

# 應用獨立成份分析與支援向量迴歸於財務時間序列 預測

呂奇傑

清雲科技大學工業工程與管理系

李天行

輔仁大學管理學研究所

陳學群

輔仁大學應用統計研究所

## 摘要

由於財務時間序列資料具有高頻率、雜訊、非定態與混沌等性質，因此在現今時間序列預測領域中，向來被認為是一極具挑戰性的應用領域。本研究提出一結合獨立成份分析(Independent component analysis, ICA)與支援向量迴歸(Support Vector Regression, SVR)之財務時間序列預測模式，先利用 ICA 具有將混合訊號分離出個別獨立之來源訊號之能力，從預測變數中估計出獨立成份，並在去除代表雜訊的獨立成份及保留主要的獨立成份作為預測變數後，再使用 SVR 以濾除雜訊後之預測變數建構預測模式。期望可以讓 SVR 在建構模式時不受雜訊影響，進而提升預測結果的準確度。為驗證所提方法之有效性，本研究將以日經 225 現貨開盤指數及台股現貨收盤指數之預測進行實證研究，並與單純使用 SVR 模式及隨機漫步模式的預測結果作比較。實證結果顯示，所提之方法不論是在預測誤差或是預測準確度的表現上均較單純使用 SVR 及隨機漫步模式為佳。

**關鍵詞：**獨立成份分析，支援向量迴歸，財務時間序列預測，股價指數



# Using Independent Component Analysis and Support Vector Regression for Financial Time Series Forecasting

Chi-Jie Lu

Department of Industrial Engineering and Management, Ching Yung University

Tian-Shyug Lee

Graduate Institute of Management, Fu-Jen Catholic University

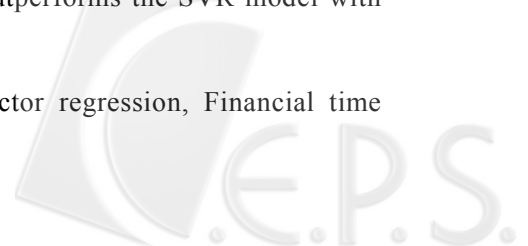
Hsueh-Chun Chen

Graduate Institute of Applied Statistics, Fu-Jen Catholic University

## Abstract

As financial time series are inherently noisy, non-stationary and deterministically chaotic, it is one of the most challenging applications of modern time series forecasting. Due to the advantages of the generalization capability in obtaining the unique and global optimal solution, support vector regression (SVR), has also been successfully applied in time series prediction, especially in the financial time series forecasting. In the modeling of financial time series using SVR, one of the key problems is the inherent high noise. Therefore, detecting and removing the noise are important but difficult tasks when building an SVR forecasting model. To alleviate the influence of noise, a two-stage approach by integrating independent component analysis (ICA) and support vector regression is proposed in this research for financial time series forecasting. ICA is a novel statistical signal processing technique that was originally proposed to find the latent source signals from observed mixture signal without knowing any prior knowledge of the mixing mechanism. The proposed approach first uses ICA to the forecasting variables for generating the independent components (ICs). After identifying and removing the ICs containing the noise, the rest of the ICs are then used to reconstruct the forecasting variables which contain less noise. The SVR is then applied to use the filtered (or denoised) forecasting variables to build the forecasting model. In order to evaluate the performance of the proposed approach, the Nikkei 225 opening index is used as the illustrative example. The experimental results show that the proposed model outperforms the SVR model with non-filtered forecasting variables and random walk model.

**Key words:** Independent component analysis, Support vector regression, Financial time series forecasting, Stock index



## 壹、緒論

隨著經濟的全球化與各種金融商品的推陳出新，現今的投資環境正急遽的變動中，為了在金融市場的操作中獲取超額報酬，投資者莫不期望能透過各種分析及預測工具，及早發現金融商品的價格變化，以作為交易時的策略依據，也因此造就財務時間序列預測的蓬勃發展。但由於金融商品的交易特性以及價格容易受政府政策、景氣循環、國際市場的漲跌及物價膨脹等眾多因素影響，使財務時間序列具有高頻率(High frequency)、雜訊(Inherently noisy)、非定態(Non-stationary)與混沌(Chaotic)等性質(Deboeck 1994; Yaser & Atiya 1996)，因而財務時間序列成為一非常複雜、極具挑戰性以及難以預測的時間序列資料。

財務時間序列預測所使用的工具可分為定性型(Qualitative)與定量型(Quantitative)兩類。定性型預測方法主要使用在缺乏歷史數據，或者是不可量化之歷史資料上，基本上運用專家的經驗、判斷和意見來進行預測，常用的模型有德爾菲法(Delphi Technique)、行政長官意見法(Jury of Executive Opinion)和歷史類推法(Historical Analogy)等；而定量型預測方法主要是藉由過去的歷史數據推估未來之預測值，其中又可以分為以統計為基礎的傳統預測模型以及人工智慧方法兩大類。傳統預測模型中常見的方法有移動平均法(Moving average)、指數平滑法(Exponential smoothing)、迴歸分析法、ARIMA、向量自我迴歸(Vector autoregression)模型及 GARCH 模式等(Antoniou & Holmes 1995; Jung & Boyd 1996; Parisi & Vasquez 2000; Balachandher *et al.* 2002; Kwon & Kish 2002)。傳統預測方法由於常需符合模式使用上的統計假設，及較適合處理資料間符合線性關係之資料型態，因而限制了傳統預測模型的實用性與精確度(Lee & Chen 2002, Lee & Chiu 2002)。

而在人工智慧方法方面，由於沒有傳統統計方法建構模式時需要滿足許多假設條件的要求，及其在建構非線性模式方面的優越能力，因此越來越受到重視，並用於各種預測問題上，其中又以倒傳遞類神經網路(Back-propagation neural networks, BPNN)最常被討論與使用(Zhang *et al.*, 1998; Vellido *et al.* 1999)。但倒傳遞類神經網路主要的缺點在於學習的時間較長、不易進行變數篩選及容易陷入局部最佳解(Local optimum)等問題；除此之外，網路架構(Topology)不易決定亦為倒傳遞類神經網路不足之處(Cao & Tay 2001; Lee *et al.* 2003; Cao 2003; Kim 2003)。

近來，以統計學習理論(Statistical learning theory)為基礎所發展之支援向量機(Support vector machine, SVM)是人工智慧領域中一新興且受到重視的機器學習系統(Vapnik *et al.* 1997; Vapnik 2000)。其主要概念是在高維度之特徵空間(High-dimensional feature space)中尋找一個擁有最大邊界之超平面(Hyperplane)，以有效分隔兩種不同類型的資料。原始的 SVM 被建立於處理二元分類問題，由於 SVM 並不要求相關研究領域之先驗知識及優異之推論績效，使其受到廣大的歡迎並且廣泛的用於不同的領域，

包含文字分類(Texture classification)、影像識別(Image recognition)、手寫數字辨識(Hand-written digit recognition)、資料探勘(Data mining)及生物資訊(Bioinformatics)等 (Burbidge *et al.* 2001; Kim *et al.* 2002; Chang *et al.* 2003; Norinder 2003; Li *et al.* 2003; Shin *et al.* 2005)。而除了應用於分類問題外,Drucker *et al.* (1997)及 Vapnik *et al.* (1997) 導入 $\epsilon$ -insensitivity 損失函數(Loss function)之概念於 SVM 中,提出了支援向量迴歸(Support vector regression, SVR)模式來處理非線性的迴歸問題。由於支援向量迴歸能捕捉特徵空間的幾何特性及時間序列資料中的特徵,已逐漸的被用於各種預測問題中並獲得良好的預測績效(Thissen *et al.* 2003; Mohandes *et al.* 2004; Koike & Takagi 2004; Karras & Mertzios 2004),因此本研究將使用 SVR 為建構財務時間序列模型之工具。

在建立財務時間序列預測模型時,普通面臨之問題就是財務時間序列資料包含許多的雜訊。當資料中存在雜訊時,會使得預測模式因在訓練過程中學習到這些雜訊的資訊,而產生過度配適(Over-fitting)或配適不足(Under-fitting)的問題,進而在測試階段喪失預測績效的可靠性與準確度,因此如何偵測以及去除財務時間序列資料中的雜訊,以降低其對 SVR 之預測結果的影響,是一項重要的議題。為因應前述之問題,一些特殊的 SVR 技術,如 robust SVR(Chuang *et al.* 2002)及 weighted Least Square SVR(Suykens *et al.* 2002),已經被提出用以強化 SVR 在面對預測資料包含雜訊時的預測能力,然而這些修正的 SVR 技術不是需要大量的運算時間與次數,就是無法確定能夠得到較佳的預測結果。

為了提升財務時間序列預測之績效及解決前述 SVR 可能面臨之問題,本研究導入近年來快速發展用於訊號分離(Signal separation)之獨立成份分析(Independent component analysis, ICA)技術於財務時間序列預測問題上,提出一個結合獨立成份分析與支援向量迴歸之兩階段時間序列預測模式,將利用 ICA 來偵測及濾除財務時間序列資料中的雜訊,以提升 SVR 之預測準確度。所提預測模式之建構流程為先使用 ICA 於預測變數中分離出個別獨立的潛在來源訊號(即獨立成份),接著從這些獨立成份中去除代表資料中雜訊的獨立成份之後,將保留的獨立成份重建時間序列資料,得到濾除雜訊後之預測變數,最後再利用 SVR 以濾除雜訊後的預測變數建構預測模式。本研究提方法能去除前述修正 SVR 模式需大量運算時間或是無法得到最佳解的缺點,也能減少 SVR 在模式建構時,因受雜訊影響而產生過度配適或配適不足的問題,進而提升 SVR 模式預測結果的準確度。為驗證所提方法之有效性,本研究將以日經 225 現貨開盤指數及台股現貨收盤指數之預測進行實證研究,並與單純使用 SVR 模式及隨機漫步模式的預測結果作比較。

本文各節區分如下:第二節將回顧獨立成份分析及支援向量迴歸之相關文獻;第三節針對獨立成份分析及支援向量迴歸進行簡單的介紹;第四節則說明所提之兩階段預測模型之建構方式;第五節為預測模型之實驗結果與討論,並與單純使用 SVR 模式及隨機漫步方法的結果進行比較;最後之第六節則為本研究之結論與建議。

## 貳、文獻回顧

### 一、獨立成份分析(Independent component analysis, ICA)

獨立成份分析(ICA)是一個新穎之訊號處理技術，主要是用來處理未知來源分離(Blind Source Separation, BSS)的問題(Lee 1998; Hyvärinen & Oja 2000; Cichocki & Amari 2002)，其目的是希望在沒有任何有關訊號混合機制的事前資訊下，從觀察到的混合訊號(Mixture signals)中找出潛在來源訊號(Latent source signals)，這些潛在來源訊號在 ICA 中稱為獨立成份(Independent Component, IC) (Hyvärinen *et al.* 2001)。

ICA 模式已廣泛的被應用在醫學訊號處理、語音訊號處理、特徵萃取及人臉辨識的問題中(Bartlett & Sejnowski 1997; Park *et al.* 1999; Vigarito *et al.* 2000; Jung *et al.* 2001; Bartlett *et al.* 2002; Jang *et al.* 2002; James & Gibson 2003; Beckmann & Smith 2004; Lin *et al.* 2004; Kim *et al.* 2004)。Visser & Lee(2003)利用 ICA 於語音增強上，假設觀察訊號是由背景噪音(Background noise)、干擾來源(Interfering point source)及發音來源訊號(Speaker signal)所混合而成，他們先透過 ICA 於觀察訊號中找出獨立成份，並於獨立成份中將干擾來源去除，之後再利用說話者活動偵測(Speaker activity detection)技術去除背景噪音以達到語音增強的目的。Ikeda & Toyama(2000)應用 ICA 於磁振造影(MEG)資料，他們先利用因素分析(Factor analysis)做前處理使磁振造影訊號非相關，再利用 ICA 於經過前處理後之訊號，可得到磁振造影資料的獨立成份。真實腦波訊號、機械上感應器之雜訊及人工訊號的特徵將各別的被凸顯在獨立成份中，進而提供觀察者更有意義的訊息。Déniz *et al.*(2003)結合 ICA 與 SVM 進行人臉辨識，先利用 ICA 對訓練影像估計出獨立成份，並以獨立成份作為人臉特徵，之後利用 SVM 於此人臉特徵建立分類模式以進行人臉分類。

由相關文獻中可知，ICA 模式迄今仍不常被用於財務資料的處理。Back & Weigend (1997)應用 ICA 模式於分析股價指數走勢的特徵，使用 ICA 模式針對日本股票交易所前 28 大公司每日收盤價的時間序列資料擷取獨立成份，接著將這些獨立成份進行排序，實驗結果顯示，排序在前的獨立成份能抓住股價走勢的特徵。Kiviluoto & Oja(1998)應用 ICA 模式分析影響某零售商現金流量的潛在因素，經由對獨立成份的分析可得知此零售商之現金流量的多寡主要受假日、季節及其他零售商競爭的影響。Oja *et al.*(2000)利用 ICA 模式進行匯率之預測，先利用 ICA 針對匯率資料獲得獨立成份，再透過線性與非線性轉換來過濾獨立成份中的雜訊，之後利用自我迴歸(Autoregression, AR)模型預測下一期各獨立成份值，最後這些預測之獨立成份被用來產生下一期的匯率預測值，結果顯示結合 AR 與 ICA 之預測結果較單純使用 AR 的結果為佳。



## 二、支援向量迴歸(Support vector regression, SVR)

支援向量機(Support vector machine, SVM)是以統計學習理論(Statistical learning theory)為基礎，運用結構風險最小化原理(Structural risk minimization inductive principle)以及最大化在群體間之點的邊界(Margin)所提出之機器學習系統(Vapnik *et al.* 1997; Vapnik 2000)。SVM 主要用來解決線性二次規劃(Quadratic programming, QP)問題，並且已經廣泛的應用於解決分類的問題。而在處理預測問題上，Drucker *et al.*(1997)及 Vapnik *et al.*(1997)利用 $\epsilon$ -insensitivity 損失函數將 SVM 的概念運用於迴歸問題中，而提出支援向量迴歸(SVR)模型。SVR 由於能捕捉特徵空間的幾何特性及時間序列資料中的特徵，因此逐漸的被用於各種預測問題中，並獲得良好的預測績效(Koike & Takagi 2004; Karras & Mertzios 2004)。Thissen *et al.* (2003)利用 SVR、ANN 及 ARMA 做為預測模型，實證兩組模擬資料及一組紡紗過程中濾清器數值之時間序列資料，結果發現 SVR 之預測結果較 ARMA 及 ANN 為佳。Mohandes *et al.* (2004)利用 SVR 及多層感知器(Multilayer Perceptrons, MLP)於風速預測，透過 RMSE 作為評估指標，結果發現 SVR 之表現優於 MLP 模型。Pai & Lin(2005)利用 SVR、SARIMA(Seasonal ARIMA)及廣義回歸神經網路(General Regression Neural Network, GRNN)預測具有季節性的工業機械總產值，結果顯示 SVR 優於 GRNN 及 SARIMA 模型。

而在財務時間序列預測方面，Tay & Cao(2001)使用支援向量機預測財金市場之數據，並且進行參數之敏感性分析以選擇預測模式之最佳參數，實驗結果顯示支援向量機模型優於倒傳遞類神經網路。Kim(2003)則利用技術性指標作為輸入變數，配合 SVR、BPNN 及案例推理法(Case-based reasoning)做為預測模型，實證每日韓國綜合股價指數(KOSPI)收盤價的漲跌，結果顯示 SVR 之表現最佳，能準確預測 57.83%的漲跌。Cao(2003)提出整合 SVR 之時間序列專家預測系統，該系統為兩階段的網路架構，首先透過自我組織特徵映射網路(Self-organizing feature map network)作為集群分析演算法，將輸入變數空間區分成數個無關聯的集群，第二階段則透過 SVR 針對各群選擇最佳預測模型，實證太陽黑子資料及聖塔菲研究院(Santa Fe Institute)之虛擬股票市場資料的結果顯示，使用此整合 SVR 預測系統之預測結果優於 SVR。

## 參、研究方法

### 一、獨立成份分析

獨立成份分析(ICA)是一種用來找出隨機變數中隱藏因子的統計方法(Comon 1994; Lee 1998; Hyvärinen & Oja 2000)。ICA 假設觀察到的隨機變數是由未知的潛在變數(Latent variables)以線性方式組合而成，而組合的機制未知。這些潛在變數在假設互為獨立的情況下，被稱為觀察資料的潛在來源(Latent sources)。ICA 就是在只有觀察到

的隨機變數資料，而未知其混合機制與來源的情況下，找出獨立成份(Independent components, ICs)以估計潛在來源的方法。

令  $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_M]^T$  為由  $\mathbf{x}_i$  所組成維度為  $M \times N$ ， $M \leq N$ ，的混合訊號矩陣，其中  $\mathbf{x}_i$  為維度  $1 \times N$  的時間序列訊號，一個基本的 ICA 模式可以表示為(Hyvärinen & Oja 2000)

$$\mathbf{X} = \mathbf{A}\mathbf{S} = \sum_{i=1}^M \mathbf{a}_i \mathbf{s}_i \quad (1)$$

其中  $\mathbf{a}_i$  是維度  $M \times M$  的未知混合矩陣(Mixing matrix)  $\mathbf{A}$  的第  $i$  個行向量(Column vector)，向量  $\mathbf{s}_i$  是維度  $M \times N$  的未知來源矩陣(Source matrix)  $\mathbf{S}$  的第  $i$  個列向量(Row vector)。 $\mathbf{s}_i$  就是無法由混合訊號矩陣  $\mathbf{X}$  所直接觀測的潛在來源訊號。在假設來源訊號  $\mathbf{s}_i$  互為統計獨立的情況下，為了估計  $\mathbf{s}_i$ ，ICA 的方法是找到一個維度  $M \times M$  的解混合矩陣(Demixing matrix)  $\mathbf{W}$ ，將所觀察到的混合訊號矩陣  $\mathbf{X}$  進行轉換，以產生維度為  $M \times N$  的矩陣  $\mathbf{Y}$ ，亦即

$$\mathbf{Y} = [\mathbf{y}_i] = \mathbf{W}\mathbf{X} \quad (2)$$

其中， $\mathbf{y}_i$  是矩陣  $\mathbf{Y}$  的第  $i$  個列向量(Row vector)。向量  $\mathbf{y}_i$  之間必須盡可能的互為統計獨立，並且被稱為稱為獨立成份(Independent component, IC)。當解混合矩陣  $\mathbf{W}$  為混合矩陣  $\mathbf{A}$  的反矩陣時，即  $\mathbf{W} = \mathbf{A}^{-1}$ ，這些獨立成份將可用來估計潛在來源訊號  $\mathbf{s}_i$ 。

ICA 模型有一個基本的假設或限制，即各獨立成份間必須統計獨立(Hyvärinen & Oja 2000)。因此，ICA 可以被描述成一個最佳化的問題，將各獨立成份的獨立性量測當成目標函數，經由最佳化的技術找出解混合矩陣  $\mathbf{W}$ 。亦即，解混合矩陣  $\mathbf{W}$  可以由一個非監督式的演算法決定，而這個演算法的目標就是使得各 IC 間的統計獨立性最大。由於 IC 具有非高斯分配(non-Gaussian distribution)隱含各個 IC 之間為統計獨立(Hyvärinen & Oja 2000; Hyvärinen *et al.* 2001)，因此 IC 的非高斯特性(non-Gaussianity)可以被用做 ICA 演算法的目標函數。在 ICA 中用來量測獨立成份的非高斯特性也有很多種方法，常見的有高階統計量法(High order cumulants)，共同資訊法(Mutual information)及負熵法(Negentropy)，其中以負熵法最常被討論與使用(Hyvärinen *et al.* 2001; David & Sanchez 2002)。

負熵法是延伸資訊理論中的熵(Entropy)的定義所提出來的方法，在資訊理論中熵為一資訊量的測量單位，用來表示隨機變數之觀察值的資訊程度，當隨機變數的結構性越差越亂(即越無法預測)時，其熵會越大。令隨機向量  $\mathbf{y}$  之機率密度函數為  $p_y(\eta)$ ，則  $\mathbf{y}$  的熵  $H$  可定義為  $H(\mathbf{y}) = -\int p_y(\eta) \log p_y(\eta) d\eta$  (Cover & Thomas 1991)。根據資訊理論，在所有具有相同變異數的隨機分佈中，呈高斯分佈的變數會具有最大的熵。因此若變數具非高斯分佈，則熵值會較其具有高斯分佈時來的小(Hyvärinen *et al.* 2001)。由此延伸，可以對熵  $H$  的定義做一修正當作非高斯的測量值。熵  $H$  修改後的定義稱為負熵  $J$  (Negentropy)，定義如下(Hyvärinen *et al.* 2001)：

$$J(\mathbf{y}) = H(\mathbf{y}_{gauss}) - H(\mathbf{y}) \quad (3)$$

其中  $\mathbf{y}_{gauss}$  表示和  $\mathbf{y}$  具有相同變異數的高斯變數向量。負熵的值永不為負值，即  $J(\mathbf{y}) \geq 0$ ，只有  $\mathbf{y}$  為高斯分佈時  $J(\mathbf{y})$  才會為零，所以 ICA 的目標函數就變成要最大化隨機向量  $\mathbf{y}$  的負熵，即 *Maximize*  $J(\mathbf{y})$ 。

然而，負熵最大的問題就是計算太過複雜，因為根據定義，要估計出負熵就必須先估計  $\mathbf{y}$  的機率密度函數  $f(\mathbf{y})$ 。為了解決這這個問題，Hyvärinen(1999)發展出負熵的近似函數

$$J(y) \propto [E\{G(y)\} - E\{G(v)\}]^2 \quad (4)$$

其中  $v$  是平均值為零且變異數為 1 的高斯分佈之隨機變數。 $G$  可為任何非二次方函數(Non-quadratic function)，因為  $G$  若為二次方函數， $J(y)$  必為零。

在眾多用來求解 ICA 模式的演算法中，Hyvärinen (1999) 所提出 *FastICA* 演算法由於具有良好的運算效率，且是處理 ICA 問題上最常被使用的演算法之一，因此本研究將使用 *FastICA* 演算法處理從財務時間序列資料中分離出獨立成份的工作。

## 二、支援向量迴歸(Support vector regression)

為說明 SVR 模式，首先考慮一個典型的迴歸模式，其目的是估計一個未知方程式  $r(\mathbf{x})$  來預測一個未知的值  $q$ ，兩者的關係如下：

$$q = r(\mathbf{x}) + \delta$$

其中  $\delta$  是隨機獨立並且平均數為零的隨機誤差(Random error)， $\mathbf{x}$  是多變數的輸入變數， $q$  是單一(Scalar)的輸出值。這個方程式可以用有限個數的訓練樣本來建立： $(\mathbf{x}_i, q_i)$ ， $(i=1, \dots, n)$ 。在迴歸問題中可分為線性迴歸與非線性迴歸問題，線性迴歸問題是較易處理，並且已有相當多的技術被提出來與成功的應用；相對而言，非線性迴歸則較不容易處理。SVR 主要就是被發展來處理非線性迴歸的議題的工具，其主要概念是將一個在低維度輸入空間(Input space)中非線性迴歸之問題，轉換成在高維度特徵空間(Feature space)中的線性迴歸問題。因此在 SVR 中，輸入變數  $\mathbf{x}$  首先被轉換(Map)到一個高維度的特徵空間( $\mathbf{F}$ )中，在這特徵空間中，輸入變數可以被修正為與輸出變數有線性的關係。以數學模式表示，SVR 模式可以表示為(Vapnik 2000)

$$f(\mathbf{x}) = (\boldsymbol{\omega} \cdot \Phi(\mathbf{x})) + b \quad (5)$$

其中， $\boldsymbol{\omega}$  為權重向量， $b$  是為偏差值(bias)， $\Phi(\mathbf{x})$  為一轉換函數，用以將輸入變數  $\mathbf{x}$  以非線性的方式轉換到高維度的特徵空間中， $(\boldsymbol{\omega} \cdot \Phi(\mathbf{x}))$  則用來描述在特徵空間  $\mathbf{F}$  中的值。

在迴歸問題中，可以在訓練資料用最小化實驗風險(Minimization of empirical risk)的準則，估計迴歸模式中的參數。典型被用在最小化實驗風險中的損失函數包含有絕對值誤差及平方誤差等，但在 SVR 中，Vapnik *et al.* (1997)則使用了一個稱為  $\epsilon$ -insensitivity 的損失函數( $L_\epsilon$ )：



$$L_\varepsilon(f(\mathbf{x})-q) = \begin{cases} |f(\mathbf{x})-q|-\varepsilon & \text{for } |f(\mathbf{x})-q| \geq \varepsilon \\ 0 & \text{otherwise} \end{cases}$$

其中， $\varepsilon$  代表一個參數，用來定義一個圍繞在迴歸方程式  $f(\mathbf{x})$  周圍的管狀區域(如圖 1 中的虛線所示)。這些管狀區域就是所謂的  $\varepsilon$ -insensitivity 區域，當預測值落在管狀區域內時其損失為零，亦即當預測誤差小於  $\varepsilon$  時，其損失成本會相當於零；反之，當預測值落在區域外時，其損失相當於預測值跟區域邊界的差，可以加入寬鬆變數(Slack variables,  $\xi, \xi^*$ )來表示之，

$$|q_i - f(\mathbf{x}_i)| - \varepsilon = \xi_i, \quad |q_i - f(\mathbf{x}_i)| - \varepsilon = \xi_i^*, \quad i = 1, \dots, n$$

其中  $\xi$  為高於區域邊界的訓練誤差值， $\xi^*$  為低於區域邊界的於訓練誤差值。一個支援向量迴歸的示意圖如圖 1 所示。

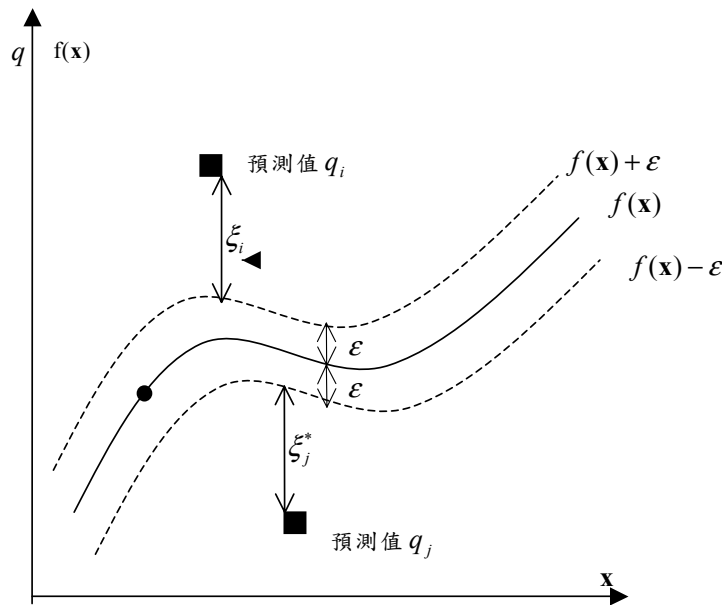


圖 1：支援向量迴歸示意圖

經由上述的說明，在使用  $\varepsilon$ -insensitivity 損失函數及一個正的調整常數 (Regularization constant)  $C$  來決定實驗誤差(Empirical error)和調整項(Regularization term)之間的轉換，SVR 模式的目的是使下面的方程式最小化(Vapnik 2000)

$$\text{Minimize: } \frac{1}{2} \|\omega\|^2 + C \sum_{i=1}^n (\xi_i + \xi_i^*)$$

$$\text{Subject to } \begin{cases} q_i - (\omega \cdot \Phi(\mathbf{x}_i)) - b \leq \varepsilon + \xi_i \\ (\omega \cdot \Phi(\mathbf{x}_i)) + b - q_i \leq \varepsilon + \xi_i^* \\ \xi_i, \xi_i^* \geq 0, \text{ for } i = 1, \dots, n \end{cases}$$

(6)



公式(6)為二次規劃(QP)的最佳化問題，可以用拉式乘數(Lagrange Multipliers)轉換成對偶問題來求解極大值，公式(6)可以轉換為下式

Maximize :

$$L_d(\alpha, \alpha^*) = -\varepsilon \sum_{i=1}^n (\alpha_i^* + \alpha_i) + \sum_{i=1}^n (\alpha_i^* - \alpha_i) q_i - \frac{1}{2} \sum_{i,j=1}^n (\alpha_i^* - \alpha_i)(\alpha_j^* - \alpha_j) \mathbf{K}(\mathbf{x}_i, \mathbf{x}_j)$$

$$\text{Subject to } \begin{cases} \sum_{i=1}^n (\alpha_i^* - \alpha_i) = 0 \\ 0 \leq \alpha_i \leq C, i=1, \dots, n \\ 0 \leq \alpha_i^* \leq C, i=1, \dots, n \end{cases} \quad (7)$$

其中， $\alpha$ 及 $\alpha^*$ 為拉式乘數，只有少數不為零，而不為零的 $\alpha$ 及 $\alpha^*$ 所對應的資料點即稱為支援向量(Support vectors)； $\mathbf{K}(\mathbf{x}_i, \mathbf{x}_j)$ 為核心函數(Kernel function)，用來將輸入空間的資料轉換成特徵空間的資料，其等於兩個向量的內積， $\mathbf{K}(\mathbf{x}_i, \mathbf{x}_j) = \Phi(\mathbf{x}_i) \times \Phi(\mathbf{x}_j)$ ，且必須需滿足 Mercer's 條件(Vapnik 2000)。常用的核心函數有多項式(Polynomial)函數及輻射基底函數(Radial basis function, RBF)(Lin *et al.* 2003; Cherkassky & Ma 2004)。

訓練一個 SVR 模式等同於根據公式(7)的限制式下，最佳化拉式乘數 $\alpha$ 及 $\alpha^*$ 以找出適當的支援向量個數來建立模式，目前已有許多的方法被提出來處理式(7)的最佳化問題(Platt 1999; Joachim 1999; Vapnik 2000; Collobert & Bengio 2001; Trafalis & Ince 2002)，其中 SMO 演算法(Sequential Minimum Optimization)是最常用的技術之一，詳細的資訊請參考 Platt(1999)及 Vapnik(2000)。

## 肆、結合 ICA 與 SVR 預測模式之建構

本研究提出一個結合獨立成份分析與支援向量迴歸之財務時間序列預測模式。此預測模式之建構可以分成利用 ICA 對預測變數去雜訊的前處理，及以 SVR 使用去除雜訊後的預測變數建構預測模式等兩部分。首先在 ICA 部分，在將  $M$  個維度為  $1 \times N$  的預測變數  $\mathbf{x}_i$  組成混合矩陣後，利用 ICA 模式從矩陣  $\mathbf{X}$  中估計出  $M$  個維度為  $1 \times N$  的獨立成份( $\mathbf{y}_i$ )及維度為  $M \times M$  的解混合矩陣  $\mathbf{W}$ 。接著，為找出代表雜訊的獨立成份(IC)，本研究利用 Cheung & Xu (2001)所提出的測試接受法(Testing-and-Acceptance, TnA)對 IC 在趨勢特徵擷取之表現進行評估。TnA 法是基於資料還原(Data reconstruction)的排序方法(ICA 中還原的公式如下面式(8)所示)，使用 Relative hamming distance(RHD)還原誤差值為指標，衡量還原後的序列資料與原始序列間的趨勢相似程度。RHD 值越小代表兩個序列資料之趨勢越相似，當趨勢完全一致時其值為 0，完全相反則為 4 (RHD 之公式請參閱 Cheung & Xu 2001)。

使用 TnA 法排序的步驟為先計算每一個獨立成份各自進行還原後與原始資料之間的 RHD 值，有最小 RHD 值的獨立成份是最能夠捕捉到原始序列資料中主要趨勢結

構的獨立成份，因此被排序為第一個；接著將排序第一的獨立成份與其餘的獨立成份兩兩組合成一組後，再次進行還原並計算 RHD 值，同樣有最小 RHD 值的獨立成份組合是表現最佳的組合，組合內排序第一的獨立成份外的另一個獨立成份將是排序第二的獨立成份，以此類推，重複此組合還原及比較選擇的程序，一次排序一個獨立成份，直到所有的獨立成份被排序完成。

舉例說明，圖 2 為 4 個維度均為  $1 \times 794$  的財務時間序列訊號，將其組成維度為  $4 \times 794$  的混合訊號矩陣後，利用 ICA 模式於此混合訊號矩陣，可以估計 4 個維度為  $1 \times 794$  的獨立成份，此 4 個獨立成份(IC1-IC4)的波形圖列示於圖 3 中。接下來利用 TnA 法來對 4 個獨立成份的特徵擷取表現進行排序，表 1 為 TnA 法排序過程的 RHD 還原誤差值，排序過程為先對 4 個 IC 各別還原，由表 1 中可知，4 個 IC 中以 IC1 的 RHD 值 1.7531 最低，因此被排序為第一個。接著將 IC1 與其餘的獨立成份兩兩組合成一組後再進行還原，同樣由表 1 中可以發現，各組合的還原結果中以 IC1 與 IC3 組合的 RHD 值 1.4346 最低，因此除了 IC1 為排序第一外，IC3 將排序第二。重複此一操作，最後之排序結果依序為 IC1、IC3、IC2 及 IC4，其中以 IC1 最能夠捕捉原始資料結構中的趨勢特徵，而 IC4 的特徵擷取表現為最差。表 1 中必須備註說明的是，在 4 個獨立成份的排序中，決定了排名前三的獨立成份後，剩下的必定排名最後，因此第 4 次之比較在實際操作中是不需要的。

在獨立成份排序後，我們將選擇排序在前的數個獨立成份，代表預測變數資料中的主要趨勢結構，其餘的獨立成份則代表預測變數資料中的雜訊。在保留代表主要趨勢結構的獨立成份及去除(將其值設為 0)代表雜訊的獨立成份並且還原後，還原的預測變數資料中將存在最少的雜訊資訊。還原的式子如下：

$$\hat{\mathbf{X}} = \sum_{i=1, i \neq k}^M \mathbf{a}_i \mathbf{y}_i, 1 \leq k \leq M \quad (8)$$

其中  $\hat{\mathbf{X}} = [\hat{\mathbf{x}}_1, \hat{\mathbf{x}}_2, \dots, \hat{\mathbf{x}}_M]^T$  為維度  $M \times N$  的還原預測變數矩陣， $\hat{\mathbf{x}}_i$  是還原後的預測變數， $\mathbf{a}_i$  為混合矩陣  $\mathbf{A}$  的第  $i$  個行向量(Column vector)， $\mathbf{A} = \mathbf{W}^{-1}$ ， $\mathbf{y}_i$  是第  $i$  個獨立成份， $k$  是代表雜訊的獨立成份。

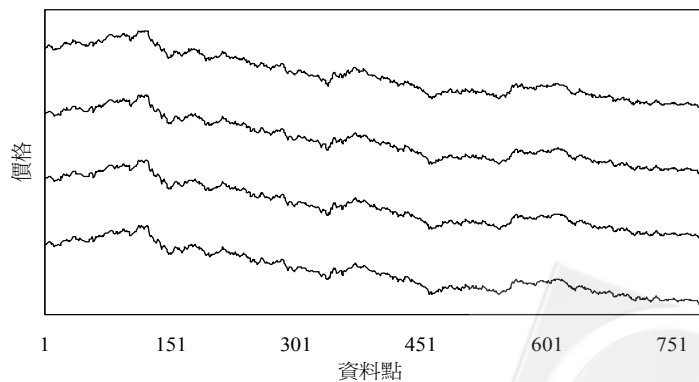


圖 2：4 個維度均為  $1 \times 794$  的財務時間序列訊號

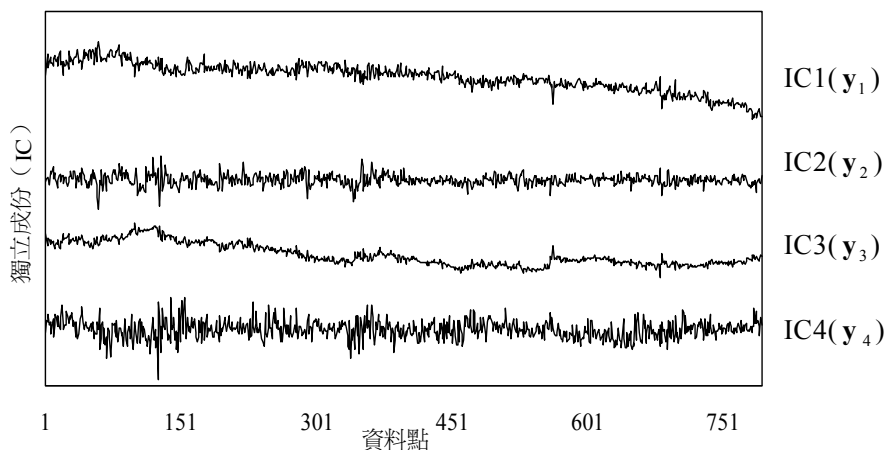


圖 3：圖 2 之財務時間序列訊號的 4 個獨立成份

在還原的過程中，重要的是選擇適當的獨立成份個數來代表預測變數資料中的主要趨勢結構，以達到最佳的雜訊去除效果，因選擇過多獨立成份會使得雜訊的濾除效果不佳，反之選擇太少則可能會喪失預測變數的趨勢結構。本研究根據依序還原排序第一的獨立成份、前兩個獨立成份及所有獨立成份的 RHD 值之變化，選擇適當的獨立成份個數。由於還原的獨立成份越多，還原的訊號會與原始訊號越相似，所以此一 RHD 值的變化是將是由大到小的排序。而選擇的準則為若 RHD 值在某個獨立成份加入還原組合後之下降幅度不大，表示此一獨立成份包含較少的趨勢特徵資訊，加入還原後並無法使還原後訊號更近似原始訊號，所以此一及排序在其之後的獨立成份將被用來代表原始資料中的雜訊。反之，若 RHD 值在某個獨立成份加入還原組合後有顯著的下降，則此一獨立成份代表預測變數中的主要趨勢結構。除此之外，若有  $M$  個獨立成份，則最多選擇  $M-1$  個獨立成份代表預測變數中的主要趨勢結構，或是最少會選擇一個獨立成份(亦即排序最後獨立成份)代表預測變數中的雜訊，因為財務時間序列資料存在雜訊(Deboeck 1994; Yaser & Atiya 1996)。

表 1：測試接受法排序過程的 RHD 還原誤差值

比較次數	還原之獨立成份 (IC)	RHD 還原誤差值
第 1 次	IC1	1.7531
	IC2	1.8528
	IC3	1.9779
	IC4	1.9044
第 2 次	IC1、IC2	1.7137
	IC1、IC3	1.3346
	IC1、IC4	1.7373
第 3 次	IC1、IC3、IC2	0.7662
	IC1、IC3、IC4	1.4294
第 4 次	IC1、IC3、IC2、IC4	0

以圖 3 中的 4 個獨立成份為例，圖 4 為還原排序第一的獨立成份(即 IC1)、前二個獨立成份(即 IC1 及 IC3)到前四個獨立成份(即 IC1-IC4)的 RHD 值，這些 RHD 值可以從 TnA 法的排序過程中(即表 1 中)整理而得。由圖 4 中可知 RHD 值從還原第一個到還原前 4 個 IC 均明顯下降，雖然如此，由於最少會選擇排序最後的獨立成份代表預測變數資料中的雜訊，因此前三個獨立成份(即 IC1、IC3 及 IC2)將被用來代表預測變數中的主要趨勢結構，排序最後的 IC4 則代表預測變數中的雜訊部分。為說明代表主要趨勢之 IC 與代表雜訊之 IC 在還原結果上之差異，圖 5 為以圖 2 中  $x_1$  前 201 個資料點為例，還原代表主要趨勢結構的 IC(即 IC1、IC3 及 IC2)以及代表雜訊的 IC(即 IC4)之結果。由圖 5 中可以發現，還原 IC1、IC3 及 IC2 的結果會與原始訊號(即  $x_1$ )在主要趨勢及結構(Shape)上非常近似，而代表雜訊的 IC 之還原結果除了在趨勢上與原始訊號不符外，並且在整體結構的資訊貢獻度上不佳。

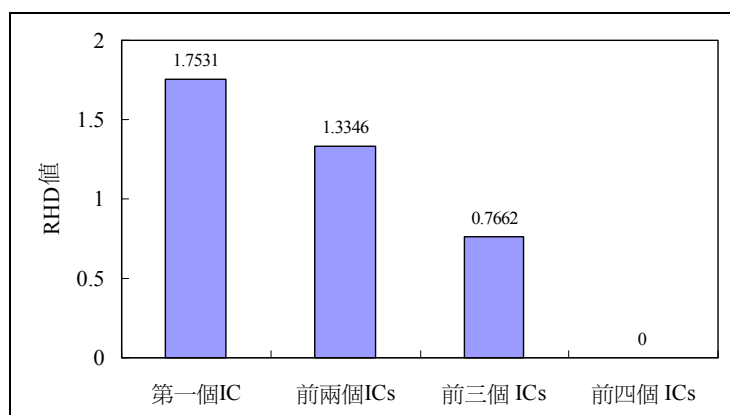


圖 4：獨立成份依排序前後組合並還原的 RHD 值

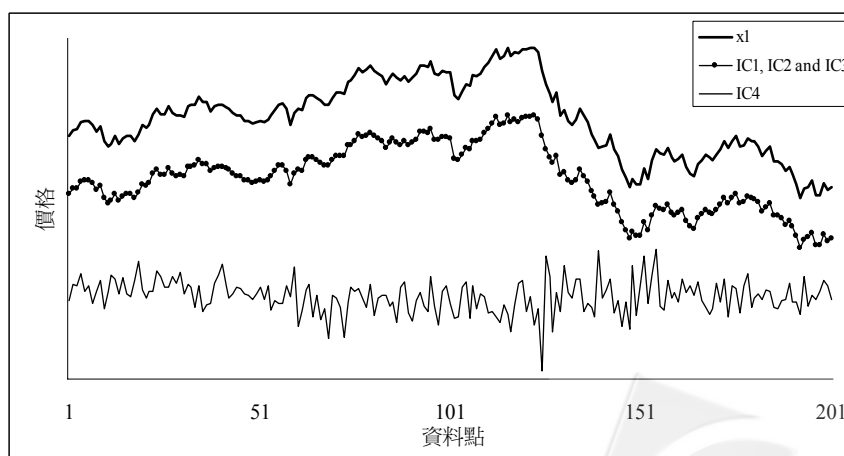


圖 5：還原代表主要趨勢結構的 IC (即 IC1、IC3 及 IC2) 以及代表雜訊的 IC (即 IC4) 之結果-以圖 2 中  $x_1$  前 201 個資料點為例

在利用 ICA 模式得到雜訊濾除之預測變數後，接著 SVR 將以前述步驟所得之變數建立預測模式。在使用 SVR 時首先需進行核心函數的選擇，由於 RBF 函數是 SVR 中常用的核心函數之一(Lin *et al.* 2003; Cherkassky & Ma 2004)，本研究將使用 RBF 函數為 SVR 中的核心函數。接著，由於 SVR 的預測品質主要受其模式中的調整常數  $C$  及損失函數  $\varepsilon$  的影響(Lin *et al.* 2003; Cherkassky & Ma 2004)，因此使用 SVR 時最主要的議題是如何選擇適當的調整常數  $C$  及損失函數  $\varepsilon$  的值以得到良好的預測結果。

在調整常數  $C$  及損失函數  $\varepsilon$  的決定上，目前並無一般性的方法可以使用，通常會依問題的特性直接判斷或使用試誤法決定最適的參數值。本研究參考 Lin *et al.*(2003) 所提出的網格搜尋法(Grid search)決定參數  $C$  及  $\varepsilon$  的數值，其建議在 SVR 的參數選擇過程中，使用指數成長的序列搜尋能得到好的參數結果，(例如， $C = 2^{-5}, 2^{-3}, 2^{-1}, \dots, 2^{15}$ )。但由於網格搜尋法建議的試誤次數過多並且沒有範圍，為了減少決定參數的搜尋時間並得到較佳的結果，本研究將先利用 Cherkassky & Ma(2004)所建議之方法得到 SVR 之參數  $C$  及  $\varepsilon$  的建議值(即參數組合)後，再根據此建議值做適當次數的指數序列搜尋，並以能產生最小均方誤差(Mean Square Error, MSE)之預測結果的參數組合為最適的組合。此方法將可去除單一參數組合建議值的風險，並降低試誤法計算次數過多的缺點。

## 伍、實證結果

### 一、資料收集及績效衡量準則

在實證部分，為驗證本研究所提之結合 ICA 與 SVR 之兩階段預測模式(簡稱兩階段模式)之預測能力，將以日經 225 指數為研究標的，對日經 225 之現貨開盤指數進行預測。並且兩階段預測模式之預測結果將與單純使用 SVR 模式及隨機漫步方法的預測結果進行比較，以驗證所提方法的有效性，其中本研究所使用的隨機漫步模式為利用標的物前一日價格( $t-1$ )，預測當日的標的物價格( $t$ )。而在兩階段模式及 SVR 模式的預測變數方面，由於期貨市場反映市場資訊的速度較現貨為快，並且市場的收盤成交價會率先反映出盤後訊息(Lee & Chen 2002; Lee & Chiu 2002)，因此本研究選擇三個以日經 225 指數為標的期貨指數，分別為大阪、新加坡和芝加哥商品交易所之日經 225 期貨，及前一日現貨收盤指數為 4 個預測變數。

在實證資料部分，蒐集的資料期間為 1999 年 10 月 4 日至 2004 年 9 月 30 日共五年 1144 個交易日，實證資料之走勢圖如圖 6 所示。實證資料中從 1999 年 10 月 4 日至 2003 年 3 月 27 日共 794 筆資料將作為訓練數據集，另保留 2003 年 3 月 28 日至 2004 年 9 月 30 日共 350 筆資料作為測試資料集。在預測績效的衡量準則方面，本研究以均方根誤差(Root Mean Square Error, RMSE)、標準化平均平方誤差(Normalized mean square error, NMSE)、平均絕對離差(Mean Absolute Difference, MAD)、方向對稱性(Directional Symmetry, DS)、正確上漲趨勢率(Correct Up trend, CP)及正確下跌趨勢率

(Correct Down trend, CD)等 6 個常用之指標評估預測模式之預測結果準確度，這些指標的定義列示於表 2 中。在這些指標中，RMSE、NMSE 及 MAD 是用來衡量預測值與實際值的誤差(Deviation)，此三個指標均是越小代表預測值與實際值越接近。DS、CP 及 CD 則都是用來衡量預測的漲跌方向與實際方向的準確率，其中 DS 衡量的是整體(亦即不管上漲或下跌)的預測準確率，而 CP 與 CD 則是分別用來評估上漲及下跌時的預測準確率。

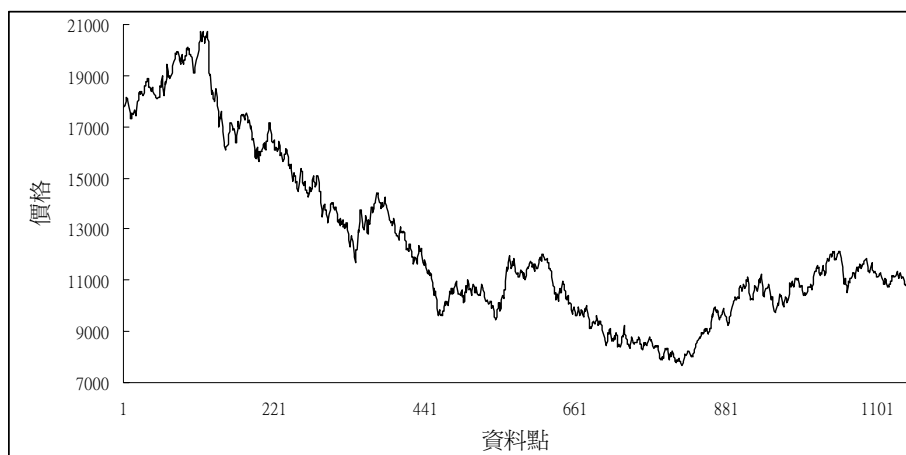


圖 6：日經 225 開盤指數走勢圖

表 2：績效衡量指標及其公式

指標	公式
RMSE	$RMSE = \sqrt{\frac{\sum_{i=1}^N (T_i - A_i)^2}{N}}$
NMSE	$NMSE = 1/(\sigma^2 N) * \sum_{i=1}^N (T_i - A_i)^2$ ，其中 $\sigma^2 = 1/(N-1) * \sum_{i=1}^N (A_i - \bar{A})^2$
MAD	$MAD = \frac{\sum_{i=1}^N  T_i - A_i }{N}$
DS	$DS = \frac{100}{N} * \sum_{i=1}^N d_i$ ，其中 $d_i = \begin{cases} 1 & (A_i - A_{i-1})(T_i - T_{i-1}) \geq 0 \\ 0 & otherwise \end{cases}$
CP	$CP = \frac{100}{N} * \sum_{i=1}^N d_i$ ，其中 $d_i = \begin{cases} 1 & (T_i - T_{i-1}) > 0 \text{ and } (A_i - A_{i-1})(T_i - T_{i-1}) \geq 0 \\ 0 & otherwise \end{cases}$
CD	$CD = \frac{100}{N} * \sum_{i=1}^N d_i$ ，其中 $d_i = \begin{cases} 1 & (T_i - T_{i-1}) < 0 \text{ and } (A_i - A_{i-1})(T_i - T_{i-1}) \geq 0 \\ 0 & otherwise \end{cases}$

\*上面所有式子中， $T$  代表實際值， $A$  代表預測值， $N$  代表資料筆數。

## 二、實驗結果

首先在單純使用 SVR 模式方面，我們使用 Chang & Lin (2001)所發展的 LIBSVM 建立預測模式，預測標的物的數值(即日經 225 開盤價)將先經過尺度化(Scaling)的處理。而在 SVR 的參數 C 的  $\epsilon$  設定上，利用 Cherkassky & Ma (2004)所建議之方法得到  $\epsilon=0.0019$ ， $C=1.25$ ，由於  $\epsilon$  接近  $2^{-9}$ ， $C$  接近  $2^1$ ，因此我們將以  $(\epsilon=2^{-9}, C=2^1)$  的參數組合為起點進行網格搜尋，以試誤出最佳的參數組合產生最好的預測結果。表 3 為單純以 SVR 工具在不同參數組合下的訓練及測試結果。由表 3 中可知當  $\epsilon=2^{-9}, C=2^1$  時有最小的測試資料 MSE 值，是單純 SVR 模式的最佳參數組合。

而在兩階段預測模式方面，由於所使用的 4 個預測變數之資料即為圖 3 中的 4 個財務時間序列訊號，因此利用 ICA 模式濾除預測變數中雜訊等同於濾除圖 3 中財務時間序列訊號的雜訊，其過程與結果已說明於第三節中，在此就不贅述。而在 SVR 預測模式的建立上， $(\epsilon=2^{-9}, C=2^1)$  的參數組合仍是網格搜尋的起始點。表 4 為兩階段預測模式在不同參數組合下的預測結果。由表 4 中可以觀察到， $\epsilon=2^{-7}, C=2^3$  有最小的測試資料 MSE 值，是兩階段預測模式的最佳參數組合。

表 5 為隨機漫步、單純使用 SVR 及兩階段模式三種預測模式的預測結果彙總表。由表 5 可知，所提兩階段模式的 RMSE、NMSE 及 MAD 值分別為 56.76、0.0026 及 40.86，均小於 SVR 模式及隨機漫步模式，代表兩階段模式所預測的數值與實際值最接近，預測誤差最小。除此之外，相對 SVR 模式及隨機漫步模式，兩階段模式亦有最高的 DS、CD 及 CP 值，分別為 87.53%、88.77%及 86.09%，代表其在股價趨勢的預測上亦較其他兩個方法準確。而為了進一步解釋預測模型間的差異，圖 6 為 3 個預測模型在測試樣本中的最後 50 筆之預測值與實際值比較圖。由圖 6 中可以觀察到，兩階段模式與單純使用 SVR 模式在趨勢的捕捉上均較隨機漫步為佳，但兩階段模式在一些轉折點的精確度又比 SVR 模式為佳(如樣本點 307, 322 與 349)，且 SVR 模式相較而言易高估預測值。備註說明的是雖然圖 6 只是測試樣本中的 50 筆資料，但用整個測試資料集來觀察仍可獲得相同的結論。整體而言，本研究所提之結合 ICA 與 SVR 之兩階段預測模式，相較於 SVR 模式及隨機漫步模式，能提供最小的預測誤差以及最高的趨勢預測準確率。





表 3：SVR 模式在不同參數組合下的預測結果

$\varepsilon$	C	訓練資料 MSE	測試資料 MSE
$2^{-13}$	$2^{-1}$	0.0000319491	0.0000248156
	$2^1$	0.0000291671	0.0000214102
	$2^3$	0.0000279661	0.0000240622
$2^{-11}$	$2^{-1}$	0.0000319861	0.0000247999
	$2^1$	0.0000291736	0.0000214540
	$2^3$	0.0000280469	0.0000239275
$2^{-9}$	$2^{-1}$	0.0000323087	0.0000273171
	$2^1$	0.0000294424	0.0000230685
	$2^3$	0.0000282242	0.0000221947
$2^{-7}$	$2^{-1}$	0.0000426816	0.0000335855
	$2^1$	0.0000346765	0.0000342352
	$2^3$	0.000031965	0.0000290501

表 4：兩階段預測模式在不同參數組合下的預測結果

$\varepsilon$	C	訓練資料 MSE	測試資料 MSE
$2^{-11}$	$2^{-1}$	0.0000547405	0.0000279631
	$2^1$	0.0000535331	0.0000304249
	$2^3$	0.0000528961	0.0000396468
	$2^5$	0.0000525349	0.0000469127
$2^{-9}$	$2^{-1}$	0.0000546991	0.0000258684
	$2^1$	0.0000537528	0.0000284480
	$2^3$	0.0000529974	0.0000344395
	$2^5$	0.0000526340	0.0000400822
$2^{-7}$	$2^{-1}$	0.0000554437	0.0000202124
	$2^1$	0.0000544119	0.0000198950
	$2^3$	0.0000538615	0.0000191902
	$2^5$	0.0000533897	0.0000202790
$2^5$	$2^{-1}$	0.0001088290	0.0001002030
	$2^1$	0.0001415800	0.0001092080
	$2^3$	0.0001368980	0.0000906426
	$2^5$	0.0001514900	0.0001442630

表 5：隨機漫步、SVR 及兩階段模式於 Nikkei225 開盤價的預測結果彙總表

預測模型 \ 指標	RMSE	NMSE	MAD	DS	CD	CP
隨機漫步模式	137.85	0.1431	105.77	50.43%	54.26%	45.96%
單純使用 SVR 模式	60.53	0.0285	43.71	83.67%	86.17%	80.75%
結合 ICA 與 SVR 模式	56.76	0.0226	40.86	87.53%	88.77%	86.09%

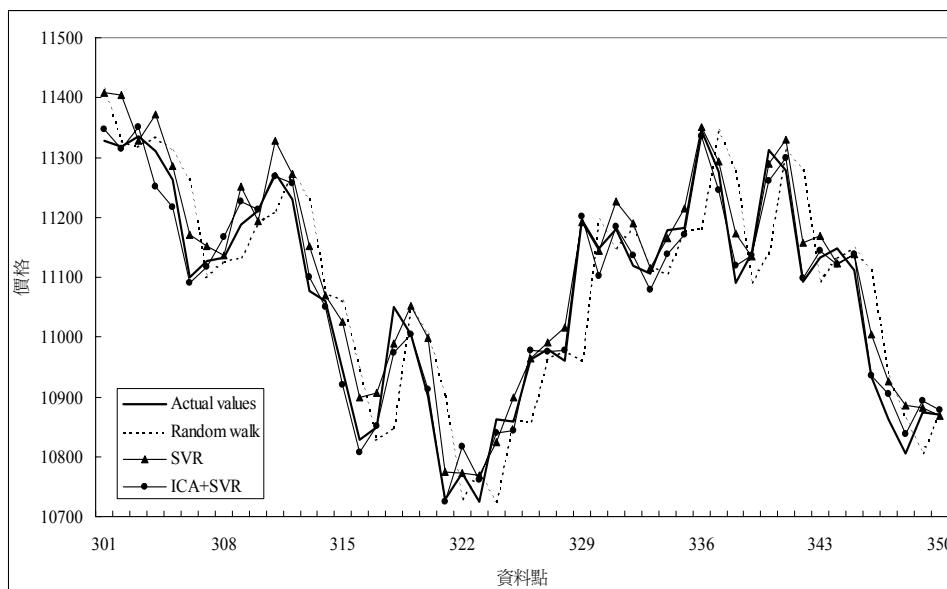


圖 6：3 個預測模式之預測值與實際值比較圖-以測試樣本中的最後 50 筆為例

為進一步呈現所提兩階段預測模式之預測效果，本研究另以台灣加權股價指數 (Taiwan stock exchange capitalization weighted stock index, TAIEX) 為標的，對 TAIEX 之現貨收盤價進行預測。採用的資料期間為 2003 年 1 月 2 日至 2006 年 2 月 27 日共四年 781 個交易日，其中在訓練樣本的部分為 2003 年 1 月 2 日至 2005 年 3 月 17 日共 546 筆資料 (佔資料期間 69.9%)，另保留 2005 年 3 月 18 日至 2006 年 2 月 27 日共 235 筆資料 (佔資料期間 30.1%) 作為測試樣本，圖 7 為資料期間的 TAIEX 收盤指數走勢圖。在預測變數部分，由於技術指標及以該現貨市場為標的之期貨指數是股價預測時常用的預測變數 (Balachandher *et al.* 2002; Leigh *et al.* 2005)，因此本研究使用現貨指數之前一日最高價、前一日最低價、前一日成交值、前一日成交量、現貨當日開盤價、六日相對強弱指標 (6 days-relative strength indicator, RSI6)、10 日指數點成交值指標 (Total amount weight stock price Index, TAPI)、新加坡交易所摩台指開盤期貨指數 (SGX-DT MSCI Taiwan opening index futures) 以及台灣期貨交易所台指期開盤期貨指數 (TAIFEX TAIEX Taiwan opening index futures) 等 9 個預測變數建構 TAIEX 收盤價之預測模型。

表 6 為結合 ICA 與 SVR 模式、單純使用 SVR 與隨機漫步模式於 TAIEX 收盤價的預測結果。由表 6 中各評估指標之值可以發現，不論在離差衡量指標或趨勢正確率指標方面，所提結合模式之結果均較單純使用 SVR 與隨機漫步模式為佳，有較低的預測誤差及較高的預測準確度。

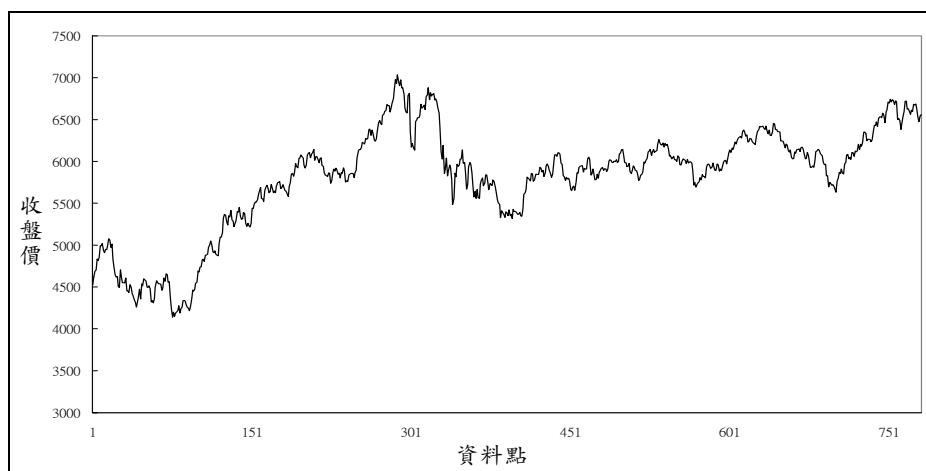


圖 7：TAIEX 收盤價走勢圖

表 6：隨機漫步、SVR 及兩階段模式於 TAIEX 收盤價的預測結果彙總表

預測模型 \ 指標	RMSE	NMSE	MAD	DS	CD	CP
隨機漫步模式	53.21	0.1692	39.88	46.15%	45.22%	47.93%
單純使用 SVR 模式	46.60	0.0330	34.63	55.98%	60.00%	52.10%
結合 ICA 與 SVR 模式	41.09	0.0297	31.70	60.15%	60.44%	62.61%

## 陸、結論

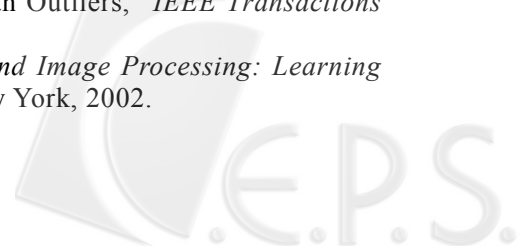
本研究提出一結合獨立成份分析與支援向量迴歸之兩階段財務時間序列預測模式，先利用 ICA 具有將混合訊號分離出個別獨立的來源訊號之能力，從預測變數中估計出獨立成份，並在去除代表雜訊的獨立成份及保留剩餘的獨立成份進行重建時間序列後，SVR 使用濾除雜訊後的預測變數建構財務時間序列預測模式。所提之方法能減少 SVR 在模式架構時因受雜訊影響而產生過度配適或配適不足的問題，並希望提升 SVR 模式的預測準確度。研究中以日經 225 之現貨開盤價格以及台灣加權股價指數之現貨收盤價格預測為實例，將所提方法之預測結果與單純使用 SVR 模式及隨機漫步方法的預測結果作比較。實驗結果顯示，所提之方法不論是在預測誤差或是預測準確度的表現上均較單純使用 SVR 模式及隨機漫步方法模式為佳，驗證本研究所提方法的有效性。在未來研究方向上，可以本研究之研究成果及預測模式擬定投資操作策略，並以投資績效作為預測結果之衡量，進一步增加研究結果之實用性。

## 致謝

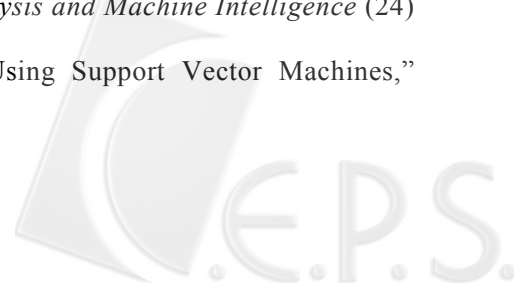
本研究部分承行政院國科會（計畫編號：NSC 95-2221-E-030-002）經費贊助，特此致謝。

## 柒、參考文獻

1. Antoniou, A. and Holmes, P. "Futures Trading, Information and Spot Price Volatility: Evidence for the FTSE-100 Stock Index Futures Contract Using GARCH," *Journal of Banking and Finance* (19) 1995, pp: 117-129.
2. Back, A. and Weigend, A. "Discovering Structure in Finance Using Independent Component Analysis," *Proceeding of 5th International Conference on Neural Networks in Capital Market* 1997, pp: 15-17.
3. Balachandher, K.G., Fauzias, M.N. and Lai, M.M. "An Examination of the Random Walk Model and Technical Trading Rules in the Malaysian Stock Market," *Quarterly Journal of Business & Economics* (41) 2002, pp: 81-104.
4. Barlett, M. and Sejnowski, T.J. "Viewpoint Invariant Face Recognition Using Independent Component Analysis and Attractor Networks," *Advances in Neural Information Processing Systems* (9) 1997, pp: 817-823.
5. Bartlett, M.S., Movellan, J.R. and Sejnowski, T.J. "Face Recognition By Independent Component Analysis," *IEEE Transactions on Neural Networks* (13) 2002, pp: 1450-1464.
6. Beckmann, C.F. and Smith, S.M. "Probabilistic Independent Component Analysis for Functional Magnetic Resonance Imaging," *IEEE Transactions on Medical Imaging* (23) 2004, pp: 137-152.
7. Burbidge, R., Trotter, M., Buxton, B. and Holden, S. "Drug Design by Machines Learning: Support Vector Machines for Pharmaceutical Data Analysis," *Computer & Chemistry* (26) 2001, pp: 5-14.
8. Cao, L.J. "Support Vector Machines Experts for Time Series Forecasting," *Neurocomputing* (51) 2003, pp: 321-339.
9. Cao, L. and Tay, F.E.H. "Financial Forecasting Using Support Vector Machines," *Neural Computing & Applications* (10) 2001, pp: 184-192.
10. Chang, C.C. and Lin, C.J. "LIBSVM: A Library for Support Vector Machines," 2001, *Software available at <http://www.csie.ntu.edu.tw/~cjlin/libsvm>.*
11. Chang, R.F., Wu, W.J., Moon, W.K. and Chen, D.R. "Improvement in Breast Tumor Discrimination by Support Vector Machines and Speckle-Emphasis Texture Analysis," *Ultrasound in Medicine and Biology* (29) 2003, pp: 679-686.
12. Cheung, Y.M. and Xu, L. "Independent Component Ordering in ICA Time Series Analysis," *Neurocomputing* (41) 2001, pp: 145-152.
13. Chuang, C.C., Su, S.F., Jeng, J.T. and Hsiao, C.C. "Robust Support Vector Regression Networks for Function Approximation with Outliers," *IEEE Transactions on Neural Networks* (13) 2002, pp: 1322-1330.
14. Cichocki A. and Amari, S.I. *Adaptive Blind Signal and Image Processing: Learning Algorithms and Applications*, John Wiley & Sons, New York, 2002.



15. Cherkassky, V. and Ma, Y. "Practical Selection of SVM Parameters and Noise Estimation for SVM Regression," *Neural Networks* (17) 2004, pp: 113–126.
16. Collobert, R. and Bengio, S. "SVM-Torch: Support Vector Machines for Large-Scale Regression Problems," *Journal of Machine Learning Research* (1) 2001, pp: 143-160.
17. Comon, P. "Independent Component Analysis: A New Concept?" *Signal Processing* (36) 1994, pp: 287-314.
18. Cover, T.M. and Thomas, J.A. *Elements of Information Theory*, John Wiley and Sons, New York, 1991.
19. David, V. and Sanchez, A. "Frontiers of Research in BSS/ICA," *Neurocomputing* (49) 2002, pp: 7-23.
20. Deboeck, G.J. *Trading on the Edge: Neural, Genetic, and Fuzzy Systems for Chaotic Financial Markets*, Wiley, New York, 1994.
21. Déniz, O., Castrillón, M. and Hernández, M. "Face Recognition Using Independent Component Analysis and Support Vector Machines," *Pattern Recognition Letters* (24) 2003, pp: 2153-2157.
22. Drucker, H., Burges, C.J.C., Kaufman, L., Smola, A. and Vapnik, V.N. "Support Vector Regression Machines," *Advances in Neural Information Processing Systems* (9) 1997, pp: 155.
23. Hyvärinen, A. "Fast and Robust Fixed-Point Algorithms for Independent Component Analysis," *IEEE Transactions on Neural Networks* (10) 1999, pp: 626-634.
24. Hyvärinen, A. and Oja, E. "Independent Component Analysis: Algorithms and Applications," *Neural Networks* (13) 2000, pp: 411-430.
25. Hyvärinen, A., Karhunen, J. and Oja, E. *Independent Component Analysis*, John Wiley & Sons, New York, 2001.
26. Ikeda, S. and Toyama, K. "Independent Component Analysis for Noisy Data-MEG Data Analysis," *Neural Networks* (13) 2000, pp: 1063-1074.
27. James, C.J. and Gibson, O.J. "Temporally Constrained ICA: An Application to Artifact Rejection in Electromagnetic Brain Signal Analysis," *IEEE Transactions on Biomedical Engineering* (50) 2003, pp: 1108-1116.
28. Jang, G.J., Lee, T.W. and Oh, Y.H. "Learning Statistically Efficient Features for Speaker Recognition," *Neurocomputing* (49) 2002, pp: 329-348.
29. Joachim, T. "Making Large-Scale SVM Learning Practical," *Advances in Kernel Methods-Support Vector Learning* (Scholkopf, B., Burges, C.J.C. and Smola, A.J., eds.), MIT Press, Cambridge, MA, 1999.
30. Jung, C. and Boyd, R. "Forecasting UK Stock Prices," *Applied Financial Economics*, (6) 1996, pp: 279-286.
31. Jung, T.P., Makeig, S., McKeown, M.J., Bell, A.J., Lee, T.W. and Sejnowski, T.J. "Imaging Brain Dynamics Using Independent Component Analysis," *Proceedings of the IEEE*, California University, San Diego, La Jolla, CA, 2001, pp: 1107-1122.
32. Karras, D.A. and Mertzios, B.G. "Time Series Modeling of Endocardial Border Motion in Ultrasonic Images Comparing Support Vector Machines, Multilayer Perceptrons and Linear Estimation Technique," *Measurement* (36) 2004, pp: 331-345.
33. Kim, K.I., Jung, K., Park, S.H. and Kim, H.J., "Support Vector Machines for Texture Classification," *IEEE Transactions on Pattern Analysis and Machine Intelligence* (24) 2002, pp: 1542-1550.
34. Kim, K.J. "Financial Time Series Forecasting Using Support Vector Machines," *Neurocomputing* (55) 2003, pp: 307-319.



35. Kim, T.K., Kim, H., Hwang, W. and Kittler, J. "Independent Component Analysis in A Local Facial Residue Space for Face Recognition," *Pattern Recognition* (37) 2004, pp: 1873-1885.
36. Kiviluoto, K. and Oja, E. "Independent Component Analysis for Parallel Financial Time Series," *Proceedings of the Fifth International Conference on Neural Information*, Tokyo, Japan, 1998, pp: 895-898.
37. Koike, A. and Takagi, T. "Prediction of Protein-Protein Interaction Sites Using Support Vector Machines," *Protein Engineering Design & Selection* (17) 2004, pp: 165-173.
38. Kwon, K.Y. and Kish, J.R. "Technical Trading Strategies and Return Predictability: NYSE," *Applied Financial Economics* (12) 2002, pp: 639-653.
39. Li, S., Kwok, J.T., Zhu, H. and Wang, Y. "Texture Classification Using The Support Vector Machines," *Pattern Recognition* (36) 2003, pp: 2883-2893.
40. Lin, Q.H., Zheng, Y.R., Yin F. and Liang, H.L. "Speech Segregation Using Constrained ICA," *Lecture Notes in Computer Science* (3173) 2004, pp: 755-760.
41. Lin C.J., Hsu, C.W. and Chang, C.C. "A Practical Guide to Support Vector Classification," *Technical Report*, Department of Computer Science and Information Engineering, National Taiwan University, 2003.
42. Lee, T.W. *Independent Component Analysis: Theory and Application*, Kluwer Academic Publishers, Boston, 1998.
43. Lee, T.S. and Chen, N.J. "Investigating the Information Content of Non-Cash-Trading Index Futures Using Neural Networks," *Expert Systems with Applications* (22) 2002, pp: 225-234.
44. Lee, T.S. and Chiu, C.C. "Neural Network Forecasting of An Opening Cash Price Index," *International Journal of Systems Science* (33) 2002, pp: 229-237.
45. Lee, T.S., Chen, N.J. and Chiu, C.C. "Forecasting the Opening Cash Price Index Using Gray Forecasting and Neural Networks: Evidence from the SGX-DT MSCI Taiwan Index Futures Contracts," *Computational Intelligence in Economics and Finance* (Wang, P and Chen, S. S., Eds), Springer, 2003.
46. Mohandes, M.A., Halawani, T.O., Rehman, S. and Hussain, A.A. "Support Vector Machines for Wind Speed Prediction," *Renewable Energy* (29) 2004, pp: 939-947.
47. Norinder, U. "Support Vector Machine Models in Drug Design: Applications to Transport Processes and QSAR Using Simplex Optimisations and Variable Selection," *Neurocomputing* (55) 2003, pp: 337-346.
48. Oja, E., Kiviluoto, K. and Malaroiu, S. "Independent Component Analysis for Financial Time Series," *Proceedings of the IEEE 2000 Adaptive Systems for Signal Processing, Communications, and Control Symposium*, Lake Louise, Canada, 2000, pp: 111-116.
49. Pai, P.F. and Lin, C.S. "Using Support Vector Machines in Forecasting Production Values of Machinery Industry in Taiwan," *International Journal of Advanced Manufacturing Technology* (27) 2005, pp: 205-210.
50. Parisi, F. and Vasquez, A. "Simple Technical Trading Rules of Stock Returns: Evidence from 1987 to 1998 in Chile," *Emerging Markets Review* (1) 2000, pp: 152-164.
51. Park, H.M., Jung, H.Y., Lee, T.W. and Lee, S.Y. "On Subband-based Blind Signal Separation for Noisy Speech Recognition," *Electronic Letters* (35) 1999, pp: 2011-2012.

52. Platt, J.C. "Fast Training of Support Vector Machines Using Sequential Minimal Optimization," *Advances in Kernel Methods -- Support Vector Learning* (Scholkopf, B., Burges, C.J.C. and Smola, A.J., eds.), MIT Press, Cambridge, MA, 1999.
53. Shin, K.S., Lee, T.S. and Kim, H.J. "An Application of Support Vector Machines in Bankruptcy Prediction Model," *Expert Systems with Applications* (28) 2005, pp: 127-135.
54. Suykens, J.A.K., De Brabanter, J., Lukas, L. and Vandewalle, J. "Weighted Least Squares Support Vector Machines: Robustness and Sparse Approximation," *Neurocomputing* (48) 2002, pp: 85-105.
55. Tay, F.E.H. and Cao, L. "Application of Support Vector Machines in Financial Time Series Forecasting," *Omega* (29) 2001, pp: 309-317.
56. Thissen, U., Van Brakel, R., De Weijer, A.P., Melssen, W.J. and Buydens, L.M.C. "Using Support Vector Machines for Time Series Prediction," *Chemometrics and Intelligent Laboratory Systems* (69) 2003, pp: 35-49.
57. Trafalis, T.B. and Ince, H. "Benders Decomposition Technique for Support Vector Regression," *Proceedings of the International Joint Conference on Neural Networks (IJCNN)*, Honolulu, 2002, pp: 2767-2772.
58. Vapnik, V.N., Golowich, S. and Smola, A.J. "Support Vector Method for Function Approximation, Regression Estimation, and Signal Processing," *Advances in Neural Information Processing Systems* (Mozer, M., Jordan, M. and Petsche, T., eds), MIT Press, Cambridge, MA, 1997.
59. Vapnik, V.N. *The Nature of Statistical Learning Theory*, Springer, New York, 2000.
60. Vellido, A., Lisboa, P.J.G. and Vaughan, J. "Neural Networks in Business: A Survey of Applications (1992-1998) ," *Expert Systems with Applications* (17) 1999, pp: 51-70.
61. Vigarario, R., Sarela, J., Jousmaki, V., Hamalainen, M. and Oja, E. "Independent Component Approach to the Analysis Of EEG And MEG Recordings," *IEEE Transactions on Biomedical Engineering* (47) 2000, pp: 589-593.
62. Visser, E., Lee, T.W. "Speech Enhancement Using Blind Source Separation and Two-Channel Energy Based Speaker Detection," *Proceedings of 2003 IEEE International Conference on Acoustics, Speech, and Signal*, La Jolla, CA, 2003, pp: 884-887.
63. Yaser, S. and Atiya, A.F. "Introduction to Financial Forecasting," *Applied Intelligence* (6) 1996, pp: 205-213.
64. Zhang, G., Patuwo, B.E. and Hu, M.Y. "Forecasting with Artificial Neural Networks: the State of the Art," *International Journal of Forecasting* (14) 1998, pp: 35-62.

