Noname manuscript No.
(will be inserted by the editor)

A novel multivariate volatility modeling for risk management to stock markets

Zi-Kai Wei · Ka-Fai Cedric Yiu · Heung Wong · Kit-Yan Chan

Abstract Volatility modeling is crucial for risk management and asset allocation; this is an influential area in financial econometrics. The central requirement of volatility modeling is to be able to forecast volatility accurately. The literature review of volatility modeling shows that the approaches of model averaging estimation are commonly used to reduce model uncertainty in order to achieve a satisfactory forecasting reliability. However, those approaches attempt to forecast more reliable volatilities by integrating all forecasting outcomes equally from several volatility models. Forecasting patterns generated by each model may be similar. This may cause redundant computation without improving forecasting reliability. The proposed multivariate volatility modeling method which is called the Fuzzy-method-involving Multivariate Volatility Model (abbreviated as FMVM) classifies the individual models into smaller scale clusters and selects the most representative model in each cluster. Hence, repetitive but unnecessary computational burden can be reduced, and forecasting patterns from representative models can be integrated. The proposed FMVM is benchmarked against existing multivariate volatility models on forecasting volatilities of Hong Kong Hang Seng Index (HSI) constituent stocks. Numerical results show that it can obtain relatively lower forecasting errors with less model complexity.

Keywords: Multivariate Volatility Models; Risk Management to Future Markets; Generalized Autoregressive Conditional Heteroscedastic Modelling; Model Averaging Techniques.

Zi-Kai Wei
Department of Applied Mathematics, The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong, PR China
E-mail: nathaniel.wei@connect.polyu.hk

Ka-Fai Cedric Yiu
Department of Applied Mathematics, The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong, PR China
Tel.: +852-34008981
Fax: +852-23629045
E-mail: cedric.yiu@polyu.edu.hk

Heung Wong
Department of Applied Mathematics, The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong, PR China
E-mail: heung.wong@polyu.edu.hk

Kit-Yan Chan
Department of Electrical and Computer Engineering, Curtin University, Perth, Australia
E-mail: Kit.Chan@curtin.edu.au

1 Introduction

The global financial crisis of 2008 has led investors to reassess the forecasting adequacy of financial models against soaring volatilities. The use of volatility models in quantitative risk management has gained increasing importance among academics and practitioners concerned with measuring and managing financial risks [1–4]. There are also important applications in insurance [5,6] and supply chains management [7,8]. These models can be used to forecast volatility in order to assist investors in making financial decisions. Generally, these models can be classified into univariate volatility models and multivariate ones. The most widely used univariate model namely generalized autoregressive conditional heteroscedastic (GARCH) model was developed in [9].

The multivariate volatility models attempt to specify the dynamic process of diagonal elements of the volatility matrix or variance-covariance matrix. The distinction of each multivariate volatility model derives from the differences between their specifications of the conditional correlation processes. In the literature, diverse multivariate GARCH models have been developed, including BEKK GARCH model [10], the constant conditional correlation (CCC) model [11], the orthogonal GARCH (OGARCH) model [12], the dynamic conditional correlation model (DCC) [13], and the asymmetric dynamic conditional correlation model (ADCC) [14].

In order to improve the forecasting accuracy and moderate model uncertainty in multivariate volatility models, the model averaging strategy deploying multi-models namely the average multivariate volatility model (AMVM) [15] has been developed. The AMVM generates a multivariate volatility model by integrating the forecasting outcomes of all models equally. However, using forecasting outcomes from all selected multivariate models may be not the most effective, as some models may generate similar forecasting patterns with the others. Repetitive but unnecessary computation is used to analyze the similar forecasting patterns. The approach increases computational burden of the multivariate volatility models however the approach does not improve forecasting accuracy.

To avoid redundant computation, multivariate volatility models with similar forecasting patterns can be clustered in a group and only the most significant model in each group is integrated into the final volatility model. Optimal weights for all the groups are determined in order to reduce forecasting errors. These processes can be performed by fuzzy classification operations. Indeed, fuzzy classification operations have been developed successfully in many practical classification systems [16, 17]. They are also widely used in finance [18, 19]. The fuzzy c-means (FCM) clustering algorithm is applicable to a wide variety of numerical data and it is also accessible to generate fuzzy partitions for any set of numerical data; also, the FCM clustering introduces functions namely membership functions to describe the similarity a data point shares with each cluster [20]. An integrated approach that uses the FCM, mixture models, and the collaborative clustering framework to classify the mixed data which contains both numerical and categorical attributes. This novel clustering framework not only uses the FCM as a component of it but also highlights that the FCM is a very effective and popular algorithm for numerical data [21]. A new segmentation approach named I-Ching spatial shadowed FCM is very efficient not only in tackling the overlapping segments but also in suppressing the noise in images. This method is also founded on the FCM with some reversion of the FCM: the fuzzy set in the FCM is replaced by the shadowed set and I-Ching operators are used to find the optimal cluster centers of the shadowed FCM [22]. Fuzzy methods are not only widely used in clustering but also in solving programming and transportation problem. An interactive fuzzy goal programming algorithm is proposed to solve decentralized bilevel multiobjective fractional programming problem, and it can make a fuzzy decision by taking most satisfactory solution for all decision makers at the both levels [23]. In solving a transportation problem, an interval programming model using the nearest interval approximation of trapezoidal fuzzy numbers is developed to obtain the optimal solution of the multiobjective multi-item solid transportation problem under uncertainty [24, 25]. The fuzzy methods are widely applied to different areas. In numerical data clustering, the FCM is very effective and common algorithm. Thus, our framework handling redundant computation is inspired by the FCM clustering initially.

In this paper, we propose a novel multivariate volatility models namely the Fuzzy-method-involving Multivariate Volatility Model (FMVM) to cluster multivariate volatility models with similar forecasting patterns. Here we propose the FMVM as this is effective on clustering similar patterns for forecasting stock exchanges which are involved with uncertainty [19]. In order to evaluate the performance of the forecasted volatility matrix, the Frobenius norm is used to evaluate forecasting errors of different multivariate volatility models relative to the realized covariance matrix [26]. First, the FMVM classifies all the individual models into smaller scale clusters by using the fuzzy C-means clustering algorithm where the optimal number of clusters is given. Hence, models which generate similar forecasting patterns can be grouped into a cluster. It then selects the model with the lowest tracking error from each cluster. Subsequently, it determines the optimal weight for each selected model. As only significant forecasting patterns are used by the proposed multivariate volatility model, repetitive and unnecessary computational burden can be reduced in the final multivariate volatility model. The effectiveness of the proposed FMVM is evaluated based on two cases with either 4 or 15 HSI constituent stocks. The empirical result shows that the FMVM is able to improve the forecasting accuracy compared with the AMVM. It also shows that the computational cost required by the FMVM is less than that required by the AMVM.

The paper contributes to the following research issues: (1) this paper is the vanguard which combines the fuzzy clustering technique and multivariate volatility models in order to perform forecasting; (2) this research proposes a weighting mechanism to select models from different clusters and aggregate forecasting power of those selected models to perform forecasting; (3) our proposed FMVM overcomes the limitation of the existing AMVM which requires excessive utilization and computation burden, and existing AMVM is likely to generate unnecessary forecasting errors; (4) the proposed FMVM can achieve better forecasting accuracy with smaller computational efforts.

The proposed FMVM is applied to risk management and portfolio management. As an example, a fund manager needs to determine the predetermined weights of the assets from the same portfolio. When the other constraints are unchanged, the change of the volatility matrix affects the optimal weights of all the assets in the same portfolio. To address the uncertainty, one needs to process risk management in order to adjust portfolio in advance. However, some widely used optimal portfolio models are very sensitive to the volatility matrix. Therefore, a better forecasting method is essential.

The rest of the paper is organized as follows: Section 2 gives a brief introduction of some widely used univariate and multivariate volatility models; Section 3 discusses the particularization of the proposed FMVM; Section 4 presents the empirical results and analyzes. Several well-known and widely used volatility models were applied to 15 high weighted HSI constituent stocks from Nov 2010 to Oct 2014. The effectiveness of the FMVM is compared with those tested volatility models; the appendix presents different types of multivariate volatility models including the CCCs, the DCCs, the ADCCs, the OGARCHs, and the BEKK.

2 Volatility Models

This section first introduces the univariate GARCH model, and it then shows a general form of the multivariate GARCH models.

2.1 Univariate Volatility Model

Bollerslev [9] proposed a univariate volatility model namely GARCH model. For a log return time series r_t, let $\varepsilon_t = r_t - \mu_t$ be innovation at time t. Denote the information set available at time $t-1$. Given F_t, $\mu_t = E(r_t|F_t)$ is the conditional mean of r_t and $\sigma_t^2 = Var(r_t|\mu_t) = E[(r_t-\mu_t)^2|F_t]$ is conditional variance of r_t . ε_t is formulated in the model GARCH(p,q) as

$$r_t = \mu_t + \varepsilon_t,\ \varepsilon_t = \sigma_t z_t$$
$$\sigma_t^2 = c + \sum_{i=1}^{p} \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^2$$

where $\{z_t\}$ is a sequence of independent and identically distributed ($i.i.d.$) random variables with mean equal to 0 and variance equal to 1, $c > 0$, $\alpha_i \geq 0$, $\beta_j \geq 0$, and $\sum_{i=1}^{max(p,q)}(\alpha_i + \beta_i) < 1$. For $i > p$ and $j > q$, a_i and b_j are given as $\alpha_i = 0$ and $\beta_j = 0$, respectively. Given the GARCH(1,1) model and assume that the forecast origin is t, we define 1-step-ahead forecast form can be defined as

$$\sigma_{t+1}^2 = c + \alpha_1 \varepsilon_t^2 + \beta_1 \sigma_t^2$$

where ε_{t_0} and $\sigma_{t_0}^2$ are at time index t_0.

2.2 Multivariate Volatility Models

Almost all the multivariate volatility models can be represented as the decomposition of the conditional volatility matrix, $\mathbf{H}_t$ [11]:

$$\mathbf{H}_t = \mathbf{D}_t \mathbf{R}_t \mathbf{D}_t \tag{1}$$

where $\mathbf{R}_t$ is the one-step-ahead conditional correlation matrix with its (i,j)th entry denoted by $\rho_{ij,t}$, conditional linear correlation of $r_{i,t}$ and $r_{j,t}$, and $\mathbf{D}_t$ is a diagonal matrix with $\sqrt{\sigma_{ii,t}}$ on the (i,i)th entry. Equation (1) is a convenient decomposition, which allows separate specification of the conditional volatilities and conditional cross-asset returns correlations. The specification of $\sqrt{\sigma_{ii,t}}$ and $\rho_{ij,t}$ is varied among those multivariate GARCH models which is given in the following part. Given that we have n financial assets and $\varepsilon_t = r_t - \mu_t$, is the $n \times 1$ vector of residuals from the ordinary least square regressions (OLS) of the predictor variables. For n financial assets, D_t can be given as

$$\mathbf{D}_t = \begin{bmatrix} \sigma_{1,t} & & & \\ & \sigma_{2,t} & & \\ & & \ddots & \\ & & & \sigma_{n,t} \end{bmatrix}$$

and $\mathbf{R}_t$ is given as

$$\mathbf{R}_t = \begin{bmatrix} 1 & \rho_{12,t} & \cdots & \rho_{1n,t} \\ \rho_{21,t} & 1 & \cdots & \rho_{2n,t} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{n1,t} & \rho_{n2,t} & \cdots & 1 \end{bmatrix}$$

where $\rho_{ij,t} = \rho_{ji,t}, i \neq j$. $\mathbf{H}_t$ in Equation (1) can be elaborated as

$$\begin{aligned}\mathbf{H}_t &= \mathbf{D}_t\mathbf{R}_t\mathbf{D}_t \\ &= (\mathbf{D}_t\mathbf{R}_t)\mathbf{D}_t \\ &= \begin{bmatrix} \sigma_{11,t} & \sigma_{12,t} & \cdots & \sigma_{1n,t} \\ \sigma_{21,t} & \sigma_{22,t} & \cdots & \sigma_{2n,t} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{n1,t} & \sigma_{n2,t} & \cdots & \sigma_{nn,t} \end{bmatrix}\end{aligned}$$

where $\sigma_{ij,t} = \sigma_{ji,t}, i \neq j$. Based on the one-step-ahead forecast, Equation (1) are given as $\hat{\mathbf{H}}_{t+1} = \hat{\mathbf{D}}_{t+1}\hat{\mathbf{R}}_{t+1}\hat{\mathbf{H}}_{t+1}$.

Given the different specifications of $\mathbf{D}_t$ or/and $\mathbf{R}_t$, the general form can be represented as different type of multivariate volatility models:

1) When $\mathbf{R}_t$ is simply assumed as the constant conditional correlation matrix without varying with time, $\mathbf{R}$, $\mathbf{R}_t = \mathbf{R}$. The general form can be written into the CCC which is detailed in [11].
2) When $\mathbf{R}_t$ is a time varying matrix, the CCC can be reformulated as the DCC which can be referred to [13].
3) When the possibility of asymmetric effects can be allowed on conditional variance and correlations, the DCC can be relaxed into the the Asymmetric Dynamic Conditional Correlation Model (ADCC) which can be referred to [14].
4) When the standardized return is defined and the static principle component decompositions of standardized residuals is used, the OGARCH can be formulated, where the OGARCH can be referred to [12].
5) When $\mathbf{H}_t$ is assumed as the positive definite, the BEKK is somehow a quartic form of the general which can be referred to [10].

3 Methodology

In order to moderate the model uncertainty and to improve the forecasting accuracy simultaneously, Pesaran et al. [15] developed the average volatility models by applying model average techniques. The average multivariate volatility model is given as

$$\mathbf{H}_t^{amvm} = \frac{1}{m}\sum_{i=1}^{m}\mathbf{H}_{i,t}$$

where $\mathbf{H}_{i,t}$ is the forecasted conditional volatility matrix in the multivariate volatility model and m is the total number of the multivariate volatility models. Figure 1 illustrates the mechanism of the model averaging technique which generates an identical weight to m multivariate volatility models to forecast the one-day ahead volatility matrix. However, this method has several limitations:

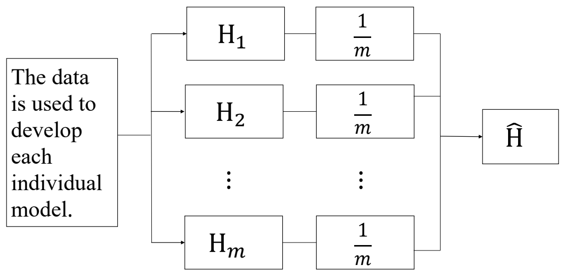


Fig. 1 The Mechanism of the Simple Model Averaging Technique

1) The excessive utilization ascribes to the similarity among different multivariate volatility models. These models sharing the similarity are classified into the same cluster. One forecasted volatility matrix generated by a volatility model may have linear correlation with another forecasted matrix generated by another similar model. The model averaging technique does not take this phenomenon into consideration.
2) The computation burden derived from the utilization of large-scale multi-models. Although this large amount of computation helps to adapt more situation and it performs indeed better than a single model, it occupies a lot of computation resources. Also, they may share similar properties or attributes with some other models, which can be grouped together to release the computation burden.
3) There exist some individual models which may involve higher tracking errors comparing with other considered models. Although it is a forthright way to weight each model equally, these models with higher tracking errors still affect the averaging tracking errors by model averaging. A more proper way is to give smaller weights to those models with higher tracking errors and to give larger weights to those models with higher forecasting accuracy.

To tackle the limitations of the existing model averaging technique, we propose the FMVM which attempts to give an alternative way to improve the forecasting accuracy. As illustrated in Figure 2, the mechanism of the FMVM is divided by training and operation phases. In the training phase, the FMVM inputs the

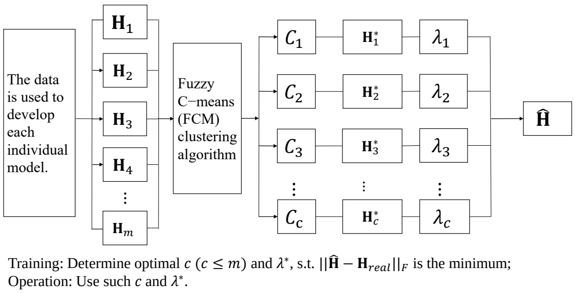


Fig. 2 The Mechanism of the FMVM

data and then applies the data to all individual models; it groups such individual models into different clusters using fuzzy C-means (FCM) clustering algorithm [19], when the number of the clusters is given; it selects the model with minimum tracking errors from each cluster; it determines the optimal weights to those selected models such that the mixed model has the minimum tracking errors. The FMVM decides the optimal number of the clusters, c, and selects the model with minimum tracking errors in each cluster. Subsequently, in the operation phase, the FMVM performs a one-day-ahead forecasting based on c and the resolved optimal weights of the selected models. It uses the sifted individual models with their optimal weights to forecast the volatility matrix.

3.1 Training Phase

The mechanism of the training phase is shown in Figure 3.1. In the training phase, the FMVM attempts to solve model-averaging problem based on fuzzy clustering and quadratic programming. We assume that m

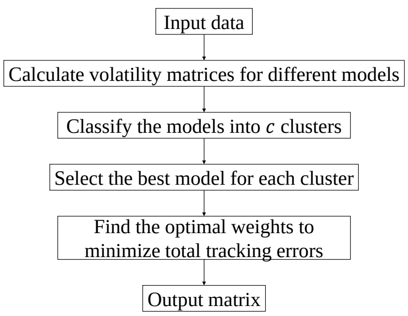


Fig. 3 The Mechanism of the Training Phase

multivariate volatility models are used and the data length is $3k-1$. In both training and operation phases, $2k-1$ data are involved. Define

$$\mathbf{H}_t^{real}(k) = (\mathbf{H}_{t-k+1}^{real}, \ldots, \mathbf{H}_t^{real})$$
$$\mathbf{H}_{i,t}(k) = (\mathbf{H}_{i,t-k+1}, \ldots, \mathbf{H}_{i,t}),\ i = 1, \ldots, m$$
$$L_{i,t}(k) = \|\mathbf{H}_{i,t}(k) - \mathbf{H}_t^{real}(k)\|_F,\ i = 1, \ldots, m,$$

where $\mathbf{H}_t^{real}(k)$, $\mathbf{H}_{i,t}(k)$, and $L_{i,t}(k)$ are the $n \times (k \times n)$ realized volatility matrices, m input time-varying $n \times (k \times n)$ volatility matrices for the ith model, and the tracking error for the ith model, respectively. Assume that there is a data set with m input time-varying $n \times (k \times n)$ volatility matrices $\mathbf{H}_{i,t}(k)$ $(i = 1, \ldots, m)$ and one output time-varying matrix $\mathbf{H}_t^{fmvm}$. The training phase performs three tasks: (1) classify the input models with the fuzzy C-means classification and select the model with the least tracking errors from each cluster, (2) determine the optimal weight for each selected model, (3) determine the optimal number of clusters, c.

3.1.1 Model Selection

The FMVM first transforms the input matrices into the vectors using the half-vectorization. The half-vectorization, $vech(\mathbf{H}_{j,t})$, processes the symmetric $n \times n$ matrix $\mathbf{H}_{j,t}$ of the $n(n+1)/2 \times 1$ columns $vech(\mathbf{H}_{j,t})$ only processes the lower triangular part of $\mathbf{H}_{j,t}$, where $vech(\mathbf{H}_{j,t})$ is given by:

$$vech(\mathbf{H}_{j,t}) = [H_{j,t}(1,1), \ldots, \mathbf{H}_{j,t}(n,1), \mathbf{H}_{j,t}(2,2), ..., \mathbf{H}_{j,t}(n,2), ..., \mathbf{H}_{j,t}(n,n)],\ j = 1, \ldots, m. \tag{2}$$

Let $X_j = vech(\mathbf{H}_{j,t}(K)), j = 1, \ldots, m$. Define the number of clusters as c. The fuzzy C-means (FCM) clustering algorithm [19] is used to partition m training samples, $X_j = vech(\mathbf{H}_{j,t}(k))$, into clusters C_1, C_2, ..., and C_c, where (5) computes the cluster center V_i of the cluster C_i and the membership grade u_{ij} of the training sample X_j belonging to C_i, with $1 \leq i \leq c$ and $1 \leq j \leq m$. We used the FCM clustering algorithm as this is widely used in pattern recognition [19]. In FCM, it denotes a fussiness index as s, $1 \leq s < \infty$. This fuzziness index s [20] is usually chosen to be 2. The FCM clustering algorithm merges m matrices X_j) into clusters C_i $(i = 1, 2, \ldots, c,\ 2 \leq c \leq m)$ by solving the following minimization problem [20]:

$$J_s = \sum_{i=1}^{c} \sum_{j=1}^{m} (u_{ij})^s ||V_i - X_j||_F^2, \tag{3}$$

where $||V_i - X_j||_F$ is the Frobenius norm between matrix X_j and the cluster center V_i $(i = 1, 2, \ldots, c)$; u_{ij}^s is the membership grade of X_j belonging to cluster C_i, in which s is the weighting exponent controlling the relative weights placed on each of the $||V_i - X_j||_F^2$, $1 \le s < \infty$. The FCM clustering algorithm is summarized as follows:

Step 1: Initialize $c = 2$.

Step 2: Generate the random number as the membership grade $u_{ij}^{(0)}$ of the sample X_j to the cluster C_i, where $0 \le u_{ij}^{(0)} \le 1$, $\sum_{i=1}^{c} u_{ij}^{(0)} = 1$, $1 \le i \le c$ and $1 \le j \le m$. Set $k = 1$.

Step 3: Compute the cluster center V_i of C_i based on

$$V_i^{(k)} = \frac{\sum_{j=1}^{n} (u_{ij}^{(k-1)})^s X_j}{\sum_{j=1}^{n} (u_{ij}^{(k-1)})^s} \tag{4}$$

where $1 \le i \le c$.

Step 4: Update u_{ij} based on:

$$u_{ij}^{(k)} = \frac{1}{\sum_{d=1}^{c} \left(\frac{||V_i^{(k)} - X_j||}{||V_d^{(k)} - X_j||} \right)^{\frac{2}{s-1}}} \tag{5}$$

where $1 \le i \le c$, $1 \le j \le m$, and $\sum_{i=1}^{c} u_{ij} = 1$.

Step 5: Set $k = k + 1$. Repeat **Step** 3 and **Step** 4 until J_s is no longer decreasing [19].

Step 6: Let $c = c + 1$. Go to **Step** 1 until $c \ge m + 1$.

3.1.2 Weight Assignment

The optimal weights of the selected models can be obtained by solving the following minimization problem:

$$\begin{aligned} &\min ||\mathbf{H}_t^{real}(k) - \sum_{i=1}^{c} \lambda_i^* \mathbf{H}_{i,t}^*(k)||_F^2 \\ &s.\,t. \;\; \sum_{i=1}^{c} \lambda_i^* = 1 \end{aligned} \tag{6}$$

where $\mathbf{H}_{i,t}^*(k)$ is the best prediction generated by the optimal model from C_i. Equation (6) can be reformulated as a quadratic form as

$$\begin{aligned} &\min \;\; \boldsymbol{\lambda}^T \mathbf{A} \boldsymbol{\lambda} + \boldsymbol{\lambda}^T B + C \\ &s.\,t. \;\; \sum_{i=1}^{c} \lambda_i = 1 \end{aligned} \tag{7}$$

where

$$\boldsymbol{\lambda} = (\lambda_1, \lambda_2, \ldots, \lambda_c)^T,$$

$$\mathbf{A} = \begin{pmatrix} \sum_{i=1}^{n} \sum_{j=1}^{nk} x_{ij}^{(1)} x_{ij}^{(1)} & \ldots & \sum_{i=1}^{n} \sum_{j=1}^{nk} x_{ij}^{(1)} x_{ij}^{(c)} \\ \vdots & \ddots & \vdots \\ \sum_{i=1}^{n} \sum_{j=1}^{nk} x_{ij}^{(c)} x_{ij}^{(1)} & \ldots & \sum_{i=1}^{n} \sum_{j=1}^{nk} x_{ij}^{(c)} x_{ij}^{(c)} \end{pmatrix},$$

$$B = -2((x_{ij}^{(1)} y_{ij}), (x_{ij}^{(2)} y_{ij}), \ldots, (x_{ij}^{(c)} y_{ij}))^T,$$

$$C = \sum_{i=1}^{n} \sum_{j=1}^{nk} y_{ij}^2 = ||\mathbf{H}_t^{real}(k)||_F,$$

$x_{ij}^{(c)}$ is the entry in the i^{th} row and j^{th} column of $\mathbf{H}_{i,t}^*(k)$, and y_{ij} is the element in the i^{th} row and j^{th} column of $\mathbf{H}_t^{real}(k)$.

Based on the FMVM, the forecasted volatility matrix $\mathbf{H}_t^{fmvm}$ can be developed as

$$\mathbf{H}_t^{fmvm} = \sum_{i=1}^{c} \lambda_i^* \mathbf{H}_{i,t}^*, \; i = 1, \ldots, c, \; 2 \le c \le m. \tag{8}$$

where n_M is the amount of the one-day-ahead volatility matrices determined by each individual multivariate volatility model; $\mathbf{H}_{i,t}^*$ is the one-day-ahead volatility matrix of the selected model i^*; and λ_i^* is the fuzzy optimal weight the selected model i^*.

3.1.3 Determination of the Number of Clusters

Given the number of cluster c, the tracking error can be evalulated. Therefore, the optimal number of clusters, namely c^*, can be determined as follows:

$$\min_c \;\; L_c = ||\mathbf{H}_t^{fmvm}(k, c) - \mathbf{H}_t^{real}(k)||_F. \tag{9}$$

3.2 Operation Phase

Given c^*, and the optimal weights for the i^*th selected model from the i^*th cluster which are determined in the training phase, the forecasted volatility matrix, namely $\mathbf{H}_t^{fmvm}$ is

$$\mathbf{H}_t^{fmvm} = \sum_{i=1}^{c^*} \lambda_i^* \mathbf{H}_{i,t}^*, \tag{10}$$

where $\mathbf{H}_{i,t}^*$ is the one-day-ahead volatility matrix of the selected model i, and λ_i^* is the optimal weight for the ith cluster.

Assume that we had n financial assets and the rolling windows size of realized covariance matrix is k. For the ith asset, $\varepsilon_{i,t} = r_{i,t} - \mu_{i,t}$, is the $n \times 1$ vector of residuals from the orthogonal least square (OLS) regressions. Given a $k \times n$ residual matrix

$$\mathbf{E}_t^{real} = \begin{bmatrix} \varepsilon_{1,t-k+1} & \varepsilon_{2,t-k+1} & \cdots & \varepsilon_{n,t-k+1} \\ \varepsilon_{1,t-k+2} & \varepsilon_{2,t-k+2} & \cdots & \varepsilon_{n,t-k+2} \\ \vdots & \vdots & \ddots & \vdots \\ \varepsilon_{1,t} & \varepsilon_{2,t} & \cdots & \varepsilon_{n,t} \end{bmatrix},$$

the realized covariance matrix with k time-series samples is defined as

$$\mathbf{H}_t^{real} = \mathbf{E}_t^{real^T} \mathbf{E}_t^{real} / k.$$

$\mathbf{H}_t^{real}$ is the benchmark for each $\mathbf{H}_{i,t}$, $i = 1, \ldots, m$.

The proposed FMVM and the AMVM generate the tracking errors L_t^{fmvm} and L_t^{amvm} at t which are denoted respectively as

$$L_t^{fmvm} = \|\mathbf{H}_t^{real} - \mathbf{H}_t^{fmvm}\|_F$$
$$L_t^{amvm} = \|\mathbf{H}_t^{real} - \mathbf{H}_t^{amvm}\|_F$$

where $\|.\|_F$ is the Frobenius norm. The covariance matrices generated by the FMVM and AMVM using k time-series samples can be written respectively as

$$\mathbf{H}_t^{fmvm}(k) = [\mathbf{H}_{t-k+1}^{fmvm}, \ldots, \mathbf{H}_t^{fmvm}]$$
$$\mathbf{H}_t^{amvm}(k) = [\mathbf{H}_{t-k+1}^{amvm}, \ldots, \mathbf{H}_t^{amvm}],$$

and, the proposed FMVM and the AMVM generate the tracking errors L_t^{fmvm} and L_t^{amvm} in k-length time period at t which are denoted respectively as

$$L_t^{fmvm}(k) = \|\mathbf{H}_t^{real}(k) - \mathbf{H}_t^{fmvm}(k)\|_F$$
$$L_t^{amvm}(k) = \|\mathbf{H}_t^{real}(k) - \mathbf{H}_t^{amvm}(k)\|_F.$$

Consequently, we denote Δ_t as the evaluation indicator of the FMVM which is relative to the AMVM at time index t, and $\Delta_t(k)$ is denoted as those with a k-length-time-period data. Based on the tracking errors of the FMVM and AMVM, Δ_t and $\Delta_t(k)$ can be elaborated as

$$\Delta_t = L_t^{fmvm} - L_t^{amvm}$$
$$\Delta_t(k) = L_t^{fmvm}(k) - L_t^{amvm}(k).$$

The algorithm of the operation phase is summarized as follows:

1: If $\Delta_t > 0$, the FMVM lags behind the AMVM with time t, else the FMVM leads the AMVM with time t.
2: If $\Delta_t(k) > 0$, the FMVM lags behind the AMVM with time t with a k-length-time-period data, else the FMVM leads the AMVM with time t. If $\Delta_t(k) = 0$ or $\Delta_t = 0$, the performance of the FMVM is equivalent to that of the AMVM.
3: Prepare the time series samples, based on a particular decision period, of n financial assets. Determine the types and number of the individual multivariate volatility models, m, and the rolling window size k;
4: Determine the parameters of multivariate volatility models based on the time series samples of the log-return of assets. Generate the $n \times n$ covariance matrix $\mathbf{H}_{i,t}$ with a k-length rolling window size at each time point of the index t. Generate the realized volatility based on another k-length period samples.
5: Generate a multivariate volatility model, $\mathbf{H}_{i,t}(k)$, by plugging $\mathbf{H}_{t-k+1}, \mathbf{H}_{t-k+2}, \ldots, \mathbf{H}_t$, in chronological order. Hence, the realized form $\mathbf{H}_t^{real}(k)$ can be developed.
6: For $i = 1, 2, \ldots, m$, the matrix A is recombined by the elements of those $\mathbf{H}_{i,t}(k)$ using the optimization algorithm; determine matrix B, based on the elements of $\mathbf{H}_{i,t}(k)$ and $\mathbf{H}_t^{real}(k)$; determine C based on the Frobenius norm of $\mathbf{H}_t^{real}(k)$; based on A, B and C, $\Lambda = (\lambda_1, \lambda_2, \ldots, \lambda_m)^T$ in (7) can be determined by harnessing the quadratic programming.
7: Obtain the L_t^{fmvm} and L_t^{amvm} and make decision based on FMVM using Δ_t if this is necessary.

4 Empirical Results

In this section, we demonstrate with examples how the proposed FMVM overcomes the limitations of the average multivariate volatility model (AMVM), namely that: (1) the FMVM selects fewer models without sacrificing forecasting powers; (2) the FMVM can achieve more accurate forecasting results than those achieved by AMVM. Indeed, the proposed FMVM can achieve better forecasting accuracy with smaller computational efforts.

The proposed FMVM is first evaluated based on 15 high weighted HSI constituent stocks. Table 1 shows the selected samples in both the low-dimensional cases and the high-dimensional cases of the volatility matrix forecasting and 17 different widely-used multivariate volatility models. In the low-dimensional cases, top 4 weighted HSI constituent stocks are chosen, and those stocks are ‘5.HK’, ‘939.HK’, ‘1299.HK’, and ‘9413.HK’. In the high-dimensional cases, top 15 weighted HSI constituent stocks, including the 4 identical stocks in the low-dimensional cases, are selected. Another 11 stocks are ‘1398.HK’, ‘3988.HK’, ‘883.HK’, ‘857.HK’, ‘386.HK’, ‘2628.HK’, ‘3.HK’, ‘151.HK’, ‘762.HK’, ‘494.HK’, and ‘3328.HK’. Table 1 specifies 17 most widely used multivariate volatility models. The models from the same class with different p or q is distinct from other models in the same class. For example, $CCC(1,1)$ and $CCC(2,1)$ come from the same family; however, $CCC(1,1)$ is a distinctive model and $CCC(1,1)$ is a independent model from $CCC(2,1)$. These daily trading samples, from 29 October 2010 to 7 October 2014, are obtained from the Bloomberg database which is collected from 29 October 2010 to 7 October 2014. The forecasting results of the proposed FMVM are compared with those obtained by the 17 commonly used forecasting models which are listed in listed in Table 1. All the models are used to forecast the one-day-ahead volatility matrices based on different rolling window size with $k = 20, 30, \ldots, 120$. Also, we compared the results obtained by the proposed FMVM with the existing multivariate volatility approach, the AMVM [15].

The CCC, DCC, ADCC, BEEK, OGARCH, AMVM models and the proposed FMVM were implemented using MATLAB with toolboxes including the Oxford MFE MATLAB toolbox [27]. The one-day-ahead volatility matrices for AMVM is given as:

Table 1 Selected stocks and models

The Low-Dim Case	The High-Dim Case	Multivariate Volatility Models
‘5.HK’	‘5.HK’	$\underline{CCC(p,q)}$
‘939.HK’	‘939.HK’	$CCC(1,2)$, $CCC(1,2)$
‘1299.HK’	‘1299.HK’	$CCC(2,1)$, $CCC(2,2)$
‘941.HK’	‘941.HK’	
	‘1398.HK’	$\underline{DCC(p,q)}$
	‘3988.HK’	$DCC(1,1)$, $DCC(1,2)$
	‘883.HK’	$DCC(2,1)$, $DCC(2,2)$
	‘857.HK’	
	‘386.HK’	$\underline{ADCC(p,q)}$
	‘2628.HK’	$ADCC(1,1)$, $ADCC(1,2)$
	‘3.HK’	$ADCC(2,1)$, $ADCC(2,2)$
	‘151.HK’	
	‘762.HK’	$\underline{OGARCH(p,q)}$
	‘494.HK’	$OGARCH(1,1)$, $OGARCH(1,2)$
	‘3328.HK’	$OGARCH(2,1)$, $OGARCH(2,2)$
		$BEKK$
4 assets	15 assets	17 models

$$\mathbf{H}_t^{amvm} = \frac{1}{17}\sum_{i=1}^{17}\mathbf{H}_{i,t} \qquad (11)$$

where $\mathbf{H}_{1,t}, \mathbf{H}_{2,t}, \ldots, \mathbf{H}_{17,t}$ are the conditional volatility matrix of each model listed in Table 1.

We evaluate the models with different rolling window sizes. The window size is varied from $k = 20$ to 120 in a step of 10. The date, 7 October 2014, is used as the end window. We used the last k trading days as the operation period and the preceding k trading days as the training period. Therefore each model can be evaluated $2k$ times. For instance, in the training phase, the forecasted results are $\hat{Y}_t, \ldots, \hat{Y}_{t+k}$, where $\hat{Y}_t$ is a forecasted value at time t based on the information the time series samples of $t-k+1, \ldots, t-1$. Similarly, in the operation phase, the forecasted results are $\hat{Y}_{t+k+1}, \ldots, \hat{Y}_{t+2k}$, where $\hat{Y}_{t+k+1}$ is forecasted based on the time series samples of $t+1, \ldots, t+k$. In training phase, we try to use k forecasted results $\hat{Y}_t, \ldots, \hat{Y}_{t+k}$ to train the FMVM model (where the optimal number of clusters and the optimal weights of the selected models are predefined); in operation phase, we use the trained FMVM to forecasted k forecasted results $\hat{Y}_{t+k+1}, \ldots, \hat{Y}_{t+2k}$. We need the operation phase to validate the performance of the proposed FMVM. Both low-dimensional and high-dimensional circumstances are considered. For the clustering results, the low-dimensional case is illustrated, and the complete results of models in each cluster are shown in Appendix A.

4.1 The low-dimensional case

Here, we consider a simple case with 4 stocks which are shown in the first column of Table 1. The tracking errors $\Delta_I(k)$ and the differences $\Delta_{II}(k)$ are shown in Table 2. $L_I^F(k)$ and $L_I^F(A)$ are the total tracking errors for the training phase obtained by the FMVM and the AMVM respectively. $L_{II}^F(k)$ and $L_{II}^A(k)$ are the total tracking errors for the operation phase obtained by the FMVM and the AMVM respectively.

Table 2 Tracing Errors in the Training and the Operation Phase

RWS k	clusters	$L_I^F(k)$	$L_I^A(k)$	$\Delta_I(k)$	$L_{II}^F(k)$	$L_{II}^A(k)$	$\Delta_{II}(k)$
20	3	3.1162	7.7491	-4.6329	4.8583	7.9454	-3.0870
30	3	7.5967	15.0727	-7.4761	4.4505	7.3815	-2.9310
40	5	5.0178	16.6071	-11.5893	6.6014	9.6715	-3.0701
50	4	7.3173	18.0725	-10.7552	6.7466	10.8176	-4.0710
60	7	8.9751	18.9535	-9.9784	6.8483	10.3731	-3.5247
70	4	8.0716	19.3060	-11.2343	8.0451	12.0133	-3.9682
80	6	7.0084	16.1564	-9.1481	7.9594	13.0397	-5.0804
90	6	7.1209	13.4263	-6.3055	10.2467	14.1335	-3.8868
100	4	9.9625	12.9350	-2.9725	10.5589	14.4315	-3.8726
110	7	9.5720	14.0492	-4.4772	10.1283	14.7852	-4.6569
120	7	8.7855	15.5307	-6.7453	9.5383	13.6743	-4.1360

This results for both training and operation phases show that the tracking errors of the FMVM are smaller than those of the AMVM as the differences are negative in all cases. Hence, the FMVM can surpass the AMVM not only in training but also in operation phase.

Table 2 shows that the number of clusters with respect to different rolling windows. Table 2 shows that the FMVM required less computational efforts, as three to seven models are only involved of the FMVM. As the number of clusters is identical to the number of the individual models, the FMVM does not need to involve all 17 models. Therefore, this low-dimensional cases study shows that the proposed FMVM can achieve better forecasting accuracy with smaller computational efforts comparing with the AMVM.

4.2 The high-dimensional case

In the high dimensional case, we consider 15 stocks from the Hang Seng Index, as shown in the second column of Table 1. Table 3 shows the tracking errors and the differences between the FMVM and the AMVM.

Table 3 Tracing Errors in the Training and the Operation Phase

RWS k	clusters	$L_I^F(k)$	$L_I^A(k)$	$\Delta_I(k)$	$L_{II}^F(k)$	$L_{II}^A(k)$	$\Delta_{II}(k)$
20	5	15.6529	49.2121	-33.5592	18.4214	33.4936	-15.0722
30	3	26.8251	50.9704	-24.1454	17.1039	33.4648	-16.3608
40	4	34.6771	85.3520	-50.6749	20.5884	40.1193	-19.5309
50	4	25.2633	97.2940	-72.0308	31.0473	48.5467	-17.4994
60	5	27.5873	94.5644	-66.9771	29.7626	51.8209	-22.0584
70	4	25.1795	93.0677	-67.8883	30.0937	57.6992	-27.6054
80	3	23.2929	87.4429	-64.1500	26.4884	58.3126	-31.8242
90	5	18.1365	84.6518	-66.5153	29.4302	64.1372	-34.7070
100	3	21.0107	82.8058	-61.7951	22.8918	68.6020	-45.7102
110	3	19.7757	78.7352	-58.9595	21.6531	69.0370	-47.3840
120	4	20.1628	76.0412	-55.8784	21.2291	70.0627	-48.8337

The result shows that the tracking errors of the FMVM are smaller than those of the AMVM. It also shows that the FMVM can surpass the AMVM not only in the training phase but also in the operation phase. Also, a smaller computational effort is required by the FMVM as three to five models are involved among all the seventeen models. Therefore, the empirical results show that the FMVM can obtain better forecasting accuracy and the FMVM requires much less computational effort than that required by the AMVM.

5 Conclusion

In this paper, a novel multivariate volatility model FMVM is proposed to improve forecasting accuracy. The proposed FMVM overcomes the disadvantages of incurring redundant computation when all multivariate volatility models are employed in an averaging manner. The proposed FMVM classifies individual volatility model into smaller scale clusters by using fuzzy C-means clustering algorithm and selects the most representative model with the lowest tracking error from each cluster. Optimal weight for each selected model can then be sought via training. The effectiveness of the proposed FMVM is evaluated based on 15 stocks from the Hang Seng Index. Empirical results have shown that the proposed FMVM can obtain better accuracy in volatility forecasting with less number of volatility models. Hence, a less computational effort is required by the proposed FMVM than that required by the AMVM. For the future work, we will apply the proposed FMVM to other financial products such as the commodity market.

References

1. P. F. Christoffersen and F. X. Diebold, “How relevant is volatility forecasting for financial risk management?” *Review of Economics and Statistics*, vol. 82, pp. 12–23, 2000.
2. K. F. C. Yiu, “Optimal portfolios under a value-at-risk constraint,” *Journal of Economic Dynamics and Control*, vol. 28, no. 7, pp. 1317–1334, 2004.
3. P. Jorion, *Value at risk: the new benchmark for managing financial risk, 3rd ed.* McGraw Hill Professional, 2006.
4. K. F. C. Yiu, J. Z. Liu, T. K. Siu, and W. K. Ching, “Optimal portfolios with regime switching and value-at-risk constraint,” *Automatica*, vol. 46, no. 6, pp. 979–989, 2010.
5. J. Z. Liu, K. F. C. Yiu, T. K. Siu, and W. K. Ching, “Optimal investment-reinsurance with dynamic risk constraint and regime switching,” *Scandinavian Actuarial Journal*, vol. 2013, no. 4, pp. 263–285, 2013.
6. K. van den Broek, “Long-term insurance products and volatility under the solvency II framework,” *European Actuarial Journal*, vol. 4, no. 2, pp. 315–334, 2014.
7. K. F. C. Yiu, S. Y. Wang, and K. L. Mak, “Optimal portfolios under a value-at-risk constraint with applications to inventory control in supply chains,” *Journal of Industrial and Management Optimization*, vol. 4, no. 1, p. 81, 2008.
8. S. Y. Wang, K. F. C. Yiu, and K. L. Mak, “Optimal inventory policy with fixed and proportional transaction costs under a risk constraint,” *Mathematical and Computer Modelling*, vol. 58, no. 9, pp. 1595–1614, 2013.
9. T. Bollerslev, “Generalized autoregressive conditional heteroskedasticity,” *Journal of Econometrics*, vol. 31, no. 3, pp. 307–327, 1986.
10. R. Engle and K. F. Kroner, “Multivariate simultaneous generalized ARCH,” *Econometric theory*, vol. 11, no. 01, pp. 122–150, 1995.
11. T. Bollerslev, “Modelling the coherence in short-run nominal exchange rates: a multivariate generalized arch model,” *The Review of Economics and Statistics*, pp. 498–505, 1990.
12. C. Alexander, “Orthogonal GARCH,” in *Mastering Risk, Financial Times.* Prentice Hall, 2001, vol. 2, pp. 21–38.
13. R. Engle, “Dynamic conditional correlation: A simple class of multivariate generalized autoregressive conditional heteroskedasticity models,” *Journal of Business & Economic Statistics*, vol. 20, no. 3, pp. 339–350, 2002.
14. L. Cappiello, R. F. Engle, and K. Sheppard, “Asymmetric dynamics in the correlations of global equity and bond returns,” *Journal of Financial Econometrics*, vol. 4, no. 4, pp. 537–572, 2006.

15. M. H. Pesaran, C. Schleicher, and P. Zaffaroni, “Model averaging in risk management with an application to futures markets,” *Journal of Empirical Finance*, vol. 16, no. 2, pp. 280–305, 2009.
16. R. Batuwita and V. Palade, “Fsvm-cil: Fuzzy support vector machines for class imbalance learning,” *IEEE T. Fuzzy Systems*, vol. 18, no. 3, pp. 558–571, 2010.
17. R. Batuwita, V. Palade, and D. C. Bandara, “Fsvm-cil: Fuzzy support vector machines for class imbalance learning,” *International Journal on Artificial Intelligence Tools*, vol. 20, no. 3, pp. 425–455, 2011.
18. A. Serguieva and J. Hunter, “Fuzzy interval methods in investment risk appraisal,” *Fuzzy Sets and Systems*, vol. 142, no. 3, pp. 443–466, 2004.
19. S. M. Chen and Y. C. Chang, “Multi-variable fuzzy forecasting based on fuzzy clustering and fuzzy rule interpolation techniques,” *Information Sciences*, vol. 180, pp. 4772–4783, 2010.
20. J. C. Bezdek, R. Ehrlich, and W. Full, “Fcm: The fuzzy c-means clustering algorithm,” *Computers & Geosciences*, vol. 10, no. 2-3, pp. 191–203, 1984.
21. A. Pathak and N. R. Pal, “Clustering of mixed data by integrating fuzzy, probabilistic, and collaborative clustering framework,” *International Journal of Fuzzy Systems*, pp. 1–10.
22. T. Zhang, L. Chen, and C. P. Chen, “Clustering algorithm based on spatial shadowed fuzzy c-means and i-ching operators,” *International Journal of Fuzzy Systems*, vol. 18, no. 4, pp. 609–617, 2016.
23. H. Dalman, “An interactive fuzzy goal programming algorithm to solve decentralized bi-level multiobjective fractional programming problem,” *arXiv preprint arXiv:1606.00927*, 2016.
24. H. Dalman, N. Güzel, and M. Sivri, “A fuzzy set-based approach to multi-objective multi-item solid transportation problem under uncertainty,” *International Journal of Fuzzy Systems*, pp. 1–14, 2015.
25. H. Dalman, “Uncertain programming model for multi-item solid transportation problem,” *International Journal of Machine Learning and Cybernetics*, pp. 1–9, 2016.
26. T. G. Andersen, T. Bollerslev, F. X. Diebold, and H. Ebends, “Fuzzy interval methods in investment risk appraisal,” *Journal of Financial Economics*, vol. 61, pp. 43–76, 2001.
27. K. Sheppard, “MFE matlab function reference,” *Financial Econometrics [Computer software manual]*, 2009.
28. L. R. Glosten, R. Jagannathan, and D. E. Runkle, “On the relation between the expected value and the volatility of the nominal excess return on stocks,” *The journal of Finance*, vol. 48, no. 5, pp. 1779–1801, 1993.

A Model Classification Example of Low Dimensional Cases

Assume that we have 4 assets and use one-day-ahead forecasting. The classification results are shown in Table 4 to 12. The first row in each table shows our benchmark. Other rows show different clustering results with the first model in each row/cluster being the representative model for this cluster; for instance, in Table 12, $DCC(2,2)$ and $ADCC(2,2)$ are grouped in cluster 6, and the $DCC(2,2)$ is the representative model.

Table 4 The 2-classes case

The index of the cluster	The benchmark and the models
1	Realized Volatility
2	Others (17 models mentioned in this paper)

Table 5 The 3-classes case

The index of the cluster	The benchmark and the models
1	Realized Volatility
2	DCC(2,2), OGARCH(1,1), OGARCH(1,2), OGARCH(2,1), OGARCH(2,2))
3	CCC(1,1), CCC(1,2), CCC(2,1), DCC(1,1), DCC(1,2), DCC(2,1), DCC(2,2), ADCC(1,1), ADCC(1,2), ADCC(2,1), ADCC(2,2), BEKK

Table 6 The 4-classes case

The index of the cluster	The benchmark and the models
1	Realized Volatility
2	CCC(2,2), DCC(2,2), ADCC(2,2)
3	OGARCH(1,1), OGARCH(1,2), OGARCH(2,1), OGARCH(2,2)
4	CCC(1,1), CCC(1,2), CCC(2,1), DCC(1,1), DCC(1,2), DCC(2,1), ADCC(1,1), ADCC(1,2), ADCC(2,1), BEKK

Table 7 The 5-classes case

The index of the cluster	The benchmark and the models
1	Realized Volatility
2	CCC(2,2)
3	DCC(2,2), ADCC(2,2)
4	OGARCH(1,1), OGARCH(1,2), OGARCH(2,1), OGARCH(2,2)
5	CCC(1,1), CCC(1,2), CCC(2,1), DCC(1,1), DCC(1,2), DCC(2,1), ADCC(1,1), ADCC(1,2), ADCC(2,1), BEKK

Table 8 The 6-classes case

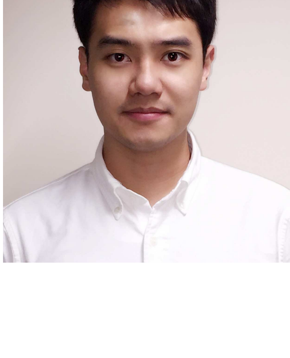
The index of the cluster	The benchmark and the models
1	Realized Volatility
2	CCC(2,2)
3	OGARCH(2,2)
4	OGARCH(1,1), OGARCH(1,2), OGARCH(2,1)
5	CCC(1,1), CCC(1,2), CCC(2,1), DCC(1,1), DCC(1,2), DCC(2,1), ADCC(1,1), ADCC(1,2), ADCC(2,1), BEKK
6	CCC(1,1), CCC(1,2), CCC(2,1), DCC(1,1), DCC(1,2), DCC(2,1), ADCC(1,1), ADCC(1,2), ADCC(2,1), BEKK

Table 9 The 7-classes case

The index of the cluster	The benchmark and the models
1	Realized Volatility
2	CCC(2,2)
3	OGARCH(2,2)
4	OGARCH(2,1)
5	DCC(2,2),ADCC(2,2)
6	OGARCH(1,1), OGARCH(1,2)
7	CCC(1,1), CCC(1,2), CCC(2,1), DCC(1,1), DCC(1,2), DCC(2,1), ADCC(1,1), ADCC(1,2), ADCC(2,1), BEKK

Table 10 The 8-classes case

The index of the cluster	The benchmark and the models
1	Realized Volatility
2	CCC(2,2)
3	OGARCH(2,2)
4	OGARCH(2,1)
5	BEKK
6	DCC(2,2),ADCC(2,2)
7	OGARCH(1,1), OGARCH(1,2)
8	CCC(1,1), CCC(1,2), CCC(2,1), DCC(1,1), DCC(1,2), DCC(2,1), ADCC(1,1), ADCC(1,2), ADCC(2,1)



Zi-kai Wei received his B.A. degree in Finance and Banking from Shenzhen Univesity, Shenzhen, China, in 2013. He is currently an M.Phil. candidate in the Department of Applied Mathematics at the Hong Kong Polytechnic University. His research interests are in

Table 11 The 9-classes case

The index of the cluster	The benchmark and the models
1	Realized Volatility
2	CCC(2,2)
3	OGARCH(2,2)
4	OGARCH(2,1)
5	BEKK
6	DCC(2,2),ADCC(2,2)
7	OGARCH(1,1), OGARCH(1,2)
8	CCC(2,1), DCC(2,1), ADCC(2,1)
9	CCC(1,1), CCC(1,2), DCC(1,1), DCC(1,2), ADCC(1,2), ADCC(2,1)

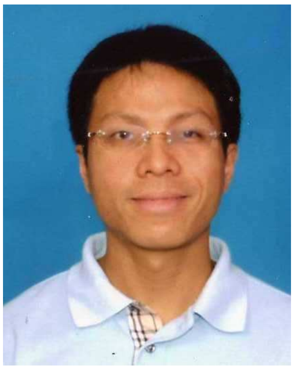
Table 12 The 10-classes case

The index of the cluster	The benchmark and the models
1	Realized Volatility
2	CCC(2,2)
3	OGARCH(2,2)
4	OGARCH(2,1)
5	BEKK
6	DCC(2,2),ADCC(2,2)
7	OGARCH(1,1), OGARCH(1,2)
8	CCC(2,1), DCC(2,1), ADCC(2,1)
9	CCC(1,2), DCC(1,2), ADCC(1,2)
10	OGARCH(1,1), OGARCH(1,2)

the areas of volatility models, fuzzy system, pattern recognition, and trading strategies.

Ka-Fai Cedric Yiu received his M.Sc. from University of Dundee and University of London, and D.Phil. from University of Oxford. Over the century, he had worked closely with the industry on different projects in University of Oxford and University College of London. He started his lecturing career in the University of Hong Kong. He is currently working in the Hong Kong Polytechnic University. He has served on the program committee and the organizing committee of a number of conferences and has organized a number of special sessions in conferences. He has published over 100 journal publications and given over 30 conference presentations. He holds two U.S. patents in signal processing. He received the third prize of Chongqing Natural Science Foundation Award in 2014. He also received the Donald Julius Groen Prize back in 2002. He is currently working on several research projects related to optimization and data analysis, signal processing system, risk management and high frequency trading. His current research interests include optimization and optimal control, signal processing and financial risk management.

Heung Wong is a professor in the Department of Applied Mathematics, The Hong Kong Polytechnic University. He graduated with B.Sc. in Mathematics, M.Sc. in Statistics, and Ph.D. in Statistics from The Chinese University of Hong Kong, Newcastle University (U.K.), and University of Hong Kong respectively. His current research interests include time series analysis, graphical modeling, and application of mathematics and statistics to finance, economics, environmental science and hydrology.



Kit Yan Chan is a Senior Lecturer in the Department of Electrical and Computer Engineering, Curtin University, Australia. He received his PhD in Computing in London South Bank University, U.K., in 2006. He was a full time researcher in The Hong Kong Polytechnic University (2004-2009) and Curtin University (2009-2013). He has edited a research book entitled Computational Intelligence Techniques for New Product Development. He has published 1 research book, 64 international journal papers, 43 international conference papers and 2 book chapters. His current research interests are in machine learning and affective computing.