Article

A Network-Based Analysis for Evaluating Conditional Covariance Estimates

Carlo Drago and **Andrea Scozzari** *

Facoltà di Economia, Università degli Studi Niccolò Cusano Roma, 00166 Roma, Italy
* Correspondence: andrea.scozzari@unicusano.it

Abstract: The modeling and forecasting of dynamically varying covariances has received a great deal of attention in the literature. The two most widely used conditional covariances and correlations models are BEKK and the DCC. In this paper, we advance a new method based on network analysis and a new targeting approach for both the above models with the aim of better estimating covariance matrices associated with financial time series. Our approach is based on specific groups of highly correlated assets in a financial market and assuming that those relationships remain unaltered at least in the long run. Based on the estimated parameters, we evaluate our targeting method on simulated series by referring to two well-known loss functions introduced in the literature. Furthermore, we find and analyze all the maximal cliques in correlation graphs to evaluate the effectiveness of our method. Results from an empirical case study are encouraging, mainly when the number of assets is not large.

Keywords: network analysis; mining financial data; maximal cliques; multivariate GARCH models; BEKK; DCC; targeting

MSC: 62M10; 91B84

1. Introduction

Modeling financial volatility is a crucial issue in empirical finance. *Heteroscedasticity* is a peculiar feature of a financial time series, as it is characterized by periods more volatile than others due to unpredictable outer events. The existence of heteroscedasticity is a major concern in a univariate approach aiming to estimate and forecast a given (financial) phenomenon. Estimating financial volatility becomes a more difficult task if we consider N time series (multivariate approach) and we are interested in understanding the co-movements of financial returns. In this regard, different multivariate general auto-regressive conditional heteroscedasticity (MGARCH) models have been presented in the literature. In finance, the main aim of MGARCH models is to predict the future values of the variance and covariance matrix of asset returns. Readers may refer to the very recent paper [1] and the references therein for a comprehensive review of the existing literature on MGARCH approaches. These methods play an important role in different fields of finance such as, for instance, in portfolio optimization [2–4], option pricing [5,6], energy markets [7], and analysis of contagion and volatility spillovers [8–10]. In the literature, MGARCH models are classified into different categories (see, e.g., [11]): (1) direct generalizations of the univariate models; (2) linear combinations of univariate models, such as, generalized orthogonal models and latent factor models; (3) nonlinear combinations of univariate models; and (4) nonparametric and semi-parametric models that constitute an alternative to the parametric estimation of the financial volatility and do not impose a particular structure on the data. In this paper, we will focus on categories (1) and (3).

In general, estimating a time-varying covariance matrix is difficult due to the problem's dimension and the fact that a covariance matrix must be positive definite. The most

renowned MGARCH model of category (1) is BEKK, which has the attractive property that the conditional covariance matrices are positive definite by construction. However, estimation of a BEKK model involves somewhat heavy computations due to the fact that it contains a large number of parameters. To overcome this problem, one approach was to decompose the conditional covariance matrix into conditional standard deviations and a conditional correlation matrix (models of category (3)). In [12,13], the authors introduced the dynamic conditional correlation model (DCC-GARCH), for which the conditional correlation matrix is designed to vary over time. This approach has one important advantage, that is, the number of parameters to be simultaneously estimated is reduced, as a complex optimization problem is disaggregated into simpler ones.

Several authors have provided a number of different modifications of the MGARCH models, some of which are mentioned in [1]. Other meaningful approaches explicitly incorporate the effect of measurement errors and time-varying attenuation biases into the covariance estimations and forecasts (see, e.g., [14]). Engle and Kelly [15] studied a special case of the DCC-GARCH model, the dynamic equicorrelation model, in which all correlations are identical. This can be a useful model in situations where the number of parameters to estimate is large.

A specific modification of the MGARCH models is provided in [16]. The authors directly modified the structure of the BEKK and DCC models by introducing a "target" matrix and compared their modified models to the original ones. The authors considered the long-run covariance matrix as a target matrix, which can be consistently estimated by the corresponding sample estimator and considered to be different versions of the BEKK and DCC models. They point out that by imposing positive definiteness and covariance stationarity for the different versions of BEKK is extremely complicated. For the DCC models, targeting can be useful to reduce the number of parameters to estimate. However, in [17] the authors proved that the estimation of DCC models with targeting may be inconsistent. From a theoretical viewpoint, the paper [16] is worthy of interest because the authors study the availability of the analytical forms for the sufficient conditions for consistency, asymptotic normality of the appropriate estimators, and computational tractability for large-dimension problems. However, in [16], no empirical tests were presented, and the authors claimed that it is not possible to provide an appropriate evaluation of which of the different MGARCH models is preferred.

In this paper, we advance a novel approach based on network analysis for evaluating the estimates of the time-varying correlation matrices in financial markets. We refer to the BEKK and DCC models and propose a variant of these two models by suitably modifying the log-likelihood function to maximize. We call the two resulting models the *modified* BEKK and *modified* DCC models, respectively. This modification consists of introducing a term in the function incorporating a loss measure based on the difference between the time-varying covariance matrices and the covariance matrix estimated with respect to the whole in-sample period. In actuality, as observed in [18], a financial market is characterized by some stylized facts. More precisely, there are often specific groups of assets that are highly correlated in such a way that positive price changes of one asset in the group determine positive price changes of all the other assets in the group, and these relationships remain unaltered over time. Hence, the idea behind our modification is that (extremely) high/low values of correlations observed between pairs of assets do not change too much during time. In particular, any pair of highly positive or negative correlated assets in the market remains highly (positive/negative) correlated during the given observed time period.

In a financial market, the correlation between assets can be represented via a *correlation graph* where, given a time period, assets are identified by vertices of a complete graph G and distances (weights) assigned to pairs of assets (edges) incorporate the dependency structure of returns. Mantegna [19] was one of the first who constructed asset graphs based on stock price correlations in order to detect the hierarchical organization inside a stock market. In order to highlight only groups of assets whose correlation is above or below a given threshold δ, one can extract a (sub)graph $G(\delta)$ of G that contains only

a subset of the connections in G between pairs of assets. In this new graph $G(\delta)$, two stocks are linked if and only if they exhibit high/low values of correlation. In other words, the *observed* subgraph $G(\delta)$ reveals the strong relationships between groups of assets in a financial market during a given time period T. Examples of such interconnected groups are, for instance, firms belonging to a specific industrial sector. We introduce an additional term in the log-likelihood function that takes into account the fact that the time-varying covariance or correlation matrices must not alter the clusters formed by highly correlated assets observed during T. Hence, with our method we estimate a modified MGARCH model, simulate T realizations of returns of each of the N assets, compute the correlation matrix with respect to the whole simulated in-sample period and, given the threshold δ, we obtain the corresponding *simulated* graph $G_S(\delta)$. Then, we compare the observed and simulated graphs $G(\delta)$ and $G_S(\delta)$, respectively. In particular, we compare all the *maximal cliques* of the two graphs. Given a general graph G, a *clique* is a set of interconnected stocks that forms a complete subgraph of G. Hence, maximal cliques in $G(\delta)$ and $G_S(\delta)$ represent highly (positive/negative) correlated assets, that is, a stock that belongs to a clique is highly correlated with all other stocks in the clique. The comparisons allow us to verify if a modified MGARCH model has been able to correctly reproduce the volatility and all the strong relationships among assets of the given financial market and if a modified model is able to outperform the corresponding original MGARCH model.

In light of the above discussion, the contribution of the present paper is threefold. First, we propose a new alternative targeting method that does not alter the models' structure. In fact, we modify only the log-likelihood objective function by introducing a suitable loss or distance measure between the long-run sample covariance or correlation matrix and the corresponding conditional covariance or conditional correlation matrices, respectively. Hence, we do not have to impose additional constraints for covariance stationarity or to guarantee that the matrices are definite positive. Second, for the first time, we provide an empirical analysis for evaluating modified and original models. We do this with the caveat that it is not our goal to definitely decide which of the different models is the best; rather, we seek to understand whether our modified models allow to better capture the strong relationships between assets in a stock market. We hope that the modified models can be more accurate than the original ones. Third, we advance a new method for evaluating the effectiveness of the estimated models through some specific tools commonly used in network analysis.

Our findings enable future analysis on volatility. In actuality, the importance of identifying the volatility effect and correlation in a dynamic way is crucial for efficiently managing the investment portfolios and executing optimal diversification of assets. This will reduce the risk and controls the changes that may occur in financial markets due to the economic situation in general. Our approach can build a precise mathematical model for managing financial data, which can be also incorporated in artificial intelligence algorithms to avoid the problems associated with the problems' dimensions.

The paper is organized as follows: Section 2 reports notation and some basic definitions. Sections 2.1 and 2.2 summarize the two MGARCH models considered in the paper along with the new log-likelihood functions used in the optimization/estimation phase. Section 3 reports the structure of the empirical financial dataset we used in the experimental phase. Finally, some conclusions and further research are depicted in Section 4.

2. Notation and Definitions

Consider a financial market formed by N assets. Let P_{jt} be the daily closing price of asset j at time t, $t = 0, \ldots, T$ and $r_{jt} = \log(\frac{P_{jt}}{P_{jt-1}})$ be the corresponding log-return of asset j. In the following, Pearson correlation coefficients are used to detect dependencies between assets returns. A general MGARCH model is defined:

$$\begin{cases} r_t & = & \mu_t + \epsilon_t \\ \epsilon_t & = & H_t^{\frac{1}{2}} \eta_t, \end{cases} \tag{1}$$

where r_t is the $N \times 1$ vector of log-returns at time t, ϵ_t is the $N \times 1$ vector of mean-corrected returns of N assets at time t, $\mathrm{E}[\epsilon_t] = 0$, and $\mathrm{Cov}[\epsilon_t] = H_t$. The vector μ_t represents the expected value of r_t. Observe that μ_t may be modeled as a constant vector or as a time series model. In this paper, we assume μ_t constant.

H_t is the $N \times N$ matrix of conditional variances and covariances of the unpredictable component ϵ_t at time t. Finally, η_t is the $N \times 1$ vector of i.i.d errors such that $\mathrm{E}[\eta_t] = 0$ and $\mathrm{E}[\eta_t \eta_t'] = I$.

In a general MGARCH model we have to estimate the conditional covariance matrix H_t, which, in addition, has to be positive definite for all t. Depending on the possible specifications for H_t, there are different MGARCH models each belonging to one of the four categories mentioned in Section 1.

In order to estimate the conditional covariance matrix H_t, a common issue is to resort to the maximization of an appropriate log-likelihood function $L(\theta)$, where θ denotes the vector of all the parameters to estimate. Depending on the MGARCH model, the function $L(\theta)$ assumes a different appearance. It is well-known that the quality of the maximum likelihood (ML) estimation relies also on the assumed data distribution. In general, when dealing with models with conditional heteroscedasticity, the estimates are known to be asymptotically normal [20]. In our approach, we consider a multivariate Gaussian distribution for the standardized error η_t, even if, in principle, our method can be applied assuming different distributions.

Given a stock market, consider H_t the conditional variance and covariance matrix of the returns of the N assets at time t, and let $R = [\rho_{ij}]$, with $\rho_{ii} = 1$, $i = 1, \ldots, N$, be the global correlation matrix with respect to the sample period. The idea is that if asset i is highly positively/negatively correlated with asset j, then the (high positive/negative) correlation does not change during the observed period. Furthermore, we assume that the relationship based on the correlation between two assets i and j does not change during time whereas the value of the correlation coefficient ρ_{ij}, can (obviously) vary.

Let $\delta > 0$ be a threshold. We construct the correlation graph $G(\delta) = (V, E)$ where V is the set of vertices each representing an asset, and E is the set of edges of $G(\delta)$, that is, the set of connections between pairs of assets. We assume that there is an edge (i, j) between assets i and j if and only if the (global) correlation coefficient $|\rho_{ij}| > \delta$. Let $A(\delta)$ be the adjacency matrix of $G(\delta)$. Observe that the generic element $a_{ij}(\delta)$ can be 0 or 1 if we assume an unweighted graph, otherwise $a_{ij}(\delta) = w_{ij}$ if we assume that the graph $G(\delta)$ is weighted with weights w_{ij} assigned to each edge (i, j) and obtained as a function of the correlation coefficients ρ_{ij}, that is, $w_{ij} = f(\rho_{ij})$. Given a general graph $G = (V, E)$, the following definitions hold [21]:

Definition 1. *A subset $C \subseteq V$ is called a clique of G if any two vertices in C are connected by an edge. The order q of a clique is the cardinality of C.*

Definition 2. *A subset $C \subseteq V$ is called a maximal clique of order q of G if C is not included in any other clique C' of order $q + 1$.*

Other concepts not defined in the paper can be found in the book [21].

In a modified MGARCH model, the modified likelihood function to maximize is:

$$L(\theta) - \sum_{t=1}^{T} [Dist(A(\delta), A_t(\delta)], \tag{2}$$

where $Dist(\cdot)$ is (any) distance or divergence measure between the adjacency matrix $A(\delta)$ referred to the global correlation matrix R, and the adjacency matrix $A_t(\delta)$ related to the conditional correlation matrix R_t. Because $A(\delta)$ is the correlation matrix computed with respect to the sample period, it can be considered to be a target matrix.

In our framework, as a measure we can consider any distance or divergence measure between two $N \times N$ positive definite matrices P and Q. In this paper, we consider the

well-known Kullback–Leibler (KL) distance (or divergence) [22] between P and Q, namely $KL(P, Q)$. As a statistical measure, it is a measure of the distance between two probability densities P and Q. In the case of Gaussian multivariate distributions, this distance is completely defined by the correlation matrices of the whole system. Thus, it can be interpreted as how a multivariate probability distribution represented by the matrix Q is different from a reference multivariate probability distribution represented by the matrix P. It is also well-known that it is not a metric, as this measure is not symmetric and does not satisfy the triangle inequality. Given matrices P and Q, the $KL(P, Q)$ measure is:

$$KL(P, Q) = \frac{1}{2} \left[\log \left(\frac{|Q|}{|P|} \right) + Tr(Q^{-1}P) - N \right], \tag{3}$$

where the operator $|\cdot|$ computes the determinant of a matrix and $Tr(\cdot)$ is the trace of a square matrix, which is the sum of its diagonal elements. In our approach, $Dist(A(\delta), A_t(\delta)) = KL(A(\delta), A_t(\delta))$.

Note that assuming the identity function as the function of the graph's weights, i.e., $w_{ij} = \rho_{ij}$, $i \neq j$, the matrix $A(\delta)$ corresponds, in fact, to the unconditional correlation matrix where, given δ, $a_{ij}(\delta) = \rho_{ij} \neq 0$, $i \neq j$, if and only if $|\rho_{ij}| > \delta$, $a_{ij}(\delta) = 0$, $i \neq j$, otherwise (i.e., $|\rho_{ij}| \leq \delta$), and further imposing $a_{ij}(\delta) = 1$, $i = 1, \ldots, N$. Denote by $\widehat{Z}$ the matrix $A(\delta)$ defined so far. Under the above hypothesis on the graph's weights, $A_t(\delta)$ can also be considered to be a conditional correlation matrix at time t. Hence, in this special case, in order to preserve the strong relationships among specific groups of assets, in the maximization of the modified likelihood function (2) we are, in fact, requiring that $a_{ij}(\delta)$ and $a_{ij,t}(\delta)$ be as close as possible, that is, the corresponding correlation values $|\rho_{ij}|$ and $|\rho_{ij,t}|$ be as close as possible. In the rest of the paper, we assume $w_{ij} = \rho_{ij}$, $i \neq j$.

2.1. The BEKK Model

Ding and Engle [23] introduced the *diagonal* BEKK model, which is:

$$H_t = C \cdot C' + \sum_{k=1}^{K} \sum_{i=1}^{p} A_{ki} \cdot (\epsilon_{t-i} \epsilon'_{t-i}) \cdot A'_{ki} + \sum_{k=1}^{K} \sum_{j=1}^{q} B_{kj} \cdot H_{t-j} \cdot B'_{kj}, \tag{4}$$

where A_{ki} and B_{kj} are parameter matrices, C is a lower triangular matrix, p and q represent the lagged error term and the number of conditional covariance lags, respectively. K determines the generality of the process. We assume that $p = q = 1$ and $K = 1$, so that the diagonal BEKK model can be written in a compact form as:

$$H_t = C \cdot C' + A \cdot (\epsilon_{t-1} \epsilon'_{t-1}) \cdot A' + B \cdot H_{t-1} \cdot B', \tag{5}$$

with A and B diagonal matrices. Positive definiteness of conditional covariance matrices is guaranteed by construction (see [23]). The procedure used in estimating the parameters of the model is the maximization of a likelihood function constructed under the assumption of an i.i.d. of the errors η_t. Under the further assumption of conditional normality, the set of all parameters θ of the multivariate diagonal BEKK model can be estimated by maximizing the following sample log-likelihood function:

$$L(\theta) = -\frac{TN}{2} \log(2\pi) - \frac{1}{2} \sum_{t=1}^{T} \left(\log |H_t| + \epsilon'_t H_t^{-1} \epsilon_t \right), \tag{6}$$

with T the number of returns observations. Note that in the BEKK model, Equation (5) refers to the conditional covariance matrices. Thus, in the following modified log-likelihood function we have to use covariance matrices in place of the correlation matrices. Hence, under the assumption $w_{ij} = \rho_{ij}$, $i \neq j$, given δ, we first compute matrix $\widehat{Z}$ and then the corresponding unconditional covariance matrix $\widehat{\Sigma}$ with respect to the whole sample period T, that is, $\widehat{\Sigma} = \Gamma \widehat{Z} \Gamma$, with Γ the diagonal matrix of standard deviations with respect to the

whole sample period. Finally, in the Kullback–Leibler divergence measure of Formula (7) we compute the difference between matrices $\widehat{\Sigma}$ and H_t. The modified log-likelihood function is:

$$
\begin{aligned}
L(\theta) \quad &= -\frac{TN}{2}\log(2\pi) - \frac{1}{2}\sum_{t=1}^{T}\left(\log|H_t| + \epsilon_t'H_t^{-1}\epsilon_t\right) - KL(\widehat{\Sigma}, H_t) \\[2mm]
&= -\frac{TN}{2}\log(2\pi) - \frac{1}{2}\sum_{t=1}^{T}\left(\log|H_t| + \epsilon_t'H_t^{-1}\epsilon_t\right) - \frac{1}{2}\sum_{t=1}^{T}\left(\log(\tfrac{|H_t|}{|\widehat{\Sigma}|}) + Tr(H_t^{-1}\widehat{\Sigma}) - N\right) \\[2mm]
&= -\frac{1}{2}\left[N(T\log(2\pi)) + \sum_{t=1}^{T}\left(2\log(|H_t|) + \epsilon_t'H_t^{-1}\epsilon_t - \log(|\widehat{\Sigma}|) + Tr(H_t^{-1}\widehat{\Sigma})\right)\right].
\end{aligned}
\tag{7}
$$

Observe that the minimization of the distance between the target unconditional covariance matrix $\widehat{\Sigma}$ and H_t forces the values $\sigma_{ij,t}$ to be as close as possible to the values σ_{ij}, that is, forces the values $\rho_{ij,t}$ and $i \neq j$ to be as close as possible to the values ρ_{ij}.

2.2. The DCC Model

The idea of this model, introduced and analyzed in [12], is that the conditional covariance matrix $\boldsymbol{H_t}$ can be decomposed into the conditional standard deviations $\boldsymbol{D_t}$ of each of the N series and a conditional correlation matrix of the returns $\boldsymbol{R_t}$. The dynamic of the model is described by Equation (1) and:

$$
\boldsymbol{H_t} = \boldsymbol{D_t R_t D_t}.
\tag{8}
$$

The matrix $\boldsymbol{D_t}$ is a diagonal matrix and consists of the N univariate GARCH models. Because it is a diagonal non-negative matrix with all diagonal elements positive, $\boldsymbol{D_t}$ is positive definite. To ensure that $\boldsymbol{H_t}$ is positive definite, it is necessary that the matrix $\boldsymbol{R_t}$ is positive definite with the additional constraint that all its elements must be equal to or less than 1 by definition. The dynamic of correlation matrix $\boldsymbol{R_t}$ is derived from another matrix $\boldsymbol{Q_t}$ of the form:

$$
\boldsymbol{R_t} = \bar{\boldsymbol{Q}}_t^{-1}\boldsymbol{Q}_t\bar{\boldsymbol{Q}}_t^{-1},
\tag{9}
$$

where $\bar{\boldsymbol{Q}}_t = diag(\boldsymbol{Q_t})$ with $diag(\cdot)$ is the diagonal of a square matrix. The form of $\boldsymbol{Q_t}$ determines the dynamic of the model and its complexity (see, e.g., [13,24]). For example, following [13] we have:

$$
\boldsymbol{Q_t} = (1 - \theta_1 - \theta_2)\hat{\boldsymbol{Q}}_t + \theta_1(\epsilon_{t-1}\epsilon_{t-1}') + \theta_2\boldsymbol{Q_{t-1}},
\tag{10}
$$

with $\hat{\boldsymbol{Q}}_t = Cov[\epsilon_t\epsilon_t'] = E[\epsilon_t\epsilon_t']$. In order to ensure that $\boldsymbol{R_t}$ is positive definite, the parameters θ_1 and θ_2 must satisfy:

$$
\theta_1 \geq 0, \ \theta_2 \geq 0, \ \theta_1 + \theta_2 < 1.
$$

The parameter estimation phase is rather difficult, and hence for the DCC model a two-stage estimation procedure is provided. In the first stage, the parameters of the univariate GARCH models are estimated for each asset series. In the second stage, a second set of parameters are estimated given the parameters found in the previous phase. Referring to the dynamics described in (10) and assuming multivariate Gaussian distributed errors, after the first step only the parameters θ_1 and θ_2 are unknown so they are estimated in the second stage. In this second phase, the log-likelihood function is:

$$
L(\theta) = -\frac{1}{2}\sum_{t=1}^{T}\left(\log(|R_t|) + \epsilon_t'R_t^{-1}\epsilon_t\right).
\tag{11}
$$

In our approach, we are interested in the second stage of the process where the log-likelihood function takes into account the correlation matrix of the assets returns at time t.

Hence, under the assumption $w_{ij} = \rho_{ij}$, $i \neq j$, given δ, we consider the matrix $\widehat{Z}$ and the matrix R_t. The corresponding modified log-likelihood function is:

$$
\begin{aligned}
L(\theta) &= -\frac{1}{2} \sum_{t=1}^{T} \left(\log(|R_t|) + \epsilon_t' R_t^{-1} \epsilon_t \right) - KL(\widehat{Z}, R_t) \\
&= -\frac{1}{2} \Big[\sum_{t=1}^{T} \left(2\log(|R_t|) + \epsilon_t' R_t^{-1} \epsilon_t - \log(|\widehat{Z}|) + Tr(R_t^{-1} \widehat{Z}) - N \right) \Big].
\end{aligned}
\tag{12}
$$

As for the modified BEKK model, to highlight the strong relationships among groups of assets in the market, the minimization of the distance between the (target) matrix $\widehat{Z}$ and R_t forces the values $\rho_{ij,t}$ and $i \neq j$ to be as close as possible to the values ρ_{ij}.

To conclude this section, we observe that if we set $\delta = 0$, and under the assumption that $w_{ij} = \rho_{ij}$, $i \neq j$, the correlation matrix $\widehat{Z}$ and the unconditional covariance matrix $\widehat{\Sigma}$ correspond exactly to the long-run correlation and covariance matrices used in the model provided in [16]. Hence, on the one hand, our method can be precisely considered as an alternative targeting method with respect to the one proposed in [16]. On the other hand, assuming $\delta > 0$, with our method we are, in fact, "forcing" our modified MGARCH models to find clusters of highly correlated assets in the simulated series.

3. Data Analysis and Test

In order to demonstrate the effectiveness of our new estimation approach, which aims at correctly identifying highly (positive or negative) correlated groups of assets given a financial dataset, we first performed a preprocessing phase by a method of clustering financial time series (see [25]). Time series clustering creates groups based on similarities between time series without using pre-existing categorization labels [26]. Various approaches were proposed in time series clustering, and various authors attempted to synthesize the literature on the field (see, e.g., [27–29]). According to [29], there is a wide range of applications for time series clustering that can be found in a variety of domains (e.g., economic or medical time series analysis, visual processing, and anomalous activity detection).

Time series clustering is particularly relevant in finance. There is a particular interest in discovering typical dynamics of financial markets and the impact of different shocks on time series and portfolio allocation groups. In this respect, according to [30] clustering of financial time series is essential to determine how much wealth should be allocated to financial assets and opportunities. Therefore, financial time series need to be clustered to select an appropriate portfolio and analyze an economic system. When studying highly complicated phenomena such as financial time series, one has to deal with substantial heterogeneity and peculiar characteristics and features (see [31,32]). Therefore, robustness measures should address this challenge [33]. Moreover, studying their behavior over time in a dynamic framework is also essential, as these systems are associated with uncertainty and for this reason, other approaches representing this uncertainty were proposed in [34].

In this section, we provide a statistical framework for classifying time series and an example of applying the proposed method to a group of time series. Thus, our whole approach is based on two stages: In the first stage, we explicitly define the different groups of time series using a clustering procedure, which is helpful for identification of the different groups. In the second stage, we apply our new estimation method to correctly recognize the groups obtained so far. In our example, we perform the first stage on the following set of stocks: Facebook, Apple, Google, Boeing, Microsoft, Amazon, General Motors, Goldman Sachs, JPMorgan Chase, Intel, Verizon Communications, Visa, Cisco, Coca-Cola Company, and Salesforce. They were selected without any particular pattern from the stocks listed on the NYSE, the New York Stock Exchange. The period considered is from 1 January 2020 to 1 January 2021, and the resulting 253 observations refer to the daily closing prices of each financial time series.

Following [29], we have to clarify what is the claim in what two data are similar, and in what way we can obtain a good similarity and dissimilarity measure between

two different observations (in this case, two different financial time series). These are the central questions of cluster analysis. In this regard, we have explicitly considered a hierarchical clustering algorithm that uses a Pearson correlation-based distance and the complete method (see [35]) as a helpful approach to distinguish the different groups that we can identify. Concerning financial time series, Pearson's correlation coefficient is widely used in the related financial literature to quantify the degree of similarity or dissimilarity between two time series [30]. If we denote y_t and x_t as two different time series, we can consider their correlations [29]:

$$\text{COR}(x_t, y_t) = 1 - \frac{\sum_{t=1}^{T}(y_t - \overline{x}_t)(y_t - \overline{y}_t)}{\sqrt{\sum_{t=1}^{t}(x_t - \overline{x}_t)^2}\sqrt{\sum_{t=1}^{t}(y_t - \overline{y}_t)^2}}, \tag{13}$$

Then, we can compute the following distance, and we have:

$$d_{Cor}(x_t, y_t) = 1 - COR(x_t, y_t) \tag{14}$$

We apply the distance and cluster approach to develop the dendrogram and interpret the different clusters (see Figure 1). Then, we follow an exploratory approach for identification of the groups of different stocks that are included in each cluster. We use the dendrogram to collect information about the variability of a cluster structure and the overall group's structure.
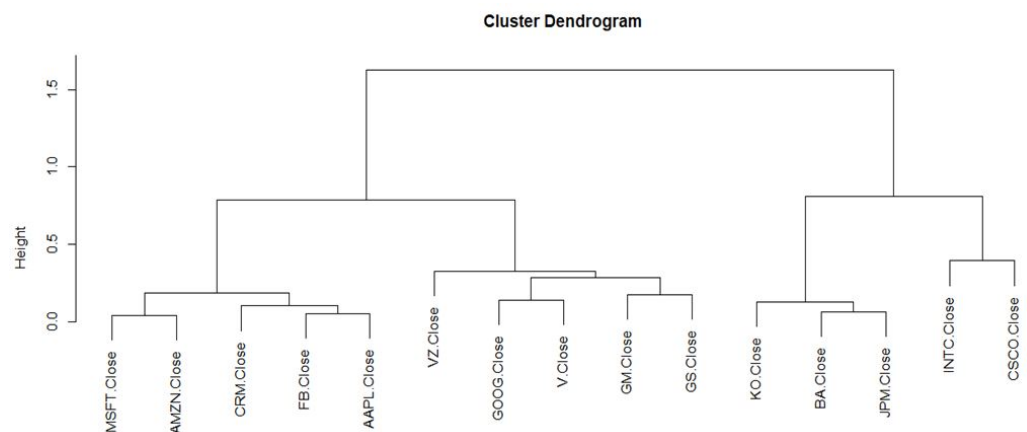

Figure 1. Clusters obtained in the first phase.

No prior knowledge is required for this approach, and it allows for the understanding of patterns in financial time series data without the need for additional information [36]. This approach can be summarized as follows: use similarity measures to identify subgroups (see [37]) and determine which generic grouping each individual belongs to. Through the above clustering procedure, we can identify four general clusters (see Figure 1). First, we can identify Microsoft's most significant joint movement (MSFT) and Amazon (AMZN). At the same time, Facebook (FB), Apple (AAPL), and finally, Salesforce (CRM) belong to the same cluster. This cluster is clearly based on the technological innovations that can have an overarching impact on different sectors and companies (see also [34]). Finally, at the same time, Coca Cola (KO) and, more importantly, Boeing (BA) and JPMorgan Chase (JPM) are in a different cluster. Furthermore, Intel (INTC) and Cisco (CSCO) can also be considered part of the same cluster. We can further observe the critical role of economic similarities for these two stocks in this case. In this respect, it is possible to suppose relevant economic shocks that affect both financial time series. Finally, in the third cluster, we can observe a slightly different behavior for Verizon Communications (VZ), but also daily solid joint movements for Google (GOOG) and Visa (V) as well as for General Motors (GM) and Goldman Sachs (GS). Overall, through observation and visual analysis of the dendrogram, the distinct

clusters obtained will be used as a reference for the evaluation of the effectiveness of our new estimation approach (second stage).

In the following, we present some scenarios characterized by a different number N of time series considered with the aim of evaluating the differences between the simulated series obtained by the BEKK and DCC models and their modified versions. All MGARCH models have been implemented in Matlab R2018 version using the MFE Toolbox code repository by Kevin Sheppard [38]. In the resulting correlation graphs, we detected all the maximal cliques using the Bron–Kerbosch algorithm [39]. The experiments were conducted on a PC equipped with an Intel Core i7-3632MQ processor with 2.20 Ghz.

Scenario 1

In this first scenario, we consider only the group of technological innovation assets formed by MSFT, AMZN, CRM, FB, and AAPL (see Figure 1). We compute the 252 log-return values and the corresponding unconditional correlation matrix, which is (see Table 1):

Table 1. The correlation matrix for $N = 5$ assets.

	MSFT	AMZN	CRM	FB	AAPL
MSFT	1	0.7712	0.7690	0.6855	0.6914
AMZN	0.7712	1	0.8435	0.7028	0.6458
CRM	0.7690	0.8435	1	0.7413	0.7508
FB	0.6855	0.7028	0.7413	1	0.6121
AAPL	0.6914	0.6458	0.7508	0.6121	1

Let us now consider the (sub)graph $G(\delta)$ of the complete correlation graph G whose vertex set contains only the subset of pairs of assets such that $|\rho_{ij}| > \delta$, with $\delta = 0.5$ (see Figure 2). In the figure, the numbers associated with the vertices correspond to the following labeling: 1, MSFT; 2, AMZN; 3, CRM; 4, FB; and 5, AAPL.
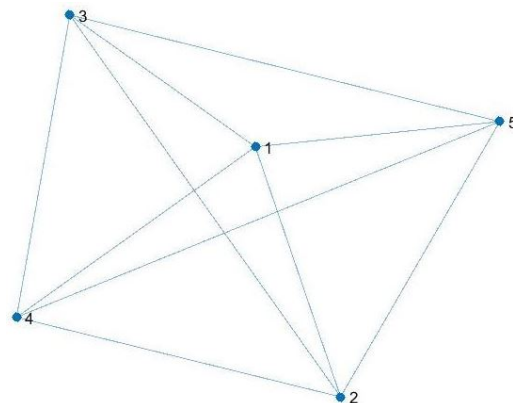

Figure 2. The graph $G(\delta)$ with $\delta = 0.5$. Vertex number refers to the associated asset.

Note that the graph $G(\delta)$ corresponds to the complete graph G, because for all pairs of assets i and j we have $\rho_{ij} > 0.5$.

Considering the whole original dataset consisting of $T = 252$ return observations, we first estimate the parameters of the BEKK and DCC models and then the parameters of each of the modified versions of the two models using the log-likelihood functions (7) and (12), respectively. Then, using these sets of parameters we simulate (new) $T = 252$ log-return observations of the five assets. Our aim is to verify whether our modified models allow to better capture the strong relationships between assets in the stock market. In other words, we want to compare the graph $G(\delta)$ with the simulated graphs $G_S(\delta)$ (see Figure 3).
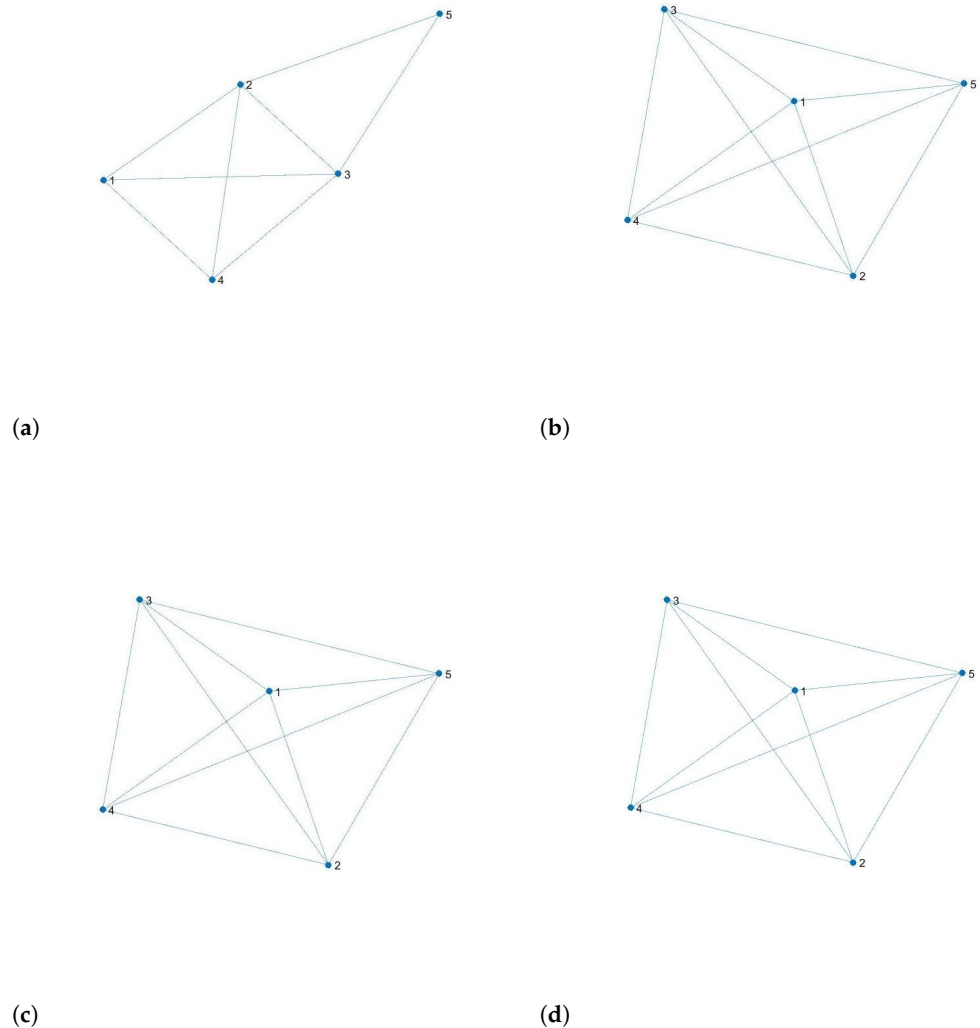
(**a**)   (**b**)

(**c**)   (**d**)

Figure 3. The simulated graphs $G_S(\delta)$ with $\delta = 0.5$. (**a**) The graph $G_S(\delta)$ resulting from the BEKK model; (**b**) the graph $G_S(\delta)$ resulting from the modified BEKK model; (**c**) the graph $G_S(\delta)$ resulting from the DCC model; (**d**) the graph $G_S(\delta)$ resulting from the modified DCC model. Vertex number refers to the associated asset.

Comparing the five graphs in Figures 2 and 3, only the original BEKK model was not able to correctly reproduce the original series. In this case, vertex 5, corresponding to asset AAPL, is not connected to all the other vertices despite its correlation value with other assets is greater that 0.5. Observe that for the DCC approach the modification brings no gain. This can occur when the series' dimension is small, and, in this specific case, the modification has no effect. In addition to the graphs $G_S(\delta)$, in Table 2 we report the Frobenius distance between the fitted covariance matrices H_t with respect to the unconditional covariance matrix $\widehat{\Sigma}$, and the Kullback–Leibler divergence between the matrix $\widehat{\Sigma}$ and the unconditional covariance matrix of the new simulated series of log-returns, denoted by $\widehat{\Sigma}_S$, with respect to all the considered models. The Frobenius norm is:

$$F = \sqrt{\sum_{t=1}^{T} Tr[(H_t - \widehat{\Sigma})'(H_t - \widehat{\Sigma})]}, \tag{15}$$

and the Kullback–Leibler divergence is:

$$KL = \frac{1}{2}\Big[\log\big(\frac{|\widehat{\Sigma}_S|}{|\widehat{\Sigma}|}\big) + Tr(\widehat{\Sigma}_S^{-1}\widehat{\Sigma}) - N\Big]. \tag{16}$$

Table 2. Values of the Frobenius and the Kullback–Leibler loss functions: case $N = 5$ and $\delta = 0.5$.

	BEKK	Modified BEKK	DCC	Modified DCC
F	0.0669	0.0555	0.0881	0.0870
KL	0.0855	0.0229	0.0162	0.0154

Both functions F and KL measure loss, so that lower values are preferable. We note that the values of the modified MGARCH models are better than the corresponding original models. Here, we are not interested in comparing the values of the two loss functions between all the models, as, again, it is not our goal to find a single winner.

Assume now that we are interested in detecting groups of assets with a higher value of correlation. On the basis of the values in Table 1, we set $\delta = 0.71$. The corresponding graph $G(\delta)$ is (see Figure 4):
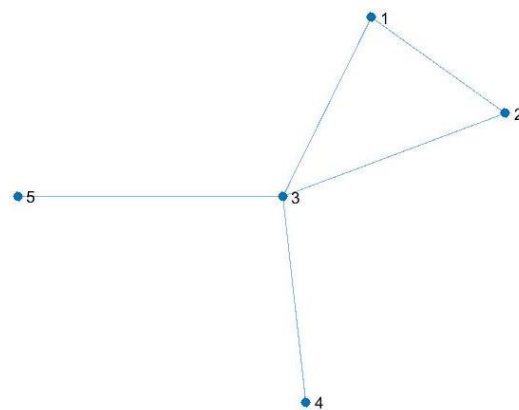

Figure 4. The graph $G(\delta)$ with $\delta = 0.71$. Vertex number refers to the associated asset.

We observe that with this value of δ, only assets $\{1, 2, 3\}$ form a clique. The goal is to find simulated graphs $G_S(\delta)$, with $\delta = 0.71$, as similar as possible to the subgraph $G(\delta)$ of Figure 4. The results are reported in Figure 5 along with the corresponding table (see Table 3) reporting the values of the two loss functions considered.

The simulated graph of the modified BEKK model exactly replicates $G(\delta)$. The two graphs associated with the DCC and modified DCC models correctly find the clique $C = \{1, 2, 3\}$, but the DCC model determines a new clique $C' = \{1, 2, 3, 4\}$ that contains C, whereas the modified DCC model introduces the additional edge $(4, 2)$. Thus, these two models "overestimate" the correlation between assets pointing out correlations that do not really exist. The original BEKK model has the lowest ability to correctly identify groups of highly correlated assets. On the other hand, the values of the loss functions reported in Table 3 are in line with the values in Table 2.
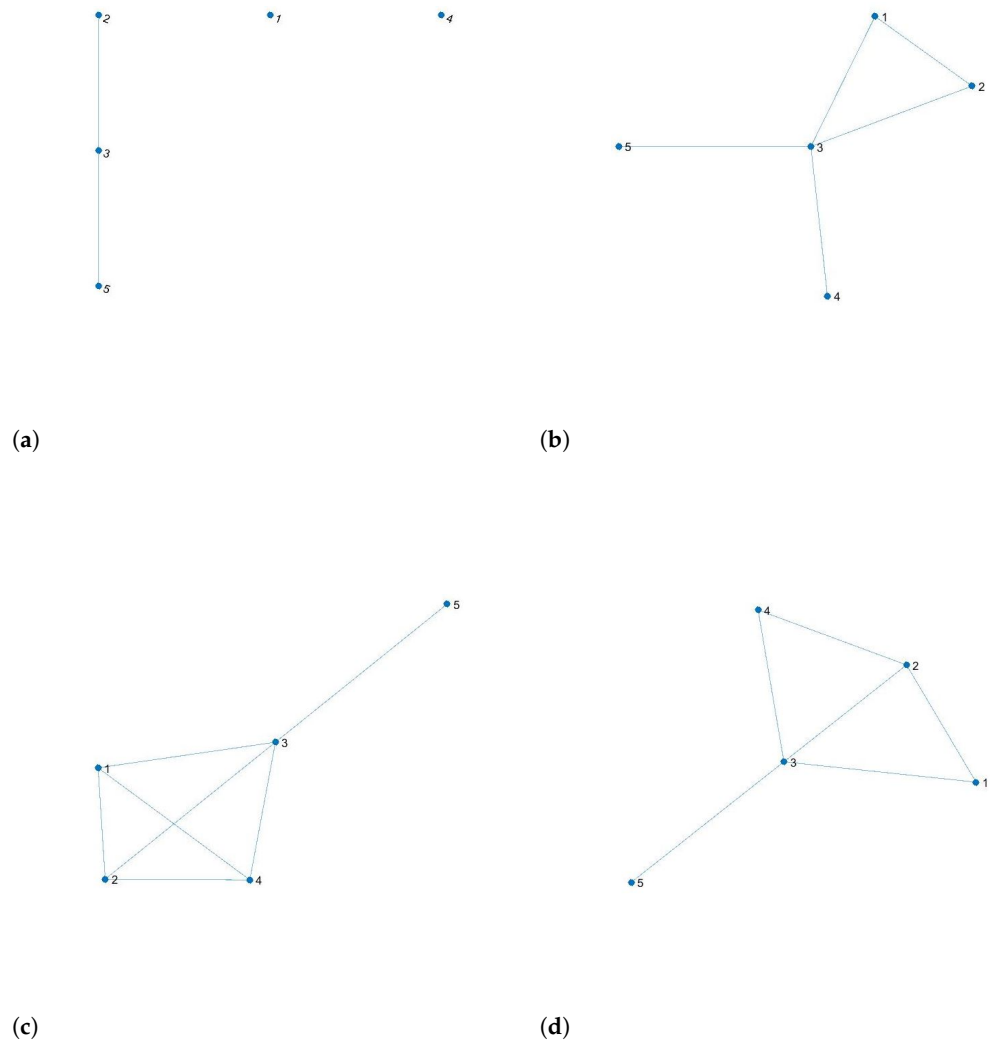
(**a**)   (**b**)

(**c**)   (**d**)

Figure 5. The simulated graphs $G_S(\delta)$ with $\delta = 0.71$. (**a**) The graph $G_S(\delta)$ resulting from the BEKK model; (**b**) the graph $G_S(\delta)$ resulting from the modified BEKK model; (**c**) the graph $G_S(\delta)$ resulting from the DCC model; (**d**) the graph $G_S(\delta)$ resulting from the modified DCC model. Vertex number refers to the associated asset.

Table 3. Values of the Frobenius and the Kullback–Leibler loss functions: case $N = 5$ and $\delta = 0.71$.

	BEKK	**Modified BEKK**	**DCC**	**Modified DCC**
F	0.0669	0.0575	0.0881	0.0840
KL	0.0976	0.0269	0.0306	0.0262

Scenario 2

In this second scenario, we consider $N = 8$ assets along with the corresponding correlation matrix reported in Table 4. The set of stocks is formed by the five previously considered assets and the group formed by KO, BA, and JPM (see Figure 1). In the corresponding graphs, the numbers associated with the vertices now correspond to the following labeling: 1, MSFT; 2, AMZN; 3, CRM; 4, FB; 5, AAPL; 6, KO; 7, BA; 8, JPM. We choose this new group of three stocks because the two clusters are adequately separated

from each other, so that we expect finding two well-separated groups of assets in the following correlation graphs. The unconditional correlation matrix is:

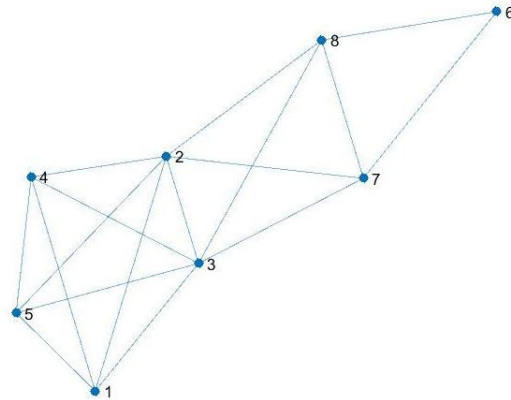Consider $\delta = 0.5$, the graph $G(\delta)$ is (see Figure 6):

Figure 6. The graph $G(\delta)$ with $\delta = 0.5$ for $N = 8$ assets. Vertex number refers to the associated asset.

In $G(\delta)$, we clearly detect the three maximal cliques $C_1 = \{1, 2, 3, 4, 5\}$; $C_2 = \{6, 7, 8\}$; and $C_3 = \{2, 3, 7, 8\}$. In particular, cliques C_1 and C_2 refer to the two clusters of assets in Figure 1. The simulated graphs $G_S(\delta)$ are presented in Figure 7.

In Table 5, we report the values of the loss functions F and KL and the additional information on the number of maximal cliques in each of the simulated graphs $G_S(\delta)$.

The graphs $G_S(\delta)$ related to the two modified MGARCH models are exactly the same as the graph $G(\delta)$. It is worth noting that this does not mean that the correlation matrices of the simulated returns series with the modified BEKK and DCC models are equal to the correlation matrix of the original series, but that both the modified models are able to detect the same (original) clusters of highly correlated assets. In actuality, assuming highly (positive) correlated assets remain unaltered over time, by forecasting future values of the series using the results provided by the modified BEKK and/or DCC models, one might expect to simulate returns series that do not alter too much the correlation structure among assets as well as the strong relationships among stocks. This can be useful in a portfolio selection problem from a diversification viewpoint. It is well-known that correlation represents the degree of relationship between the price movements of different assets included in the portfolio for diversified portfolios. Thus, choosing pairs of assets less correlated decreases the portfolio's overall risk. Consider, for example, the graph $G_S(\delta)$ in Figure 7c referred to the original DCC model, which is very similar to the graph $G(\delta)$ (in fact there is just one additional edge in $G_S(\delta)$, that is, edge $(2, 6)$). It includes the maximal clique $C = \{2, 6, 7, 8\}$, demonstrating that asset 2 now shows a correlation value with asset 6 greater than 0.5 (precisely, equal to 0.536), whereas from Table 4 we have $\rho_{26} = 0.4810$. Hence, the original DCC model inserts a strong correlation between these two assets, which is not truthful, and, without reason, this might prevent from choosing assets 2 and 6 in an optimal portfolio. Finally, note that these considerations are difficult to obtain by examination of only the values of the loss functions, which demonstrates that the modified models behave better than the original models.
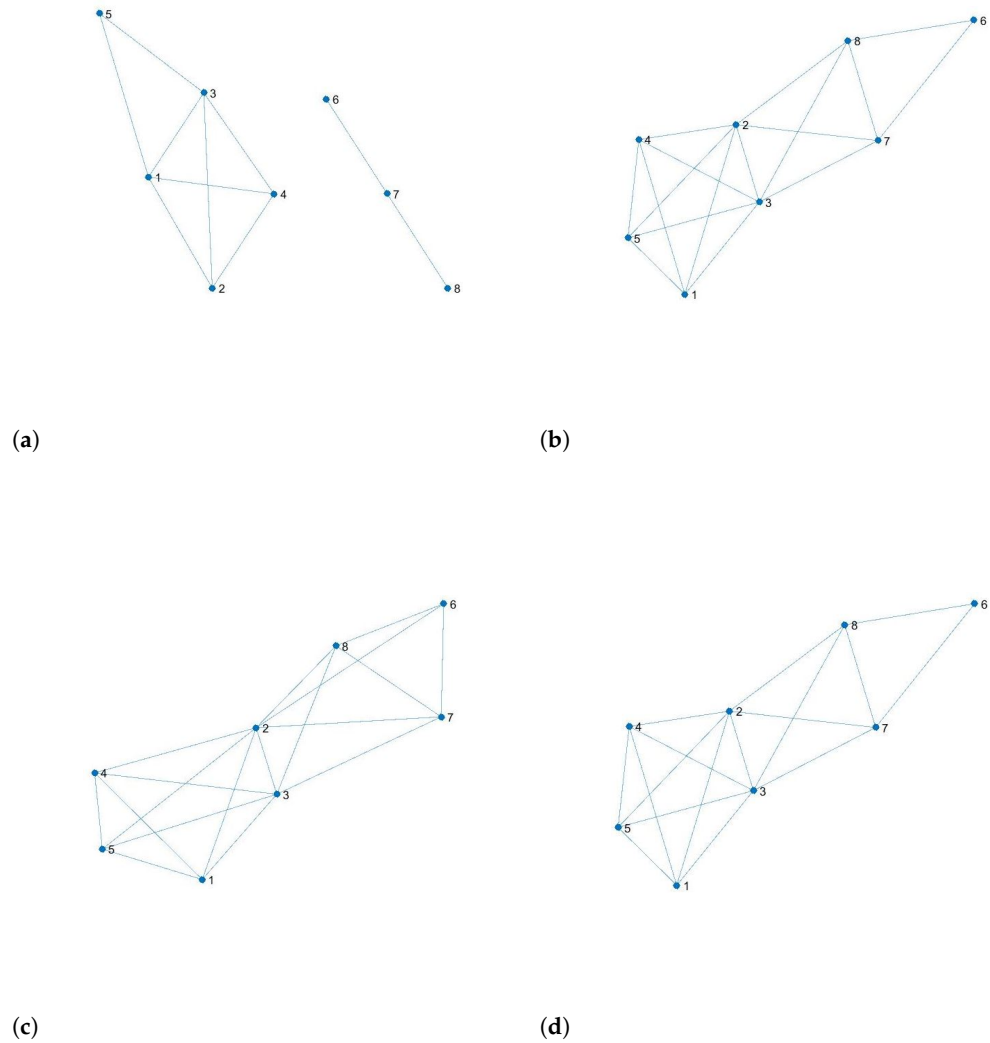
(a)

(b)

(c)

(d)

Figure 7. The simulated graphs $G_S(\delta)$ with $\delta = 0.5$ for $N = 8$ assets. (**a**) The graph $G_S(\delta)$ resulting from the BEKK model; (**b**) the graph $G_S(\delta)$ resulting from the modified BEKK model; (**c**) the graph $G_S(\delta)$ resulting from the DCC model; (**d**) the graph $G_S(\delta)$ resulting from the modified DCC model. Vertex number refers to the associated asset.

Table 4. The correlation matrix for $N = 8$ assets.

	MSFT	AMZN	CRM	FB	AAPL	KO	BA	JPM
MSFT	1	0.7712	0.7690	0.6855	0.6914	0.4038	0.4834	0.4387
AMZN	0.7712	1	0.8435	0.7028	0.6458	0.4810	0.5412	0.5580
CRM	0.7690	0.8435	1	0.7413	0.7508	0.4798	0.5912	0.5838
FB	0.6855	0.7028	0.7413	1	0.6121	0.2451	0.2697	0.3162
AAPL	0.6914	0.6458	0.7508	0.6121	1	0.3995	0.4562	0.4472
KO	0.4038	0.4810	0.4798	0.2451	0.3995	1	0.7211	0.6461
BA	0.4834	0.5412	0.5912	0.2697	0.4562	0.7211	1	0.7304
JPM	0.4387	0.5580	0.5838	0.3162	0.4472	0.6461	0.7304	1

Table 5. Values of the Frobenius and the Kullback–Leibler loss functions and the maximal cliques: case $N = 8$ and $\delta = 0.5$.

	BEKK	Modified BEKK	DCC	Modified DCC
F	0.0669	0.0575	0.0881	0.0840
KL	0.0976	0.0269	0.0306	0.0262
Maximal Cliques	{1,2,3,4}, {1,3,5}, {6,7}, {7,8}	{1,2,3,4,5}, {2,3,7,8}, {6,7,8}	{1,2,3,4,5}, {2,3,7,8}, {2,6,7,8}	{1,2,3,4,5}, {2,3,7,8}, {6,7,8}

Finally, observe that the graph $G_S(\delta)$, related to the original BEKK model (see Figure 7a), reports a very distorted relationship between the assets in the market. Hence, on the basis of the two above scenarios, we can state that it seems that the BEKK model benefits more effectively from the introduction of the modified log-likelihood function.

Scenario 3

In this third scenario, we consider all $N = 15$ assets described in Section 3. In the graphs, the new labels associated with each asset are (see Figure 1): 1, MSFT; 2, AMZN; 3, CRM; 4, FB; 5, AAPL; 6, VZ; 7, GOOG; 8, V; 9, GM; 10, GS; 11, KO; 12, BA; 13, JPM; 14, INTC; 15, CSCO. The corresponding correlation matrix is in Table 6:

Table 6. The correlation matrix for $N = 15$ assets.

	MSFT	AMZN	CRM	FB	AAPL	VZ	GOOG	V	GM	GS	KO	BA	JPM	INTC	CSCO
MSFT	1	0.7712	0.7690	0.6855	0.6914	0.8067	0.4671	0.5619	0.6656	0.4566	0.4038	0.4834	0.4387	0.5493	0.5777
AMZN	0.7712	1	0.8435	0.7028	0.6458	0.7581	0.4539	0.6225	0.7123	0.5523	0.4810	0.5412	0.5580	0.6333	0.6596
CRM	0.7690	0.8435	1	0.7413	0.7508	0.8558	0.4849	0.6484	0.7830	0.6080	0.4798	0.5912	0.5838	0.7187	0.7171
FB	0.6855	0.7028	0.7413	1	0.6121	0.6837	0.2288	0.3825	0.4519	0.3955	0.2451	0.2697	0.3162	0.5065	0.4991
AAPL	0.6914	0.6458	0.7508	0.6121	1	0.6846	0.4191	0.5118	0.6353	0.4143	0.3995	0.4562	0.4472	0.5170	0.5358
VZ	0.8067	0.7581	0.8558	0.6837	0.6846	1	0.5379	0.6644	0.7903	0.5877	0.5163	0.6020	0.6172	0.6535	0.6838
GOOG	0.4671	0.4539	0.4849	0.2288	0.4191	0.5379	1	0.7488	0.7014	0.4135	0.6972	0.7590	0.5881	0.4913	0.5131
V	0.5619	0.6225	0.6484	0.3825	0.5118	0.6644	0.7488	1	0.7604	0.5899	0.6963	0.8907	0.7001	0.5981	0.6526
GM	0.6656	0.7123	0.7830	0.4519	0.6353	0.7903	0.7014	0.7604	1	0.6280	0.6541	0.7773	0.7363	0.6404	0.7295
GS	0.4566	0.5523	0.6080	0.3955	0.4143	0.5877	0.4135	0.5899	0.6280	1	0.4366	0.5975	0.7133	0.5698	0.6544
KO	0.4038	0.4810	0.4798	0.2451	0.3995	0.5163	0.6972	0.6963	0.6541	0.4366	1	0.7211	0.6461	0.4908	0.5061
BA	0.4834	0.5412	0.5912	0.2697	0.4562	0.6020	0.7590	0.8907	0.7773	0.5975	0.7211	1	0.7304	0.5663	0.6145
JPM	0.4387	0.5580	0.5838	0.3162	0.4472	0.6172	0.5881	0.7001	0.7363	0.7133	0.6461	0.7304	1	0.5542	0.5984
INTC	0.5493	0.6333	0.7187	0.5065	0.5170	0.6535	0.4913	0.5981	0.6404	0.5698	0.4908	0.5663	0.5542	1	0.6700
CSCO	0.5777	0.6596	0.7171	0.4991	0.5358	0.6838	0.5131	0.6526	0.7295	0.6544	0.5061	0.6145	0.5984	0.6700	1

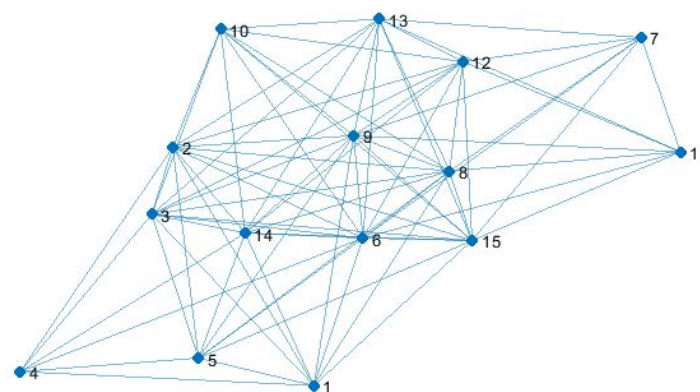The correlation graph for $\delta = 0.5$ is depicted in Figure 8.

Figure 8. The graph $G(\delta)$ with $\delta = 0.5$ for $N = 15$ assets. Vertex number refers to the associated asset.

The maximal cliques in $G(\delta)$ with $\delta = 0.5$ are: $C_1 = \{1,2,3,4,5,6,14\}$; $C_2 = \{1,2,3,5,6,8,9,14,15\}$; $C_3 = \{2,3,6,8,9,10,12,13,14,15\}$; and $C_4 = \{6,7,8,9,11,12,13,15\}$. These cliques show the most interconnected groups of assets at a correlation level greater

than 0.5. In the following, we report a table of the values of the loss functions F and KL and the information on the maximal cliques in each of the simulated graphs $G_S(\delta)$. More precisely, in Table 7 we report in bold the cliques of the simulated graphs that coincide to the ones of the original graph $G(\delta)$. The modified BEKK model seems to be the best performing both with respect to the loss function values and the number of the maximal cliques of $G(\delta)$. We also note that the first clique in the graph, which refers to the modified BEKK model, has the same number of nodes of clique C_1 except with vertex 15 in place of vertex 14. We observe that the correlation values of the first six assets in Table 6 with respect to vertices 14 and 15 are very close to each other. Thus, the difference in those two cliques might be due to numerical reasons in the estimation and simulation phase of the modified BEKK model. In any case, all models introduce correlation values, and a larger number of quite different cliques, which do not appear in the graph $G(\delta)$.

Table 7. Values of the Frobenius and the Kullback–Leibler loss functions and the maximal cliques: case $N = 15$ and $\delta = 0.5$. In bold the cliques of the simulated graphs that coincide to the ones of the original graph $G(\delta)$.

	BEKK	Modified BEKK	DCC	Modified DCC
F	0.1708	0.1682	0.1995	0.2028
KL	0.4035	0.3002	0.3309	0.3091
Maximal Cliques	{1,2,3,4,5,6}, {7,8,9,11,12,13}, **{1,2,3,5,6,8,9,14,15}**, {1,2,3,6,8,9,10,14,15}, **{2,3,6,8,9,10,12,13,14,15}**, {7,8,9,12,13,14,15}	{1,2,3,4,5,6,15}, **{1,2,3,5,6,8,9,14,15}**, **{2,3,6,8,9,10,12,13,14,15}**, **{6,7,8,9,11,12,13,15}**, {6,8,9,11,12,13,14,15}	{1,2,3,4,5,6}, {1,2,3,5,6,8,9}, {2,3,6,8,9,10,15}, {3,6,8,9,10,13,15}, {3,8,9,10,12,13,15}, {3,8,9,10,12,14,15}, {7,8,9,11,12,13}, {2,8,9,11,14,15}, {8,9,11,12,13,15}, {8,9,11,12,14,15}	{1,2,3,4,6}, {1,3,4,5,6}, {1,3,5,6,9}, {1,2,3,6,8,9,10,13,14,15}, {1,3,6,8,9,10,12,13,14,15}, {1,6,7,8,9,12,13}, {7,8,9,11,12,13}

To evaluate the effectiveness of our method in finding groups of strongly correlated assets we consider a higher value for the threshold, that is, $\delta = 0.65$. The corresponding correlation graph $G(\delta)$ is in Figure 9:
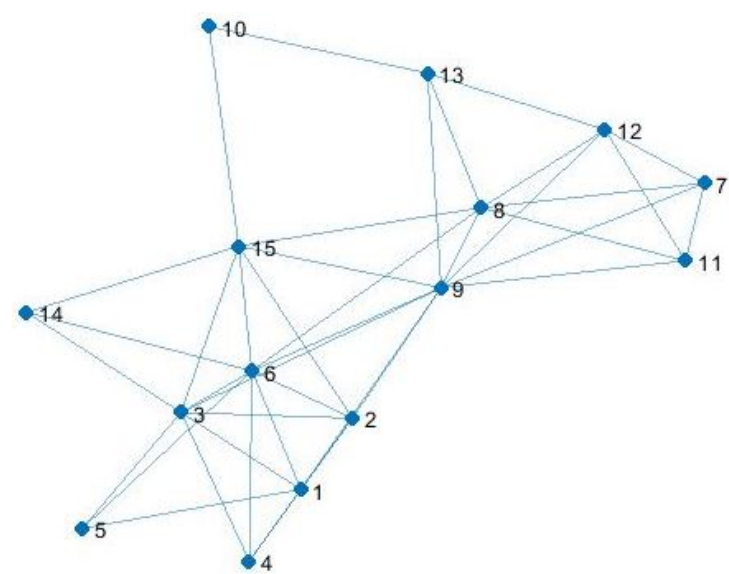

Figure 9. The graph $G(\delta)$ with $\delta = 0.65$ for $N = 15$ assets. Vertex number refers to the associated asset.

As expected, the graph $G(\delta)$, with $\delta = 0.65$, is sparser than the graph $G(\delta)$, with $\delta = 0.5$. In the new graph, we identify a larger number of cliques, but of cardinality less than the cardinality of the maximal cliques when $\delta = 0.5$.

In Table 8, we report in bold the cliques of the simulated graphs that coincide to those in $G(\delta)$ with $\delta = 0.65$. It is evident that when N grows, the results are less clear-cut with respect to the two previous examples. On the one hand, this can be due to the fact that the log-likelihood functions to optimize are highly nonlinear functions, and it is difficult to find provably optimal solutions. Local solutions may be problematic, and this creates difficulties in the estimation of the models. Additionally, because in the optimization phase matrices have to be inverted in each iteration, it makes the overall computation demanding unless N is small. In any case, Tables 7 and 8 still highlight that the modified MGARCH models performs slightly better than the original ones, in particular, when δ increases. For example, for $\delta = 0.65$ the simulated series obtained with the modified BEKK and DCC models are still able to better capture the strong relationships among assets in the market than the corresponding original models.

Table 8. Values of the Frobenius and Kullback–Leibler loss functions and the maximal cliques: case $N = 15$ and $\delta = 0.65$. In bold the cliques of the simulated graphs that coincide to the ones of the original graph $G(\delta)$.

	BEKK	Modified BEKK	DCC	Modified DCC
F	0.1708	0.1659	0.1995	0.2038
KL	0.3721	0.2982	0.2627	0.2433
Maximal Cliques	**{1,2,3,4,6}**, {1,2,3,5,6}, {2,3,5,6,9}, {2,3,6,9,14}, {3,6,9,14,15}, **{7,8,9,11,12}**, {8,9,11,12,13}, {6,8,9}, {9,10,13}	**{1,2,3,4,6}**, **{1,2,3,6,9}**, {1,2,3,5,6}, {2,3,6,9,14,15}, {3,6,8,9,14,15}, **{7,8,9,11,12}**, {8,9,11,12,13}, {8,9,12,14}, {9,10,12,13}, {9,10,15}	**{1,2,3,4,6}**, {2,3,6,9}, {3,5}, {3,14}, {7,8,9,12}, {7,8,11,12}, **{10,13}**	**{1,2,3,4,6}**, {1,2,3,5,6}, **{1,2,3,6,9}**, {7,8,9,12}, **{8,9,12,13}**, **{10,13}**, {7,8,11,12}, {3,15}

From the above results, we can state that the application of our method to empirical case studies is encouraging, particularly when the number of assets is not large and we are able to find (global) optimal solutions in the maximization of the log-likelihood objective functions. In fact, in these cases, we observed a significant improvement in the ability of the modified MGARCH models to replicate the correlation structure of the assets in the market compared with the original models. When the number of assets increases, the estimation of the models involves somewhat heavy computations because they contain a large number of parameters, and there is no guarantee of finding provably global optimal solutions but only local optimal solutions. This undermines the ability of the estimated models to replicate the correlation structure of the assets correctly. Despite this, even when the number of assets is large, the modified MGARCH models seem to perform better than the original ones. Overall, our experiments demonstrate that the BEKK model benefits most from the modification of the log-likelihood function.

4. Conclusions and Further Research

In this paper, we advance a method for improving the estimation phase of financial time series with the aim of improving the analysis and evaluation of portfolios of financial assets whose performance strictly depends on the correlation among assets. Several different estimation models have been proposed in the literature; among others, the family of models known as multivariate general auto regressive conditional heteroscedasticity (MGARCH) models are the most widely used. These models consider that a financial time series suffers from heteroscedasticity. This paper considers two such models, namely the BEKK and DCC models, where we modified the log-likelihood objective function assuming that there are specific groups of assets that are highly correlated in a financial market,

and these relationships remain unaltered over time. Hence, in the log-likelihood function we introduced a term referring to a loss measure computed on the difference between the time-varying covariance/correlation matrices and the covariance/correlation matrix estimated with respect to the whole in-sample period. Given the set of the estimated parameters, we use them for simulating new time series to evaluate the effectiveness of our modified estimation phase. We also propose a new approach for the evaluation of the results based on network analysis and, more precisely, on detecting maximal cliques in correlation graphs.

On the basis of the results reported in Section 3, we cannot state whether the modified models always outperform the original ones, and much more experiments are needed. Therefore, this leaves plenty of opportunity for further research. On the one hand, one can experiment with other loss functions to introduce in the log-likelihood objective functions to improve the estimates. In fact, we observe that our approach is extremely flexible. Different loss functions can be considered without imposing additional constraints on covariance or correlation matrices. On the other hand, much attention should be paid in order to improve the optimization phase. This involves developing ad hoc strategies for finding global optimal solutions or solutions close to the optimal ones. In this regard, the development of metaheuristic procedures could be a further line of research to improve the optimization phase. Finally, finding new network indexes that better highlight other peculiar aspects of a financial market is worth investigating.

Author Contributions: Methodology, C.D. and A.S.; Software, A.S.; Data curation, C.D.; Writing—original draft, A.S.; Funding acquisition, A.S. All authors have read and agreed to the published version of the manuscript.

Data Availability Statement: The data presented in this study are available on request from the corresponding author.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Silvennoinen, A.; Terasvirta, T. Consistency and asymptotic normality of maximum likelihood estimators of a multiplicative time-varying smooth transition correlation GARCH model. *Econom. Stat.* 2021, *in press*. [CrossRef]
2. Engle, R.; Colacito, R. Testing and valuing dynamic correlations for asset allocation. *J. Bus. Econ. Stat.* **2006**, *24*, 238–253. [CrossRef]
3. Hsiao, C.W.; Chan, Y.C.; Lee, M.Y.; Lu, H.P. Heteroscedasticity and Precise Estimation Model Approach for Complex Financial Time-Series Data: An Example of Taiwan Stock Index Futures before and during COVID-19. *Mathematics* **2021**, *9*, 2719. [CrossRef]
4. Lei, B.; Zhang, B.; Song, Y. Volatility forecasting for high-frequency financial data based on web search index and deep learning model. *Mathematics* **2021**, *9*, 320. [CrossRef]
5. Posedel Simovic, P.; Tafro, A. Pricing the Volatility Risk Premium with a Discrete Stochastic Volatility Model. *Mathematics* **2021**, *9*, 2038. [CrossRef]
6. Rombouts, J.; Stentoft, L.; Violante, F. The value of multivariate model sophistication: An application to pricing Dow Jones Industrial Average options. *Int. J. Forecast.* **2014**, *30*, 78–98. [CrossRef]
7. Gargallo, P.; Lample, L.; Miguel, J.A.; Salvador, M. Co-movements between EU ETS and the energy markets: A VAR-DCC-GARCH approach. *Mathematics* **2021**, *9*, 1787. [CrossRef]
8. Billio, M., Caporin, M. Market linkages, variance spillovers, and correlation stability: Empirical evidence of financial contagion. *Comput. Stat. Data Anal.* **2010**, *54*, 2443–2458. [CrossRef]
9. Chang, C.L.; Liu, C.P.; McAleer, M. Volatility spillovers for spot, futures, and ETF prices in agriculture and energy. *Energy Econ.* **2019**, *81*, 779–792. [CrossRef]
10. Mata, M.N.; Razali, M.N.; Bentes, S.R.; Vieira, I. Volatility Spillovers Effect of Pan Asia's Property Portfolio Markets. *Mathematics* **2021**, *9*, 1418. [CrossRef]
11. Bauwens, L.; Laurent, S.; Rombouts, J.V. Multivariate GARCH models: A survey. *J. Appl. Econom.* **2006**, *21*, 79–109. [CrossRef]
12. Engle, R.F.; Sheppard, K. *Theoretical and Empirical Properties of Dynamic Conditional Correlation Multivariate GARCH*; NBER Working Papers, No. 8554; 2001, National Bureau of Economic Research (NBER.org). Available online: https://www.nber.org/papers/w8554 (accessed on 4 February 2022).

13. Engle, R.F. Dynamic conditional correlation: A simple class of multivariate generalized autoregressive conditional heteroskedasticity models. *J. Bus. Econ. Stat.* **2002**, *20*, 339–350. [CrossRef]
14. Bollerslev, T.; Patton, A.J.; Quaedvlieg, R. Modeling and forecasting (un)reliable realized covariances for more reliable financial decisions. *J. Econom.* **2018**, *207*, 71–91. [CrossRef]
15. Engle, R.; Kelly, B. Dynamic equicorrelation. *J. Bus. Econ. Stat.* **2012**, *30*, 212–228. [CrossRef]
16. Caporin, M.; McAleer, M. Do we really need both BEKK and DCC? A tale of two multivariate GARCH models. *J. Econ. Surv.* **2012**, *26*, 736–751. [CrossRef]
17. Aielli, G.P. Dynamic Conditional Correlations: On Properties and Estimation. 2009. Available online: http://ssrn.com/abstract=1507743 (accessed on 2 February 2022).
18. Stoyanov, S.V.; Rachev, S.T.; Racheva-Yotova, B.; Fabozzi, F.J. Fat-tailed models for risk estimation. *J. Portf. Manag.* **2011**, *37*, 107–117. [CrossRef]
19. Mantegna, R.N. Hierarchical structure in financial markets. *Eur. Phys. J. B-Condens. Matter Complex Syst.* **1999**, *11*, 193–197. [CrossRef]
20. Bollerslev, T.; Wooldridge, J.M. Quasi-maximum likelihood estimation and inference in dynamic models with time-varying covariances. *Econom. Rev.* **1992**, *11*, 143–172. [CrossRef]
21. Bondy, A.; Murty, U.S.R. *Graph Theory*; Springer: New York, NY, USA, 2008.
22. Kullback, S.; Leibler, R.A. On information and sufficiency. *Ann. Math. Stat.* **1951**, *22*, 79–86. [CrossRef]
23. Ding, Z.; Engle, R. Large scale conditional covariance modelling, estimation and testing. Academia *Econ. Pap.* **2001**, *29*, 157–184.
24. Tse, Y.K.; Tsui, A.K.C. A multivariate generalized autoregressive conditional heteroscedasticity model with time-varying correlations. *J. Bus. Econ. Stat.* **2002**, *20*, 351–362. [CrossRef]
25. Sarda-Espinosa, A. Comparing time-series clustering algorithms in r using the dtwclust package. *R Package Vignette* **2017**, *12*, 41.
26. Radhimeenakshi, D.S.; Latha, K. Similarity Measure Selection for Clustering Stock Market Time Series Databases. *Int. J. Eng. Sci. Comput.* **2017**, *7*, 13879–13882.
27. Aghabozorgi, S.; Shirkhorshidi, A.S.; Wah, T.Y. Time-series clustering—A decade review. *Inf. Syst.* **2015**, *53*, 16–38. [CrossRef]
28. Liao, T.W. Clustering of time series data—A survey. *Pattern Recognit.* **2005**, *38*, 1857–1874. [CrossRef]
29. Montero, P.; Vilar, J.A. TSclust: An R Package for Time Series Clustering. *J. Stat. Softw.* **2014**, *62*, 1–43. [CrossRef]
30. Piccardi, C.; Calatroni, L.; Bertoni, F. Clustering financial time series by network community analysis. *Int. J. Mod. Phys. C* **2011**, *22*, 35–50. [CrossRef]
31. Sewell, M. *Characterization of Financial Time Series*; Rn, 11(01), 01. Working Paper; 2011, UCL Department of Computer Science, University College London. Available online: http://www.cs.ucl.ac.uk/fileadmin/UCL-CS/images/Research_Student_Information/RN_11_01.pdf (accessed on 2 December 2021).
32. Shi, Y.; Li, B.; Du, G.; Dai, W. Clustering framework based on multi-scale analysis of intraday financial time series. *Phys. A Stat. Mech. Appl.* **2021**, *567*, 125728. [CrossRef]
33. D'Urso, P.; De Giovanni, L.; Massari, R. GARCH-based robust clustering of time series. *Fuzzy Sets Syst.* **2016**, *305*, 1–28. [CrossRef]
34. Drago C.; Lauro C.; Scepi G. Visualization and Analysis of Large Datasets by Beanplot PCA. In *Advances in Latent Variables*; Carpita, M., Brentari, E., Qannari, E.M., Eds.; Vita e Pensiero; Università Cattolica del S. Cuore: Milano, Italy, 2013. Available online: https://www.researchgate.net/profile/Carlo-Drago (accessed on 30 June 2013).
35. Everitt B.S.; Landau S.; Leese, M. *Cluster Analysis*; Arnold: London, UK, 2001; ISBN 0-340-76119-9.
36. Ahn, J.; Lee, J.H. Clustering Method for Financial Time Series with Co-Movement Relationship. In Proceedings of the 2018 3rd International Conference on Computational Intelligence and Applications (ICCIA), Hong Kong, China, 28–30 July 2018; pp. 260–264.
37. Wang, X.; Smith, K.; Hyndman, R. Characteristic-based clustering for time series data. *Data Min. Knowl. Discov.* **2006**, *13*, 335–364. [CrossRef]
38. Sheppard, K. MFE Toolbox. 2013. Avaiable online: https://www.kevinsheppard.com/MFE_Toolbox (accessed on 16 August 2018).
39. Bron, C.; Kerbosch, J. Algorithm 457: Finding all cliques of an undirected graph. *Commun. ACM* **1973**, *16*, 575–577. [CrossRef]