國立政治大學理學院應用物理研究所

碩士論文

Graduate Institute of Applied Physics

College of Science

National Chengchi University

Master Thesis

基於 EEMD 與類神經網路方法進行台指期貨高頻交易研究

A Study of TAIEX Futures High-frequency Trading by using EEMD-based Neural Network Learning Paradigms

黃仕豪

Sven Shih-hao Huang

指導教授：蕭又新博士

Advisor: Yuo-Hsien Shiau, Ph.D.

中華民國 一零三 年 七 月

July, 2014

國立政治大學
應用物理研究所

碩士論文

基於EEMD與類神經網路方法進行台指期貨高頻交易研究

黃仕豪 撰

一零三
七

國立政治大學理學院應用物理研究所

黃仕豪君所撰之碩士學位論文

基於 EEMD 與類神經網路方法進行台指期貨高頻交易研究

A Study of TAIEX Futures High-frequency Trading by using EEMD-based Neural Network Learning Paradigms

業經本委員會審議通過

論文考試委員會委員：

蕭又新

指導教授： 蕭又新

中華民國一〇三年七月二十八日

Acknowledgments

本篇論文的完成，代表了三年碩士生涯的告一段落。在這段時間裡，因著許多師長、學長姐、同學們在各方面的幫助，使得論文能夠順利完成。

首先要感謝的是我的指導老師蕭又新老師，老師在研究與生涯規劃上給我很大的彈性，使我在經濟/金融物理這個跨領域研究學門上得以自由發揮探索，充實跨領域研究所需的知識與技術。老師在專業上的指導使我在邏輯理論上更趨紮實，老師豐富的跨領域研究經驗不只使我看見跨領域研究的趣味所在，更是奠定本論文學術內涵的基石，是本論文得以完成的最關鍵原因。

感謝在復旦大學交換期間東京大學陳昱老師指導的期權定價研究，雖然研究主題與本論文無直接相關，但透過此次研究，陳老師讓我以更全觀有系統的方式以科學人的視角來看金融領域，並進一步了解科學研究者在金融量化研究領域的可能性。

感謝復旦大學魏心源老師提供 Cluster 運算資源，雖權限不大但讓我對 Matlab® 在 Linux Cluster 上的佈署與操作有更進一步的經驗與認識。

感謝中央大學數據研究中心的王淵弘老師對於 Fast EEMD 方法上的協助指導，使本論文在運算效率上獲得大幅的提升，是得以成功處理大型高頻數據的關鍵之一。

感謝第一證券 (香港) 的陳思聰先生，在數據資料的收集上陳先生給予許多實質有效的建議與幫助。

感謝謝佳宏學長在基本研究方法的指導與鞭策，讓論文的進度與架

構方法能更趨完整。

感謝陳原孝學長以其在證券市場豐富的實務操作經驗提供許多中肯的意見與想法，對本論文的研究方向有相當大的幫助。

感謝同學徐哲仁，哲仁在 IT 技術與 LaTeX 方面給我相當大的幫助，並在撰寫的最後階段彼此惕勵打氣，使得本論文得以順利撰寫完成。

感謝同學林丁順，不論是在上海期間還是回到政大後，丁順在行政方面給了我許多幫助，讓我在準備 CFA® 考試與論文撰寫上都沒有後顧之憂得以全力衝刺。

感謝陳敏霞助理，敏霞姐在所上許多事務適時的提醒與關心使我們不治錯失許多行政上的重要時程。敏霞姐像媽媽一樣把所上打理的井井有條，使大家能夠專心的在研究與學習上。

感謝同學帥文與學弟妹帥舞昱君在撰寫論文時的一同打拼與鼓勵。

更要感謝我的父母親與家人，除了三年來經濟與親情上的支持, 我所想完成的目標他們在能力所及都給予支持鼓勵和建議。沒有家人做為我的後盾我將無法順利完成碩士生涯，這篇論文同時也獻給我的父母。

在碩士三年生涯之中，接觸了許多人事物，不論是不是直接與學業研究相關，因著你們的幫助與分享，使我不論在知識、研究、思考或者生命身量上都有很大的進步。這篇論文仍不夠成熟，但透過這篇論文不論是技能或是思考概念上我學習到許多有價值的東西，它也為我的碩士生涯做了階段性的註腳。最後我要感謝神在碩士三年中，祂讓我遇到大大小小的挑戰，在開發模型時遇到的瓶頸、在苦於硬體限制於法提升運算效能時、在同時準備衝刺 CFA® 與論文時看不到的盡頭以及其他不足道的困難，都因著祂的恩典與賜下的智慧使我能突破重重難關。三年前祂領我進到政大，三年後祂也領著我從這裡畢業。在碩士生涯，如同使徒保羅所說：那美好的仗我已經打過了，當跑的路我已經跑盡了。[1]接下來的兵役與求職仍充滿種種不確定，我只有忘記背後，努力面前，向著標竿直跑[2]，神必帶領前面的路。

[1] 聖經提摩太後書 4:7

[2] 聖經腓立比書 3:13-14

中文摘要

金融市場是個變化莫測的環境，看似隨機，在隨機中卻隱藏著某些特性與關係。不論是自然現象中的氣象預測或是金融領域中對下一時刻價格的預測, 都有相似的複雜性。時間序列的預測一直都是許多領域中重要的項目之一, 金融時間序列的預測也不例外。在本論文中我們針對金融時間序列的非線性與非穩態關係引入類神經網路 (ANNs) 與集合經驗模態分解法 (EEMD), 藉由 ANNs 處理非線性問題的能力與 EEMD 處理時間序列信號的優點，並進一步與傳統上使用於金融時間序列分析的自回歸滑動平均模型 (ARMA) 進行複合式的模型建構，引入燭型圖概念嘗試進行高頻下的台指期貨 TAIEX 交易。在不計交易成本的績效測試下本研究的高頻交易模型有突出的績效，證明以 ANNs、EEMD 方法與 ARMA 組成的混合式模型在高頻時間尺度交易下有相當的發展潛力，具有進一步發展的價值。在處理高頻時間尺度下所產生的大型數據方面，引入平行運算架構 SPMD(single program, multiple data) 以增進其處理大型資料下的運算效率。本研究亦透過分析高頻時間尺度的本質模態函數 (IMFs) 探討在高頻尺度下影響台指期貨價格的因素。

關鍵字：類神經網路方法、燭型圖 (K 線圖)、自回歸滑動平均模型、集合經驗模態分解法、高頻交易、平行運算、時間序列分析、大型數據處理

Abstract

Financial market is complex, unstable and non-linear system, it looks like have some principle but the principle usually have exception. The forecasting of time series always an issue in several field include finance. In this thesis we propose several version of hybrid models for candlestick charts forecast, they combine Ensemble Empirical Mode Decomposition (EEMD), Back-Propagation Neural Networks(BPNN) and Autoregressive Moving Average(ARMA) model, try to improve the forecast performance of financial time series forecast. We also found the physical means or impact factors of IMFs under high-frequency time-scale. For processing the massive data generated by high-frequency time-scale, we pull in the concept of big data processing, adopt parallel computing method ”single program, multiple data (SPMD)” to construct the model improve the computing performance. As the result of backtesting, we prove the enhanced hybrid models we proposed outperform the standard EEMD-BPNN model and obtain a good performance. It shows adopt ANN, EEMD and ARMA in the hybrid model configure for high-frequency trading modelling is effective and it have the potential of development.

Key words: Artificial Neural Networks, Candlestick Charts, Autoregressive Moving Average model, Ensemble Empirical Mode Decomposition, High-Frequency Trading, Parallel Computing, Time series analysis, Big Data Processing

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Overview of The Development Track of Quantitative Analysis, Econophysics and High-Frequency Trading

Fischer Sheffey Black (January 11, 1938 – August 30, 1995) was an American economist, he is one of the authors of the famous Black–Scholes equation [2], he also the one of first pull in the quantitative analysis method to economic field, it impact markets and investment banks deeply. Black is not a traditional economist, he had highly interest in different field. When he started his college life in Harvard University, he study across social science, biology, chemistry and physics, Black began his graduate student life in the Physics Institute of Harvard University in 1959, after one year he transferred from the Physics Institute to the Applied Mathematics Institute. But he expulsion from Harvard University because of his interest is too wide to definite his PhD thesis proposal. Then he join a high tech consulting firm and develop artificial intelligence program. Finally, he backed to Harvard University and got his PhD degree in 1964. Similar background with Black's colleague Jack Treynor in Arthur D. Little, Treynor major in Mathematics and enter the Commerce College of Harvard University [3]. Treynor join Merrill Lynch in 1966 after work in Arthur D. Little.

Not just Black and Treynor, many researchers started proposed quantitative method or applied quantitative method in financial market field in 1960’s, such as Edward O. Thorp. He got an Mathematics PhD and Physics master degree, especially in Quantum Mechanics. He was a pioneer in modern applications of probability theory. He developed and applied effective hedge fund techniques in the financial markets, and collaborated with Claude Shannon in creating the first wearable computerc̃iteintro3.

Even quantitative method already used in financial field in 1960’s, using mathematical model and physical concept in financial field still not a well-known field in population, this field generally just known in related Econometrics or investment banks in Wall Street.

In 1990’s, Harry Eugene Stanley proposed ”Econophysics” describe the research in stock market, corporation growth and other economic or financial problem by physicists [4]. Econophysics use methods, models, theories which relate with physics to make solution of Economical and Financial problem. It also called ”Financial Physics”. The concept of econophysics is already exist, such as pull-in the Brownian motion in financial field and become the random walk theory in finance field. One reasons of propose Econophysics is the theoretical economics become more and more purely mathematics, it cause some problem cannot be explain, such as ”fat tail” phenomenon in the distribution of stock profitability, it cannot be explain by standard financial theory. In the standard financial theory it shall be normal distribution, and complex system of Physics have similar characteristic with fat tail. Actually financial system is also a kind of complex system, this is the second reason of propose econophysics. Although economic usually be known as one of social sciences, but many characteristics in economic and finance are similar with nature science. For instance, fractal is a natural phenomenon and a mathematical set, we can see this phenomenon in nature environment. Romanesco broccoli, for example, showing its self-similar form approximating a natural fractal. Not just in nature environment, in stock market we also can see this phenomenon, in stock price, fractal concept already applied in ”Elliott wave principle” [5] proposed by Ralph Nelson Elliott in his book <The Wave Principle> [6] in 1938. Figure1.1 demonstrate the multifractal simulate stock price generated by the modified model from Michael Hanchak(2010). Fractal is one of basic concept for high-frequency trading, we will mention it in following section. Another example is the uncertainty principle of Quantum Physics also reflect in financial market in a certain extent [7]. Before econophysics proposed, using mathematical and physical knowledge in financial filed majority for investment and related research, now also focus on the relationship and behaviour in financial and economical field, or observe the characteristic of financial market in scientific view.

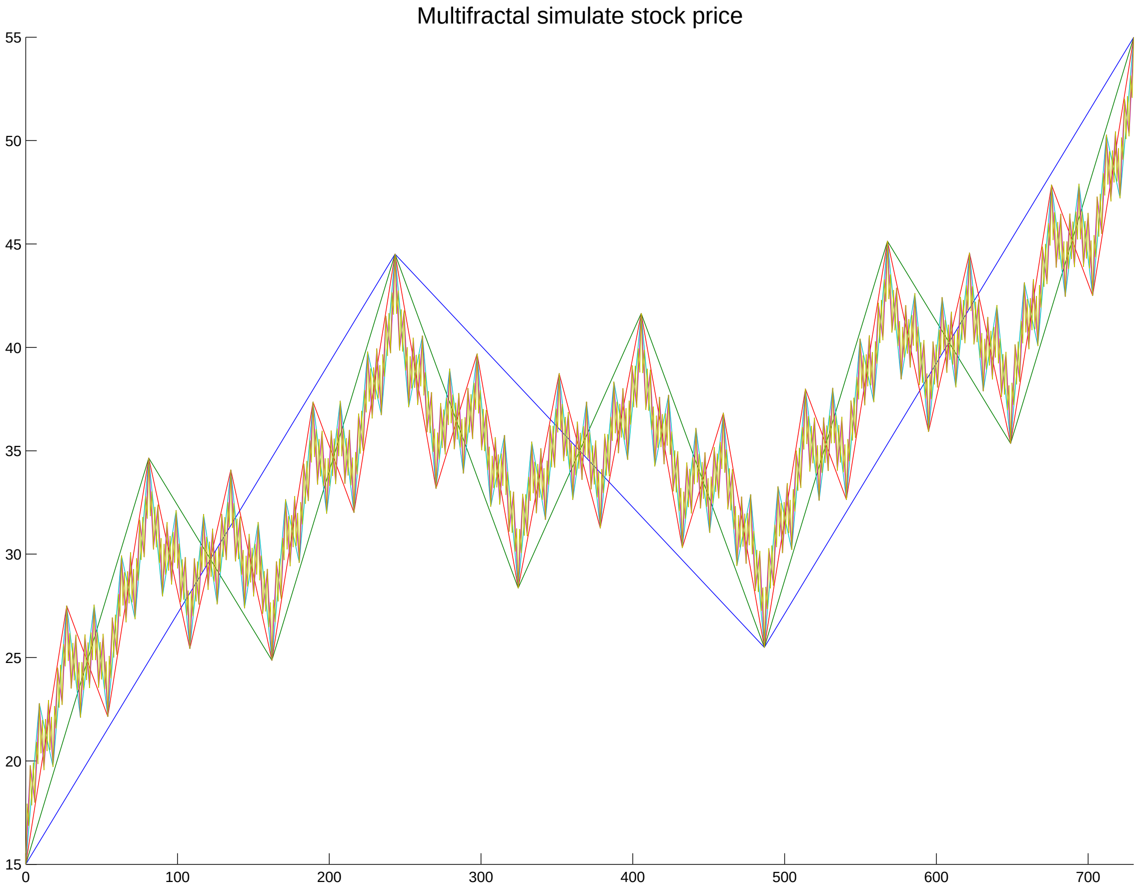


Figure 1.1: A multifractal simulate stock price [1]

Algorithmic Trading become more and more popular because the universalness of electronic trading(As figure1.3 [8]). Algorithmic Trading can trading by setting rules without sentiment delay or mistake. It developed for use buy-side to manage orders and to reduce the impact to market by optimizing trade. High-frequency trading is a major branch of algorithmic trading, it focus on trading in short time scale. Figure1.3 compare the different time scale of different strategies. High-frequency trading become well known

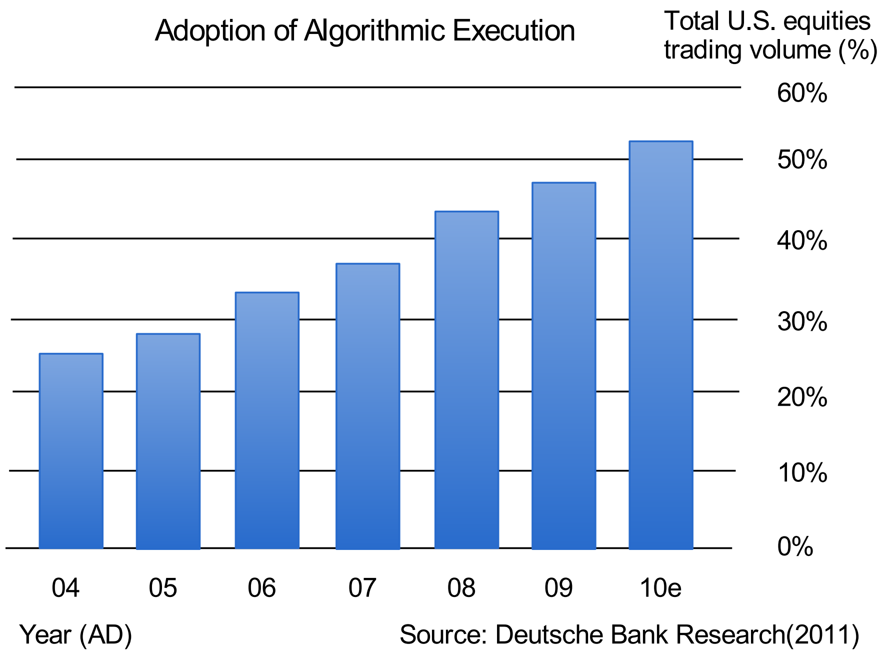
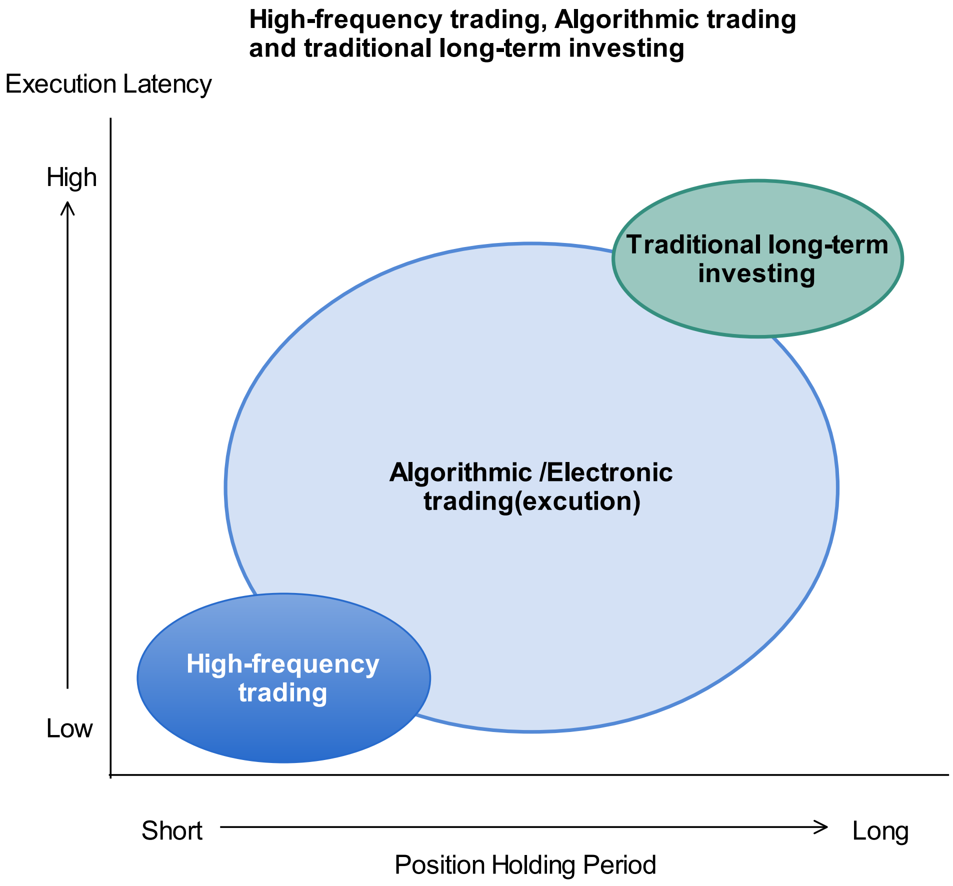


Figure 1.2: Growth of Algorithmic Trading

in recent decades, especially after financial crisis in 2008 [9]. In 2008, James Harris ”Jim” Simons of Renaissance Technologies Corp. is the highest investment return fund manager, he is also the founder of Renaissance Technologies Corp. Simons is a famous mathematician, introduced the famous Chern-Simons theory with famous Chinese-born American mathematician Shiing-Shen Chern [10]. Simons’ company don’t hire traditional finance analysts or MBAs, he hire scientists from mathematics or physics and other different scientific research field. They use scientific knowledge to analyze financial market and build trading model. In Wall Street they called ”Quant”, as the most successful quant, Simons is known as the ”Quant King”.

In quantitative investment, high-frequency trading is a major part and it also a kind of algorithmic trading method, it can be realized because traditional trading process replaced by electronic trading, reduce the trading cost and increase the trading efficiency. Ideal high frequency trading strategy catch the inefficiency of market in short time scale no matter the market environment is good or not. This is the first challenge for high frequency trading [11]. The second challenge is processing big data. Shorter time-scale means more data


High-frequency trading, Algorithmic trading and traditional long-term investing
Execution Latency
High
Low
Traditional long-term investing
Algorithmic /Electronic trading(excution)
High-frequency trading
Short
Long
Position Holding Period

Source: Source: Deutsche Bank Research(2011)

Figure 1.3: Compare the different time scale trading strategies

we get. Big data also a well-known issue recently. Similar as other big data analysis, high frequency trading also need to processing dig data during modelling and test [12]. More addition, when execute trading, high-frequency trading model have to process real-time data, it need highly computing speed for processing single data unit. To make a long story short, high-frequency trading not only highly require the IT programming and hard ware performance, but also need the knowledge across finance and science.

1.2 EEMD and ANN in Forecasting

Stock market forecast always been an issues for financial market. Like the same property of the complex system in physics, financial market is a complex, evolutionary, and non-linear dynamical system (Abu-mostafa and Atiya, 1996) [13]. The challenges of financial market forecasting are non-stationary, intensity, noise, highly uncertainty, etc. Quantitative methods are the crucial method for forecasting in the stock market and improved decisions or investments. It is already been proved about predicting stock data with traditional time series analysis is difficult. Artificial Intelligence we mentioned in last section, or Machine Learning provide different ways to analyze the stock market and Artificial Neural Networks (ANNs) is one of Machine Learning/Artificial Intelligence. [14] Primarily, ANNs don't need any assumption prior to forecasting, it has the ability to extract useful information from large sets of data, which is often required for a satisfying description of a financial time series. [15]

There is no need to specify a particular model form for ANNs, the model is adaptively formed based on the features presented in the data. Hamid et al. (2004) [16] discussed the ability to foretell the volatility of the markets is critical to analysts. They present a primer for using neural networks for financial forecasting. Huang et al. (1998) [17] proposed the Empirical Mode Decomposition (EMD) method, it can help the prediction more efficient for the economic dilemma. EMD decomposes the original data into several independent and nearly periodic intrinsic modes based on a local characteristic scale, it help us not only discover the characteristics of the data but also understand the underlying rules of reality. Wu and Huang (2004) [18] improved the EMD method by sifting an ensemble of white noise-added signal and treating the mean as the final true result. This new method called Ensemble Empirical Mode Decomposition (EEMD) is finite and not infinitesimal. With the ensemble mean, we can separate scales without any prior subjective criterion in the intermittence test for the original EMD method, it represents a substantial improvement over the original EMD, it is a typical noise-assisted data analysis (NADA) method. Klevecka and Lelis (2008) [19] created a functional algorithm of preprocessing of input data taking into account the specific aspects of teletraffic and properties of neural networks. The preprocessing algorithm for input data of neural networks can use in several real-time series with different lengths representing the intensity of telephone traffic and international outgoing traffic of the IP network. They found a obvious advantage of neural networks: they can work successfully with non-normally distributed data.

En Tzu Li (2011) found EEMD-ANN Network was more proper to use in predicting index and she predicted the future index price. As a result, the FK indicator display a signal of buy or sell, make buy or sell Call-Put decisions of TAIEX options by neural network.

Tsai Yu Ching (2011) [20] applied the Ensemble Empirical Mode Decomposition (EEMD) based Back-propagation Neural Network (BPNN) learning paradigm into two forecasting topics: the electricity consumption per hour in National Chengchi University and the historical daily gold price.He used moving-window method in the prediction process. He also used the ensemble average method compared with the results without applying ensemble average method. Through ensemble average, the outcome was more precise with smaller errors. It resulted from the procedure of finding minimum error function in the BPNN training.

Chen Yuan Hsiao(2013) used the developed the EEMD-BPNN based model, which input different IMFs into the individual BPNN and summate these output as forecast result. He use moving average indicator to build trading model and got a good performance in back test in 2010 TAIEX. He also found the physical means of monthly IMFs.

We look back some important develop track of EEMD and ANNs above. In this thesis, we develop the model form Chen Yuan Hsiao (2013), construct different EEMD-based ANN combine to traditional ARMA algorithm for per minute high-frequency data of TAIEX futures and pull in the concept of candlestick to build the trading model and test the returns of TAIEX futures and find out the physical means or the impact indicators of the per minute IMFs.

Chapter 2

Methodology

2.1 Empirical Mode Decomposition (EMD)

The empirical mode decomposition (EMD, Huang et al, 1998) is a intuitive, direct and adaptive method. Compared with other traditional analysis method, EMD can use to analysis non-linear and non-stationary time-series data, such as use in seismology or meteorology, etc. This decomposition has an assumption that the data you want to decompose must be consisted from different intrinsic mode of oscillations. Each of intrinsic mode that represents an oscillation, we called them ”intrinsic mode function (IMF)”. IMF have two conditions must be followed:

1. The same number of extreme and zero-crossings, or differ at most by one
2. The mean value of the defined upper and lower envelopes will be symmetric with respect to the local mean.

In general, we can acquires several IMFs from original raw data series through the following process:

1. Find all the local maxima and minima points included in the time series x_t that we want to decompose
2. Connect all the local maxima points by cubic spline interpolation as the upper envelops $e_{t,\max}$ and generate the lower envelops $e_{t,\min}$ as the same way

3. Calculate the mean $m_{t,1}$ of upper and lower envelopes

$$m_{t,1} = \frac{e_{t,\max} + e_{t,\min}}{2} \tag{2.1}$$

4. Calculate the difference between original time series x_t and the mean $m_{t,1}$ as $h_{t,1}$:

$$h_{t,1} = x_t - m_{t,1} \tag{2.2}$$

5. If $h_{t,1}$ still doesn't satisfy the request condition of IMF, repeat the process $1.4.K$times, until $h_{t,k}$ achieve the criterion :

$$\begin{aligned} & h_{t,1} = x_t - m_{t,1} \\ & h_{t,2} = h_{t,1} - m_{t,2} \\ & h_{t,3} = h_{t,2} - m_{t,3} \\ & h_{t,k} = h_{t,k-1} \\ & \Longrightarrow 0.1 \leq SD = \sum_{t=1}^{T} \frac{(h_{t,k-1} - h_{t,k})^2}{h_{t,k-1}^2} \leq 0.3 \end{aligned} \tag{2.3}$$

We typically setting the SD value in the range $0.1 \sim 0.3$. Once $h_{t,k}$ achieve the criterion, it will be classified as the IMFs, the first one denote as $IMF1(c_{t,1})$ and the second one denote as $IMF2(c_{t,2})$, and so on. Until the raw data cannot decompose IMFs anymore, the raw time series can be expressed as

$$x_t = \sum_{t=1}^{n} C_{t,i} + r_{t,n} \tag{2.4}$$

Where n is the number of IMFs and $r_{t,n}$ is the final residue, it also the trend of x_t. And $c_{t,1}$ presents IMFs that are nearly orthogonal to each other.

After the process above, the raw data decompose to IMFs which represent high to low frequency, and each IMF contains physical mean themselves. Figure2.1 illustrate the process of decompose EMD by flowchart.

2.1.1 Ensemble Empirical Mode Decomposition (EEMD)

EMD provides a useful analysis method for extracting signals, especially the raw data is non-linear and non-stationary data. But even so, EMD still has its defects, such as ”mode mixing” problem. A single IMF may contains of signals with widely disparate scales or similar scale signal belonging multiple IMF components. This phenomenon called ”mode mixing”. When mode mixing occurred, IMF may not present its physical meaning suitably.

To solve this problem, Wu and Huang (2004) modified EMD method, proposed Ensemble Empirical Mode Decomposition (EEMD)method. EEMD provides a method which based on statistical comparison with the white noise, it can determine whether an IMF contains true signals. We know that each observed data are mixed with true time series and noise. Even if data is collected by separate observations with different white noise, the ensemble mean can eliminate the random effect from white noise and the result will close to the true property of the time series. It means we can extract the meaningful signal from data by adding white noise in EMD randomly. In EEMD method, adding white noise provide a uniformly distributed reference scale, then help EMD to overcome the mode mixing phenomenon. The EEMD process are following:

1. Add a white noise series to the raw data.

2. Decompose this white noise-added data into IMFs.

3. Iterative the two steps above, add different white noise every time.

4. Finally we through the ensemble means of corresponding IMFs acquire final result.

Wu and Huang (2004) prove a statistical rule to control the adding white noise effect:

$$\varepsilon_n = \frac{\varepsilon}{\sqrt{N}} \tag{2.5}$$

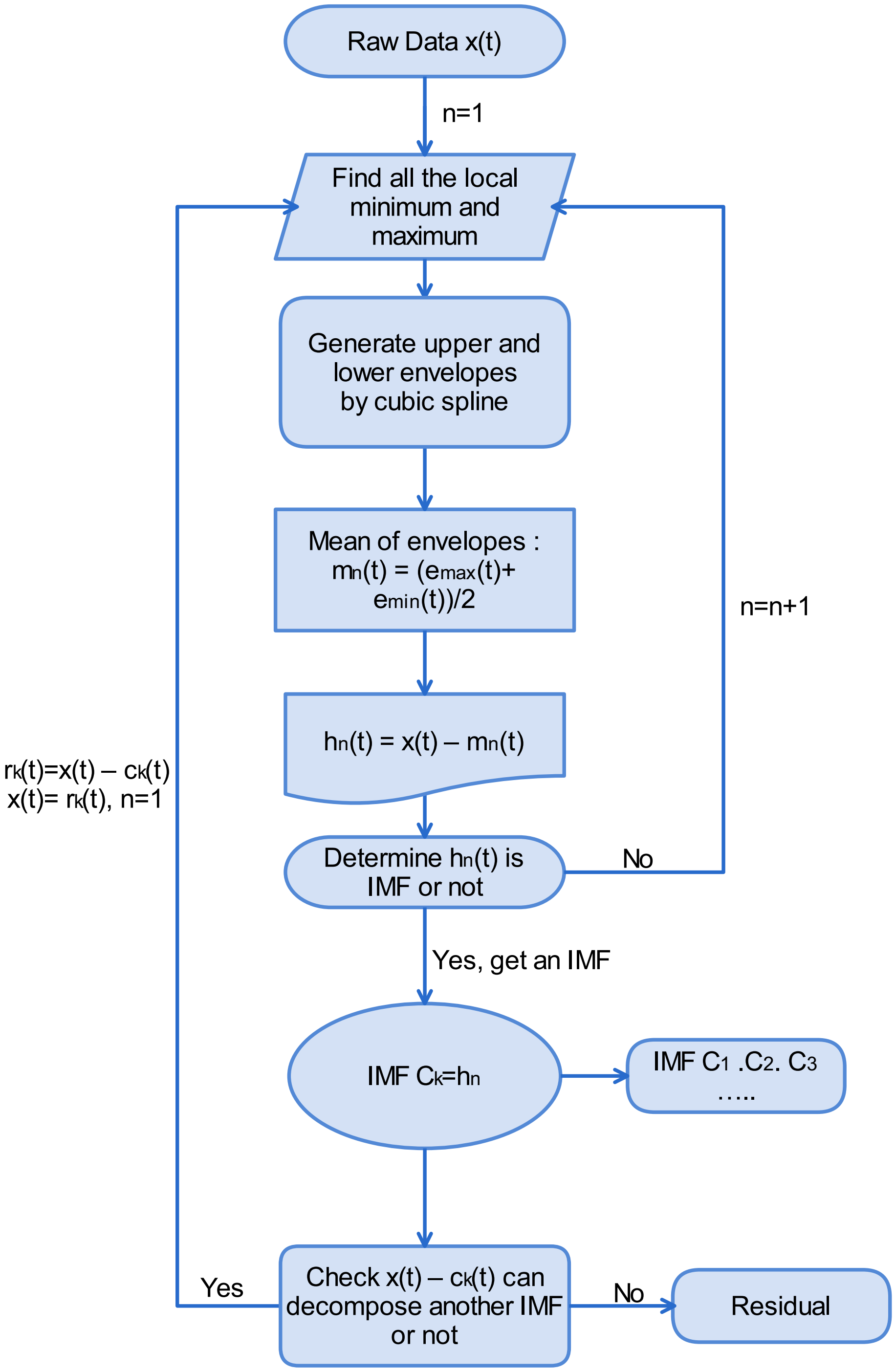

Raw Data x(t)
n=1
Find all the local minimum and maximum
Generate upper and lower envelopes by cubic spline
Mean of envelopes : $m_n(t) = (e_{max}(t) + e_{min}(t))/2$
$h_n(t) = x(t) – m_n(t)$
Determine $h_n(t)$ is IMF or not
No
n=n+1
Yes, get an IMF
IMF $C_k = h_n$
IMF C_1 .C_2. C_3 …..
$r_k(t) = x(t) – c_k(t)$
$x(t) = r_k(t)$, n=1
Yes
Check $x(t) – c_k(t)$ can decompose another IMF or not
No
Residual

Figure 2.1: Flowchart of EMD process

Where n is the number of ensemble members, ε is the amplitude of the added noise and ε_n is the standard deviation error, it measure the difference between input signal and the corresponding IMFs. Empirically, the ensemble number N usually set to 100 and the white noise level may set to 0.1, 0.2 (Zhang et al., 2008)or 0.3 (En Tzu Li, 2011). Adding white noise can makes signals into the comparable scales to collate into one useful IMF, therefore the EEMD reduces the occurrence of mode mixing and it prove it's a effective way of improve the standard EMD performance. In this thesis we adopt the EEMD and modified fast-EEMD subroutine released by Research Center for Adaptive Data Analysis(RCADA) of National Central University(NCU)(Taiwan) as EEMD decompose core [21].

Here we use the multifractal price simulator which we use in figure1.1 test the EEMD method. As simplified , we choose a part of 2nd scale from multifractal series for as figure 2.2. In figure2.3 and figure2.4, each scale of IMFs match every multifractal simulate price turning points, but IMFs smoother than fractal scale (For example the red points in the subfigure). The amplitude in large scale IMFs is smaller than original large scale fractal, because the larger scale fractal series more deviate final simulate price. In large scale component , IMFs are more close to final simulate price.

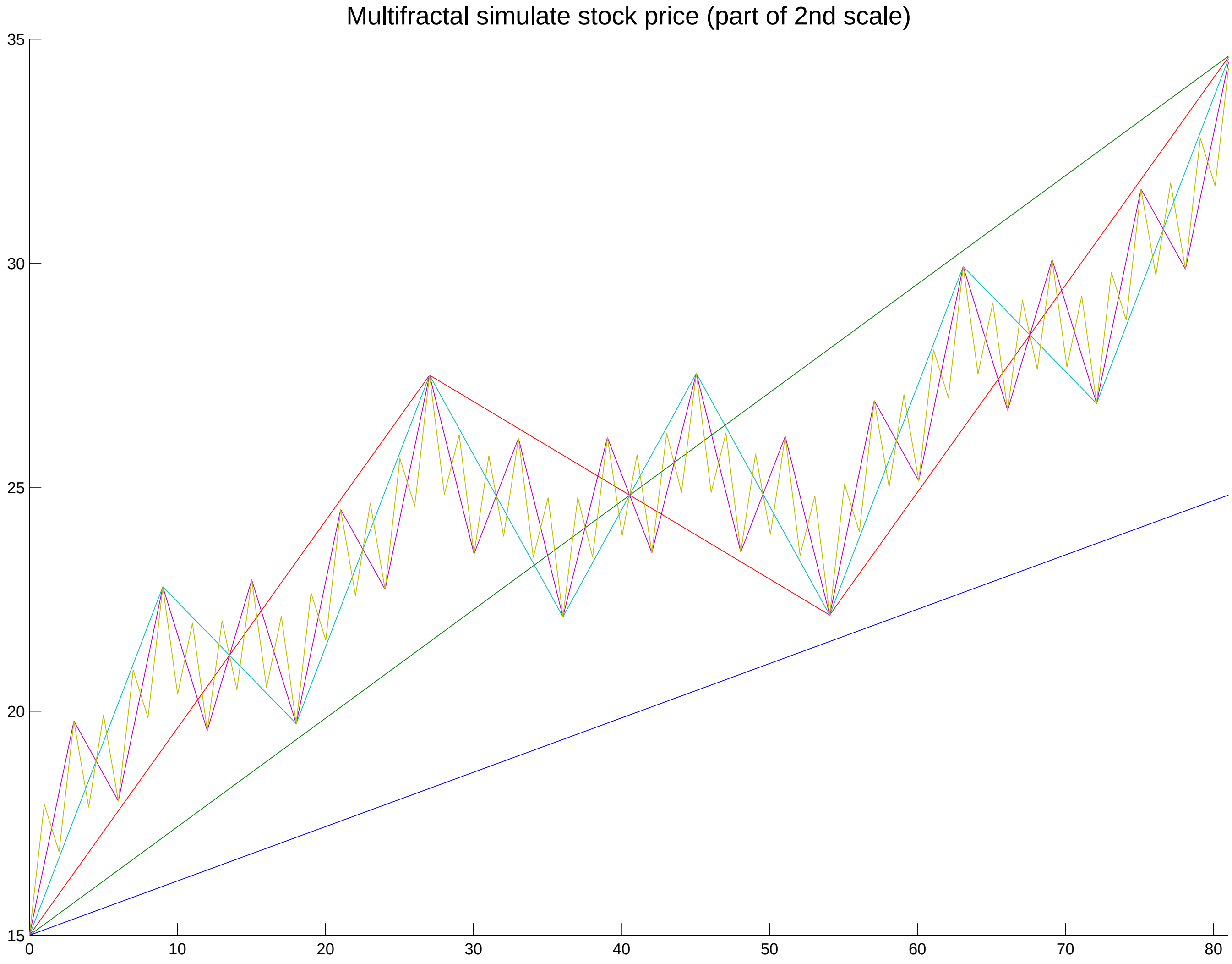


Figure 2.2: A part of multifractal simulate price, from 2nd to 5th scale [1]

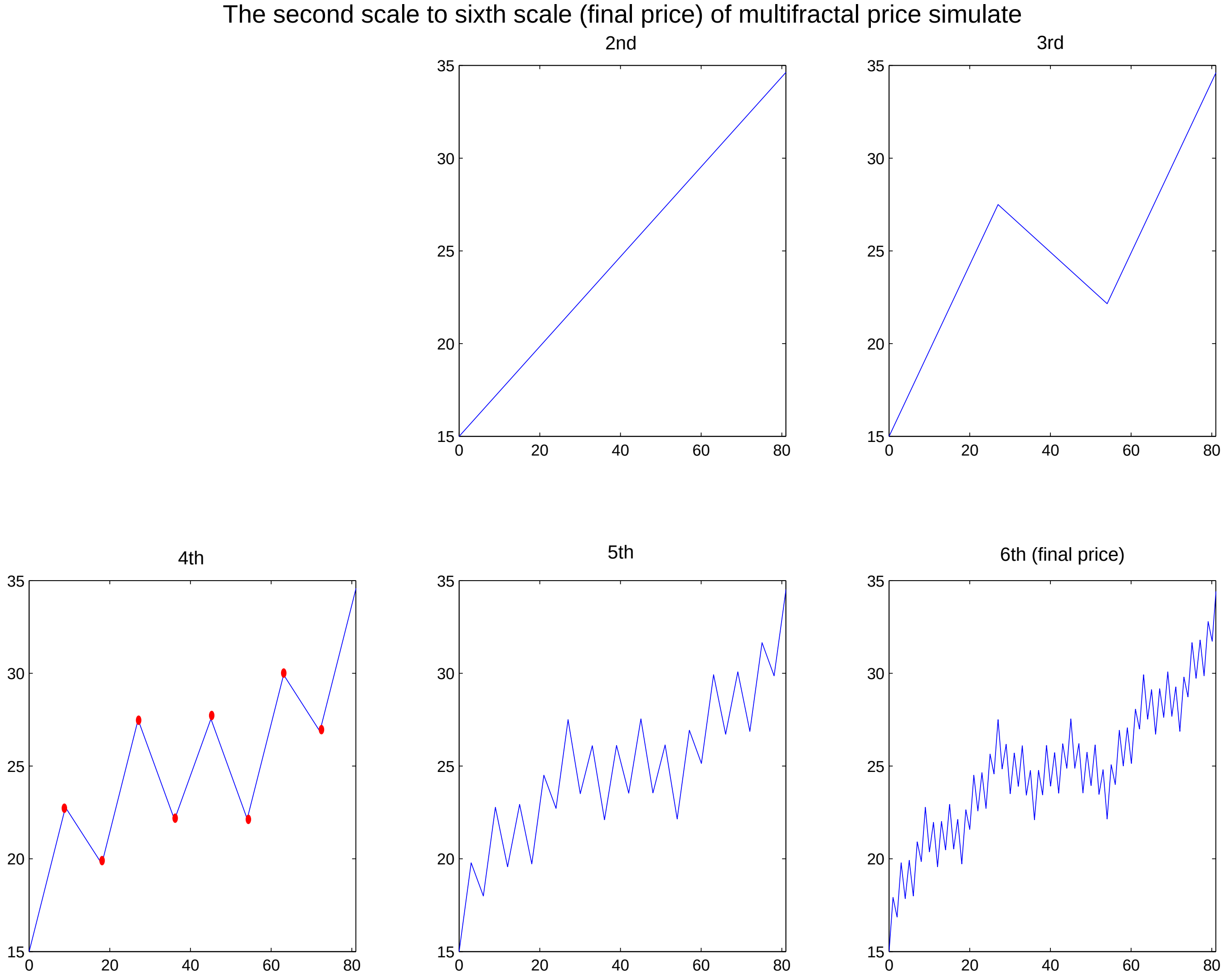


Figure 2.3: The second to sixth scale(finale price) of multifractal simulate price

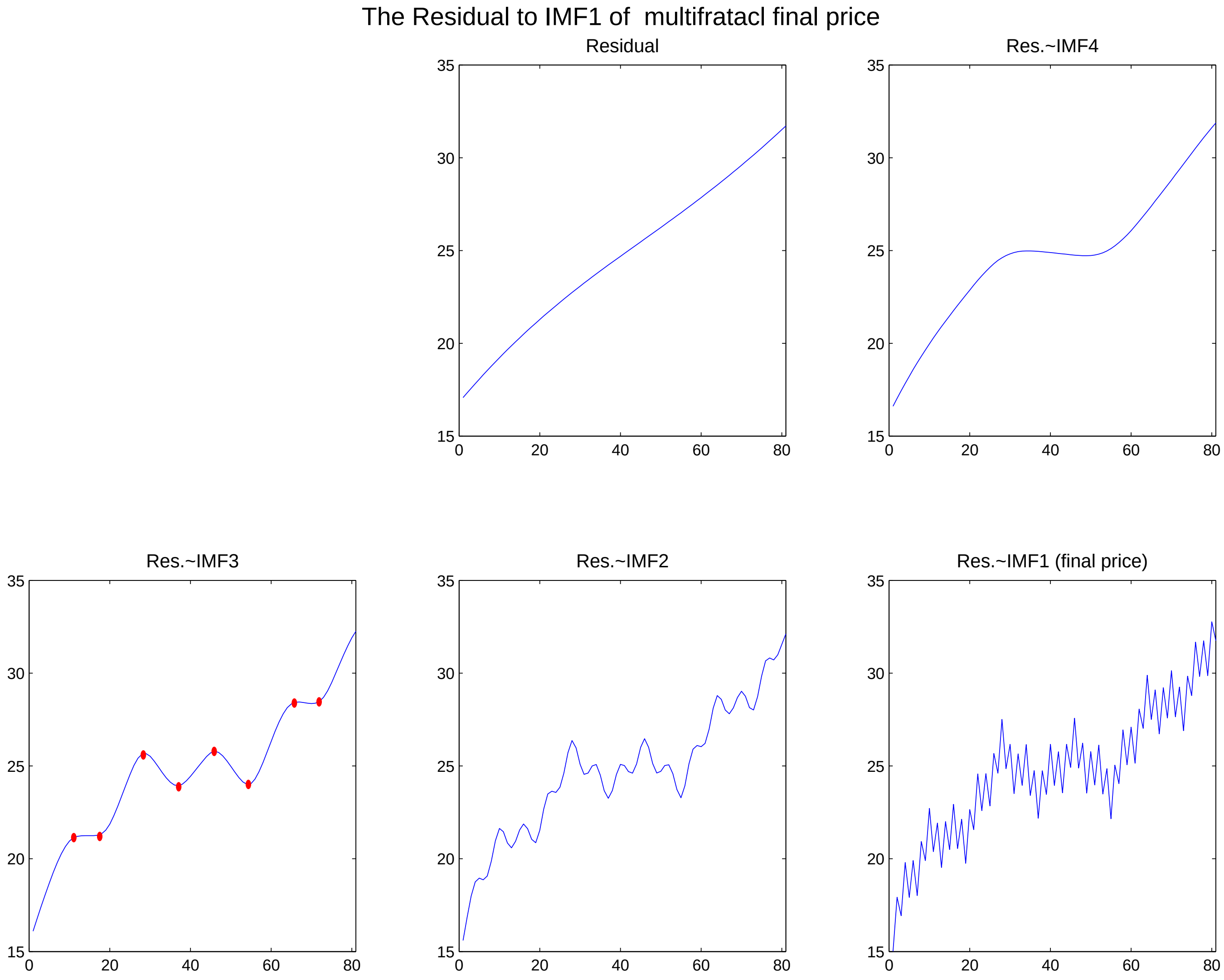


Figure 2.4: The IMFs of the multifractal simulate price

2.2 The Artificial Neural Networks (ANNs)

Artificial Neural Networks (ANNs) is a kind of machine learning theory, it simulate the real neural network configure in order to flexible computing a broad range of non-linear problems, and it contain with several variant. ANNs have been developed over 50 years, the original model called Perceptron was proposed by Frank Rosenblatt in 1957. [22] Until today, feed-forward back-propagation neural network (BPNN) is one of the widely used configure in ANN family, especially in processing time-series data, which we study in this thesis. ANNs are designed to imitate the biological neural system; it can seen as one kind of bionics method [23] Figure2.5 illustrate a typical structure of ANNs.

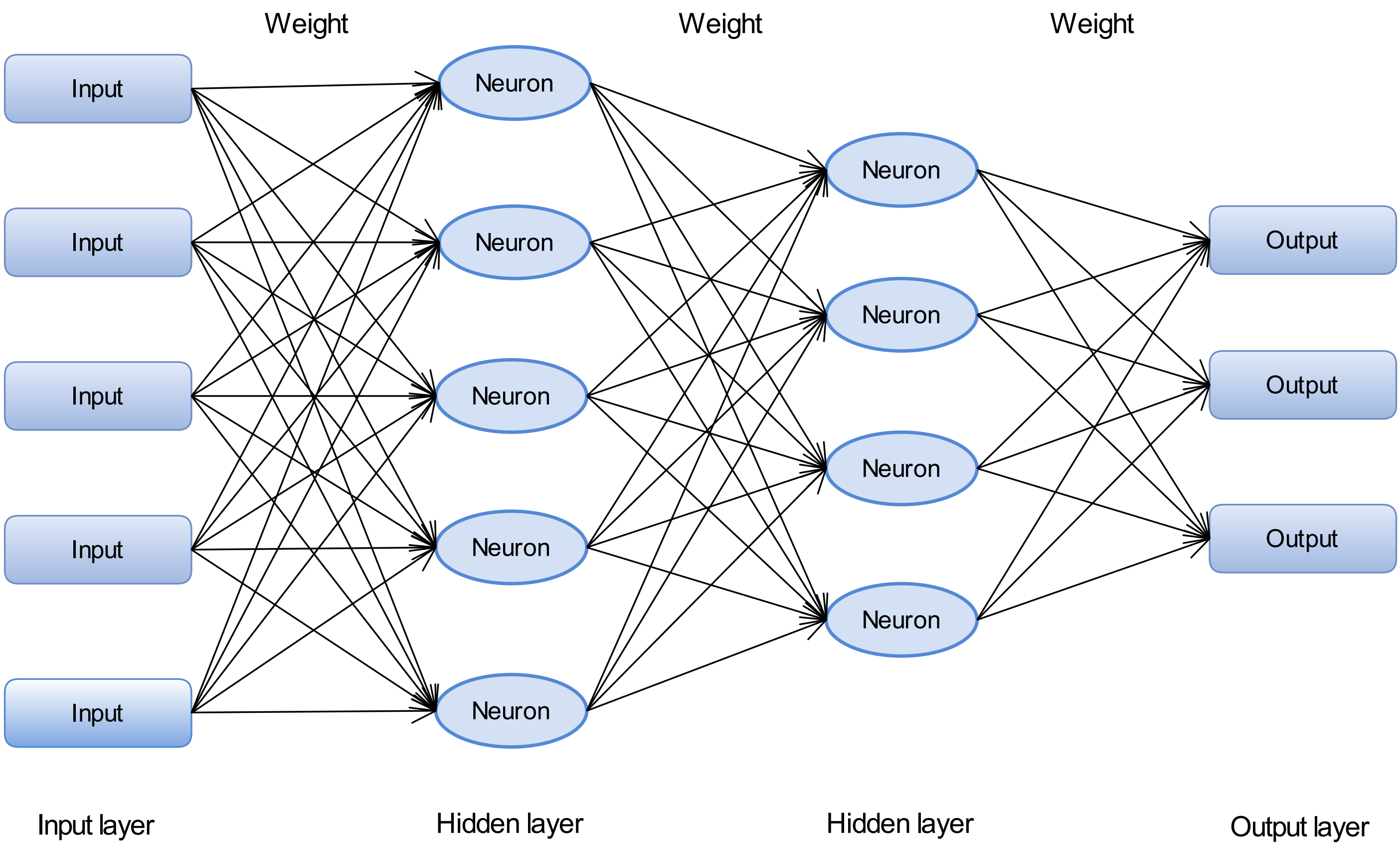


Figure 2.5: A basic structure of typical Artificial Neural Networks

ANN can widely use in some practical application such as forecasting, estimating, diagnosis, recognition etc., it can solve different problem in varied domains. The basic structural concept of ANN can be described as a type of multiple regressions with neurons arranged into layers and the connected neurons classified in different layers according to its connection strength called weights. [24] As a processing element, a neuron received input and generate an output. Following figure exhibit a typical three-layer feed forward neural network, defined three different property layers:

1. Input layer: input data into the network.

2. Hidden layer: creates an internal mapping of the input data.

3. Output layer: receives data from the hidden layer and make output.

Each layer contains different number of neuron. The number of neuron in input or output layer implies the number of input or output variables. The hidden layers provide network the ability of generalize. There are no rules determine the suitable number of hidden layers would given network a better generalization. In conclusion of previous practice, one of the widely used and well performed neural networks configure is one and occasionally two hidden layers. Increasing the number of hidden layers also increases computing time and may cause over-fitting. Over-fitting occurs when the forecasting model has relatively few observations in relation to its parameters and therefore it has the ability to memorize individual points rather than learn the general pattern, it will limit the forecasting performance [25] (I. Kaastra, M. Boyd, 1996).

Normally, if setting too few neurons in the model, it will limit the network ability to correctly mapping the input and output; if setting too many neurons in the model, it may cause the network to memorize trivial patterns, the ability to demonstrate expected features or trends and make an appropriate generalization will be weak. The empiric principle of setting neuron number can range from one and half to twice (Mendelsohn, 1993) the extent of geometric pyramid rule depending on the complexity of the problem.

Transfer function is a important component for neuron units, it is a mathematical formula used to determine the output of neurons. Transfer function have four typical forms:

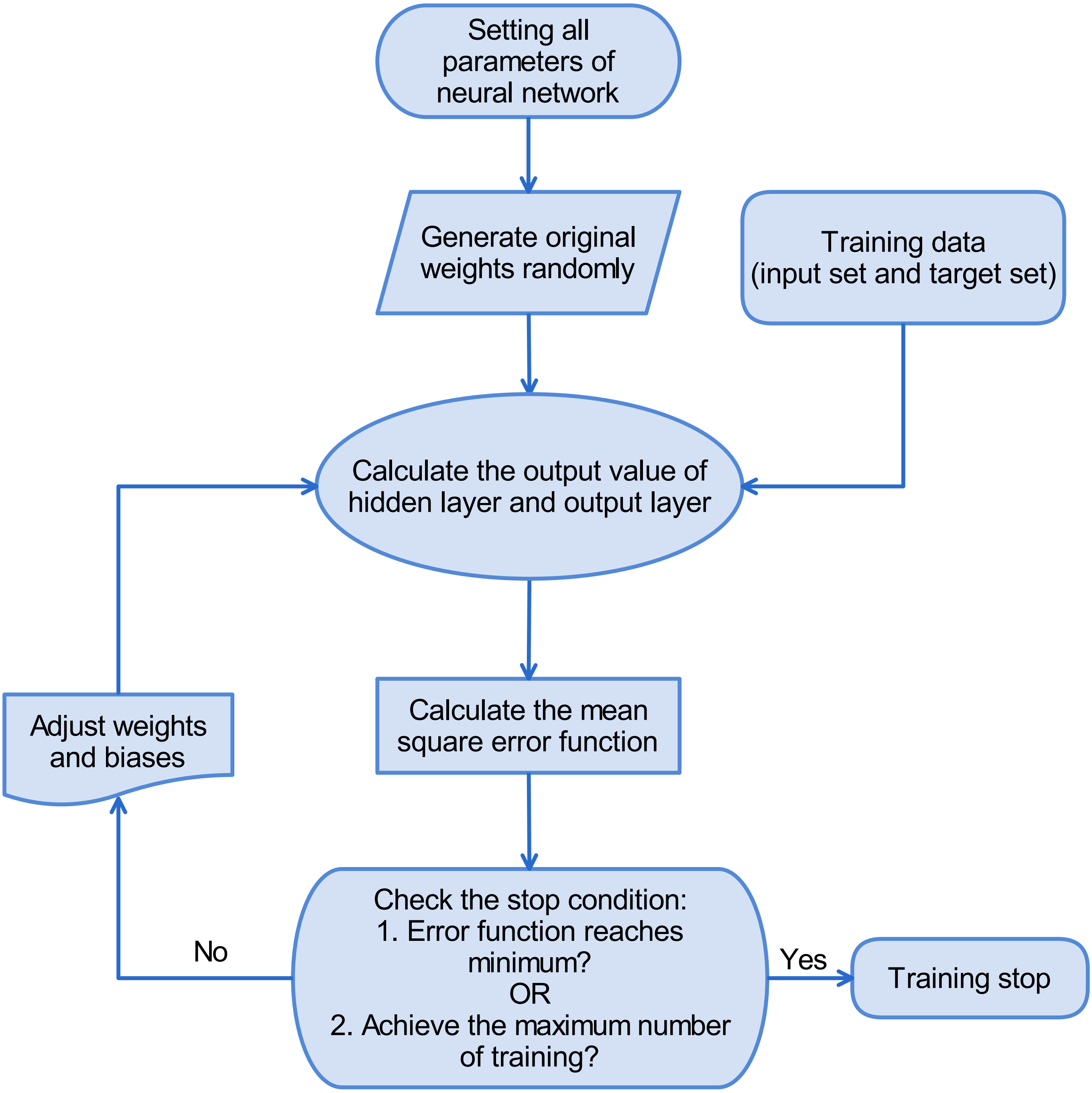

Setting all parameters of neural network
Generate original weights randomly
Training data (input set and target set)
Calculate the output value of hidden layer and output layer
Adjust weights and biases
Calculate the mean square error function
No
Check the stop condition:
1. Error function reaches minimum?
OR
2. Achieve the maximum number of training?
Yes
Training stop

Figure 2.6: Flowchart of ANN training procedure

$linearfunction$, $logisticfunction$, $hyperbolicfunction$ and $sigmoidfunction$.

$$\begin{aligned}
&linearfunction: && Purelin(x) = x \\
&logisticfunction: && Logsig(x) = \frac{1}{1+e^{x}} \\
&hyperbolicfunction: && Tansig(x) = \frac{e^{x}-e^{-x}}{e^{x}+e^{-x}} \\
&sigmoidfunction: && Sigmoid(x) = \frac{1}{1+e^{-x}}
\end{aligned}$$

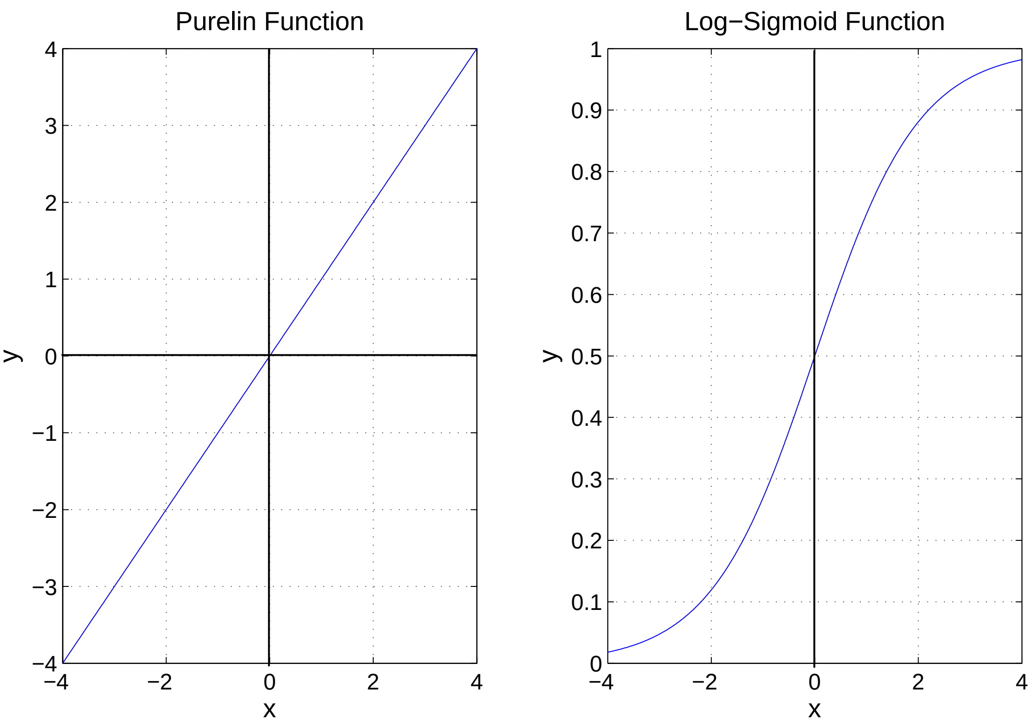


Figure 2.7: Linear(left) and logistic(right) function

$Logisticfunction$ is often used in a hidden layer for gather units from different layers. They map out the neurons and deliver the signals to the neurons of output layer (or next hidden layer, if setting more than one hidden layer in the model). $Hyperbolicfunction$ has an upper and lower bound between 1 and -1, and sigmoid function's upper and lower bound between 0 and -1. Consider about the nonlinear property and momentum effect of financial market, nonlinear transfer functions are more favored.

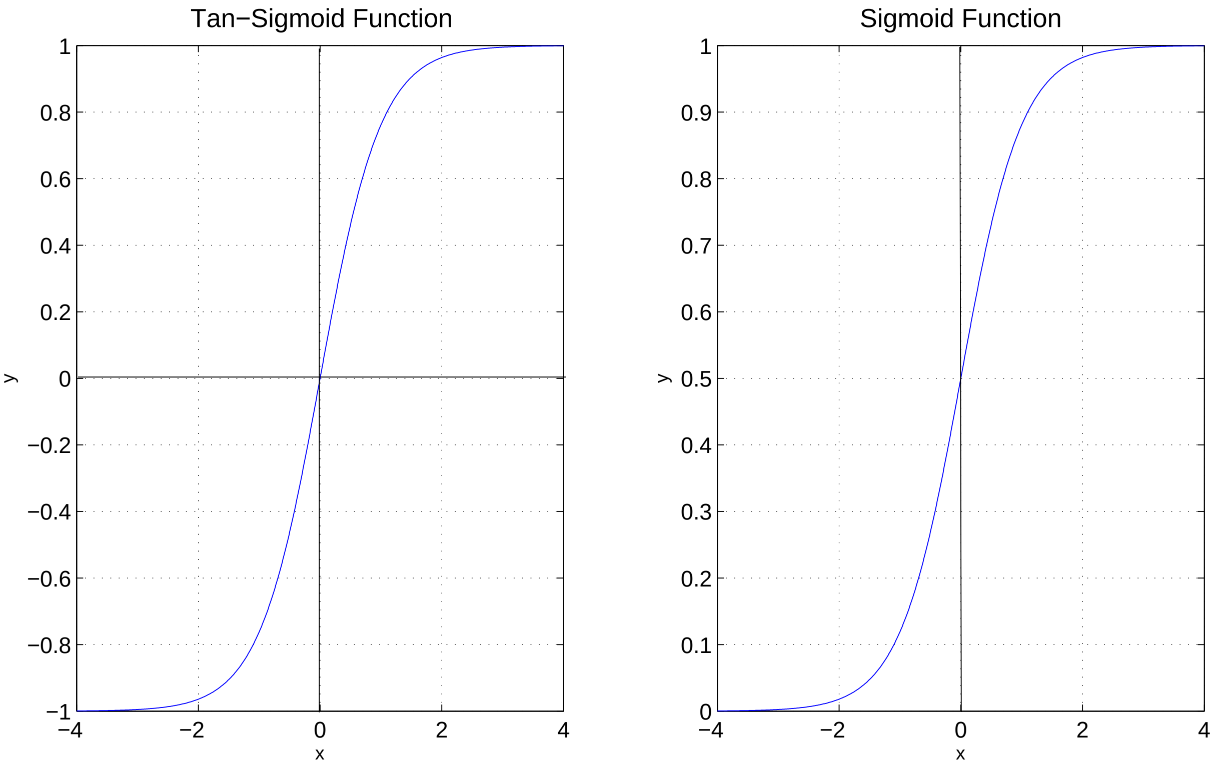


Figure 2.8: Hyperbolic(left) and sigmoid(right) function

2.2.1 Operation of Back Propagation Neuron

Here we focus on how to work in single neuron of Back Propagation Neural Network(BPNN). BPNN is one of the ANN family, it can sent back the error in reverse direction to adjust the weight and bias so we called it ”back propagation”. We adopt BPNN in this thesis. If a neuron of ANNs want to operate normally, it shall contains four major components: Summation Function, Bias/Threshold, Transfer Function/Activation Function and weight. Bias is the initial state of neuron, summation function is the summation of after weighted input values. The value after summation process, it will output through the treatment of transfer function, transfer function we already discussed above. It can present as following formulas [26]:

$$x = \sum_{i}^{N} W_{ij} X_i - \Theta_j \tag{2.6}$$

$$Y_j = f(x) \tag{2.7}$$

Where:

- $X_i =$ the i^{th} input variable.
- $Y_i =$ the output value of the j^{th} neuron.
- $f(x) = transferfunction$, which pull in the influence of non-linear effect.
- W_{ij} is the connect weight between the i^{th} input and the the j^{th} neuron.
- Θ_j is the threshold value of the j^{th} neuron.
- $i = 1...N, j = 1...M$; N is the number of input variables and M is the number of neurons.

If the output value doesn't achieve either stop condition, neuron will send back the error term to modify weights and biases, it's back-propagation process [27].

In back-propagation process the square error function present as following:

$$E = \frac{1}{2}\sum_{j}^{M}(T_j - Y_j)^2 \tag{2.8}$$

- $T_j =$ target value of the J^th output neuron.
- $T_j =$ output value from the J^th output neuron.
- $M =$ the number of neurons.

setting coefficient value $1/2$ in equation2.8 is for convenience in derivative process.

Then we can use gradient decent method to find the modify value of weightΔW.

$$\Delta W = -\eta\frac{\partial E}{\partial W} \tag{2.9}$$

- $W =$ connecting weight between neurons.
- $\partial E =$ the modification value of weights.
- $\eta =$learning rate, it control the modification range of weight, too large or too small both affect the convergence of network.

Using chain rule in equation2.9 of output layer, then:

$$\frac{\partial E}{\partial W_{jk}} = \frac{\partial E}{\partial Y_j} \cdot \frac{\partial Y_j}{\partial net_j} \cdot \frac{\partial net_j}{\partial W_{jk}} \tag{2.10}$$

$$\frac{\partial E}{\partial Y_j} = \frac{\partial[\frac{1}{2}\sum\limits_j^M (T_j - Y_j)^2]}{\partial Y_j} = -(T_j - Y_j) \tag{2.11}$$

$$\frac{\partial Y_j}{\partial net_j} = \frac{\partial[f(net_j)]}{\partial net_j} = f(net_j) \cdot [1 - f(net_j)] = Y_j(1 - Y_j) \tag{2.12}$$

$$\frac{\partial net_j}{\partial W_{jk}} = \frac{\partial(\sum\limits_k^{N_{hidden}} W_{jk}H_k - \Theta_j)}{\partial W_{jk}} = H_k \tag{2.13}$$

We substitution equation2.11 2.122.13 into 2.10, then obtain:

$$\frac{\partial E}{\partial W_{jk}} = (T_j - Y_j)Y_j(1 - Y_j)H_k \tag{2.14}$$

then we substitution equation2.14 into 2.9 and setting $\eta = 1$, redetermine ΔW as δ_j, $\delta_j =$ the j^th output layer neuron's modification value:

$$\delta_j = (Y_j - T_j)Y_j(1 - Y_j) \tag{2.15}$$

Similar way for neurons of hidden layer, the error function for the connecting weight of i^th input unit and k^{th} hidden neuron can be written as follow:

$$\frac{\partial E}{\partial W_{ki}} = \frac{\partial E}{\partial H_k} \frac{\partial H_k}{\partial net_k} \frac{\partial net_k}{\partial W_{ki}} \tag{2.16}$$

Using chain rule again, we get:

$$\frac{\partial E}{\partial W_{ki}} = (\sum_j^M \frac{\partial E}{\partial Y_j} \cdot \frac{\partial Y_j}{\partial net_j} \cdot \frac{\partial net_j}{\partial H_k}) \cdot \frac{\partial H_k}{\partial net_k} \cdot \frac{\partial net_k}{\partial W_{ki}} \tag{2.17}$$

$$\frac{\partial net_j}{\partial H_k} = \frac{\partial(\sum_k^{N_{hidden}} W_{jk}H_k - \Theta_j)}{\partial H_k} = W_{jk} \tag{2.18}$$

$$\frac{\partial H_k}{\partial net_k} = \frac{\partial[f(net_k)]}{\partial net_k} = f(net_k) \cdot [1 - f(net_k)] = H_k(1 - H_k) \tag{2.19}$$

$$\frac{\partial net_k}{\partial W_{ki}} = \frac{\partial(\sum_i^N W_{ki}X_i - \Theta_k)}{\partial W_{ki}} = X_i \tag{2.20}$$

Substitution equation2.6 2.7 of hidden and output layer the into equation2.22, we obtain:

$$\begin{aligned}\frac{\partial E}{\partial W_{ki}} &= [\sum_j^{N_{output}} (T_j - Y_j)Y_j(1 - Y_j)W_{jk}]H_k(1 - H_k)X_i \\ &= [\sum_j^{N_{output}} \delta_j W_{jk}]H_k(1 - H_k)X_i\end{aligned} \tag{2.21}$$

determine Δk as the $k^t h$ neuron's modification value of hidden layer:

$$\Delta k = [\sum_j^{N_{output}} \delta_j W_{jk}]H_k(1 - H_k) \tag{2.22}$$

$\sum_j^{N_{output}} \delta_j W_{jk}$ represent as the weighted summation of modification value of output layer. From equation2.22, we can see the modification value of hidden layer relate with the modification value of output layer, the modification value of output layer back propagate to the hidden layer of BPNN, this property let it called ”Back Propagation Neural Network”. We can use similar way to obtain the modification value of bias:

$$\Delta\Theta_j = -\eta\frac{\partial E}{\partial \Theta_j} = -\eta\delta_j \tag{2.23}$$

Above is the modification value of bias of output neuron. And for hidden neuron:

$$\Delta\Theta_k = -\eta\frac{\partial E}{\partial \Theta_k} = -\eta\Delta k \tag{2.24}$$

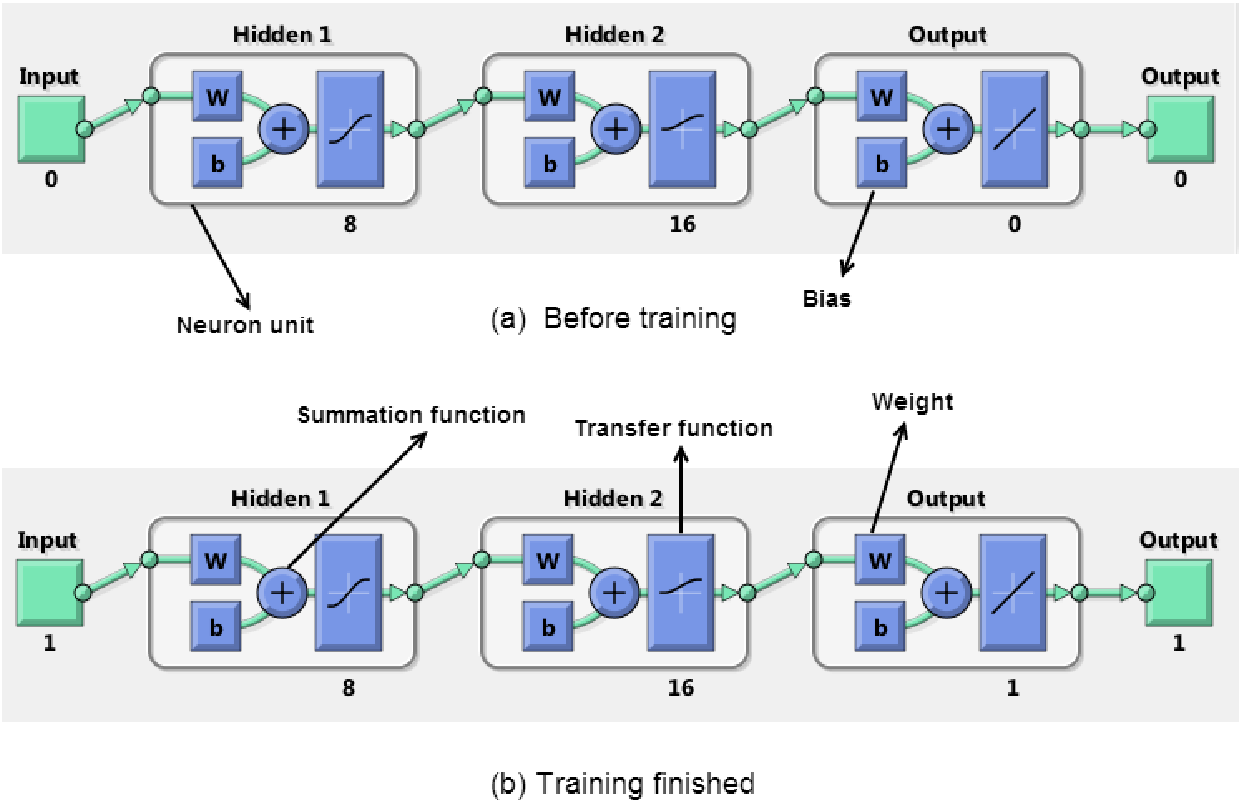
Then renew the weight and bias:

$$W_{(n)} = W_{(n-1)} + \Delta W_{(n)} \tag{2.25}$$

$$\Theta_{(n)} = \Theta_{(n-1)} + \Delta\Theta_{(n)} \tag{2.26}$$

$W_{(n)}$ and $W_{(n-1)}$ are the n^{th} and $(n-1)^{th}$ weight; $\Theta_{(n)}$ and $\Theta_{(n-1)}$ are the n^{th} and $(n-1)^{th}$ bias. Adding the modification into the current value, then we get the new weight and bias value after modification.

The flowchart figure2.10 illustrate this process. Figure2.9 demonstrate the configuration of BPNN networks we adopt in this research, and (a) is before training, (b) is training finished.



(a) Before training

(b) Training finished

Figure 2.9: Configuration of BPNN networks we adopt in this research

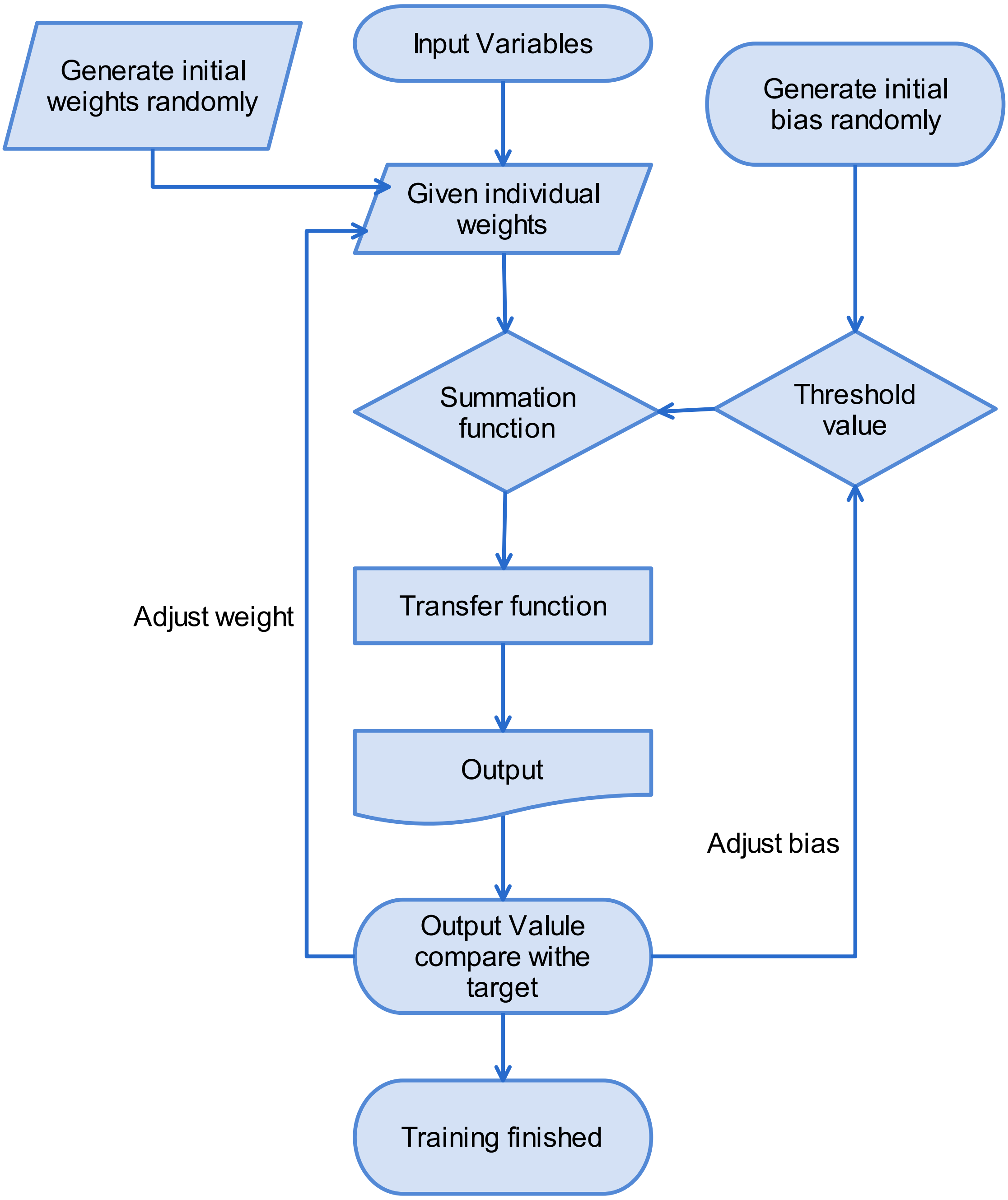


Figure 2.10: Flowchart of signal BPNN neuron training process

2.3 Autoregressive Moving Average model (ARMA)

Autoregressive Moving Average model (ARMA) is one of the standard econometrics models, it also can use in other domains for processing time series data. ARMA also have several variants, such as GARCH model, VAR model and ARIMA model etc. The general ARMA model proposed by Peter Whittle in 1951 thesis, <Hypothesis testing in time series analysis> [28], and it became well-known through the book by George E. P. Box and Gwilym Jenkins in 1971 [29]. The basic concept of ARMA assume the previous value will affect present value, and combined AR (Autoregressive) model and MA (moving average) model together to modified normal AR model [30].

2.3.1 The Autoregressive Model (AR)

An autoregressive model (AR) describes a certain time-varying processes. The ”autoregressive” means the output value have a relationship with its own previous values on some extent. For time series processing, the autoregressive model is a special case of the ARMA model.

The notation AR(p) indicates p order autoregressive model, it defined as following:

$$X_t = c + \sum_{i=1}^{p} \varphi_i X_{t-i} + \varepsilon_t \tag{2.27}$$

p is the lag operator produced by the previous element,$\varphi_i \ldots \varphi_i$ are parameters of AR(p) model, C is a constant and ε_t is white noise [31].

2.3.2 The Moving-Average Model (MA)

The moving-average (MA) model defined as following:

$$X_t = \mu + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \ldots + \theta_q \varepsilon_{t-q} \tag{2.28}$$

thus

$$X_t = \mu + \varepsilon_t + \sum_{i=1}^{q} \theta_i \varepsilon_{t-i} \tag{2.29}$$

where μ is the series mean, $\theta_1 \ldots \theta_q$ are the parameters of the model,$\varepsilon_t, \varepsilon_{t-1}, \ldots$are white noise error terms. The lag operator q is called the order of the MA model.

The moving-average (MA) model is a common approach for modelling univariate time series models, it is a linear regression between against current and previous white noise error terms or random shocks, it hides and unobserved. The random shocks in each point are considered mutually independent and the same distribution, usually a normal distribution, with zero location and constant scale. Fitting MA model is more difficult than fitting autoregressive models (AR models) because the lagged error terms are unobservable.

2.3.3 ARMA(p,q) Model

In equation2.27, we can see the white noise term ε_{t-1} not appear in the AR(p) equation, because it included in X_{t-1} term, thus the impact of ε_{t-1} is indirectly.

ARMA model is the evolution type of AR model and become more general use than AR model. ARMA model also consider the target data series exist the relationship with historical data. The impact factor of ARMA model classified to two types: one relate the previous data itself, it's AR model; another relate about random shock(white noise error) currently and previously, it's MA model [32].

Following equation combine AR model(equation2.27) and MA model(equation2.29) together:

$$X_t = c + \varepsilon_t + \sum_{i=1}^{p} \varphi_i X_{t-i} + \sum_{i=1}^{q} \Theta_i \varepsilon_{t-i} \tag{2.30}$$

Which is ARMA(p,q) model, this equation with p autoregressive orders and q moving-average orders.

If the data affected by more previously data(lags), then the order of p and q will larger. The popular form of ARMA(p,q) model is ARMA(1,1) model, the order of p and q are one. Following is ARMA(1,1) model:

$$X_t = c + \varepsilon_t + \varphi_1 X_{t-1} + \Theta_1 \varepsilon_{t-1} \tag{2.31}$$

In this thesis we build ARMA(1,1)model as a configure component to modify standard EEMD-BPNN model and try to find better configure between these three method.

2.4 High Frequency Trading, Big Data Processing and Parallel computing

2.4.1 High Frequency Data

High frequency trading is a major part of quantitative investment and recently decades become more and more popular. High frequency normally define as per minute data or shorter time-scale. During financial crisis in 2008, quantitative investment outperform traditional invest strategy, it give a reason for put into this subject. Based on the concept of fractal(section1.1), different time scale have similar price fluctuation, but the impact factor should be different. Shorter time scale, the impact factors of high frequency trading are different than traditional day trading [33], we will try to find it. Shorter time scale also generate massive data, high frequency data of one trading day in TAIEX 300 times more than daily data. And further more, the highest frequency data is the tick data(also called ”ultra-high frequency data”, figure2.11), its size bigger than per minute data hundreds to thousands times, it require higher performance of computing.

2.4.2 Parallel Computing for high-frequency trading backtesting

Higher frequency generate bigger data, how to process big data is another issue of high frequency trading. Big data will let the same thing become more difficult, you can do a simple thing 20 times, but it’s hard to doing it 20000 times, it also waste time. High-frequency data is a typical concept of big data, how to analyse high-frequency data is a problem, we use the batch process and parallel computing as solution.

In this thesis, we use Matlab® R2013a/b build parallel programming [34] by single program, multiple data (SPMD) method, which is a kind of parallel computing method [35], SPMD distribute the different parts of data stream to different workers, the worker can be

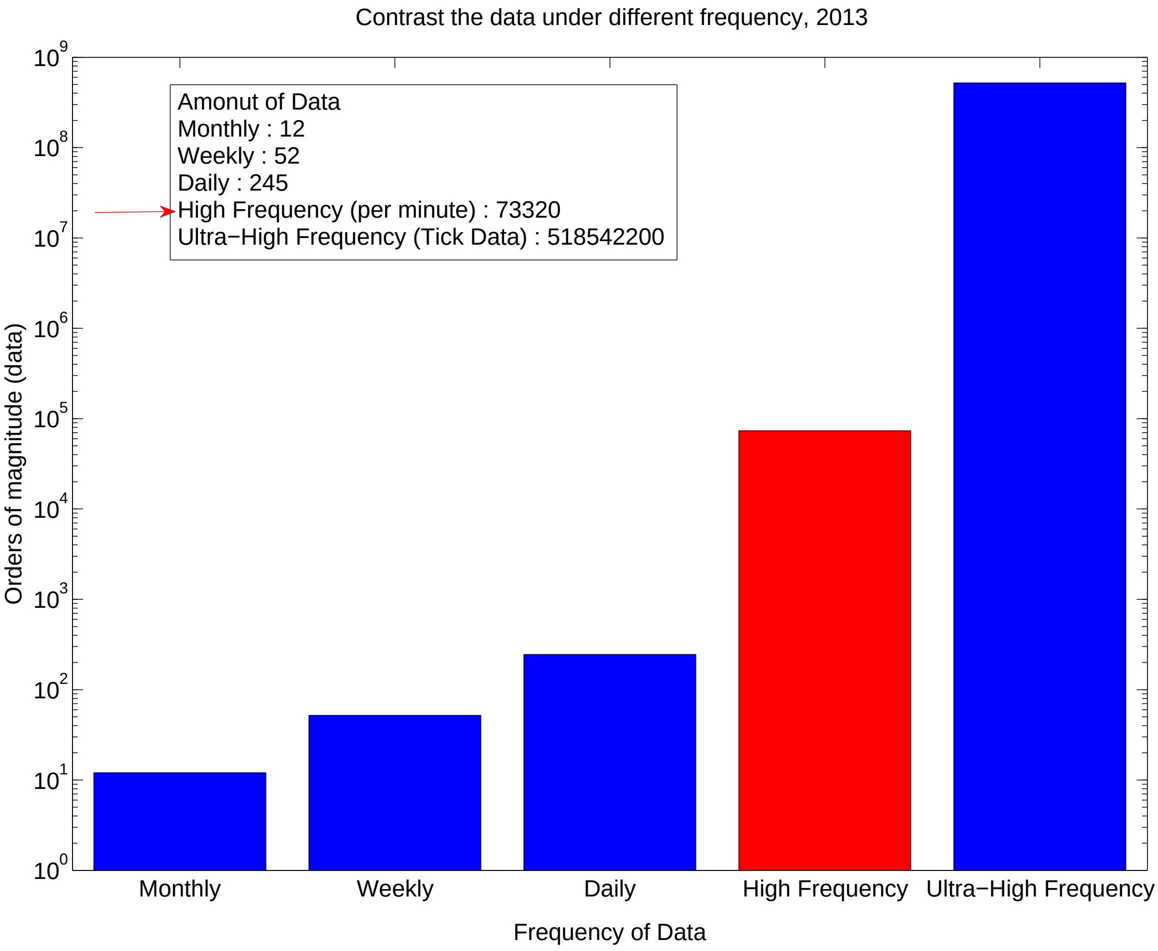


Figure 2.11: Contrast the different frequency of data

a computing node of a cluster sever, or a core of CPU. Here we adopt Intel® Core™i7-950 for SPMD parallel method in processing big data. Core™i7-950 contain four cores and eight threads in a single CPU, it can support two Matlab® program use four works(cores) the same time. Thought this construction, we can accelerate up to four times computing speed. Figure2.14 is a speed compare in one of our model.

Note that one principle of high-frequency trading is you cannot trading too many contracts at one time [36]. Stock market have some property similar as quantum system, one of these is uncertainty principle. In stock market, uncertainty principle can described as: stock price will affected by your trading, or your trading will affect stock price. The more volume you trade, the price you affect more [37]. This effect will cause back-testing become invalid. To avoid this situation, the volume of every trading time shall be reduce the level to doesn't affect the price.

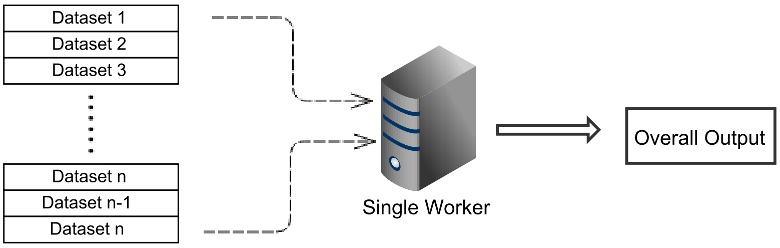


Figure 2.12: Normal single-threaded computing process

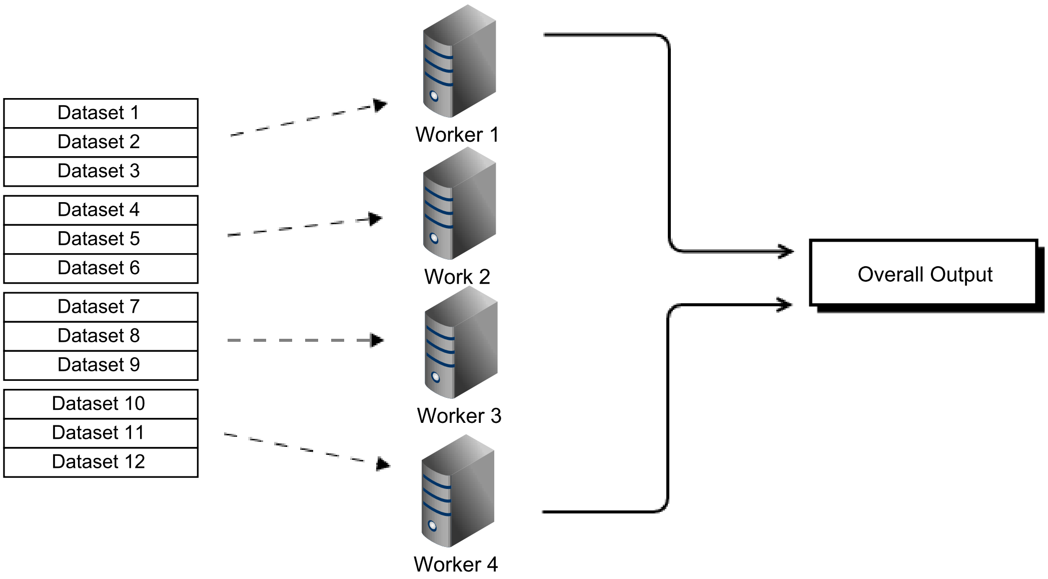


Figure 2.13: SPMD parallel computing process

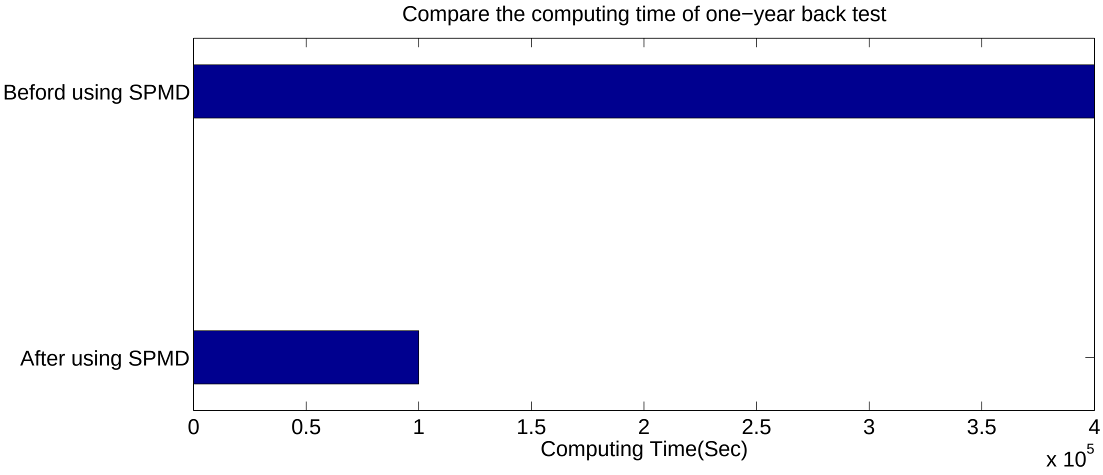


Figure 2.14: Contrast of SPMD computing speed: before and after use

2.5 Candlestick Charts

The candlestick chart first seen in 18^{th} century Japanese rice trading market, and then this unique pattern extension into western world [38] This pattern illustrate the price information and price trend, it general use to describe a derivative or currency price movements of a designated span of time. Candlestick charts usually use in the technical analysis, especially the item cannot adopt fundamental analysis [39]. The shape of candlestick chart looks like box plots superficially, but actually they are different and unrelated patterns [40]. The following figure illustrate the candlestick charts construction.

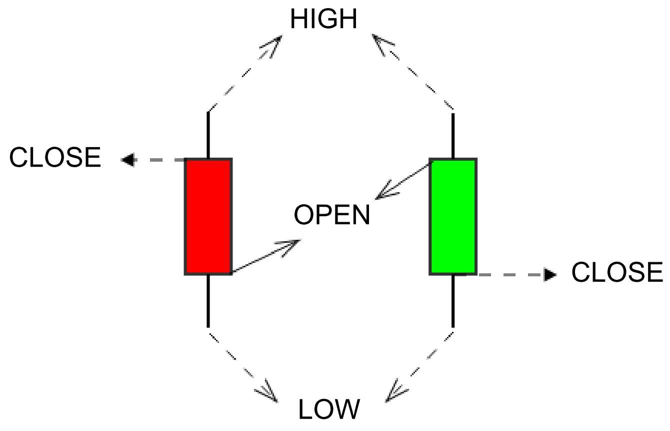


Figure 2.15: Scheme of bullish (left side)and bearish (right side)Candlestick Charts,the blocks with colors are realbodies. Usually we use red represent as bullish and green as bearish in Taiwan, but in other country may use different or opposite color to represent bullish and bearish.

Candlestick charts present the four important price value : open, high, low and close of a designated time unit. The red color of the candlestick indicates that closing price higher than opening price. The shadows represent the price fluctuation during the time unit, we will apply this concept later. The size and color of real-body offer the bearish or bullish information, the shadows offer the message of potential volatility. We can use these information deduce traders' thinking, it direct cause the change of supply and demand.

2.5.1 Supply and Demand Principle with Candlestick Chart

A price will adjust to higher or lower prices based on supply and demand principle [41]. The candlestick's color and size provide important clues regarding the trader's sentiment toward a given future price, it direct cause the change of supply and demand.In short term trading, understanding other traders' thinking is critical for trader. The most direct way to get that understanding is through proper interpretation of the candlestick charts. In figure 2.16 show a example, which opened at 100 and closed at 120.

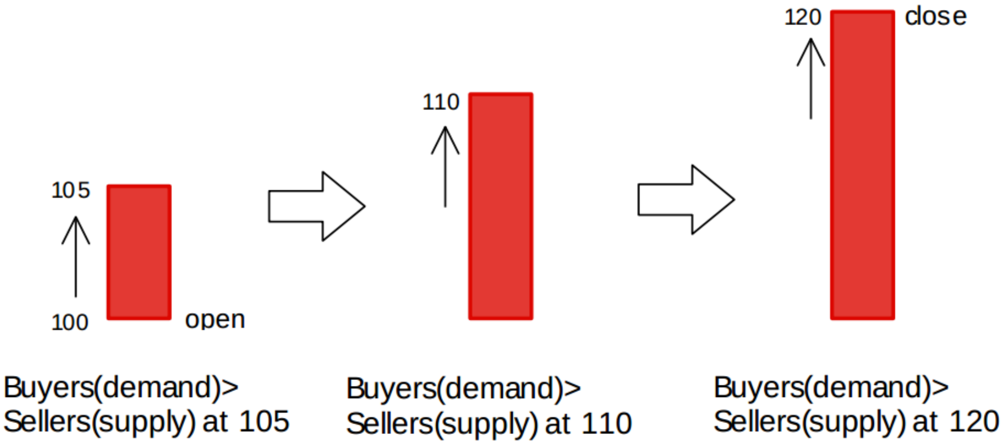

105
100
open
110
120
close
Buyers(demand)> Sellers(supply) at 105
Buyers(demand)> Sellers(supply) at 110
Buyers(demand)> Sellers(supply) at 120

Figure 2.16: The supply and demand mechanism of candlestick charts

In figure 2.16, we see the price opens at 100, and then quickly rallies to 105. The reason the price moves to 105 is because there is a high demand to buy the stock at 105, and a short supply of offering stock at 105. Once all of the stock available at price 105 is snatched up, the next seller group of will offer their stock at 110 price. This process will repeat itself until the buyers loose interest in buying(in this example is 120). It resulting in a reduction of demand. At this moment the price will stop the rising period. During the rally period, the astute candlestick reader observe the long red color of the candlestick, and deduce that buyer demand is high. In bearish(green color candlestick in Taiwan) also follow the same principle.

The reason why traders would increase demand by stepping up to buy, and that is because they think that the stock will rise in the near future. In this thesis we training computer try to using BPNN to learning the price change mechanism, act like a astute candlestick reader's trading. We build the candlestick chart's real-body time series from as raw data to forecast the bullish or bearish trend of the coming next time span.

Chapter 3

Market Efficiency and Physical Mean of IMFs

3.1 The Efficient-Market Hypothesis(EMH)

Efficient-market hypothesis is one of the most important fundamental concept in finance. Although we call it ”hypothesis”, some people consider it as ”theory”. Efficient market assume that financial markets are ”informationally efficient”. Base on this concept, more liquid market will cause price reflect new information more efficiently, and reducing trading cost. [42] Under the efficient-market hypothesis, investor can not earn supernormal profit at an efficient market. In other words, efficient market cannot exist the opportunity of consistently earn excess returns in long period. it’s a indicator of examine efficient-market hypothesis. It should note that investors can temporary earn excess returns in short run under the efficient-market hypothesis, but earn excess returns in long run will violate the EMH.

Efficient-market hypothesis can classified to three different intensity ”weak-from”, ”semi-strong-form”, and ”strong-form”, describe as following:

- **Weak-form efficiency:** In weak-form efficient market, analysing past prices cannot predict future prices value. in the long run , cannot earned excess returns by using any investment strategies based on historical share prices or other historical data.

Technical analysis techniques In weak-form efficiency market will not be able to produce excess returns consistently, but investor still can earn excess returns through some forms of fundamental analysis.

- **Semi-strong-form efficiency:** In semi-strong-form efficient market, it implies share prices which adjusted from public material in an unbiased fashion very rapidly. As the result, no excess returns can be earned by trading on the new public material information. Semi-strong-form efficiency implies neither fundamental analysis nor technical analysis can reliably produce excess returns.

- **Strong-form efficiency:** In strong-form efficient market, share prices reflect all kinds information, no matter the information is public or private, no investor can earn excess returns. Note that although investment strategies cannot help investors earn excess returns under the strong-form efficiency market, it still can reduce the non-systematic risk of the portfolios.

3.2 Market Inefficiency and High Frequency Trading

In nowadays, most of public-trading securities markets are a certain extent efficient, it means the time period of excess returns that investors can earn will decreasing with time. This inefficient ”short run” is the high frequency trading target. High frequency trading aim to use scientific or statistic method developing investment strategies to catch these very short period inefficient moments to make profit. Following figures illustrate the short-term market inefficient movement. Several reasons can cause market inefficient, include delay for information delivery, delay for trading process, difference of individual market regulation, different trading market/time, and irrational behaviour of investors, etc. [43]

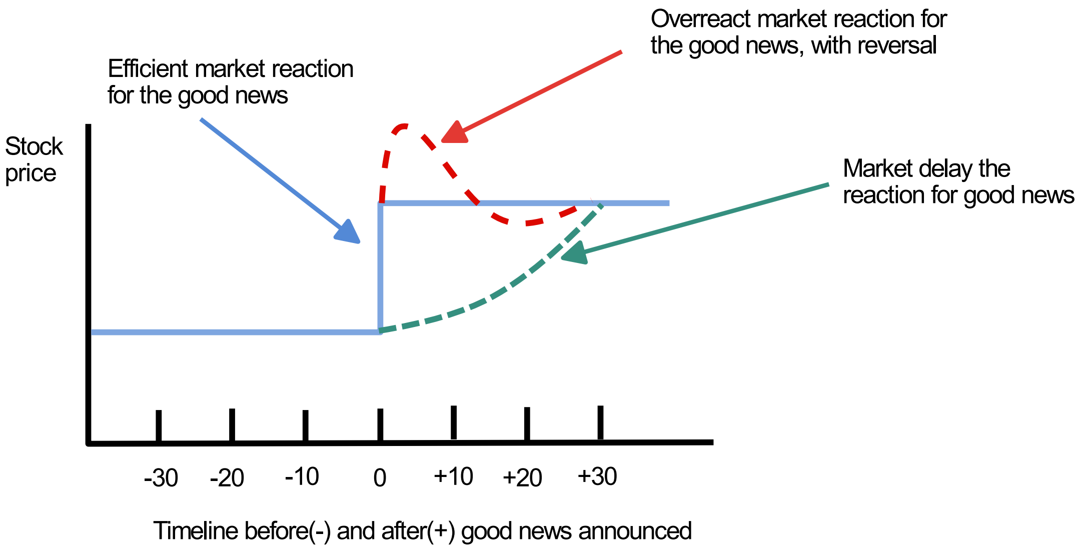

Overreact market reaction for
the good news, with reversal
Efficient market reaction
for the good news
Stock
price
Market delay the
reaction for good news
-30
-20
-10
0
+10
+20
+30
Timeline before(-) and after(+) good news announced

Figure 3.1: Reaction of stock price to new ”good” information in efficient and inefficient markets. Here we demonstrate two typical inefficient reactions: overreaction and delayed.

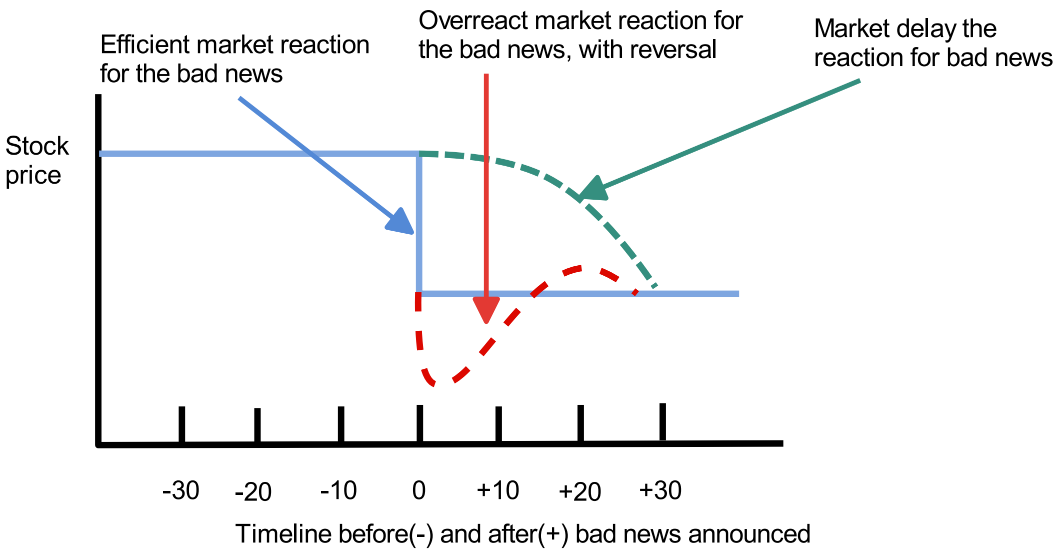

Efficient market reaction
for the bad news
Overreact market reaction for
the bad news, with reversal
Market delay the
reaction for bad news
Stock
price
-30
-20
-10
0
+10
+20
+30
Timeline before(-) and after(+) bad news announced

Figure 3.2: This figure illustrate the reactions of stock price to new ”bad” information in efficient and inefficient markets. Similar as figure 3.1, here we also demonstrate the two typical inefficient reactions. The causes of overreaction primary is irrational investment behaviour and over expect for market. And the reason of delayed usually because the limitation of informations.

3.3 The Randomness of Data

In the section of efficient-market hypothesis, we know if the market are weak-form efficient, we cannot earn excess returns through analysing historical price. It means the price is unpredictable by historical price. In statistic, more randomly will cause more difficult to analyse and determine the property. In this section we try to find out the randomness of data. The methodology we use is Non-parametric run test [44].

Non-parametric run test have been proposed the earliest (Louis Bachelier, 1900 [45]) in several test method. This method measure the probability of a series positive or negative fluctuation (or called ”run”). Just like the throw a coin, probability of appearing the same side two times continuously is $1/(2^2) = 0.25$, appearing three times continuously is $1/(2^3) = 0.125$, and appearing four times continuously is $1/(2^4) = 0.0625$ or 6.25% , and so on. The more times of appearing the same side continuously, the appear probability of this situation will decreasing. No matter in which time scale, this kind of non-randomness provide opportunities of trading. The testing procedure as following:

1. For the data's frequency we require, record the number of series of moving to the same direction. If the frequency we require is per minute, the run is the series of the number of moving to the same direction under the gap of one minute.

2. Record the total number of runs, including positive and negative as u. Then we record the number of positive direction of per minutes as n_1 and negative direction of per minutes as n_2.

3. Then the expect value of total number of a randomly sample $\bar{x}$, and the standard deviation s:

$$\bar{x} = \frac{2n_1n_2}{n_1 + n_2} + 1 \tag{3.1}$$

$$s = \sqrt{\frac{2n_1n_2(2n_1n_2 - n_1 - n_2)}{(n_1 + n_2)^2(n_1 + n_2 - 1)}} \tag{3.2}$$

4. Then we test whether the run number can represent the non-random characteristic in statistic. If the $\bar{x}$ greater than 1.645 times of standard deviation s, then we may

consider under 95% confidence level it is predictable in this frequency, otherwise it is non-random. If the two-tails test based on Z-value have rejected by the run number, then the data of run is non-random. In other words, in 95% confident level if Z greater than 1.645 we can reject the null hypothesis(the data series is not random); if Z-value less than 1.645, we cannot reject the null hypothesis (the data series is random)

Table 3.1: Different Stock Index With Different Time Scale

Item	Z-value	Skewness	Kurtosis	Random
2007∽ 2009 S&P500 weekly	0.9442	-0.12	2.1	Y
1997∽ 2010 S&P500 monthly	0.341355	-0.08	2.22	Y
2000∽ 2008 SH000001 daily	0.809884	1.94	6.08	Y
2010/07/20 TAIEX per minute	3.496811	0.15	1.35	N
2001/01/02 SH000001 per minute	4.935185	-0.12	7.03	N
2000∽ 2008 SH000001 weekly	2.566168	1.94	6.06	N
2000∽ 2008 SH000001 monthly	1.852582	1.96	6.22	N
2010/07/20 TAIEX per minute rise/fall range	9.606336	0.36	4.47	N
1997/07∽ 2009/04 TAIEX daily	0.867647	0.23	2.19	Y

Table3.1 list different market indexes and different time scales. Here we can see all the test under per minute scale reject the null hypothesis. It means exist serious inefficient characteristic of market under one-minute time scale. We can see the shorter time scale shows more inefficiency of market,table3.1supports this conclusions.

Here we introduce a indicator ”correct trend”, it will mentioned again later. Depend on Chen Yuan Hsiao(2013), correct trend defined as follows:

$$Correct\ Trend = \frac{Correct\ signal}{Total\ point} \tag{3.3}$$

the Correct Trend measure the ratio of forecasting model’s predict accuracy.

Table3.2 present the correct trend of TAIEX and SH00001 the Standard Model. Standard Model is the benchmark model adopted in this thesis later. We can see SH00001 have much higher correct trend ratio, and the Z-value test also reject the random hypothesis. The reason is SH000001 adopt (T+1) rule, it restrict the intra-day trading. This kind of restrict let market have lower liquidity and lower price fluctuation (we can see the shadow

Table 3.2: Correct Trend compare between TAIEX and SH00001

TAIEX		SH000001		
Trading Date	Standard	Trading Date	Standard	Z-value of SH000001
2010/2/1	0.45	2010/2/1	0.73	4.4568
2010/2/2	0.45	2010/2/2	0.7	3.5328
2010/2/3	0.45	2010/2/3	0.77	5.4852
2010/2/4	0.42	2010/2/4	0.74	3.1505
2010/2/5	0.41	2010/2/5	0.72	1.9804
2010/2/6	0.45	2010/2/8	0.66	2.554
2010/2/8	0.36	2010/2/9	0.67	3.1401
2010/2/9	0.45	2010/2/10	0.6	2.2754
2010/2/10	0.45	2010/2/11	0.64	1.7151
2010/2/22	0.35	2010/2/12	0.56	1.4295
2010/2/23	0.38	2010/2/22	0.69	2.4925
2010/2/24	0.38	2010/2/23	0.72	2.3126
2010/2/25	0.43	2010/2/24	0.68	2.0211
2010/2/26	0.40	2010/2/25	0.67	0.1569
N/A	N/A	2010/2/26	0.71	1.6985
Average	0.42	Average	0.68	2.5601

of candlestick charts in figure3.3), it will cause more market inefficiency and more predictable.

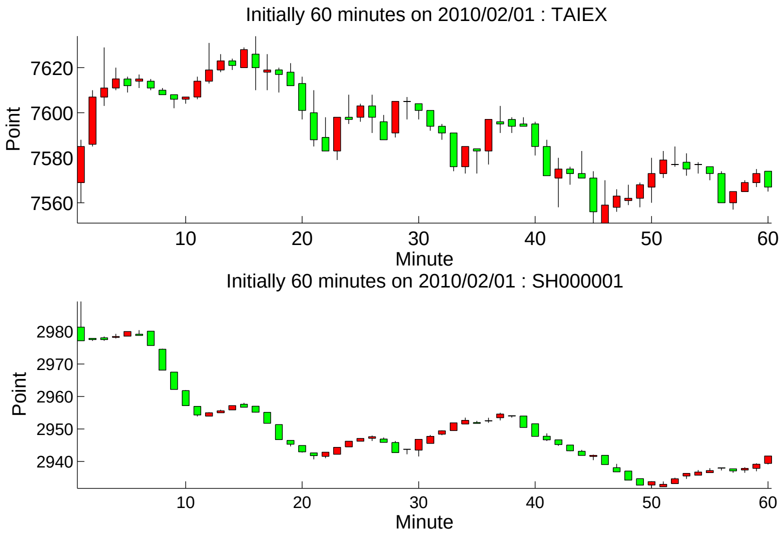


Figure 3.3: Contrast of different price fluctuation between TAIEX and SH000001

3.4 Physical Meaning of IMFs in High Frequency Data

Previously, we already discuss at section2.1 and section2.1.1 about how to decompose raw data into IMFs through EEMD method. In this section we will discuss the physical meaning or impact factor of IMF through try to find out the correlation of IMFs and other time series data. Based on the characteristic of fractal we introduce in section1.1, different time scale IMFs have different correlate object, Chen Yuan Hsiao(2013) describe the monthly scale IMFs meaning, he concluded that high frequency impact factor relate about foreign institutional investors for three-month period; mid frequency impact factor is the annual reports from stock listing corporations, which are the material that long-term investors depend on to determine their investment strategies. In low frequency, the business cycle is the significant impact factor of low frequency monthly IMFs.The business cycle presents as macroeconomic leading indicator, it highly correlate with low frequency monthly IMFs in four-year period. The residual term is overall trend, Chen concluded that real GDP is the significant impact factor of residual term, and it present long-term economic growth.

In this thesis, the time scale of raw data is per minute, consider about the different periodic time scale with Chen, the impact factor will different with Chen's conclusion for monthly-scale IMFs. Most of regular economic indicator release in annually, quarterly, monthly, daily release is the shortest release frequency, but all of them cannot suitable for intra-day high frequency data. In order to comparable with the intra-day raw data, we compare some quintessential object in financial market such as gold futures and exchange rates in comparable time scale with raw data.

Table 3.3: Decomposed IMFs of raw data. Note that the power percentage is the percentage of individual IMF amplitude contrast with raw data amplitude, the summation amplitude of all IMFs may not be perfectly coincident with raw data amplitude.

	Mean period	Correlation	Power percentage
IMF1	3.24	0.74	53.87
IMF2	5.61	0.55	17.4
IMF3	15.48	0.36	12.41
IMF4	30.11	0.23	1.38
Residual	67.37	0.25	8.78
Sum	N/A	N/A	93.84

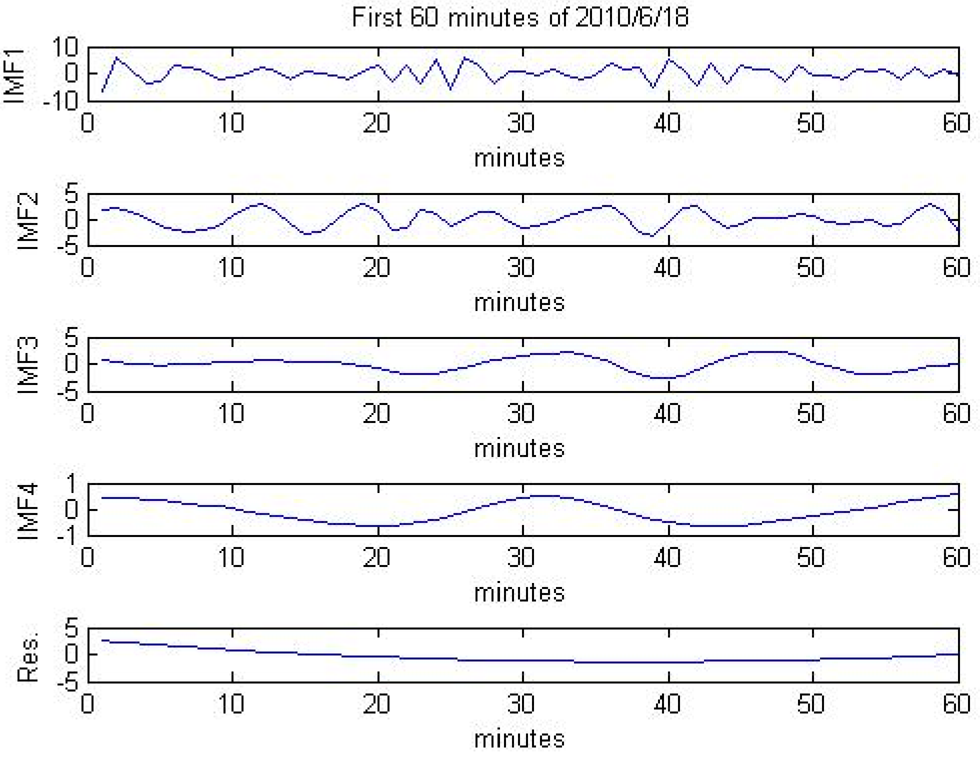


Figure 3.4: By definition of decompose IMF, in general more data length will decompose into more IMFs. Here we use 60 minutes length, the same as the training size in our trading model.

The following figure 3.4 and table 3.3 present the decomposed IMFs of raw data. We can see the mean periodic of IMFs around three minutes, five minutes, fifteen minutes, thirty minutes, and sixty minutes. These periodic represent as different time scale price data release, different investors have different favor time scale of data on their own to build different investment strategies for different purpose. Therefore different IMF period represent as different released time scale.

3.4.1 Long-Short Position

In high frequency terms of IMFs, we intend to compare with the bearish and bullish trend under the same time scale. Here we defined an index to measure the bearish/bullish trend of candlestick charts:

$$\text{Long-short position} = (C - L) - (H - C) \tag{3.4}$$

C is close price, H is highest value and L is lowest value in single candlestick chart. This formula represents overall trend between bearish and bullish. Note that bullish is long position and bearish is short position. In candlestick charts, the difference between highest price and close price represent the price-rising resistance, or bearish. Follow the same concept, the difference between lowest price and close price represent the price-rising support, or bullish. This concept is extend from basic concept of shadows we mentioned in section2.5. Following figure illustrate this relation.

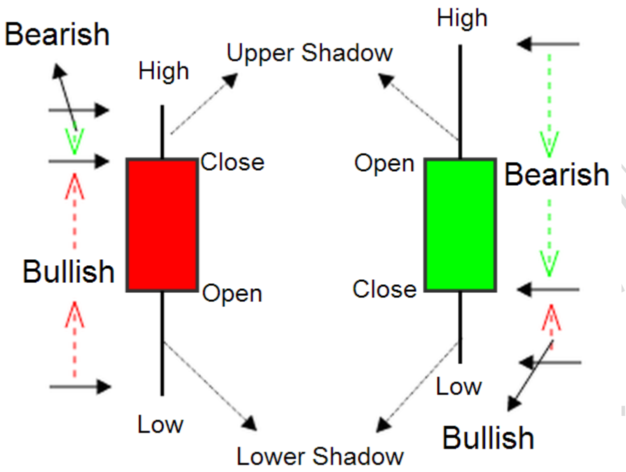


Figure 3.5: This figure illustrate the Long-short position of candlestick. Left side bullish > bearish, represent as long position; right side bullish < bearish, represent as short position. Note that the long or short position doesn't have absolute correlate with direction of realbody, if the bearish (upper shadow) greater than bullish(realbody + lower shadow), we still possible to obtain a short position result, even the realbody is rise (represent as red in Taiwan)

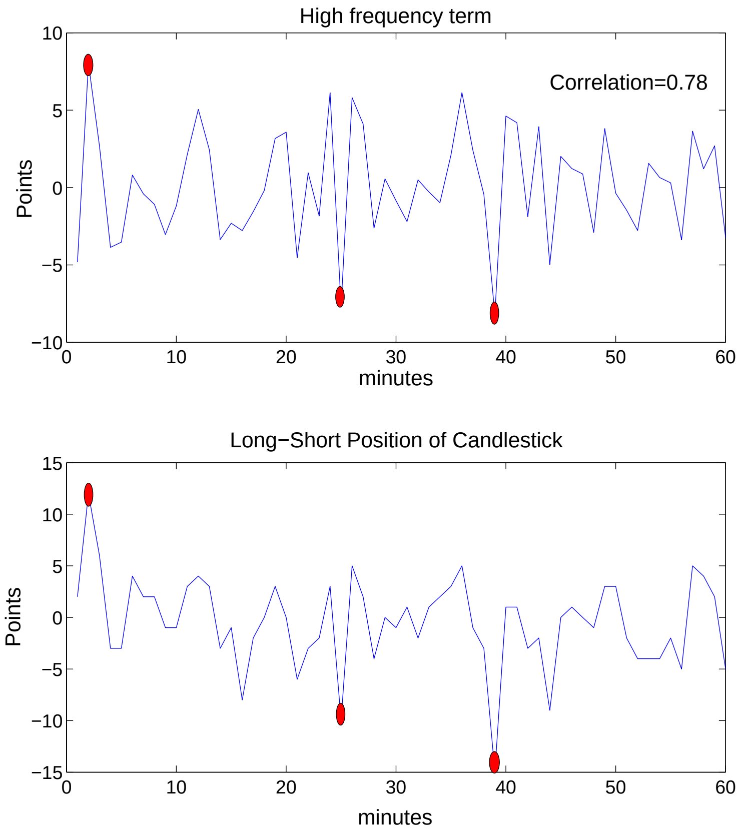


Figure 3.6: This figure illustrate the correlation between high frequency terms IMFs and the long-short position of raw candlestick. This result represent that the realbody data have a relationship in certain extent with long-short position in high frequency terms.

3.4.2 Volume

In this section we will discuss the relationship between volume and price. Volume is the quantitative measurement number of deal. Volume provide a measurement of market liquidity. Liquidity is one of the basic assumption of efficient markets hypothesis, more liquidity of market, the market will more efficient. Trading price change should supported by a certain extent volume, otherwise the price change trend will unavailable (because of little deal number). This concept called ”volume-price relationship”. [46]

Based on volume-price relationship, some technical analyst use volume determined the price rise/fall trend. Actually, the relationship of price and volume is not always correlation, when volume is low, the price may not react for volume, significant correlation appear usually when volume increase rapidly, but not everytime. Some analysts use moving average of volume to filter the volume peaks which affect price(illustrate as figure3.7).

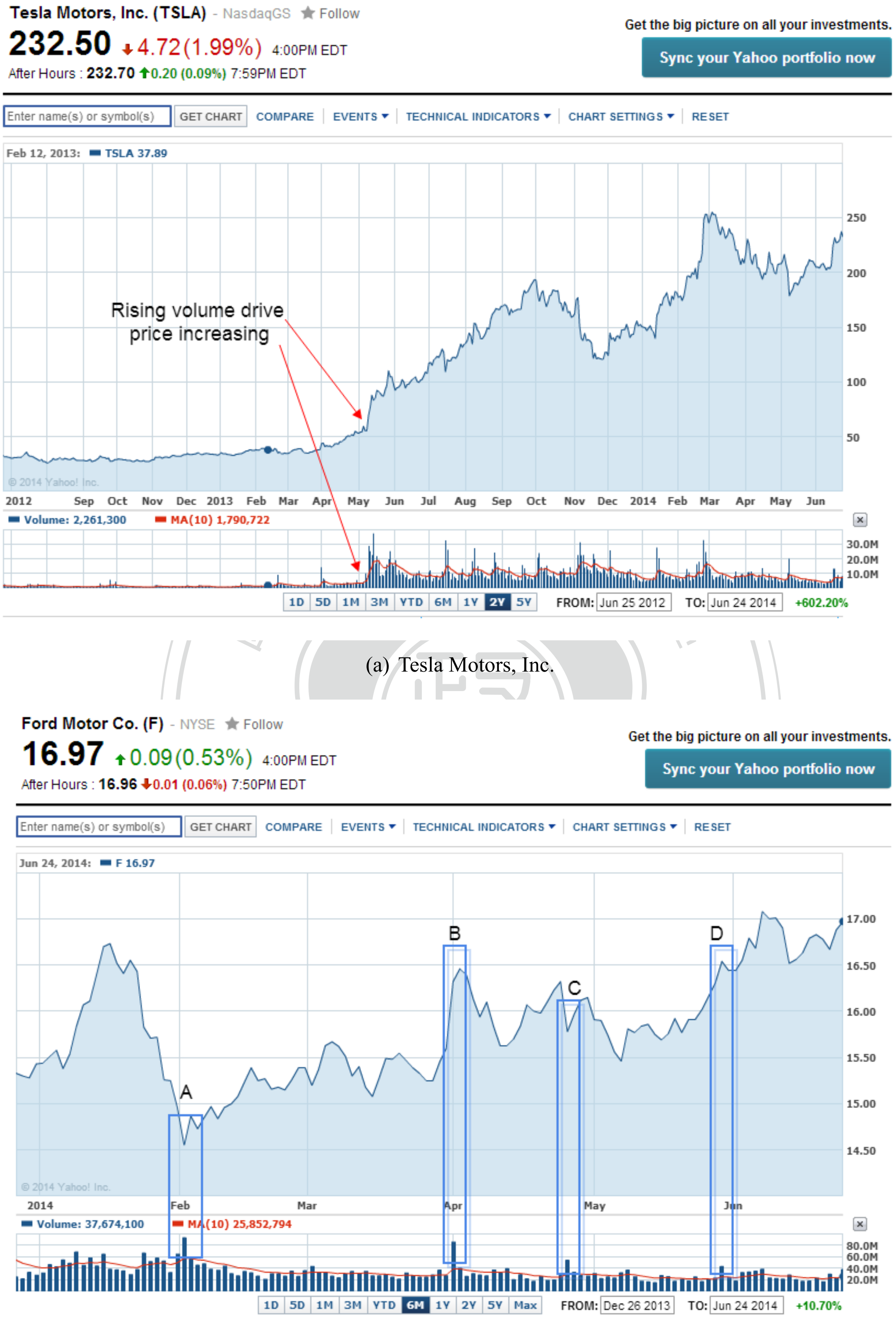
Table 3.4: IMFs of Volume

Volume	Mean period	Correlation	Power percentage
IMF1	3.57	0.37	36.99
IMF2	8.59	0.63	11.42
IMF3	22.74	0.37	26.52
IMF4	31.35	0.51	20.20
Residual	145.22	0.36	16.90
Sum	N/A	N/A	112.04

Then we decomposed the volume time series, intent to find out more clearly relationship between Price and Volume. We find out in low and mid frequency IMFs have high correlation between Volume and real body. Residual also have the same phenomenon, the reason is volume time series usually fluctuate in a specific range level, this property similiar to realbody.(the reason of realbody is fluctuation limits, for volume because the capacity/number of deal contracts are limited. This two reasons limit the fluctuate range.) This characteristic indicate the relationship between candlestick charts and volume more close than close price and volume, in figure3.13 we can seen this relationship. Close price more present long term price rising or falling characteristic, this is the reason why we adopt candlestick rather than close price in this thesis.

Figure3.8 is contrast between realbody and volume, we can also observe the characteristic mentioned above.

Figure3.9 3.10 3.11 describe the correlation phenomenon between volume and price, we can see the characteristic we mentioned above. In figure3.12 we can see volume leading price about five minutes, because volume is the basic support for price change, this phenomenon present the volume-price relationship [47].



(a) Tesla Motors, Inc.

(b) Ford Motor Co.

Figure 3.7: Volume-Price Relationship: in fig.(a), Tesla's price rise after the volume increasing; in fig.(b),all four points A B C D, their volume significant over MA(10) line, and cause the price rapidly rise or turning.

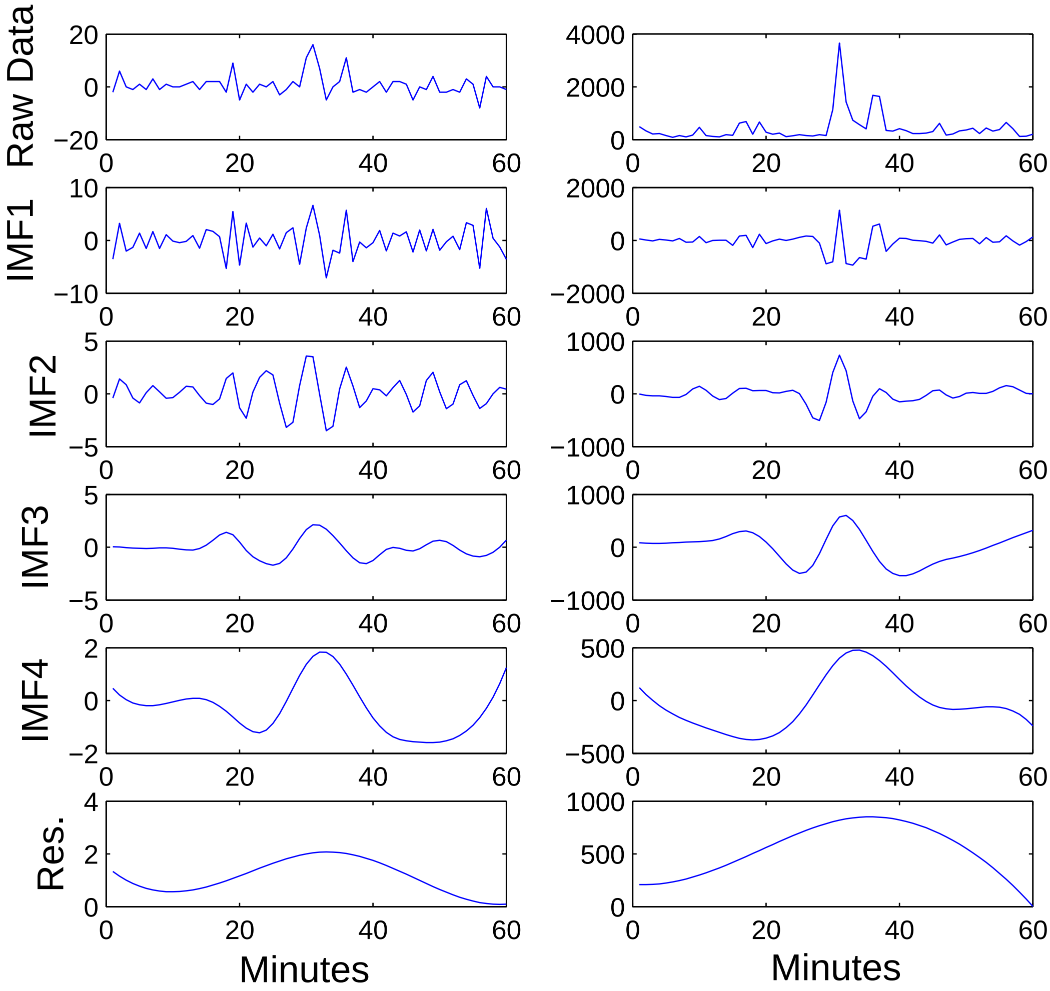


Figure 3.8: Contrast of Volume and Realbody

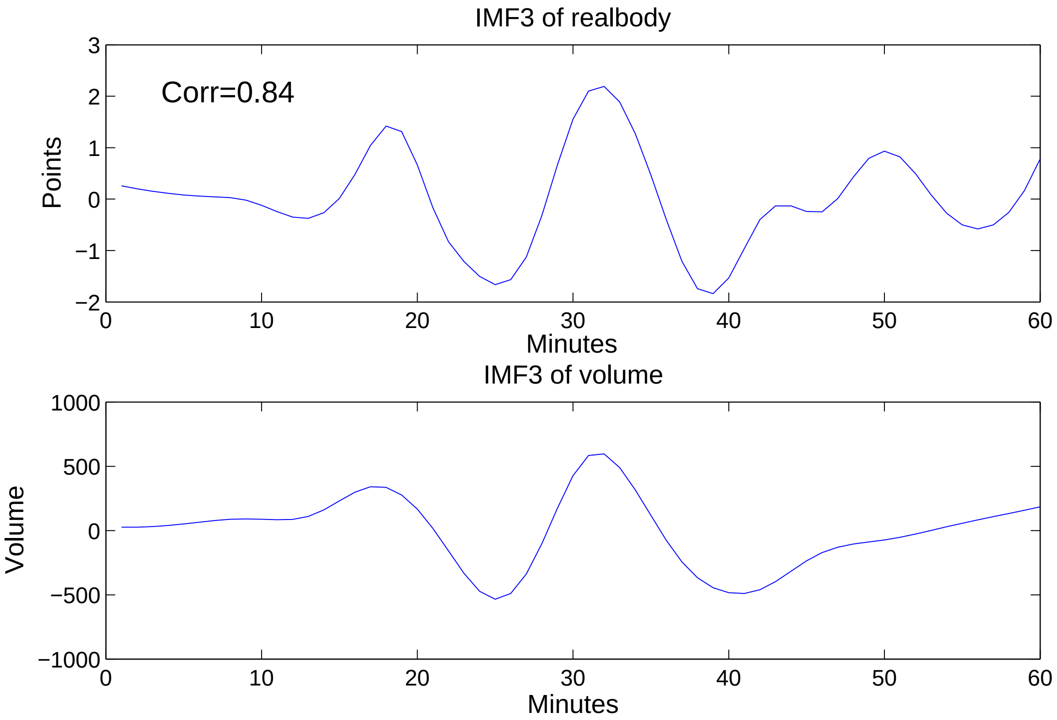


Figure 3.9: Contrast of IMF3

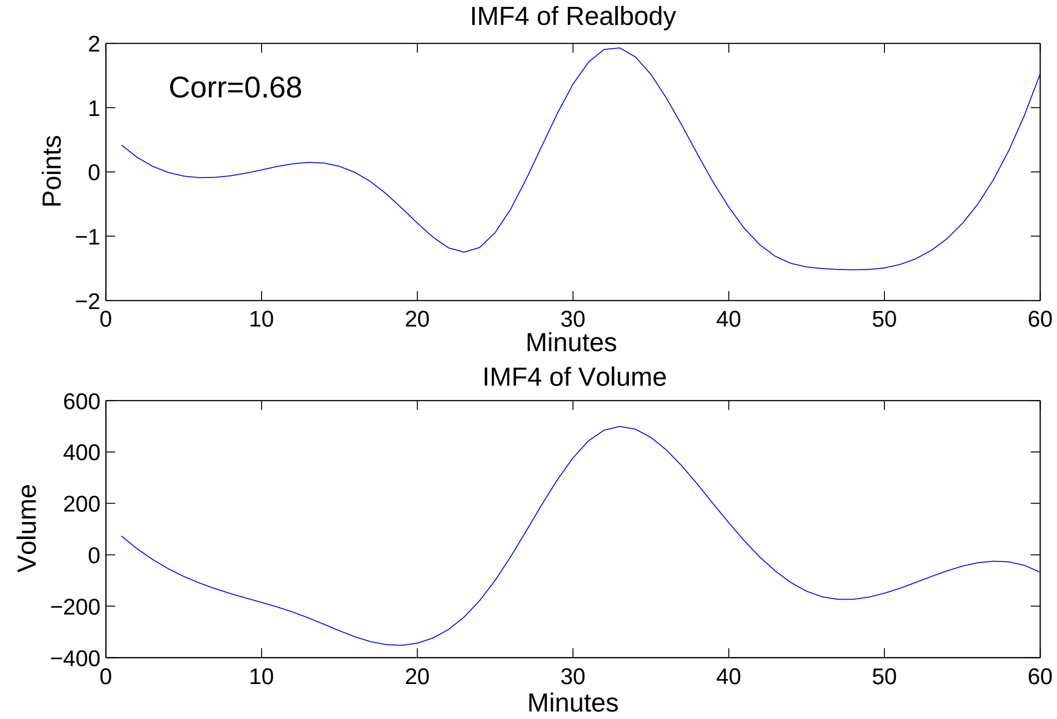


Figure 3.10: Contrast of IMF4

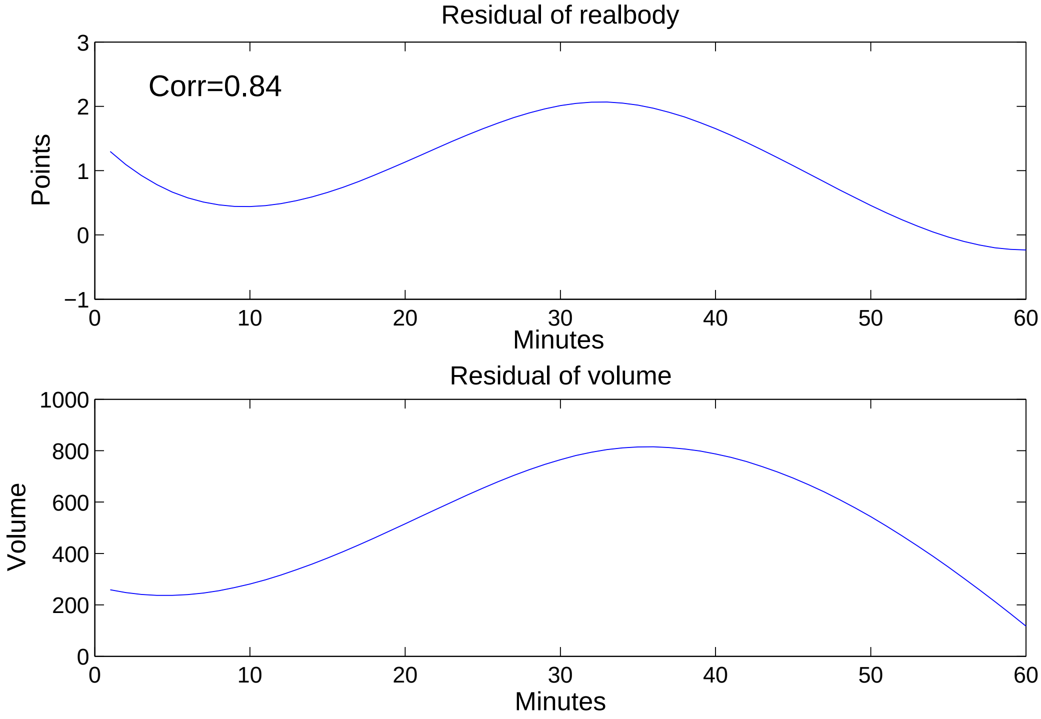


Figure 3.11: Contrast of Volume and Realbody: Residual

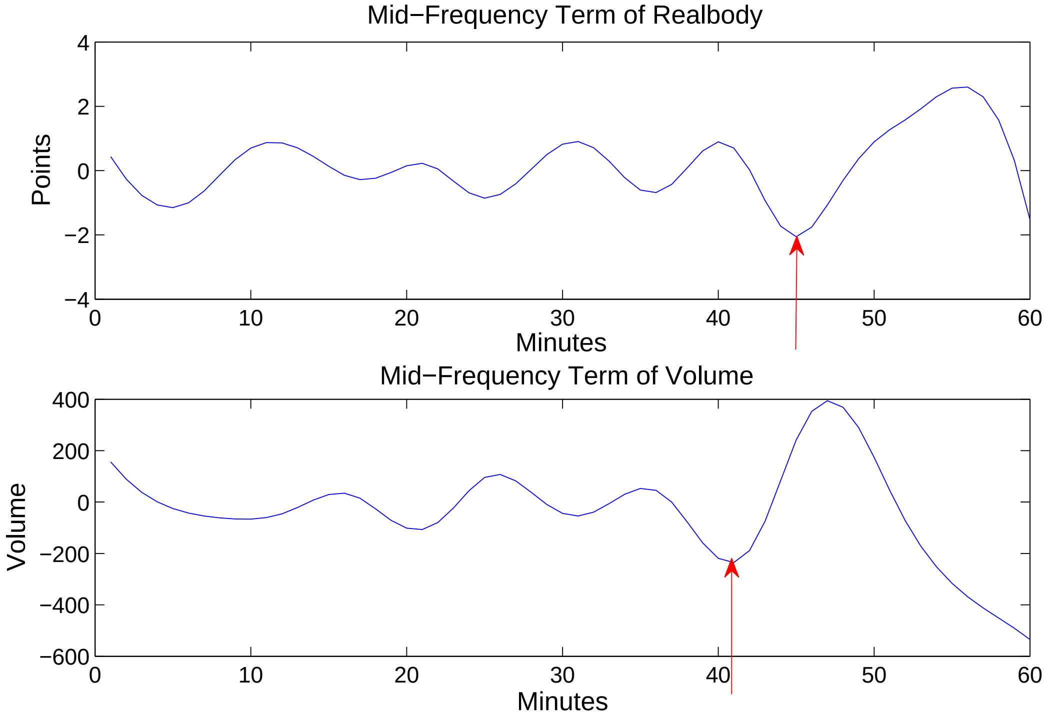


Figure 3.12: This figure describe the relationship of realbody and volume in mid-frequency term. from the arrow we can see the volume leading realbody about 5 minutes, the correlation coefficient after shifting adjust is 0.68

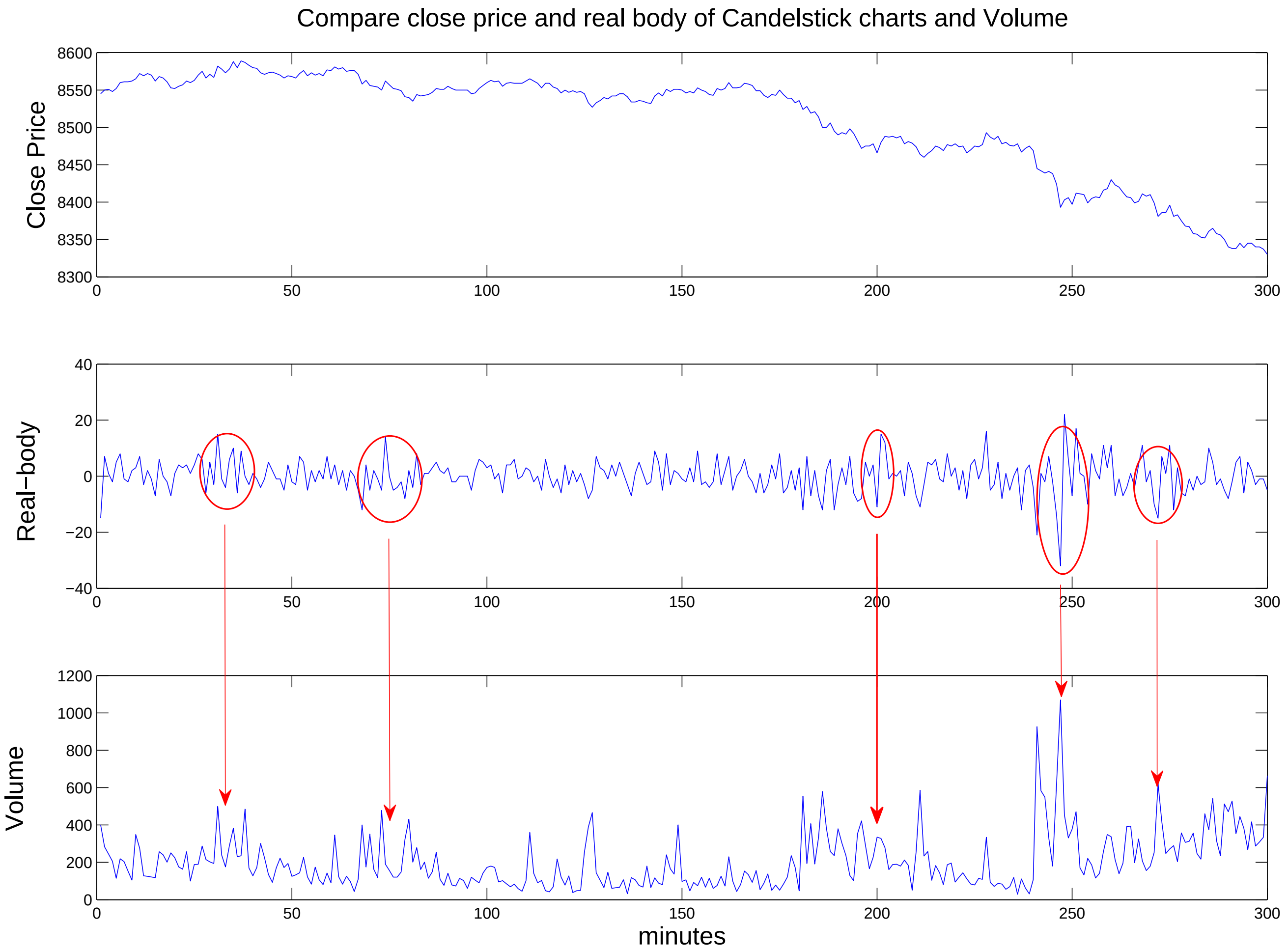


Figure 3.13: Compare of close price, candlestick realbody and volume. We can see the rapid rise of candlestick and the rapid rise of volume have high correlative level(indicate by arrows and circles).

3.4.3 MSCI Taiwan

In mid-frequency, we use the MSCI Taiwan index futures as contrast with mid-frequency IMF. The MSCI Taiwan Index is a free-float adjusted market capitalization weighted index made by Morgan Stanley Capital International(MSCI) [48]. MSCI Taiwan is aim to track Taiwanese securities market performance which listed on Taiwan Stock Exchange and GreTai Securities Market. The MSCI Taiwan Index is constructed based on the MSCI Global Investable Market Indexes Methodology, use in a free-float market capitalization coverage of 85 percent. The base date is December 31, 1987 [49].

MCSI Taiwan trading at Singapore Exchange (SGX), charge unit is U.S. Dollar, MCSI Taiwan provide a way to reduce the currency risk for foreign investment institutional investors. Additional, its weighting and constituent stock both constructed by foreign investment institutional investor, it has caused investment institutional investors usually used MSCI Taiwan futures to hedge risk or build hedge investment strategies for TAIEX. MSCI Taiwan has significant influence for TAIEX movement [50].

For comparability, we choose the overlap trading time between TAIEX futures and MSCI Taiwan futures and compare their IMFs. The following figure and table represent IMFs of MSCI Taiwan.

Table 3.5: IMFs of MSCI Taiwam (SGX)

MSCI Taiwan(SGX)	Mean period	Correlation	Power percentage
IMF1	3.65	0.14	0.67
IMF2	6.68	0.11	0.4
IMF3	27.4	0.13	2.35
IMF4	33.98	0.56	7.94
Residual	14226.5	0.94	78.67
Sum	N/A	N/A	90.02

Here we used the second hour close price time series to avoid the opening volatility. Based on table3.5, the IMF periods are around three minutes, five minutes and thirty minutes.

Figures 3.15 and 3.16 represent the analyze result. In figure3.15, Both IMF3 of TAIEX and MSCI Taiwan have high correlation coefficient at 0.78. In this analysis, IMF3 classified as mid-frequency IMF, MSCI Taiwan (SGX) is delay around 5 minutes than TAIEX.

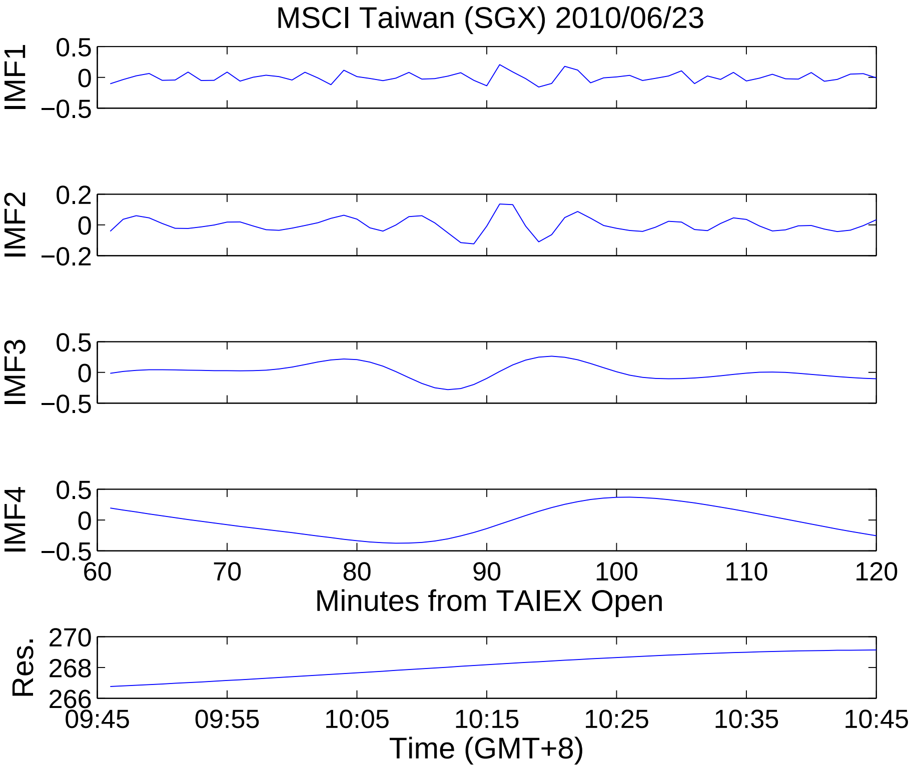


Figure 3.14: This figure illustrate IMFs compoment of MSCI Taiwan (SGX). Note that the time line at X-axis is all the same but represent as two different form.

In figure3.16, the residual of TAIEX realbody has high correlation coefficient at 0.66 with IMF4 of MSCI Taiwan (SGX) close price. In previous IMFs analysis we concluded the realbody have a certain extent represent the close price trend, here we also prove this property. MSCI Taiwan is lagging behind TAIEX in both figure3.15 and 3.16 around five to ten minutes, one of the reasons is information delivery delay. Generally in Taiwan, if you trading or monitoring real-time foreign index or price, stock broker firm usually announce it will be lagging behind few minutes (for ordinary individual investors), the range of delaying minutes from ten to thirty, it depends on the regions where you want to trading or monitoring. Depend on this fact, we can concluded MSCI Taiwan trading at SGX will lagging behind domestic TAIEX for few minutes is reasonable.

The other characteristic of the correlation between MSCI Taiwan and TAIEX is period. From figure3.15 and3.16, the frequency range of high correlation is mid-frequency, the periodic about thirty minutes. The investors trading at MSCI Taiwan are foreign insti-

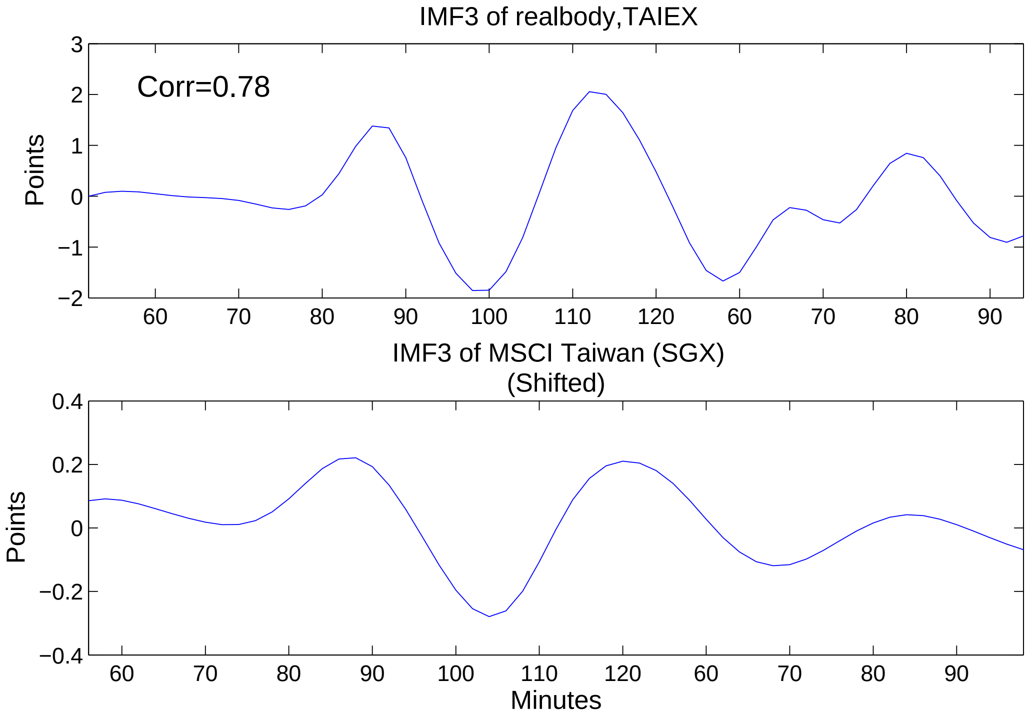


Figure 3.15: The IMF3 correlation between TAIEX and MSCI Taiwan(SGX). MSCI Taiwan lagging around 5 minutes. Note that X-axis is minutes after TAIEX opening.

tutional investors, because of these foreign institutional investors holding massive volume both in MSCI Taiwan and stock market in Taiwan, they cannot trading too frequently, so these foreign institutional investors usually use MSCI Taiwan as a tool to hedge risk of Taiwan's stock market in the mid-term trading periodic.

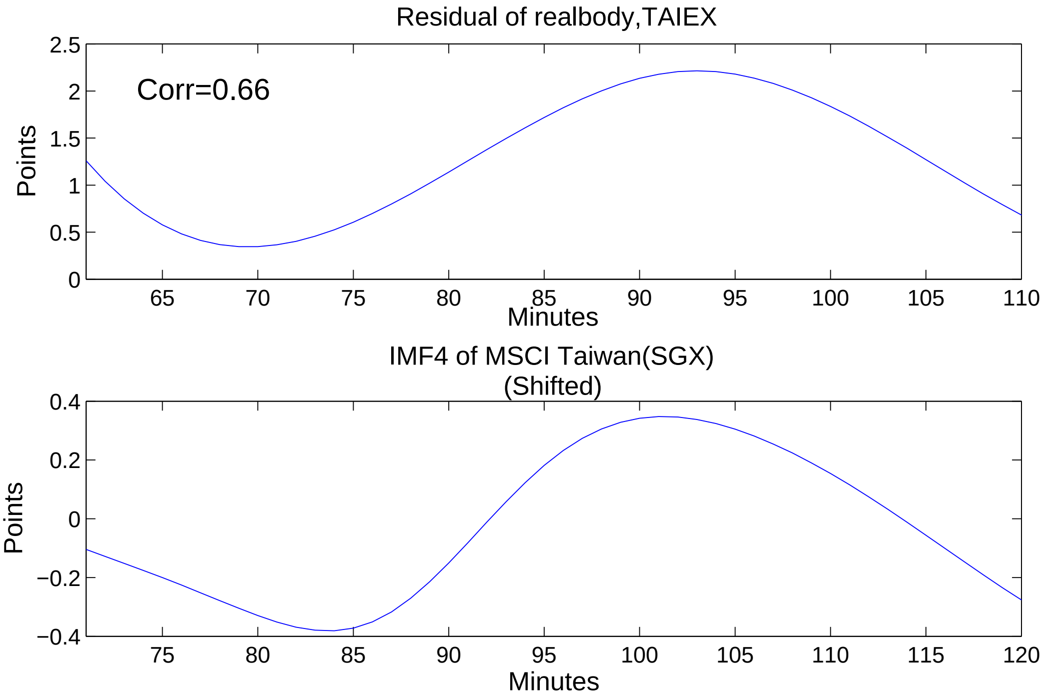


Figure 3.16: MSCI Taiwan lagging than TAIEX around 10 minutes at low-frequency terms.

3.4.4 Gold Futures

As a typical hard currency,since ancient time gold represent wealth. In nowadays, although we don't use gold as currency, it still play an important role of financial market. Unlike other currency and stock price, the value of gold is from the rarity itself (silver also the same), not just determine by simple supply-demand relationship or law. It means gold is expected to serve as a reliable and stable store, which is a fundamental characteristic of hard currency.

The indicator of enhance the hard status of currency include the long-term purchasing power stability, country's political and fiscal condition and outlook, and the policy posture of the central bank. And for gold, its purchasing power is stability. Compare with other currencies, a soft currency considered to fluctuate erratically or depreciate. ”Hard” or ”soft” currency is a relative concept. These hard currency characteristic cause gold become a useful hedging tool especially in low prosperity, financial crisis, runaway inflation (hyperinflation) and in time of war. In these situations, the value of currency is unstable or decreasing, other investment returns may little. For investors, holding gold have less risk than holding currency or other investment especially in these situations, it will cause increasing demand of gold, have an directly effect on gold price rising [51].

Based on the characteristic of gold discussed above, gold price as a indicator present the overall financial environment and gold futures reflect this anticipation of investors. In addition, holding gold futures contracts can reduce the holding physical gold risk, and increasing trading liquidity and market efficiency. Gold futures are significant indicator for financial market and they have similar risk-hedged function with index futures [52], the relationship between gold futures and index futures is what we interest. Following analysis we used the gold future which trading in the Tokyo Commodity Exchange (TOCOM) as contrast. TOCOM is one of the largest precious metal exchange in the world. We choose the overlapping trading time between TOCOM and TAIEX to observe their correlation.

Table 3.6 illustrate IMFs of the TOCOM gold futures, the high frequency period is about three minutes, mid frequency period is about eight minutes and low frequency period is about thirty minutes, it is similar to TAIEX realbody in IMF1 IMF2 IMF4 but lack of fifteen minutes period. We can see in figure3.17 and figure3.18 , in mid-frequency and low-frequency have highly correlation coefficient at 0.75 and 0.73. In the mean time we noticed that TOCOM leading beyond about thirty minutes, it display the opposite situation with TAIEX and MSCI Taiwan (SGX). This difference because TOCOM gold is the major market/commodity, TAIEX is a minor market index of global financial market. In another words, TOCOM gold futures affect TAIEX index futures. But the relationship between TAIEX and MSCI Taiwan (SGX) is opposite because both TAIEX and MSCI Taiwan (SGX), their underlying market are Taiwan stock market, although they construct in different way. For Taiwan stock market, the real-time information TAIEX(domestic) is prior than MSCI Taiwan(foreign).

Table 3.6: IMFs of TOCOM-Gold futures

TOCOM-Gold	Mean period	Correlation	Power percentage
IMF1	3.56	0.19	4.58
IMF2	7.54	0.35	7.84
IMF3	29.29	0.03	24.65
IMF4	32.75	0.55	12.22
Residual	42835.21	0.7	87.71
Sum	N/A	N/A	136.99

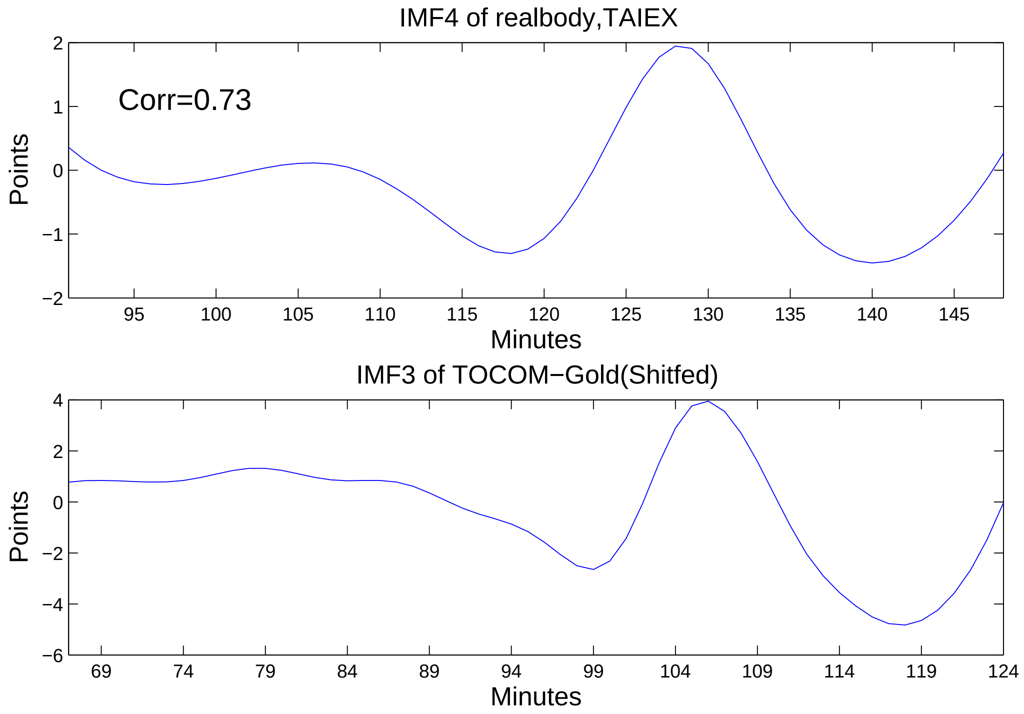


Figure 3.17: Contrast of Low-frequency

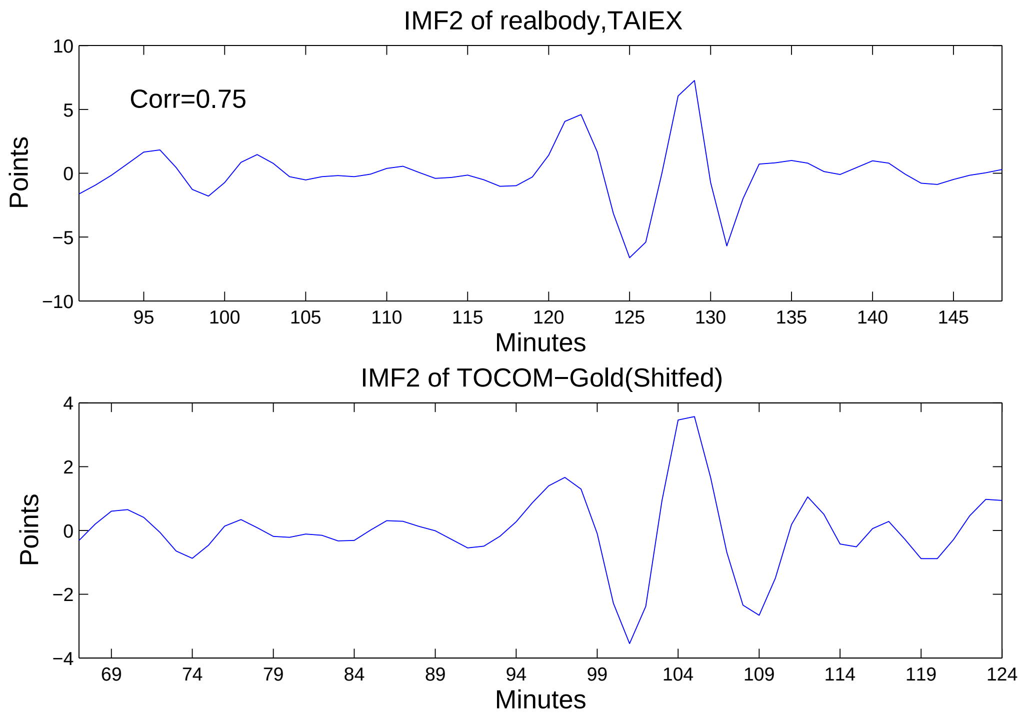


Figure 3.18: Contrast of Mid-frequency

3.4.5 TWD/USD Currency Rate

In the final stage of this section, we discuss the relationship between currency rate and TAIEX. We may know US Dollar exchange rate affects TAIEX, but how? Through analyze TWD/USD Currency Rate, we try to figure out their relationship. Currency Rate have two different rate: selling rate and buying rate. Selling rate is the currency rate for import customs declaration and buying rate is for export customs declaration, consider about Domestic Corporation export is a part of GDP, we use buying rate as contrast [53]. Unlike other ordinary Intraday trading data, the raw data of currency rate seldom change (illustrate at figure3.19), in order to analyze this situation, we use full trading day three hundred trading minutes close price decompose IMFs instead of training length sixty minutes of trading model. Figure3.19, figure3.20 and table3.7 illustrate the IMF of TWD/USD Currency Rate (on 2010/05/21 as represent).

In figure3.21 and figure3.22 we can see the highly inverse correlation in low-frequent IMF6 and IMF7 terms. Actually, it still positive correlation, because of the unit of currency rate. Higher TWD/USD currency rate means you have to pay more NTD to get one US

Dollar, NTD reduce its value essentially. On the other hand, TWD/USD currency rate become lower means you can pay less NTD to get one US Dollar, essentially value of NTD increase.

Because of the frequency of raw data of TWD/USD currency rate is low, the correlation of IMFs also appears in the low-frequency IMF7 and IMF8, we can see when NTD price falling (the TWD/USD currency rate rise), The TAIEX also fall. In monetary economics, money supply greater than its demand will cause the currency value decrease, if money demand greater than its supply will cause the currency value increase.

When foreign institutional investors intent to invest in Taiwan, they shall exchange the US Dollar into NTD, it will directly increase the demand of NTD, the currency value of NTD will rise (amount of TWD/USD will fall); if foreign institutional investors retreat in Taiwan's investment market, the demand of NTD decrease, currency value of NTD will fall (amount of TWD/USD will rise). In conclusion, low-frequency exchange rate IMF terms can represent currency value and the tendency of foreign institutional investors in a certain extent, and the low-frequency IMF of TAIEX futures reflect this phenomenon. [54] [55]

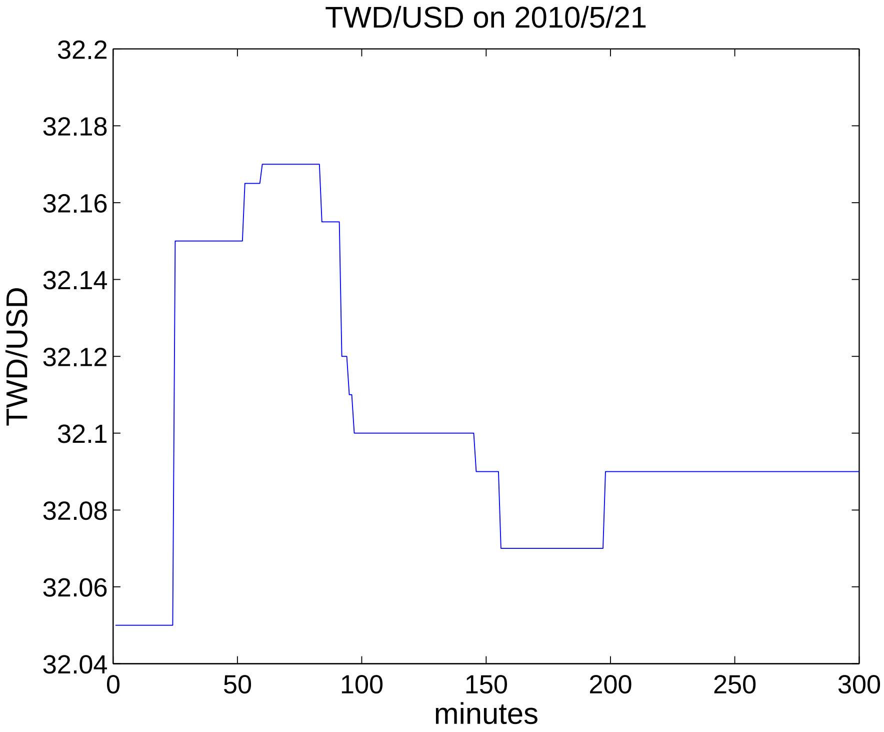


Figure 3.19: Raw Data of TWD/USD Currency Rate

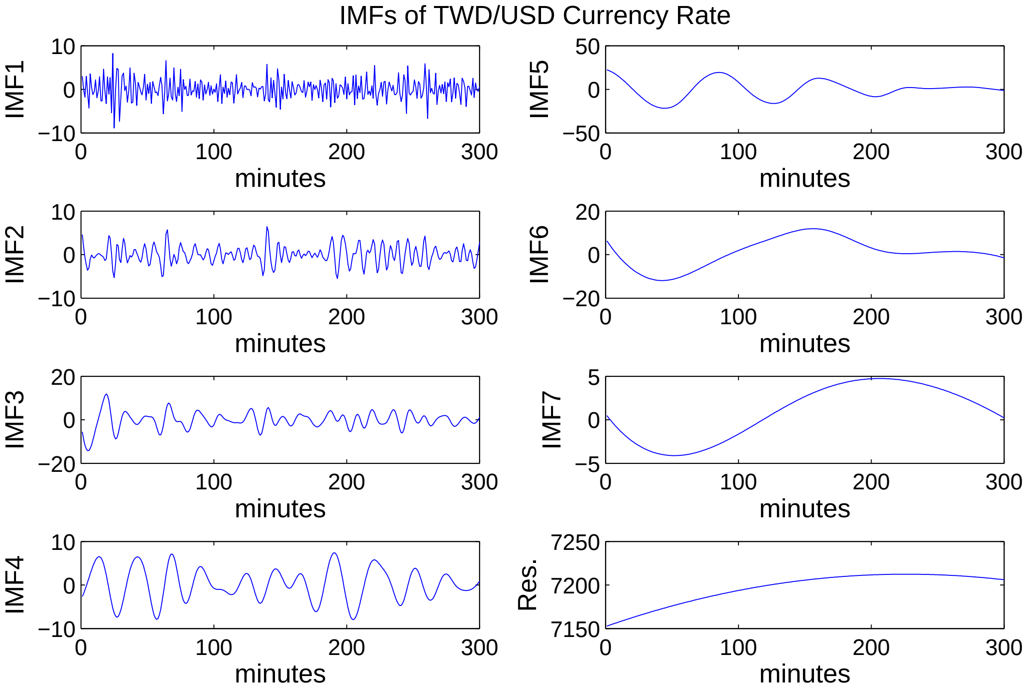


Figure 3.20: All IMFs of TWD/USD Currency Rate

Table 3.7: IMFs of TWD/USD Currency Rate

TWD/USD	Mean period	Correlation	Power percentage
IMF1	2.99	0.03	0.69
IMF2	6.94	0.13	0.69
IMF3	16.97	-0.04	2.01
IMF4	19.66	0.2	5.08
IMF5	84.53	0.75	10.09
IMF6	90.36	0.89	36.62
IMF7	284.05	0.7	3.96
Residual	1494515.39	0.33	0.94
Sum	N/A	N/A	60.08

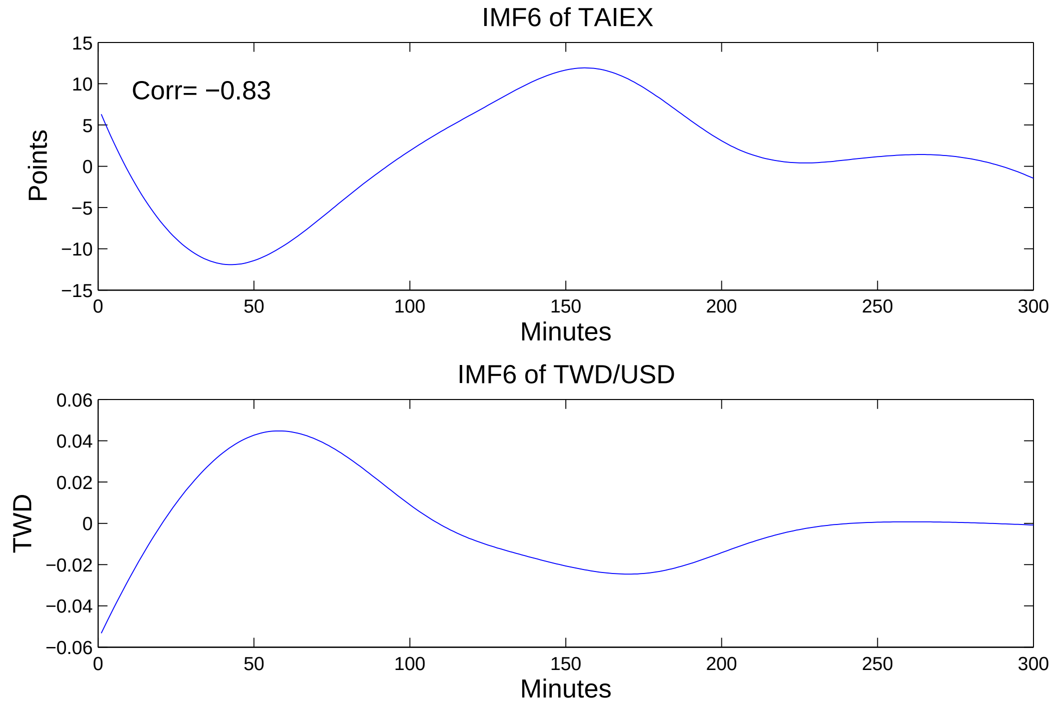


Figure 3.21: Compare of IMF6

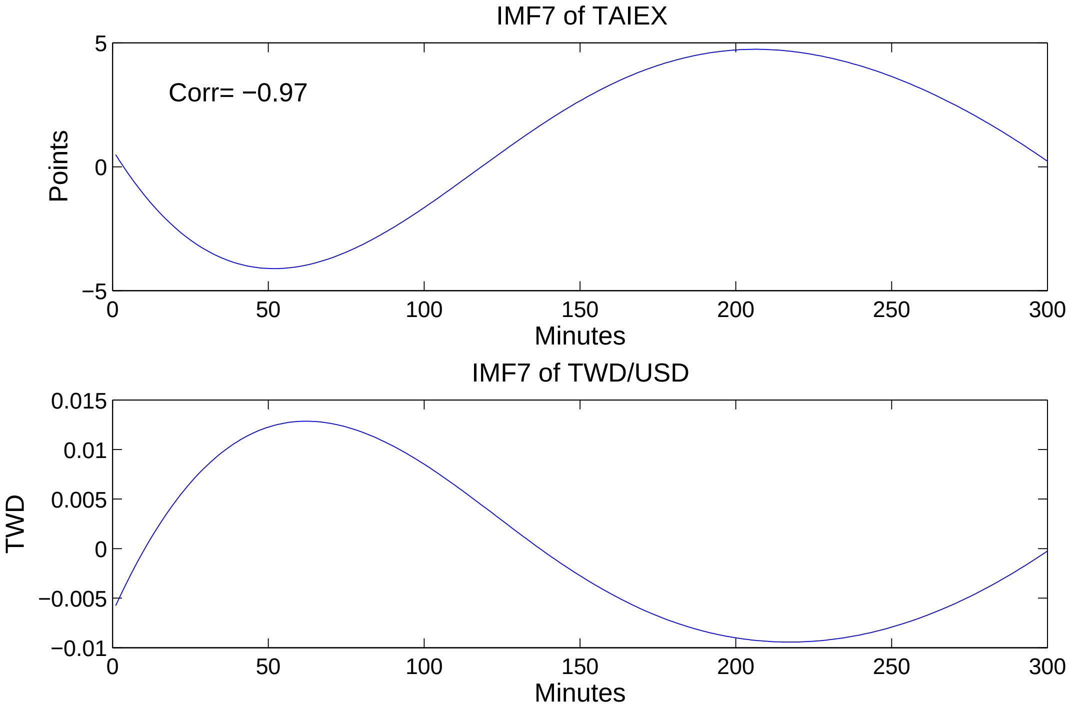


Figure 3.22: Compare of IMF7

Chapter 4

Hybrid Algorithmic Trading Model

4.1 Constructing Model

4.1.1 The Moving Window Method

In previous section we introduce the operating principle of BPNN network, but how to use BPNN network in the time series forecasting? Here we introduce the moving window concept as the principle of network training [56]. Following figure4.1 present the training process of moving window. In this thesis we choose 60 minutes as the length of training data. In the beginning of training, we choose a training set meanwhile as input set and target set. Then we eliminate the first data point of input data and the last data point(here is the 60^{th} data point) of target data. Then we paired the dataset, the 1st input pair up 2nd target; the 2nd input pair up 3rd target, and so on. After training, the network obtain the capacity of simulate the $n + 1$ point from n^{th} data point. Since we choose 60 units as training length, we use 60 minutes predicting 61^{st} minute. As similar concept, we use 60 minutes training to estimate ARMA model parameters, then we used this fitted ARMA model forecasting 61^{st} data. Thus, both in BPNN and ARMA, we use 60 minutes predict the 61 minute and then shift one unit to predict next (here is 62^{nd})minute. Use this concept, we can forecast candlestick charts real-body time-series, adopt the predicted value's sign as forecast direction of price trend of next time unit.

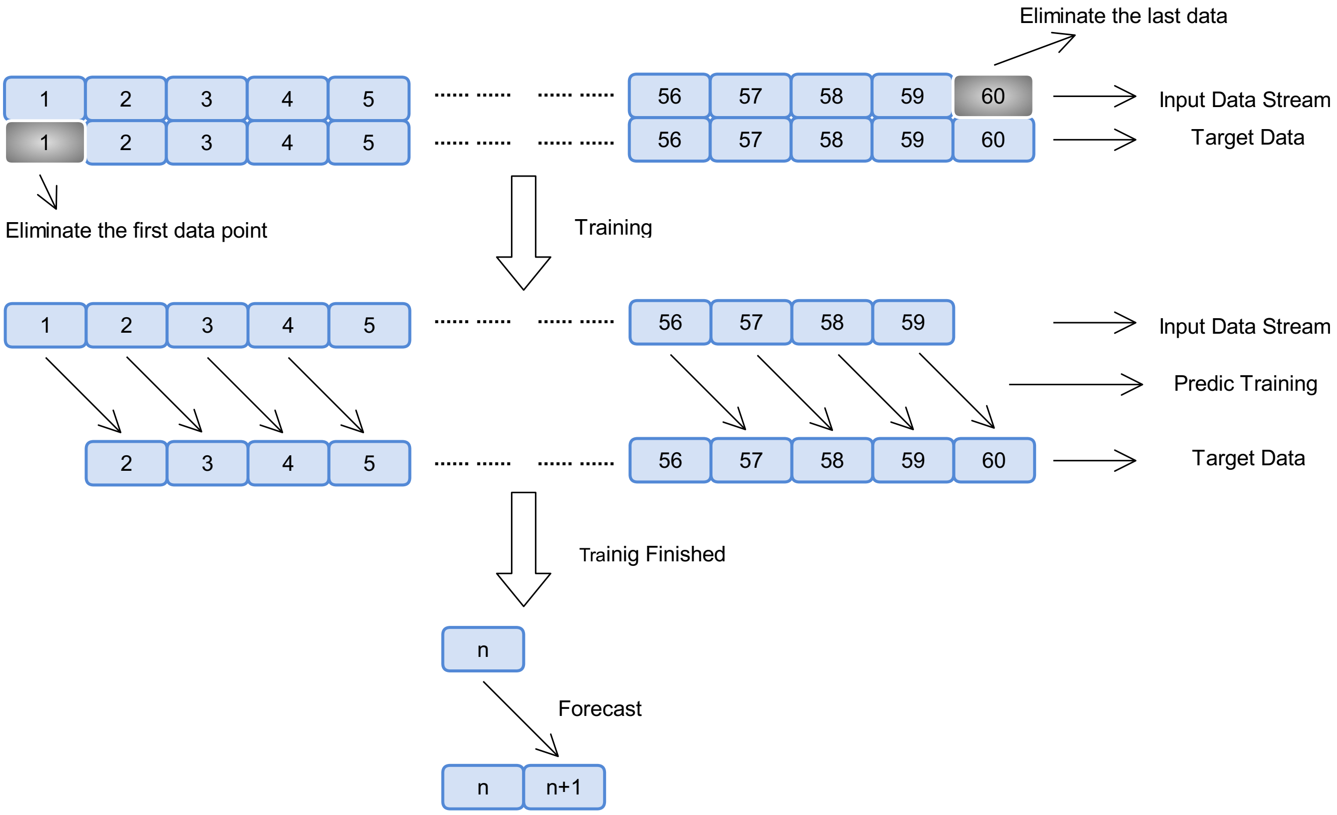


Figure 4.1: The Moving Window Method

4.1.2 The construction models

In this thesis, we develop the EEMD-BPNN structure from En Tzu Li (2011) [57] and Chen Yuan Hsiao (2013) [58]. We adopt the standard structure as following figure4.4. In this structure, we decompose the raw data stream into individual IMFs and residual. And then we use individual BPNNs training specific IMFs and the residual one to one, one BPNN training one data stream. Chen proved this structure have better forecast performance than training all IMFs in one BPNN network, which is the structure Li used in her thesis. The number of IMFs we obtain generally depend on the length we decompose, if we decompose longer data stream, we obtain more IMFs(if the stop criteria is fixed). In this thesis, based on Li and Chen's study, we compare the different decomposed length of data, found out the 135 and 120 data unit length are not so suitable for per minute high-frequent data, then we adopt 60 units(minutes) as the decomposed length. Thus we start to generate forecast result at the 61^{st} minute of a trading day.

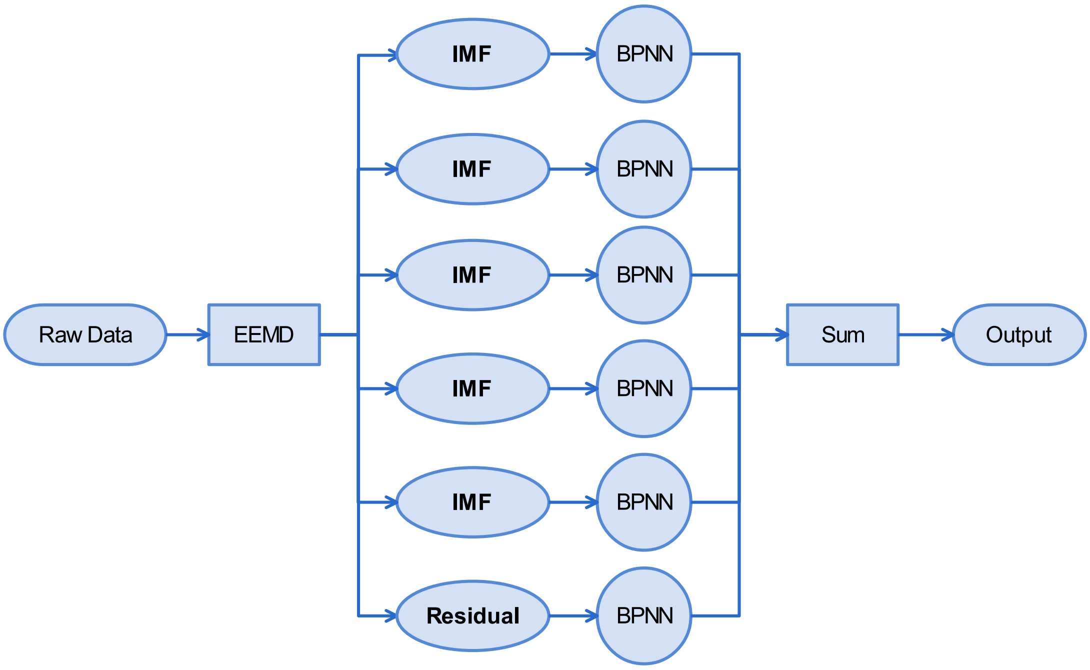


Figure 4.2: The standard EEMD-BPNN construction

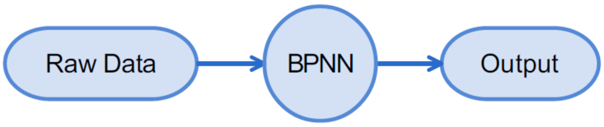


Figure 4.3: The Single BPNN model

Based on Chen's standard EEMD-BPNN structure, we also develop other model structure form. Figure4.4 demonstrate the structure of model Mk-1. This model using plural data stream as the input set in single BP-network. We believe the upper and lower shadow data streams should put in the high frequency terms of IMFs, because shadows indicate the price fluctuation during the candlestick's time interval, this concept we mentioned in previous section2.5 and3.4.1. After the preliminary test in following section4.2.2, we confirmed this concept is effective. Next, in model Mk-2 we consider the different characteristic between ANNs and ARMA[1], using the shadow-NN network for high frequency IMFs, and adopt ARMA(1,1) in mid-frequency and low-frequency terms of IMFs, this structure also effective. In model Mk-3 and Mk-4, we adopt the normal single input data stream in mid-frequency IMF in BPNN as contrast set, try to find more effective structure in mid-frequency IMFs.

During performance backtesting, we found out ARMA(1,1) take more estimate time than BPNN training[2]. In addition, since ARMA model is designed for estimate stationary time series [32], low frequency IMFs and residual have longer trends, it may cause ARMA model reduce the effectiveness. In Mk-5, we use BPNN replace the ARMA in low frequency IMFs and residual, since BPNN have advantage for processing non-stationary series. Consider about the characteristic we discussed above, Mk-5 adopt shadow-NN in high-frequency IMFs, ARMA(1,1) in mid-frequency IMFs and normal BPNN in low-frequency IMFs, try to find out better hybrid EEMD-based forecast model configure. We also adopt single BPNN configure as contrast benchmark.

[1] In Chen's study, he thought ARMA model is more suitable for the data series with the property of more linearity and less fluctuation.

[2] In one-year backtesting, adding one more ARMA(1,1) in hybrid model will take longer computing time, usually about 15 thousand seconds long

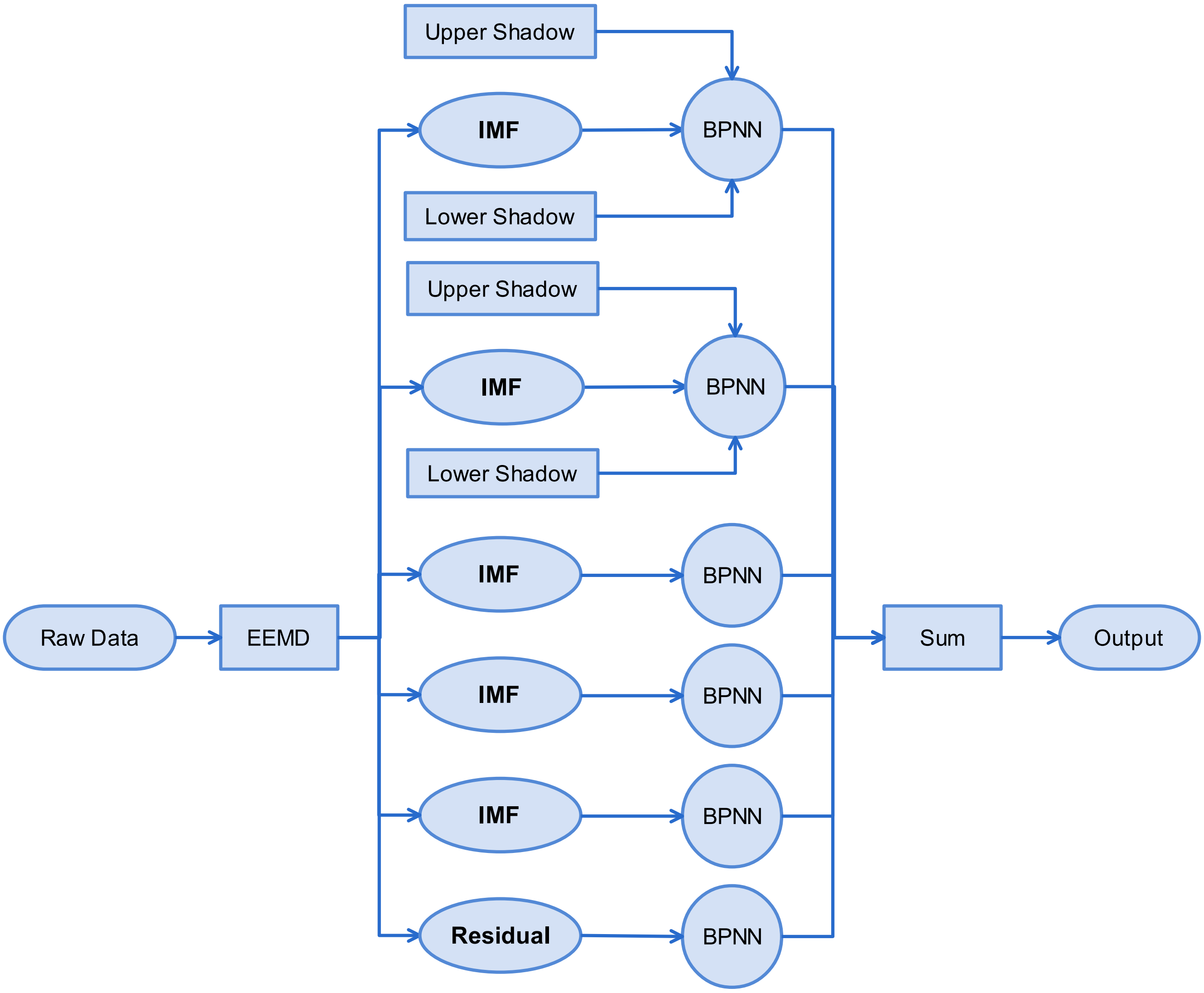


Figure 4.4: The Shadow-NN configure, model Mk-1

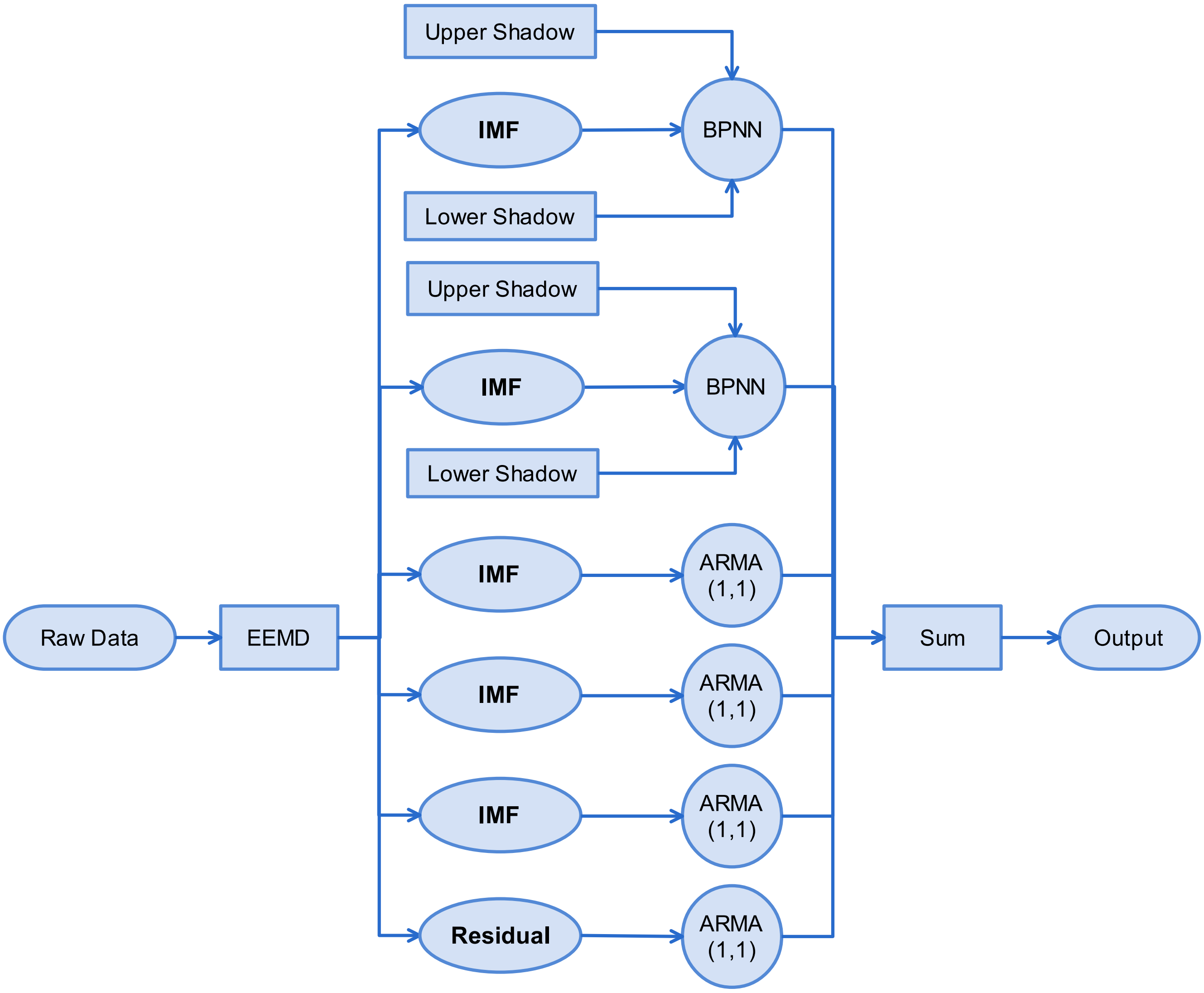


Figure 4.5: The Shadow-NN-ARMA configure, model Mk-2

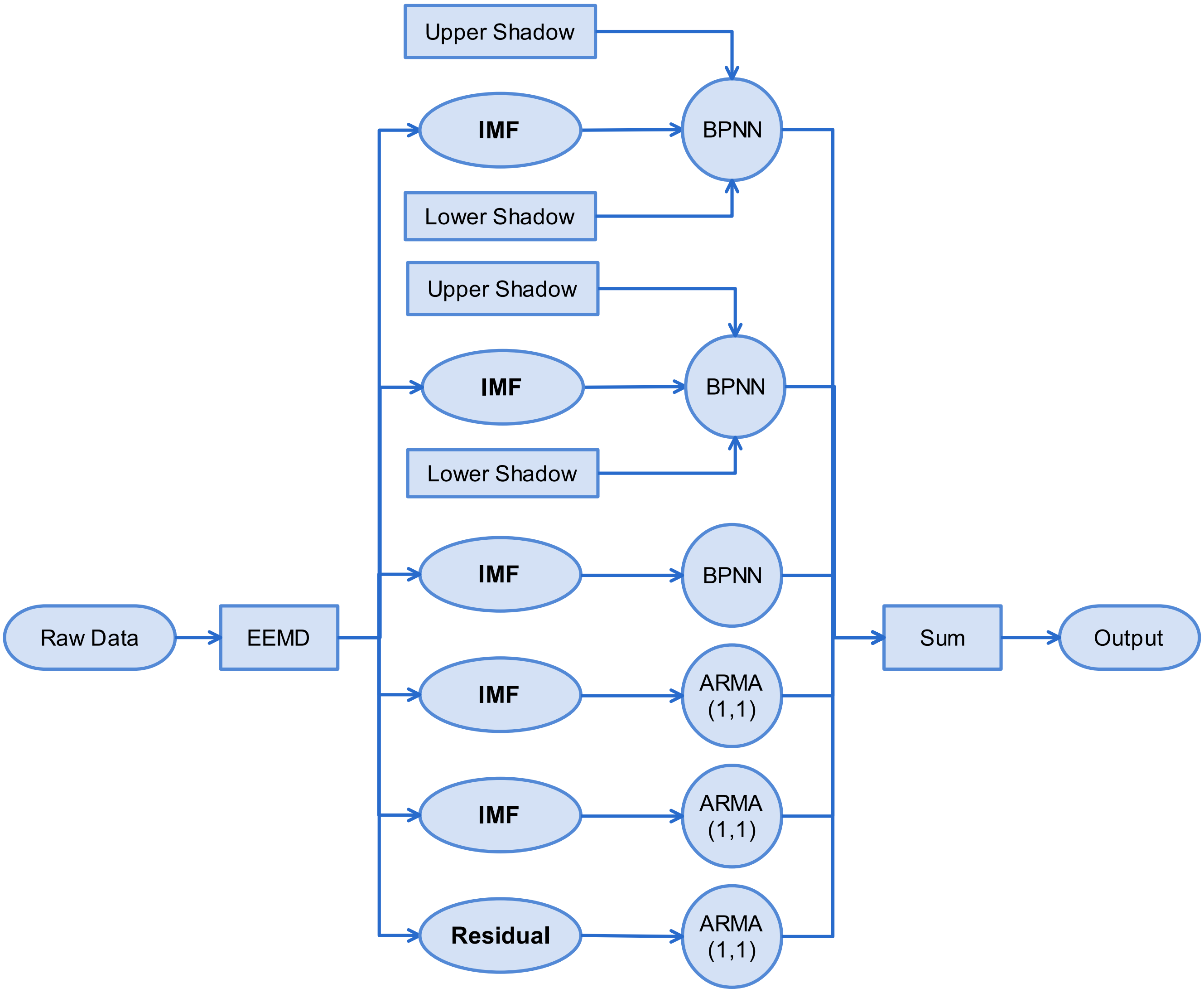


Figure 4.6: The Shadow-NN-ARMA configure (Ver.2), model Mk-3

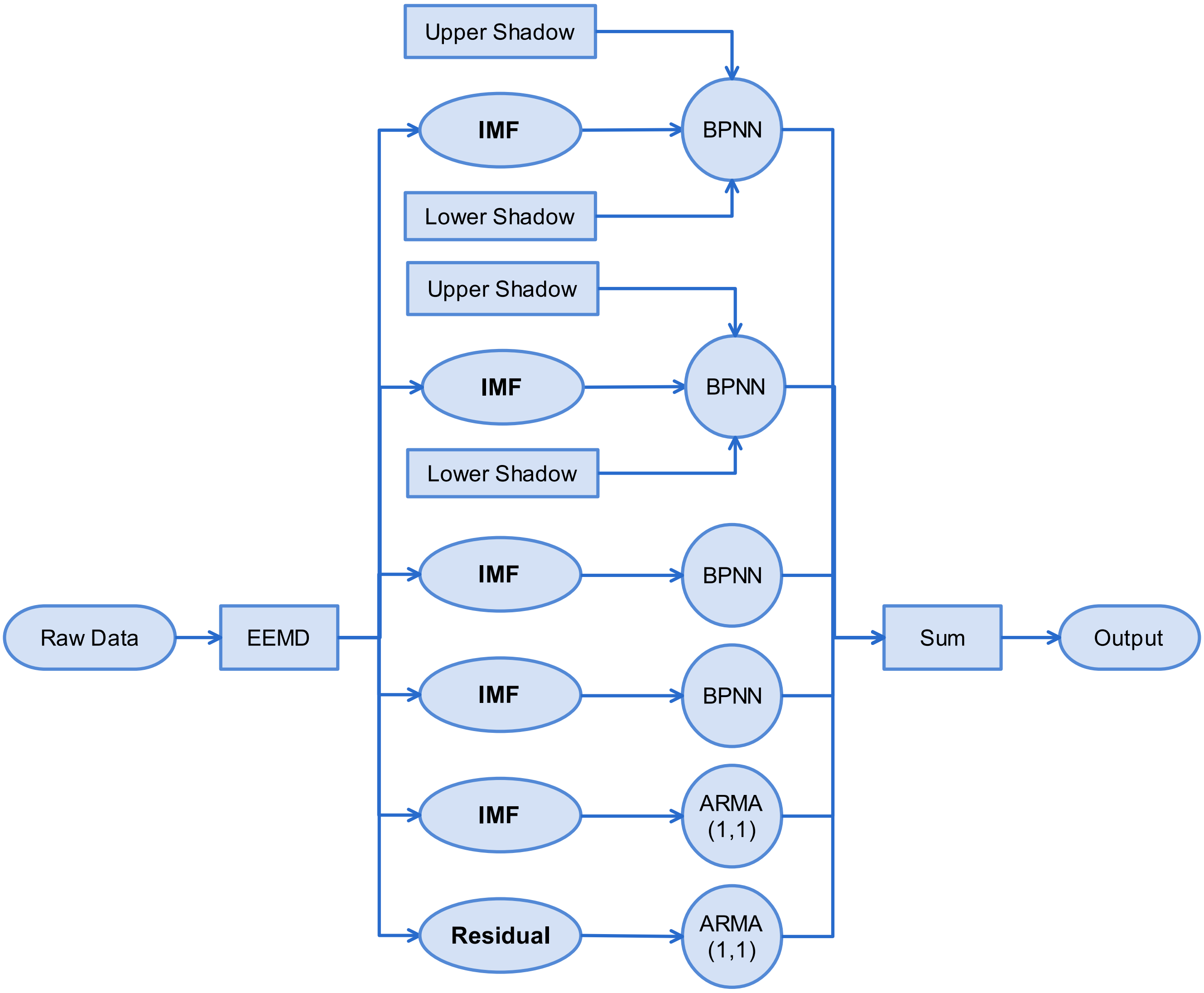


Figure 4.7: The Shadow-NN-ARMA configure (Ver.3), model Mk-4

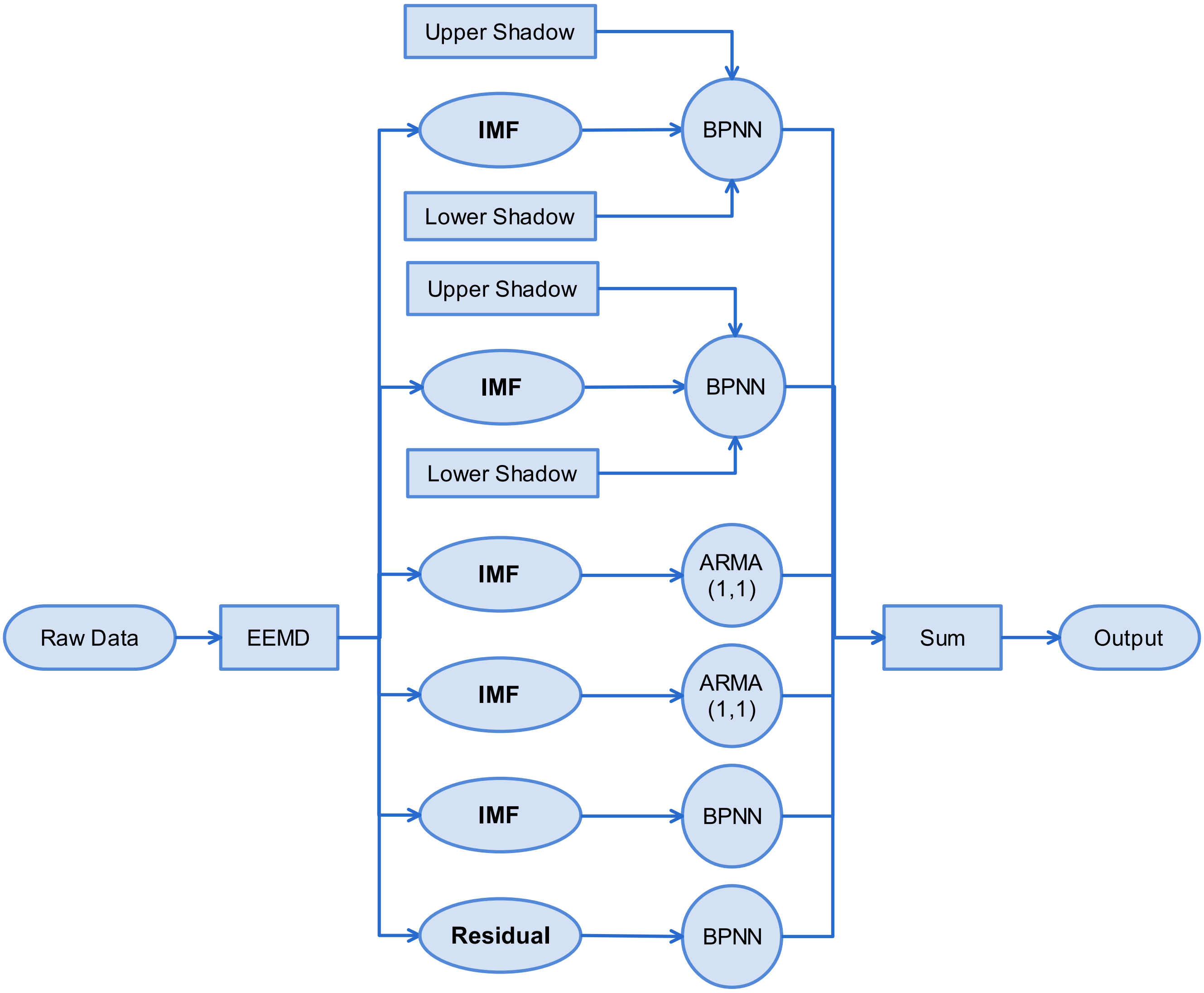


Figure 4.8: The Shadow-NN-ARMA configure (Ver.4), model Mk-5

4.2 Performance

We use 2007 2011 TAIEX futures as testing data. In this five-year period, contains 60 months,256 weeks, 1237 trading days, 371315 trading minutes. The minimum holding period we calculate is one month. Here we adopt correct trend, holding period return (HPR), effective annual yield(EAY) and Sharpe ratio. Note that to simplify the performance compare, we don't consider the margin and other contract detail in following performance computing. Following section we will introduce these indicators.

4.2.1 Evaluation Indexes

Here we review $Correct\ Trend$ again. Correct trend measure the success ratio of forecast the future price trend, in section3.3 we used once to measure the randomness of market, here we rewrite its form. Its definition is quite simple:

$$Correct\ Trend = \frac{Number\ of\ Successfully\ Forecast\ Price\ Trend}{Total\ Number\ of\ Forecast\ Units} \tag{4.1}$$

We shall note that the higher correct trend doesn't direct indicate higher performance, because the scope of price change influence the performance more than correct trend ratio. For instance, a lower correct trend model may outperform a higher correct trend model, because the higher correct trend model forecast many price change successfully but usually within small scope; although the lower correct trend model forecast successful not many, but it forecast larger scope of price change successfully. Obviously in this example, this lower correct trend model have higher forecast quality performance. This phenomenon we can see in next section.

Following we introduce $Holding\ Period\ Return$(HPR):

$$HPR = \frac{Income + (P_f - P_i)}{P_i} \tag{4.2}$$

Where the P_i is initial value, P_f is end-of-period value, income is any intermediate gains during the holding period.

Here we don't have gains during the holding period, we eliminate the unused income term, then:

$$HPR = \frac{P_f - P_i}{P_i} \tag{4.3}$$

The $Geometric\ Mean$ is often used in calculating investment returns over multiple periods or when measuring compound growth rates.

$$1 + R_G = \sqrt[n]{(1 + R_1) \cdot (1 + R_2) \cdots (1 + R_n)} \tag{4.4}$$

Where $R_t =$ the return of period t and each items should be non-negative. Note that when you compute geometric mean or other compound rates, you should add 1 (represent as principle) before calculate and subtract 1 back to get the result.

Base on HPR and the concept of Geometric Mean, we can extend to $Effective\,Annual\,Yield$ (EAY):

$$EAY = (1 + HPR)^{\frac{365}{t}} - 1 \tag{4.5}$$

EAY is annualized value, based on a 365-day year, calculate in compound interest. By definition, actually EAY is the annualized holding period return.

The $Sharp\ Ratio$ is widely used for investment performance measurement of excess return per unit of risk, William Forsyth Sharpe proposed it in 1966[3]. The definition of Sharp ratio as following:

$$Sharp\ ratio = \frac{\overline{r_p} - r_f}{\sigma_p} \tag{4.6}$$

Where:

- $\overline{r_p} =$ portfolio return
- $r_f =$ risk-free return
- $\sigma_p =$ standard deviation of portfolio returns(investment risk)

The quantity $\overline{r_p} - r_f$ referred to as the ”excess return” of Portfolio P, it means the extra

[3] William Forsyth Sharpe is a Nobel Memorial Prize laureates in Economics(1990)

reward that investors receive for exposing themselves to risk. The basic concept is : If you want to creative a risky portfolio, the required return should greater than risk-free rate, and maximize the return, minimize the risk. Larger Sharp ratio is better [59]. In general, sharp ratio greater than 2 is very well, greater than 3 is excellent performance.

4.2.2 Performance Analysis

Table 4.1: Average Correct Trend Ratio

	Standard	Mk-1	Mk-2	Mk-3	Mk-4	Mk-5	Single BPNN
Average	0.428	0.429	0.429	0.428	0.428	0.430	0.435
SD	0.023	0.023	0.024	0.023	0.023	0.024	0.023
Rank	5	3	4	5	5	2	1

In table4.1 we can see the average correct trend ratio . We find out the single BPNN model higher than all other models, and the rest have another correct trend level themselves. But single BPNN is the worst performance model we can see in following analysis result. It is because single BPNN model predict correct trend more, but the correct trend it predict usually are little price change amplitude. Other model have lower correct trend, but they predict higher price change amplitude successfully. In other words, other models have better predict quality, just like we discussed in last subsection.

HPR is the simplest way to measure the performance of portfolio and it still effective. In table4.9 we illustrate the result of nine different holding periods in the total five-year period we test, choose 25 times average performance and standard deviation. In model Mk-1 we can see the strategies of input shadow in high-frequency IMFs; and more further, in Mk-2 put ARMA in mid-frequency and low-frequency IMFs. Both strategies enhance the performance successfully.

In model Mk-2 to Mk-5 we can see ARMA in mid-frequency more suitable than low-frequency, the same result as we discussed in section4.1.2. On the other hand, since ARMA take more time to estimate variable, less ARMA means faster computing speed. Model Mk-5 save around thirty thousand seconds than Mk-2 but their performance just little different. In future work we can further improve efficiency both in performance and computing speed.

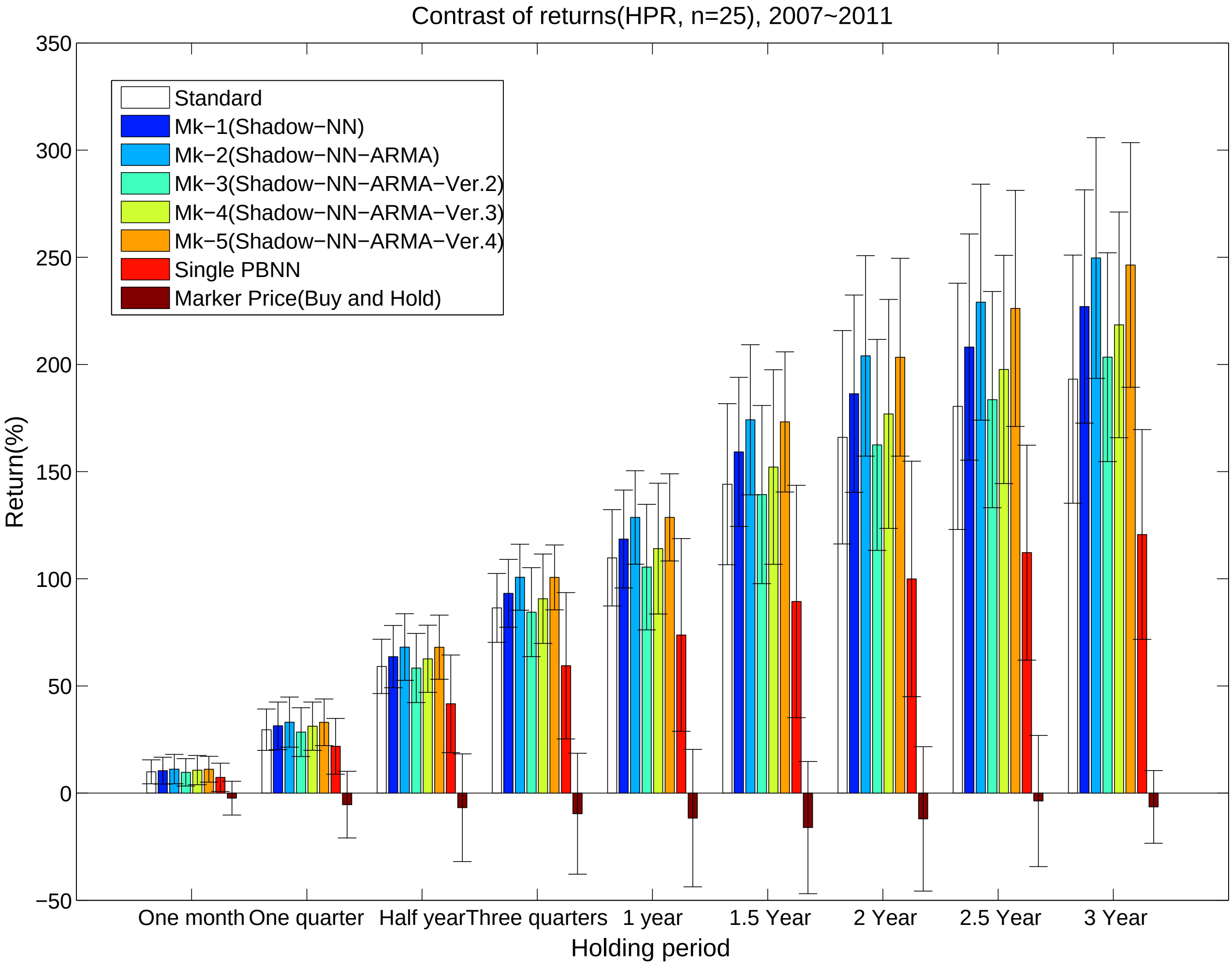


Figure 4.9: Contrast by HPR, error bar is standard deviation

Here we introduce another indicator: maximum drawdown. Maximum drawdown is a simple concept, it is the largest range of price or portfolio deficit. Figure4.10 is the monthly TAIEX price in overall test period, we find the maximum drawdown during 2008/04 to 2009/01, it's 4671.95. Then table4.2 is the performance during this period. We can find besides single BPNN model, all other models overtake the amount of maximum drawdown. One of the major reason of this phenomenon is the important property of high-frequency trading. In the fractal theory, stock price series have similar shape in different scale (see section1.1), it means in short term still have many small price fluctuation, although they may not affect long term price trend. High-frequency trading aim to catch these short term small price fluctuation, make little profit in small scale many times,

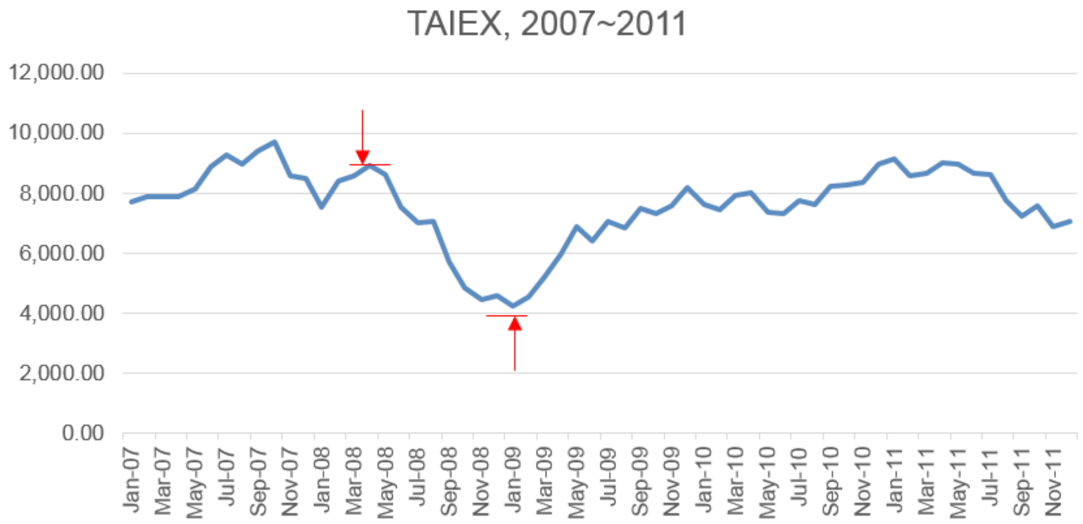


Figure 4.10: 2008/04-2009/01, Maximum Drawdown:4671.95

Table 4.2: Compare with Maximum Drawdown, 2008/04-2009/01

	Standard	Mk-1	Mk-2	Mk-3	Mk-4	Mk-5	Single BPNN	Max Drawdown
Performance	6647	7707	7803	5829	7287	7617	2279	4671.95
Rank	5	2	1	6	4	3	N/A	N/A

finally the overall profit can overtake the long term price trend.

Table 4.3: HPR of individual year

Year	2007	2008	2009	2010	2011
Standard	124%	107%	95%	12%	22%
Mk-1	131%	113%	100%	29%	33%
Mk-2	135%	126%	113%	30%	38%
Mk-3	131%	101%	67%	32%	40%
Mk-4	136%	120%	65%	34%	38%
Mk-5	134%	126%	113%	30%	35%
Single BPNN	104%	79%	28%	18%	8%

EAY using the geometric concept of compound interest normalize the over one-year period HPR convertible to a comparable one-year period return. In figure4.11 we normalize the over one-year period HPR to EAY, we can see the longer holding period, the less EAY return. Why? We can find the reason in table4.3. Table4.3 list all the models annual return per year, we can see performance before 2010 are extremely high, but in 2010 and 2011 are relatively low, although it still have relatively higher return. During 2007 to 2010 is the financial crisis period, the price fluctuated wildly in this period, it offer high-frequency trading excellent chance to earn excess return. In fact, in this pe-

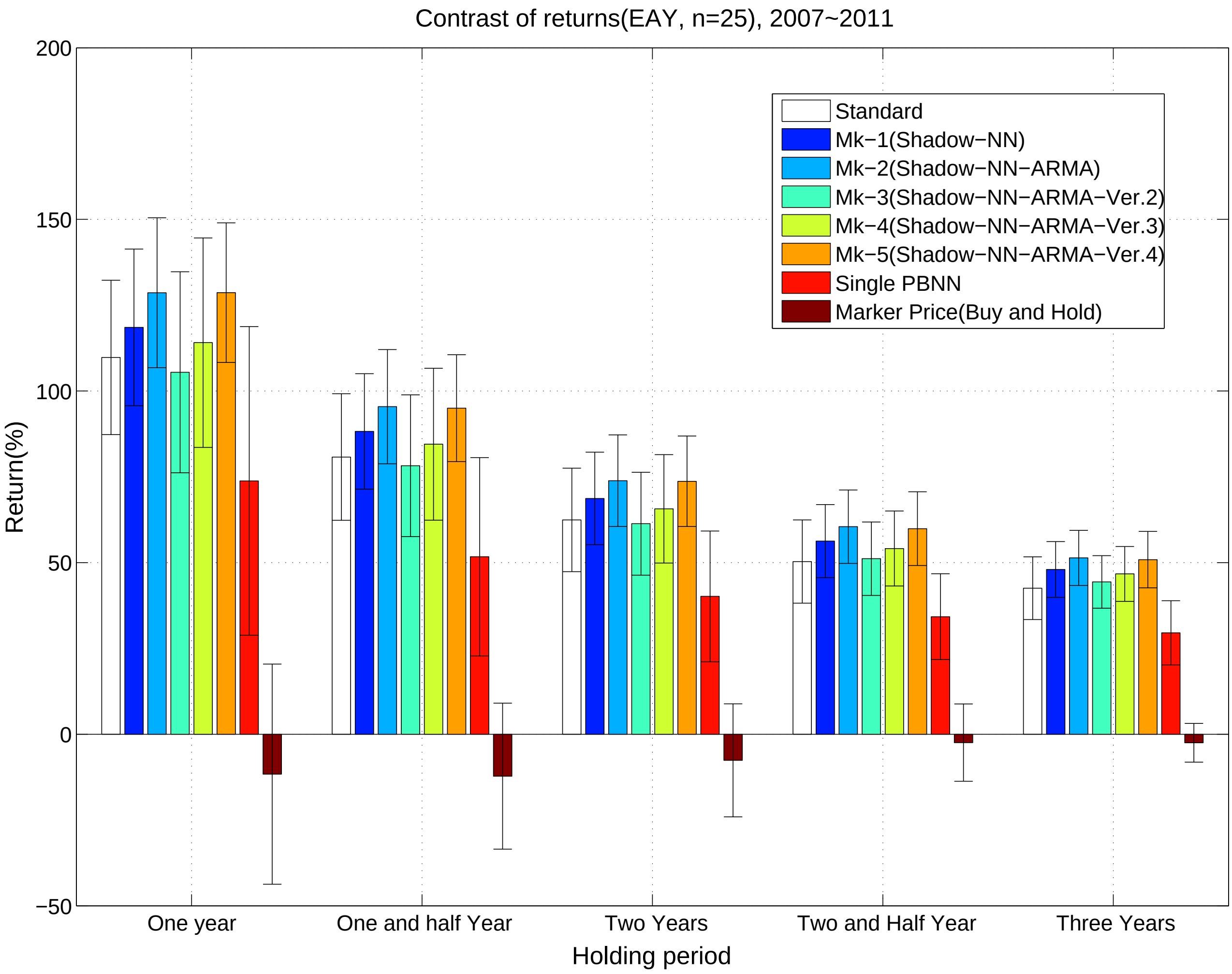


Figure 4.11: Contrast by EAY, error bar is standard deviation

riod the hedge funds adopt HFT and other quantitative strategies (just like the Medallion Fund of Renaissance Technologies, the company funded by James Simons, we introduce in section1.1)usually outperform adopt traditional strategy hedge funds. But in 2010 and 2011 the price backing and filling, we still can earn relative higher return, but cannot earn like fluctuated wildly period, it pull down the overall EAY in longer period. The reason described above caused the EAY in longer period less than shorter period.

Then we see the Sharpe Ratio. Sharpe Ratio is a useful indicator for performance measure and widely used in many different investment portfolio, it consider the portfolio risk and risk-free rate. The concept of Sharp Ratio is how much excess return can obtain through every risk unit increased. Since the different choosing of risk-free rate, Sharp Ra-

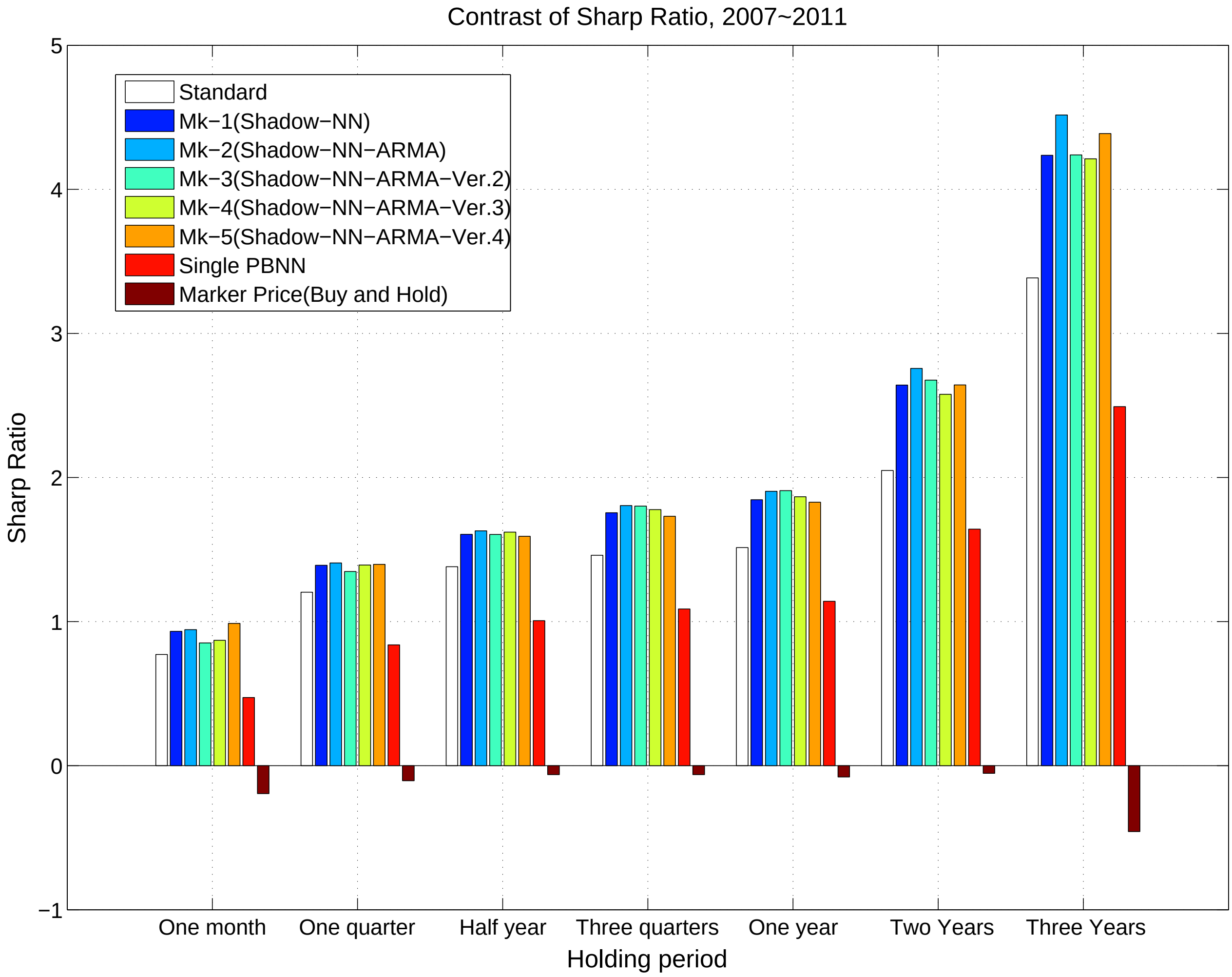


Figure 4.12: Contrast by Sharp Ratio

tio in different standard cannot compare directly. In United States, risk-free rate usually adopt Treasury Bond, but in Taiwan government seldom issue Treasury Bond, adopt Treasury Bond as the risk-free rate in Taiwan is not suitable. Here we adopt average deposit interest rates of four major banks of Taiwan as risk-free rate, it's one of the general used risk-free rate in Taiwan. Figure4.12 compare the Sharp Ratio between different model. The two major factors of higher Sharp Ratio are higher return and lower risk. In general, Sharpe Ratio higher than 2 or 3 is good performance. We can see shorter period have lower Sharpe Ratio, until over two years holding period the Sharpe Ratio finally greater than two, it is opposite trend with EAY, and we just discuss EAY above. Why Sharpe Ratio and EAY show the opposite trend? Because Sharpe Ratio consider about the impact of risk. In shorter holding period, the occasionally appear monthly deficit, it cause higher risk and pull down the Sharpe Ratio value; in longer period, the impact of these sporadic monthly deficit diluted through longer holding period, decrease their impact (risk), then pull up its Sharpe Ratio value.

Here we can seen although Mk-1, Mk-3 and Mk-4 have different return in HPR, but in long term their Sharp Ratio quite close , because their risk is different. Especially compare the return between standard model and Mk-3, in HPR these two models have little difference but they have huge gap in Sharp Ratio, the main reason is Mk-3 have lower standard deviation(risk). These phenomena indicate Sharpe Ratio is highly sensitive about risk, it is a useful indicator to determine the overall performance of portfolio which included their risk.

Chapter 5

Conclusion

5.1 Summary

In this research, we use the high-frequency data to build trading model, find out the correlative relationship between IMFs and the other typical indicators such as gold, volume, currency rate and try to apply some properties of these result into forecasting models. For analyze the big data which generated from high-frequency time scale, we apply parallel process SPMD and batch process to execute the backtesting process. SPMD and batch process accelerate computing speed and improve the efficiency of processing big data, plus adopt the fast-EEMD algorithm, thus we can execute five-year period backtesting. We build models for high-frequency trading based on candlestick charts, and find the meaning and impact indicators of IMFs in high-frequency time scale, try to use some of these properties to build model; based on combine EEMD method and ANN, we introduce shadows and ARMA(1,1) into the hybrid model configure, aim to enhance the performance of forecasting. As the performance result, we develop the hybrid model configure and the concept we used in build models both have potential to develop further in high-frequency trading field.

5.2 Future Works

By focus on the future works, we still have much improve space: at first, the major problem is computing speed. In high-frequency trading have high require for computing performance. Although we adopt the fast-EEMD and SPMD configure, it still take around eighty thousand to one hundred thousand seconds for annual backtesting of each model. Long computing speed limit the experimental test and develop. Aim to solve this problem, we have three solutions: software, hardware and cloud computing. In software, we should develop the hybrid programming between C code and Matlab, for increase computing performance. The fast-EEMD subroutine already apply this skill, but it only for Windows platform, it must also develop for Linux platform such that can applied in high performance computing cluster.

Second, in hardware we can develop performance by adopt GPU(Graphics Processing Unit) parallel process. The parallel platform we use in this thesis is multi core CPU , GPU parallel can accelerate computing more efficiency than normal CPU parallel in several situation with relative low cost, and GPU parallel also can apply under Matlab environment [60] [61].

Third, we can adopt the cloud computing in data processing. Recently cloud computing is a popular solution for big data processing, it's started from Google, and the open source version Hadoop debut, some company like Amazon Web Service (AWS) offer the cloud computing platform to clients. Based on the performance of processing big data, cloud computing is a potential solution for processing high-frequency even ultra-high frequency data [62].

Since in this thesis we focus on enhance the forecasting performance by different configure. Besides this topic, in the next work we can develop in four major direction: programming structure, optimize the forecast core model, reducing trading cost and collocate with efficient trading strategy.

For programming, SPMD is a efficient way for processing big data, it is suitable for backtesting but not for real-time trading. Since SPMD separate the dataset to different worker, it accelerate the overall dataset processing not the (real-time)single data unit processing. When we execute real-time trading, we should using the parallel structure that accelerate single unit processing speed. For forecasting model structure, no matter EEMD, BPNN or ARMA model, we still have much space to enhance their individual performance, for instance optimise the weights and bias of Neural Network through genetic algorithm and find the more suitable lag operator of ARMA(p,q) model; allocate BPNN and ARMA to more appropriate IMFs, improve the accuracy of allocate.

For trading cost and trading strategy, in this thesis we don't consider the contract detail in the performance measure, but trading cost most included in the contract. Here we don't aim to reduce trading cost, we aim to improve the forecasting performance of hybrid model. But we still remember in high-frequency trading, how to reduce the trading cost is another important issue. Although we don't discuss this in here, it still should be considered in future works. At last, we can collocate efficient trading strategy to enhance the overall performance. These future direction help us improve the hybrid model configure become better.

Appendix A

Data tables

We list empirical data here for reference.

Table A.1: Risk-free rate(Average deposit interest rates of four major banks(Bank of Taiwan, Taiwan Cooperative Bank (TCB), First Commercial Bank, Hua Nan Bank.))

Unit:(%)	One Month	One Quarter	Two Quarters	Three Quraters	One Year	Two years	Three Years
2007/01	1.74	1.80	1.95	2.08	2.21	2.28	2.30
2007/02	1.74	1.80	1.95	2.08	2.21	2.28	2.30
2007/03	1.74	1.80	1.95	2.08	2.21	2.28	2.30
2007/04	1.77	1.83	1.98	2.11	2.24	2.29	2.30
2007/05	1.77	1.83	1.98	2.11	2.24	2.29	2.30
2007/06	1.97	2.03	2.18	2.31	2.44	2.50	2.51
2007/07	1.97	2.03	2.18	2.31	2.44	2.50	2.51
2007/08	1.97	2.03	2.18	2.31	2.44	2.50	2.51
2007/09	2.03	2.10	2.26	2.39	2.52	2.57	2.59
2007/10	2.03	2.10	2.26	2.39	2.52	2.57	2.59
2007/11	2.03	2.10	2.26	2.39	2.52	2.57	2.59
2007/12	2.09	2.17	2.34	2.47	2.60	2.65	2.66
2008/01	2.09	2.17	2.34	2.47	2.60	2.65	2.66
2008/02	2.09	2.17	2.34	2.47	2.60	2.65	2.66
2008/03	2.09	2.17	2.34	2.47	2.60	2.65	2.66
2008/04	2.14	2.22	2.39	2.52	2.64	2.69	2.71
2008/05	2.14	2.22	2.39	2.52	2.64	2.69	2.71
2008/06	2.15	2.24	2.40	2.54	2.66	2.71	2.73
2008/07	2.20	2.28	2.45	2.58	2.70	2.75	2.77
2008/08	2.20	2.28	2.45	2.58	2.70	2.75	2.77
2008/09	2.17	2.26	2.42	2.56	2.68	2.73	2.75
2008/10	1.98	2.07	2.24	2.38	2.49	2.54	2.55
2008/11	1.68	1.77	1.94	2.07	2.18	2.23	2.25
2008/12	1.01	1.07	1.24	1.37	1.48	1.53	1.55
2009/01	0.56	0.62	0.79	0.92	1.03	1.08	1.10
2009/02	0.50	0.56	0.71	0.85	0.93	0.98	1.00
2009/03	0.50	0.56	0.71	0.85	0.93	0.98	1.00
2009/04	0.50	0.56	0.71	0.85	0.93	0.98	1.00
2009/05	0.50	0.56	0.71	0.85	0.93	0.98	1.00
2009/06	0.50	0.56	0.71	0.85	0.93	0.98	1.00
2009/07	0.50	0.56	0.71	0.85	0.93	0.98	1.00
2009/08	0.50	0.56	0.71	0.85	0.93	0.98	1.00
2009/09	0.50	0.56	0.71	0.85	0.93	0.98	1.00
2009/10	0.53	0.58	0.74	0.87	0.95	1.01	1.03
2009/11	0.53	0.58	0.74	0.87	0.95	1.01	1.03
2009/12	0.53	0.58	0.74	0.87	0.95	1.01	1.03
2010/01	0.53	0.58	0.74	0.87	0.95	1.01	1.03
2010/02	0.53	0.58	0.74	0.87	0.95	1.01	1.03
2010/03	0.53	0.58	0.74	0.87	0.95	1.01	1.03
2010/04	0.53	0.58	0.74	0.87	0.95	1.01	1.03
2010/05	0.53	0.58	0.74	0.87	0.95	1.01	1.03
2010/06	0.62	0.67	0.82	0.94	1.04	1.10	1.11
2010/07	0.62	0.67	0.82	0.94	1.04	1.10	1.11
2010/08	0.62	0.67	0.82	0.94	1.04	1.10	1.11
2010/09	0.62	0.67	0.82	0.94	1.04	1.10	1.11
2010/10	0.69	0.74	0.89	1.01	1.13	1.16	1.18
2010/11	0.69	0.74	0.89	1.01	1.13	1.16	1.18
2010/12	0.69	0.74	0.89	1.01	1.13	1.16	1.18
2011/01	0.75	0.79	0.95	1.07	1.18	1.21	1.23
2011/02	0.75	0.79	0.95	1.07	1.18	1.21	1.23
2011/03	0.75	0.79	0.95	1.07	1.18	1.21	1.23
2011/04	0.82	0.87	1.03	1.15	1.27	1.30	1.31
2011/05	0.82	0.87	1.03	1.15	1.27	1.30	1.31
2011/06	0.82	0.87	1.03	1.15	1.27	1.30	1.31
2011/07	0.88	0.94	1.11	1.22	1.35	1.38	1.39
2011/08	0.88	0.94	1.11	1.22	1.35	1.38	1.39
2011/09	0.88	0.94	1.11	1.22	1.35	1.38	1.39
2011/10	0.88	0.94	1.11	1.22	1.35	1.38	1.39
2011/11	0.88	0.94	1.11	1.22	1.35	1.38	1.39
2011/12	0.88	0.94	1.11	1.22	1.35	1.38	1.39

Table A.2: Overall Original Monthly Performance

Month	Standard Model	Model Mk-1	Model Mk-2	Model Mk-3	Model Mk-4	Model Mk-5	Single BPNN
7-Jan	760	678	862	842	940	942	488
7-Feb	556	456	544	498	460	538	112
7-Mar	243	189	157	45	265	161	455
7-Apr	322	456	362	448	586	382	378
7-May	151	167	205	161	11	199	407
7-Jun	441	721	607	713	453	557	1117
7-Jul	1275	1473	1487	1289	1267	1321	1057
7-Aug	2254	1550	1636	1394	1864	1608	380
7-Sep	627	923	959	883	857	1005	1211
7-Oct	1034	1224	1278	1332	1174	1144	1132
7-Nov	1368	1744	1720	1948	2184	1830	578
7-Dec	732	716	846	776	666	908	848
8-Jan	1391	691	757	689	849	1083	2083
8-Feb	211	281	237	87	277	219	917
8-Mar	1389	1947	2267	2119	1855	2061	1543
8-Apr	361	947	1033	665	881	859	-115
8-May	1213	1113	1037	961	1121	1149	483
8-Jun	452	96	290	468	222	472	224
8-Jul	1111	1673	1985	925	1805	1357	165
8-Aug	918	438	536	430	422	388	200
8-Sep	482	408	524	308	492	788	558
8-Oct	467	435	589	433	607	519	397
8-Nov	765	687	489	429	711	793	577
8-Dec	315	865	903	1023	869	931	-327
9-Jan	563	415	417	187	157	361	117
9-Feb	659	615	615	413	409	639	211
9-Mar	207	413	277	133	103	375	359
9-Apr	424	712	1022	1018	818	1044	-242
9-May	114	316	486	334	312	512	-80
9-Jun	870	544	836	392	208	838	168
9-Jul	-54	318	202	40	-8	144	64
9-Aug	449	169	163	-45	155	77	229
9-Sep	59	397	493	59	215	535	213
9-Oct	413	429	305	163	217	251	-249
9-Nov	499	419	503	373	435	559	379
9-Dec	280	-44	24	82	68	-4	162
10-Jan	18	542	572	542	482	542	-200
10-Feb	358	332	428	202	328	454	-14
10-Mar	-112	202	150	264	268	112	134
10-Apr	-156	-118	2	170	78	-32	238
10-May	134	758	474	686	876	402	434
10-Jun	11	-11	209	-121	-147	185	57
10-Jul	223	3	39	65	33	-21	101
10-Aug	63	-253	-133	67	15	-91	123
10-Sep	140	224	144	240	276	218	14
10-Oct	237	395	401	275	355	389	167
10-Nov	23	141	51	67	37	167	165
10-Dec	39	177	111	171	193	129	249
11-Jan	-220	128	98	186	162	98	374
11-Feb	195	389	737	587	483	279	307
11-Mar	434	546	634	776	714	514	522
11-Apr	235	121	135	133	177	115	-225
11-May	24	84	86	286	124	-32	346
11-Jun	-39	107	115	213	79	279	43
11-Jul	24	162	114	28	138	234	-144
11-Aug	1297	761	867	599	819	683	-155
11-Sep	-538	-36	-60	-202	-38	384	-184
11-Oct	278	350	186	398	342	318	246
11-Nov	187	317	467	509	391	325	131
11-Dec	120	22	84	64	18	-80	-508

Bibliography

[1] B. B. Mandelbrot, "A multifractalwalkdown," *Scientific American*, p. 71, 1999.

[2] F. Black and M. Scholes, "The pricing of options and corporate liabilities," *The journal of political economy*, pp. 637–654, 1973.

[3] J. L. Treynor, "How to rate management of investment funds," *Harvard business review*, vol. 43, no. 1, pp. 63–75, 1965.

[4] R. N. Mantegna, H. E. Stanley, *et al.*, *An introduction to econophysics: correlations and complexity in finance*, vol. 9. Cambridge university press Cambridge, 2000.

[5] A. J. Frost and R. R. Prechter, *Elliott wave principle: key to market behavior*. Elliott Wave International, 2005.

[6] R. N. Elliott, "The wave principle," *New York*, 1938.

[7] H. Kleinert, *Path integrals in quantum mechanics, statistics, polymer physics, and financial markets*. World Scientific, 2009.

[8] M. Chlistalla, B. Speyer, S. Kaiser, and T. Mayer, "High-frequency trading," *Deutsche Bank Research*, pp. 1–19, 2011.

[9] B. Biais and P. Woolley, "High frequency trading," *Manuscript, Toulouse University, IDEI*, 2011.

[10] S.-S. Chern and J. Simons, "Characteristic forms and geometric invariants," *Annals of Mathematics*, pp. 48–69, 1974.

[11] I. Aldridge, *High-frequency trading: a practical guide to algorithmic strategies and trading systems*. John Wiley & Sons, 2013.

[12] C. A. Goodhart and M. O'Hara, "High frequency data in financial markets: Issues and applications," *Journal of Empirical Finance*, vol. 4, no. 2, pp. 73–114, 1997.

[13] Y. S. Abu-Mostafa and A. F. Atiya, "Introduction to financial forecasting," *Applied Intelligence*, vol. 6, no. 3, pp. 205–213, 1996.

[14] 羅華強 and 通信工程, 類神經網路*: MATLAB* 的應用. 高立, 2011.

[15] 蘇木春, 張孝德, *et al.*, 機器學習*:* 類神經網路*,* 模糊系統以及基因演算法則. 臺北市: 全華科技圖書股份有限公司, 1997.

[16] S. A. Hamid and Z. Iqbal, "Using neural networks for forecasting volatility of s&p 500 index futures prices," *Journal of Business Research*, vol. 57, no. 10, pp. 1116–1125, 2004.

[17] N. E. Huang, Z. Shen, S. R. Long, M. C. Wu, H. H. Shih, Q. Zheng, N.-C. Yen, C. C. Tung, and H. H. Liu, "The empirical mode decomposition and the hilbert spectrum for nonlinear and non-stationary time series analysis," *Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences*, vol. 454, no. 1971, pp. 903–995, 1998.

[18] Z. Wu and N. E. Huang, "A study of the characteristics of white noise using the empirical mode decomposition method," *Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences*, vol. 460, no. 2046, pp. 1597–1611, 2004.

[19] I. Kļevecka and J. Lelis, "Pre-processing of input data of neural networks: the case of forecasting telecommunication network traffic," *publication. editionName*, vol. 104, pp. 168–178, 2008.

[20] Y. C. Tsai, “Forecasting electricity consumption as well as gold price by using an eemd-based back-propagation neural network learning paradigm,” Master’s thesis, National Chengchi University, Taiwan, 2011.

[21] Y.-H. Wang, C.-H. Yeh, H.-W. V. Young, K. Hu, and M.-T. Lo, “On the computational complexity of the empirical mode decomposition algorithm,” *Physica A: Statistical Mechanics and its Applications*, vol. 400, pp. 159–167, 2014.

[22] H. Demuth and M. Beale, “Neural network toolbox for use with matlab,” 1993.

[23] K. Hornik, M. Stinchcombe, and H. White, “Multilayer feedforward networks are universal approximators,” *Neural networks*, vol. 2, no. 5, pp. 359–366, 1989.

[24] M. T. Hagan, H. B. Demuth, M. H. Beale, *et al.*, *Neural network design*, vol. 1. Pws Boston, 1996.

[25] E. M. Azoff, *Neural network time series forecasting of financial markets*. John Wiley & Sons, Inc., 1994.

[26] 王奕鈞, “神經網路應用於地籍坐標轉換之研究,” 2005.

[27] 陈明, *MATLAB* 神经网络原理与实例精解. 清华大学出版社, 2013.

[28] P. Whitle, *Hypothesis testing in time series analysis*, vol. 4. Almqvist & Wiksells, 1951.

[29] G. E. Box, G. M. Jenkins, and G. C. Reinsel, *Time series analysis: forecasting and control*. John Wiley & Sons, 2013.

[30] J. D. Hamilton, *Time series analysis*, vol. 2. Princeton university press Princeton, 1994.

[31] G. E. Box and D. A. Pierce, “Distribution of residual autocorrelations in autoregressive-integrated moving average time series models,” *Journal of the American statistical Association*, vol. 65, no. 332, pp. 1509–1526, 1970.

[32] R. S. Tsay, *Analysis of financial time series*, vol. 543. John Wiley & Sons, 2005.

[33] A. Pole, *Statistical arbitrage: algorithmic trading insights and techniques*, vol. 411. John Wiley & Sons, 2008.

[34] V. Menon and A. E. Trefethen, "Multimatlab integrating matlab with high performance parallel computing," in *Supercomputing, ACM/IEEE 1997 Conference*, pp. 30–30, IEEE, 1997.

[35] B. Barney *et al.*, "Introduction to parallel computing," *Lawrence Livermore National Laboratory*, vol. 6, no. 13, p. 10, 2010.

[36] T. Hendershott and R. Riordan, "Algorithmic trading and the market for liquidity," *Journal of Financial and Quantitative Analysis*, vol. 48, no. 04, pp. 1001–1024, 2013.

[37] M. Schaden, "Quantum finance," *Physica A: Statistical Mechanics and its Applications*, vol. 316, no. 1, pp. 511–538, 2002.

[38] K. Lee and G. Jo, "Expert system for predicting stock market timing using a candlestick chart," *Expert Systems with Applications*, vol. 16, no. 4, pp. 357–364, 1999.

[39] J. H. Fock, C. Klein, and B. Zwergel, "Performance of candlestick analysis on intraday futures data," *The Journal of Derivatives*, vol. 13, no. 1, pp. 28–40, 2005.

[40] S. Nison, *Japanese candlestick charting techniques: a contemporary guide to the ancient investment techniques of the Far East*. Penguin, 2001.

[41] DayTradingCoach, "Candlestick chart course." http://www.daytradingcoach.com/daytrading-candlestick-course.htm.

[42] T. Chordia, R. Roll, and A. Subrahmanyam, "Liquidity and market efficiency," *Journal of Financial Economics*, vol. 87, no. 2, pp. 249–268, 2008.

[43] J. Brogaard, "High frequency trading and its impact on market quality," *Northwestern University Kellogg School of Management Working Paper*, p. 66, 2010.

[44] W. Hoeffding, “A non-parametric test of independence,” *The Annals of Mathematical Statistics*, pp. 546–557, 1948.

[45] L. Bachelier, “Théorie de la spéculation,” in *Annales scientifiques de l’École Normale Supérieure*, vol. 17, pp. 21–86, Société mathématique de France, 1900.

[46] J. M. Karpoff, “The relation between price changes and trading volume: A survey,” *Journal of Financial and quantitative Analysis*, vol. 22, no. 01, pp. 109–126, 1987.

[47] G. E. Tauchen and M. Pitts, “The price variability-volume relationship on speculative markets,” *Econometrica: Journal of the Econometric Society*, pp. 485–505, 1983.

[48] S.-Y. Chen, C.-C. Lin, P.-H. Chou, and D.-Y. Hwang, “A comparison of hedge effectiveness and price discovery between taifex taiex index futures and sgx msci taiwan index futures,” *Review of Pacific Basin Financial Markets and Policies*, vol. 5, no. 02, pp. 277–300, 2002.

[49] MSCI, “Msci taiwan.” http://www.msci.com/products/indexes/licensing/msci_taiwan/.

[50] C. Wang and S. Sern Low, “Hedging with foreign currency denominated stock index futures: evidence from the msci taiwan index futures market,” *Journal of Multinational Financial Management*, vol. 13, no. 1, pp. 1–17, 2003.

[51] H.-P. Spahn, *From Gold to Euro: On monetary theory and the history of currency systems*. Springer, 2001.

[52] G. Grudnitski and L. Osburn, “Forecasting s&p and gold futures prices: an application of neural networks,” *Journal of Futures Markets*, vol. 13, no. 6, pp. 631–643, 1993.

[53] T. G. Andersen and T. Bollerslev, “Intraday periodicity and volatility persistence in financial markets,” *Journal of empirical finance*, vol. 4, no. 2, pp. 115–158, 1997.

[54] I. S. Abdalla and V. Murinde, "Exchange rate and stock price interactions in emerging financial markets: evidence on india, korea, pakistan and the philippines," *Applied financial economics*, vol. 7, no. 1, pp. 25–35, 1997.

[55] C. K. Ma and G. W. Kao, "On exchange rate changes and stock price reactions," *Journal of Business Finance & Accounting*, vol. 17, no. 3, pp. 441–449, 1990.

[56] A. Lendasse, E. de Bodt, V. Wertz, M. Verleysen, *et al.*, "Non-linear financial time series forecasting-application to the bel 20 stock market index," *European Journal of Economic and Social Systems*, vol. 14, no. 1, pp. 81–92, 2000.

[57] E. T. Li, "Taiex option trading by using eemd-based neural network learning paradigm," Master's thesis, National Chengchi University, Taiwan, 2011.

[58] Y. H. Chen, "A study of trading strategies of taiex futures by using eemd-based neural network learning paradigms," Master's thesis, National Chengchi University, Taiwan, 2013.

[59] KaplanSchweser, ed., *SCHWESERNOTES 2014 CFA LEVEL I BOOK 1: ETHICAL AND PROFESSIONAL STANDARDS AND QUANTITATIVE METHODS*. Kaplan,Inc., 2013.

[60] D. Kirk, "Nvidia cuda software and gpu parallel computing architecture," in *ISMM*, vol. 7, pp. 103–104, 2007.

[61] M. Fatica and W.-K. Jeong, "Accelerating matlab with cuda," in *The High Performance Embedded Computing Workshop*, 2007.

[62] D. Agrawal, S. Das, and A. El Abbadi, "Big data and cloud computing: current state and future opportunities," in *Proceedings of the 14th International Conference on Extending Database Technology*, pp. 530–533, ACM, 2011.