型的（stochastic）。若時間序列的產生依照某種機率分佈，例如高斯分佈，則所得到的隨機型的時間序列，否則為確定型的時間序列。以統計的角度，時間序列又可分為靜定型的（stationary）或非靜定型的（non-stationary）。若時間序列經過一段時間平移後仍保持相同的統計特性（如平均值、標準差），則屬於靜定型的時間序列，反之為非靜定型的時間序列。時間序列有一種很重要的特性：相鄰的觀測資訊之間存在某一種相互依存的特性（dependence）。時間序列之資訊相互依存性可視為類似物理學中物體的慣性（inertia），此種資訊依存性具有高度實用性及研究價值，因為研究時間序列的資訊依存性可以進一步了解其背後的資訊關連機制（information relation mechanism）。



時間序列分析即是在研究資訊依存性與相關性的科學方法；時間序列分析通常涉及分析模型的發展及使用（Box and Jenkins, 1976; Tsay, 2005; Wooldridge, 2006）。時間序列的應用可分為預測、轉移函數估測、多個相關時間序列的相互關係、控制應用以及時間序列的干涉效應之評估。本研究將著重於時間序列的預測研究。一般常用時間序列的分析模型可分為自動回歸模型（autoregressive, AR）、平均移動模型（moving average, MA）、自動回歸平均移動模型（autoregressive moving average, ARMA）與自動回歸差分平均移動模型（autoregressive integrated moving average, ARIMA）。其中，AR 及 MA 模型是較為簡單而實用的模型；AR 模型是由時間序列的過去觀測資訊作為現在或未來時間序列的解釋變數而形成的線性模型；MA 模型則是用過去模型的餘差項（residuals）作為現在或未來時間序列的解釋變數而形成的線性模型。ARMA 模型是綜合 AR 及 MA 兩種模型的線性模型，其模型的結構較具彈性；ARIMA 模型基本上是 ARMA 結構，差異在於它將 AR 模型的解釋變數作差分（difference）的前處理。藉由 ARIMA 模型可以清楚地了解資料觀測點的特性並進而預測未來的可能趨勢，透過差分處理可使得非靜定時間序列具有靜定的果效，用以消除非平穩之特性，因此 ARIMA 模型對非靜定時間序列有不錯的性能（Box and Jenkins, 1976）。然而，ARIMA 受制於本身的線性函數模型結構，對於預測真實世界中非線性、複雜且巨幅的波動時間序列時，並無法獲得令人滿意的預測效能。

現實生活中時間序列通常為一個非靜定之連續資料，常會因為環境的改變而帶來急遽的變化，使得傳統的統計回歸與數學模型方法較難描述如此複雜且混沌的時間序列。近二十年來，有許多的文獻開始運用機器學習與柔性計算方法來建立智慧型時間序列預測模型（Boyacioglu and Avci, 2010; Chen et al., 2005; Cho and Wang, 1996; Tung and Quek, 2011）。其中，類神經網路（neural network）以及模糊理論（fuzzy theory）已被廣為應用於時間序列研究領域。類神經網路與模糊理論已被證實為廣泛近似器（universal approximation），能近似任何的函數並達到高準確率（Castro, 1995; Hornik et al., 1989; Wang and Mendel, 1992）。若進一步比較類神經網路與模糊理論兩者之間的差異，可以發現類神經網路具有良好的學習與資料映射能力，但其訓練後所得之模型為一個「黑盒子（black-box）」，難以用人類的知識來解釋該模型之意義。針對類神經網路的缺點，模糊推論系統可提供一個互補性的替代方案。藉由模糊邏輯推論，我們得以將人類的知識與經驗轉化為讓人易理解的模糊法則（Huwendiek and Brockmann, 1999）。Jang（1993）結合了模糊推論與類神經網路，提出了適應性類神經模糊系統（adaptive-network-based fuzzy inference system, ANFIS）。由過去文獻的研究成果中亦證實，藉由整合兩種方法的綜效能有效的發掘出時間序列資料中的資訊依存性（Castillo and Melin, 2002; Hengjie et al., 2010; Kasabov and Qun, 2002）。Ramot et al.（2002）提出了複數模糊集合（complex fuzzy



set, CFS) 的概念。複數模糊集合可視為傳統型一模糊集合 (type-1 fuzzy set) 的延伸, 將傳統模糊集合的歸屬度由 $[0, 1]$ 區間延伸至複數的平面, 其歸屬度包含於複數平面的二維單位圓盤中。亦即 CFS 中的元素已經不是傳統的實數值型態 (real-valued state), 其擴展為複數單位圓平面中的複數值型態 (complex-value state) (Dick, 2005; Ramot et al., 2002; Ramot et al., 2003)。換句話說, CFS 可以將模糊集合的歸屬度從一維區間擴展至二維複數空間, 進而提升模糊集合的適應與映射能力。Zhanget al. (2009) 針對 CFS 的特性提出相關的研究, 但並未針對複數模糊推理程序與設計多加著墨; Chen et al. (2011) 提出適應性類神經模糊推論於時間序列預測問題, 不過該研究的 CFS 設計並無法完全發揮 CFS 的特性。儘管 CFS 的概念與特性已經被提出探討 (Chen et al., 2011; Dick, 2005; Ramot et al., 2002; Ramot et al., 2003; Zhang et al., 2009), 但仍尚未有研究提出一套明確的 CFS 設計準則與方法, 因而運用此概念之相關研究仍然相當稀少。在我們過去的研究成果中 (Li and Chiang, 2010; 2011a; 2011b), 已提出複數型模糊類神經系統之模型, 並用於處理真實世界的相關應用, 其研究成果證明其擁有優良的函數映射能力。

本研究提出一套全新的智慧型複數類神經模糊預測系統, 其整合複數模糊集合與類神經模糊系統之綜效, 並應用於時間序列預測之議題。藉由複數型模糊集合來提升類神經模糊系統之適應能力。基於 CNFS, 本研究將 ARIMA 模型結合到 CNFS 的模糊法則中, 進而建置一個創新性的 CNFS-ARIMA 預測系統。藉由 CNFS 卓越的適應彈性與非線性映射能力, 以及 ARIMA 模型擅長於處理線性與非平穩性時間序列之特性來提升系統預測的準確度。為了優化 CNFS-ARIMA 系統之參數, 本研究提出一複合式學習法, 其結合了粒子群最佳化演算法 (particle swarm optimization, PSO) 與遞迴式最小平方估計法 (recursive least square estimator, RLSE)。藉由複合式學習與分而擊之 (divide and conquer) 的概念, 將 PSO 用來調整模糊法則的前鑑部參數 (premise parameters); RLSE 用來調整模糊法則的後鑑部參數 (consequent parameters)。由於粒子群最佳化演算法對於尋找全域最佳解問題是一個很受歡迎且有效率的求解法, 而遞迴最小平方估計法則是能以極快的速度找到線性模型的最佳解。透過複合式學習方法來尋找 CNFS-ARIMA 預測系統之最佳或趨近於最佳解的參數集合並達到快速收斂的目的。為了測試提出之預測模型的效能, 本研究使用兩個文獻中常見的時間序列標竿資料集作為系統的測試標的。實驗結果顯示, 本研究方法擁有卓越的預測效能, 其結果皆優於文獻所提出之方法。

本研究的主要貢獻分述如下:

- (1) 提出 CNFS-ARIMA 預測模型, 並提出一個全新的複數模糊集合之設計概念, 藉以提升模型之適應能力與彈性, 進而提升模型的整體效能。
- (2) 提出一個全新的複合式學習演算法, 用以更新 CNFS-ARIMA 模型之參數,



使該模型得以更快的找到系統參數之最佳組合。

- (3) 本研究所提出之 CNFS-ARIMA 系統結合複合式學習演算法，可獲得良好的預測能力，並優於文獻所提出之方法。

在本研究於第二節，將論述 CNFS-ARIMA 之系統方法。第三節，說明複合式學習演算法。第四節，將本研究方法應用於時間序列預測問題，並進行實驗與結果呈現。第五節，將針對實驗結果作探討與分析。最後為本研究之結論。

2. CNFS-ARIMA預測模型

2.1 複數模糊集合

模糊集合 (fuzzy set) 是從一般傳統集合 (crisp set) 延伸而來，模糊集合中的元素已不再只是絕對的「對或錯」、「屬於或不屬於」和「0 或 1」的二值狀態 (boolean-valued state)。模糊集合中的元素歸屬度可延伸為連續的實數狀態 (real-valued state)，藉以表達不明確的資料、資訊與概念。複數模糊集合 (complex fuzzy set, CFS) 是從模糊集合延伸而來，其元素之歸屬程度 (membership degrees) 不再只是定義於實數區間 $[0, 1]$ 內，其歸屬程度可進一步延伸至單位複數圓盤之複數區間 (Dick, 2005; Ramot et al., 2002; Ramot et al., 2003)。經過如此的延伸，CFS 即能進一步從經驗導向 (experience oriented) 進入經驗抽象化導向 (abstraction oriented) 的表達。假設有一個複數模糊集合 (S)，其歸屬函數 $\mu_s(h)$ 表示如下：

$$\begin{aligned}\mu_s(h) &= r_s(h) \exp(j\omega_s(h)) \\ &= \text{Re}(\mu_s(h)) + j \text{Im}(\mu_s(h)) \\ &= r_s(h) \cos(\omega_s(h)) + jr_s(h) \sin(\omega_s(h))\end{aligned}\quad (1)$$

其中， $j = \sqrt{-1}$ ； h 為數值變數 (base variable for the complex fuzzy set)； $r_s(h)$ 為複數歸屬程度之振幅函數 (amplitude function of complex membership)； $\omega_s(h)$ 為複數歸屬程度之相位函數 (phase function of complex membership)。Re(.) 和 Im(.) 分別為實部 (real-part) 與虛部 (imaginary-part)。當 $\omega_s(h)=0$ 時，可將傳統模糊集合視為複數模糊集的一種特例。複數模糊集合 (S) 可用下式表達：

$$S = \{(h, \mu_s(h)) | h \in U\} \quad (2)$$

其中， U 代表複數模糊集合之論域 (universe of discourse)。

本研究提出一複數型高斯模糊函數 (complex Gaussian membership function, cGMF)。其歸屬函數如 (3) 所示。

$$cGMF(h, m, \sigma, \lambda) = r_s(h, m, \sigma) \exp(j\omega_s(h, m, \sigma, \lambda)) \quad (3a)$$

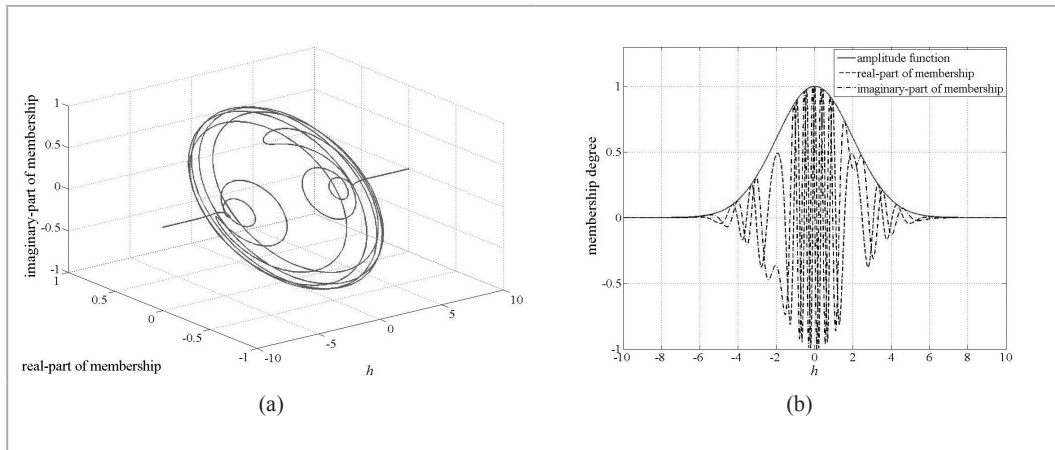


其中，

$$r_s(h, m, \sigma) = \text{Gaussian}(h, m, \sigma) = \exp\left[-0.5\left(\frac{h-m}{\sigma}\right)^2\right] \quad (3b)$$

$$\omega_s(h, m, \sigma, \lambda) = -\exp\left[-0.5\left(\frac{h-m}{\sigma}\right)^2\right] \times \left[\left(\frac{h-m}{\sigma^2}\right)\right] \times \lambda \quad (3c)$$

其中， (h, m, σ, λ) 分別代表 $cGMF$ 的數值變數、平均值、延展度及相位頻率參數。 $cGMF$ 模糊集合之圖形描繪如圖 1 所示。



▲ 圖 1 複數高斯模糊集合。(a) 複數歸屬函數三維視圖（維度分別為：數值變數 (h) 、歸屬度之實部（real-part membership）與虛部（imaginary-part membership）。(b) 實數高斯函數與複數高斯歸屬函數之實部與虛部圖形。

2.2 自動回歸差分平均移動模型

時間序列預測之基本概念如下，假設目前時刻為 t ，已知時間 t 與時間 t 之前的歷史觀測資料值，且歷史觀測資料值之間具相關性，利用已知之觀測值資料並透過預測模型對時間點 $t+i$ 進行預測，稱之第 i 步預測。在時間序列預測中，自回歸移動平均模型（autoregressivemoving average, ARMA）為一個最常被使用來預測未來趨勢的模型。根據 ARMA 模型理論，時間序列可以依據過去觀測值與隨機誤差所組合的數學線性函數表示。ARMA 模型中包含兩個部分：自回歸項（AR）與移動平均項（MA）。假設有一組時間序列資料，定義為 $\{y_\tau, \tau=1, 2, \dots\}$ ，其中 τ 為時間指標。一般型式的 $AR(p)$ 模型如下所示：



$$y_t = \varphi_0 + \varphi_1 y_{t-1} + \cdots + \varphi_p y_{t-p} = \varphi_0 + \sum_{i=1}^p \varphi_i y_{t-i} + e_t \quad (4)$$

其中， p 代表 AR 模型落後期數； $\{\varphi_i, i=0, 1, 2, \cdots, p\}$ 是 AR 之係數； e_t 為 t 時間的隨機誤差項並且符合獨立性同分佈 (independently and identically distributed, iid) 特性，其平均數為 0 且標準差為 σ^2 。而 MA(q) 模型為：

$$y_t = u + e_t - \theta_1 e_{t-1} - \cdots - \theta_q e_{t-q} = u + e_t + \sum_{j=1}^q \theta_j e_{t-j} \quad (5)$$

其中， u 為 y_t 的期望值； q 代表 MA 模型落後期數； $\{\theta_j, j=1, 2, \cdots, q\}$ 為 MA 模型的係數； $\{e_{t-j}, j=1, 2, \cdots, q\}$ 為隨機誤差，且同樣符合相互獨立且具有相同分配的特性。ARMA 為 AR 與 MA 之結合，可表示為以下之形式：

$$\begin{aligned} y_t &= \varphi_0 + \varphi_1 y_{t-1} + \cdots + \varphi_p y_{t-p} - \theta_1 e_{t-1} - \cdots - \theta_q e_{t-q} + e_t \\ &= \varphi_0 + \sum_{i=1}^p \varphi_i y_{t-i} - \sum_{j=1}^q \theta_j e_{t-j} + e_t \end{aligned} \quad (6)$$

ARIMA 是由 ARMA 延伸出的預測模型，其架構可被視為預測值、歷史觀測值與預測誤差的線性關係。以系統塑模的角度而言，ARIMA 的架構具有代表各種模型的適應性與彈性。對於非靜定型時間序列，ARIMA 利用差分的方式將非靜定的時間序列轉換成靜定序列，其相關定義如下所示：

$$B^k y_t \equiv y_{t-k}, \forall k \in \{1, 2, \cdots\} \quad (7a)$$

$$\Phi_p(B) \equiv 1 - \varphi_0 - \varphi_1 B - \varphi_2 B^2 - \cdots - \varphi_p B^p = 1 - \varphi_0 - \sum_{i=1}^p \varphi_i B^i \quad (7b)$$

$$\Theta_q(B) \equiv 1 - \theta_1 B - \theta_2 B^2 - \cdots - \theta_q B^q = 1 - \sum_{j=1}^q \theta_j B^j \quad (7c)$$

$$\nabla^d \equiv (1 - B)^d \quad (7d)$$

其中， B 為延遲運算子 (back shift operator)； ∇ 為差分運算子； d 為一個不為負的整數，通常被稱為差分階數。因此根據公式 (7)，ARIMA 模型可表示如下：

$$\Phi_p(B) \nabla^d y_t = \Theta_q(B) e_t \quad (8)$$

根據公式 (7) 的 ARIMA 模型理論，可對於時間序列建置一個 ARIMA 預測模型。由公式 (8)，ARIMA 可延伸為以下形式：



$$\nabla^d y_\tau = \varphi_0 + \sum_{i=1}^p \varphi_i B^i \nabla^d y_\tau - \sum_{j=1}^q \theta_j B^j e_\tau + e_\tau \quad (9)$$

根據 (7d) 和 (9) 可進一步定義為：

$$\psi_\tau \equiv \nabla^d y_\tau = (1 - B)^d y_\tau \quad (10)$$

根據 (9) 和 (10)，ARIMA 的一般形式可定義如下所示：

$$\psi_\tau = \varphi_0 + \sum_{i=1}^p \varphi_i \psi_{\tau-1} - \sum_{j=1}^q \theta_j e_{\tau-j} + e_\tau \quad (11)$$

根據 (11) 可以得到 ARIMA 預測模型之定義。假設隨機誤差 $\{e_{\tau-j}, j=0, 1, \dots, q\}$ 可視為預測值與觀測值之間的預測誤差，則基本形式的 ARIMA 預測模型可以表示如下：

$$\begin{aligned} \psi_\tau &= \varphi_0 + \sum_{i=1}^p \varphi_i \psi_{\tau-1} - \sum_{j=1}^q \theta_j e_{\tau-j} \\ &= \varphi_0 + \sum_{i=1}^p \varphi_i B^i \psi_\tau - \sum_{j=1}^q \theta_j B^j e_\tau \\ &\equiv \text{ARIMA}(p, d, q) \end{aligned} \quad (12)$$

其中， (p, d, q) 分別為 AR 階數、差分階數以及 MA 階數。

2.3 CNFS-ARIMA 預測模型

本質上，ARIMA 為一個線性回歸的資料導向方法 (Box and Jenkins, 1976)，擅長處理線性的應用。但面對複雜與非線性的應用，將難以獲得令人滿意的效能。因此，本研究整合複數模糊類神經系統具非線性映射能力與 ARIMA 線性迴歸的性質，提出 CNFS-ARIMA 預測模型，藉以提升預測系統的整體綜效。假設設計一個由 K 條 Takagi-Sugeno 模糊規則 (T-S fuzzy rules) 所共同組成的 CNFS-ARIMA 預測模型，其具有 M 個輸入與單一輸出，第 i 條模糊規則如下所示：

$$\begin{aligned} \text{Rule } i: & \text{IF } v_1 \text{ is } {}^i A_1(h_1) \text{ and } v_2 \text{ is } {}^i A_2(h_2) \text{ and } \dots \text{ and } v_M \text{ is } {}^i A_M(h_M) \\ & \text{THEN } {}^i \psi = {}^i \varphi_0 + \sum_{k=1}^p {}^i \varphi_k B^k \psi_\tau - \sum_{l=1}^q {}^i \theta_l B^l e_\tau \end{aligned} \quad (13)$$

其中， $i=1, 2, \dots, K$ ； v_j 為第 j 個輸入語意變數； ${}^i A_j(h_j)$ 為第 i 條規則中第 j 個條件的複數模糊集合， $j=1, 2, \dots, M$ ； h_j 為第 j 個輸入數值變數值； ${}^i \psi$ 為第 i 條模糊法則的輸出； $\{{}^i \varphi_k, k=0, 1, \dots, p\}$ 與 $\{{}^i \theta_l, l=1, 2, \dots, q\}$ 為模糊法則之後鑑部參數。

本研究提出之 CNFS-ARIMA 系統設計，使用六層式的前向式 (multilayer feed-



forward) 模糊類神經網路架構。各階層之模糊邏輯推論說明如下：

第 1 層 (Layer 1)：此階層稱為輸入層 (input layer)，此階層用以接收 t 時間的外界輸入值，並將輸入值傳輸到下一個階層中。 t 時間的輸入向量表示如下：

$$\mathbf{H}(t) = \begin{bmatrix} h_1(t) \\ h_2(t) \\ \vdots \\ h_M(t) \end{bmatrix} \quad (14)$$

第 2 層 (Layer 2)：此階層稱為模糊集合層 (fuzzy-set layer)。此階層的每一個節點代表模糊集合，透過歸屬函數的計算，可描述輸入變數對應複數模糊集合之歸屬程度。本研究之複數模糊集合之設計為高斯型態的複數歸屬函數，如公式 (3) 所示。在本層中所有複數模糊集合的平均值、延展度與相位變數，統稱為 CNFS-ARIMA 模型的前鑑部參數。

第 3 層 (Layer 3)：此階層稱為模糊法則層 (fuzzy-rule layer)，用以計算每一條法則的複數啟動強度 (complex firing strength)。第 i 條模糊規則的啟動強度 ${}^i\beta(t)$ 經交集運算 (T-norm)，表示如下：

$${}^i\beta(t) = \prod_{k=1}^M {}^i\mu_k(h_k(t)) = \prod_{k=1}^M {}^i r_k(h_k(t)) \exp(j^i \omega_k(h_k(t))) \quad (15)$$

其中， ${}^i\mu_k(h_k(t))$ 代表第 i 條模糊規則的第 k 個模糊集合對應之歸屬程度。

第 4 層 (Layer 4)：此階層稱為正規化層 (normalization layer)，用以將每條模糊規則之啟動強度進行正規化的計算，以便了解每條模糊規則對於總系統輸出之貢獻程度。第 i 條規則的啟動強度正規化方式定義如下：

$${}^i\lambda(t) = \frac{{}^i\beta(t)}{\sum_{i=1}^K {}^i\beta(t)} = \frac{\prod_{k=1}^M {}^i r_k(h_k(t)) \exp(j^i \omega_k(h_k(t)))}{\sum_{i=1}^K \prod_{k=1}^M {}^i r_k(h_k(t)) \exp(j^i \omega_k(h_k(t)))} \quad (16)$$

第 5 層 (Layer 5)：此階層為後鑑部層 (consequent layer)，根據第 4 層產生之正規化啟動強度來計算每條模糊規則的後鑑部輸出。透過 (13) 和 (16)，可以得到每條模糊規則正規化後的後鑑部輸出，如下所示：

$${}^i\xi(t) = {}^i\lambda(t) {}^i\psi(t) = {}^i\lambda(t) ({}^i\varphi_0 + \sum_{k=1}^p {}^i\varphi_k B^k \psi_\tau - \sum_{l=1}^q {}^i\theta_l B^l e_\tau) \quad (17)$$

第 6 層 (Layer 6)：此階層稱為輸出層 (output layer)。將第 5 層所得到的每條規則



輸出做結合，產生預測系統的整體輸出，輸出結果如下所示：

$$\hat{\psi}(t) = \sum_{i=1}^K \xi(t)^i = \sum_{i=1}^K \lambda(t)^i \psi(t) \quad (18)$$

其中，系統之輸出為一個複數值（complex-valued output），可表示為以下形式：

$$\begin{aligned} \hat{\psi}(t) &= \hat{\psi}_{\text{Re}}(t) + j\hat{\psi}_{\text{Im}}(t) \\ &= |\hat{\psi}(t)| \cdot \exp(j\omega_{\hat{\psi}(t)}) \\ &= |\hat{\psi}(t)| \cdot \cos(\omega_{\hat{\psi}(t)}) + j|\hat{\psi}(t)| \cdot \sin(\omega_{\hat{\psi}(t)}) \end{aligned} \quad (19)$$

$\hat{\psi}_{\text{Re}}(t)$ 代表 CNFS 系統輸出值之實數部分； $\hat{\psi}_{\text{Im}}(t)$ 代表系統輸出值之虛數部分。CNFS 複數輸出值之振幅和相位值可表示如下：

$$|\hat{\psi}(t)| = \sqrt{(\hat{\psi}_{\text{Re}}(t))^2 + (\hat{\psi}_{\text{Im}}(t))^2} \quad (20)$$

$$\omega_{\hat{\psi}(t)} = \tan^{-1} \left(\frac{\hat{\psi}_{\text{Im}}(t)}{\hat{\psi}_{\text{Re}}(t)} \right) \quad (21)$$

根據 (19)，CNFS 可以視為一個複數函數，定義如下：

$$\hat{\psi}(t) = \mathbf{F}(\mathbf{H}(t), \mathbf{W}) = \mathbf{F}_{\text{Re}}(\mathbf{H}(t), \mathbf{W}) + j\mathbf{F}_{\text{Im}}(\mathbf{H}(t), \mathbf{W}) \quad (22)$$

其中， $\mathbf{F}_{\text{Re}}(\cdot)$ 為 CNFS 輸出值之實數部分； $\mathbf{F}_{\text{Im}}(\cdot)$ 為 CNFS 輸出值之虛數部分； $\mathbf{H}(t)$ 代表 CNFS 的系統輸入； \mathbf{W} 是 CNFS 中的系統參數集合，可以表示成前鑑部 (\mathbf{W}_{If}) 與後鑑部 (\mathbf{W}_{Then}) 兩個子集合。CNFS-ARIMA 之參數集合可表示為：

$$\mathbf{W} = \mathbf{W}_{\text{If}} \cup \mathbf{W}_{\text{Then}} \quad (23)$$

3. 複合式學習演算法

3.1 粒子群最佳化演算法

本研究以粒子群最佳化演算法（particle swarm optimization, PSO）作為複合式機器學習方法之一。PSO 是一種以族群為本（population-based）的最佳化搜尋學習演算法（Eberhart and Kennedy, 1995; Kennedy and Eberhart, 1995），其結合自我認知及社會認知兩種機制，在尋找最佳解的歷程中產生了互通互補的良好作用。PSO 的發展觀念來自於鳥群（或魚群）的覓食行為的啟發。由觀察鳥群覓食的行為中發現，當個別鳥隻在群體中尋找食物（最佳解）的過程中，若有某些鳥隻發現食物時就會改變飛



行路線，往食物位置移動；由於某些鳥發現食物，其他的鳥也會產生互動並隨著改變飛行方向，最後整個鳥群會逐漸降落在食物地點（指搜尋過程逐漸收斂）。在 PSO 的方法中，每一個個體稱為粒子（particle），一整群粒子稱為群族（swarm）。在最佳化搜尋策略上，一整群粒子並行式搜尋，並彼此分享且交換搜尋資訊，藉此避免掉入局部最佳解的區域。在學習進行過程中，每一個粒子的位置代表一個潛在可能解（potential solution）。粒子群演算法之個別粒子之速度與位置更新公式如下。定義如下：

$$\vartheta_i(k+1) = w\vartheta_i(k) + c_1\rho_1(\mathbf{pbest}_i(k) - \mathbf{L}_i(k)) + c_2\rho_2(\mathbf{gbest}(k) - \mathbf{L}_i(k)) \quad (24)$$

$$\mathbf{L}_i(k+1) = \mathbf{L}_i(k) + \vartheta_i(k+1) \quad (25)$$

其中， $\vartheta_i(k) = [\vartheta_{i,1}(k), \vartheta_{i,2}(k), \dots, \vartheta_{i,Q}(k)]^T$ 為第 i 個粒子在第 k 次迭代的搜尋速度； $\mathbf{L}_i(k) = [l_{i,1}(k), l_{i,2}(k), \dots, l_{i,Q}(k)]^T$ 為第 i 個粒子在第 k 次迭代的所在的位置； $\{c_1, c_2\}$ 為粒子加速常數； w 為慣性權重值； $\{\rho_1, \rho_2\}$ 為均勻分布在 $[0, 1]$ 區間的隨機變數。PSO 透過粒子找尋各別 **pbest** 之後再辨出 **gbest** 的粒子所在位置，接著透過迭代驅使粒子往最佳解的所在位置移動，進而搜尋至最佳解或近似最佳解。

3.2 遞迴最小平方估計法

本研究使用遞迴最小平方估計法（Ljung and Ljung, 1987）調整複數模糊類神經系統後鑑部之系統參數。最小平方估計通常以線性的函數表示，定義如下：

$$d = \theta_1 f_1(u) + \theta_2 f_2(u) + \dots + \theta_m f_m(u) + \varepsilon \quad (26)$$

其中， u 為系統的輸入值； $\{f_i(u), i=1, 2, \dots, m\}$ 代表已知函數； $\{\theta_i, i=1, 2, \dots, m\}$ 可視為複數模糊類神經系統之後鑑部參數； ε 為系統誤差。假設從未知系統觀察所得之訓練資料定義如下：

$$\mathbf{TD} = \{(u_i, d_i), i=1, 2, \dots, n\} \quad (27)$$

其中， (u_i, d_i) 代表由（輸入值，目標值）所構成的第 i 組資料數據； n 代表訓練資料的大小。藉此可以得到一組由 n 個線性方程式所構成的集合，定義如下：

$$\begin{aligned} d_1 &= f_1(u_1)\theta_1 + f_2(u_1)\theta_2 + \dots + f_m(u_1)\theta_m + \varepsilon_1 \\ d_2 &= f_1(u_2)\theta_1 + f_2(u_2)\theta_2 + \dots + f_m(u_2)\theta_m + \varepsilon_2 \\ &\vdots \quad \quad \quad \vdots \quad \quad \quad \vdots \\ d_n &= f_1(u_n)\theta_1 + f_2(u_n)\theta_2 + \dots + f_m(u_n)\theta_m + \varepsilon_n \end{aligned} \quad (28)$$

$\{\varepsilon_i, i=1, 2, \dots, n\}$ 代表系統誤差。將 (28) 以陣列型態表式，如下所示：

$$\mathbf{d} = \mathbf{B}\boldsymbol{\theta} + \boldsymbol{\varepsilon} \quad (29)$$



其中， \mathbf{B} 、 $\boldsymbol{\theta}$ 、 \mathbf{d} 和 $\boldsymbol{\varepsilon}$ 表示如下：

$$\mathbf{B} = \begin{bmatrix} f_1(u_1) & f_2(u_1) & \cdots & f_m(u_1) \\ f_1(u_2) & f_2(u_2) & \cdots & f_m(u_2) \\ \vdots & \vdots & \cdots & \vdots \\ f_1(u_n) & f_2(u_n) & \cdots & f_m(u_n) \end{bmatrix} \quad (30)$$

$$\boldsymbol{\theta} = [\theta_1 \quad \theta_2 \quad \cdots \quad \theta_m]^T \quad (31)$$

$$\mathbf{d} = [d_1 \quad d_2 \quad \cdots \quad d_n]^T \quad (32)$$

$$\boldsymbol{\varepsilon} = [\varepsilon_1 \quad \varepsilon_2 \quad \cdots \quad \varepsilon_n]^T \quad (33)$$

為了最小化 $\|\boldsymbol{\varepsilon}\|^2$ ，可以藉由 RLSE 函數來最佳化參數 $\boldsymbol{\theta}$ ，定義如下：

$$\mathbf{P}_{k+1} = \mathbf{P}_k - \frac{\mathbf{P}_k \mathbf{b}_{k+1} (\mathbf{b}_{k+1})^T \mathbf{P}_k}{1 + (\mathbf{b}_{k+1})^T \mathbf{P}_k \mathbf{b}_{k+1}} \quad (34a)$$

$$\boldsymbol{\theta}_{k+1} = \boldsymbol{\theta}_k + \mathbf{P}_{k+1} \mathbf{b}_{k+1} (y_{k+1} - (\mathbf{b}_{k+1})^T \boldsymbol{\theta}_k) \quad (34b)$$

其中， $k=0, 1, 2, \dots, N-1$ ； $[\mathbf{b}_k^T, d_k]$ 代表矩陣 $[\mathbf{B}, \mathbf{d}]$ 內的第 k 列。為了執行 RLSE，必須先將 $\boldsymbol{\theta}_0$ 初始化為一個零向量；另一方面， \mathbf{P}_0 的初始化集合定義如下：

$$\mathbf{P}_0 = \alpha \mathbf{I} \quad (35)$$

其中， α 代表一個極大值， \mathbf{I} 代表單位矩陣。

3.3 複合式進化演算法

本研究提出之複合式進化演算法整合了粒子群最佳化演算法（PSO）與遞迴式最小平方估計法（RLSE），用以訓練 CNFS-ARIMA 模型之系統參數。PSO 用以更新前鑑部 (\mathbf{W}_{If}) 參數子集合，後鑑部 (\mathbf{W}_{Then}) 參數子集合則是透過 RLSE 來進行調整，兩者以協同合作的方式完成工作。RLSE 則遵循式 (34) 來進行後鑑部的參數更新，其中向量 \mathbf{b} 和 $\boldsymbol{\theta}$ 表示如下：

$$\mathbf{b}_{k+1} = [{}^1\mathbf{b}\mathbf{b}(k+1) \quad {}^2\mathbf{b}\mathbf{b}(k+1) \quad \cdots \quad {}^K\mathbf{b}\mathbf{b}(k+1)] \quad (36)$$

$${}^i\mathbf{b}\mathbf{b}(k+1) = [{}^i\lambda \quad {}^i\lambda h_1(k+1) \quad \cdots \quad {}^i\lambda h_M(k+1)] \quad (37)$$

$$\boldsymbol{\theta}_k = [{}^1\boldsymbol{\tau}_k \quad {}^i\boldsymbol{\tau}_k \quad \cdots \quad {}^K\boldsymbol{\tau}_k] \quad (38)$$



$${}^i\boldsymbol{\tau}_k = [{}^1a_0(k) \quad {}^1a_1(k) \quad \cdots \quad {}^1a_M(k)] \quad (39)$$

其中， $i=1, 2, \dots, K$ ； $k=0, 1, 2, \dots, N-1$ ； $\boldsymbol{\tau}^i(k)$ 代表 (38) 中第 k 組資料數據輸入 RLSE 後所產生的第 i 條模糊規則之後鑑部參數更新結果。將訓練資料 (27) 全部輸入到複合式進化演算法後，可以得到實際輸出與系統輸出之間的誤差值。透過這些誤差，可以進一步以平均平方誤差（mean square error, MSE）呈現兩數據集合之間的差距，MSE 之定義如下所示：

$$\text{MSE} = \frac{1}{n} \sum_{t=1}^n (e(t))^2 = \frac{1}{n} \sum_{t=1}^n (d(t) - \hat{y}(t))^2 \quad (40)$$

其中， $d(t)$ 代表期望的輸出值； $\hat{y}(t)$ 可由 (18) 式之結果經適當轉換而得，其代表 CNFS-ARIMA 模型之預測值。本研究使用 MSE 做為複合式進化學習演算法之成本函數。

4. 實驗實作與結果

本研究使用兩個時間序列資料作為時間序列預測問題之範例，用以測試本研究提出之 CNFS-ARIMA 預測系統與複合式進化學習演算法之系統效能。CNFS-ARIMA 模型作為預測模型，並藉由 PSO-RLSE 來調整其系統參數。本研究亦實作了傳統的類神經模糊系統（NFS）作為預測器，用以比較 CNFS 之系統映射能力。為求實驗結果的客觀性，本研究亦與文獻提出之研究結果進行比較。

4.1 太陽黑子時間序列

太陽黑子時間序列（sunspot time series）為時間序列相關研究中常使用之標竿資料集。該資料記錄從 1700 到 1979 年的太陽黑子平均數量資料，總共 280 筆資料（Hyndman, 2011）。本實驗將資料的範圍正規化至 $[0, 1]$ 區間，該時間序列可以表示為 $\{y(t), t=1, 2, \dots, 280\}$ ， t 為時間指標。將時間序列資料整理為輸入值與目標值的資料對（input-target data pair），可表示為 $\{(\mathbf{H}(i), d(i)), i=1, 2, \dots, 278\}$ ， $t=i+1$ ，

$$\mathbf{H}(i) = [y(t-1), y(t)]^T \quad (41)$$

$$d(i) = y(t+1) \quad (42)$$

其中， $\mathbf{H}(i)$ 為 CNFS-ARIMA(p, d, q) 的輸入向量； $d(i)$ 為實際目標值。本實驗將 1700 到 1920 年間的太陽黑子時間序列資料作為系統訓練資料，剩下的部分作為測試資料。CNFS 使用複數高斯模糊集合 (3a)，而 NFS 使用傳統的高斯模糊集合 (3b) 設



計。系統共有兩個輸入且每個輸入之語意變數皆由三個模糊集合組成，並採用格狀分割（grid partition），所以會產生 9 條 T-S 模糊法則。系統參數學習階段，本研究提出 PSO-RLSE 複合式學習演算法，並用 MSE 作為系統訓練的成本函數。本研究提出之複合式 PSO-RLSE 學習演算法之設定如表 1 所示。

▼ 表 1 PSO-RLSE 複合式學習方法之參數設定

PSO	
粒子維度	18
粒子數目	300
初始粒子位置	Random in $[0, 1]^{18}$
初始粒子速度	Random in $[0, 1]^{18}$
學習速率 (c_1, c_2)	(2, 2)
權重值 w	0.729
最大迭代數	300
RLSE	
後鑑部參數個數	27
θ_0	27×1 zero vector
\mathbf{P}_0	$\alpha \mathbf{I}$
α	10^8
\mathbf{I}	27×27 identity matrix

此外，本研究加入標準均方根誤差（normalized mean square error, NMSE）、平均絕對值誤差（mean absolute percentage error, MAPE）與泰勒不等係數（Thiel's inequality coefficient, Theil' U）三個評估指標，藉以獲得更為客觀的系統效能評估與比較。NMSE、MAPE 與 Theil' U 分別定義如下：

$$\text{NMSE} = \frac{1}{\delta^2 n} \sum_{t=1}^n (d(t) - \hat{y}(t))^2 \quad (43a)$$

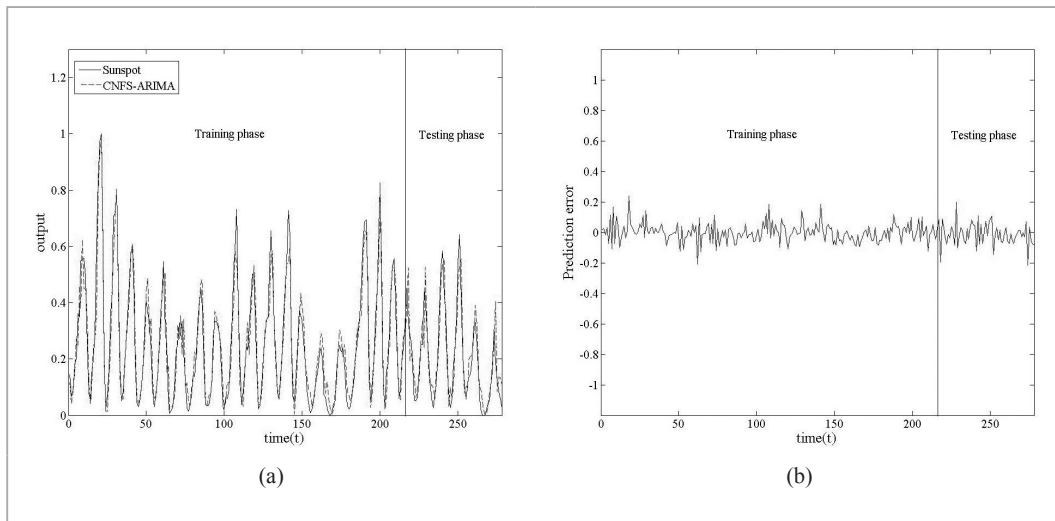
$$\delta = \frac{1}{n-1} \sum_{t=1}^n (d(t) - \bar{d}(t))^2 \quad (43b)$$



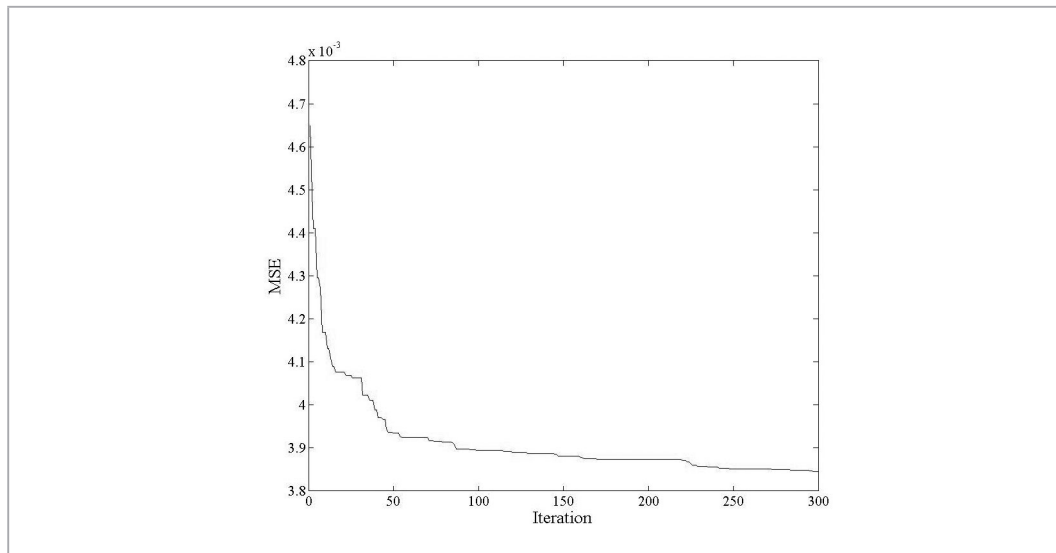
$$\text{MAPE} = \frac{1}{n} \sum_{t=1}^n \left| \frac{d(t) - \hat{y}(t)}{d(t)} \right| \times 100\% \quad (44)$$

$$\text{Theil' U} = \frac{\sqrt{\frac{1}{n} \sum_{t=1}^n (d(t) - \hat{y}(t))^2}}{\sqrt{\frac{1}{n} \sum_{t=1}^n d(t)^2 + \frac{1}{n} \sum_{t=1}^n \hat{y}(t)^2}} \quad (45)$$

其中， n 代表資料集的樣本數目； $\bar{d}(t)$ 表示目標值的平均； $\hat{y}(t)$ 代表預測值。NMSE 指標等同於在文獻（Lijuan, 2003）中所使用平均相對變動值（average relative variance, ARV）。而 MAPE 指標（Lewis, 1982）所計算之結果為百分比，因此不會因數值之大小而產生比較基礎不穩固的問題。MAPE 值愈小，表示模型預測能力愈好；Theil'U（Theil, 1958）用以衡量相對之誤差，其值介於 0 至 1 之間，當 U 之值越接近 0，表示預測值愈接近實際值。本研究所提出之方法與文獻所提出之方法效能比較如表 2 所示。其中，CNFS-ARIMA(2, 1, 0) 預測系統擁有最佳的系統測試誤差，其測試階段之 NMSE 為 0.1115；MSE 值為 4.50×10^{-3} 。圖 2 分別為 CNFS-ARIMA(2, 1, 0) 預測系統之預測結果與誤差；圖 3 為其學習曲線。



▲ 圖 2 (a)CNFS-ARIMA(2, 1, 0) 預測結果，其中黑色實線為實際觀測資料；紅色虛線為模型預測資料。(b)CNFS-ARIMA(2, 1, 0) 系統預測誤差。(太陽黑子時間序列)



▲ 圖 3 CNFS-ARIMA(2, 1, 0) 學習曲線。(太陽黑子時間序列)

▼ 表 2 效能比較 (太陽黑子時間序列)

方法	NMSE	MSE	MAPE	Theil' U
Threshold Autoregression(Tong and Lim, 1980)	0.28	N/A	N/A	N/A
ARMA (Cholewo and Zurada, 1997)	0.252	N/A	N/A	N/A
Elman (Cholewo and Zurada, 1997)	0.348	N/A	N/A	N/A
Exten. Elman(Cholewo and Zurada, 1997)	0.162	N/A	N/A	N/A
FIR (Cholewo and Zurada, 1997)	0.115	N/A	N/A	N/A
SVR (Deng et al., 2005)	0.64	N/A	N/A	N/A
Bagging SVR (Deng et al., 2005)	0.58	N/A	N/A	N/A
Boosting SVR 1 (Deng et al., 2005)	0.27	N/A	N/A	N/A
Boosting SVR 2 (Deng et al., 2005)	0.33	N/A	N/A	N/A
Ensemble SVR (Lijuan, 2003)	0.154	N/A	N/A	N/A
NFS-ARIMA(2, 0, 0)	0.1853	7.51×10^{-2}	6.47	0.1279
CNFS-ARIMA(2, 0, 0)(proposed approach)	0.1392	5.63×10^{-3}	6.03	0.0109
CNFS-ARIMA(2, 1, 0) (proposed approach)	0.1115	4.50×10^{-3}	5.88	0.0089



4.2 日經 225 指數時間序列

日經 225 (Nikkei 225, N225) 指數是日本股票市場最具代表性的股票指數，由東京股票交易所第一組掛牌的 225 種股票的價格所組成。由於指數延續時間較長，具有很好的可比性，成為考察日本股票市場股價長期演變及最新變動最常用和最可靠的指標。本實驗從 1999 年 10 月 4 日起至 2004 年 9 月 30 日止，共蒐集 1227 筆每日開盤價之日經 225 開盤指數資料 (Yahoo! Finance, 2011)，可表示為 $\{y(t), t=1, 2, \dots, 1227\}$ ，其中， $y(t)$ 為日經 225 指數的開盤價； t 為時間指標。時間序列資料可以被整理為輸入值與目標值的資料對，表示為 $\{(\mathbf{H}(i), d(i)), i=1, 2, \dots, 1225\}$ ，其中 $t=i+1$ ，輸入向量與預測目標值如下所示：

$$\mathbf{H}(i) = [y(t-1), y(t)]^T \quad (46)$$

$$d(i) = y(t+1) \quad (47)$$

其中， $\mathbf{H}(i)$ 為輸入向量； $d(i)$ 為實際目標值。本實驗將前 850 筆開盤指數資料作為系統訓練資料，剩下的部分作為測試資料。針對本研究所提出之 CNFS-ARIMA 模型結構設計，本研究使用格狀分割模式，每個輸入之語意變數皆由三個模糊集合組成。由於本實驗有兩個輸入值，因此會產生 9 條 T-S 模糊法則。系統參數學習階段，使用本研究所提出之複合式 PSO-RLSE 學習演算法對 CNFS-ARIMA 模型進行參數調整，並用 MSE 作為系統訓練的成本函數。複合式學習演算法之參數設定與實驗一相同，如表 1 所示。此外，為了與文獻 (Lu et al., 2009) 之研究結果進行比較，此範例採用與文獻相同之均方根誤差 (root mean square error, RMSE) 與平均絕對誤差 (mean absolute difference, MAD) 之指標設計。其中，RMSE 為 MSE 的均方根，而 MAD 定義如下：

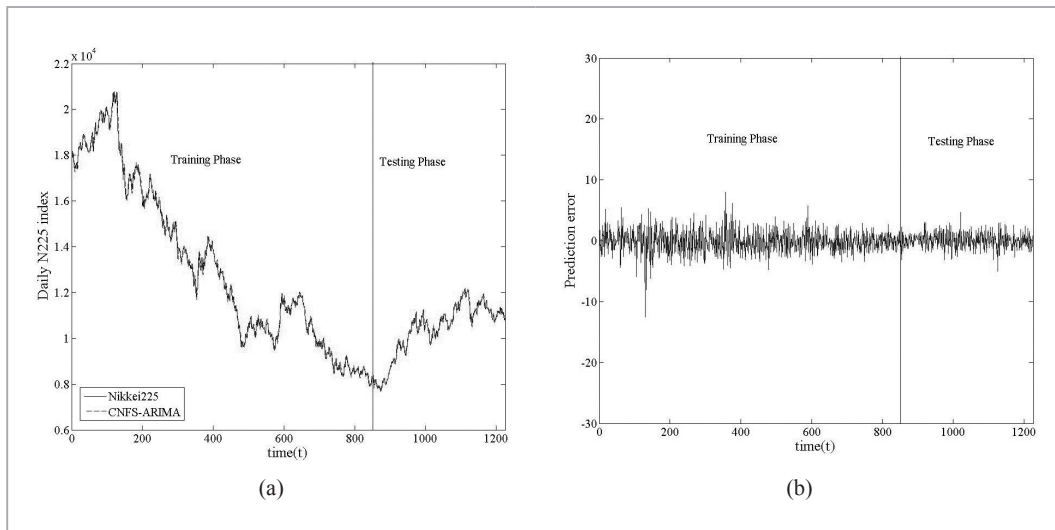
$$\text{MAD} = \frac{1}{n} \sum_{t=1}^N |d(t) - \hat{y}(t)| \quad (48)$$

其中， n 代表資料集的樣本數目； $d(t)$ 表示目標值； $\hat{y}(t)$ 代表預測值。本研究所提出之方法與文獻所提出之方法效能比較如表 3 所示。其中，CNFS-ARIMA(2, 1, 1) 預測系統擁有最佳的系統測試誤差。圖 4 為 CNFS-ARIMA(2, 1, 1) 預測系統之預測結果與誤差。



▼ 表 3 效能比較（日經 225 指數）

方法	RMSE	MAD	MAPE	Theil' U
Random walk (Lu et al., 2009)	137.85	105.77	N/A	N/A
SVR (Lu et al., 2009)	60.53	43.71	N/A	N/A
ICA-SVR (Lu et al., 2009)	56.76	40.86	N/A	N/A
NFS-ARIMA(2, 0, 0)	14.65	11.18	2.13	0.0907
CNFS-ARIMA(2, 0, 0) (proposed approach)	13.25	10.52	2.00	0.0751
CNFS-ARIMA(2, 0, 1) (proposed approach)	13.01	10.13	0.99	0.0717
CNFS-ARIMA(2, 1, 0) (proposed approach)	13.01	10.23	1.26	0.0717
CNFS-ARIMA(2, 1, 1) (proposed approach)	12.97	10.06	0.98	0.0712



▲ 圖 4 (a) CNFS-ARIMA(2, 1, 1) 預測結果，其中黑色實線為實際觀測資料；紅色虛線為模型預測資料。(b) CNFS-ARIMA(2, 1, 1) 系統預測誤差。（日經 225 指數時間序列）

5. 實驗結果分析與討論

本研究提出 CNFS-ARIMA 作為預測模型，其整合複數模糊集合之空間彈性與優良的資料映射能力，以及 ARIMA 對於非平穩性資料的處理能力，進而形成一個具高準確率的預測系統。由於過去文獻並未針對複數模糊集合提出明確的設計方法。因此，本研究提出一個全新的複數模糊集合之設計方法，並應用於本研究所提出之 CNFS-ARIMA 預測模型中。CNFS-ARIMA 以複數模糊規則為系統的架構，系統參數



可分為前鑑部與後鑑部參數。為了調整預測模型之系統參數，本研究提出 PSO-RLSE 複合式進化學習演算法。PSO 用以更新前鑑部；RLSE 用以調整後鑑部參數。此種複合式學習的優點在於能夠快速到達靠近最佳解的區域，以便尋找問題的最佳解答，並且降低演算法搜尋時的負載程度。藉由複合式進化演算法可以有效的找到預測模型參數集合之最佳解，進而提升系統的整體效能。

為了測試本研究所提出之系統方法，本研究使用文獻上常使用的兩個時間序列標準資料集來驗證 CNFS-ARIMA 模型的預測效能。由實驗結果可發現，使用複數模糊集之 CNFS-ARIMA 預測效能皆優於使用傳統模糊集合之 NFS-ARIMA 之預測率。例如在太陽黑子時間序列之實驗中，使用高斯複數模糊集合的 CNFS-ARIMA(2, 0, 0) 之 NMSE 值為 0.1392，其預測效能優於使用傳統高斯模糊集合 NFS-ARIMA(2, 0, 0) 約 24%。同樣的結果亦可於日經 225 指數實驗中得到驗證。此結果可證明本研究之論點，透過本研究所設計之複數模糊集合可有效提升預測模型的適應性與可塑性，進而達到較佳的系統預測能力。實驗結果顯示，本研究方法擁有卓越的預測效能，其結果皆優於文獻所提出之方法。在太陽黑子時間序列集（sunspot time series）的實驗中，CNFS-ARIMA(2, 1, 0) 預測系統擁有最佳的系統測試誤差，其測試階段之 NMSE 值為 0.1115，預測效能優於文獻中最佳的 Ensemble SVR 模型約 28%。在日經 225 股票指數（Nikkei 225 index）之預測實驗中，本研究所提出之 CNFS-ARIMA(2, 1, 1) 擁有優越的預測準確率（12.97 RMSE; 10.06 MAD）。其結果遠優 ICA-SVR（56.76 RMSE; 40.86 MAD）之預測效能約 77%。ARIMA 亦是 CNFS-ARIMA 預測模型能獲得良好映射能力的重要元件。ARIMA 包含誤差項與差分項的預測模型，進而延伸傳統 T-S 模糊法則型態。透過 ARIMA 差分之運算能力於非平穩資料上，能將資料進行平穩化的動作，進而提升預測的效能。依據實驗結果可以發現差分項的預測模型比單純 T-S 型態的預測模型擁有更佳的預測效能。值得注意的是加入差分項雖能增進預測效能，但二次差分之後的預測效能反而下降。造成此情形的原因是因為二次差分會將時間序列過度平穩化，使得預測模型無法準確判定該時間序列的趨勢，進而導致較大預測誤差。此外，本研究提出 PSO-RLSE 複合式學習演算法應用於最佳化 CNFS-ARIMA 模型的系統參數調整。其中，PSO 用於調整系統前鑑部參數；RLSE 用以處理並調整系統後鑑部參數，使整體系統參數在少數迭代即可讓成本函數值趨於收斂。如實驗一之學習曲線（圖 3）之結果可發現，前 10 個迭代即可獲得良好的效能，第 50 個迭代之後即趨於收斂，隨後的迭代變化量僅在 10^{-3} 之微小區間內變化。由此可證明 PSO-RLSE 複合式學習演算法的確具有良好的學習速率，而且透過兩個學習演算法之間搭配，亦能夠有效避免單一演算法求解時落入局部最佳解的問題。本研究亦與其他相關文獻進行比較，由實驗結果可得知，本研究所提出之系統方法皆可比文獻所提出之方法獲得更為優良的效能。



6. 結論

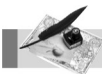
本研究提出一個 CNFS-ARIMA 模型作為預測系統，並應用於時間序列預測之研究。此模型結合複數模糊集、複數模糊推論、類神經模糊計算、差分自回歸移動平均模型以及最佳化演算法。模糊邏輯可以用來描述模糊或不精確的資訊並透過模糊推論過程產生具決定性的決策資訊。類神經網路架構具有相當良好的學習能力與容錯能力，在系統塑模上是廣泛被應用的模型。複數模糊集合可以提升類神經模糊系統對於非線性複雜問題的映射能力，達到較佳的系統彈性與適應能力。而差分自回歸移動平均模型對於線性時間序列資料有良好的評估能力，並透過差分的運算將非平穩之時間序列轉換為平穩性的時間序列資料，進而提升預測效能。本研究結合上述各理論之綜效，使 CNFS-ARIMA 模型成為一個擁有良好彈性與適應性的預測模型，並提出複合式學習演算法，能有效的找到預測模型參數的最佳解，進而提升系統的整體效能；由實驗結果可證實本研究所提出之系統方法可以獲得良好的預測效能。

參考文獻

- Box, G. E. P., & Jenkins, G. M. (1976). *Time series analysis, forecasting and control*. San Francisco: Holden-day.
- Boyacioglu, M. A., & Avci, D. (2010). An adaptive network-based fuzzy inference system (ANFIS) for the prediction of stock market return: The case of the Istanbul stock exchange. *Expert Systems with Applications*, 37(12), 7908-7912.
- Castillo, O., & Melin, P. (2002). Hybrid intelligent systems for time series prediction using neural networks, fuzzy logic, and fractal theory. *IEEE Transactions on Neural Networks*, 13(6), 1395-1408.
- Castro, J. L. (1995). Fuzzy logic controllers are universal approximators. *IEEE Transactions on Systems, Man and Cybernetics, Part A: Systems and Humans*, 25(4), 629-635.
- Chen, Y., Yang, B., Dong, J., & Abraham, A. (2005). Time-series forecasting using flexible neural tree model. *Information Sciences*, 174(34), 219-235.
- Chen, Z., Aghakhani, S., Man, J., & Dick, S. (2011). ANCFIS: A neurofuzzy architecture employing complex fuzzy sets. *IEEE Transactions on Fuzzy Systems*, 19(2), 305-322.
- Cho, K. B., & Wang, B. H. (1996). Radial basis function based adaptive fuzzy systems and their applications to system identification and prediction. *Fuzzy Sets and Systems*, 83(3), 325-339.
- Cholewo, T. J., & Zurada, J. M. (1997). Sequential network construction for time series



- prediction. *Proceedings of the IEEE International Joint Conference on Neural Networks*, Houston, Texas, USA.
- Deng, Y. F., Jin, X., & Zhong, Y. X. (2005). Ensemble SVR for prediction of time series. *Proceedings of the Fourth International Conference on Machine Learning and Cybernetics*, Guangzhou, China.
- Dick, S. (2005). Toward complex fuzzy logic. *IEEE Transactions on Fuzzy Systems*, 13(3), 405-414.
- Eberhart, R., & Kennedy, J. (1995). A new optimizer using particle swarm theory. *Proceedings of the Sixth International Symposium on Micro Machine and Human Science*, Nagoya, Japan.
- Hengjie, S., Chunyan, M., Roel, W., Zhiqi, S., & Catthoor, F. (2010). Implementation of fuzzy cognitive maps based on fuzzy neural network and application in prediction of time series. *IEEE Transactions on Fuzzy Systems*, 18(2), 233-250.
- Hornik, K., Stinchcombe, M., & White, H. (1989). Multilayer feedforward networks are universal approximators. *Neural Networks*, 2(5), 359-366.
- Huwendiek, O., & Brockmann, W. (1999). Function approximation with decomposed fuzzy systems. *Fuzzy Sets and Systems*, 101(2), 273-286.
- Hyndman, R. J. (2011). *Time series data library, physics, Wolf's sunspot numbers*. Retrieved June 01, 2010, from <http://www-personal.buseco.monash.edu.au/hyndman/TSDL/>
- Jang, J. S. R. (1993). ANFIS: Adaptive-network-based fuzzy inference system. *IEEE Transactions on Systems, Man, and Cybernetics*, 23(3), 665-685.
- Kasabov, N. K., & Qun, S. (2002). DENFIS: Dynamic evolving neural-fuzzy inference system and its application for time-series prediction. *IEEE Transactions on Fuzzy Systems*, 10(2), 144-154.
- Kennedy, J., & Eberhart, R. (1995). Particle swarm optimization. *Proceedings of the IEEE International Conference on Neural Networks*, Piscataway, NJ, USA.
- Lewis, C. D. (1982). *Industrial and business forecasting method*. London: Butterworth.
- Li, C., & Chiang, T. W. (2010). Complex neuro-fuzzy self-learning approach to function approximation. *Lecture Notes in Artificial Intelligence*, 5991, 289-299.
- Li, C., & Chiang, T. W. (2011a). Complex fuzzy computing to time series prediction-A multi-swarm PSO learning approach. *Lecture Notes in Artificial Intelligence*, 6592, 242-251.
- Li, C., & Chiang, T. W. (2011b). Complex fuzzy model with PSO-RLSE hybrid learning



- approach to function approximation. *International Journal of Intelligent Information and Database Systems*, 5(4), 409-430.
- Lijuan, C. (2003). Support vector machines experts for time series forecasting. *Neurocomputing*, 51, 321-339.
- Ljung, L., & Ljung, E. (1987). *System identification: Theory for the user*. Englewood Cliffs, NJ: Prentice-Hall.
- Lu, C. J., Lee, T. S., & Chiu, C. C. (2009). Financial time series forecasting using independent component analysis and support vector regression. *Decision Support Systems*, 47(2), 115-125.
- Ramot, D., Friedman, M., & Kandel, A. (2002). Complex fuzzy sets. *IEEE Transactions on Fuzzy Systems*, 10(2), 171-186.
- Ramot, D., Friedman, M., Langholz, G., & Kandel, A. (2003). Complex fuzzy logic. *IEEE Transactions on Fuzzy Systems*, 11(4), 450-461.
- Theil, H. (1958). *Economic forecasts and policy*. Amsterdam: North Holland.
- Tong, H., & Lim, K. S. (1980). Threshold autoregression, limit cycles and cyclical data. *Journal of the Royal Statistical Society. Series B (Methodological)*, 42(3), 245-292.
- Tsay, R. S. (2005). *Analysis of financial time series*. NJ: Wiley-Interscience.
- Tung, W. L., & Quek, C. (2011). Financial volatility trading using a self-organising neural-fuzzy semantic network and option straddle-based approach. *Expert Systems with Applications*, 38(5), 4668-4688.
- Wang, L. X., & Mendel, J. M. (1992). Fuzzy basis functions, universal approximation, and orthogonal least-squares learning. *IEEE Transactions on Neural Networks*, 3(5), 807-814.
- Wooldridge, J. M. (2006). *Introductory econometrics: A modern approach*. Mason, OH: Thomson/South-Western.
- Yahoo! Finance (2011). *Nikkei 225 Stock Index*. Retrieved June 5, 2011, from <http://finance.yahoo.com/q?s=^N225>
- Zhang, G., Dillon, T. S., Cai, K. Y., Ma, J., & Lu, J. (2009). Operation properties and d-equalities of complex fuzzy sets. *International Journal of Approximate Reasoning*, 50(8), 1227-1249.