# 國立交通大學

## 資訊管理研究所

## 博 士 論 文

以計算智慧為基礎之新的避險比例決定方法

A Novel Approach for Hedge Ratio Decision Based on Computational Intelligence

研 究 生：許育嘉

指導教授：陳安斌 博士

中華民國 九十九 年 六 月

# 以 計 算 智 慧 為 基 礎 之 新 的 避 險 比 例 決 定 方 法
# A Novel Approach for Hedge Ratio Decision Based on Computational Intelligence


研 究 生： 許育嘉　　　　Student: Yu-Chia Hsu

指導教授： 陳安斌　　　　Advisor: An-Pin Chen

國 立 交 通 大 學

資 訊 管 理 研 究 所


博 士 論 文

A Dissertation

Submitted to Institute of Information Management

College of Management

National Chiao Tung University

in Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy in Information Management

June 2010

Hsinchu, Taiwan, the Republic of China

中華民國九十九年六月

**以計算智慧為基礎之新的避險比例決定方法**

學生：許育嘉　　　　　　　　　　　　　　指導教授：陳安斌 博士

國立交通大學資訊管理研究所博士班

# 摘　要

本研究提出了一個整合計算智慧與統計方法學的最適避險比例決定方法，用來改善不同避險區間下最小變異避險比例之預測準確度。透過衡量金融市場現貨及期貨商品報酬時間序列之變異數、共變數、價差及其他們的一階、二階變量，市場波動的動態行為可以被擷取出來，之後以增長階層式自我組織圖進行階層式的分群。經過分群，這些位在相同集群裡具有相似行為的時間序列資料，經過給予不同的權重進行重新取樣後，會被蒐集起來用來取代原先估算最適避險比例的資料樣本。我們將這個方法運用在台灣加權股價指數、標準普爾 500 指數、金融時報 100 指數、以及日經 255 指數之避險實證研究上，對於避險區間之長短與避險效果的關係進行研究。實驗結果顯示，這個方法所估算之避險比例，在中、長期避險區間下可以顯著地得到優於傳統最小平方法模型及天真避險模型之表現，決定出各種避險期間下之最適避險比例。

關鍵字: 最適避險比例，財務時間序列，成長階層式自組織映射圖，集群分析

**A Novel Approach for Hedge Ratio Decision Based on Computational Intelligence**

Student: Yu-Chia Hsu Advisors: Dr. An-Pin Chen

Institute of Information Management
National Chiao Tung University

## ABSTRACT

In this study, a novel procedure combining computational intelligence and statistical methodologies is proposed to improve the accuracy of minimum -variance optimal hedge ratio (OHR) estimation over various hedging horizons. The time series of financial asset returns are clustered hierarchically using a growing hierarchical self-organizing map (GHSOM) based on the dynamic behaviors of market fluctuation extracted by measurement of variances, covariance, price spread, and their first and second differences. Instead of using original observations, observations with similar patterns in the same cluster and weighted by a resample process are collected to estimate the OHR. Four stock market indexes and related futures contracts, including Taiwan Weighted Index (TWI), Standard & Poor's 500 Index (S&P 500), Financial Times Stock Exchange 100 Index (FTSE 100), and NIKKEI 255 Index, are adopted in empirical experiments to investigate the correlation between hedging horizon and performance. Results of the experiments demonstrate that the proposed approach can significantly improve OHR decisions for mid-term and long-term hedging compared with traditional ordinary least squares and naïve models.

*Keywords: optimal hedge ratio; financial time series; GHSOM; cluster analysis.*

# 誌　謝

這份博士論文能夠圓滿完成，順利取得博士學位，除了指導教授自謙所說的要歸功於「祖上有積德」之外，其實最要感謝的人之一，就是我的恩師—陳安斌教授。十年前，從碩士班就開始接受陳安斌教授殷切的指導、諄諄的教誨，不僅僅是獨立的研究能力，從老師身上，也學習到許多處事的態度、人生的哲理，這段期間也非常感謝老師給了我許多機會而得以成長茁壯，在接受磨練與挑戰同時，逐漸培養出專業的能力以及開闊的視野，這些體驗與見解是一輩子受益不盡的，在此致上心中無盡的感激。

此外，也要感謝論文口試委員：翁崇雄教授、蔡垂雄教授、王淑芬教授、劉敦仁教授、李永銘教授，在百忙之中撥冗指導並提供許多寶貴意見，使得論文得以完備。而歷年來 APC Lab 一起共同研究、參與討論的學長姊、學弟妹，也非常感激你們，有了你們的協助與貢獻，這一段路才能這麼平順。而資策會的欲盛組長、馬偕的秀玲主任，也謝謝你們過去在工作上所給予的協助，才能使我在有限的時間下完成這不可能的任務。

最後，最要感激的是慈愛的母親，以及僅差五個月遺憾未能來的及看到這份成就的父親，感謝你們的養育與教導，才能有今日的我。也要感謝我的老婆，有了妳愛的支持以及背後辛苦付出照料這個家庭，才能讓我安心地離家工作、讀書，完成學位。在此僅將這份榮耀，獻給我摯愛的家人，我愛你們！

許育嘉　謹誌

2010 Jun.

# Table of Contents

# List of Tables

# List of Figures

# Chapter 1 Introduction

## 1.1 Background of the Study

With the emergence of financial derivatives markets in the past two decades, hedging has been of interest to both academicians and practitioners. The goal of hedging is to minimize exposure to unwanted risk. This is carried out by establishing the position of a derivative instrument to offset exposure to price fluctuations opposite to that of underlying assets, such as using futures to hedge a portfolio of risky assets. The primary objective is to estimate the size of the short position that must be held in the futures market (i.e., a proportion of the long position held in the spot market) with minimal risk and specific risk aversion of the hedged portfolio. Aside from hedge ratio, hedge horizon should also be considered simultaneously because investors, such as regulators and speculative investors, as well as individuals and institutions participating in the stock and futures markets have different hedging horizon and decision making over different time scales. Therefore, ignoring the dependence of the optimal hedge ratio on hedging horizon could lead to investors making inadequate decisions (Geppert, 1995; Lien and Shrestha, 2007), suggesting problems on the optimal hedge ratio (OHR) decision.

Many methods have been used to decide the OHR. Most studies adopt the mean-variance framework, which measures the risk of the hedged portfolio by standard deviation, and which assumes that OHR simply minimizes the variance of hedged portfolios. Many applications of optimal hedging use the criterion of minimum variance to estimate OHR, such as by regressing the spot market return on the futures market return using ordinary least squares (OLS) (Ederington, 1979; Hill and Schneeweis, 1982; Sener, 1998). However, OHRs estimated via the conventional approach is constant over time. The classical time-invariant OHR appears inappropriate with the time-varying nature of many financial markets. Improvements were made to capture time-varying features, such as by adopting dynamic hedging strategies based on the bivariate generalized autoregressive conditional heteroskedasticity (GARCH) framework (Kroner and Sultan, 1993; Lien and Luo 1994; Moschini and Myers, 2002; Choudhry, 2003; Wang and Low, 2003) or the stochastic volatility (SV) model (Anderson and Sorensen 1996; Lien and Wilson 2001). Although these studies are successful in capturing time-varying features, many give little attention to OHR decisions over different time scales.

## 1.2 Statement of the Problem

The models presented by the authors have several limitations in estimating the multiscale hedge ratio. All these approaches to estimating the abovementioned OHR are based on sample variance and covariance estimators of returns without considering the underlying distribution of data. The conventional OLS approach ignores the conditional distribution of most financial asset returns, which tends to vary over time. To obtain recent information, most research adopt a rolling window scheme to estimate the variance and covariance of spot and futures returns. However, rolling window estimators use an equally weighted moving average of past squared returns and their cross products. Observations have equal weight in the variance-covariance matrix estimator of the arbitrarily defined estimation period, but they have zero weight beyond the estimation period. GARCH class models are successful in capturing time-varying features for estimating conditional variance-covariance matrices, but they place too much weight on extreme observations (Nelson and Foster, 1996) when the distribution of data is leptokurtic and fat-tailed.

Furthermore, disregarding the dependence of OHR on the hedging horizon is problematic in these conventional approaches for estimation. Only a few studies consider different hedging horizons for hedge ratio estimation, including Howard and D’Antonio (1991), Lien and Luo

(1993, 1994), Geppert (1995), Lien and Wilson (2001), Chen, Lee, and Shrestha (2004), In and Kim (2006), and Lien and Shrestha (2007). However, these models have three problems in incorporating the hedging horizon in OHR estimation. First, the long-horizon OHR estimator based on a handful independent observations generated from long-horizon return series is unreliable (Geppert, 1995). This is because the frequency of data must match the hedging horizon (e.g., weekly or monthly data must be used to estimate the hedge ratio where the hedging horizon is one week or one month, respectively). Low data frequency would result in a substantial reduction in sample size (Lien and Shrestha, 2007). Second, the assumption for the error term of the GARCH/SV model would lead to inaccurate results when estimating the multiperiod hedge ratio (Lien and Wilson, 2001). Third, the assumption for the underlying data-generating process, such as a unit root process, is unsuitable when the assumed condition does not hold true, as evidenced in many commodities markets (Chen, Lee and Shrestha, 2004).

## 1.3 Purpose of the Study

The main purpose of this paper is to introduce a novel approach for deciding the OHR of different hedging horizons using computational intelligence technique. The new approach uses the growing hierarchical self-organizing map (GHSOM) of Rauber *et al.* (2002) to cluster time series data, which could decompose financial data into a hierarchical architecture consisting of several familiar clusters. Several applications of cluster analysis to economics and finance time series have been documented in recent literature, including identification of mutual funds styles by analyzing the time series of past returns (Pattarin *et al.*, 2004), discovery of companies that share similar behavior with the Dow Jones industrial average (DJIA) index (Basalto *et al.*, 2007), prediction of value at risk (Karandikar *et al.*, 2007), prediction of oil futures price (Zhu, 2008), and determination of optimal tracking portfolio (Focardi and Fabozzi, 2009).

In this paper, our work employs a different weight for observations in a rolling window OLS estimator of the variance-covariance matrix subsequent to the clustering time series using GHSOM. The weights of observations are determined by the measurement of similar patterns, which are correlated with the sample size of the cluster they belong to in the hierarchy architecture. The observations with different weights in clusters are then used to predict the conditional distribution of spot and futures returns for different hedging horizons in the future. When the conditional distribution of spot and futures returns is predictable, a more efficient estimate of the OHR can be obtained by conditioning on recent information (Harris and Shen, 2003).

## 1.4 Significance of the Study

The application of GHSOM to clustering time series to improve the conventional OHR estimator has at least three salient advantages. First, clustering time series does not suffer from the sample reduction problem when matching data frequency to hedging horizon. Second, observations within the cluster with similar patterns provide a way for examining the dependency of observations, which are generated from long-horizon return series and can provide predictable time patterns. The conditional distribution of returns in the next hedging horizon is predictable by aggregating these clustered observations, which are inspired by the well-established features of many asset returns that their conditional distribution is time-varying and tendency display volatility clustering. The final advantage is that the proposed computational intelligence (CI) approach is a non-parametric method, which can avoid too many inappropriate assumptions and restrictions found in conventional parametric models.

By doing this, OHR estimation for different horizons can be achieved. This proposed approach is also categorized as an analysis tool to investigate the relationship between hedge ratio and hedging horizon, which provides valuable information for reference in OHR decision making.

## 1.5 Theoretical Framework

Our study focuses on OHR estimation over various hedging horizons to support decision making. Conventional approaches on OHR estimation are based on parametric models that may encounter many issues and be restricted by many inappropriate assumptions. The proposed CI approach is a non-parametric model designed to overcome issues found in conventional models without the underlying assumptions. The theoretical framework is shown in Figure 1-1.

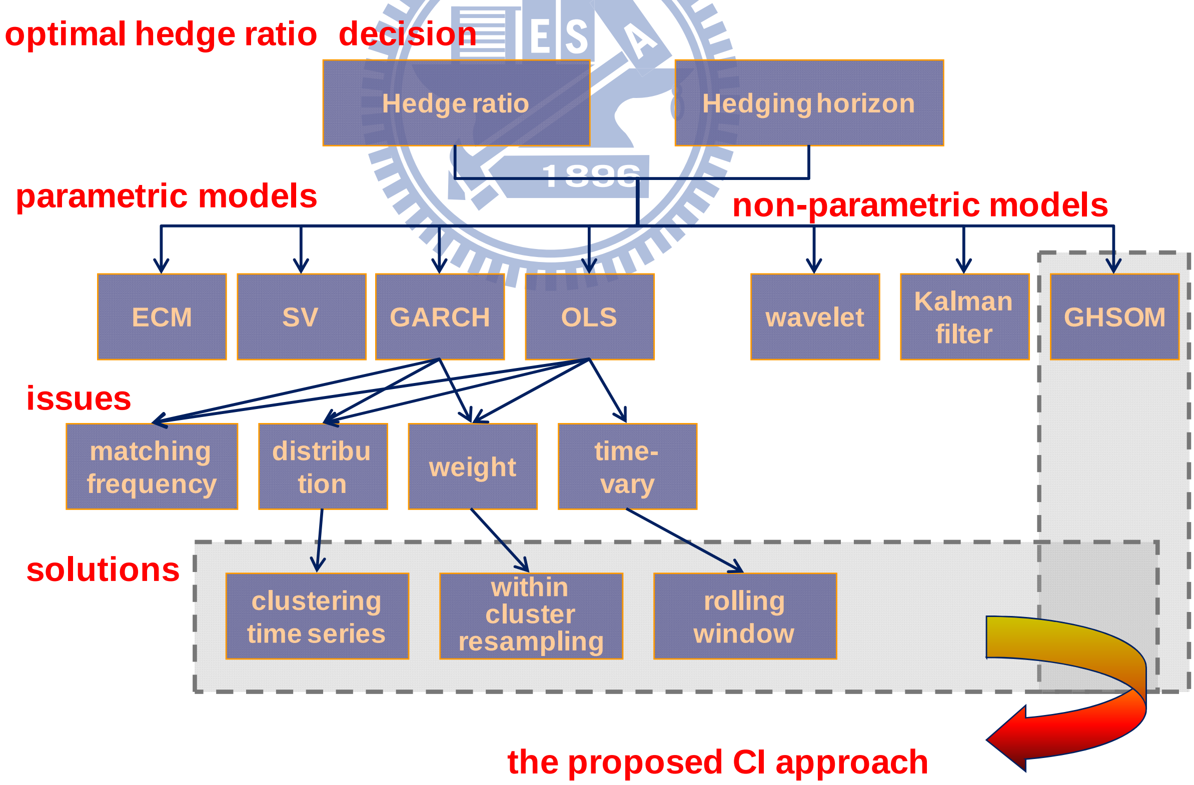



**Figure 1-1. Theoretical framework**

## 1.6 Organization of the Dissertation

This dissertation is presented in five chapters. Chapter 1 includes the background, statement of the problem, purpose of the study, significance of the study, and the theoretical framework. Chapter 2 presents a review of literature. Chapter 3 introduces the proposed model used for this study. Chapter 4 presents the experiments design and results analysis. The concluding remarks and recommendations for further work are provided in Chapter 5.

# Chapter 2 Literature Review

This chapter presents the rationale for conducting research on OHR decision making for different hedging horizons using computational intelligence technique. The following sections describe the theoretical background and examination of previous research presenting relevant issues on the minimum variance hedge ratio, clustering time series, and GHSOM.

## 2.1 Hedging Theorem and Hedge Ratio

The most widely used hedging strategy is to adopt the minimum-variance hedge ratio (Ederington, 1979; Myers and Thompson, 1989), which reduces the variance of portfolio to attain minimum risk for the hedger.

Traditionally, two approaches have been suggested to minimize portfolio risk. The first approach, naïve hedge, simply sets the hedge ratio equal to 1 over the whole hedging horizon. The correlation between spot and futures prices is assumed to be perfect, but it challenges the fact that the spot and futures prices are naturally stochastic and time variant. The second approach is the static OLS hedge, which accurately recognizes the correlation between futures and spot prices using the OLS coefficient of a regression of spot return on futures return (Ederington, 1979; Figlewski, 1984). However, it considers the joint distribution of spot and futures return as constant, and hence leads to suboptimal hedging decisions in periods of high basis volatility. The naïve and OLS approaches do not require any adjustment in hedge ratio once the decision is taken, which fails to consider current available information.

Recently, numerous works have focused on improving hedging performance using the dynamics in the joint distribution of returns and the time-varying nature of OHRs. Optimal hedge ratios are estimated using the family of GARCH models proposed by Engle (1982), Engle and Kroner (1995), and Bollerslev (1986, 1990). Various GARCH models are studied in literature to investigate hedge ratio and hedging performance, including bivariate GARCH model with diagonal vech parameterization for commodity futures contracts (Baillie and Myers, 1991), bivariate constant-correlation GARCH (CC-GARCH) model for foreign currency futures (Kroner and Sultan, 1993) and stock index futures (Park and Switzer, 1995), GARCH model with Baba-Engle-Kraft-Kroner (BEKK) parameterization for interest-rate futures (Gagnon and Lypny, 1995), augmented GARCH model for the freight futures market (Kavussanos and Nomikos, 2000), and orthogonal GARCH and CC-GARCH for the electricity futures market (Bystrom, 2003).

The GARCH family model can capture the dynamic behavior of a time series for OHR estimation, but these approaches have three drawbacks. First, GARCH models require the time series to be stationary, such that the price series of a financial asset are usually transformed to the return series by a differential. However, this eliminates much information and ignores co-integrated properties. Second, many research have attempted to improve the GARCH model by adding the error term or other variables to the model. These improvements can increase accuracy, but the model has become increasingly complicated and its variables are difficult to determine. Third, when these models work to deal with different hedging horizons, the original time series is required for sampling, which is based on data frequency. However, the information and property of the original time series may be eliminated after data sampling.

More recently, other approaches based on non-parametric models have been proposed to avoid undue restrictions. Alizadeh and Nomikos (2004) introduce the Markov regime switching (MRS) model to estimate the time-varying minimum-variance hedge ratio for stock index futures. Hatemi-Ja and Roca (2006) utilize the Kalman Filter approach for estimating time-varying hedge ratio. Gencay *et al.* (2003), Kim and In (2005), and In and Kim (2005, 2006) adopt wavelet analysis to study the relationship between hedge ratio and hedging horizon.

Azevedo *et al.* (2007) propose a particle swarm optimization (PSO) approach to support electricity producers for multiperiod optimal contract allocation. Although some of these models are more burdensome in computing, the accuracy of results have been improved and better hedging performance is obtained.

## 2.2 Cluster Analysis for Time Series

Since the proposal of the famous cluster analysis algorithms, k-means, 50 years ago, the cluster analysis has been widely used as a data analyzing tool in various domains. In the past two decades, time series clustering has shown effective results in providing useful information in various domains. In the financial field, clustering financial time series is a new approach to analyze the dynamic behavior of time series, and to forecast any future tendency of the time series for purposes of decision making. Many financial problems have been studied based on cluster analysis via computational intelligence approach instead of the conventional approach. Pattarin *et al.* (2004) propose a classification algorithm for mutual funds style analysis, which combines different statistical techniques and exploits information readily available at low cost. In their analysis, time series of past returns is used to retrieve synthetic and informative description of contexts characterized by high degrees of complexity, which is useful in identifying the styles of mutual funds. Gafnychuk *et al.* (2004) use the self-organizing methods to investigate the time series data of the Dow Jones index. Basalto *et al.* (2007) use a novel clustering procedure, which is applied to the financial time series of the Dow Jones industrial average (DJIA) index to find companies that share similar behaviors. The techniques proposed could extract relevant information from raw market data and yield meaningful hints as to the mutual time evolution of stocks. Karandikar *et al.* (2007) develop a volatility clustering model to forecast value at risk (VaR). The feasibility and benefits of the model are demonstrated in an electricity price return series. Zhu (2008) propose a new model based on cluster analysis for oil futures price forecasting. This model is demonstrated using the historical data from NYMEX market, and shows that the proposed model can effectively and stably improve the precision of oil futures price forecasting. Focardi and Fabozzi (2009) adopt a clustering-based methodology to determine optimal tracking portfolio to track indexes. Papanastassiou (2009) discuss classification and clustering of financial time series data based on a parametric GARCH (1,1) representation to assess their riskiness.

In spite of the prevalence of numerous clustering algorithms, including their success in a number of different application domains, clustering remains difficult. When applying the clustering analysis on time series, the method of data processing, feature extraction, similarity measurement, and topology of cluster construction should be determined. Features extracted from the time series are organized by past research into three groups (i.e., according to data used) (Liao, 2005): working directly with the data either in the time or frequency domain; working indirectly with features extracted from the raw data; and working indirectly with models built from raw data. Defining an appropriate similarity measure and objective function is difficult when choosing clustering algorithm. Jain (2010) emphasizes that “there is no best clustering algorithm” when comparing the results of different clustering algorithms. Furthermore, the clustering method can be classified into two categories depending on whether the data objects are grouped into a tree of clusters or not (i.e., hierarchical and non-hierarchical). There are generally two types of hierarchical clustering methods: agglomerative and divisive. Agglomerative methods start by placing each object in its own cluster and then merging clusters into larger and larger clusters, until all objects are in a single cluster or until certain termination conditions, such as the desired number of clusters, are satisfied. Divisive methods do the opposite. Determining the number of clusters automatically is one of the most difficult problems in data clustering. Most methods for automatically determining the number of clusters cast it into the problem of model selection.

Although clustering remains a difficult problem, in time series, it offers two benefits. First, clustering can avoid the improper assumption and restriction of data. Gershenfeld *et al.* (1999) propose a cluster-weighted model for time series analysis, which is a simple special case of the general theory of probabilistic networks but one that can handle most of the limitations of practical data sets without unduly constraining either data or user. They show that are nonlinear, non-stationary, non-Gaussian, and discontinuous signals can be described by expanding the probabilistic dependence of the future depending on past relationships of local models. Second, data objects with similar dynamic behavior in their evolution over time are pooled and can thus help in data modeling. Fruhwirth-Schnatter and Kaufmann (2008) propose to pool multiple time series into several groups using finite-mixture models. Within a panel of time series, only those that display "similar" dynamic properties are pooled to estimate the parameters of the generating process. They estimate the groups of time series simultaneously with group-specific model parameters using Bayesian Markov chain Monte Carlo simulation methods. They document the efficiency gains in estimation, and forecasting is realized relative to the overall pooling of the time series. D'Urso and Maharaj (2009) suggest that time series often display dynamic behavior in their evolution over time, which should be taken into account when attempting to cluster the time series. They proposed a fuzzy clustering approach based on autocorrelation functions to determine and represent the underlying structure in the given time series.

Based on literature, we apply the cluster analysis to find the dynamic behavior of financial time series based on computational intelligence approach. The method of data processing, feature extraction, and similarity measurement, as well as the topology of clusters constructed, are easy to determine.

## 2.3 GHSOM

Kohonen's self-organizing feature map (SOM) is an unsupervised two-layer network that organizes a topological map. The resulting map shows the natural relationships among patterns given in the network. SOM is suitable for clustering analysis and has been applied to time series pattern discovery (Fu *et al.*, 2001; Tsao and Chen, 2003) and time series forecasting (Senjyu, 2000; Simon *et al.* 2005; Afolabi and Olude, 2007). However, when applying the SOM for cluster analysis, the topology of the SOM describing the number of clusters needs to be determined in advance. Moreover, the topology of the SOM lacks the ability to represent hierarchical relations of the data.

The hierarchical relations of the data are treated as trends of different time scales in time series analysis. These concepts are utilized in decomposition analyses, such as Fourier analysis and wavelet analysis. Time series data are decomposed into many components that can easily show the detailed properties of long-term, mid-term, and short-term tendencies.

Proposed by Rauber *et al.* (2002), GHSOM has a hierarchical architecture of multiple layers. Each layer comprises several independent clusters representing the growing SOM. The breadth of each SOM and the depth of the hierarchy are adjusted according to the characteristics of the input data during the unsupervised training process. Each SOM undergoes training via an unsupervised and competitive learning algorithm, as proposed by Kohonen (1989). The training steps include competitive and weight adaptation processes.

In many research, GHSOM has been used to perform clustering, including presentation of a content-based and easy-to-use map hierarchy for legal documents in the securities and futures markets (Shih *et al.*, 2008), knowledge discovery from multilingual text documents

(Yang *et al.*, 2009), and pattern discovery of time series data from robot execution failures and gene expression data (Liu *et al.*, 2006).

In this study, GHSOM is used for the time series analysis to deal with variance and covariance data, which have not been studied in detail. Using the GHSOM algorithm, time series data with similar patterns are clustered together. If the similarity of data in the same cluster is below a certain threshold, data are clustered once again by breadth or depth, thus expanding the SOM clusters. The topology of the clusters is automatically determined by the data and related with the threshold setting for width and depth expansion. Finally, the hierarchical clustering results, which represent the data relationship, are obtained and used to modify the mean-variance model for OHR estimation.

Based on previous literature, as described in previous sections, we consider adopting the GHSOM for financial series clustering. The method utilizes an unsupervised hierarchical clustering algorithm with easily determined cluster numbers, and is less sensitive in model parameter selection. The dynamic behaviors of financial time series are modeled by cluster analysis. The model can also forecast data distribution, which can improve the accuracy of OHR estimation.

# Chapter 3 Research Methodology

## 3.1 Design of Computational Intelligence-based Model

The conventional approach to OHR estimation is simply to regress the spot and futures series. The basic operating steps are shown in Figure 3-1. The first step is to collect the market price of spots and futures as original data. Next, the original price series is sampled so it coincides with the hedging horizon and then transformed into a return series by differencing. Finally, these data are used to estimate variance and covariance using OLS to obtain the OHR.

In this study, two modifications of the conventional approach are proposed based on computational intelligence, as shown in the top part of Figure 3-1. First, the data sampling process in the conventional approach is omitted from the conventional approach because it causes reduction of sample size, which decreases the accuracy of OHR estimation. Second, the original composition of data for OHR estimation is modified by the selected data with a similar pattern, which is performed in two phases. Phase I is clustering time series, and Phase II is modifying the probability distribution. The philosophy of the proposed approach is that data with similar dynamic behaviors may appear in the future with higher probability than dissimilar ones. Therefore, the objective of Phase I is to identify higher probability data, which would occur in the next hedging horizon based on the whole data set, and ignore lower probability data. In Phase II, the data composed by the higher probability data are expected to be more approximate to the normal distribution than the original data, suggesting decreased inaccuracy caused by leptokurtic and fat-tailed distributions. Details of the proposed model are described in the following sections.

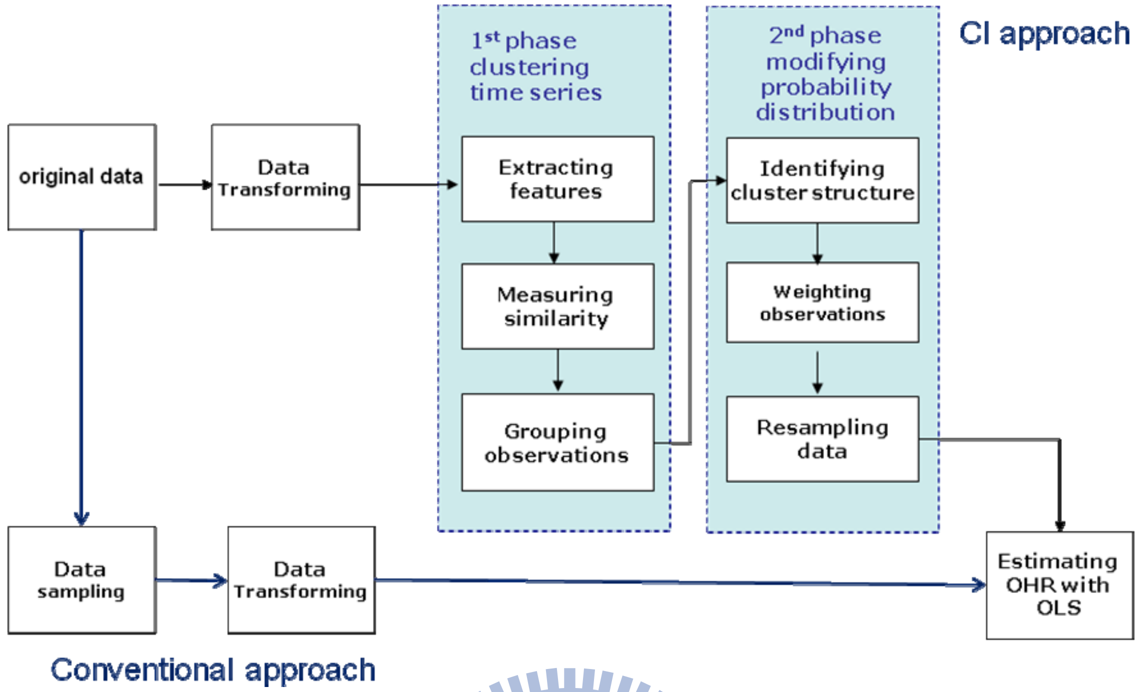



**Figure 3-1. The computational intelligence approach**

### 3.1.1 Clustering Time Series

Phase I of our proposed approach is clustering time series. Cluster analysis is an unsupervised learning method that can gather similar data in the same group by feature extraction and similarity measurement. Consequently, the features of time series and the algorithm for measuring similarity should be determined when applying the proposed approach. Many research works indicate that dynamic behaviors exist in financial time series, and these dynamic behaviors are helpful for time series forecasting (Fruhwirth-Schnatter and Kaufmann, 2008; D'Urso and Maharaj, 2009). The dynamic behaviors often refer to the velocity and the acceleration of a moving object, which are computed by the first-order and second-order differencing of the object position, respectively. In other words, dynamic behaviors can be defined as the speed of change and the acceleration of change. OHR estimation is relative to bivariate random variable analysis, which considers the joint probability distribution of spot and futures return series, and focuses on variance and covariance analysis. Dynamic behaviors—the interest of this study—are the variance of spot and futures return series, as well as the speed and acceleration of variance change. Furthermore, the time series of financial asset returns often exhibits the volatility clustering property. As noted by Mandelbrot (1963), "large changes tend to be followed by large changes, of either sign, and small changes tend to be followed by small changes." Therefore, the variance, speed of variance change, and acceleration of variance change are adopted as dynamic behaviors, which represent the features of time series and are extracted for clustering. The other features of the time series that are helpful for OHR estimation, such as price spread (Lien and Yang, 2006) and covariance of joint distribution, are also considered and tested in this study.

The cluster algorithm chosen in this study is GHSOM, but not k-mean, SOM, and others, because of three main reasons. The first relates to the benefits of the hierarchical structure. Cluster algorithms can be classified into two categories, non-hierarchical and hierarchical. For the non-hierarchical structure, the degree of similarity for each cluster is obtained by measuring the distance between cluster centers. When we want to collect a certain number of observations (i.e., the most similar), the distance of cluster centers for measuring the similarity is hardly determined. However, for the hierarchical structure, the degree of similarity for each cluster can be obtained in a more natural manner depending on the layer it belongs to in the hierarchical structure. Second, the result of GHSOM is stable regardless of cluster number. Many cluster algorithms should determine the number of clusters prior to their application. However, the best number of clusters for analysis is unknown, and the clustering results are often unstable with the cluster number. GHSOM can grow and expand the hierarchy of a cluster by parameter setting, which can determine the number of clusters automatically and is not sensitive to clustering results. The last reason is that the similarity measurement function of various cluster algorithms are not sensitive to the clustering results (Jain, 2010). Therefore, GHSOM is adopted in this study for clustering time series.

### 3.1.2 Modifying the Probability Distribution

Many properties of financial time series are time variant. We suggest that the probability distribution should be time variant, and estimated and updated by the latest time series data. We are also interested in the accuracy of forecasting. Observations with similar behavior may occur more frequently in the future and should be more emphasized than the dissimilar ones. However, when data are grouped by cluster analysis, the original data are divided into several groups, with each group only containing partial data. The number of similar data is far less than the original data. Reducing sample size causes inaccuracy when OLS for OHR estimation is employed (Lien and Shrestha, 2007). To overcome this problem, we propose to adopt with-in cluster resampling. With-in cluster resampling has been used for solving sample-reduced problems in the biometric field (Hoffman, 2001; Rieger and Weinberg, 2002). The observations of the cluster are replicated to expand the sample size. This idea is inspired by the stratified resampling scheme and bootstrap resampling method. The architecture of hierarchical cluster is very similar to the hierarchical stratified resampling scheme, in which the observations are divided into several groups according to their properties. Each group is weighted by the number of observation replications. Bootstrap method, which replicates the observation randomly to simulate the status in the future based on few observations, has been commonly used in finance and economics models. Therefore, in this study, observations in the cluster are randomly replicated until the sample size reaches the number size of the population.

## 3.2 Procedure of the Proposed Model

### 3.2.1 Data Transformation

In the process of data transformation, the original time series are transformed from prices to returns. These return series are then sequentially segmented into several windows with a fixed length in order to perform dynamic hedging strategy and out-of-sample testing.

The original data for OHR estimation gathered from the financial market are the daily closing (or settlement) prices of spot and futures. Generally, these price series are transformed into return series in consideration of payoffs in finance. The return series can be obtained by differencing the price series. We consider continuously compounded data and magnify the scale by multiplying by 100 to avoid a small scale. The return series is expressed as the price change:

$$\Delta S_t, \Delta F_t = \ln(P_t / P_{t-1}) \times 100 \qquad (3\text{-}1)$$

where $\Delta S$ and $\Delta F$ are price changes of spot and futures, respectively; $P$ is the price series; and $t$ refers to the time at present.

These return series are then divided into two parts, in-sample estimating period and out-of-sampling testing period. The hedge portfolio in this study is adjusted every hedging horizon according to the latest estimated OHR until the out-of-sample testing period is due. A rolling window scheme is applied to achieve the dynamic hedging strategy. The rolling windows scheme estimates the OHR at time $t$ according to the conditioning on the time $t$-1 data set, which is exhibited in Figure 3-1. Herein, $e$ denotes the in-sample estimating period while $h$ is the hedging horizon. The length of the rolling window is $e$+$h$. OHR is estimated based on the observations in the in-sample estimating period, from $t$-$e$ to $t$, then used for hedging from $t$ to $t$+$h$. Next, the window is rolled one hedging horizon ahead in order to reestimate the OHR based on the observations from $t+h-e$ to $t+h$. Then, we use the new OHR for the next hedging horizon, from $t+h$ to $t+2h$. OHR is reestimated every $h$ day, and then used to adjust the hedging portfolio until the out-of-sample testing period is due.

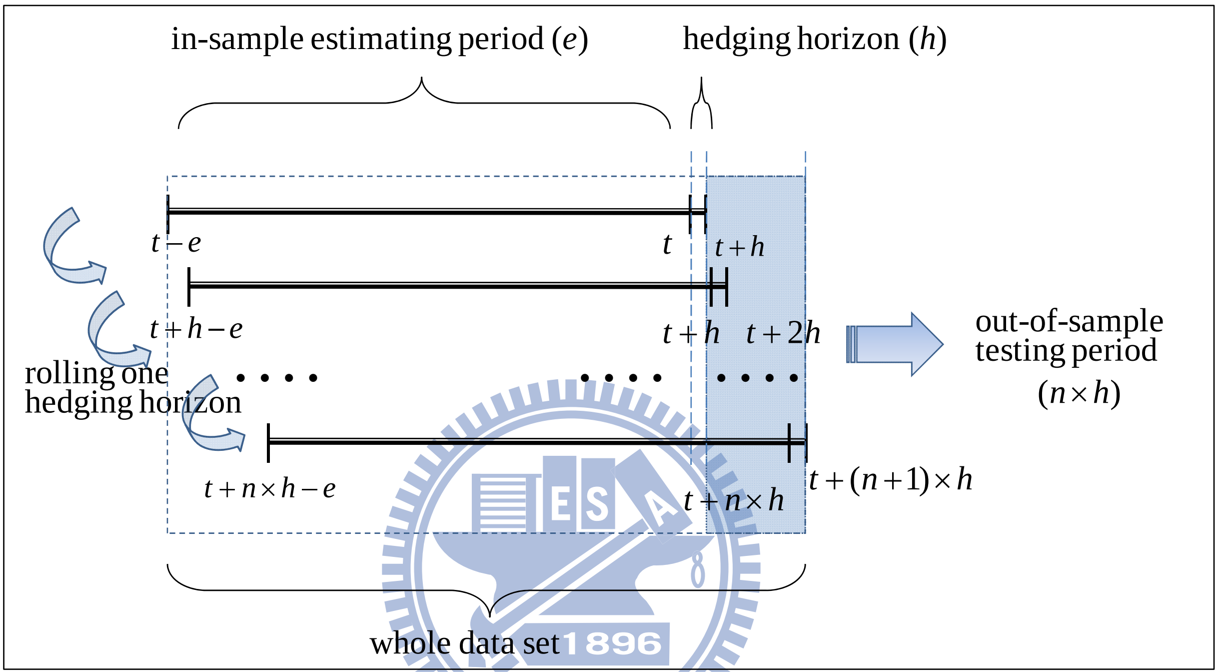



**Figure 3-2. The rolling windows scheme**

### 3.2.2 Feature Extraction for Various Horizons

In this study, variance, covariance, price spread, and their first and second differencing are adopted as the features of time series. These features are calculated using the data in the most recent hedging horizons just before the present; it is denoted by $h$. These features are calculated as follows:

$$Var(\Delta S_t) = Var[\Delta S_{t-h}, ..., \Delta S_t] \tag{3-2}$$

$$Var(\Delta F_t) = Var[\Delta F_{t-h}, ..., \Delta F_t] \tag{3-3}$$

$$Cov(\Delta S_t, \Delta F_t) = Cov\begin{bmatrix} \Delta S_{t-h},\ldots,\Delta S_t \\ \Delta F_{t-h},\ldots,\Delta F_t \end{bmatrix} \tag{3-4}$$

$$Spread(S_t, F_t) = F_t - S_t \tag{3-5}$$

The first and second order difference of these features are shown as

$$X_t' = \frac{X_t - X_{t-1}}{X_{t-1}} \tag{3-6}$$

$$X_t'' = \frac{X_t' - X_{t-1}'}{X_{t-1}'} \tag{3-7}$$

where $X$ represents the functions of $Var$, $Cov$, and $Sperad$. These extracted features from a period of data can represent the dynamic behavior of time series in the recent hedging horizon. Twelve values are extracted to describe an observation and used as the input variables of GHSOM (Table 3-1).

**Table 3-1. Features of the observation**

| Input Variables | Notations |
|---|---|
| **<u>Variance</u>** | |
| Variance of spot return series | $Var(\Delta S)$ |
| First order differential of $Var(\Delta S)$ | $Var'(\Delta S)$ |
| Second order differential of $Var(\Delta S)$ | $Var''(\Delta S)$ |
| Variance of futures return series | $Var(\Delta F)$ |
| First order differential of $Var(\Delta F)$ | $Var'(\Delta F)$ |
| Second order differential of $Var(\Delta F)$ | $Var''(\Delta F)$ |
| **<u>Covariance</u>** | |
| Covariance of spot and futures return series | $Cov(\Delta S, \Delta F)$ |
| First order differential of $Cov(\Delta S, \Delta F)$ | $Cov'(\Delta S, \Delta F)$ |
| Second order differential of $Cov(\Delta S, \Delta F)$ | $Cov''(\Delta S, \Delta F)$ |
| **<u>Price spread</u>** | |
| Spread of spot and futures price series | $Spread(S, F)$ |
| First order differential of $Spread(S, F)$ | $Spread'(S, F)$ |
| Second order differential of $Spread(S, F)$ | $Spread''(S, F)$ |

### 3.2.3 Clustering by GHSOM

Each observation can extract a feature vector from the data from the previous hedging horizon. The feature vectors of the observations in the estimation interval include input matrix of GHSOM for OHR estimation. The GHSOM algorithm in this study is implemented in MATLAB (Chan and Pampalk, 2002). When using the GHSOM, the parameters related to network topology, such as breadth (e.g., breadth of map) and depth (e.g., depth of GHSOM), need to be determined first. If the similarity of data in the same cluster is below a certain threshold, the data will be clustered once again by breadth or depth, hence expanding the SOM clusters. To emphasize the hierarchical relationship of the clusters and to avoid data from being too concentrated on some clusters, we set the depth parameter as 0.001 and the breadth parameter as 0.8. Items that cannot be expanded are restricted when they are less than 100. The topology of the clusters is automatically determined by the data and related with the threshold setting of width and depth expansion.

After the GHSOM is initialized, we input the features vector extracted from the historical time series to the GHSOM. The feature vectors are then processed by min-max normalization, which maps the value of the vector from -1 to 1. Normalization can ensure stable results. The approach has been widely used in computational intelligence, such as neural network and genetic algorithm. Using competitive learning, the output of a neuron is determined by calculating a similarity measure between the weight of the neuron and the external input. Furthermore, the topology can grow and form a hierarchical architecture when the neuron exceeds the quality measurement request. Finally, the input data can be hierarchically clustered. Figure 3-3 shows an example of the hierarchical clusters obtained from the GHSOM. The data are grouped into several groups in a hierarchical structure of four layers. The largest rectangle is Layer 0, which contains 2X2 SOM clusters in Layer 1. The upper-left cluster in Layer 1 also contains 2X2 SOM clusters, and so on.

Figure 3-3. Result of GHSOM clustering

### 3.2.4 Identifying Cluster Structure

The historical financial time series data were hierarchically clustered by the GHSOM with similar patterns. Results show that the hierarchical architecture consists of many clusters distributed in different layers. The relations of hierarchical clusters are illustrated in Figure 3-4. The sample size of each cluster is different. Clusters in the upper layers of the hierarchical architecture contain more samples of observations than those in the lower layers. The hierarchical architecture also represents the degree of similarity. Any observation can be identified on the cluster based on the layer it belongs to. The host cluster in the lowest layer contains the least data but has the highest similarity with the forecasting data. In addition, similarity with data is decreased in the upper layer clusters. Figure 3-5 is a real example which exhibits an observation and its similar observations in the clusters. The observation of the date 21 July, 2000 can find its similar observations in three layers. Each layer has a cluster it belonged to. The number of observations is decreased when the depth of layer increase.

With regard to the latest observation in the estimation interval, similar observations can be obtained based on the group it belongs to in each layer. To forecast the fluctuation of the spot and futures for the next hedging horizon, we collect the observations which are one hedging horizon ahead the similar observations in the same clusters. Figure 3-6 illustrates the observations collection for forecasting. These observations are weighted by similarity based on the cluster they belong to, and are used to modify the probability distribution for OHR estimation.

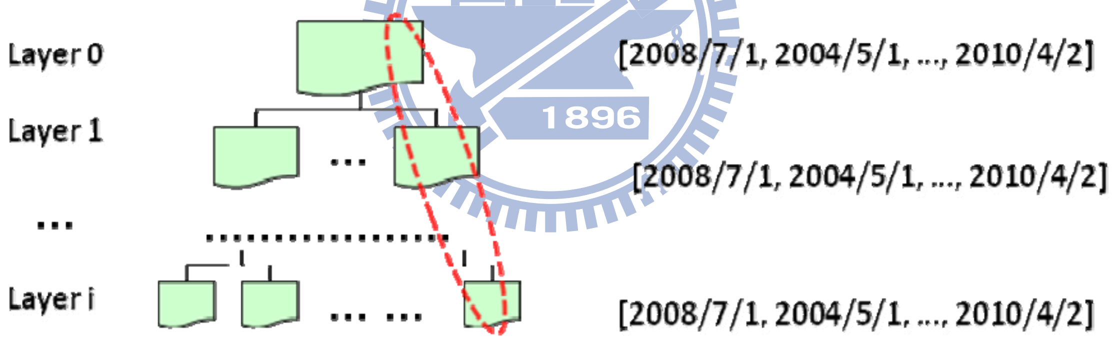



**Figure 3-4. An example of the hierarchical clustered data**

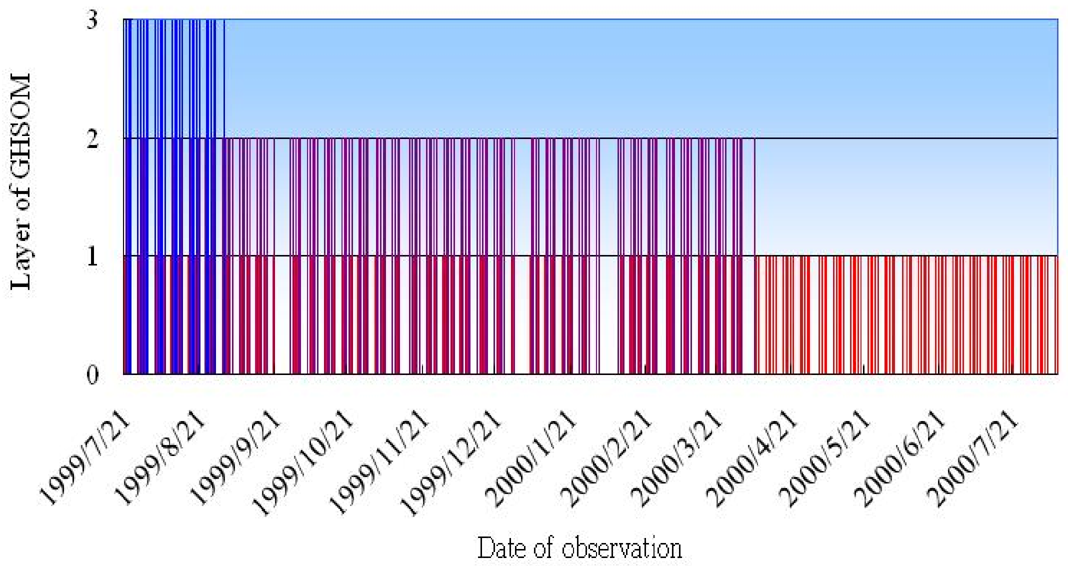



**Figure 3-5. An observation and its similar observations in the clusters**

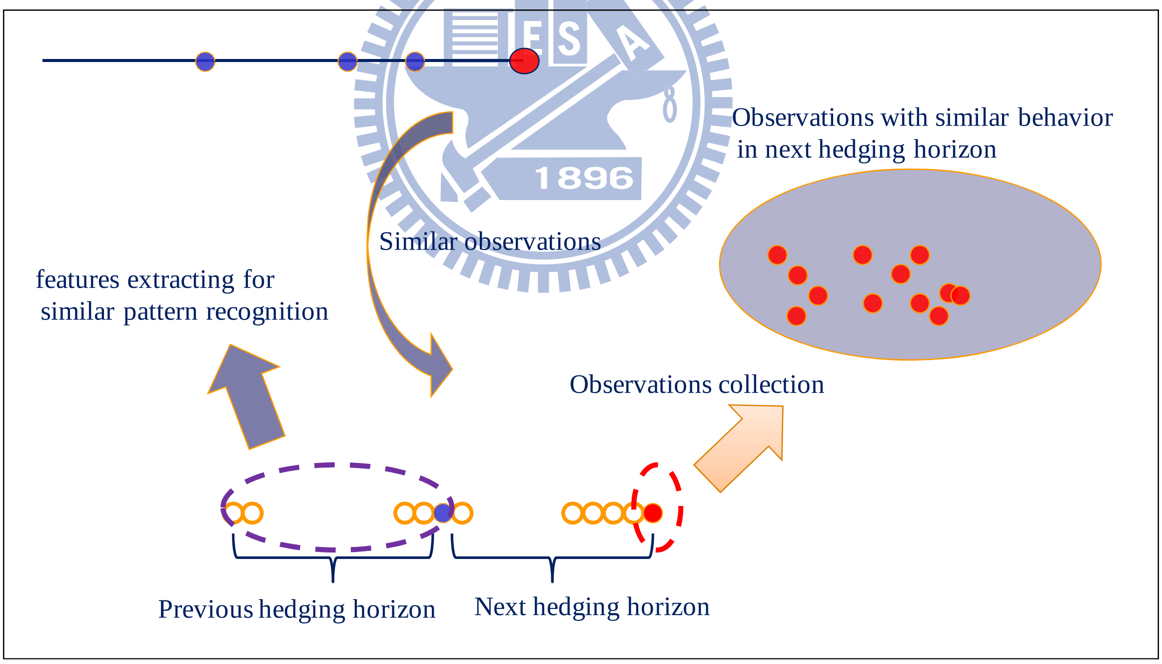



**Figure 3-6. The observations collection for forecasting**

### 3.2.5 Data Resampling and Weighting

Regardless of similarity in data, every observation in the estimation interval is given equal weight for OHR estimation in the conventional model. In our proposed CI approach, for OHR estimation, the observation is given a different weight according to similarities. Although the size of the clusters obtained by GHSOM is different, the more similar data in the lower level clusters would be given more weight than the upper level clusters in the resampling process. The weights of observation are suggested for data replication. The more similar data will be replicated more frequently, thus increasing the occurrence probability in the whole population.

For each layer in the hierarchical cluster, the data in the same cluster are replicated randomly until the sample size coincides with the original sample size of the estimation period. The sample size is expanded by multiplying the layer of the hierarchical architecture. For example, in the conventional approach, if the size of observations in the estimating period is 1000, the OHR for the next hedging horizon is estimated based on 1000 observations. In the proposed CI approach, if 1000 observations are clustered into three layers by GHSOM, the OHR will be estimated based on triple observations, all generated by resampling. The expansion of sample size avoids the small sample effect and increases the accuracy of OHR estimation.

After data resampling, we can obtain a collection of observations composed of similar data in the cluster and consequent observations in the following hedging horizon with different weight. The pseudo code used to form the collection is described in Figure 3-7. The collection can include the original observation used for OHR estimation in order to modify the probability distribution of the original time series with unequal weight measured by similarity. Finally, the OHR can be estimated by the traditional OLS method. Calculating the variance and covariance of spot and futures based on the collection of observations is carried out.

```
Function modifying distribution ()
    For (each layer in the GHSOM cluster result)
        Identify the cluster number clus_{l,c} where the latest observation obs_{latest} belongs to
        For (each observation obs_i in the cluster clus_{l,c})
            Insert the observations obs_{i+h} into the collection col_l
        End for
        Resample the collection col_l randomly with repetition until the size of col_l equals the
            estimation interval
    End for
```

$clus_{l,c}$ : the cluster number, where $l$ is the number of layers, and $c$ is the number of the cluster in layer $l$

$obs_{latest}$ : the latest observation in a rolling window

$obs_i$ $obs_{i+h}$ : the observations of a cluster, where $i$ is the number of observations, and $h$ is the hedging horizon

$col_l$ : the collection of observations for layer $l$

**Figure 3-7. The pseudo code of data resampling and weighting**

### 3.2.6 Estimating OHR with OLS

The basic concept of hedging involves the elimination of fluctuations in the value of a spot position by tracking futures contracts that are opposite to the position held by the spot market. For a long position in the spot market, the return of a hedged portfolio is given by

$$\Delta HP = \Delta S - r \times \Delta F \tag{3-8}$$

where $\Delta HP$ is the change in the value of the hedge portfolio; $\Delta S$ and $\Delta F$ are the changes in the spot and futures prices, respectively; and $r$ is the hedge ratio. Changes in spot and futures prices are also considered as returns, which are calculated by Equation (3-1). OHR is the value of $r$ that maximizes the expected utility of the investor; it is defined as the expected return and risk of the hedged portfolio. The expected return of futures is 0 when the futures price follows a martingale; hence, the futures position will not affect the expected return of the portfolio.

The risk of the hedge portfolio is defined by its variance in the mean-variance framework. Therefore, OHR is simply the value of $r$ that minimizes the variance of Equation (3-8), which is given by

$$\frac{\partial Var(\Delta HP)}{\partial r} = 2r \times Var(\Delta F) - 2Cov(\Delta S, \Delta F) = 0 \tag{3-9}$$

where $Var(\Delta F)$ is the variance of the futures return and $Cov(\Delta S, \Delta F)$ is the covariance between the spot return and the futures return. OHR is determined by solving Equation (3-9):

$$r^* = \frac{Cov(\Delta S, \Delta F)}{Var(\Delta F)} \tag{3-10}$$

The OHR given by Equation (3-10) can be estimated by regressing the spot return on the futures return using OLS, which corresponds to conventional OHR.

In this study, OHR estimation is improved by replacing the original observations in the estimation period with the collection of observations, which is composed of similar data in the cluster. Their consequent observations in the following hedging horizon have different weights. The traditional OLS method for OHR estimation, expressed by Equation (3-10), is modified to Equation (3-11), in which $\Delta\tilde{S}$ and $\Delta\tilde{F}$ refer to the collection of observations derived from spot and futures return series, respectively.

$$r^* = \frac{Cov(\Delta\tilde{S}, \Delta\tilde{F})}{Var(\Delta\tilde{F})} \tag{3-11}$$

## 3.3 Model Evaluating Criteria

### 3.3.1 Hedging Effectiveness

Hedging performance is typically evaluated by hedging effectiveness (HE). The degree of hedging effectiveness is measured by the percentage reduction in the variance of portfolio after hedging (Geppert, 1995). The variance of hedge portfolio with estimated OHR can be expressed as

$$Var_{hedged} = Var(\Delta HP) = Var(\Delta S_t - r \times \Delta F_t) \quad (3\text{-}12)$$

where $r$ is the OHR. Therefore, HE can be expressed as

$$HE = \frac{Var_{un-hedged} - Var_{hedged}}{Var_{unhedged}} \times 100\% = \frac{Var(\Delta S) - Var(\Delta HP)}{Var(\Delta S)} \times 100\% = (1 - \frac{Var(\Delta HP)}{Var(\Delta S)}) \times 100\% \quad (3\text{-}13)$$

The value of HE can be used to evaluate the model of OHR estimation. A higher HE represents better OHR estimation, and vice versa.

### 3.3.2 White's Reality Check

In comparing the different OHR estimation models and to test the statistical significance of variance deduction, we apply White's Reality Check (White, 2000), which has been used to compare hedging models (Lee and Yoder, 2007). The Reality Check consists of a non-parametric test that checks if any of the numbers in the concurrent methods yield forecasts that are significantly better than a given benchmark method; then, it corrects the data snooping bias. Data snooping bias may occur when a given dataset is reused by one or more researchers for model selection. The null hypothesis that the performance of the proposed hedging model has no predictive superiority over the conventional model is not rejected. The hypotheses are as follows:

*H0: No method is better than the benchmark.*

*H1: At least one method is better than the benchmark.*

The detailed process of White's Reality Check can be found in literature (White, 2000; Lee and Yoder, 2007).

# Chapter 4 Experimental Design and Results Analysis

## 4.1 Experimental Design

The experiments in this study are designed with two objectives: feasibility of the proposed CI-based model and hedging performance over various hedging horizons for OHR decision making based on different models. Figure 4-1 shows the framework of the experiments.

First, on the left side of the figure, the feasibility of the proposed CI-based model is examined using dynamic behaviors extracted as the feature of the time series. The feature-extracting process of the proposed model is tested in different settings to achieve the best parameters. The feature vectors that represent the dynamic behaviors of time series for GHSOM similarity measurement are composed of variance, covariance, price spread, and their first and second order differences. We design six combinations of these parameters, which are adopted in the experimental models and denoted by CI_1, CI_2, CI_3, CI_4, CI_5, and CI_6, respectively, to verify the performance over various hedging horizons. Table 4-1 presents the parameter settings of these models.

Second, on the right side of the figure, the optimal hedge ratio is estimated by the proposed model concerned with the hedging horizon, and the performances are compared with conventional models. The hedging decision is evaluated by hedging effectiveness. For each hedging horizon in the testing period, the hedged portfolio is adjusted once according to the latest OHR at the beginning of a hedge horizon and lasts until the beginning of the next hedging horizon. At the end of the testing period, hedging effectiveness is calculated based on the variance of the hedging portfolio in each hedging horizon. Hedge horizons in the experiments are set at 1, 7, 14, 21, and 28 days, which cover the intervals from short-term to long-term. To compare hedging performance, the superiority of the proposed model is verified using two conventional models, the OLS and naïve models, both of which are widely used in OHR research on different hedging horizons (Chen *et al.* ,2004; In and Kim, 2006).

Furthermore, to obtain a better understanding of the properties of the proposed model, several experiments are designed separately to investigate the influence of the period of data for feature extraction and the period of experiment data. The design and result for these experiments are presented in Sections 4.6 and 4.7.

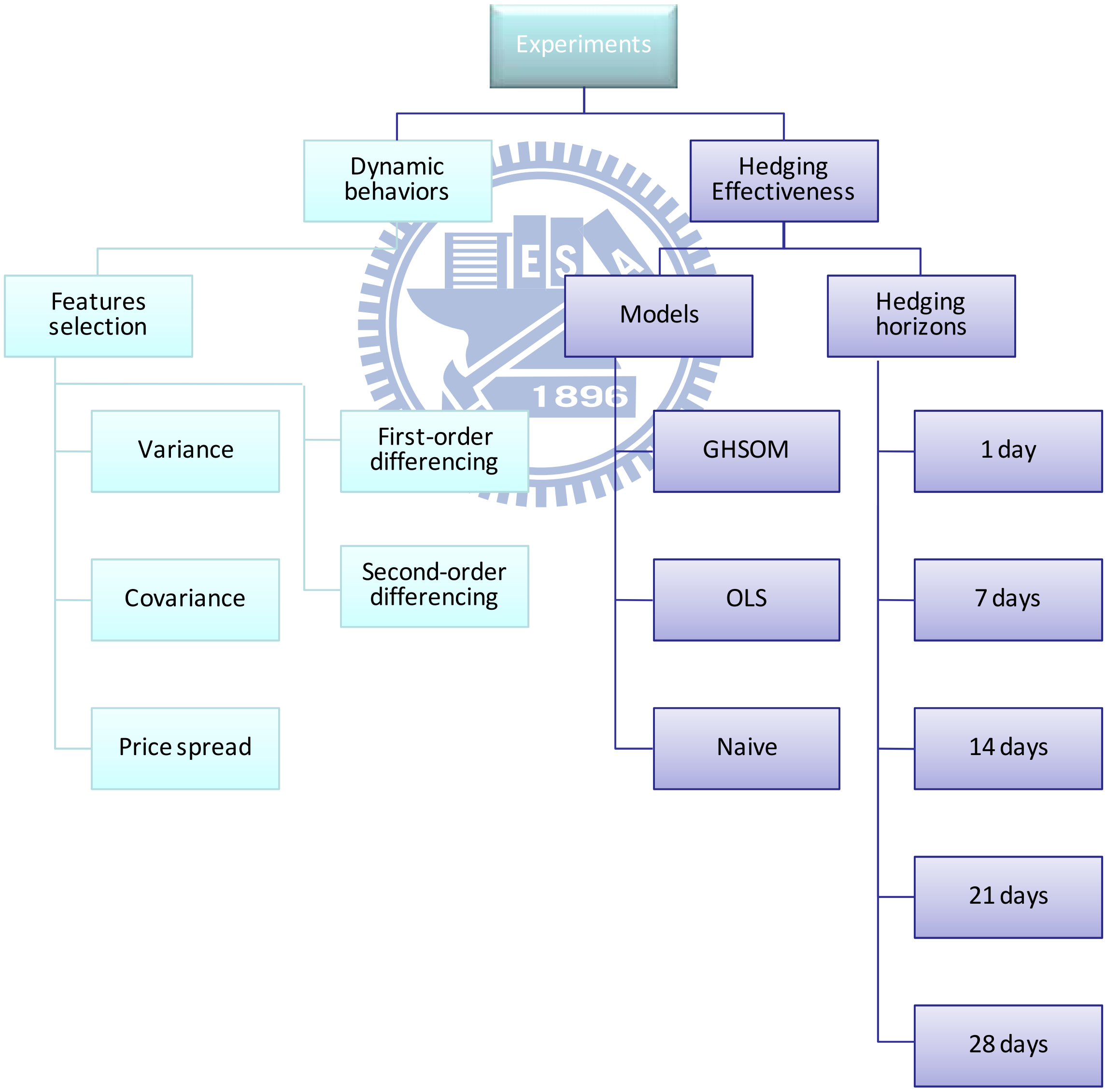



**Figure 4-1. Experiment framework**

**Table 4-1. Parameter settings for testing dynamic behavior**

| Model | Parameter setting | | |
|---|---|---|---|
| | Variance | Covariance | Price spread |
| CI_1 | $Var(\Delta S)$, $Var(\Delta F)$ | | |
| CI_2 | $Var(\Delta S)$, $Var(\Delta F)$<br>$Var'(\Delta S)$, $Var'(\Delta F)$<br>$Var''(\Delta S)$, $Var''(\Delta F)$ | | |
| CI_3 | $Var(\Delta S)$, $Var(\Delta F)$ | $Cov(\Delta S, \Delta F)$ | $Spread(S, F)$ |
| CI_4 | $Var(\Delta S)$, $Var(\Delta F)$<br>$Var'(\Delta S)$, $Var'(\Delta F)$<br>$Var''(\Delta S)$, $Var''(\Delta F)$ | $Cov(\Delta S, \Delta F)$ | $Spread(S, F)$ |
| CI_5 | $Var(\Delta S)$, $Var(\Delta F)$ | $Cov(\Delta S, \Delta F)$<br>$Cov'(\Delta S, \Delta F)$<br>$Cov''(\Delta S, \Delta F)$ | $Spread(S, F)$<br>$Spread'(S, F)$<br>$Spread''(S, F)$ |
| CI_6 | $Var(\Delta S)$, $Var(\Delta F)$<br>$Var'(\Delta S)$, $Var'(\Delta F)$<br>$Var''(\Delta S)$, $Var''(\Delta F)$ | $Cov(\Delta S, \Delta F)$<br>$Cov'(\Delta S, \Delta F)$<br>$Cov''(\Delta S, \Delta F)$ | $Spread(S, F)$<br>$Spread'(S, F)$<br>$Spread''(S, F)$ |

## 4.2 Experiment Data and Basic Statistics

This study obtained empirical trading data of the daily closing price from various stock and futures markets, including Taiwan Weighted Index (TWI), Standard & Poor's 500 Index (S&P 500), Financial Times Stock Exchange 100 Index (FTSE 100), NIKKEI 255 Index, and their correlative futures contracts. Table 4-2 lists the stock market index and exchange of their correlative futures contracts trade. All data were obtained from the Thomson Datastream database in the same period from July 21, 1999 to July 18, 2008. The futures prices series was gathered from the nearest month contracts and rolled over to the next nearest contracts on the maturity day due to the consideration of liquidity and price spread risk. The return series are defined as the logarithmic first difference of price series multiplied by 100 using Equation (3-1). The numbers of observation for each market are listed in Table 4-2. Among the total observations, the first 90% is considered the estimation period, and the remaining 10% is considered the testing period.

Table 4-3 shows some basic distributional characteristics of the spot and futures return series. All eight series show high significant skewness, kurtosis, and Jarque-Bera (JB) statistics, implying non-normal distributions with fatter tails. A comparison of the standard deviation of return, kurtosis, and JB statistics indicate that the largest and smallest discrepancy between the spot and futures data are in TWI and FTSE 100, respectively. In other words, the correlation between spot and futures is highest in FTSE 100 and lowest in TWI. The large discrepancy between the spot and futures data displays more extreme movements than would be predicted by a normal distribution. The F-test for equal variance between spot and futures also indicates different characteristics in each market. The result shows that the null hypothesis of equal variance is rejected in TWI, but cannot be rejected in S&P 500, FTSE 100, and NIKKEI 255 Index. Consequently, the data of the same period gathered from different markets may exhibit different behaviors and cause inconsistencies in the results.

**Table 4-2. Experiment data**

| **Index (Spot)** | **Exchange (Futures)** | **Observations** |
|---|---|---|
| Taiwan Weighted Index (TWI) | Taiwan Futures Exchange (TAIFEX) | 2217 |
| Standard & Poor's 500 index (S&P 500) | Chicago Mercantile Exchange (CME) | 2263 |
| Financial Times Stock Exchange 100 Index (FTSE 100) | London International Financial Futures and Options Exchange (LIFFE) | 2215 |
| NIKKEI 255 Index | Osaka Securities Exchange (OSE) | 2275 |

Note: Data period is from July 21, 1999 to July 18, 2008.

**Table 4-3. Basic distributional statistics of return series**

| | TWI | | S&P 500 | | FTSE 100 | | NIKKEI 255 | |
|---|---|---|---|---|---|---|---|---|
| | **Spot** | **Futures** | **Spot** | **Futures** | **Spot** | **Futures** | **Spot** | **Futures** |
| Mean | -0.0060 | -0.0068 | -0.0040 | -0.0042 | -0.0072 | -0.0073 | -0.0160 | -0.0158 |
| Maximum | 6.1721 | 6.7659 | 5.5744 | 5.7549 | 5.9026 | 5.9506 | 7.2217 | 8.0043 |
| Minimum | -9.9360 | -11.0795 | -6.0045 | -6.2709 | -5.8853 | -6.0625 | -7.2340 | -7.5986 |
| Std. Dev. | 1.5931 | 1.8262 | 1.1287 | 1.1404 | 1.1657 | 1.1663 | 1.4058 | 1.4346 |
| Kurtosis | 5.2942 | 5.8891 | 5.2732 | 5.4067 | 5.7929 | 5.7588 | 4.6042 | 4.6496 |
| Skewness | -0.1883 | -0.1867 | 0.0600 | 0.0274 | -0.2096 | -0.1658 | -0.2075 | -0.2122 |
| Jarque-Bera (JB) | 499.0886 *** | 783.5957 *** | 488.3711 *** | 546.2142 *** | 755.7494 *** | 731.5745 *** | 253.2913 *** | 267.633 *** |
| F-test for equal variances (p value) | 0.0000 *** | | 0.6247 | | 0.9790 | | 0.3396 | |

Note: (1) *** represents significance at the 1% level. (2) The skewness of normal distribution is zero. (3) The kurtosis of normal distribution is 3. (3) The hypothesis of F-test is that two independent samples, spot and futures return, come from normal distributions with the same variance.

## 4.3 Comparisons of Dynamic Behaviors Prediction

In this study, we apply the features extracted from the time series to represent dynamic behavior as the input variable for time series clustering by GHSOM. The variance, covariance, price spread, and first and second differences of the observations in previous hedging horizons are suggested to capture the dynamic behavior for predicting fluctuations in the next hedging horizon. The feasibility of this approach is examined using the six combinations of parameters listed in Table 4-1. Model performance is evaluated via hedging effectiveness over various hedging horizons in the four stock markets.

Table 4-4 presents the hedging effectiveness for all models. Results indicate that based on the same experiment data, the CI-based model can obtain the best performance compared with the traditional OLS and naïve models, except for short-term hedging in FTSE 100 and one day hedging in S&P 500. A comparison of the six experiment models in all market data indicates that the best CI-based model is different over different hedging horizons. For seven days hedging, the CI_1 model is the best model in all market data. However, for the 1 day and 28 days hedging, the CI_2 and CI_4 models are the best models in three of four market data, respectively.

The results imply that the ability to capture fluctuation under various timescales is different for CI-based models. Short-term dynamic behavior may be captured by variance and its first and second differences. Long-term tendency may need more variables for its description than short-term tendency by adding covariance and price spread.

**Table 4-4. Comparisons of dynamic behaviors**

| Market/model | | Hedging effectiveness | | | | |
|---|---|---|---|---|---|---|
| | | Hedging horizon (days) | | | | |
| | | 1 | 7 | 14 | 21 | 28 |
| **TWI** | | | | | | |
| **CI** | **1** | 93.3309% * | 97.1661% ** | 99.2656% | 99.3811% | 99.3131% |
| | **2** | 93.3905% ** | 97.1534% * | 99.2480% | 99.3942% | 99.3556% * |
| | **3** | 93.2715% | 96.9289% | 99.2751% | 99.3947% | 99.3160% |
| | **4** | 93.0998% | 96.8809% | 99.2879% * | 99.4342% * | 99.3802% ** |
| | **5** | 93.1081% | 97.0102% | 99.3111% ** | 99.4327% | 99.3431% |
| | **6** | 93.1798% | 96.9169% | 99.2047% | 99.4666% ** | 99.3470% |
| **OLS** | | 93.3055% | 97.0244% | 99.1612% | 99.3860% | 99.3089% |
| **Naïve** | | 90.6982% | 96.0278% | 98.5331% | 99.0888% | 98.8415% |
| **S&P 500** | | | | | | |
| **CI** | **1** | 96.6140% | 99.1678% ** | 99.3126% | 99.6413% | 99.6493% |
| | **2** | 96.6662% * | 99.1206% | 99.3777% | 99.6010% | 99.6513% |
| | **3** | 96.6401% | 99.1236% | 99.3860% ** | 99.6888% ** | 99.7094% |
| | **4** | 96.6447% | 99.0880% | 99.3856% * | 99.6806% * | 99.7310% ** |
| | **5** | 96.6069% | 99.0630% | 99.3837% | 99.6774% | 99.6793% |
| | **6** | 96.6221% | 99.0656% | 99.3667% | 99.6661% | 99.7263% * |
| **OLS** | | 96.6510% | 99.1287% * | 99.3705% | 99.5911% | 99.6131% |
| **Naïve** | | 96.7974% ** | 99.0029% | 99.3045% | 99.4826% | 99.5752% |
| **FTSE 100** | | | | | | |
| **CI** | **1** | 96.9688% | 98.5911% | 98.5323% | 99.0140% | 99.4868% ** |
| | **2** | 97.0511% * | 98.5522% | 98.5596% | 99.1438% | 99.4483% |
| | **3** | 96.9842% | 98.5673% | 98.5690% | 99.1743% | 99.4857% * |
| | **4** | 96.9828% | 98.6141% * | 98.5572% | 99.2197% * | 99.4751% |
| | **5** | 96.9945% | 98.5904% | 98.5883% | 99.1774% | 99.4475% |
| | **6** | 97.0278% | 98.5823% | 98.6069% | 99.2263% ** | 99.4764% |
| **OLS** | | 97.0130% | 98.5552% | 98.6306% * | 99.0979% | 99.4767% |
| **Naïve** | | 97.1492% ** | 98.6258% ** | 98.8094% ** | 99.1005% | 99.3223% |

| | | | | | | |
|---|---|---|---|---|---|---|
| **NIKKEI 255** | | | | | | |
| **CI** | **1** | 96.4072% | 99.4409% ** | 99.5007% | 99.5261% | 99.8709% |
| | **2** | 96.4909% | 99.3939% | 99.4533% | 99.5560% ** | 99.9036% * |
| | **3** | 96.5677% ** | 99.4311% * | 99.5075% * | 99.5499% * | 99.8948% |
| | **4** | 96.5026% | 99.4310% | 99.4954% | 99.5245% | 99.9051% ** |
| | **5** | 96.5099% | 99.4037% | 99.5107% ** | 99.5274% | 99.8899% |
| | **6** | 96.4615% | 99.3797% | 99.4698% | 99.5431% | 99.8893% |
| **OLS** | | 96.5501% * | 99.4271% | 99.5002% | 99.5072% | 99.9023% |
| **Naïve** | | 96.2222% | 99.3305% | 99.4317% | 99.4585% | 99.8711% |

Note: ** and * represent the best and second best HE among eight models at the same hedging horizon, respectively.

## 4.4 Comparison of Hedging Performance

For a comparison of hedging performance, we list the best CI-based model from the six experiments models, and the two conventional models (naïve and OLS) in Table 4-5. The hedging performance of the model is evaluated using hedging effectiveness and statistic testing for significance of superiority. The hedging effectiveness of the model is calculated using the variance reduction of the hedged portfolio (Table 4-4). Table 4-5 presents the variance of unhedged and hedged portfolios employed in White's reality check to verify the significance of superiority.

The variance of Table 4-5 shows that increasing the hedging horizon will increase the variance of unhedged portfolio but will be effectively reduced by the hedging model. The percentage of variance reduction, shown as hedging effectiveness in Table 4-4, is higher in a long hedging horizon than in a short one.

A comparison of the model using the variance of hedged portfolio in Table 4-5 shows that the CI-based model is superior to the OLS model; for TWI and NIKKEI 255, the CI-based model obtains the minimum variance in all hedging horizons. However, the conventional OLS model cannot obtain minimum variance for all markets. Notably, for FTSE 100 and S&P 500 in short-term hedging, the static naïve model obtains the minimum variance. A possible reason for this may be the high correlation of the fluctuations of spot and futures for FTSE 100. This can be observed from the closing statistics value of the spot and futures market in Table 4-2.

The value of hedging effectiveness is slightly different in these models. To test the significance of these models' performance improvements, we perform White's reality check. When OLS is treated as the benchmark, the null hypothesis of no improvement of CI-based model over benchmark is rejected for 28 days hedging in TWI, 21 and 28 days hedging in S&P 500, 21 days hedging in FTSE 100 and NIKKEI 255 at the significance level of 1%. Results of the reality check provide evidence that the proposed CI-model can improve the OLS model, especially in long-term hedging.

**Table 4-5. Variance of the portfolio**

| Market/models | Variance | | | | |
|---|---|---|---|---|---|
| | **Hedging horizon** | | | | |
| | **1** | **7** | **14** | **21** | **28** |
| **TWI** | | | | | |
| **Unhedged** | 2.7527 | 20.5443 | 41.8060 | 35.5709 | 39.9629 |
| **Naïve** | 0.2561 | 0.8161 | 0.6132 | 0.1840 | 0.1698 |
| **OLS** | 0.1843 | 0.6113 | 0.3507 | 0.1454 | 0.1546 |
| **CI-based** | 0.1819 *** | 0.5822 *** | 0.2880 *** | 0.1148 *** | 0.1056 *** |
| **Reality check p value** | 0.134 | 0.026 * | 0.015 * | 0.085 | 0.000 ** |
| **S&P 500** | | | | | |
| **Unhedged** | 1.6688 | 7.7995 | 14.9320 | 35.5709 | 39.9629 |
| **Naïve** | 0.0534 *** | 0.0778 | 0.1039 | 0.1840 | 0.1698 |
| **OLS** | 0.0559 | 0.0680 | 0.0940 | 0.1454 | 0.1546 |
| **CI-based** | 0.0556 | 0.0649 *** | 0.0917 *** | 0.1107 *** | 0.1075 *** |
| **Reality check p value** | 0.354 | 0.072 | 0.257 | 0.000 ** | 0.000 ** |
| **FTSE 100** | | | | | |
| **Unhedged** | 2.0479 | 8.3512 | 23.6463 | 35.0068 | 57.8882 |
| **Naïve** | 0.0584 *** | 0.1148 *** | 0.2815 *** | 0.3149 | 0.3923 |
| **OLS** | 0.0612 | 0.1207 | 0.3238 | 0.3158 | 0.3029 |
| **CI-based** | 0.0604 | 0.1157 | 0.3294 | 0.2709 *** | 0.2971 *** |
| **Reality check p value** | 0.039 * | 0.022 * | 1.000 | 0.002 ** | 0.050 * |
| **NIKKEI 255** | | | | | |
| **Unhedged** | 2.8548 | 17.2076 | 46.2477 | 38.9337 | 87.7568 |
| **Naïve** | 0.1078 | 0.1152 | 0.2628 | 0.2108 | 0.1131 |
| **OLS** | 0.0985 | 0.0986 | 0.2312 | 0.1919 | 0.0857 |
| **CI-based** | 0.0980 *** | 0.0962 *** | 0.2263 *** | 0.1729 *** | 0.0833 *** |
| **Reality check p value** | 0.318 | 0.078 | 0.151 | 0.004 ** | 0.136 |

Note: (1) The benchmark model for White's reality check is the OLS model. (2) * and ** represent significance at the 5% and 1% levels, respectively. (3) *** represents the minimum variance among the naïve, OLS, and CI-based hedged portfolios.

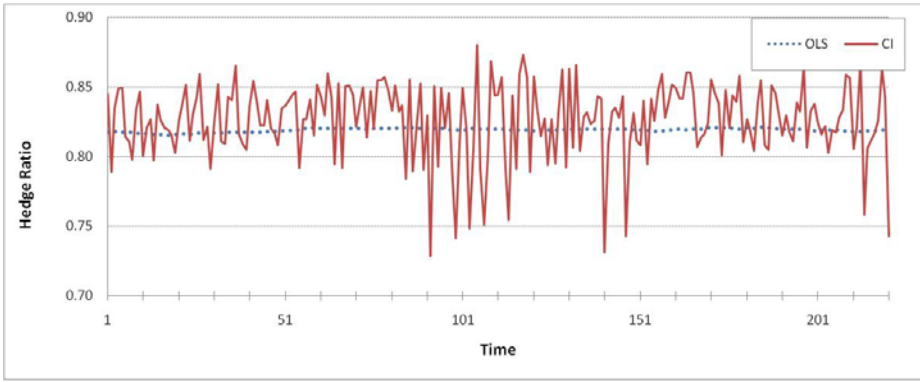
## 4.5 Comparison of OHR

Time series clustering and with-in cluster resampling are used to obtain OHR for the CI-based model. Table 4-6, which includes OLS and the best CI-based model for comparison, presents the average OHR and standard deviation for the underlying models. For all market data, the average OHR estimated using CI-based and OLS models is very close though a large discrepancy exists in the standard deviation. Maximum standard deviation of the OLS model is 0.0069 for the one day hedging for FTSE 100. However, minimum standard deviation of the CI-based model is 0.0070 for 28 days hedging for S&P 500. Results suggest that the OHR estimated using the CI-based model is more variant than that of the OLS model.

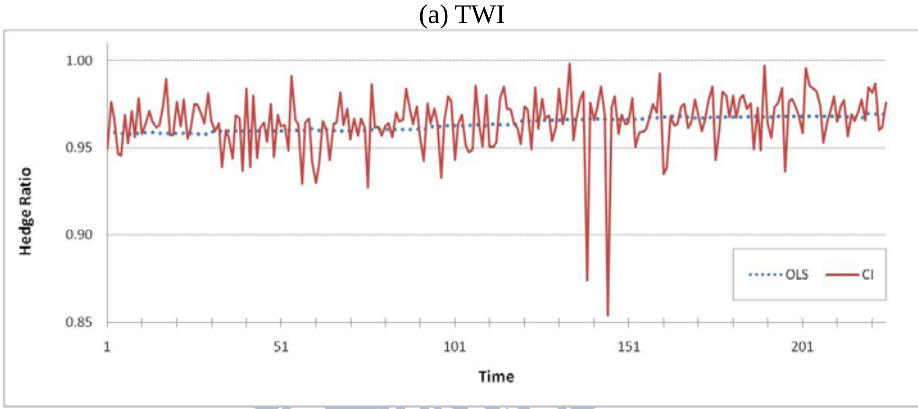
Figures 4-2, 4-3, 4-4, 4-5, and 4-6 present the plot of OHR estimated by the best CI-based and OLS models over 1, 7, 14, 21, and 28 days hedging horizons for all market data, respectively. In all figures, the OHR estimated using the traditional OLS model approximates a straight line, and the values are almost the same during the hedge period. However, the OHR estimated using the CI-based model is time-varying, which can reflect the dynamic behavior of the financial time series.

**Table 4-6. Comparison of OHR**

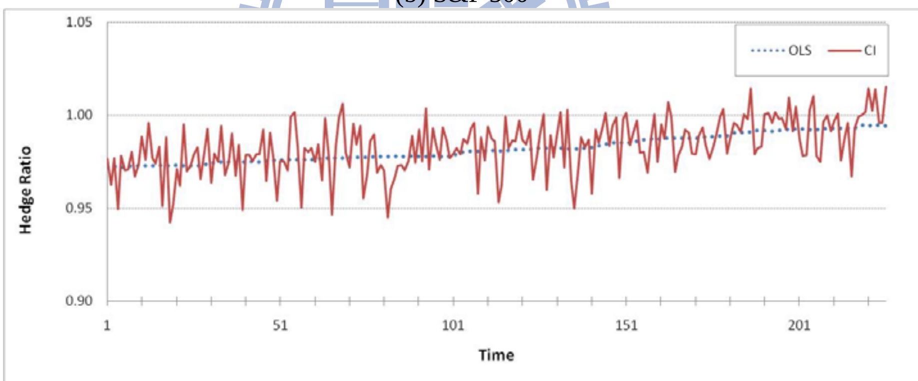
| Hedging horizon/ Model | OHR | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | TWI | | S&P 500 | | FTSE 100 | | NIKKEI 255 | |
| | Mean | Std. Dev. | Mean | Std. Dev. | Mean | Std. Dev. | Mean | Std. Dev. |
| **1** | | | | | | | | |
| **OLS** | 0.8189 | 0.0012 | 0.9636 | 0.0037 | 0.9819 | 0.0069 | 0.9429 | 0.0015 |
| **CI-based** | 0.8263 | 0.0268 | 0.9649 | 0.0163 | 0.9833 | 0.0145 | 0.9474 | 0.0175 |
| **7** | | | | | | | | |
| **OLS** | 0.9423 | 0.0019 | 0.9746 | 0.0021 | 0.9738 | 0.0040 | 0.9799 | 0.0010 |
| **CI-based** | 0.9288 | 0.0150 | 0.9751 | 0.0088 | 0.9821 | 0.0100 | 0.9826 | 0.0058 |
| **14** | | | | | | | | |
| **OLS** | 0.9570 | 0.0019 | 0.9818 | 0.0024 | 0.9746 | 0.0057 | 0.9737 | 0.0015 |
| **CI-based** | 0.9493 | 0.0134 | 0.9821 | 0.0143 | 0.9851 | 0.0159 | 0.9832 | 0.0078 |
| **21** | | | | | | | | |
| **OLS** | 0.9625 | 0.0021 | 0.9872 | 0.0039 | 0.9647 | 0.0045 | 0.9851 | 0.0019 |
| **CI-based** | 0.9598 | 0.0145 | 0.9847 | 0.0099 | 0.9701 | 0.0149 | 0.9867 | 0.0078 |
| **28** | | | | | | | | |
| **OLS** | 0.9647 | 0.0012 | 0.9960 | 0.0023 | 0.9589 | 0.0045 | 0.9764 | 0.0010 |
| **CI-based** | 0.9624 | 0.0066 | 0.9882 | 0.0070 | 0.9676 | 0.0118 | 0.9858 | 0.0083 |



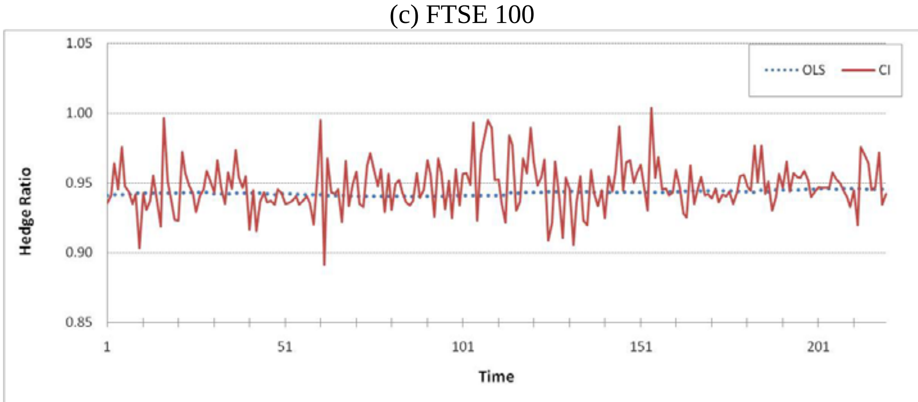

(a) TWI




(b) S&P 500

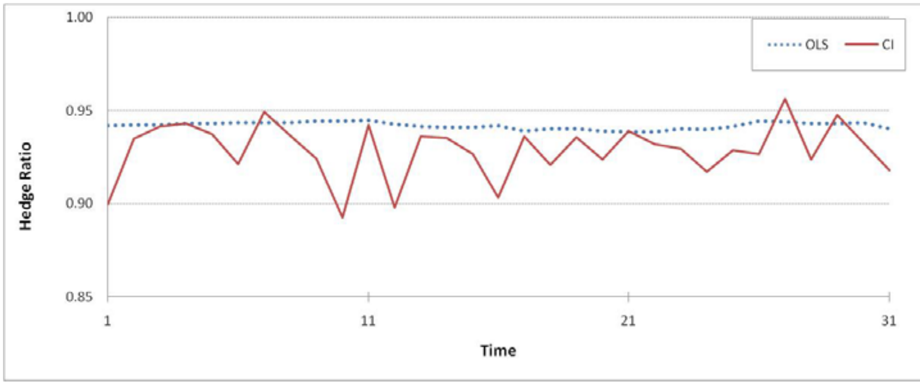



(c) FTSE 100

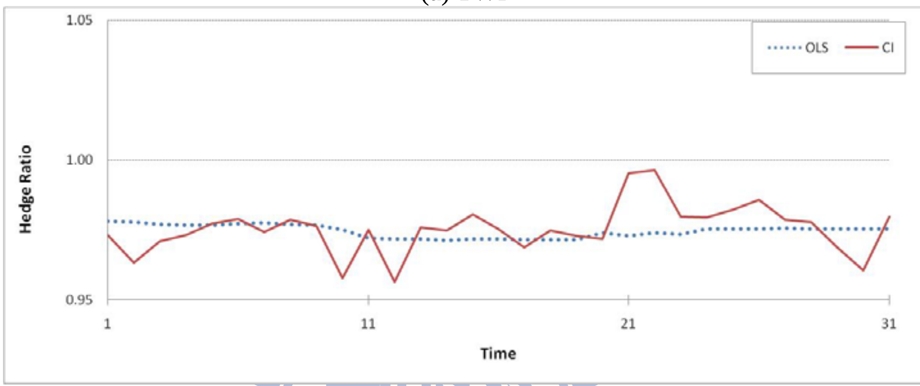



(d) NIKKEI 255

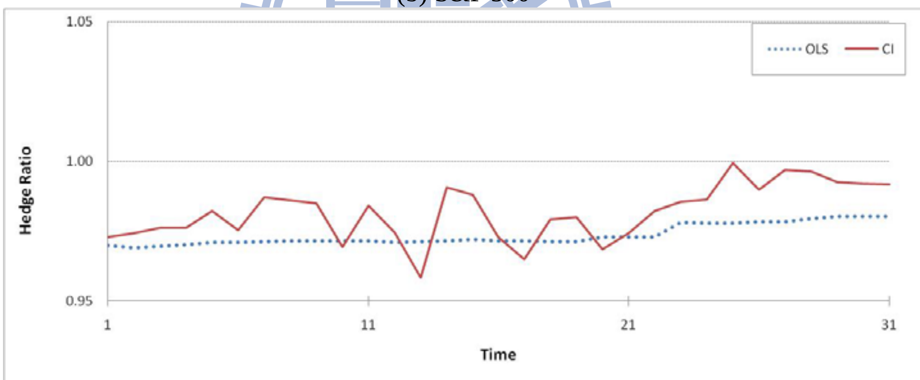
**Figure 4-2. Comparison of OHR in 1 day hedging**



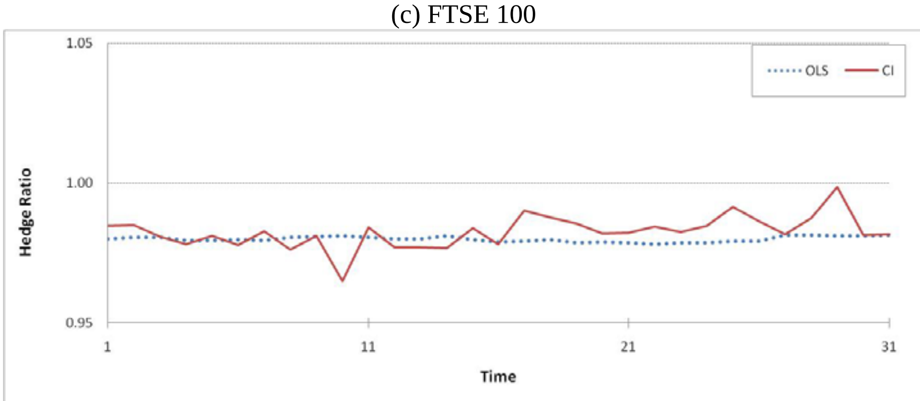

(a) TWI




(b) S&P 500

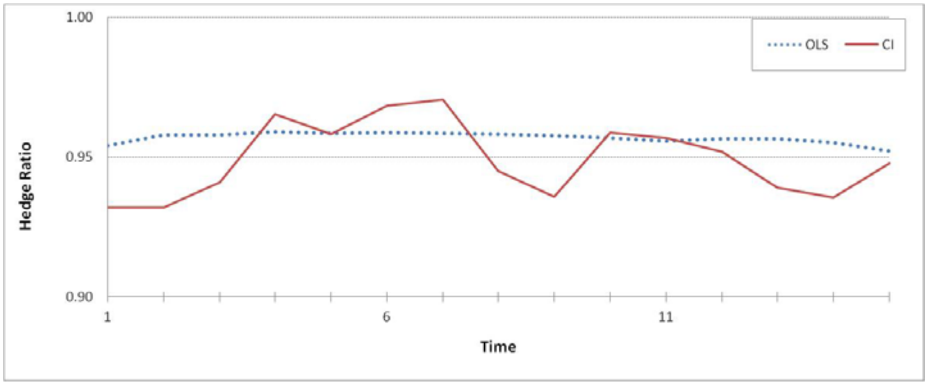



(c) FTSE 100

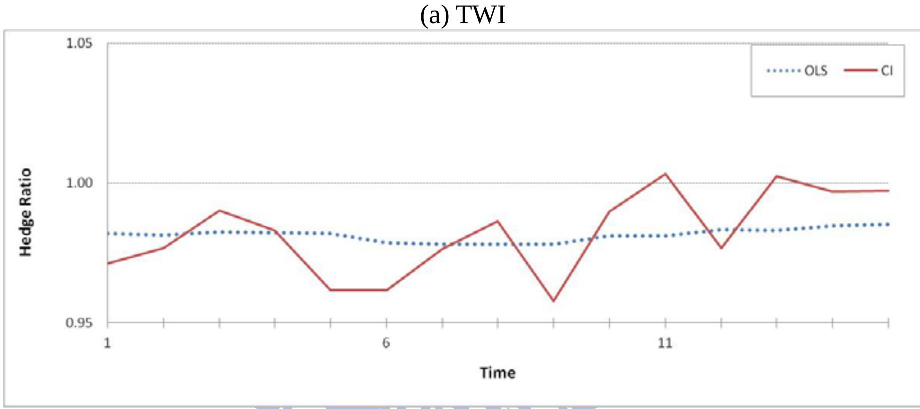



(d) NIKKEI 255

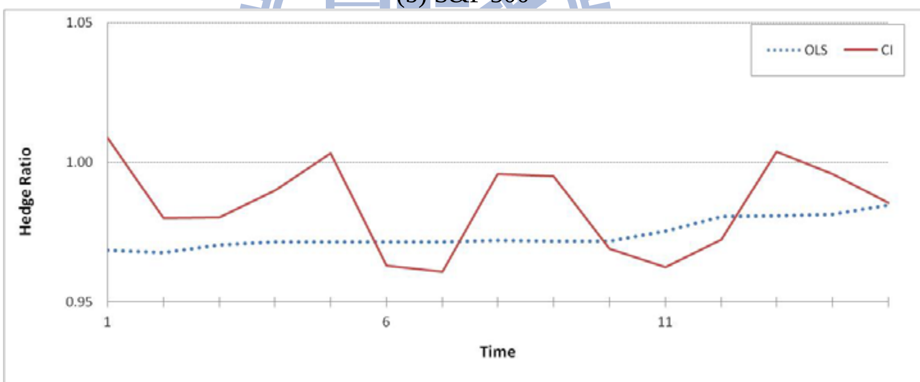
**Figure 4-3. Comparison of OHR in 7 days hedging**



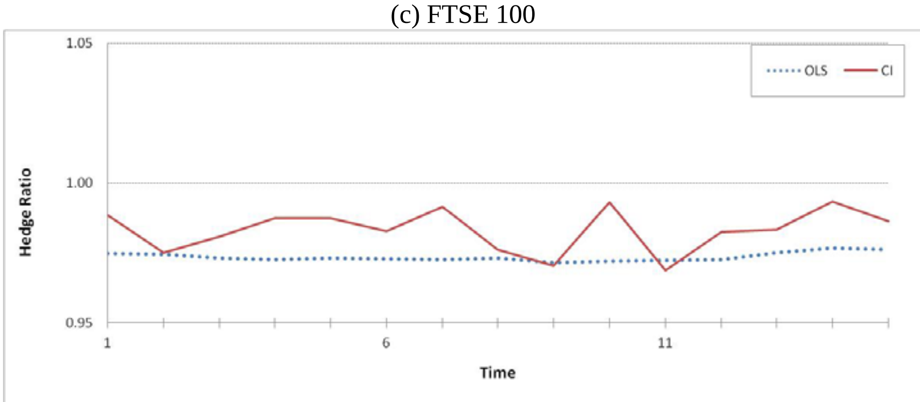

(a) TWI




(b) S&P 500

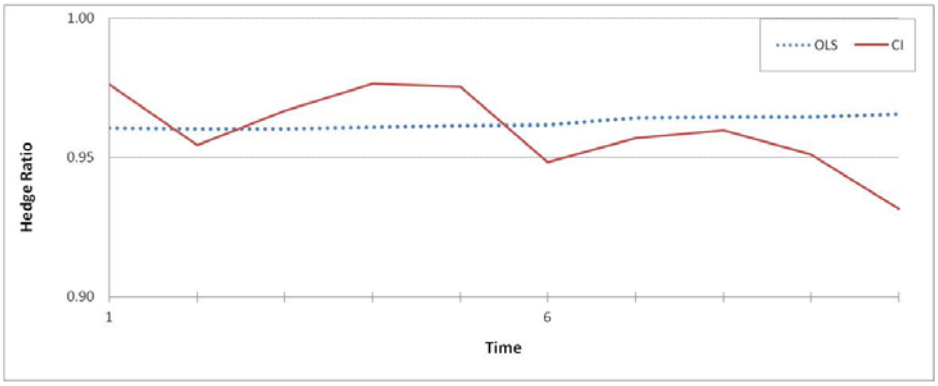



(c) FTSE 100

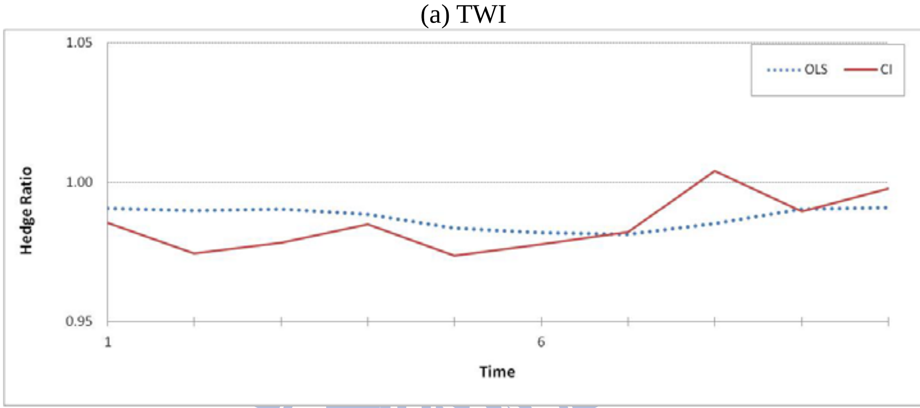



(d) NIKKEI 255

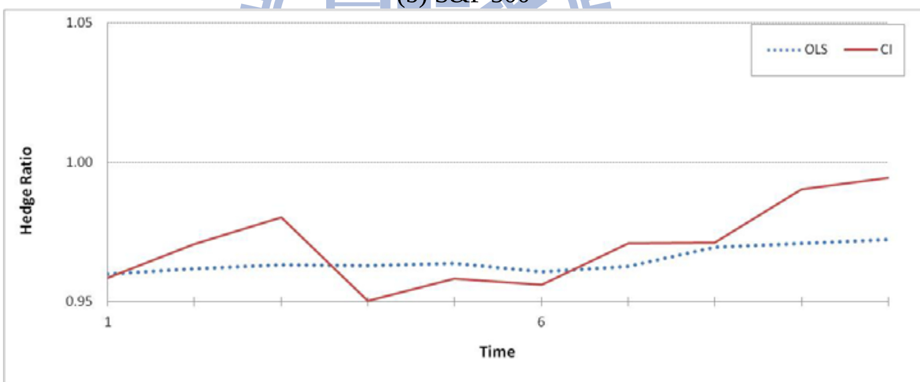
**Figure 4-4. Comparison of OHR in 14 days hedging**



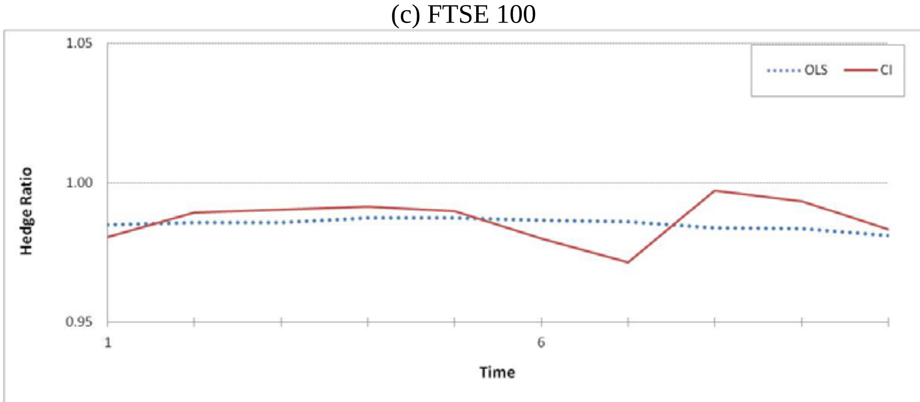

(a) TWI




(b) S&P 500

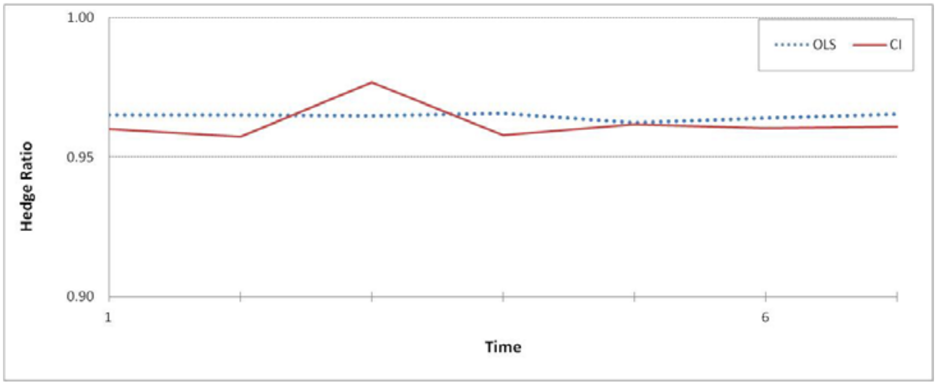



(c) FTSE 100

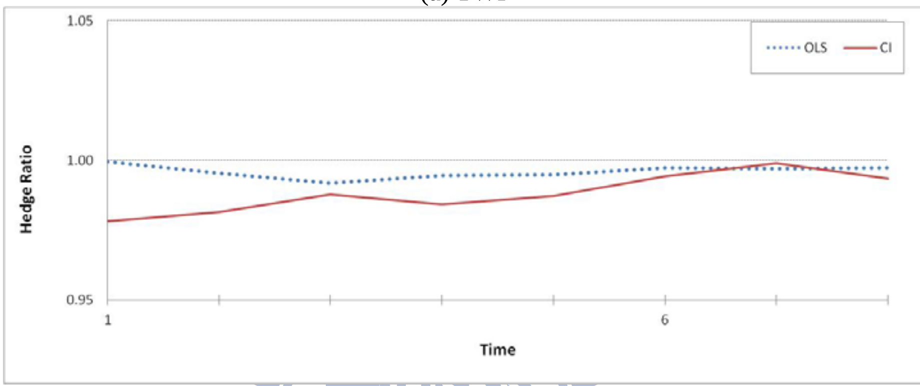



(d) NIKKEI 255

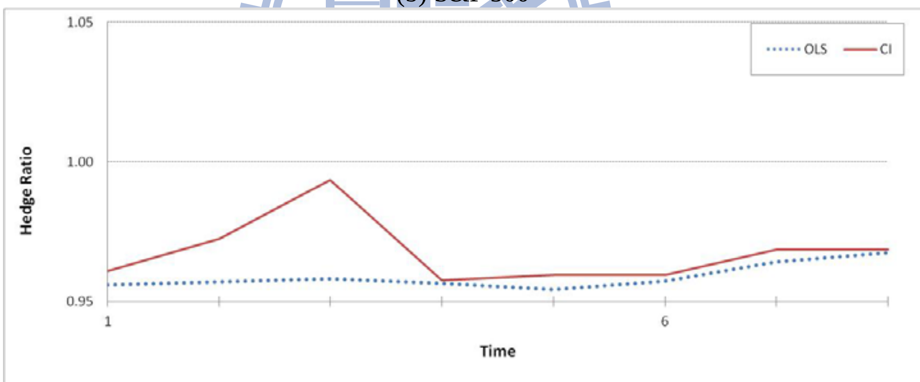
**Figure 4-5. Comparison of OHR in 21 days hedging**



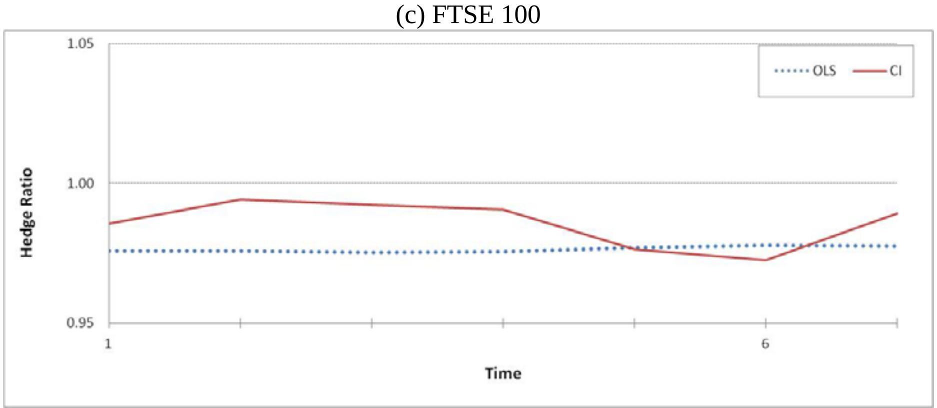

(a) TWI




(b) S&P 500




(c) FTSE 100




(d) NIKKEI 255

**Figure 4-6. Comparison of OHR in 28 days hedging**

## 4.6 Influence of the Period of Data for Feature Extraction

In this study, the proposed CI-based model adopts the period of one hedging horizon to extract the feature of the time series. Any feature of observation is represented by calculating variance, covariance, price spread, and the first and second order differences in the observations in the duration from one previous hedging horizon until present. However, the period length may influence the hedging performance of the proposed model.

Considering different hedging horizons, the proposed CI-based models, CI_1, CI_2, CI_3, CI_4, CI_5 and CI_6, are tested using different periods of data for feature extraction. The period of data for feature extraction is designed for a multiple of hedge horizons, from one to four in this study. The experiment is performed based on TWI data and compared with the conventional OLS model. Figures 4-7 to 4-11 show the experiment results in different hedging horizons. Figures 4-7 and 4-8 show that for CI_1 and CI_2, increasing the period of feature extraction will decrease hedging effectiveness. Figure 4-11, on the other hand, shows that for CI_1 and CI_2, increasing the period of feature extraction will increase hedging effectiveness. The influence of period of feature extraction is not obvious for other CI-based models, reflecting a tendency for inconsistency. The experiment results illustrate that a longer period for calculating variance and its first and second order differences is beneficial for hedging effectiveness in long-term hedging, but a shorter period is beneficial in short-term hedging.

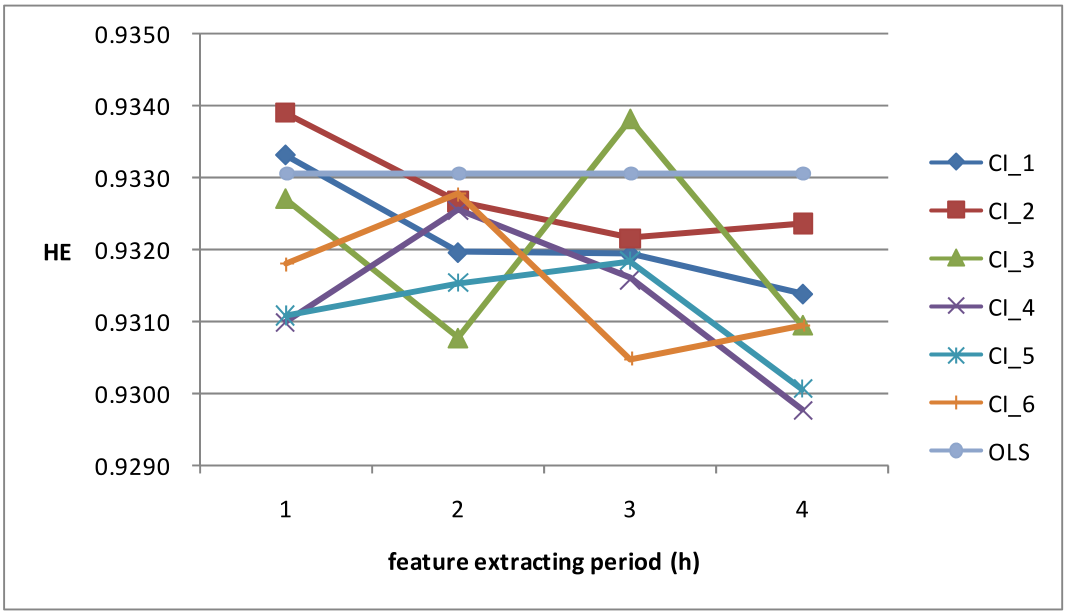



**Figure 4-7. Period of data for feature extraction in 1 day hedging**

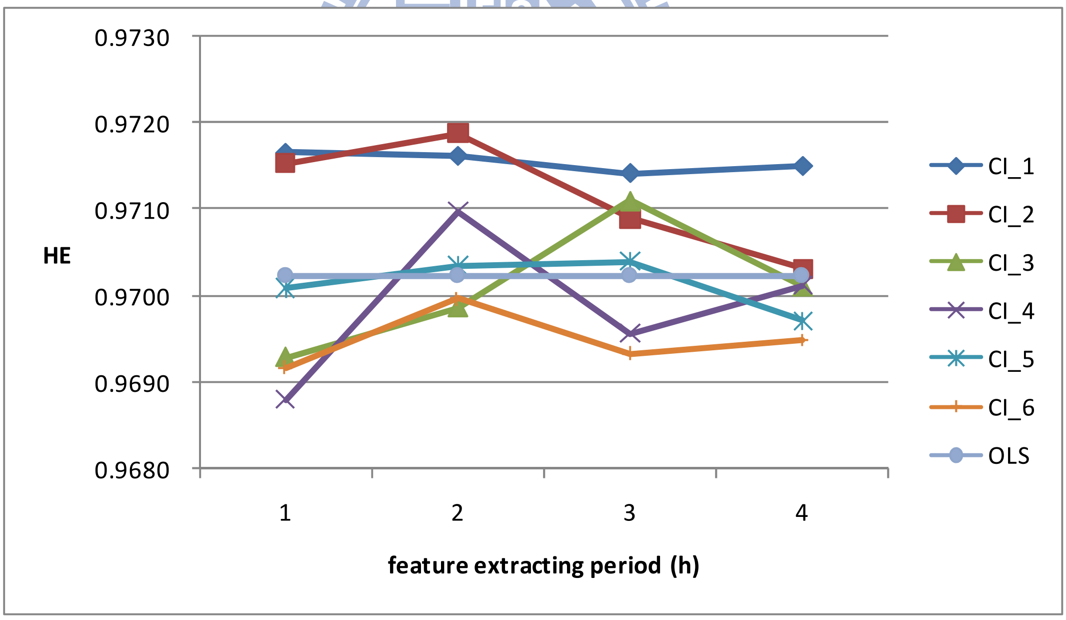



**Figure 4-8. Period of data for feature extraction in 7 days hedging**

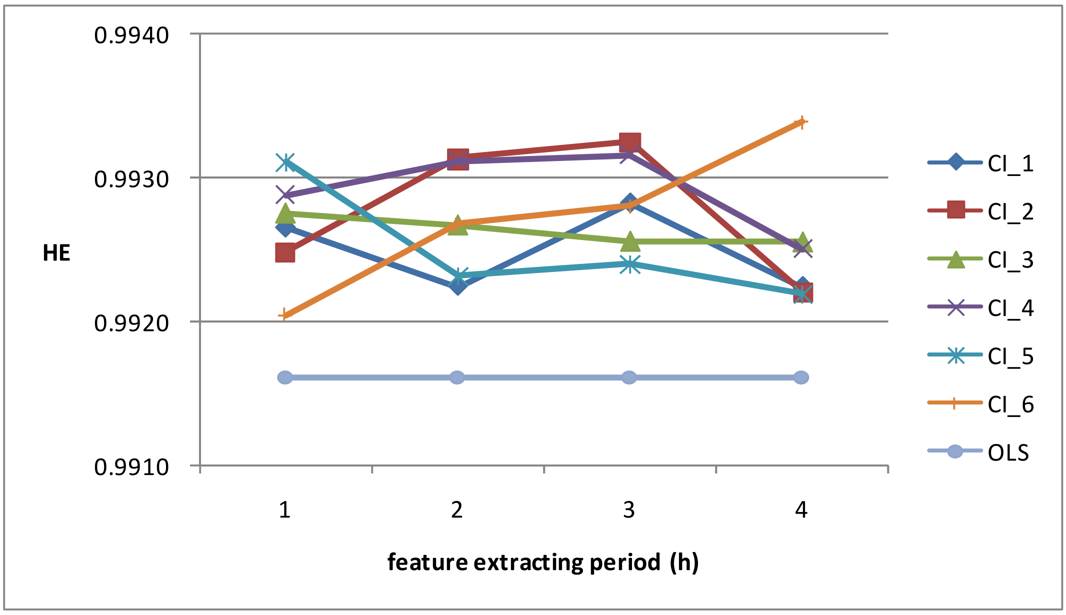



**Figure 4-9. Period of data for feature extraction in 14 days hedging**

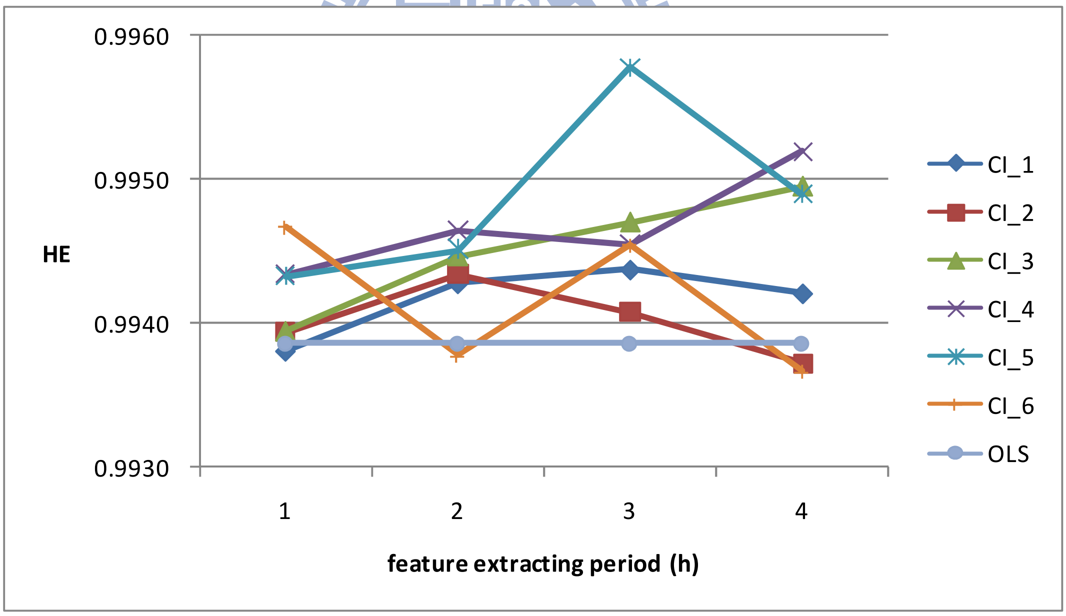



**Figure 4-10. Period of data for feature extraction in 21 days hedging**

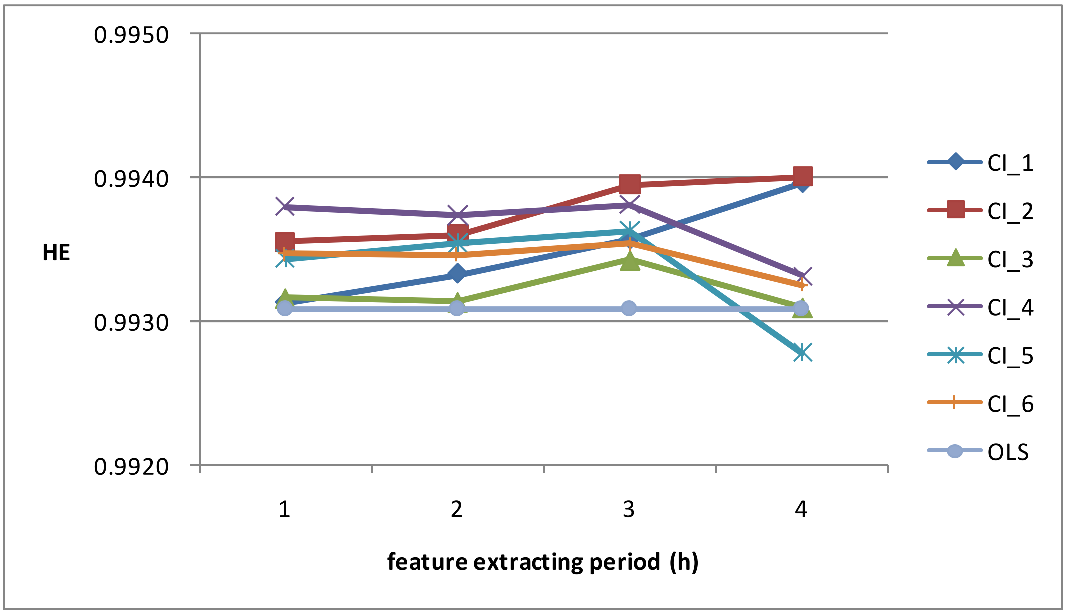


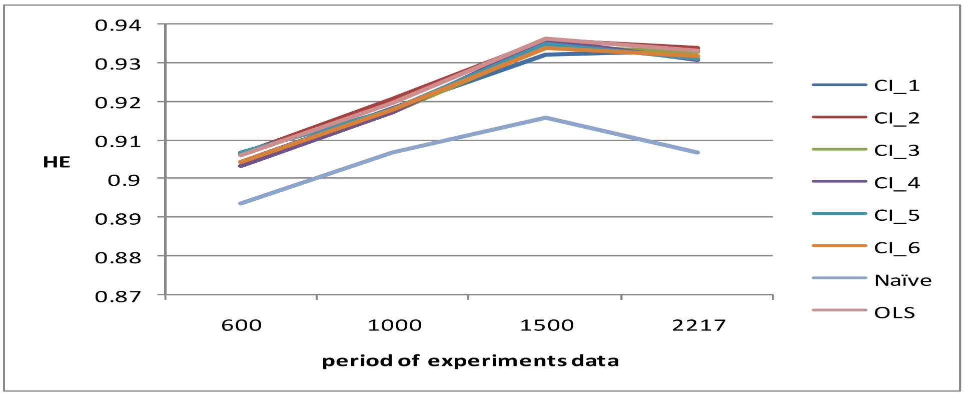

**Figure 4-11. Period of data for feature extraction in 28 days hedging**



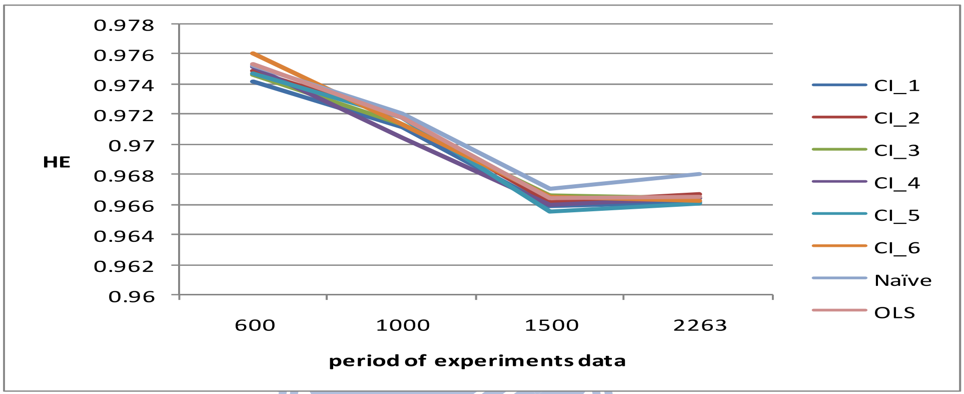
## 4.7 Influence of the Period of Data for Experiment

The dynamic behavior of a market is generally time-variant, which may influence the clustering results of the proposed model in this study. This problem can be classified to the robustness of model. In this study, we investigate the robustness of the proposed model using different periods of experiment data. The robustness of the six CI-based models are tested and compared with the two conventional models, naïve and OLS, over four periods. The different periods consist of the most recent 600, 1000, 1500, and all observations. The experiments of robustness are performed using five hedging horizons in four markets. A total of 128 experiments, with the settings composed of eight models, five hedging horizons, and four markets, are performed.

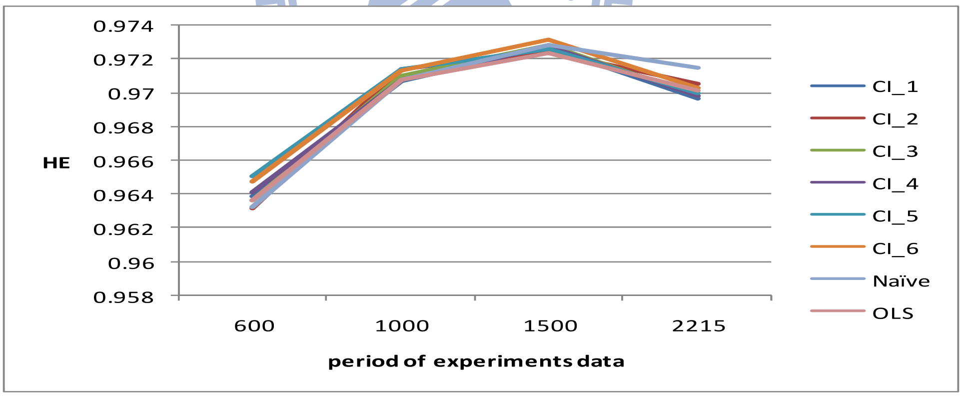
Figures 4-12 to 4-16 show the relationships between hedging effectiveness and period of experiment data for 1 day, 7 days, 14 days, 21 days, and 28 days hedging horizons, respectively. In these five figures, the relationship between hedging effectiveness and period of experiment data is neither an obviously positive correlation nor a negative one. However, it is obvious that the hedging effectiveness of these six CI-based models gets close to that of the OLS model, and they have similar change tendencies in different periods of experiment data. Table 4-7 summarizes the hedging effectiveness in these experiments and lists the mean and standard deviations of the four periods of experiment data. In Table 4-7, the standard deviation for most experiments are less than 1%, except for the experiments of 1 day and 7 days hedging in TWI. Standard deviation also decreases when hedging horizon increases. Based on the experiment results, the period of experiment data is not sensitive to model performance and the proposed CI-based model is robust.



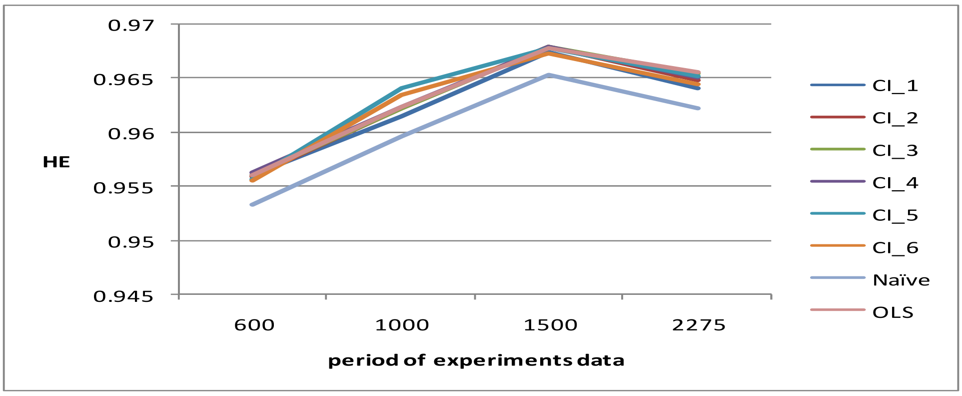

(a) TWI




(b) S&P 500

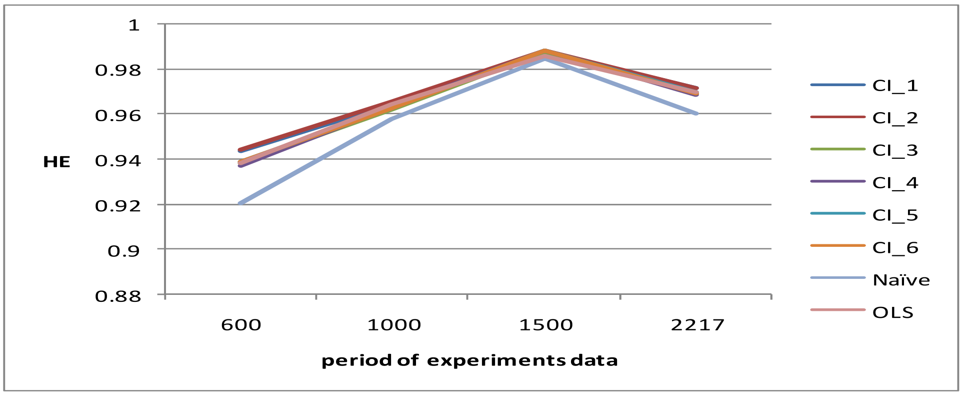



(c) FTSE 100

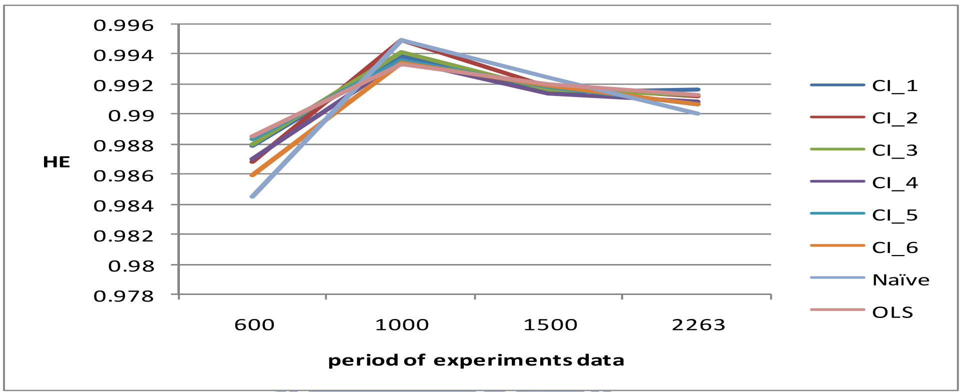



(d) NIKKEI 255

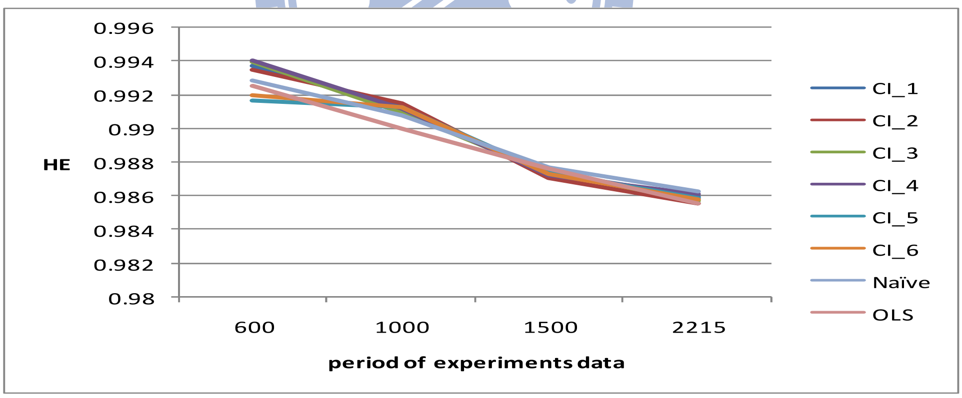
**Figure 4-12. Period of data for experiment in 1 day hedging**



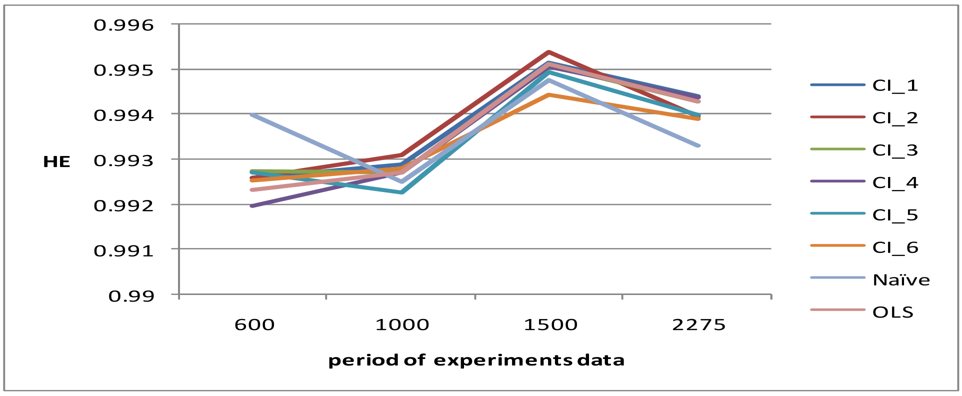

(a) TWI




(b) S&P 500

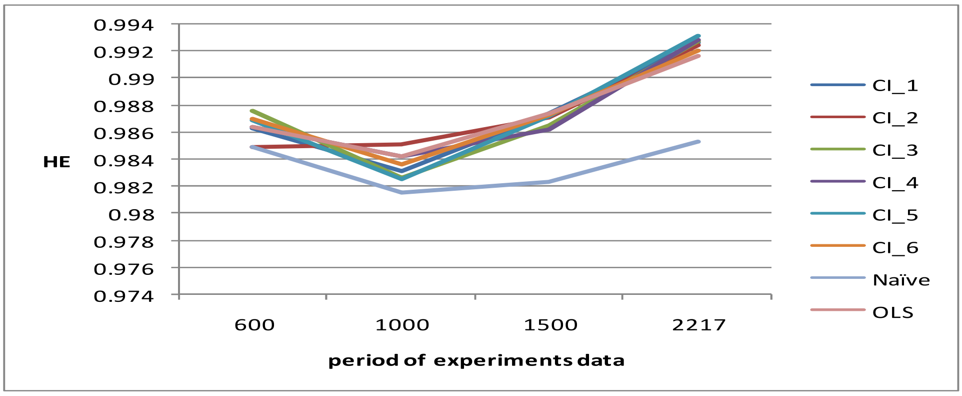



(c) FTSE 100

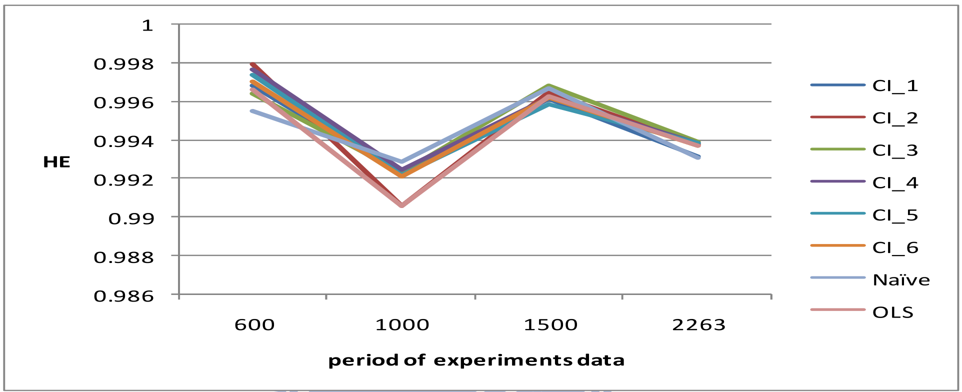



(d) NIKKEI 255

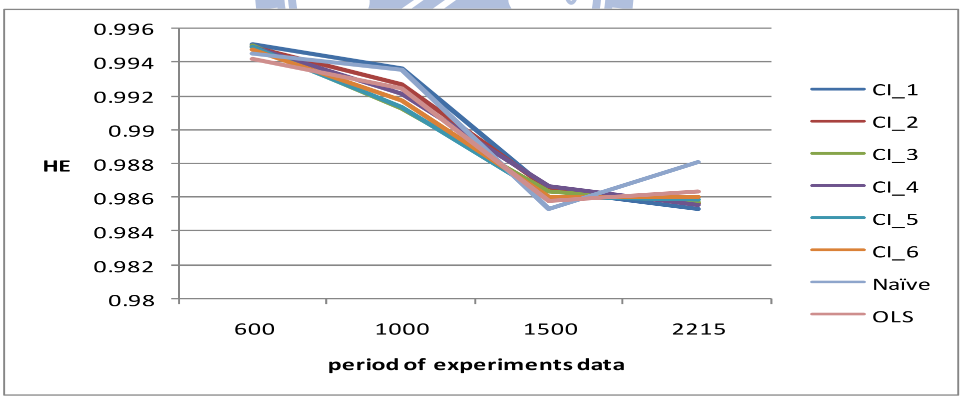
**Figure 4-13. Period of data for experiment in 7 days hedging**



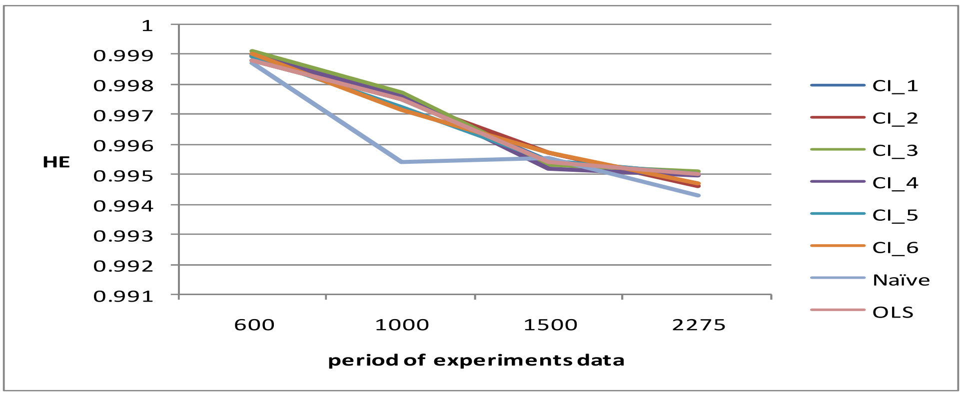

(a) TWI




(b) S&P 500

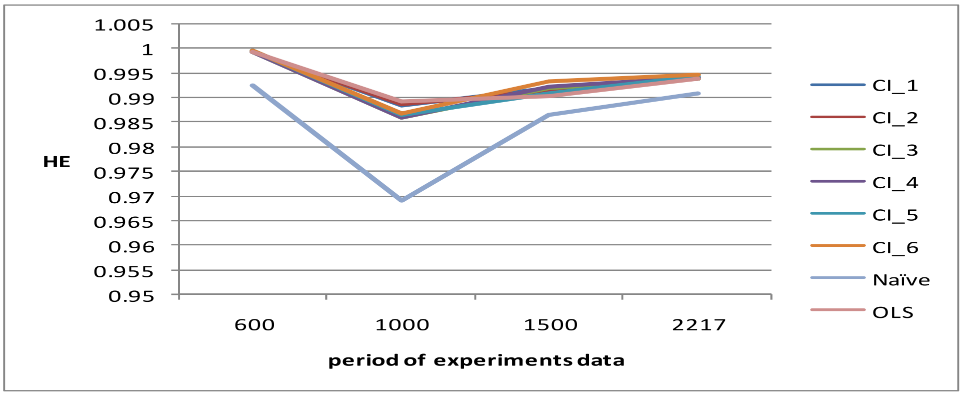



(c) FTSE 100

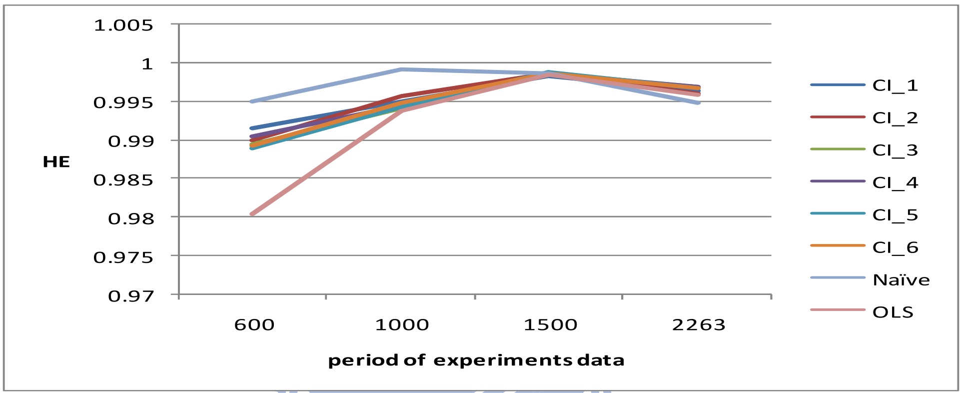



(d) NIKKEI 255

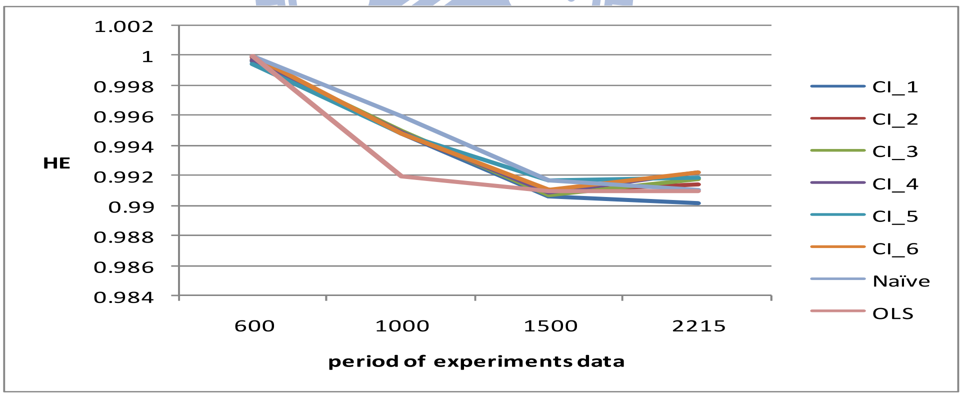
**Figure 4-14. Period of data for experiment in 14 days hedging**



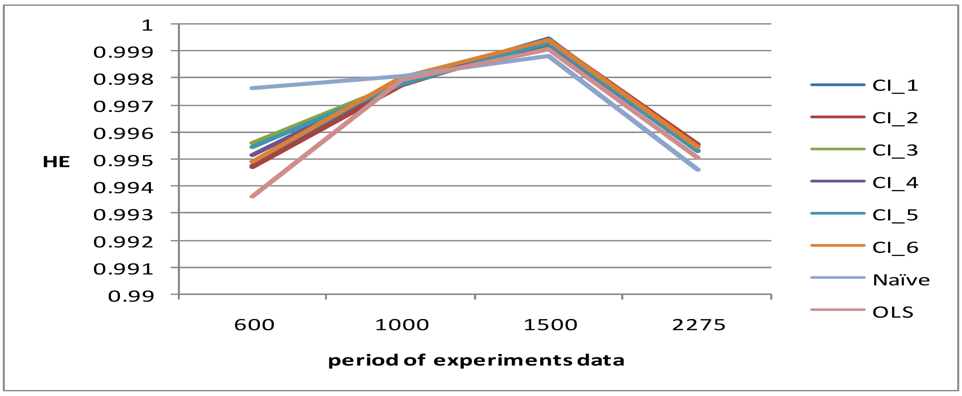

(a) TWI




(b) S&P 500

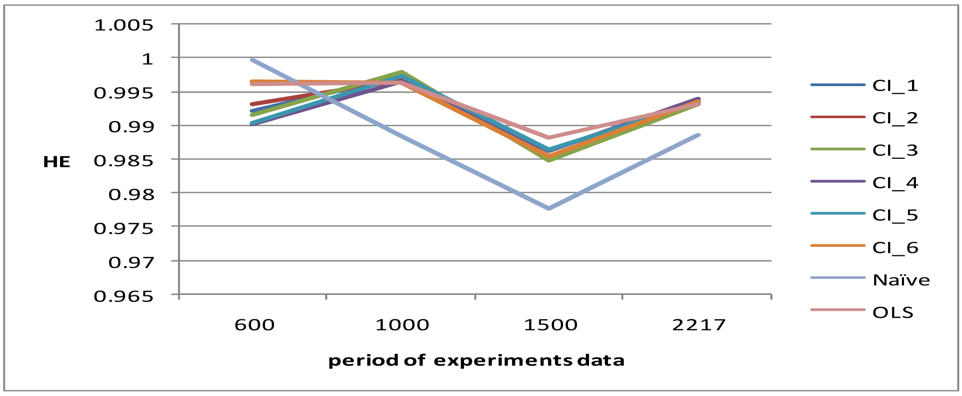



(c) FTSE 100

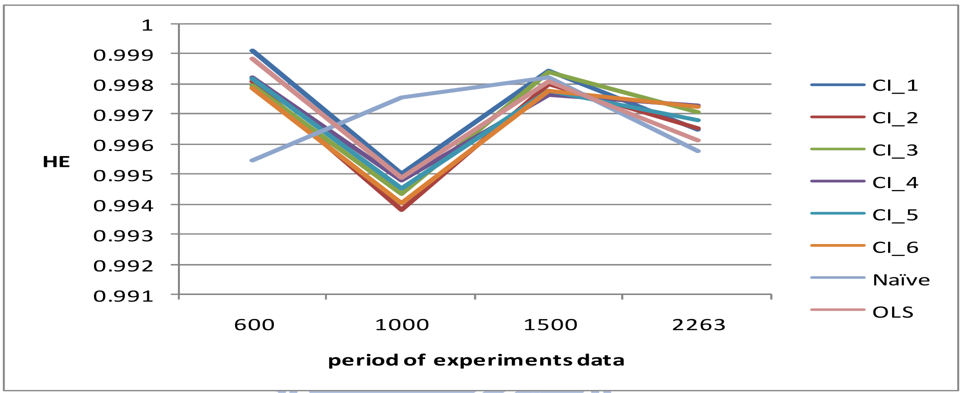



(d) NIKKEI 255

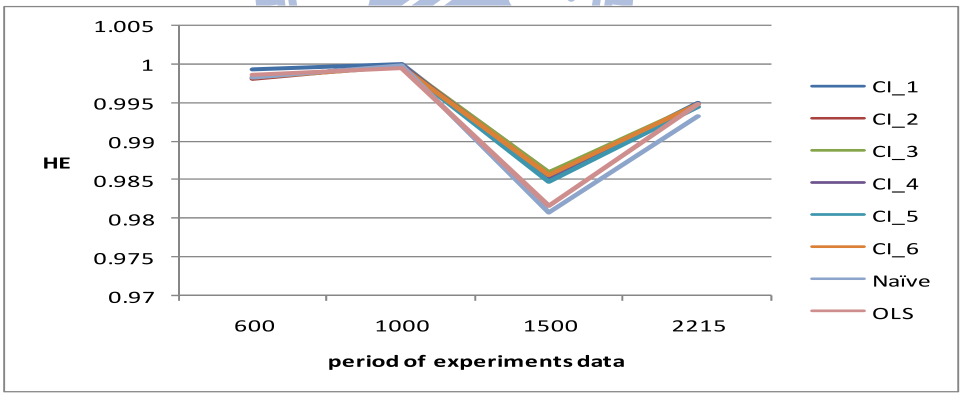
**Figure 4-15. Period of data for experiment in 21 days hedging**



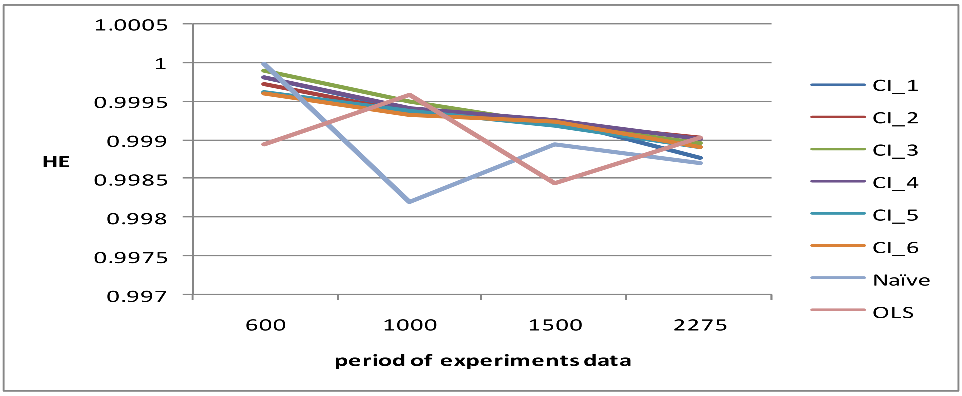

(a) TWI




(b) S&P 500




(c) FTSE 100




(d) NIKKEI 255

**Figure 4-16. Period of data for experiment in 28 days hedging**

**Table 4-7. Comparison of period of data for experiment**

| **Hedging horizon / Model** | **Hedging effectiveness** | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| **1 day** | **<u>Market</u>** | | | | | | | |
| | **TWI** | | **S&P 500** | | **FTSE 100** | | **NIKKEI 255** | |
| | **Mean** | **Std. Dev.** | **Mean** | **Std. Dev.** | **Mean** | **Std. Dev.** | **Mean** | **Std. Dev.** |
| **CI_1** | 92.2117% | 1.3529 | 96.9323% | 0.4030 | 96.9226% | 0.3768 | 96.2278% | 0.4786 |
| **CI_2** | 92.4148% | 1.3697 | 96.9887% | 0.4236 | 96.9373% | 0.4230 | 96.2738% | 0.5098 |
| **CI_3** | 92.1927% | 1.4484 | 96.9743% | 0.3998 | 96.9331% | 0.4015 | 96.2911% | 0.5151 |
| **CI_4** | 92.1968% | 1.4665 | 96.9533% | 0.4285 | 96.9382% | 0.3712 | 96.2974% | 0.4974 |
| **CI_5** | 92.2722% | 1.2920 | 96.9506% | 0.4460 | 96.9736% | 0.3315 | 96.3158% | 0.5231 |
| **CI_6** | 92.2056% | 1.3680 | 97.0034% | 0.4673 | 96.9867% | 0.3613 | 96.2734% | 0.5047 |
| **OLS** | 92.3812% | 1.3812 | 97.0026% | 0.4348 | 96.9238% | 0.3876 | 96.2918% | 0.5075 |
| **Naïve** | 90.5842% | 0.9266 | 97.0555% | 0.3800 | 96.9561% | 0.4297 | 96.0113% | 0.5148 |
| **7 days** | **<u>Market</u>** | | | | | | | |
| | **TWI** | | **S&P 500** | | **FTSE 100** | | **NIKKEI 255** | |
| | **Mean** | **Std. Dev.** | **Mean** | **Std. Dev.** | **Mean** | **Std. Dev.** | **Mean** | **Std. Dev.** |
| **CI_1** | 96.7073% | 1.8231 | 99.1244% | 0.2480 | 98.9544% | 0.3573 | 99.3760% | 0.1226 |
| **CI_2** | 96.7496% | 1.8202 | 99.1227% | 0.3337 | 98.9366% | 0.3699 | 99.3756% | 0.1216 |
| **CI_3** | 96.4375% | 2.0239 | 99.1229% | 0.2520 | 98.9472% | 0.3711 | 99.3708% | 0.1174 |
| **CI_4** | 96.4383% | 2.1028 | 99.0732% | 0.2744 | 98.9644% | 0.3631 | 99.3531% | 0.1424 |
| **CI_5** | 96.5048% | 2.0327 | 99.1113% | 0.2228 | 98.9056% | 0.2836 | 99.3479% | 0.1229 |
| **CI_6** | 96.4652% | 2.0426 | 99.0505% | 0.3229 | 98.9112% | 0.2994 | 99.3411% | 0.0899 |
| **OLS** | 96.4917% | 1.9803 | 99.1274% | 0.2020 | 98.8923% | 0.3014 | 99.3599% | 0.1312 |
| **Naïve** | 95.5952% | 2.6434 | 99.0494% | 0.4479 | 98.9388% | 0.2978 | 99.3642% | 0.0971 |
| **14 days** | **<u>Market</u>** | | | | | | | |
| | **TWI** | | **S&P 500** | | **FTSE 100** | | **NIKKEI 255** | |
| | **Mean** | **Std. Dev.** | **Mean** | **Std. Dev.** | **Mean** | **Std. Dev.** | **Mean** | **Std. Dev.** |
| **CI_1** | 98.7365% | 0.3965 | 99.4558% | 0.2268 | 99.0111% | 0.4963 | 99.6759% | 0.1856 |
| **CI_2** | 98.7389% | 0.3530 | 99.4714% | 0.3252 | 98.9922% | 0.4595 | 99.6646% | 0.1863 |
| **CI_3** | 98.7357% | 0.4181 | 99.4833% | 0.2140 | 98.9565% | 0.4385 | 99.6832% | 0.1929 |
| **CI_4** | 98.7439% | 0.3760 | 99.5046% | 0.2347 | 98.9850% | 0.4448 | 99.6697% | 0.1931 |
| **CI_5** | 98.7427% | 0.4333 | 99.4839% | 0.2290 | 98.9556% | 0.4409 | 99.6645% | 0.1796 |
| **CI_6** | 98.7454% | 0.3486 | 99.4751% | 0.2295 | 98.9651% | 0.4324 | 99.6658% | 0.1889 |
| **OLS** | 98.7389% | 0.3106 | 99.4300% | 0.2820 | 98.9691% | 0.4284 | 99.6689% | 0.1793 |
| **Naïve** | 98.3504% | 0.1911 | 99.4534% | 0.1890 | 99.0369% | 0.4418 | 99.6004% | 0.1883 |

| 21 days | Market | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | TWI | | S&P 500 | | FTSE 100 | | NIKKEI 255 | |
| | Mean | Std. Dev. | Mean | Std. Dev. | Mean | Std. Dev. | Mean | Std. Dev. |
| CI_1 | 99.3358% | 0.4658 | 99.5302% | 0.2834 | 99.3779% | 0.4439 | 99.6829% | 0.2209 |
| CI_2 | 99.3372% | 0.4615 | 99.5034% | 0.3679 | 99.4264% | 0.4122 | 99.6846% | 0.2142 |
| CI_3 | 99.2791% | 0.5583 | 99.4779% | 0.3958 | 99.4290% | 0.4044 | 99.7030% | 0.1802 |
| CI_4 | 99.3078% | 0.5553 | 99.5093% | 0.3535 | 99.4417% | 0.3878 | 99.6900% | 0.2022 |
| CI_5 | 99.2817% | 0.5497 | 99.4686% | 0.4249 | 99.4394% | 0.3648 | 99.6970% | 0.1933 |
| CI_6 | 99.3548% | 0.5247 | 99.4775% | 0.4061 | 99.4538% | 0.3964 | 99.6951% | 0.2146 |
| OLS | 99.3238% | 0.4541 | 99.2173% | 0.8086 | 99.3451% | 0.4382 | 99.6419% | 0.2517 |
| Naïve | 98.4771% | 1.0831 | 99.6859% | 0.2232 | 99.4630% | 0.4196 | 99.7275% | 0.1862 |

| 28 days | Market | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | TWI | | S&P 500 | | FTSE 100 | | NIKKEI 255 | |
| | Mean | Std. Dev. | Mean | Std. Dev. | Mean | Std. Dev. | Mean | Std. Dev. |
| CI_1 | 99.2280% | 0.4503 | 99.7272% | 0.1871 | 99.4817% | 0.6783 | 99.9301% | 0.0430 |
| CI_2 | 99.2258% | 0.4833 | 99.6604% | 0.2005 | 99.4651% | 0.6198 | 99.9329% | 0.0296 |
| CI_3 | 99.1847% | 0.5388 | 99.6953% | 0.1827 | 99.4760% | 0.6213 | 99.9387% | 0.0398 |
| CI_4 | 99.1613% | 0.4496 | 99.6990% | 0.1505 | 99.4579% | 0.6449 | 99.9379% | 0.0334 |
| CI_5 | 99.1885% | 0.4577 | 99.6817% | 0.1626 | 99.4280% | 0.6766 | 99.9272% | 0.0300 |
| CI_6 | 99.2907% | 0.5135 | 99.6739% | 0.1821 | 99.4609% | 0.6310 | 99.9273% | 0.0286 |
| OLS | 99.3386% | 0.3793 | 99.6997% | 0.1812 | 99.3655% | 0.8257 | 99.8999% | 0.0465 |
| Naïve | 98.8472% | 0.8987 | 99.6743% | 0.1353 | 99.3010% | 0.8634 | 99.8964% | 0.0751 |

# Chapter 5 Discussion and Conclusions

## 5.1 Discussion

The goal of this study is to investigate hedge ratio and hedging horizon using a novel CI approach and to compare the results obtained with the conventional OLS model. Many factors may influence the performance of the proposed CI-based model. These factors are widely discussed in other computational intelligence techniques, such as neural network and genetic algorithm, but are not extensively studied in this research. Moreover, empirical data also influence the performance of the proposed model, leading to the problem of model selection and robustness. Remarks on the proposed model and results presented in previous chapters are to be set forth and discussed.

First, the parameter setting in the proposed CI approach may influence the results slightly. In GARCH family models, the auto-correlation and partial correlation functions can be used to determine the lag of data for model construction. The fitting of model can be evaluated using the Akaike Information Criterion and Bayesian Information Criterion. In the proposed model, hedging horizon is adopted as the lag of data in the CI-based model. The feature of dynamic behavior is extracted using whole observations in the previous hedging horizon. Moreover, the hierarchical topology of the cluster may be affected by the settings for breadth and depth parameters in GHSOM. In this study, these parameter settings are determined by taking a systematic trial-and-error approach, which is commonly used in neural network and genetic algorithms.

Second, the lengths of data are important factors, including the length of estimation period, the length of testing period, and the period of empirical data. These factors will influence the stability of the results. The length of estimation period, which represents the size of observations and the length of rolling window, is different from previous research on conventional OLS and GARCH family models. Some studies adopt 900 of 1000 observations to estimate OHR (e.g., Yang and Lai, 2009) while other studies adopt 1500 observations (e.g., Lien *et al.*, 2002; Li, 2010). In this study, we initially adopt about 2000 observations to estimate the OHR, and then about 200 observations to test its performance. In the following estimations, 600, 1000, and 1500 observations are used to test the robustness of the model. Although the experiment results imply that the length of data does not obviously influence model performance, the examinations in this study do not provide statistical verification and thus require further study.

Finally, the results of model comparisons may differ in different markets. Some studies indicate GARCH family models to be superior to the OLS model in a specific market (e.g., Myers, 2000). However, other studies indicate opposing opinions, stating that the OLS hedge ratio performs better than other popular multivariate GARCH models (e.g., Lien *et al.*, 2002; Moon *et al.*, 2009). The naïve hedge ratio of 1 is suggested as the optimal hedge ratio when the hedging horizon is long (Chen *et al.*, 2004). The superiority of the hedging model can be evaluated using White's reality check. However, this evaluation is not significant for model comparisons in one day hedging (Bystrom, 2003; Lee and Yoder, 2007). This phenomenon may be due to the dissimilar behavior of markets: the behavior of an emerging market differs from a mature market. For example, hedging effectiveness can be enhanced by a certain model in emerging markets such as the Hungarian BSI market, but not for developed markets such as the US S&P 500 market (Li, 2010). A similar result can be observed in this study, that is, the hedging effectiveness in TWI is different from that of UK FTSE 100. However, the proposed CI-based model is capable of discovering similar behavior in the same market and can adapt to the characteristics of a particular market. Therefore, the long-term tendency of markets can be captured easily and the statistical significance when compared with OLS model in this study can be obtained.

## 5.2 Conclusions

In this study, we propose a novel computational intelligence approach to estimate the time-varying minimum variance hedge ratio. This investigation is one of the first studies on the computational intelligence techniques of cluster analysis to dynamic hedge ratio on stock index futures. Clustering time series are employed to recognize the observations with similar time series patterns. Observations with a high possibility of occurrence in the future are selected when hedging. These observations are used to modify the distribution probability of time series data using a resampling process with different weights given based on cluster result. This novel CI approach can overcome sample reduction problems and avoid undue assumptions in typical models.

The empirical findings in this study are consistent with the following notations. First, hedging horizon will increase hedging effectiveness. When hedge horizon is increased, hedging effectiveness is also increased. Second, the proposed CI-based model can improve the typical OLS model, especially in long-term hedging. Third, the present findings lend support to the superiority of the CI-based model in enhancing hedging effectiveness for emerging markets, but not for developed markets such as the US S&P 500 and UK FTSE 100 markets. Finally, the OHR estimated using the CI-based model is more volatile than the OHR estimated using the OLS model, which implies that the CI-based model can rapidly reflect the time-variant property of financial time series and provide accurate estimation for dynamic hedging decision.

This study evidences that the proposed model is superior to the conventional OLS model in hedging effectiveness, but the usability of this computational approach is worse than conventional OLS model. Many factors, such as the breadth and depth parameters of GHSOM, the period of data for feature extraction, which may influence the performance of the proposed model, should to be determined appropriately in practical use. Further, most non-parametric models based on computational intelligence are challenged that the experiment results of non-parametric model are more unstable than parametric model when repeating the experiments.

Although this research still have some restriction of model parameters selection, this novel approach based on computational intelligence can improve the performance of traditional approach without too many inappropriate assumptions and restrictions. Consequently, the proposed model can also be considered as a powerful tool to investigate any financial market, in which the probability distribution of data is unrestricted and not necessary to fit any type of probability distribution.

## 5.3 Recommendations for Future Works

This study proposed a novel approach for clustering the time series data and using the similarity-clustered data for OHR estimation. The empirical results showed results with different hedge intervals of the proposed models, and compared these with the hedge intervals of traditional models. The findings, although significant, have some limitations and are expected to be investigated further. The recommendations for future works are summarized as follows.

1. This research simply adopts a default value to set the GHSOM parameter. Thus, the sensitivity of parameters setting of GHSOM should be investigated further to provide a definite guide in determining optimal parameter settings.

2. The robustness of the proposed model is expected to be verified using different periods of data from various markets.
3. This research only conducts model and OHR estimations on stock index futures. However, the model has the potential to be applied to other futures market, such as foreign exchange futures or commodity futures.
4. The proposed CI-based model is expected to be used as a tool for investigating the relevant issue of volatility in financial engineering, such as volatility forecasting, modifying beta coefficient in capital asset pricing model (CAPM), and estimating value of risk (VaR).

# References

Afolabi, M. O., & Olude, O. (2007). Predicting stock prices using a hybrid Kohonen self organizing map (SOM). *Proceedings of the 40th Annual Hawaii International Conference on System Sciences*, 8 pp.|CD-ROM.

Alizadeh, A., & Nomikos, N. (2004). A Markov regime switching approach for hedging stock indices. *Journal of Futures Markets, 24*(7), 649-674.

Andersen, T. G., & Sorensen, B. E. (1996). GMM estimation of a stochastic volatility model: A Monte Carlo study. *Journal of Business & Economic Statistics, 14*(3), 328-352.

Basalto, N., Bellotti, R., De Carlo, F., Facchi, P., Pantaleo, E., & Pascazio, S. (2007). Hausdorff clustering of financial time series. *Physica a-Statistical Mechanics and Its Applications, 379*(2), 635-644.

Bollerslev, T., Chou, R. Y., & Kroner, K. F. (1992). Arch modeling in finance - a review of the theory and empirical-evidence. *Journal of Econometrics, 52*(1-2), 5-59.

Bystrom, H. N. E. (2003). The hedging performance of electricity futures on the Nordic power exchange. *Applied Economics, 35*(1), 1-11.

Chan, A., & Panipalk, E. (2002). Growing Hierarchical Self Organising Map (GHSOM) toolbox: Visualisations and enhancements. *Iconip'02: Proceedings of the 9th International Conference on Neural Information Processing*, 2537-2541.

Chen, S. S., Lee, C. F., & Shrestha, K. (2004). An empirical analysis of the relationship between the hedge ratio and hedging horizon: A simultaneous estimation of the short- and long-run hedge ratios. *Journal of Futures Markets, 24*(4), 359-386.

Choudhry, T. (2003). Short-run deviations and optimal hedge ratio: evidence from stock futures. *Journal of Multinational Financial Management*, 13, 171–192.

D'Urso, P., & Maharaj, E. A. (2009). Autocorrelation-based fuzzy clustering of time series. *Fuzzy Sets and Systems, 160*(24), 3565-3589.

Ederington, L. H. (1979). Hedging performance of the new futures market. *Journal of Finance, 34*(1), 157-170.

Engle, R. (2002). Dynamic conditional correlation: A simple class of multivariate generalized autoregressive conditional heteroskedasticity models. *Journal of Business & Economic Statistics, 20*(3), 339-350.

Focardi, S. M., & Fabozzi, F. J. (2004). A methodology for index tracking based on time-series clustering. *Quantitative Finance, 4*(4), 417-425.

Fruehwirth-Schnatter, S., & Kaufmann, S. (2008). Model-based clustering of multiple time series. *Journal of Business & Economic Statistics, 26*(1), 78-89.

Gafiychuk, V. V., Datsko, B. Y., & Izmaylova, J. (2004). Analysis of data clusters obtained by self-organizing methods. *Physica a-Statistical Mechanics and Its Applications, 341*, 547-555.

Geppert, J. M. (1995). A statistical-model for the relationship between futures contract hedging effectiveness and investment horizon length. *Journal of Futures Markets, 15*(5), 507-536.

Gershenfeld, N., Schoner, B., & Metois, E. (1999). Cluster-weighted modelling for time-series analysis. *Nature, 397*(6717), 329-332.

Ghysels, E., Santa-Clara, P., & Valkanov, R. (2006). Predicting volatility: getting the most out of return data sampled at different frequencies. *Journal of Econometrics, 131*(1-2), 59-95.

Hill, J., & Schneeweis, T. (1981). A note on the hedging effectiveness of foreign-currency futures. *Journal of Futures Markets, 1*(4), 659-664.

Hoffman, E. B., Sen, P. K., & Weinberg, C. R. (2001). Within-cluster resampling. *Biometrika, 88*(4), 1121-1134.

Howard, C. T., & Dantonio, L. J. (1991). Multiperiod hedging using futures - a risk minimization approach in the presence of autocorrelation. *Journal of Futures Markets, 11*(6), 697-710.

In, F., & Kim, S. (2006). Multiscale hedge ratio between the Australian stock and futures markets: Evidence from wavelet analysis. *Journal of Multinational Financial Management*, 16, 411–423.

Jain, A. K. (2010). Data clustering: 50 years beyond K-means. *Pattern Recognition Letters, 31*(8), 651-666.

Karandikar, R. G., Deshpande, N. R., Khaparde, S. A., & Kulkarni, S. V. (2009). Modelling volatility clustering in electricity price return series for forecasting value at risk. *European Transactions on Electrical Power*, *19*(1), 15-38.

Kohonen, T. (1989). Self-organization and Associative Memory, 3rd ed., Springer, New York.

Kroner, K.F., & Sultan, J. (1993). Time-varying distribution and dynamic hedging with foreign currency futures. *Journal of Financial and Quantitative Analysis,* 28, 535–551.

Lee, H. T., & Yoder, J. K. (2007). A bivariate Markov regime switching GARCH approach to estimate time varying minimum variance hedge ratios. *Applied Economics*, *39*(10-12), 1253-1265.

Li, M. Y. L. (2010). Dynamic hedge ratio for stock index futures: application of threshold VECM. *Applied Economics, 42*(11), 1403-1417.

Liao, T. W. (2005). Clustering of time series data - a survey. *Pattern Recognition, 38*(11), 1857-1874.

Lien, D., & B. K. Wilson. (2001). Multiperiod hedging in the presence of stochastic volatility. *International Review of Financial Analysis*, 10, 395–406.

Lien, D., & Luo, X. D. (1993). Estimating multiperiod hedge ratios in cointegrated markets. *Journal of Futures Markets*, *13*(8), 909-920.

Lien, D., & Luo, X. D. (1994). Multiperiod hedging in the presence of conditional heteroskedasticity. *Journal of Futures Markets, 14*(8), 927-955.

Lien, D., & Shrestha, K. (2007). An empirical analysis of the relationship between hedge ratio and hedging horizon using wavelet analysis. *Journal of Futures Markets, 27*(2), 127-150.

Lien, D., Tse, Y. K., & Tsui, A. K. C. (2002). Evaluating the hedging performance of the constant-correlation GARCH model. *Applied Financial Economics*, *12*(10), 791-798.

Lien, D., & Yang, L. (2006). Spot-futures spread, time-varying correlation, and hedging with currency futures. *Journal of Futures Markets*, *26*(10), 1019-1038.

Liu, S., Lu, L., Liao, G., & Xuan, J. (2006). Pattern discovery from time series using growing hierarchical self-organizing map. *Neural Information Processing. 13th International Conference, ICONIP 2006. Proceedings, Part I (Lecture Notes in Computer Science Vol. 4232)*, 1030-1037|xlvi+1153.

Liu, S., & Lv, L. (2007). Clustering method of time series based on growing hierarchical self-organizing map. *Computer Engineering, 33*(6), 208-210.

Mandelbrot, B. B., (1963). The variation of certain speculative prices. *Journal of Business*, XXXVI , 392–417.

Matsumoto, H., Masumoto, R., & Kuroda, C. (2009). Feature extraction of time-series process images in an aerated agitation vessel using self organizing map. [Proceedings Paper]. *Neurocomputing, 73*(1-3), 60-70.

Moon, G. H., Yu, W. C., & Hong, C. H. (2009). Dynamic hedging performance with the evaluation of multivariate GARCH models: evidence from KOSTAR index futures. *Applied Economics Letters, 16*(9), 913-919.

Moschini, G., and R. J. Myers. (2002). Testing for constant hedge ratios in commodity markets:A multivariate GARCH approach. *Journal of Empirical Finance,* 9, 589–603.

Myers, R.J. (2000). Estimating time-varying optimal hedge ratios on futures markets. *Journal of Futures Markets, 20*(1), 73-87.

Nelson, D. B., & Foster, D. P. (1994). Asymptotic filtering theory for univariate arch models. *Econometrica, 62*(1), 1-41.

Pattarin, F., Paterlini, S., & Minerva, T. (2004). Clustering financial time series: an application to mutual funds style analysis. *Computational Statistics & Data Analysis, 47*(2), 353-372.

Rauber, A., Merkl, D., & Dittenbach, M. (2002). The growing hierarchical self-organizing map: Exploratory analysis of high-dimensional data. *IEEE Transactions on Neural Networks, 13*(6), 1331-1341.

Rieger, R. H., & Weinberg, C. R. (2002). Analysis of clustered binary outcomes using within-cluster paired resampling. *Biometrics, 58*(2), 332-341.

Sener, T. (1998). Objectives of hedging and optimal hedge ratios: US vs. Japanese investors. *Journal of Multinational Financial Management,* 8, 137–153.

Senjyu, T., Tamaki, Y., & Uezato, K. (2000). Next day load curve forecasting using self organizing map. *PowerCon 2000. 2000 International Conference on Power System Technology. Proceedings (Cat. No.00EX409)|PowerCon 2000. 2000 International Conference on Power System Technology. Proceedings (Cat. No.00EX409),* 10.1109/ICPST.2000.897176.

Shih, J. Y., Chang, Y. J., & Chen, W. H. (2008). Using GHSOM to construct legal maps for Taiwan's securities and futures markets. *Expert Systems with Applications, 34*(2), 850-858.

Simon, G., Lendasse, A., Cottrell, M., Fort, J. C., & Verleysen, M. (2005). Time series forecasting: Obtaining long term trends with self-organizing maps. *Pattern Recognition Letters, 26*(12), 1795-1808.

Tay, F. E. H., & Li Juan, C. (2001). Improved financial time series forecasting by combining support vector machines with self-organizing feature map. *Intelligent Data Analysis, 5*(4), 339-354.

Tsao, C.-Y., & Chen, S.-H. (2003). Self-organizing maps as a foundation for charting or geometric pattern recognition in financial time series. *2002 IEEE International Conference on Computational Intelligence for Financial Engineering. Proceedings (Cat. No.03TH8653)*, 387-394|x+436.

Wang, C., Low, S.S. (2003). Hedging with foreign currency denominated stock index futures: evidence from the MSCI Taiwan index futures market. *Journal of Multinational Financial Management*, 13, 1–17.

White, H. (2000). A reality check for data snooping, *Econometrica, 68*, 1097-126.

Yang, H. C., Lee, C. H., & Chen, D. W. (2009). A method for multilingual text mining and retrieval using growing hierarchical self-organizing maps. *Journal of Information Science, 35*(1), 3-23.

Yang, M. J. & Lai, Y. (2009). An out-of-sample comparative analysis of hedging performance of stock index futures: Dynamic versus static hedging. *Applied Financial Economics, 19*(13), 1059-1072.

Zhu, J. R., & Ieee. (2008). A new model for oil futures price forecasting based on cluster analysis. *2008 4th International Conference on Wireless Communications, Networking and Mobile Computing, Vols 1-31* (pp. 11456-11459).