# Experimentation

本研究總共有3個實驗，每個實驗目標相關性高，如實驗一開盤價及收盤價的關係；實驗二為指標型指數，其中，台股加權指數(The Taiwan Stock Exchange Capitalization Weighted Stock Index, TAIEX)為台灣上市的股票中經過加權計算出的指標，代表著台灣上市股票的波動，道瓊工業指數(Dow Jones Industrial Average Index, DJI)涵蓋著財務等9大產業，為一股價加權指標，納斯達克(National Association of Securities Dealers Automated Quotation, NASDAQ)為超過三千檔股票所組合成的市值加權指標，大多以科技產業為例，標準普爾500(Standard and Poor’s ,S&P 500)為美國前500大公司的市值加權，當中包含IT等11個產業。上述皆為非常著名的股票指標，故能精準預測可帶來一定的效用。第三個實驗使用的目標多為大規模的科技公司，如IBM、APPLE、DELL、Microsoft等。為了與其他文獻比較模型的好壞，我們將透過兩種指標對模型評估。前兩個實驗使用RMSE做為評估指標，第3個實驗，基於其他文獻所提供的數據資料，我們將以平均絕對誤差百分比(Mean average percent errors, MAPE)做為評估指標，MAPE公式如下。

|  |  |  |
| --- | --- | --- |
|  | , |  |

其中，為資料總筆數，為第筆真實數據值，為第筆模型輸出。

在結構學系部分，本研究中所有實驗的上限皆設定為4，下限皆設定為2，故模型輸入維度會介於2~4之間。

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

在這個實驗中，我們使用真實世界的時間序列數據來驗證模型的效能。使用的資料為納斯達克(National Association of Securities Dealers Automated Quotation, NASDAQ) 2007年1月3號至2010年12月20號每天的開盤和收盤價以及標準普爾500(Standard and Poor’s ,S&P 500) 2007年1月3號至2010年12月20號每天的開盤和收盤價。此實驗的原始資料為1029筆，經過一次差分得到1028筆，並以4組差分資料取出30個特徵，共有120個特徵，每個特徵會有998筆資料，前500筆資料為訓練資料，剩下的為測試資料。第1至30個特徵為NASDAQ開盤價，第31至60個特徵為NASDAQ收盤價，第61至90個特徵為S&P 500 開盤價，第91至120個特徵為S&P 500收盤價，這120個特徵與目標形成資料矩陣，資料矩陣中以S&P500距離目標最近，目標排序為NASDAQ開盤價、NASDAQ收盤價、S&P500開盤價、S&P500收盤價。資料矩陣經過多目標特徵挑選[22]後，選出特徵作為模型輸入，複數型態目標有兩個，第一組複數型態的目標實部部分為NASDAQ開盤價，虛部部分為NASDAQ收盤價，第二組複數型態的目標實部部分為S&P500開盤價，虛部部分為S&P500收盤價。結構學習部分，則將每個被挑選的特徵，用減數分群演算法[11]分群。並透過第二章所介紹的神經元挑選方法，從原本的54個神經元篩選到剩下8個神經元。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。本篇論文所提出的模型可以一次有四個複數型態的輸出，故可以預測四個複數型態目標，在此實驗中只同時預測兩個目標。

實驗一模型設定

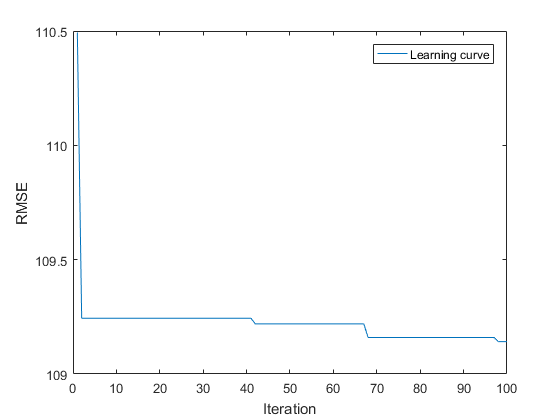
|  |  |
| --- | --- |
| **Parameter** | **Value** |
| Feature variables as model inputs | {} |
| Number of input fuzzy sets | {2, 3, 3, 3} |
| Type of fuzzy sets | SCFS |
| Number of complex–valued targets\* | 2 |
| Number of neurons | 8 |
| Number of parameters in the SCFS layer | 44 |
| Number of aim object neurons | 5 |
| Number of T–S neurons | 5 |
| Number of parameters in the T–S layer | 25 |

\* Each complex-valued target whose real and imaginary parts contain two real-valued targets, respectively.

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 particle positions | By SC algorithm in section II-C |
| Initial particle velocities | 0 |
| **RLSE** | |
|  |  |
|  | 25-by-1 zero vector |
|  | **I** |
| **I** | 25-by-25 identify matrix |

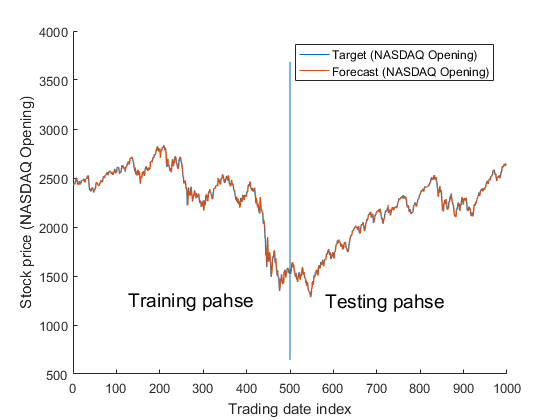
本次實驗的結果將與其他文獻[23]所提的方法做比較。其中除了SVR之外的模型，皆可以同時預測兩個實數目標。所以我們將使用模型的第一組複數型態的輸出與其他文獻做比較，結果如表III所示，十次重複實驗效能統計如表IV所示。模型的學習曲線，如Fig. 5所示，可看出在前10個迭代時，學習逐漸穩定；目標與模型輸出的結果，如Fig. 6所示，有著不錯的效果；預測誤差如Fig. 7所示，誤差呈現亂數狀態，範圍介於-50至50，代表模型預測能力穩定。



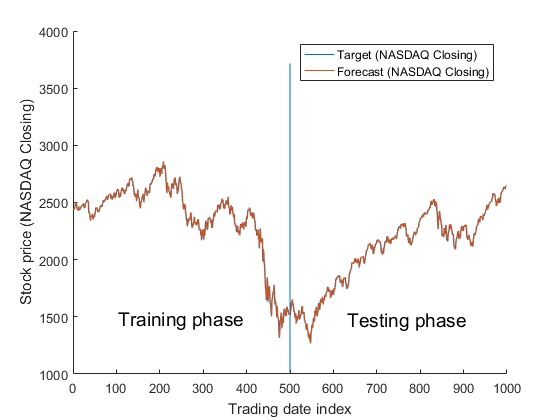
1. 學習曲線。可從此曲線看出RMSE的變化量，在前10迭代時，學習逐漸穩定。(實驗一)

效能比較(NASDAQ, 實驗一)

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Method | RMSE | | | | |
| Training phase |  |  | Testing phase |  |
| Opening index | Closing index |  | Opening index | Closing index |
| SVR (two models, each with single output) [23] | 35.18 | 35.24 |  | 37.23 | 40.24 |
| ANFIS (two models, each with single output) [23] | 37.83 | 38.66 |  | 38.80 | 42.36 |
| ANFIS (one model with two outputs) [23] | 62.75 | 71.51 |  | 72.52 | 85.08 |
| RBF (two models, each with single output) [23] | 37.59 | 33.89 |  | 37.52 | 44.08 |
| RBF (one model with two outputs) [23] | 178.57 | 179.87 |  | 261.37 | 258.89 |
| CNFS(5)-ARIMA (one model with two outputs) [23] | 21.56 | 20.81 |  | 32.52 | 33.70 |
| SCNFS(proposed) | 38.58 | 38.59 |  | 27.73 | 27.44 |

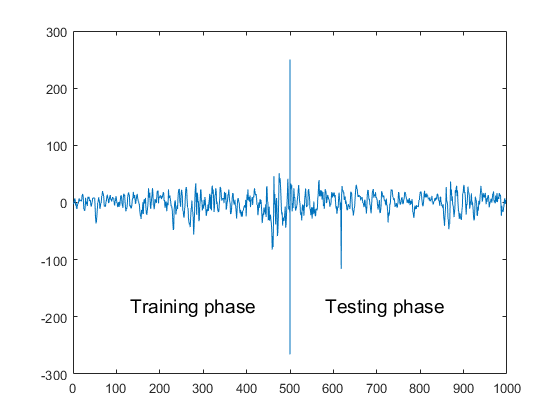


(a)



(b)

1. NASDAQ的實際數值和模型輸出(a)每日開盤價(b)每日收盤價。x軸為交易日，y軸為股票價格，可看出訓練和測試階段皆有不錯的效果。(實驗一)



1. 預測誤差。誤差呈現亂數狀態，範圍介於-50至50，代表模型預測能力穩定。(實驗一)

十次重複實驗效能統計(實驗一)

|  |  |  |
| --- | --- | --- |
|  | Performance (RMSE) | |
| Trials | Opening index | Closing index |
| 1 | 27.67 | 27.75 |
| 2 | 29.62 | 28.04 |
| 3 | 33.85 | 33.99 |
| 4 | 28.09 | 28.77 |
| 5 | **27.73** | **27.44** |
| 6 | 33.32 | 42.27 |
| 7 | 27.68 | 27.61 |
| 8 | 29.35 | 27.93 |
| 9 | 27.67 | 27.91 |
| 10 | 27.72 | 27.76 |

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

在這個實驗中，我們也是使用真實世界的時間序列數據來驗證模型的效能，與實驗一不同的是，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實驗一目標曲線的相似度高。使用的資料為台股加權指數(The Taiwan Stock Exchange Capitalization Weighted Stock Index, TAIEX)、道瓊工業指數(Dow Jones Industrial Average Index, DJI)、納斯達克(National Association of Securities Dealers Automated Quotation, NASDAQ)以及標準普爾500(Standard and Poor’s ,S&P 500)，資料區間為2001年至2004年的收盤價，每年收盤價筆數如下，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，每年的做一次模型預測，前十個月的資料當作訓練資料，剩餘的當作測試資料。每年四組原始資料會做一次差分，並以4組差分資料取出30個特徵，共有120個特徵。第1至30個特徵為TAIEX收盤價，第31至60個特徵為DJI收盤價，第61至90個特徵為NASDAQ收盤價，第91至120個特徵為S&P 500收盤價，這120個特徵與目標形成資料矩陣，資料矩陣中以S&P500距離目標最近，目標排序為TAIEX收盤價、DJI收盤價、NASDAQ收盤價、S&P500收盤價。資料矩陣經過多目標特徵挑選[22]後，將選出的特徵作為模型輸入，複數型態目標有兩個，第一組複數型態的目標實部部分為TAIEX收盤價，虛部部分為DJI收盤價，第二組複數型態的目標實部部分為NASDAQ收盤價，虛部部分為S&P500收盤價。結構學習部分，則將每個被挑選的特徵，用減數分群演算法[11]分群。並透過第二章所介紹的神經元挑選方法，減少神經元數目。整體模型在結構學習後的參數，如表V-表VIII所示。PSO–RLSE混合方法的機器學習參數設定，如表IX所示。本篇論文所提出的模型可以一次有四個複數型態的輸出，故可以預測四個複數型態目標，但在此實驗中只同時預測兩個目標。

實驗二模型設定(2001年)

|  |  |
| --- | --- |
| **Parameter** | **Value** |
| Feature variables as model inputs | {} |
| Number of input fuzzy sets | {3, 4, 3, 4} |
| Type of fuzzy sets | SCFS |
| Number of complex–valued targets\* | 2 |
| Number of neurons | 15 |
| Number of parameters in the SCFS layer | 56 |
| Number of aim object neurons | 3 |
| Number of T–S neurons | 3 |
| Number of parameters in the T–S layer | 15 |

\* Each complex-valued target whose real and imaginary parts contain two real-valued targets, respectively.

實驗二模型設定(2002年)

|  |  |
| --- | --- |
| **Parameter** | **Value** |
| Feature variables as model inputs | {} |
| Number of input fuzzy sets | {3, 3, 3, 3} |
| Type of fuzzy sets | SCFS |
| Number of complex–valued targets\* | 2 |
| Number of neurons | 13 |
| Number of parameters in the SCFS layer | 48 |
| Number of aim object neurons | 5 |
| Number of T–S neurons | 5 |
| Number of parameters in the T–S layer | 25 |

\* Each complex-valued target whose real and imaginary parts contain two real-valued targets, respectively.

實驗二模型設定(2003年)

|  |  |
| --- | --- |
| **Parameter** | **Value** |
| Feature variables as model inputs | {} |
| Number of input fuzzy sets | {3, 3, 3, 3} |
| Type of fuzzy sets | SCFS |
| Number of complex–valued targets\* | 2 |
| Number of neurons | 11 |
| Number of parameters in the SCFS layer | 48 |
| Number of aim object neurons | 4 |
| Number of T–S neurons | 4 |
| Number of parameters in the T–S layer | 20 |

\* Each complex-valued target whose real and imaginary parts contain two real-valued targets, respectively.

實驗二模型設定(2004年)

|  |  |
| --- | --- |
| **Parameter** | **Value** |
| Feature variables as model inputs | {} |
| Number of input fuzzy sets | {5, 4, 4, 4} |
| Type of fuzzy sets | SCFS |
| Number of complex–valued targets\* | 2 |
| Number of neurons | 15 |
| Number of parameters in the SCFS layer | 68 |
| Number of aim object neurons | 8 |
| Number of T–S neurons | 8 |
| Number of parameters in the T–S layer | 40 |

\* Each complex-valued target whose real and imaginary parts contain two real-valued targets, respectively.

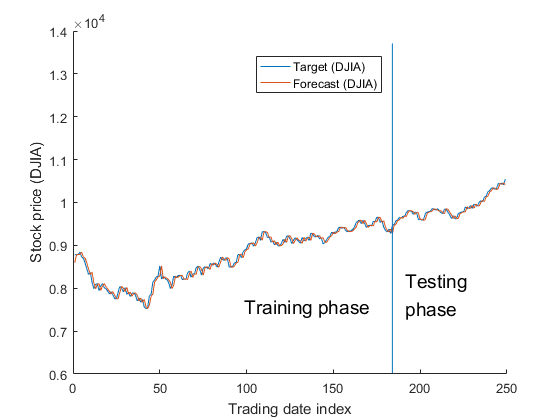
機器學習參數設定

|  |  |
| --- | --- |
| **PSO** | |
| Swarm size | 64 |
| Iterations | 100 |
|  | {0.8, 2.0, 2.0} |
|  | Random in [0,1] |
| Initial particle positions | By SC algorithm in section II-C |
| Initial particle velocities | 0 |
| **RLSE** | |
|  |  |
|  | By number of parameters in the T–S layer |
|  | -by-1 zero vector |
|  | **I** |
| **I** | -by- identify matrix |

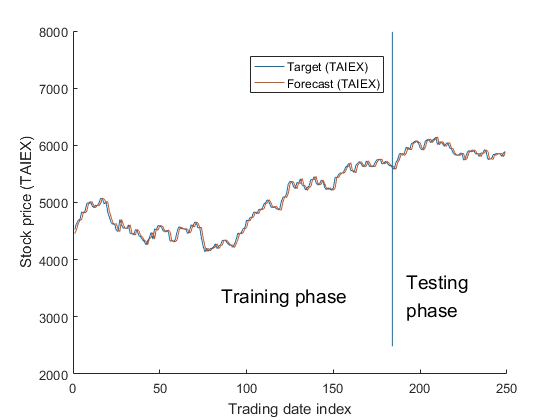
本次實驗的結果將與其他文獻[23]所提的方法做比較。其中除了SVR之外的模型，皆可以同時預測兩個實數目標。所以我們將使用模型的第一組複數型態的輸出與其他論文做比較，結果如表XI和表XII所示，十次重複實驗效能統計如表X所示。目標與模型輸出的結果，如Fig. 8所示，有著不錯的效果；預測誤差如Fig. 9所示，誤差呈現亂數狀態，範圍介於-100至100，代表模型預測能力穩定。

十次重複實驗效能統計(實驗二)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Performance (RMSE) | | | |
| Trials | 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 |

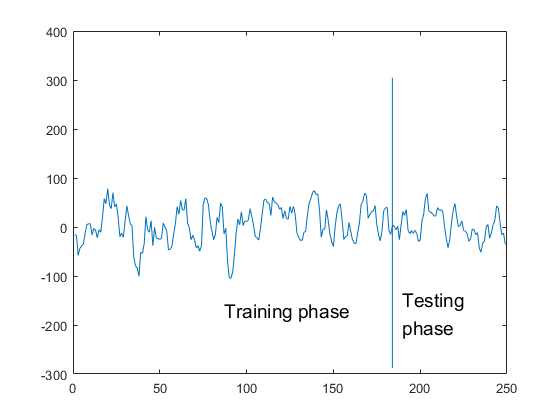


(a)



(b)

1. 實際數值和模型輸出(a)DJIA (2003年) (b)TAIEX (2003年)。x軸為交易日，y軸為股票價格，可看出訓練和測試階段皆有著不錯的效果。(實驗二)



1. 預測誤差。誤差呈現亂數狀態，範圍介於-100至100，代表模型預測能力穩定。(實驗二)

效能比較(DJIA, 實驗二)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | RMSE | | | |
| Method | **2001** | **2002** | **2003** | **2004** |
| SVR (two models, each with single output) [23] | 101.44 | 117.95 | 82.76 | 71.49 |
| ANFIS (two models, each with single output) [23] | 105.56 | 111.69 | 72.09 | 68.00 |
| ANFIS (one model with two outputs) [23] | 128.20 | 142.05 | 90.37 | 83.69 |
| RBF (two models, each with single output) [23] | 106.33 | 131.24 | 97.58 | 81.79 |
| RBF (one model with two outputs) [23] | 181.79 | 136.28 | 154.14 | 148.11 |
| CNFS(5)-ARIMA (one model with two outputs) [23] | 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 |

效能比較(TAIEX, 實驗二)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | RMSE | | | |
| Method | **2001** | **2002** | **2003** | **2004** |
| SVR (two models, each with single output) [23] | 162.46 | 67.72 | 59.47 | 58.81 |
| ANFIS (two models, each with single output) [23] | 147.36 | 70.17 | 72.61 | 65.33 |
| ANFIS (one model with two outputs) [23] | 151.62 | 78.27 | 81.69 | 70.54 |
| RBF (two models, each with single output) [23] | 134.32 | 65.15 | 60.41 | 102.86 |
| RBF (one model with two outputs) [23] | 137.58 | 78.54 | 115.92 | 126.48 |
| CNFS(5)-ARIMA (one model with two outputs) [23] | 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

在這個實驗中，我們使用真實世界的時間序列數據來驗證模型的效能。使用的資料為蘋果電腦(Apple computer inc.)、IBM(International business machines corporation)、戴爾(Dell inc.)、微軟(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。此實驗的原始資料為492筆，經過一次差分得到491筆，並以4組差分資料取出30個特徵，共有120個特徵，每個特徵會有460筆資料，前433筆資料為訓練資料，剩下的為測試資料。第1至30個特徵為APPLE收盤價，第31至60個特徵為IBM收盤價，第61至90個特徵為Dell收盤價，第91至120個特徵為Microsoft收盤價，這120個特徵與目標形成資料矩陣，資料矩陣中以Microsoft距離目標最近，目標排序為APPLE、IBM、Dell和Microsoft。資料矩陣經過多目標特徵挑選[22]後，選出特徵作為模型輸入，複數型態目標有兩個，第一組複數型態的目標實部部分為APPLE收盤價，虛部部分為IBM收盤價，第二組複數型態的目標實部部分為Dell收盤價，虛部部分為Microsoft收盤價。結構學習部分，則將每個被挑選的特徵，用減數分群演算法[11]分群。並透過第二章所介紹的π神經元挑選方法，從原本的81個π神經元篩選到剩下13個π神經元。整體模型在結構學習後的參數，如表XIII所示。PSO–RLSE混合方法的機器學習參數設定，如表XIV所示。本篇論文所提出的模型可以一次有四個複數型態的輸出，故可以預測四個複數型態目標，在此實驗中只同時預測兩個目標。

實驗三模型設定

|  |  |
| --- | --- |
| **Parameter** | **Value** |
| Feature variables as model inputs | {} |
| Number of input fuzzy sets | {3, 3, 3, 3} |
| Type of fuzzy sets | SCFS |
| Number of complex–valued targets\* | 2 |
| Number of neurons | 13 |
| Number of parameters in the SCFS layer | 48 |
| Number of aim object neurons | 3 |
| Number of T–S neurons | 3 |
| Number of parameters in the T–S layer | 15 |

\* Each complex-valued target whose real and imaginary parts contain two real-valued targets, respectively.

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 particle positions | By SC algorithm in section II-C |
| Initial particle velocities | 0 |
| **RLSE** | |
|  |  |
|  | 15-by-1 zero vector |
|  | **I** |
| **I** | 15-by-15 identify matrix |

本次實驗的結果將與其他論文所提的方法做比較，像是HiMMI[13]、ANN-GA-HMM-Interpolation[13]、ANN-GA-HMM-WA[13]、ARIMA[41]、Bayesian ANN[41]。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XVI所示，十次重複實驗效能統計如表X所示。模型的學習曲線，如Fig. 10所示，可看出於45迭代時，學習逐漸穩定；目標與模型輸出的結果，如Fig. 11所示，有著不錯的效果，APPLE股價介於5美元左右，故預測曲線看起來有波動;預測誤差如Fig. 12所示，誤差呈現亂數狀態，範圍介於-2至2，代表模型預測能力穩定。



1. 學習曲線。由此圖可看出RMSE於45迭代時，模型學習已逐漸穩定。(實驗三)

十次重複實驗效能統計(實驗三)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Performance (MAPE) | | | |
| Trials | APPLE | IBM | DELL | Microsoft |
| 1 | 1.8909 | 0.8156 | 0.6173 | 0.7007 |
| 2 | 2.3692 | 0.8286 | 0.5957 | 0.7111 |
| 3 | **1.8453** | **0.8051** | **0.6187** | **0.8591** |
| 4 | 2.0273 | 0.8172 | 0.6184 | 0.7048 |
| 5 | 2.0476 | 0.8183 | 0.6100 | 0.7099 |
| 6 | 2.2186 | 0.8005 | 0.6270 | 0.7399 |
| 7 | 2.0239 | 0.8313 | 0.6219 | 0.7128 |
| 8 | 2.2186 | 0.8005 | 0.6270 | 0.7399 |
| 9 | 2.0239 | 0.8313 | 0.6219 | 0.7128 |
| 10 | 1.8636 | 0.8159 | 0.6159 | 0.7045 |



(a)



(b)



(c)

1. 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。x軸為交易日，y軸為股票價格，可看出訓練和測試階段皆有不錯的效果。(實驗三)



1. 預測誤差。誤差呈現亂數狀態，範圍介於-2至2，代表模型預測能力穩定。(實驗三)

效能比較 (實驗三)

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | **MAPE** | | | | |
| **Method** | **APPLE** |  | **IBM** |  | **DELL** |
| HiMMI[13] | 2.8373 |  | 1.2186 |  | 1.0117 |
| ANN-GA-HMM-Interpolation[13] | 2.1649 |  | 1.0555 |  | 0.8446 |
| ANN-GA-HMM-WA[13] | 1.9247 |  | 0.8487 |  | 0.6992 |
| Bayesian ANN[41] | 1.9688 |  | 0.7441 |  | - |
| ARIMA[41] | 1.8009 |  | 0.9723 |  | 0.6604 |
| SCNFS(proposed) training phase | 2.4175 |  | 1.1758 |  | 0.8955 |
| SCNFS(proposed) testing phase | 1.8453 |  | 0.8051 |  | 0.6187 |

## Example 4—Quadruple Time Series of Daily Dow Jones Industrial Average Index

在這個實驗中，我們也是使用真實世界的時間序列數據來驗證模型的效能，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實驗一目標曲線的相似度高。使用的資料為台股加權指數(The Taiwan Stock Exchange Capitalization Weighted Stock Index, TAIEX)、HSI恆生指數、Nikkei 225日經指數以及SSEC 上證指數，資料區間為2000年的收盤價，共246筆，每年的做一次模型預測，前十個月的資料當作訓練資料，剩餘的當作測試資料。每年四組原始資料會做一次差分，並以4組差分資料取出30個特徵，共有120個特徵。第1至30個特徵為TAIEX收盤價，第31至60個特徵為HSI收盤價，第61至90個特徵為Nikkei 225收盤價，第91至120個特徵為SSEC，這120個特徵與目標形成資料矩陣，資料矩陣中以SSEC距離目標最近，目標排序為TAIEX收盤價、HSI收盤價、Nikkei 225收盤價、SSEC收盤價。資料矩陣經過多目標特徵挑選[22]後，將選出的特徵作為模型輸入，複數型態目標有兩個，第一組複數型態的目標實部部分為TAIEX收盤價，虛部部分為HSI收盤價，第二組複數型態的目標實部部分為Nikkei 225收盤價，虛部部分為SSEC收盤價。結構學習部分，則將每個被挑選的特徵，用減數分群演算法[11]分群。並透過第二章所介紹的神經元挑選方法，減少神經元數目。整體模型在結構學習後的參數，如表V-表VIII所示。PSO–RLSE混合方法的機器學習參數設定，如表IX所示。本篇論文所提出的模型可以一次有四個複數型態的輸出，故可以預測四個複數型態目標，但在此實驗中只同時預測兩個目標。

實驗四模型設定

|  |  |
| --- | --- |
| **Parameter** | **Value** |
| Feature variables as model inputs | {} |
| Number of input fuzzy sets | {3, 3, 3, 3} |
| Type of fuzzy sets | SCFS |
| Number of complex–valued targets\* | 2 |
| Number of neurons | 15 |
| Number of parameters in the SCFS layer | 56 |
| Number of aim object neurons | 3 |
| Number of T–S neurons | 3 |
| Number of parameters in the T–S layer | 15 |

\* Each complex-valued target whose real and imaginary parts contain two real-valued targets, respectively.

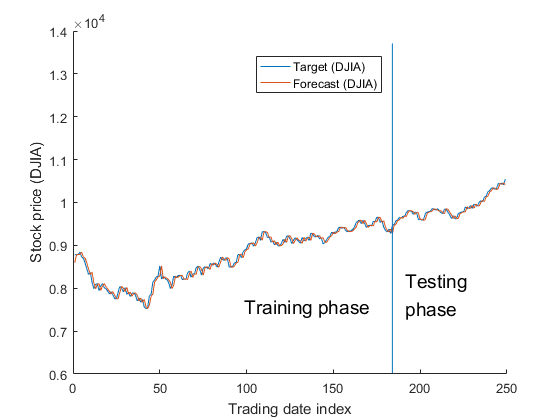
機器學習參數設定

|  |  |
| --- | --- |
| **PSO** | |
| Swarm size | 64 |
| Iterations | 100 |
|  | {0.8, 2.0, 2.0} |
|  | Random in [0,1] |
| Initial particle positions | By SC algorithm in section II-C |
| Initial particle velocities | 0 |
| **RLSE** | |
|  |  |
|  | By number of parameters in the T–S layer |
|  | -by-1 zero vector |
|  | **I** |
| **I** | -by- identify matrix |

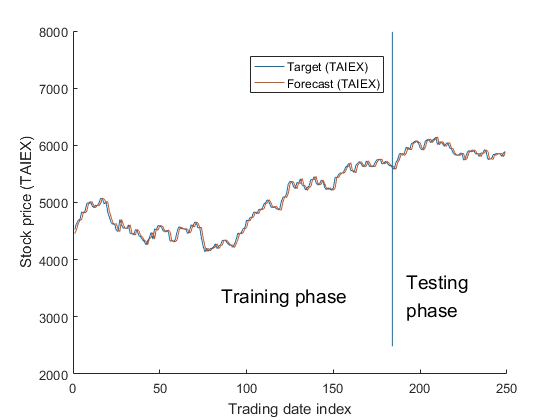
本次實驗的結果將與其他文獻[23]所提的方法做比較。其中除了SVR之外的模型，皆可以同時預測兩個實數目標。所以我們將使用模型的第一組複數型態的輸出與其他論文做比較，結果如表XI和表XII所示，十次重複實驗效能統計如表X所示。目標與模型輸出的結果，如Fig. 8所示，有著不錯的效果；預測誤差如Fig. 9所示，誤差呈現亂數狀態，範圍介於-100至100，代表模型預測能力穩定。

十次重複實驗效能統計(實驗四)

|  |  |
| --- | --- |
| Trials | Performance (RMSE) |
| 1 | 622.758638179047 |
| 2 | 621.259712779326 |
| 3 | 623.839057550794 |
| 4 | 615.384592548673 |
| 5 | 616.704401921082 |
| 6 | 607.068546703722 |
| 7 | 596.220743516664 |
| 8 | 620.100239926252 |
| 9 | 619.709993120866 |
| 10 | 613.458211924098 |
| 平均 | 615.6504 |
| 標準差 | 8.449023 |
| Best | 596.2207 |
| Worst | 623.8391 |

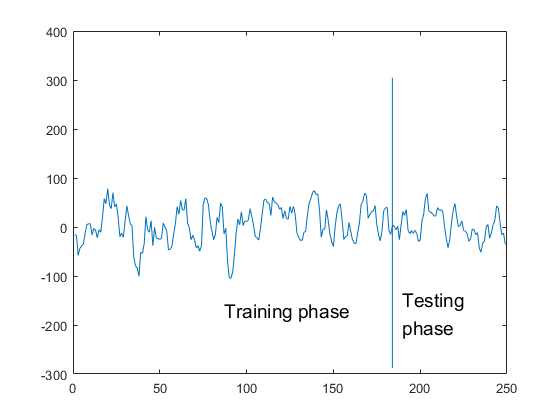


(a)



(b)

1. 實際數值和模型輸出(a)DJIA (2003年) (b)TAIEX (2003年)。x軸為交易日，y軸為股票價格，可看出訓練和測試階段皆有著不錯的效果。(實驗四)



1. 預測誤差。誤差呈現亂數狀態，範圍介於-100至100，代表模型預測能力穩定。(實驗四)

效能比較(DJIA, 實驗四)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | RMSE | | | |
| Method | **2001** | **2002** | **2003** | **2004** |
| SVR (two models, each with single output) [23] | 101.44 | 117.95 | 82.76 | 71.49 |
| ANFIS (two models, each with single output) [23] | 105.56 | 111.69 | 72.09 | 68.00 |
| ANFIS (one model with two outputs) [23] | 128.20 | 142.05 | 90.37 | 83.69 |
| RBF (two models, each with single output) [23] | 106.33 | 131.24 | 97.58 | 81.79 |
| RBF (one model with two outputs) [23] | 181.79 | 136.28 | 154.14 | 148.11 |
| CNFS(5)-ARIMA (one model with two outputs) [23] | 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 |

效能比較(TAIEX, 實驗四)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | RMSE | | | |
| Method | **2001** | **2002** | **2003** | **2004** |
| SVR (two models, each with single output) [23] | 162.46 | 67.72 | 59.47 | 58.81 |
| ANFIS (two models, each with single output) [23] | 147.36 | 70.17 | 72.61 | 65.33 |
| ANFIS (one model with two outputs) [23] | 151.62 | 78.27 | 81.69 | 70.54 |
| RBF (two models, each with single output) [23] | 134.32 | 65.15 | 60.41 | 102.86 |
| RBF (one model with two outputs) [23] | 137.58 | 78.54 | 115.92 | 126.48 |
| CNFS(5)-ARIMA (one model with two outputs) [23] | 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 5—Quadruple Time Series of Daily Dow Jones Industrial Average Index

在這個實驗中，我們也是使用真實世界的時間序列數據來驗證模型的效能，與實驗一不同的是，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實驗一目標曲線的相似度高。使用的資料為巴西股指指數BVSP、道瓊工業指數(Dow Jones Industrial Average Index, DJI)、日經指數(Nikkei 225)以及上證指數(SSEC)，資料區間為2006年3月14日至2010年4月1日，其中2006年3月14日至2009年5月27日當作訓練資料，剩餘的當作測試資料。每年四組原始資料會做一次差分，並以4組差分資料取出30個特徵，共有120個特徵。第1至30個特徵為BVSP收盤價，第31至60個特徵為DJI收盤價，第61至90個特徵為Nikkei 225收盤價，第91至120個特徵為SSEC收盤價，這120個特徵與目標形成資料矩陣，資料矩陣中以S&P500距離目標最近，目標排序為BVSP收盤價、DJI收盤價、Nikkei 225收盤價、SSEC收盤價。資料矩陣經過多目標特徵挑選[22]後，將選出的特徵作為模型輸入，複數型態目標有兩個，第一組複數型態的目標實部部分為BVSP收盤價，虛部部分為DJI收盤價，第二組複數型態的目標實部部分為Nikkei 225收盤價，虛部部分為SSEC收盤價。結構學習部分，則將每個被挑選的特徵，用減數分群演算法[11]分群。並透過第二章所介紹的神經元挑選方法，減少神經元數目。整體模型在結構學習後的參數，如表V-表VIII所示。PSO–RLSE混合方法的機器學習參數設定，如表IX所示。本篇論文所提出的模型可以一次有四個複數型態的輸出，故可以預測四個複數型態目標，但在此實驗中只同時預測兩個目標。

實驗五模型設定

|  |  |
| --- | --- |
| **Parameter** | **Value** |
| Feature variables as model inputs | {} |
| Number of input fuzzy sets | {3, 3, 3, 3} |
| Type of fuzzy sets | SCFS |
| Number of complex–valued targets\* | 2 |
| Number of neurons | 15 |
| Number of parameters in the SCFS layer | 48 |
| Number of aim object neurons | 2 |
| Number of T–S neurons | 2 |
| Number of parameters in the T–S layer | 10 |

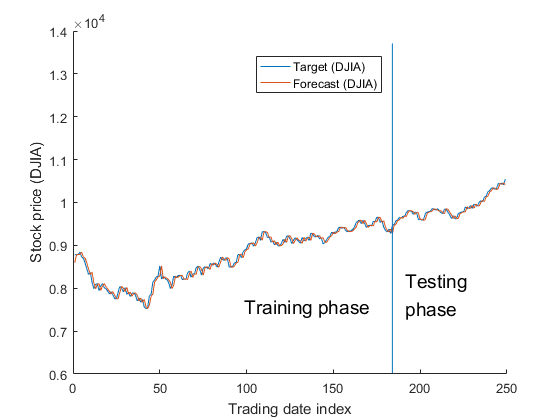
\* Each complex-valued target whose real and imaginary parts contain two real-valued targets, respectively.

機器學習參數設定

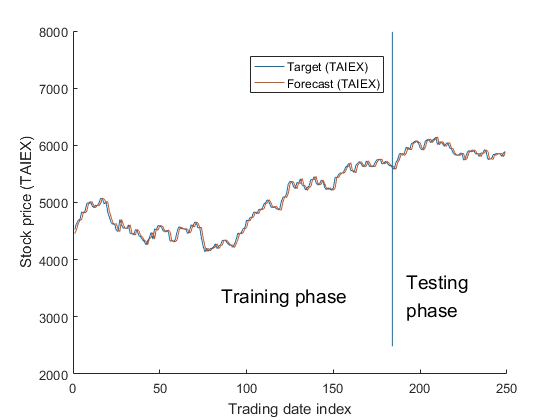
|  |  |
| --- | --- |
| **PSO** | |
| Swarm size | 64 |
| Iterations | 100 |
|  | {0.8, 2.0, 2.0} |
|  | Random in [0,1] |
| Initial particle positions | By SC algorithm in section II-C |
| Initial particle velocities | 0 |
| **RLSE** | |
|  |  |
|  | By number of parameters in the T–S layer |
|  | -by-1 zero vector |
|  | **I** |
| **I** | -by- identify matrix |

十次重複實驗效能統計(實驗五)

|  |  |
| --- | --- |
| Trials | Performance (RMSE) |
| 1 | 1357.36305358175 |
| 2 | 1356.14250112063 |
| 3 | 1353.97942328238 |
| 4 | 1351.06486410248 |
| 5 | 1354.06643783019 |
| 6 | 1882.55675699264 |
| 7 | 1354.60934003776 |
| 8 | 1358.15750856816 |
| 9 | 1358.25392899641 |
| 10 | 1391.28831800211 |
| 平均 | 1411.748 |
| 標準差 | 165.8227 |
| Best | 1351.065 |
| Worst | 1882.557 |

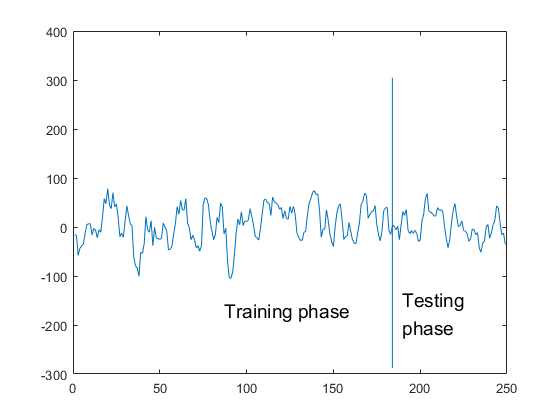


(a)



(b)

1. 實際數值和模型輸出(a)DJIA (2003年) (b)TAIEX (2003年)。x軸為交易日，y軸為股票價格，可看出訓練和測試階段皆有著不錯的效果。(實驗五)



1. 預測誤差。誤差呈現亂數狀態，範圍介於-100至100，代表模型預測能力穩定。(實驗五)

效能比較(DJIA, 實驗五)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | RMSE | | | |
| Method | **2001** | **2002** | **2003** | **2004** |
| SVR (two models, each with single output) [23] | 101.44 | 117.95 | 82.76 | 71.49 |
| ANFIS (two models, each with single output) [23] | 105.56 | 111.69 | 72.09 | 68.00 |
| ANFIS (one model with two outputs) [23] | 128.20 | 142.05 | 90.37 | 83.69 |
| RBF (two models, each with single output) [23] | 106.33 | 131.24 | 97.58 | 81.79 |
| RBF (one model with two outputs) [23] | 181.79 | 136.28 | 154.14 | 148.11 |
| CNFS(5)-ARIMA (one model with two outputs) [23] | 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 |

效能比較(TAIEX, 實驗五)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | RMSE | | | |
| Method | **2001** | **2002** | **2003** | **2004** |
| SVR (two models, each with single output) [23] | 162.46 | 67.72 | 59.47 | 58.81 |
| ANFIS (two models, each with single output) [23] | 147.36 | 70.17 | 72.61 | 65.33 |
| ANFIS (one model with two outputs) [23] | 151.62 | 78.27 | 81.69 | 70.54 |
| RBF (two models, each with single output) [23] | 134.32 | 65.15 | 60.41 | 102.86 |
| RBF (one model with two outputs) [23] | 137.58 | 78.54 | 115.92 | 126.48 |
| CNFS(5)-ARIMA (one model with two outputs) [23] | 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 6—Quadruple Time Series of Daily Dow Jones Industrial Average Index

在這個實驗中，我們也是使用真實世界的時間序列數據來驗證模型的效能，與實驗一不同的是，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實驗一目標曲線的相似度高。使用的資料為印尼雅加達綜合指數 (JKSE)以及恆生指數 (HSI)，資料區間為2000年8月1日至2015年12月31日，其中2000年8月1日至2009年12月31日當作訓練資料，剩餘的當作測試資料。每組原始資料會做一次差分，並以2組差分資料取出30個特徵，共有60個特徵。第1至30個特徵為JLSE收盤價，第31至60個特徵為HSI收盤價，這60個特徵與目標形成資料矩陣，資料矩陣中以HSI距離目標最近，目標排序為JKSE收盤價、HSI收盤價。資料矩陣經過多目標特徵挑選[22]後，將選出的特徵作為模型輸入，複數型態目標有兩個，第一組複數型態的目標實部部分為JKSE收盤價，虛部部分為HSI收盤價。結構學習部分，則將每個被挑選的特徵，用減數分群演算法[11]分群。並透過第二章所介紹的神經元挑選方法，減少神經元數目。整體模型在結構學習後的參數，如表V-表VIII所示。PSO–RLSE混合方法的機器學習參數設定，如表IX所示。

實驗六模型設定

|  |  |
| --- | --- |
| **Parameter** | **Value** |
| Feature variables as model inputs | {} |
| Number of input fuzzy sets | {3, 3, 3, 3} |
| Type of fuzzy sets | SCFS |
| Number of complex–valued targets\* | 1 |
| Number of neurons | 11 |
| Number of parameters in the SCFS layer | 56 |
| Number of aim object neurons | 3 |
| Number of T–S neurons | 3 |
| Number of parameters in the T–S layer | 15 |

\* Each complex-valued target whose real and imaginary parts contain two real-valued targets, respectively.

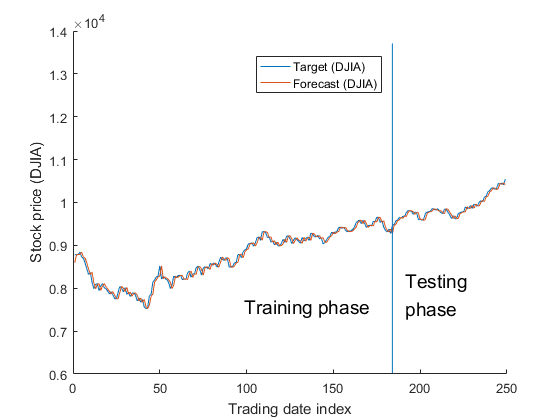
機器學習參數設定

|  |  |
| --- | --- |
| **PSO** | |
| Swarm size | 64 |
| Iterations | 100 |
|  | {0.8, 2.0, 2.0} |
|  | Random in [0,1] |
| Initial particle positions | By SC algorithm in section II-C |
| Initial particle velocities | 0 |
| **RLSE** | |
|  |  |
|  | By number of parameters in the T–S layer |
|  | -by-1 zero vector |
|  | **I** |
| **I** | -by-15 identify matrix |

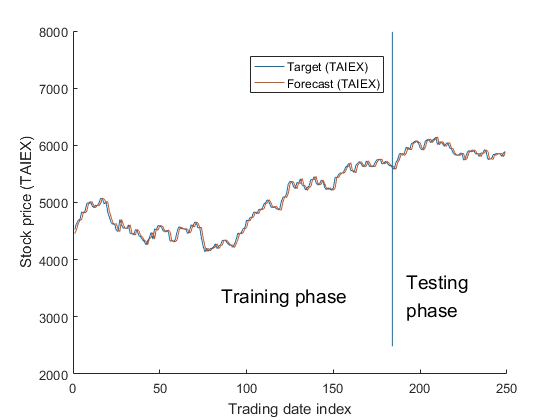
本次實驗的結果將與其他文獻[23]所提的方法做比較。其中除了SVR之外的模型，皆可以同時預測兩個實數目標。所以我們將使用模型的第一組複數型態的輸出與其他論文做比較，結果如表XI和表XII所示，十次重複實驗效能統計如表X所示。目標與模型輸出的結果，如Fig. 8所示，有著不錯的效果；預測誤差如Fig. 9所示，誤差呈現亂數狀態，範圍介於-100至100，代表模型預測能力穩定。

十次重複實驗效能統計(實驗六)

|  |  |
| --- | --- |
| Trials | Performance (RMSE) |
| 1 | 0.119278544167292 |
| 2 | 0.118560195333155 |
| 3 | 0.120291136485088 |
| 4 | 0.111073996890621 |
| 5 | 0.116487533405275 |
| 6 | 0.120543144929471 |
| 7 | 0.116874567908697 |
| 8 | 0.119626311897708 |
| 9 | 0.115248262340888 |
| 10 | 0.114752783085563 |
| 平均 | 0.117274 |
| 標準差 | 0.002994 |
| Best | 0.111074 |
| Worst | 0.120543 |

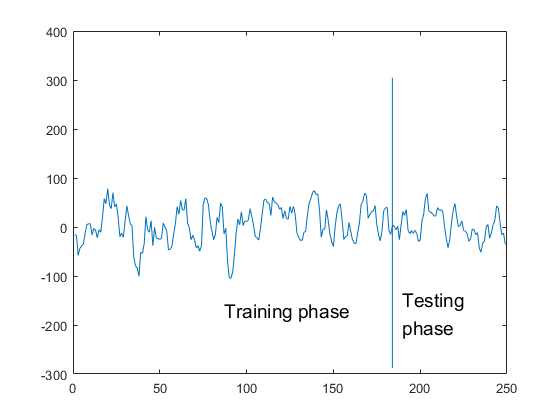


(a)



(b)

1. 實際數值和模型輸出(a)DJIA (2003年) (b)TAIEX (2003年)。x軸為交易日，y軸為股票價格，可看出訓練和測試階段皆有著不錯的效果。(實驗六)



1. 預測誤差。誤差呈現亂數狀態，範圍介於-100至100，代表模型預測能力穩定。(實驗六)

效能比較(DJIA, 實驗六)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | RMSE | | | |
| Method | **2001** | **2002** | **2003** | **2004** |
| SVR (two models, each with single output) [23] | 101.44 | 117.95 | 82.76 | 71.49 |
| ANFIS (two models, each with single output) [23] | 105.56 | 111.69 | 72.09 | 68.00 |
| ANFIS (one model with two outputs) [23] | 128.20 | 142.05 | 90.37 | 83.69 |
| RBF (two models, each with single output) [23] | 106.33 | 131.24 | 97.58 | 81.79 |
| RBF (one model with two outputs) [23] | 181.79 | 136.28 | 154.14 | 148.11 |
| CNFS(5)-ARIMA (one model with two outputs) [23] | 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 |

效能比較(TAIEX, 實驗六)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | RMSE | | | |
| Method | **2001** | **2002** | **2003** | **2004** |
| SVR (two models, each with single output) [23] | 162.46 | 67.72 | 59.47 | 58.81 |
| ANFIS (two models, each with single output) [23] | 147.36 | 70.17 | 72.61 | 65.33 |
| ANFIS (one model with two outputs) [23] | 151.62 | 78.27 | 81.69 | 70.54 |
| RBF (two models, each with single output) [23] | 134.32 | 65.15 | 60.41 | 102.86 |
| RBF (one model with two outputs) [23] | 137.58 | 78.54 | 115.92 | 126.48 |
| CNFS(5)-ARIMA (one model with two outputs) [23] | 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 |