Paper Title\* **(**use style: ***paper title***)

Subtitle as needed ***(paper subtitle)***

Chi Feng Lin

Information Management. National Central University

National Central University

Taoyuan, Taiwan

j8888888871@yahoo.com.tw

Authors Name/s per 2nd Affiliation (*Author*)

line 1 (of *Affiliation*): dept. name of organization

line 2-name of organization, acronyms acceptable

line 3-City, Country

line 4-e-mail address if desired

*Abstract*—時間序列的預測是一個很廣泛且重要的議題。我們提出了球型複數神經模糊集系統 (SCNFS)來對時間序列進行預測。透過一般的複數神經模糊集系統 (CNFS)，我們可得到一組複數型態的輸出，實數部分和虛數部分可針對不同目標做預測。本篇的模型對原本的CNFS中的複數模糊集 (CFSs)進行改良，歸屬程度仍然是在複數型態的值，但是能擁有多組輸出，即可同時針對兩個以上目標預測。模型設計上，前鑑部使用高斯型態的球型複數神經模糊集 (SCFSs)，後鑑部則使用Takagi-Sugeno的線性函式，前鑑部和後鑑部透過箭靶層 (Aim Object)結合，使得兩者有別於傳統的IF-THEN規則，數目可以不一樣。此外，為了最佳化模型的預測結果，我們前鑑部使用Particle Swarm Optimization (PSO)進行參數優化，後鑑部則使用Recursive Least Squares Estimator (RLSE)進行參數優化。最後我們將透過三個實驗與不同的方法做效能比較。

Keywords—Sphere complex neurofuzzy system (SCNFS); complex fuzzy set (CFS); complex neurofuzzy system (CNFS), time-series forecasting.

# Introduction

人工智慧在現實中有很多的應用，像是財經方面的預測[1]，醫療急診方面[2]-[3]、入學預測[4]等。上述這些有一部分屬於時間序列的預測，時間序列的預測是一個很重要的議題，因為他在真實世界中的應用範圍非常廣。過去已經有很多學者提出不同的方法來針對時間序列的預測問題，像是ARIMA[5][6]、模糊理論、神經網路運算、神經模糊混合系統等。其中，最常被提出的就是神經模糊混合系統[7]-[10]。

神經模糊混合系統 (NFSs)一直是被廣泛研究的模型，其中，類神經網路系統有所謂的IF-THEN規則，這些規則如同我們人類的經驗法則。通常可以將IF-THEN的規則結合模糊理論，使整體的架構更有彈性，我們稱之為神經模糊混合系統。如上述所提，神經模糊混合系統的特性使他對時間序列的預測有著不錯的效果。所以現在有關時間預測這方面的研究，大多採用類神經網路為模型架構。

本實驗是採用神經模糊混合系統的作法，將整體模型效仿IF-THEN規則，建造成多層神經元架構。為了使模型彈性增加，有別於傳統的IF-THEN規則方法，本實驗在前鑑部與後鑑部間，採用箭靶層 (Aim Object Layer)連接，使得前鑑部與後鑑部個數得以不相同。在模型實作方面，我們將模糊理論結合類神經網路系統，形成神經模糊混合系統，前鑑部使用高斯型態的球型複數神經模糊集，後鑑部則使用Takagi-Sugeno的線性函式[17]，前鑑部和後鑑部則透過箭靶層結合。透過此模型和機器學習我們預期對時間序列的預測可以更加精準。

關於模糊集，在1965年由Zadeh學者，最先提出模糊集的概念[11]，使資料可以透過某個函式，得到介於0到1之間的歸屬程度 (Membership degree)。之後在2002年，有另一篇研究提出了複數模糊集 (CFSs)[12]的概念，歸屬程度可以透過函式得到一個複數型態的值，這使得歸屬程度可以呈現在一個半徑為1的複數單位圓盤中。這個概念使原本能表示的歸屬程度更加的豐富。一般來說我們可以透過複數型神經模糊集系統 (CNFS)[13][14]，得到一組複數型態的輸出，而實數部分和虛數部分可針對不同目標做預測，所以可針對兩個不同目標。目前兩個目標的預測已經有很多的研究產出[6][13][14]。而為了同時對更多的目標做預測，本篇論文將原本複數型神經模糊集系統進行改良，將原本使用的複數模糊集 (CFSs)改成球型複數模糊集 (SCFSs)，其中歸屬程度仍然是複數型態，但能呈現在3-D的立體空間中，能擁有更多組複數型態的輸出，意即可以一次對多個目標進行預測。

在本研究中，為了使資料可以有效的被應用，在資料前處理的部分，我們根據將原始資料的30個漲跌值作為特徵，並透過夏農資訊熵 (Shannon Entropy)[15]，去計算他們個別對目標的資訊貢獻量，除此之外，我們透過了多目標特徵選取的概念[16]，算出每個特徵對目標的有效資訊量，以此作為挑選訓練資料的依據。從資料中萃取出最有效的資料，除了降低模型的運算負擔，也能有效的提升預測的效能。最後，機器學習部分，我們使用知名的Particle Swarm Optimization (PSO)[18]和廣為人知的Recursive Least Square Estimator (RLSE)[19]進行參數優化，並將它們整合成一種方法我們稱之為PSO-RLSE方法[20]。我們將前鑑部和後鑑部參數藉由不同的演算法訓練，想透過divide-and-conquer原理，降低搜尋的維度，使模型更容易找到最佳解，提高整體的效能。

# Methodology

## Multi-Target Feature Selection

為了使得資料被有效的應用，本實驗將對資料做前處理，以降低運算成本及提升模型精準度。一開始原始資料標記為，其中，為每組資料的總筆數，為資料的總組數。接著將所有原始資料做一次差分，公式如下:



其中，為資料的總筆數，為第組數據。將差分過後的資料漲跌形成30個特徵值，第一組數據，擁有的特徵值標記為到，而第二組數據的特徵值標記為到，依此類推。透過多目標的特徵選取方法[16]，我們可以從特徵中取得訓練資料。多目標的特徵選取方法與資訊熵有關，一般來說，資訊熵 (Entropy)是指接收的每條消息中包含的資訊的平均量，若資訊的隨機性越高，則資訊熵值會越高。公式如下:

2

其中，為的期望值，為的機率密度，但若的機率密度大於1，則部分會是負數，會影響到整體的期望值，所以我們對公式做了一些更改，更改後的公式如下:

3

4

其中，。

由於我們是特徵的選擇是針對目標，所以我們透過資訊熵的概念，計算每個特徵與目標之間的影響資訊量，公式如下:

5

其中，為特徵對目標的影響資訊量，為特徵為正數以及目標的互資訊，為特徵為負數以及目標的互資訊，公式如下:

6

7

其中，為目標的期望值，為特徵為正數時所對應的目標的期望值，為特徵為負數時所對應的目標的期望值。公式如下:

8

9

10

11

其中，，為特徵為正數時的機率密度，為特徵為正數時所對應的目標的機率密度，為特徵為負數時的機率密度，為特徵為負數時所對應的目標的機率密度。

透過上述影響資訊量的公式，可以得到每個特徵對每個目標的影響資訊量，此時就能依據這些影響資訊量做多目標的特徵選取，步驟如下:

Step1 : 算出第個特徵對第個目標的selection gain標記為，其中，為第個特徵變數，為第個目標變數。selection gain 公式如下:

12

其中，為對的影響資訊量，為對中已存在特徵的冗餘資訊量，為第個已選特徵池。冗餘資訊量公式如下:

13

其中，代表內的特徵個數，為對內的第個特徵變數的影響資訊量，為內的第個特徵變數對的影響資訊量。經過上述計算若大於0，則將特徵加入第個已選特徵池 中。

Step2 :無論重疊與否，將所有已選特徵池中出現過的特徵變數記錄下來，儲存成，，其中，是中第個特徵變數。計算每個特徵出現在所有SP的次數，標記為。

Step3 :透過即可計算覆蓋率，公式如下:

14

計算的平均標記為。

Step4 :累加每個SP裡，特徵的selection gain:

15

計算的平均標記為。

Step5 :根據累加後的資訊增益量和覆蓋率，計算出特徵的有效貢獻量:

16

Step6 :測試中所有的特徵變數，若，則將累加。

Step7 :設定上下界，標記為和，透過上下界找出，，表示最後選取的特徵數目。本研究中所有實驗皆設定為4，皆設定為2。若介於上下界間，則將設定成;若小於下界，則將設定成;若大於上界則將設定成。

Step8 :將排序，並選取前個特徵變數加入最後的特徵池(FP)中，當作多目標的特徵挑選結果。

## Structure Learning

結構化學習是為了將訓練資料可以更有邏輯的應用到模型建造中。經過多目標特徵選取後可以得到一些被挑選過的特徵，我們把這些特徵當作模型的訓練資料，標記為，為輸入維度的數量。在本研究中，這些不同輸入維度的訓練資料，用MATLAB裡所提供的subclust()函式做分群，每一個輸入維度係數皆用0.3。並將分群後的群中心配合每個維度的標準差形成高斯型態的模糊集，高斯函數的公式如下:

17

其中，為輸入變數，和為群中心和標準差的參數。基於各個輸入維度的模糊集，我們可以得到個前鑑部，如下:

*Premise :*

*IF*  18

其中，為第個輸入的語意變數;為第個前鑑部中第個輸入語意變數的模糊集;為第個輸入變數，。

為了模型的運算效率，我們透過啟動強度來篩選前鑑部，降低前鑑部的數量。步驟如下:

Step1 :從各個輸入維度與模糊集可以得到每個前艦部的啟動強度，以第個前鑑部為例，啟動強度如下:

19

其中，為第個輸入維度的第筆資料;為第個前鑑部中第個輸入維度的模糊集。

Step2 :將每個前鑑部的啟動強度累加標記為，公式如下:

20

其中，為資料總筆數。計算平均值標記為，標準差標記為。

Step3 :查看每個前鑑部，若，則將累加。設定上下界，標記為和，透過上下界找出，表示最後選取的前鑑部數目。本研究中所有實驗皆設定為15，皆設定為4。若介於上下界間，則將設定成;若小於下界，則將設定成;若大於上界則將設定成。

Step4 :將排序，並保留前個前鑑部，當作之後模型運作的前鑑部層神經元。

假設目前訓練資料集合，標記為，為資料總筆數。將各個輸入維度的資料併在一起，用MATLAB裡所提供的subclust() 函式，係數設定為0.3做分群得到群數。決定了群數之後對目標用fuzzy c-mean分群，可以得到個群中心以及標準差，第個輸出的第群中心標記為}，第個輸出的第群標準差標記為{, }。用得到的群中心以及標準差製作箭靶，每個箭靶後面連接著一個後鑑部，後鑑部為T-S function，T-S function公式如下:

21

其中，{}是第個後鑑部的參數，是第個輸入。

箭靶是用以承接前鑑部的輸出，但是本研究中前鑑部使用的是球型複數模糊集，故輸出會是在複數單位圓盤中，所以箭靶要進行轉換，以確保箭靶層輸出也是複數型態。箭靶轉換公式如下:

22a

22b

22c

其中，為第個輸出第個前鑑部射在第個箭靶的值;。為第個箭靶轉換後的中心，為第個箭靶轉換後的靶寬，公式如下:

23a

23b

23c

24a

24b

24c

其中，為的變數;為的變數;為第個目標的平均;為第個目標的標準差。

## Model Structure and I/O Relationship

本實驗的模型為一個六層的神經網路。訓練資料集合標記為，為資料總筆數，是的輸入向量，為輸入維度數量;為的目標向量，為複數型態目標的數量。透過模型可以得到輸出。

Layer 1:這層稱之為輸入層，是將原始資料做差分後，並透過多目標特徵選取，將最後挑出的特徵當作訓練資料，我們將時間序列第個點的輸入向量標記為:



Layer 2:這層為球型複數模糊集合層，透過前面結構學習的分群，可在不同維度上建構數個模糊集，而每個不同維度的輸入都可經由模糊集得到歸屬程度。傳統高斯模糊集只能得到一組實數型的歸屬程度，而透過球型複數模糊集合可得到多組複數型態的歸屬程度，且會落在半徑為1的單位複數圓球圖(1)內。不同的歸屬程度可以給不同的模型輸出做應用，以達到多目標預測的效果，透過球型複數模糊集的概念，從中可得到:







其中，為高斯函數(17);;。藉由球型的概念，可得到至少四組的複數型態歸屬程度，包含了保維法得到的歸屬程度:



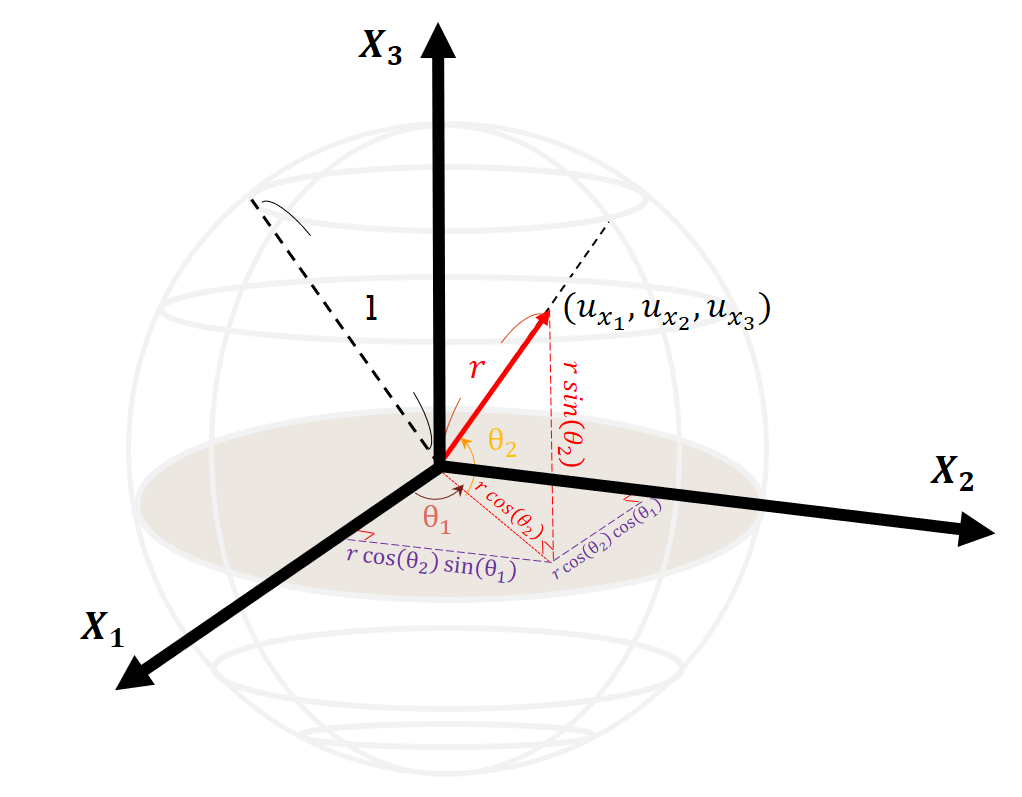
以及降維法得到的歸屬程度:







其中，。



圖(1) 單位複數圓球

Layer 3:此層稱之為前鑑部層，經過前面的結構學習，我們可以篩選出個對模型較有用的前鑑部，標記為:

*Premise :*

*IF*  

其中，為第個輸入的語意變數;為第個前鑑部中第個輸入語意變數的模糊集;為第個輸入變數，。將每個輸入維度的歸屬程度相乘後，可得到每個前鑑部的啟動強度:



Layer 4:此層稱之為箭靶層，我們藉由箭靶層連接前鑑部以及後鑑部使得兩者可以擁有不一樣個數。但由於前鑑部使用的是球型複數模糊集，故輸出會是在複數單位圓盤中，所以箭靶要透過公式(25)-(30)進行轉換，以確保箭靶層輸出也是複數型態。此層的輸出如下:



其中，為第個輸出第個前鑑部射在第個箭靶的值;。經過此層將會輸出個以對應個後鑑部。

Layer 5:此層稱之為後鑑部層，經過此層的運算可以得到個模型輸出，公式如下:





其中，{}是第個後鑑部的參數。

38a

38b

38c

38d

38e

38f

Layer 6:此層稱之為輸出層，將上一層得到的個模型輸出結合，即為我們的模型輸出:



## Parameter Learning

根據divide-and-conquer的概念，我們將使用不同的機器學習演算法，對各層的參數作優化，以便更容易找到最佳解。對於前鑑部的參數優化，我們使用知名的PSO演算法[18]學習，其原理類似鳥群在尋找食物，每回合透過自身的最佳位置和全群最佳位置調節速度，特性為收斂快速，演算法公式如下:

40

41

其中，為第回合時第個粒子的位置，為第回合時第個粒子的速度，為第回合時第個粒子的最好位置，為第回合時全部粒子中最好的位置，為PSO的參數，、為介於0到1的隨機數。在本實驗中，粒子的位置代表前鑑部的參數，其中包含了每個維度的分群中心、標準差以及、。

在本實驗中RLSE[19]是用來更新後鑑部參數，一般來說LSE問題可以被指定成一個線性的模型，如下:



其中，y是目標;u是模型的輸出;{}是u已知的方程式;{, =1,2,…,m}是我們估計的未知參數，則是整個模型的誤差LSE的問題也可以被寫成矩陣的方式表達，如下:

43a

其中:

43b

43c

43d

43e

是輸入的矩陣，是我們估計的未知參數矩陣，是目標矩陣，是誤差的向量。要最佳化，可透過RLSE的等式[19]運算:

a

b

其中，是的第行，再開始RLSE演算法時，我們會設定為0，則設定為，為單位矩陣。我們cost function使用RMSE，定義如下:



其中，()’為複數的共軛運算;，其中為目標向量，為模型輸出向量。整個PSO-RLSE的運作流程如下:

Step 1:準備訓練資料。

Step 2:用PSO粒子位置計算啟動強度。

Step 3:用RLSE更新後鑑部參數，RLSE算式中的和向量如下:

a

b

c

其中。

Step 4:更新完所有參數後，計算出模型的輸出。

Step 5:計算RMSE，更新PSO的gbest和pbest。

Step 6:對所有粒子重複Step2~Step5，直到回合結束。

# Experimentation

## Example 1—Quadruple Time Series of Daily National Association of Securities Dealers Automated Quotation Composite Index

在這個範例中，我們使用真實世界的時間序列數據來驗證模型的效能。使用的資料為National Association of Securities Dealers Automated Quotation(NASDAQ)每天的開盤和收盤價以及Standard and Poor’s (S&P 500)每天的開盤和收盤價。為了和其他論文比較效能，NASDAQ採用2007年1月3號至2010年12月20號的開盤和收盤價當作第一個複數型態的目標，總共998筆資料，而S&P 500的部分則以2010年12月20號往前取998筆開盤和收盤價當作第二組複數型態的目標。本篇論文所提出的模型可以一次有多個複數型態的輸出，故可以預測多個目標，我們這個範例中使用的第一個複數目標實數部分為NASDAQ的開盤價，複數部分則為NASDAQ的收盤價，第二個目標實數部分為S&P 500的開盤價，複數部分為S&P 500的收盤價。由於本實驗中有做多目標特徵選取[16]，故此範例的原始資料為1029筆，經過一次差分得到1028筆，並以30天漲跌為特徵，最後每個特徵會有998筆資料，經過篩選後的每個特徵前500筆資料為訓練資料，剩下的為測試資料。

表I

範例一模型設定

|  |  |
| --- | --- |
| **Parameter** | **Value** |
| Feature variable as model input | {} |
| Number of fuzzy sets of each input | 2、3 |
| Number of targets (real-valued) | 4 |
| Number of outputs (complex-valued) | 2 |
| Number of premises (before selection) | 9 |
| Type of premises | Sphere complex fuzzy set |
| Number of premises (after selection) | 4 |
| Number of premise parameters | 20 |
| Number of aim object | 3 |
| Type of consequences | Takagi-sugeno |
| Number of consequences | 3 |
| Number of consequence parameters | 9 |

結構學習部分，則將每個被挑選的特徵，用MATLAB提供的subclust()函式分群，係數皆為0.3。並透過前鑑部篩選機制，從原本的9個前鑑部篩選到剩下4個前鑑部。整體模型在結構學習後的參數，如表I所示。PSO-RLSE混合方法的機器學習參數設定，如表II所示。

表II

機器學習參數設定

|  |  |
| --- | --- |
| **PSO** | |
| Swarm size | 64 |
| Iterations | 100 |
|  | 0.8 2.0 2.0 |
|  | Random in [0,1] |
| Initial position | Given by the subtractive clustering algorithm |
| Initial velocity | 0 |
| **RLSE** | |
|  |  |
|  | 20x1 zero vector |
|  | **I** |
| **I** | 20x20 identify matrix |

本次範例的結果將與其他論文所提的方法做比較，像是ANFIS[21]，CNFS-ARIMA[6]，RBF network[22]和SVR[23][24]。其中除了SVR之外的模型，皆可以同時預測兩個實數目標。所以我們將使用模型的第一組複數型態的輸出與其他論文做比較，結果如表III所示。模型的機器學習曲線，如圖(2)所示;目標與模型輸出的結果，如圖(3)所示。

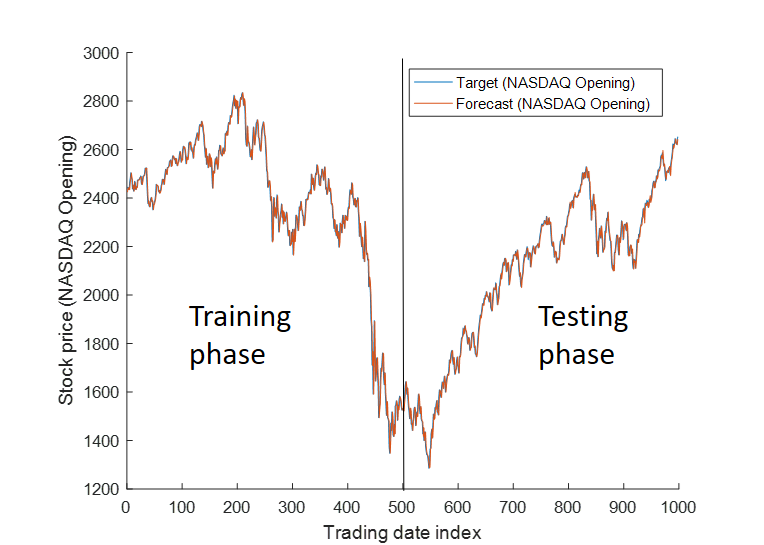


圖(2) 範例一學習曲線

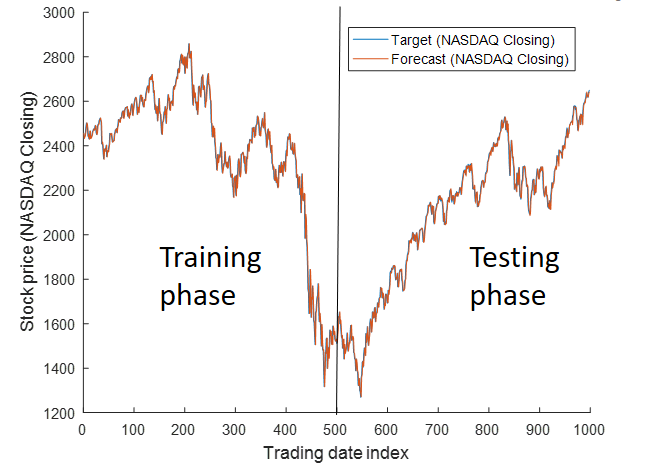
表III

NASDAQ 雙目標預測效能比較表(RMSE)

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Method** | **RMSE** | | | | |
| Training phase | |  | Testing phase | |
| Opening index | Closing index |  | Opening index | Closing index |
| SVR (two models, each with single output) | 35.18 | 35.24 |  | 37.23 | 40.24 |
| ANFIS (two models, each with single output) | 37.83 | 38.66 |  | 38.80 | 42.36 |
| ANFIS (one model with two outputs) | 62.75 | 71.51 |  | 72.52 | 85.08 |
| RBF (two models, each with single output) | 37.59 | 33.89 |  | 37.52 | 44.08 |
| RBF (one model with two outputs) | 178.57 | 179.87 |  | 261.37 | 258.89 |
| CNFS(5)-ARIMA (one model with two outputs) | 21.56 | 20.81 |  | 32.52 | 33.70 |
| SCNFS(proposed) | 38.58 | 38.59 |  | 27.79 | 27.83 |



(a)



(b)

圖(3) NASDAQ的實際數值和模型輸出(a)每日開盤價(b)每日收盤價

表 IV

十次測試效能表(實驗一)

|  |  |  |
| --- | --- | --- |
|  | Performance(RMSE) | |
| Triers | Opening index | Closing index |
| 1 | 29.56 | 28.49 |
| 2 | 130.08 | 109.50 |
| 3 | 27.79 | 27.83 |
| 4 | 30.52 | 30.46 |
| 5 | 28.16 | 27.75 |
| 6 | 33.86 | 29.16 |
| 7 | 45.77 | 33.34 |
| 8 | 28.20 | 28.53 |
| 9 | 28.54 | 28.93 |
| 10 | 27.88 | 27.56 |

## Example 2—Quadruple Time Series of Daily Dow Jones Industrial Average Index

在這個範例中，我們也是對現實世界的時間序列數據來驗證模型的效能，與範例一不同的是，四個目標之間不是收盤價與開盤價的關係，意即目標的曲線相似度沒有像範例一目標曲線的相似度高。範例二的四個目標使用的資料為The Taiwan Stock Exchange Capitalization Weighted Stock Index (TAIEX)、Dow Jones Industrial Average Index (DJI) 、National Association of Securities Dealers Automated Quotation (NASDAQ)、Standard & Poor's 500 (S&P500)，每年的收盤價。為了和其他論文比較效能，我們對2001年至2004年的收盤價做預測，2001年資料為245筆，2002年資料為248筆，2003年資料為249筆，2004年資料為250筆，每年的四組資料做一次預測，每年的前十個月的資料當作訓練資料，剩餘的當作測試資料，2001年測試資料為181筆，2002年為184筆，2003年為185筆，2004年為185筆。本篇論文所提出的模型可以一次有多個複數型態的輸出，故可以預測多個目標，我們在範例中使用的第一個複數目標實數部分為TAIEX當年的收盤價，複數部分則為DJI當年的收盤價，第二個目標實數部分為NASDAQ當年的收盤價，複數部分為S&P500當年的收盤價。

結構學習部分，則將每個被挑選的特徵，用MATLAB提供的subclust()函式分群，係數皆為0.3。整體模型在結構學習後的參數，如表IV至表VII所示。PSO-RLSE混合方法的機器學習參數設定，如表VIII所示。

表V

範例二模型設定(2001年)

|  |  |
| --- | --- |
| **Parameter** | **Value** |
| Feature variable as model input | { } |
| Number of fuzzy sets of each input | 3、4、3、4 |
| Number of targets (real-valued) | 4 |
| Number of outputs (complex-valued) | 2 |
| Number of premises (before selection) | 144 |
| Type of premises | Sphere complex fuzzy set |
| Number of premises (after selection) | 15 |
| Number of premise parameters | 56 |
| Number of aim object | 3 |
| Type of consequences | Takagi-sugeno |
| Number of consequences | 3 |
| Number of consequence parameters | 15 |

表VI

範例二模型設定(2002年)

|  |  |
| --- | --- |
| **Parameter** | **Value** |
| Feature variable as model input | { } |
| Number of fuzzy sets of each input | 3、3、3、3 |
| Number of targets (real-valued) | 4 |
| Number of outputs (complex-valued) | 2 |
| Number of premises (before selection) | 81 |
| Type of premises | Sphere complex fuzzy set |
| Number of premises (after selection) | 13 |
| Number of premise parameters | 48 |
| Number of aim object | 5 |
| Type of consequences | Takagi-sugeno |
| Number of consequences | 5 |
| Number of consequence parameters | 25 |

表VII

範例二模型設定(2003年)

|  |  |
| --- | --- |
| **Parameter** | **Value** |
| Feature variable as model input | { } |
| Number of fuzzy sets of each input | 3、3、3、3 |
| Number of targets (real-valued) | 4 |
| Number of outputs (complex-valued) | 2 |
| Number of premises (before selection) | 81 |
| Type of premises | Sphere complex fuzzy set |
| Number of premises (after selection) | 11 |
| Number of premise parameters | 48 |
| Number of aim object | 3 |
| Type of consequences | Takagi-sugeno |
| Number of consequences | 3 |
| Number of consequence parameters | 15 |

表VIII

範例二模型設定(2004年)

|  |  |
| --- | --- |
| **Parameter** | **Value** |
| Feature variable as model input | { } |
| Number of fuzzy sets of each input | 5、4、4、4 |
| Number of targets (real-valued) | 4 |
| Number of outputs (complex-valued) | 2 |
| Number of premises (before selection) | 320 |
| Type of premises | Sphere complex fuzzy set |
| Number of premises (after selection) | 15 |
| Number of premise parameters | 68 |
| Number of aim object | 8 |
| Type of consequences | Takagi-sugeno |
| Number of consequences | 8 |
| Number of consequence parameters | 40 |

表IX

機器學習參數設定

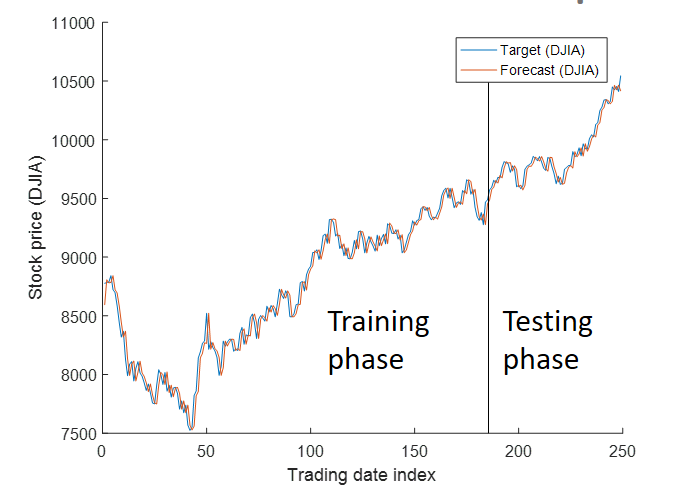
|  |  |
| --- | --- |
| **PSO** | |
| Swarm size | 64 |
| Iterations | 100 |
|  | 0.8 2.0 2.0 |
|  | Random in [0,1] |
| Initial position | Given by the subtractive clustering algorithm |
| Initial velocity | 0 |
| **RLSE** | |
|  |  |
|  | Given by Number of consequence parameters |
|  | Dx1 zero vector |
|  | **I** |
| **I** | DxD identify matrix |

本次範例的結果將與其他論文所提的方法做比較，像是ANFIS[21]，CNFS-ARIMA[6]，RBF network[22]和SVR[23][24]。其中除了SVR之外的模型，皆可以同時預測兩個實數目標。所以我們將使用模型的第一組複數型態的輸出與其他論文做比較，結果如表IX和表X所示。以2003年為例目標與模型輸出的結果，如圖(4)所示。

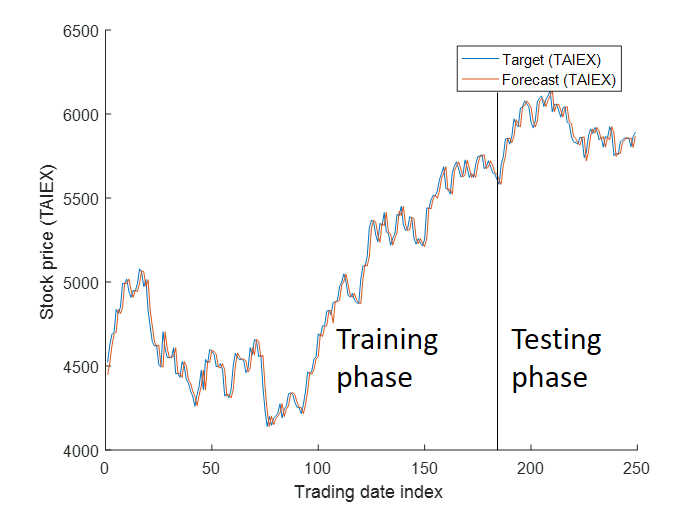
表 X

十次效能表(實驗二)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Performance(RMSE) | | | |
| triers | 2001 | 2002 | 2003 | 2004 |
| 1 | 259.74 | 279.93 | 196.91 | 282.82 |
| 2 | 259.27 | 286.02 | 198.10 | 283.94 |
| 3 | 260.99 | 280.34 | 199.06 | 286.83 |
| 4 | 260.44 | 280.56 | 199.54 | 284.57 |
| 5 | 263.29 | 281.85 | 196.94 | 278.60 |
| 6 | 263.57 | 282.00 | 197.85 | 288.52 |
| 7 | 259.86 | 279.46 | 200.05 | 275.31 |
| 8 | 262.02 | 282.35 | 198.33 | 285.69 |
| 9 | 261.23 | 282.11 | 197.53 | 288.15 |
| 10 | 260.59 | 280.93 | 197.62 | 286.68 |



(a)



(b)

圖(4) 範例二實際數值和模型輸出(a)DJIA (2003年) (b)TAIEX (2003年)

表XI

DJIA效能比較表(RMSE)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Method Year** | **2001** | **2002** | **2003** | **2004** |
| SVR (two models, each with single output) | 101.44 | 117.95 | 82.76 | 71.49 |
| ANFIS (two models, each with single output) | 105.56 | 111.69 | 72.09 | 68.00 |
| ANFIS (one model with two outputs) | 128.20 | 142.05 | 90.37 | 83.69 |
| RBF (two models, each with single output) | 106.33 | 131.24 | 97.58 | 81.79 |
| RBF (one model with two outputs) | 181.79 | 136.28 | 154.14 | 148.11 |
| CNFS(5)-ARIMA (one model with two outputs) | 103.06 | 103.42 | 70.70 | 66.55 |
| SCNFS(proposed) training phase | 91.95 | 98.69 | 69.66 | 97.99 |
| SCNFS(proposed) testing phase | 94.31 | 85.52 | 56.26 | 61.64 |

表XII

TAIEX效能比較表(RMSE)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Method Year** | **2001** | **2002** | **2003** | **2004** |
| SVR (two models, each with single output) | 162.46 | 67.72 | 59.47 | 58.81 |
| ANFIS (two models, each with single output) | 147.36 | 70.17 | 72.61 | 65.33 |
| ANFIS (one model with two outputs) | 151.62 | 78.27 | 81.69 | 70.54 |
| RBF (two models, each with single output) | 134.32 | 65.15 | 60.41 | 102.86 |
| RBF (one model with two outputs) | 137.58 | 78.54 | 115.92 | 126.48 |
| CNFS(5)-ARIMA (one model with two outputs) | 115.82 | 64.34 | 57.69 | 55.56 |
| SCNFS(proposed) training phase | 92.03 | 100.26 | 69.96 | 99.11 |
| SCNFS(proposed) testing phase | 89.59 | 86.81 | 55.34 | 60.30 |

## Example 3—Quadruple Time Series of Daily Taiwan Stock Exchange Capitalization Weighted Stock Index

在這個範例中，我們也是對現實世界的時間序列數據來驗證模型的效能，這次預測的目標為APPLE Computer Inc.、International Business Machines Corporation (IBM)、Dell Inc.、Microsoft Inc.，四個股票的收盤價，期間為2003年2月10號至2005年1月21號，總共為492筆。為了和其他論文比較效能，我們用2003年2月10號至2004年9月10號的資料當作訓練資料，共433筆，剩餘的當作測試資料，共59筆。本篇論文所提出的模型可以一次有多個複數型態的輸出，故可以預測多個目標，我們這個範例中使用的第一個複數目標實數部分為IBM的收盤價，複數部分則為APPLE的收盤價，第二個目標實數部分為Dell的收盤價，複數部分為Microsoft的收盤價。

結構學習部分，則將每個被挑選的特徵，用MATLAB提供的subclust()函式分群，係數皆為0.3。並透過前鑑部篩選機制，從原本的81個前鑑部篩選到剩下15個前鑑部。整體模型在結構學習後的參數，如表XI所示。PSO-RLSE混合方法的機器學習參數設定，如表XII所示。

表XIII

範例三模型設定

|  |  |
| --- | --- |
| **Parameter** | **Value** |
| Feature variable as model input | { } |
| Number of fuzzy sets of each input | 3、3、3、3 |
| Number of targets (real-valued) | 4 |
| Number of outputs (complex-valued) | 2 |
| Number of premises (before selection) | 81 |
| Type of premises | Sphere complex fuzzy set |
| Number of premises (after selection) | 9 |
| Number of premise parameters | 48 |
| Number of aim object | 3 |
| Type of consequences | Takagi-sugeno |
| Number of consequences | 3 |
| Number of consequence parameters | 15 |

表XIV

機器學習參數設定

|  |  |
| --- | --- |
| **PSO** | |
| Swarm size | 64 |
| Iterations | 100 |
|  | 0.8 2.0 2.0 |
|  | Random in [0,1] |
| Initial position | Given by the subtractive clustering algorithm |
| Initial velocity | 0 |
| **RLSE** | |
|  |  |
|  | 15x1 zero vector |
|  | **I** |
| **I** | 15x15 identify matrix |

本次範例的結果將與其他論文所提的方法做比較，像是HiMMI、ANN-GA-HMM-Interpolation、ANN-GA-HMM-WA、ARIMA。所以我們將使用模型的第一組輸出的實數及複數部分和第二輸出的實數部份與其他論文做比較，結果如表XIII所示。模型的機器學習曲線，如圖(5)所示;目標與模型輸出的結果，如圖(6)所示。

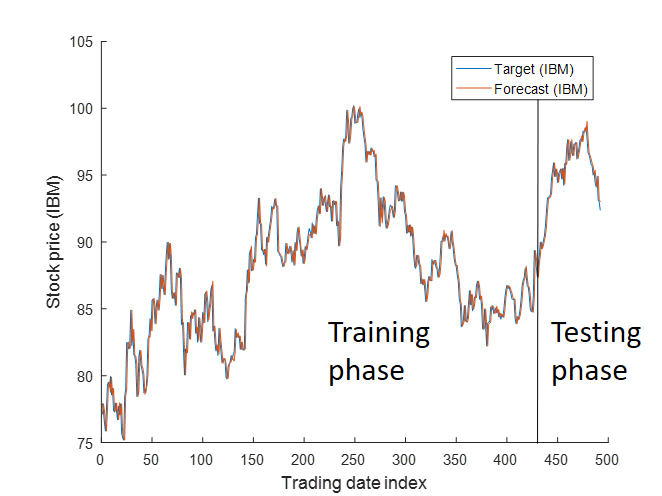
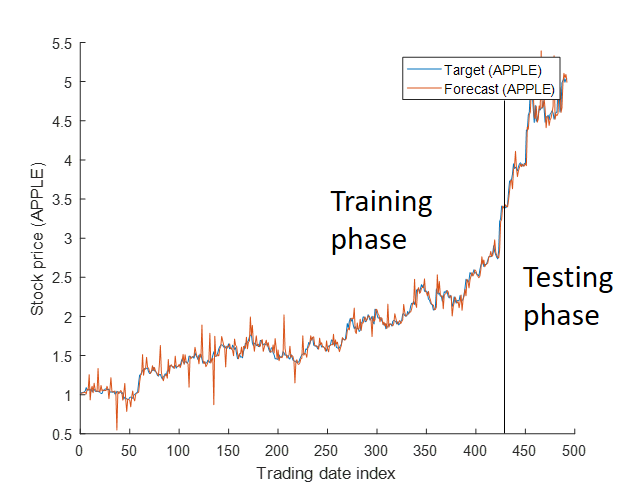
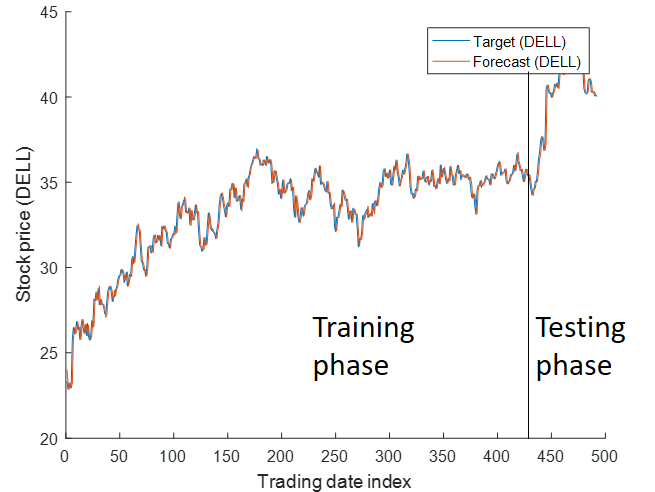


圖(5) 範例三學習曲線

表 IV

十次效能表(實驗二)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Performance(RMSE) | | | |
| triers | APPLE | IBM | DELL | Microsoft |
| 1 | 0.6633 | 0.6633 | 0.6627 | 0.6525 |
| 2 | 0.6573 | 0.6655 | 0.6682 | 0.6606 |
| 3 | 0.6852 | 0.6627 | 0.6665 | 0.6607 |
| 4 | 0.6760 | 0.6563 | 0.5407 | 0.6611 |
| 5 | 0.6643 | 0.6725 | 0.6748 | 0.6572 |
| 6 | 0.6627 | 0.6631 | 0.6670 | 0.6629 |
| 7 | 0.6655 | 0.6617 | 0.6693 | 0.6553 |
| 8 | 0.7070 | 0.6834 | 0.6494 | 0.6504 |
| 9 | 0.6706 | 0.66057 | 0.6511 | 0.6558 |
| 10 | 0.6587 | 0.6528 | 0.6628 | 0.6530 |

 (a)  (b)  (c)

圖(6) DJI的實際數值和模型輸出(a) IBM收盤價 (b) APPLE收盤價 (c) DELL收盤價

表XV

效能比較表(MAPE)

|  |  |  |  |
| --- | --- | --- | --- |
| **Method Company** | **APPLE** | **IBM** | **DELL** |
| HiMMI | 2.8373 | 1.2186 | 1.0117 |
| ANN-GA-HMM-Interpolation | 2.1649 | 1.0555 | 0.8446 |
| ANN-GA-HMM-WA | 1.9247 | 0.8487 | 0.6992 |
| ARIMA | 1.8009 | 0.9723 | 0.6604 |
| SCNFS(proposed) training phase | 0.7828 | 0.7883 | 0.7833 |
| SCNFS(proposed) testing phase | 0.6760 | 0.6563 | 0.6507 |

# Conclusion

透過三個實驗可以發現，本論文提出的模型SCFNS確實有多目標預測的能力。透過多目標特徵挑選，可以根據不同的資料去萃取原始資料中有用的資料，並且控制進入模型的資料大小。結構化學習的部分，可以將輸入的資料自動地根據數據做調整，在面臨不同的數據可以自己生成不同的結構。從實驗中可以發現透過一次對四個目標做時間序列的預測，各個目標的效果不亞於其他論文所提出的方法，甚至更好。證明了，不同資料都能在此模型中，被有效的預測。也代表PSO-RLSE的混合方法，有著一定的水準。目前可以發現模型對多目標的時間序列預測是可以運作的，而機器學習的部分可能會受限於PSO本身的特性，收斂快速、較容易掉入區域最佳解，從圖(2)和不難發現，PSO演算法在前幾個回合就已經接近最後一回合的RMSE，故在資料量龐大的情況下，所需搜尋維度增多，PSO效能就會較差，可能會限制住模型的整體發揮。在未來，可以透過不同的機器學習演算法結合SCFNS，像是Random Search[25]、CPSO[26]等，以克服目前PSO所遇到的問題。最後本篇論文所提供的模型與傳統類神經網路的IF-THEN規則有所差別，透過箭靶層的連結，前鑑部(IF-part)和後鑑部(THEN-part)可以不是一對一的關係，模型效果仍然是有達到預期，這個變動將使得模型更有彈性，在未來可以增添更多的隱藏層。

##### Acknowledgment

##### References

1. A. J. Patton, “A review of copula models for economic time series,” Journal of Multivariate Analysis, vol. 110, pp. 4-18, 2012.
2. S. S. Jones, R. S. Evans, T. L. Allen, A. Thomas, P. J. Haug, S. J. Welch and G. L. Snow, “A multivariate time series approach to modeling and forecasting demand in the emergency department,” Journal of Biomedical Informatics, vol. 42, pp. 123-139, 2009.
3. P. Aboagye-Sarfo, Q. Mai, F. M. Sanfilippo, D. B. Preen, L. M. Stewart and D. M. Fatovich, “A comparison of multivariate and univariate time series approaches to modeling and forcasting emergency department demand in Western Australia,” Journal of Biomedical Informatics, vol. 57, pp. 62-73, 2015.
4. S.-T. Li and T.-C. Cheng, “Deterministic fuzzy time series model for foercasting enrollments,” Computers and Mathematics with Application, vol. 53, pp. 1904-1920, 2017.
5. C. Li and J.-W. Hu, “A new ARIMA-based neuro-fuzzy approach and swarm intelligence for time series forecasting,” Engineering Applications of Artificial Intelligence, vol. 25, pp. 295-308, 2012.
6. C. Li and T.-W. Chiang, ”Complex neurofuzzy ARIMA forecasting—a new approach using complex fuzzy sets,” IEEE Transtraction on fuzzy systems. vol. 21.no. 3, pp.567-584, June 2013.
7. L. J. Herrera, H. Pomares, I. Rojas, A. Guillen, J. Gonzalez, M. Awad and A. Herrera, ”Multigrid-based fuzzy systems for time series prediction:CATS competition,” Neurocomputing, vol.70, pp. 2410-2425,2007.
8. I. Sugiarto and S. Natarajan, “Parameter estimation using least square method for MIMO Takagi-Sugeno neuro-fuzzy in time series forecasting,” J. Tek. Elektro, vol. 7(2), pp. 82-87, 2007.
9. M. Z.-Kermani and M. Teshnehlab, ”Using adaptive neuro-fuzzy inference system for hydrological time series prediction.” Appl. Soft Comput, vol. 8,pp. 928-936, 2008.
10. H. J. Rong,N. Sundararajan,G. B. Huang and P. Saratchandran, “Sequential adaptive fuzzy inference system (SAFIS) for nonlinear system identification and prediction,” Fuzzy Set Syst, vol.157, pp. 1260-1275, 2006.
11. Zadeh, “Fuzzy sets”, Inf. Control , vol. 8, pp. 338-353, 1965.
12. D. Ramot, R. Milo, M. Friedman, and A, Kandel, “Complex fuzzy sets,” IEEE Trans. Fuzzy syst., vol . 10, no.2, pp. 171-186, Apr. 2002
13. C. Li, T.-W. Chiang. J.-W. Hu, and T. Wu, “Complex neuro-fuzzy intelligent approach to function approximation,” in Proc. 2010 3rd Int.Workshop Adv. Comput. Intell., pp. 151-156, 2010.
14. C. Li and T.-W. Chiang, “Complex fuzzy computing to time series prediction—A multi-swarm PSO learning approach,” Lect. Notes Artif. Intell., vol. 6592, pp. 242-251, 2011.
15. E. C. Shannon, ”A mathematical theory of communication,” Bell System Technical Journal, vol. 27, pp. 379-423, 1948.
16. C. Li , ”Multi-target feature selection,” unpublished, 2017.
17. T. Takagi and M. Sugeno, “Fuzzy identification of systems and its applocations to modeling and control,” IEEE Trans. Systems, Man, and Cybernetics, vol. SMC-15, no. 1, pp. 116-132, 1985.
18. J. Kennedy and R. Eberhart, “Particle swarm optimization,” Proceedings of IEEE International Conference on Neural Networks, vol. 4, pp. 1942-1948, 1995.
19. J. S. R. Jang, C. T. Sun and E. Mizutani, “Neuro-fuzzy and soft computing: A computational approach to learning and machine intelligence prentice hall,” Upper Saddle River, 1997.
20. C. Li and T.-W. Chiang, “Complex neuro-fuzzy self-learning approach to function approximation,” Lect. Notes Artif. Intell., vol. 5991, pp. 289-299, 2010.
21. J. S. R. Jang, “ANFIS: Adaptive-network-based fuzzy inference system,” IEEE Trans. Syst., Man, Cybern., vol. 23, no. 3, pp. 665-685, May/Jun, 1993.
22. D. S. Broomhead and D. Lowe, “Multivariable functional interpolation and adaptive networks,” Complex Syst., vol. 2, pp. 321-355, 1998.
23. A. J. Smola and B. Scholkopf , “A tutorial on support vector regression,” Static. Comput., vol. 14, no. 3, pp. 199-222, 2004.
24. N. I. Sapankevychand and R. Sankar, “Time series prediction using support vector machines: A survey,” IEEE Comput. Intell. Mag., vol.4, no. 2, pp. 24-38, May 2009.
25. Z. B. Zabinsky, ”Random search algorithms.” Technical report, University of Washington, Seattle, 2009.
26. F. van den Bergh and A.P. Engelbrecht, ”A cooperative approach to particle swarm optimization,” IEEE. Trans. On Evolutionary Computation, vol. 8, pp. 225-239, 2004.