Ballestra *et al. Financial Innovation* (2025) 11:37
https://doi.org/10.1186/s40854-024-00683-7

Financial Innovation

RESEARCH Open Access

Multivariate GARCH models with spherical parameterizations: an oil price application

Luca Vincenzo Ballestra[1], Riccardo De Blasis[2*] and Graziella Pacelli[2]

*Correspondence:
r.deblasis@univpm.it

[1] Department of Statistical Sciences "Paolo Fortunati", Alma Mater Studiorum University of Bologna, Bologna, Italy
[2] Department of Management, Marche Polytechnic University, Ancona, Italy

Abstract

In popular Baba-Engle-Kraft-Kroner (BEKK) and dynamic conditional correlation (DCC) multivariate generalized autoregressive conditional heteroskedasticity models, the large number of parameters and the requirement of positive definiteness of the covariance and correlation matrices pose some difficulties during the estimation process. To avoid these issues, we propose two modifications to the BEKK and DCC models that employ two spherical parameterizations applied to the Cholesky decompositions of the covariance and correlation matrices. In their full specifications, the introduced Cholesky-BEKK and Cholesky-DCC models allow for a reduction in the number of parameters compared with their traditional counterparts. Moreover, the application of spherical transformation does not require the imposition of inequality constraints on the parameters during the estimation. An application to two crude oils, WTI and Brent, and the main exchange rate prices demonstrates that the Cholesky-BEKK and Cholesky-DCC models can capture the dynamics of covariances and correlations. In addition, the Kupiec test on different portfolio compositions confirms the satisfactory performance of the proposed models.

Keywords: BEKK, Cholesky-GARCH, Crude oils, DCC, Exchange rates, Spherical parameterization

Introduction

Over the past three decades, multivariate modeling of volatility has gained significant interest from researchers. It is commonly understood that financial volatilities exhibit interdependent behavior across markets and over time. Therefore, most studies address this issue in a multivariate setting instead of working with separate univariate specifications. In particular, with multivariate generalized autoregressive conditional heteroskedasticity (MGARCH) models, researchers can study the relationships between the volatilities of different markets (see, e.g., Kearney and Patton 2000; Karolyi 1995) or the dynamics of the correlations over time (see, e.g., Bollerslev 1990; Longin and Solnik 2001).

A general specification for the multivariate GARCH model was initially proposed by Bollerslev et al. (1988), commonly known as the VEC model, in which the authors directly model the covariance matrix over time. However, owing to the large number of parameters, this model is not easily applicable beyond the bivariate case. Many other specifications have been proposed (see, e.g., Kroner and Ng 1998; Vrontos et al. 2003); however, the most commonly employed in financial applications are the Baba, Engle, Kraft, and Kroner (BEKK) and dynamic conditional correlation (DCC) models from Engle and Kroner (1995) and Engle (2002), respectively. The BEKK model can be considered as a particular case of the general VEC specification. By contrast, the DCC model is considered a nonlinear combination of univariate GARCH models. However, both models have two disadvantages. On the one hand, in their full specifications, they require a large number of parameters for estimation. On the other hand, they impose some restrictions on the parameters to guarantee the positive definiteness of the covariance and correlation matrices.

In this study, we introduce two modifications to the BEKK and DCC models, employing two different spherical parameterizations of the covariance and correlation matrices, which can reduce the number of parameters and relax the restrictions that guarantee the requirement of positive definiteness of such matrices. Specifically, to extend the BEKK model we use the spherical parameterization of the covariance matrix by Pinheiro and Bates (1996), while to extend the DCC model we employ the spherical parameterization of the correlation matrix by Rapisarda et al. (2007) and Pourahmadi and Wang (2015). Both parameterizations are applied to the Cholesky factorization of the covariance or correlation matrix. Moreover, under simple and common assumptions, the spherical transformations are unique (i.e., they provide a one-to-one mapping from the covariance or correlation matrix to some angles and vice versa).

A similar approach, although limited to correlation matrices, has been initially introduced by Pedeli et al. (2015), who applied a spherical parameterization of the Cholesky factor of the correlation matrices. The authors indicated that the angles derived from the transformation can be modeled in terms of covariates. Unfortunately, they did not propose a specific model in this regard. Moreover, in their application, they considered a time-invariant correlation matrix that is in line with the constant conditional correlation (CCC) model of Bollerslev (1990) and is considered a special case of the DCC model. By contrast, inspired by Bernardi and Catania (2018), we propose an updating equation for the angles that allows us to deal with time-varying correlation or covariance matrices. In the spirit of Pedeli et al. (2015), we call the proposed models the Cholesky-BEKK and Cholesky-DCC models.

Therefore, our contributions are twofold: (i) we extend the standard BEKK and DCC models by means of spherical transformations that, in their full specification, allow for a reduction in the number of parameters to be estimated compared with the standard models; and (ii) we propose an updating equation for time-varying angles to allow for dynamic conditional correlations or covariance matrices.

An application to crude oils and daily exchange rate prices shows the potential of the two proposed approaches and demonstrates that the Cholesky-BEKK and Cholesky-DCC models can capture the dynamics of covariances and correlations. Moreover, the Kupiec test applied to three different portfolio combinations shows the satisfactory performance of the Cholesky multivariate models.

The remainder of this paper is organized as follows. In "Multivariate GARCH models" section we review the BEKK and DCC models, while in "Spherical parameterization of covariance and correlation matrices" section we present the spherical parameterization of the covariance and correlation matrices. "The proposed methodology" section introduces the proposed methodologies, followed by a simulation study in "A simulation study" and "Application" sections presents the application to real data with the results. Finally, "Conclusions" section concludes the paper.

Multivariate GARCH models

We model the log returns time series with the following vector process, $\boldsymbol{r}_t \in \mathbb{R}^n$,

$$\boldsymbol{r}_t = \boldsymbol{m}_t + \boldsymbol{e}_t, \tag{1}$$

where $\boldsymbol{m}_t$ is a deterministic component (i.e., the conditional mean vector), and $\boldsymbol{e}_t$ is a white noise process following a multivariate normal distribution (i.e., $\boldsymbol{e}_t \sim \mathcal{N}(\mathbf{0}, \boldsymbol{H}_t)$, where $\boldsymbol{H}_t$ is the covariance matrix at time t).

Usually, a deterministic term is specified as a function of the past, using a vector autoregressive (VAR) model. If we consider a VAR of order one, we can write

$$\boldsymbol{m}_t = \boldsymbol{c} + \boldsymbol{G}\boldsymbol{r}_{t-1},$$

where $\boldsymbol{c}$ is a vector of constants and $\boldsymbol{G}$ is an $n \times n$ matrix of coefficients, n being the number of time series.

In the remainder of the discussion, we do not focus on the conditional mean vector but consider only the residuals, $\boldsymbol{e}_t$, which enter into the definition of the multivariate GARCH models. The latter focuses on modeling the conditional covariance matrix $\boldsymbol{H}_t$ or some of its derived quantities (e.g., the correlation matrix). Following Bauwens et al. (2006), we can distinguish three different approaches to building multivariate GARCH models: (i) direct generalizations of the univariate GARCH model by Bollerslev (1986), (ii) linear combinations of univariate GARCH models, and (iii) nonlinear combinations of univariate GARCH models. This study focuses on the BEKK and DCC specifications, which belong to the first and third groups, respectively.

The BEKK model

The BEKK model in its general form, proposed by Engle and Kroner (1995), can be written as

$$\boldsymbol{H}_t = \boldsymbol{C}\boldsymbol{C}' + \boldsymbol{A}(\boldsymbol{e}_{t-1}\boldsymbol{e}'_{t-1})\boldsymbol{A}' + \boldsymbol{B}\boldsymbol{H}_{t-1}\boldsymbol{B}', \tag{2}$$

where $\boldsymbol{C}, \boldsymbol{A}$, and $\boldsymbol{B}$ are $n \times n$ parameter matrices and $\boldsymbol{C}$ is lower triangular. The full model includes $2n^2 + n(n+1)/2$ parameters and the covariances are positive definite by construction. However, to guarantee an observationally equivalent structure, Engle and Kroner (1995) demonstrate that all elements of $\boldsymbol{A}$ and $\boldsymbol{B}$ and the diagonal elements of $\boldsymbol{C}$ must be positive. In addition, the authors show that covariance stationarity is guaranteed when all the eigenvalues of $\boldsymbol{A} + \boldsymbol{B}$ are less than one in modulus.

Because of the large number of parameters to be estimated, the disadvantage of this model is that it is only feasible for small values of n. Therefore, to allow for large cross-sectional dimensions, it is common to restrict the model's parameters as proposed by Ding and Engle (2001), who considered a scalar counterpart of (2):

$$H_t = CC' + a(e_{t-1}e'_{t-1}) + bH_{t-1}. \tag{3}$$

Moreover, to reduce the number of parameters further, it is common to impose a varianceタtargeting approach on the specification. The conditional covariance matrix can be expressed in terms of the unconditional covariance matrix and other parameters. Therefore, the scalar BEKK model with variance targeting becomes

$$H_t = (1 - a - b)\bar{H} + a(e_{t-1}e'_{t-1}) + bH_{t-1}, \tag{4}$$

where $\bar{H} = \frac{1}{T}\sum_{t=1}^{T} e_t e_t^T$ is an unconditional covariance matrix estimated from the full sample. In (4), the number of parameters is reduced to two; however, we must impose the inequality constraints $a + b < 1$ and $a, b > 0$ to guarantee positive definiteness and stationarity of the conditional covariances.

The scalar version of the model has the advantage of reducing the number of parameters; however, it has the limit of imposing the same dynamics on all elements of the covariance matrix.

The DCC model

The DCC model is a class of multivariate GARCH estimators introduced by Engle (2002) as a generalization of the CCC estimator proposed by Bollerslev (1990). The model focuses on the separate modeling of conditional variances and conditional correlation, assuming the following decomposition.

$$H_t = D_t R_t D_t, \tag{5}$$

where D_t and R_t are the diagonal matrix of the standard deviations and the conditional correlation matrix of the return residuals, e_t, respectively.

According to Bauwens et al. (2006), the DCC model belongs to the group of nonlinear combinations of univariate GARCH models. Thus, it employs the univariate specification for conditional volatilities and conditional correlation decomposition as follows.

$$\begin{aligned} D_t^2 &= diag\{\omega_i\} + diag\{\alpha_i\} \circ e_{t-1}e'_{t-1} + diag\{\beta_i\} \circ D_{t-1}^2, \\ R_t &= Q_t^{*-\frac{1}{2}} Q_t \, Q_t^{*-\frac{1}{2}}, \quad Q_t^* = diag(Q_t), \end{aligned} \tag{6}$$

where $\circ$ is the Hadamard product and $diag\{a_i\}$ is a matrix with elements a_i on the main diagonal.

With this approach, the DCC does not directly model the conditional correlation R_t, but rather the quantity Q_t with the following updating scheme.

$$Q_t = (\iota\iota' - A - B) \circ \bar{Q} + A \circ (\epsilon_{t-1}\epsilon'_{t-1}) + B \circ Q_{t-1}, \tag{7}$$

where A and B are symmetric parameter matrices, $\epsilon_t = D_t^{-1} e_t$ are the standardized residuals, ι is an n-dimension vector of ones, and $\bar{Q}$ is the unconditional covariance matrix of the standardized errors estimated as $\bar{Q} = \frac{1}{T}\sum_{t=1}^{T} \epsilon_t \epsilon_t^T$. This full specification, initially proposed by Engle (2002), includes variance targeting. Moreover, to guarantee the positive semidefiniteness of the matrix Q, Ding and Engle (2001) show that A, B, and $(\iota\iota' - A - B)$ must be positive semidefinite.

The DCC specification can be expressed in scalar form as

$$\boldsymbol{Q}_t = (1 - a - b)\bar{\mathbf{Q}} + a(\boldsymbol{\epsilon}_{t-1}\boldsymbol{\epsilon}'_{t-1}) + b\mathbf{Q}_{t-1}, \tag{8}$$

where the number of parameters to be estimated is reduced to two. Similar to the scalar BEKK specification, the scalar DCC requires the imposition of inequality constraints $a + b < 1$ and $a, b > 0$ to guarantee the positive definiteness and stationarity of the correlation matrices.

The use of variance targeting in both models allows for a significant reduction in the number of parameters to be estimated. However, as reported by Caporin and McAleer (2012), the application of this approach to the DCC specification is counterintuitive because the sample correlation is not a consistent estimator of the long-run $\boldsymbol{Q}_t$ owing to the correlation decomposition in (6). Therefore, we refer to this approach as *approximate* targeting.

Spherical parameterization of covariance and correlation matrices

As seen in the previous section, one of the main challenges in modeling the volatility of multivariate time series is to guarantee the positive definiteness of the covariance matrix $\boldsymbol{H}_t$ or the correlation matrix $\boldsymbol{R}_t$, and simultaneously reduce the number of parameters to be estimated. Many authors have proposed several solutions to this issue, mainly by applying some parameterizations to those matrices. For example, Pinheiro and Bates (1996) introduced five different parameterizations for the covariance matrix to ensure positive definiteness, of which the spherical one produced the best combination of performance and interpretability for the individual parameters. Later, Rebonato and Jackel (1999) applied a similar approach to the correlation matrix by employing spherical and spectral decompositions. The correlation approach has been further studied by other authors (see, e.g., Rapisarda et al. 2007; Pourahmadi and Wang 2015).

In this study, we consider two spherical parameterizations, one for the covariance matrix and the other for the correlation matrix. In particular, because the BEKK specification directly models covariance matrices, we extend it using the spherical parameterization of the covariance matrix. On the contrary, as the DCC directly models the correlations, for such a model, we employ the spherical parameterization of the correlations. Both transformations allow us to guarantee positive definiteness requirements and reduce the number of parameters. In the following sections, we first introduce the two parameterizations, and then demonstrate how to extend both models.

Spherical parameterization of the covariance matrix

Following Pinheiro and Bates (1996), we first apply Cholesky decomposition to the covariance matrix

$$\boldsymbol{H}_t = \boldsymbol{L}_t\boldsymbol{L}'_t, \tag{9}$$

where $\boldsymbol{L}_t$ is a lower triangular matrix. Then, we apply parameterization to the matrix $\boldsymbol{L}_t$ as follows.

$$l_{ij,t} = \begin{cases} \theta_{i1,t}\cos\theta_{i2,t}, \quad i=1,\dots,n, \quad j=1 \\ \theta_{i1,t}\cos\theta_{ij+1,t}\prod_{k=2}^{j}\sin\theta_{ik,t}, \quad i=3,\dots,n,\ j=2,\dots,i-1\ , \\ \theta_{i1,t}\prod_{k=2}^{j}\sin\theta_{ik,t}, \quad i=2,\dots,n,\ j=i \end{cases} \tag{10}$$

where $\theta_{ij,t}$ are some angles.

Therefore, by omitting the time dependency, the matrix $\boldsymbol{L}_t$ can be parameterized as

$$\begin{pmatrix} \theta_{11}\cos\theta_{12} & 0 & 0 & \cdots & 0 \\ \theta_{21}\cos\theta_{22} & \theta_{21}\sin\theta_{22} & 0 & \cdots & 0 \\ \theta_{31}\cos\theta_{32} & \theta_{31}\sin\theta_{32}\cos\theta_{33} & \theta_{31}\sin\theta_{32}\sin\theta_{33} & \cdots & 0 \\ \theta_{41}\cos\theta_{42} & \theta_{41}\sin\theta_{42}\cos\theta_{43} & \theta_{41}\sin\theta_{42}\sin\theta_{43}\cos\theta_{44} & \cdots & 0 \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ \theta_{n1}\cos\theta_{n2} & \theta_{n1}\sin\theta_{n2}\cos\theta_{n3} & \theta_{n1}\sin\theta_{n2}\sin\theta_{n3}\cos\theta_{n4} & \cdots & \theta_{n1}\prod_{k=2}^{n}\sin\theta_{nk} \end{pmatrix}. \tag{11}$$

To ensure the uniqueness of the parameterization we must constrain the angles

$$\begin{aligned} &\theta_{i1,t} > 0, \quad i=1,\dots,n \\ &\theta_{ij,t} \in (0,\pi), \quad i=2,\dots,n,\ j=2,\dots,i. \end{aligned} \tag{12}$$

Thus, we can define the inverse of the parameterization

$$\begin{aligned} \theta_{i1,t} &= \sqrt{h_{ii,t}} = \sqrt{\sum_{k=1}^{i} l_{ik,t}^2}, \quad i=1,\dots,n \\ \theta_{i2,t} &= \arccos\left(\frac{l_{i1,t}}{\theta_{i1,t}}\right), \quad i=1,\dots,n \\ \theta_{ij,t} &= \arccos\left(\frac{l_{ij-1,t}}{\theta_{i1,t}\prod_{k=2}^{j-1}\sin\theta_{ik,t}}\right), \quad i=3,\dots,n,\ j=3,\dots,i \end{aligned} \tag{13}$$

to obtain a lower triangular matrix of angles $\boldsymbol{\Theta}_t = (\theta_{ij,t})$, with $n(n+1)/2$ entries,

$$\boldsymbol{\Theta}_t = \begin{pmatrix} \theta_{11,t} & 0 & 0 & 0 & \cdots & 0 \\ \theta_{21,t} & \theta_{22,t} & 0 & 0 & \cdots & 0 \\ \theta_{31,t} & \theta_{32,t} & \theta_{33,t} & 0 & \cdots & 0 \\ \theta_{41,t} & \theta_{42,t} & \theta_{43,t} & \theta_{44,t} & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \cdots & \vdots \\ \theta_{n1,t} & \theta_{n2,t} & \theta_{n3,t} & \theta_{n4,t} & \cdots & \theta_{nn,t} \end{pmatrix}, \tag{14}$$

which is later employed in the updating function of the multivariate BEKK model.

Spherical parameterization of the correlation matrix

When considering the correlation matrix R_t, we can apply Cholesky decomposition to obtain

$$\boldsymbol{R}_t = \boldsymbol{L}_t\boldsymbol{L}_t',$$

where $\boldsymbol{L}_t$ is a lower triangular matrix. Following Rapisarda et al. (2007), the entries of this matrix can be parameterized as

$$l_{ij,t} = \begin{cases} 1, \quad i,j=1 \\ \cos\theta_{i1,t}, \quad i=2,\ldots,n \\ \cos\theta_{ij,t}\prod_{k=1}^{j-1}\sin\theta_{ik,t}, \quad i=3,\ldots,n,\ j=2,\ldots,i-1 \\ \prod_{k=1}^{j-1}\sin\theta_{ik,t}, \quad i=2,\ldots,n,\ j=i \end{cases}, \tag{15}$$

where $\theta_{ij,t}$ are some angles in $(0,\pi)$. Similar to the parameterization of the covariance matrix, the restriction to the interval $(0,\pi)$ ensures the uniqueness of the transformation.

By omitting the time dependency, the matrix $\boldsymbol{L}_t$ can be written as

$$\begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ \cos\theta_{21} & \sin\theta_{21} & 0 & \cdots & 0 \\ \cos\theta_{31} & \cos\theta_{32}\sin\theta_{31} & \sin\theta_{32}\sin\theta_{31} & \cdots & 0 \\ \cos\theta_{41} & \cos\theta_{42}\sin\theta_{41} & \cos\theta_{43}\sin\theta_{42}\sin\theta_{41} & \cdots & 0 \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ \cos\theta_{n1} & \cos\theta_{n2}\sin\theta_{n1} & \cos\theta_{n3}\sin\theta_{n2}\sin\theta_{n1} & \cdots & \prod_{k=1}^{n-1}\sin\theta_{nk} \end{pmatrix}. \tag{16}$$

Although the formula in (15) allows us to compute l_{ij} from the values of θ_{ij} at each time, it is possible to describe the direct transformation from $\boldsymbol{\Theta} = (\theta_{ij})$ to $\boldsymbol{R}$, which is a lower triangular matrix with zeros on the main diagonal and $n(n-1)/2$ angles,

$$r_{ij}(\theta_{ij}) = c_{i1}c_{j1} + \sum_{k=2}^{i-1} c_{ik}c_{jk}\prod_{l=1}^{k-1} s_{il}s_{jl} + c_{ji}\prod_{l=1}^{i-1} s_{il}s_{jl}, \quad 1 \le i \le j \le n, \tag{17}$$

where $c_{ij} = \cos\theta_{ij}$ and $s_{ij} = \sin\theta_{ij}$.

Moreover, the transformation from $\boldsymbol{L}$ to $\boldsymbol{\Theta} = (\theta_{ij})$ can be expressed as

$$\begin{aligned} \theta_{i1,t} &= \arccos(l_{i1,t}), && 2 \le i \le n, \\ \theta_{ij,t} &= \arccos\left(\frac{l_{ij,t}}{\prod_{k=1}^{j-1}\sin(\theta_{ik,t})}\right), && 2 \le j < i \le n, \end{aligned} \tag{18}$$

and the matrix of angles becomes

$$\boldsymbol{\Theta}_t = \begin{pmatrix} 0 & 0 & 0 & 0 & \cdots & 0 \\ \theta_{21,t} & 0 & 0 & 0 & \cdots & 0 \\ \theta_{31,t} & \theta_{32,t} & 0 & 0 & \cdots & 0 \\ \theta_{41,t} & \theta_{42,t} & \theta_{43,t} & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \cdots & \vdots \\ \theta_{n1,t} & \theta_{n2,t} & \theta_{n3,t} & \theta_{n4,t} & \cdots & 0 \end{pmatrix}. \tag{19}$$

The proposed methodology

The spherical transformations introduced in the previous sections have been employed in financial market applications to guarantee the positive definiteness of covariance or correlation matrices. Examples include Creal et al. (2011) and Bernardi and Catania (2018), who employ the spherical transformation of the correlation matrix in the generalized autoregressive score framework, and Zhang et al. (2015), who study a joint mean-variance correlation model with covariates applied to longitudinal data. Finally, Pedeli et al. (2015) applied the spherical parameterization of the correlation matrix to extend the constant conditional correlation model of Bollerslev (1990).

In this section, inspired by Pedeli et al. (2015), who assume a time-invariant correlation matrix, we propose two extensions of the BEKK and DCC models by applying spherical parameterizations of the Cholesky decomposition of the covariance matrix $\boldsymbol{H}_t$ and correlation matrix $\boldsymbol{R}_t$, respectively. Therefore, we define Cholesky-BEKK (CH-BEKK) as an extension of BEKK and Cholesky-DCC (CH-DCC) as an extension of DCC.

The Cholesky-BEKK model

Departing from the BEKK specification in (2), which directly models the conditional covariance matrix, we employ the spherical parameterization of the lower triangular matrix from the Cholesky decomposition, as in (13). Thus, because parameterization ensures a unique relationship between the covariances and angles $\boldsymbol{\Theta}_t$, instead of modeling the covariances, we define an updating function for the angles.

Following Bernardi and Catania (2018), we assume an updating equation for the angles $\boldsymbol{\Theta}_t = (\theta_{ij,t})$ as

$$\boldsymbol{\Theta}_t = (\boldsymbol{u} - \boldsymbol{A} - \boldsymbol{B}) \circ \Lambda^{-1}(\bar{\boldsymbol{H}}) + \boldsymbol{A} \circ \boldsymbol{\Xi}_{t-1} + \boldsymbol{B} \circ \boldsymbol{\Theta}_{t-1}, \tag{20}$$

where $\circ$ is the Hadamard product, Λ^{-1} is the inverse of the spherical parameterization of the covariance matrix computed as in Equation (13), $\bar{\boldsymbol{H}}$ is the sample covariance (i.e., $\bar{\boldsymbol{H}} = \frac{1}{T}\sum_{t=1}^{T} \boldsymbol{e}_t \boldsymbol{e}_t^T$), and $\boldsymbol{u}$ is a lower triangular matrix of ones.

Equation (20) includes a variance targeting approach for reducing the number of parameters to be estimated. The innovation component is driven by the forcing variable $\boldsymbol{\Xi}_{t-1}$, which, in the spirit of Tse and Tsui (2002), is based on the spherical transformation of the sample covariance matrix computed on the past h observations, $\boldsymbol{\Xi}_{t-1} = \Lambda^{-1}(\boldsymbol{H}_{t-h:t-1})$, with $\boldsymbol{H}_{t-h:t-1}$ being the sample correlation matrix of the observations $(\mathbf{e}_{t-h}, \mathbf{e}_{t-h+1}, \mathbf{e}_{t-h+2}, ..., \mathbf{e}_{t-1})$. As stated by the authors, a necessary condition is to impose $h = n$, although applications have shown that a higher number is often necessary to obtain a well-defined covariance or correlation matrix.

The Cholesky-DCC model

Unlike the CH-BEKK model, for the CH-DCC we directly model the angles computed from the spherical transformation of the correlation matrix $\boldsymbol{R}_t$ derived from the decomposition in (5). As in the classical DCC model, our proposed CH-DCC model requires specification of the univariate variance equations, as in (6). This allows us to apply the spherical transformation in (18) directly to the correlation matrix. Thus, we can model the angles from the parameterization similar to Equation (20).

$$\boldsymbol{\Theta}_t = (\boldsymbol{u} - \boldsymbol{A} - \boldsymbol{B}) \circ \Lambda^{-1}(\bar{\boldsymbol{R}}) + \boldsymbol{A} \circ \boldsymbol{\Xi}_{t-1} + \boldsymbol{B} \circ \boldsymbol{\Theta}_{t-1}, \tag{21}$$

where $\circ$ is the Hadamard product, Λ^{-1} is the inverse of the spherical parameterization of the correlation matrix computed as in Equation (18), $\bar{\boldsymbol{R}}$ is the sample correlation matrix (i.e., $\bar{\boldsymbol{R}} = \frac{1}{T}\sum_{t=1}^{T} \boldsymbol{\epsilon}_t \boldsymbol{\epsilon}_t^T$), and $\boldsymbol{u}$ is a lower triangular matrix of ones with zeros on the main diagonal.

Similar to CH-BEKK, the forcing variable, which incorporates the new information, is based on the spherical transformation of the sample correlation matrix computed on the past h observations, $\boldsymbol{\Xi}_{t-1} = \Lambda^{-1}(\boldsymbol{R}_{t-h:t-1})$, with $\boldsymbol{R}_{t-h:t-1}$ being the sample correlation matrix of the observations $(\boldsymbol{\epsilon}_{t-h}, \boldsymbol{\epsilon}_{t-h+1}, \boldsymbol{\epsilon}_{t-h+2}, ..., \boldsymbol{\epsilon}_{t-1})$.

Scalar CH-BEKK and CH-DCC models

Both CH-BEKK and CH-DCC directly model the angles derived from spherical parameterizations. Therefore, the numbers of parameters to be estimated are $n(n+1)$ and $n(n-1)$, respectively. However, for the CH-DCC model, we need to add $3n$ parameters to estimate the univariate GARCH models. To reduce the number of parameters further, it is possible to employ a scalar version of Equation (21), thus imposing the same dynamic on all angles.

In scalar versions, Equations (20) and (21) become

$$\boldsymbol{\Theta}_t = (1 - a - b)\boldsymbol{\kappa} + a\,\boldsymbol{\Xi}_{t-1} + b\,\boldsymbol{\Theta}_{t-1}, \tag{22}$$

where a and b are scalar parameters and, depending on the model specification, $\boldsymbol{\kappa}$ is the inverse of the spherical transformation of the sample covariance or correlation matrix, and $\boldsymbol{\Xi}_{t-1}$ is the inverse of the spherical transformation of the sample covariance or correlation matrix computed from the past h observations.

Parameter estimation

Table 1 compares the required number of parameters in a general representation for all discussed models, with or without variance targeting. The last three columns show the number of parameters for $n = 3$, $n = 10$, and $n = 100$. The full models refer to equations (2), (7), (20), and (21). It is clear that in the case of full parameterizations, the proposed Cholesky models require fewer parameters to estimate than their traditional counterparts. However, this advantage disappears when considering scalar

Table 1 Comparison of the number of parameters for the multivariate GARCH models. For DCC models, the number of parameters does not include the univariate GARCH parameters, i.e. $3n$

Model	Targeting	n assets	$n = 3$	$n = 10$	$n = 100$
Scalar BEKK	No	$n(n+1)/2+2$	8	57	5052
Scalar DCC	No	$n(n-1)/2+2$	5	47	4952
Scalar CH-BEKK	No	$n(n+1)/2+2$	8	57	5052
Scalar CH-DCC	No	$n(n-1)/2+2$	5	47	4952
Full BEKK	No	$n(n+1)/2+2n^2$	24	255	25050
Full DCC	No	$n(n-1)/2+n(n+1)$	15	155	24950
Full CH-BEKK	No	$3n(n+1)/2$	18	165	15150
Full CH-DCC	No	$3n(n-1)/2$	9	135	14850
Scalar BEKK	Yes	2	2	2	2
Scalar DCC	Yes	2	2	2	2
Scalar CH-BEKK	Yes	2	2	2	2
Scalar CH-DCC	Yes	2	2	2	2
Full BEKK	Yes	$2n^2$	18	200	20000
Full DCC	Yes	$n(n+1)$	12	110	10100
Full CH-BEKK	Yes	$n(n+1)$	12	110	10100
Full CH-DCC	Yes	$n(n-1)$	6	90	9900

models, as they all present only two parameters to estimate, with the cost of having the same dynamics for all elements of the matrices.

The parameters for all the discussed models can be estimated by maximizing the log-likelihood. Assuming the normality of the residuals, $\boldsymbol{e}_t \sim \mathcal{N}(\mathbf{0}, \boldsymbol{H}_t)$, the log-likelihood function can be written as

$$\mathcal{L} = -\frac{1}{2}\sum_{t=1}^{T}\left(n\log(2\pi) + \log|\boldsymbol{H}_t| + \boldsymbol{e}_t'\boldsymbol{H}_t^{-1}\boldsymbol{e}_t\right). \tag{23}$$

However, Engle (2002) shows that the parameters of the DCC model can be easily estimated using a two-stage estimation, decomposing the log-likelihood function in (23) as

$$\mathcal{L} = -\frac{1}{2}\sum_{t=1}^{T}\left(n\log(2\pi) + 2\log|\boldsymbol{D}_t| + \boldsymbol{r}_t'\boldsymbol{D}_t^{-1}\boldsymbol{D}_t^{-1}\boldsymbol{r}_t - \boldsymbol{\epsilon}_t'\boldsymbol{\epsilon}_t + \log|\boldsymbol{R}_t| + \boldsymbol{\epsilon}_t'\boldsymbol{R}_t^{-1}\boldsymbol{\epsilon}_t\right). \tag{24}$$

Equation (24) shows that the log-likelihood is composed of a volatility and a correlation term, $\mathcal{L} = \mathcal{L}_v + \mathcal{L}_c$. The volatility term is jointly maximized by separately maximizing each univariate log-likelihood. Therefore, we initially estimate the univariate GARCH parameters in (6), $\omega_i, \alpha_i, \beta_i$, by maximizing their log-likelihood functions. We then estimate the multivariate parameters in the second stage by maximizing the correlation part of the log-likelihood. For all DCC models in our analysis, we employ the two-stage estimation procedure.

It is important to note that, for the scalar BEKK and DCC models, we must perform a constrained estimation by imposing the inequalities $a, b > 0$ and $a + b < 1$ to guarantee the positive definiteness and stationarity of the covariance or correlation matrices. In contrast, our proposed Cholesky version of the models does not require any restrictions to obtain the positive definiteness of the matrices; however, it would require some restrictions on the elements of the matrix $\boldsymbol{B}$, that is, $|b_{ij}| < 1$ (or $|b| < 1$ for the scalar case), to maintain the stationarity of the process. However, to ensure a one-to-one relationship of the spherical transformation, the angles $\theta_{ij,t}$ must be bounded. Specifically, if we consider the transformation of the covariance, we impose the inequalities in (12). By contrast, considering the parameterization of the correlation, all angles must lie within the range $(0, \pi)$. These constraints are imposed in the optimization process with the inclusion of penalization during the computation of the log-likelihood. Empirical tests show that this penalization constrains the angles to vary within the specified ranges, thus eliminating the need for the above stationarity constraints.

A simulation study

To test the proposed methodology, we perform a Monte Carlo simulation in which the structures of the correlations and covariances are known. In particular, as in Engle (2002), we simulate a multivariate model with three series and generated 1,000 observations for each series. We repeated this simulation process 200 times. The data-generating process is built based on Gaussian errors, as follows.

$$\begin{aligned}
&h_{1,t} = 0.01 + 0.05r_{1,t-1}^2 + 0.94h_{1,t-1},\\
&h_{2,t} = 0.5 + 0.2r_{1,t-1}^2 + 0.5h_{1,t-1},\\
&h_{3,t} = 0.1 + 0.8r_{1,t-1}^2 + 0.1h_{1,t-1},\\
&\boldsymbol{\epsilon_t} \sim \mathcal{N}(0, \boldsymbol{R}_t),\\
&r_{i,t} = \sqrt{h_{i,t}}\epsilon_{i,t}, \quad i \in \{1,2,3\}.
\end{aligned}$$

The $n(n-1)/2$ elements of the correlation matrix $\boldsymbol{R}_t$ are generated by applying the spherical transformation as in (17) to the following angles.

$$\begin{aligned}
\theta_{ij,t} =& (1 - A_{ij} - B_{ij}) + A_{ij}\psi_{ij,t} + B_{ij}\theta_{ij,t-1}, \quad i = 2,3,\ j = 1,\dots,i-1,\\
&\text{with}\boldsymbol{A} = \begin{pmatrix} 0 & 0 & 0 \\ 0.1 & 0 & 0 \\ 0.2 & -0.3 & 0 \end{pmatrix}, \boldsymbol{B} = \begin{pmatrix} 0 & 0 & 0 \\ -0.8 & 0 & 0 \\ 0.9 & 0.7 & 0 \end{pmatrix},
\end{aligned}$$

where $\psi_{ij,t}$ is a random angle within the range $[0, \pi]$, obtained from a distribution proportional to $(\sin\psi)^{2k+n-j}$ (Pourahmadi and Wang 2015) .

By knowing the structure of the correlations, we compare the estimated correlations, $\hat{\rho}_{ij,t}$, with the true correlations resulting from the considered covariance matrix $\boldsymbol{R}_t$. To this extent, we use the mean absolute error given by

$$MAE_t = \frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n}|\hat{\rho}_{ij,t} - \rho_{ij,t}|,$$

and we average the above measure over time t and over the 200 simulations.

The results of the Monte Carlo simulation with their standard deviations are presented in Table 2. The Cholesky models perform better than their standard counterparts, in particular both models with full specifications. Moreover, owing to the simulation structure, the DCC model approaches either the traditional or Cholesky models, outperforming the BEKK models.

Application

In this section, we apply the proposed Cholesky model in its full and scalar specifications to financial data, specifically crude oil daily prices and exchange rates. Moreover, we perform a comparison with the traditional scalar BEKK and DCC specifications. The specific choice of crude oil is mainly to show the validity of our proposed approach. Nevertheless, the model can easily be applied to different types of data, including

Table 2 Mean absolute error of correlation estimates

	MAE	SD
BEKK Scalar	0.1054	0.0084
CH-BEKK Scalar	0.0859	0.0139
CH-BEKK Full	0.0658	0.0076
DCC Scalar	0.0533	0.0022
CH-DCC Scalar	0.0521	0.0021
CH-DCC Full	0.0521	0.0024

Table 3 Summary statistics of daily log returns

	WTI	Brent	AUD	CAD	EUR	GBP	JPY	MXN	NOK
Mean	0.0234	0.0258	− 0.0018	− 0.0024	− 0.0015	0.0035	0.0019	0.0134	0.0019
Median	0.0702	0.0545	− 0.029	0.0	− 0.0079	− 0.0061	0.0	− 0.0139	− 0.013
Max	42.5832	41.2023	7.6444	3.2986	3.1771	8.4095	6.2151	8.7728	7.5385
Min	− 72.0273	− 77.2684	− 8.2218	− 3.7665	− 3.7188	− 3.0427	− 3.7831	− 6.5352	− 5.6124
Std. dev.	2.9134	2.7062	0.7974	0.546	0.6001	0.5835	0.6125	0.7411	0.7691
Skewness	− 2.1333	− 3.689	0.4525	0.1433	− 0.0075	0.9026	− 0.0311	0.8837	0.3868
Excess Kurtosis	88.2326	136.4251	11.3098	3.0414	1.9238	11.4144	5.2736	11.9675	4.4623
Jarque-Bera	1841229.0***	4404453.0***	30375.0***	2202.0***	873.0***	31512.0***	6563.0***	34531.0***	4840.0***
ADF	− 12.7***	− 12.9***	− 13.0***	− 13.0***	− 76.6***	− 16.4***	− 21.5***	− 13.3***	− 13.1***
ARCH(1)	554.0***	397.5***	453.0***	288.7***	90.7***	128.1***	103.2***	302.1***	254.3***
ARCH(6)	739.5***	462.3***	1685.7***	986.9***	359.0***	218.9***	334.7***	1130.7***	878.4***
ARCH(12)	803.5***	594.9***	1893.1***	1131.3***	481.3***	266.3***	461.7***	1155.4***	909.8***
Q(6)	96.5***	16.9***	35.1***	15.5**	7.7	12.3*	14.9**	10.8*	6.7
Q(12)	167.2***	49.3***	48.1***	27.6***	18.9*	25.7**	39.3***	24.8**	15.0
$Q^2(6)$	883.3***	482.1***	4059.6***	2189.6***	596.8***	295.5***	508.3***	2256.3***	1711.8***
$Q^2(12)$	1192.0***	771.8***	6881.8***	4076.9***	1140.6***	439.0***	874.3***	2903.7***	2198.6***

Summary statistics of the percentage log returns over the period from January 2000 to February 2022. ADF is the Augmented Dickey-Fuller test. ARCH refers to the statistic of the Lagrange Multiplier test for autoregressive conditional heteroscedasticity at 1, 6, and 12 lags. Q(6), Q(12), $Q^2(6)$, and $Q^2(12)$ are the Ljung-Box statistics for serial correlation of orders 6 and 12 in returns and squared returns.

*, **, and *** indicate the rejection of the null hypothesis of the tests at the 10%, 5%, and 1% levels of significance respectively

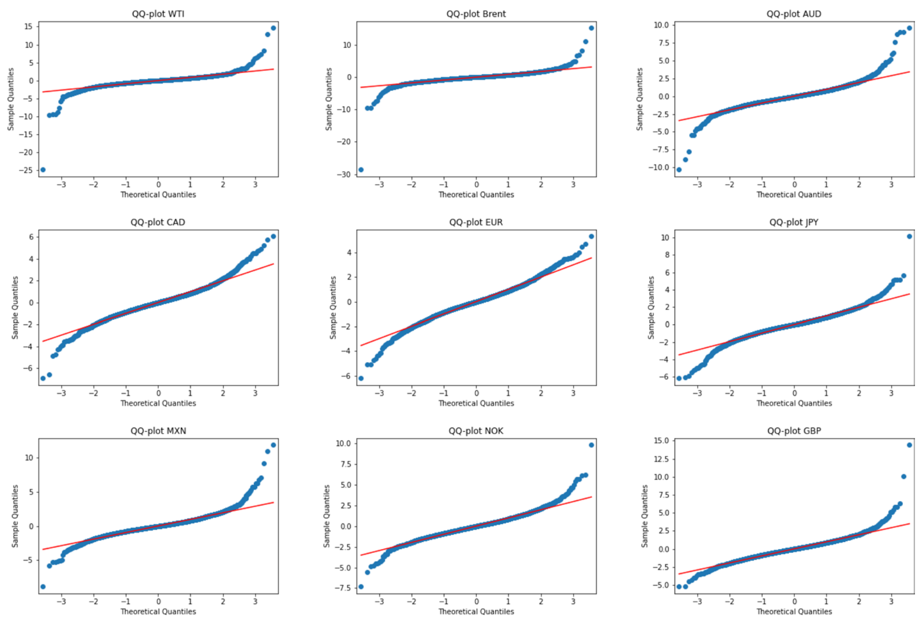

Fig. 1 Q-Q plot of the daily log returns over the period from January 2000 to February 2022

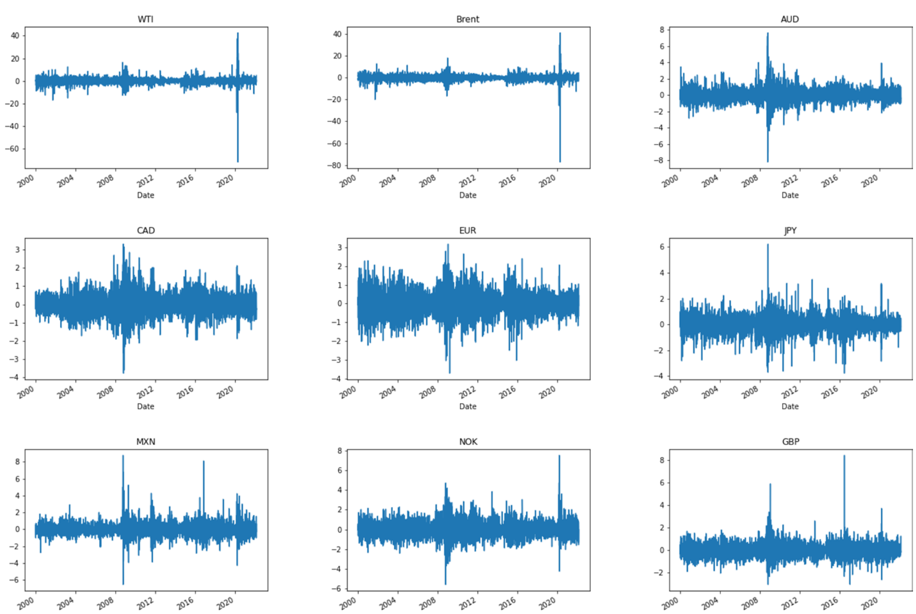

Fig. 2 Daily log returns over the period from January 2000 to February 2022

applications where the analysis of volatility over time is the key objective and the decision-making process is influenced by risk factors, for instance cryptocurrencies markets, renewable energy markets (see, e.g., Kou et al. 2023, 2024), and carbon markets (Kou et al. 2024). We also highlight that the proposed model is not constrained to daily data, but can be applied to higher or lower time frequencies, where the former are characterized by high volatility due to noise in the time series, and the latter might show volatility clustering because of the varying economic cycles when considering monthly or annual data.

Data and summary statistics

We analyze the relationship between crude oil spot prices (WTI and Brent) and seven exchange rates to the US dollar, namely the Australian dollar (AUD), the Canadian dollar (CAD), the Euro (EUR), the Japanese yen (JPY), the Mexican peso (MXN), the Norwegian crown (NOK), and the British pound (GBP). For all these quantities, we consider the daily price series from January 2000 to February 2022. The exchange rates are collected from Thomson Reuters, whereas the oil price data (both Brent and WTI) are sourced from the US Energy Information Administration.

Table 3 reports a wide range of summary statistics for the percentage log-returns of oil prices and exchange rates. All series, especially the two oil series, show small means with high standard deviations. Only EUR exhibits low skewness, and all series have a large excess kurtosis, which indicates the presence of heavy tails, as is typical of financial time series. This aspect is also confirmed by the rejection of the Jarque-Bera tests of normality at the 1% level of significance for all series. Moreover, the ARCH tests at 1, 6, and 12 lags reject the null hypothesis of homoskedasticity at the 1% level of significance. Moreover, the Ljung-Box portmanteau tests on both returns and squared returns with lags of 6 and 12 show a high serial correlation for all series except for the EUR and NOK exchange rates.

Figure 1 shows the QQ-plots against the normal distribution, which indicates that all series have heavier tails than normal, whereas in Figure 2 we observe clear volatility clustering in all the time series. This behavior of large changes, which tend to follow large changes, and small changes, which tend to follow small changes, is typical of financial time series.

Finally, Table 4 reports the results of a pairwise Granger causality test with one lag for all series (Granger 1969). Specifically, the past values of the series in columns have a statistically significant effect on the current values of the series in rows. For example, WTI Granger causes the Brent and JPY, MXN, and NOK series, whereas Brent is Granger

Table 4 Granger causality matrix at first lag

	WTI_x	Brent_x	AUD_x	CAD_x	EUR_x	GBP_x	JPY_x	MXN_x	NOK_x
WTI_y	–	2.813*	4.588**	5.852**	0.798	0.209	4.708**	3.962**	0.897
Brent_y	295.938***	–	27.034***	41.194***	3.138*	8.195***	5.603**	17.036***	14.196***
AUD_y	1.757	0.377	–	5.643**	0.504	0.525	2.849*	26.834***	0.74
CAD_y	0.002	0.035	0.017	–	3.595*	1.851	3.747*	2.759*	3.098*
EUR_y	0.146	0.579	0.025	1.176	–	0.418	1.983	1.003	0.904
GBP_y	1.929	3.7*	2.478	11.608***	0.449	–	12.725***	1.363	0.02
JPY_y	7.423***	3.584*	15.806***	12.516***	2.012	0.013	–	0.84	10.382***
MXN_y	5.147**	0.574	1.079	5.595**	2.992*	0.507	0.002	–	1.831
NOK_y	4.264**	2.217	15.616***	13.567***	3.329*	0.635	0.091	10.38***	–

The Null hypothesis for the Granger causality test is that the time series in the columns (_x suffix), do NOT Granger cause the time series in the rows (_y suffix).

*, ** and *** indicate the rejection of the null hypothesis at the 10%, 5% and 1% levels of significance respectively

Table 5 Univariate GARCH parameters

		AUD	CAD	EUR	GBP	JPY	MXN	NOK
WTI	ω	0.1149***	0.1164***	0.1172***	0.1176***	0.1166***	0.1165***	0.1157***
	α	0.0931***	0.0927***	0.0929***	0.0923***	0.0929***	0.0928***	0.0936***
	β	0.892***	0.892***	0.8916***	0.8921***	0.8918***	0.8919***	0.8914***
Brent	ω	0.0683***	0.0677***	0.0667***	0.0666***	0.0678***	0.0659***	0.0667***
	α	0.0914***	0.0907***	0.0915***	0.0908***	0.0916***	0.0913***	0.0921***
	β	0.8996***	0.9004***	0.8999***	0.9006***	0.8996***	0.9003***	0.8994***
Ex Rate	ω	0.0049***	0.0018***	0.001**	0.0041***	0.0044***	0.0066***	0.0052***
	α	0.0544***	0.0454***	0.0345***	0.0581***	0.06***	0.0975***	0.0426***
	β	0.9367***	0.9483***	0.9629***	0.9307***	0.9295***	0.8922***	0.9483***

** and *** indicate 5% and 1% levels of significance respectively

Table 6 Multivariate GARCH parameters

		AUD	CAD	EUR	GBP	JPY	MXN	NOK
BEKK Scalar	a	0.0635***	0.062***	0.057***	0.0634***	0.0652***	0.0708***	0.0631***
	b	0.925***	0.927***	0.9338***	0.9242***	0.9215***	0.9161***	0.9244***
	$\mathcal{L}$	− 28743.77	− 26675.54	− 27677.23	− 27571.02	− 27883.67	− 28368.94	− 28872.94
DCC Scalar	a	0.0073	0.0091*	0.009	0.0054	0.0278	0.0093	0.0061
	b	0.9893***	0.987***	0.9858***	0.9913***	0.9097***	0.9843***	0.9908***
	$\mathcal{L}$	− 28590.35	− 26493.11	− 27511.67	− 27398.28	− 27742.62	− 28179.21	− 28704.94
CH-BEKK Scalar	a	0.6016***	0.5644***	0.5056***	0.5032***	0.5611***	0.7221***	0.6151***
	b	0.1322	0.1839	0.2635	0.2507*	0.1559	− 0.0418	0.0972
	$\mathcal{L}$	− 29030.91	− 26944.63	− 27990.2	− 27896.62	− 28211.28	− 28762.32	− 29140.21
CH-DCC Scalar	a	0.005*	0.0073*	0.003*	0.0027	0.1407***	0.0702	0.1463
	b	0.9915***	0.9879***	0.9951***	0.995***	0.4327***	0.7671***	0.5233*
	$\mathcal{L}$	− 28612.94	− 26514.76	− 27529.94	− 27408.28	− 27763.49	− 28204.5	− 28740.12
CH-BEKK Full	a_{11}	0.9129***	0.8956***	0.9065***	0.8859***	0.9***	0.9159***	0.8906***
	a_{21}	0.8411***	0.8274***	0.8411***	0.8465***	0.8359***	0.8092***	0.826***
	a_{22}	0.2264**	0.2293**	0.2289**	0.2214**	0.2286**	0.1961**	0.2244**
	a_{31}	0.8881***	0.2106**	0.0808***	1.1643***	0.6563**	1.0727***	0.5367***
	a_{32}	0.007**	0.0101**	0.0049*	0.0037***	0.1469	0.012	0.1471**
	a_{33}	0.1065	0.1475	0.0011	− 0.0117	0.0698*	0.1705**	0.0754
	b_{11}	− 0.1686	− 0.1502	− 0.1624	− 0.1401	− 0.155	− 0.1785	− 0.142
	b_{21}	− 0.0729	− 0.0565	− 0.0674	− 0.0763	− 0.0623	− 0.0332	− 0.0509
	b_{22}	0.4184**	0.4228**	0.4112*	0.4303**	0.4214*	0.4923**	0.4194*
	b_{31}	− 0.1402	0.7375***	0.9032***	−0.6967***	0.0298	−0.3501***	0.2226
	b_{32}	0.9913***	0.987***	0.9943***	0.9956***	0.4361	0.9821***	0.6703***
	b_{33}	−0.8853***	−0.8314***	0.9966***	0.8219***	0.2512*	−0.754***	− 0.3308
	$\mathcal{L}$	− 28789.88	− 26705.09	− 27713.54	− 27615.48	− 27980.99	− 28496.34	− 28940.55
CH-DCC Full	a_{21}	0.1458*	0.1509**	0.1466*	0.1473**	0.1434	0.1291*	0.149**
	a_{31}	0.0079**	0.0123**	0.0048**	0.004**	0.147	0.0191	0.1273
	a_{32}	0.1008	0.0041**	−0.0119***	−0.0131*	0.0654	0.0936	0.0611
	b_{21}	0.5332**	0.5245***	0.5377**	0.5348***	0.5518*	0.5938**	0.5256**
	b_{31}	0.9902***	0.9847***	0.9944***	0.9949***	0.3991	0.9702***	0.7245***
	b_{32}	−0.9309***	0.9919***	0.795***	0.8615***	0.2399	−0.8668**	− 0.2326
	$\mathcal{L}$	− 28599.08	− 26499.97	− 27516.66	− 27398.26	− 27756.73	− 28182.3	− 28722.0

*, ** and *** indicate 10%, 5% and 1% levels of significance respectively

caused by all series. A similar pattern is observed for the second and third lags, although we did not tabulate the results to save space.

The results of the Granger causality test suggest that a VAR specification can be used to model the returns series. Moreover, the results of the preliminary analysis indicate that the GARCH specification is a good candidate for the modeling of conditional volatility.

Results

We apply the proposed Cholesky models, in both full and scalar settings, as well as the scalar BEKK and DCC models to different combinations of time series. Specifically, in each test, we include three time series: the two crude oils and one of the exchange rates. Thus, we obtain seven combinations of time series for the application.

For the multivariate DCC models, which include a two-stage estimation, we report the univariate GARCH parameters, ω_i, α_i, and β_i, for each combination of series in Table 5. Each column represents one combination of oils and one exchange rate. For example, the first column *AUD* includes the results of the univariate GARCH for the WTI, Brent, and AUD time series.

Table 6 reports the multivariate GARCH parameters a and b for the scalar models, and the single elements of the matrices $\boldsymbol{A}$ and $\boldsymbol{B}$ for the full models, which are described in (14) and (19). For the DCC models, the values refer to the second stage of estimation. In the Cholesky models, the innovation components $\boldsymbol{\Xi}_{t-1}$ from equations (20) and (21) are based on the past 21 observations, which correspond to one month of data using returns from business days.

An initial analysis of the log-likelihood values shows that the performance of the full Cholesky models is comparable to that of the corresponding standard scalar models. Specifically, the full log-likelihood of the CH-DCC is − 28599.08, which is very close to the value of − 28590.35 from the scalar DCC model. We note a similar behavior for the BEKK specifications. In contrast, the scalar CH-BEKK and CH-DCC models underperform compared to their traditional counterparts, with values of − 29030.91 and − 28612.94, respectively. This behavior, which is observable in all combinations of time series, does not allow us to directly discriminate one model from another. Moreover, because we observe that all log-likelihood values of the BEKK models are always lower than those of the DCC models, we tend to prefer the DCC specification over the BEKK one. However, Caporin and McAleer (2012) demonstrated that the BEKK model is theoretically more robust than the DCC model and should be preferred in applications. Owing to these ambiguities, we perform further tests to assess the performance of the proposed models.

Considering that the true values of the covariances are unknown, it is not possible to directly assess the variance–covariance output of the models. Therefore, following Engle (2002), we compute the 95% Value-at-Risk for three portfolios to evaluate model performance. The portfolios are (i) an equal weights portfolio, (ii) a mixed portfolio with 90% weight on the WTI and 5% weight on both the Brent and the exchange rate series, and (iii) a long/short portfolio with 200% weight on the WTI and two equal short positions on the Brent and the exchange rate with − 50% weight. To compare the three portfolios and the different models, we apply the Kupiec test (Kupiec 1995).

We compute the Value-at-Risk estimate for the process $\boldsymbol{e}_t$ for the i-th portfolio at time t with a coverage level α as

$$\mathrm{VaR}_{it}(\alpha) = z(\alpha) \cdot \sqrt{\boldsymbol{w}_i' \boldsymbol{H}_t \boldsymbol{w}_i}, \tag{25}$$

Table 7 Kupiec test for different portfolios and different combinations of oils and exchange rate. VaR at 95%. *p* values in brackets

	Model	Equal weights	Mix	Long/Short
AUD	BEKK	4.733 (0.03)	3.694 (0.055)	2.581 (0.108)
	CH-BEKK Scalar	3.455 (0.063)	5.601 (0.018)	6.546 (0.011)
	CH-BEKK Full	0.846 (0.358)	0.538 (0.463)	0.846 (0.358)
	DCC	0.964 (0.326)	0.133 (0.716)	0.372 (0.542)
	CH-DCC Scalar	1.09 (0.297)	0.092 (0.762)	0.237 (0.626)
	CH-DCC Full	0.372 (0.542)	0.237 (0.626)	0.301 (0.583)
CAD	BEKK	3.941 (0.047)	3.002 (0.083)	2.581 (0.108)
	CH-BEKK Scalar	3.002 (0.083)	5.601 (0.018)	7.219 (0.007)
	CH-BEKK Full	0.372 (0.542)	0.058 (0.809)	0.736 (0.391)
	DCC	0.633 (0.426)	0.181 (0.67)	0.964 (0.326)
	CH-DCC Scalar	0.633 (0.426)	0.133 (0.716)	0.736 (0.391)
	CH-DCC Full	0.452 (0.502)	0.133 (0.716)	0.237 (0.626)
EUR	BEKK	1.223 (0.269)	2.787 (0.095)	2.012 (0.156)
	CH-BEKK Scalar	4.461 (0.035)	4.733 (0.03)	7.569 (0.006)
	CH-BEKK Full	0.133 (0.716)	0.372 (0.542)	0.452 (0.502)
	DCC	0.014 (0.905)	0.237 (0.626)	0.452 (0.502)
	CH-DCC Scalar	0.004 (0.949)	0.301 (0.583)	0.372 (0.542)
	CH-DCC Full	0.058 (0.809)	0.181 (0.67)	0.301 (0.583)
GBP	BEKK	3.455 (0.063)	5.304 (0.021)	2.194 (0.139)
	CH-BEKK Scalar	3.694 (0.055)	7.219 (0.007)	9.448 (0.002)
	CH-BEKK Full	0.237 (0.626)	0.181 (0.67)	0.964 (0.326)
	DCC	0.301 (0.583)	0.372 (0.542)	0.452 (0.502)
	CH-DCC Scalar	0.133 (0.716)	0.372 (0.542)	0.633 (0.426)
	CH-DCC Full	0.133 (0.716)	0.301 (0.583)	0.058 (0.809)
JPY	BEKK	2.383 (0.123)	4.197 (0.04)	4.733 (0.03)
	CH-BEKK Scalar	5.304 (0.021)	6.223 (0.013)	9.851 (0.002)
	CH-BEKK Full	0.237 (0.626)	0.181 (0.67)	0.964 (0.326)
	DCC	0.003 (0.954)	0.452 (0.502)	0.372 (0.542)
	CH-DCC Scalar	0.004 (0.949)	0.372 (0.542)	0.181 (0.67)
	CH-DCC Full	0.016 (0.901)	0.301 (0.583)	0.133 (0.716)
MXN	BEKK	4.461 (0.035)	3.694 (0.055)	2.581 (0.108)
	CH-BEKK Scalar	6.546 (0.011)	5.014 (0.025)	7.569 (0.006)
	CH-BEKK Full	0.736 (0.391)	0.301 (0.583)	1.515 (0.218)
	DCC	0.538 (0.463)	0.452 (0.502)	0.237 (0.626)
	CH-DCC Scalar	0.846 (0.358)	0.452 (0.502)	0.736 (0.391)
	CH-DCC Full	0.372 (0.542)	0.452 (0.502)	0.372 (0.542)
NOK	BEKK	1.672 (0.196)	3.694 (0.055)	2.581 (0.108)
	CH-BEKK Scalar	3.002 (0.083)	4.461 (0.035)	5.601 (0.018)
	CH-BEKK Full	0.058 (0.809)	0.181 (0.67)	0.237 (0.626)
	DCC	0.032 (0.857)	0.058 (0.809)	0.133 (0.716)
	CH-DCC Scalar	0.014 (0.905)	0.032 (0.857)	0.014 (0.905)
	CH-DCC Full	0.0 (0.998)	0.032 (0.857)	0.014 (0.905)

where $z(\alpha)$ is the value of the inverse of the CDF of the standard normal distribution at level α and $\boldsymbol{w}_i$ is the vector containing the i-th portfolio weights.

We then define an indicator function as

$$I_{it} = \begin{cases} 1 & \text{if } \boldsymbol{e}_t'\boldsymbol{w}_i < \mathrm{VaR}_{it}(\alpha) \\ 0 & \text{if } \boldsymbol{e}_t'\boldsymbol{w}_i \geq \mathrm{VaR}_{it}(\alpha), \end{cases}$$

and compute the Coverage Ratio as

$$CR_i = \frac{1}{T}\sum_{t}^{T} I_{it}, \tag{26}$$

where T is the length of the time series.

Finally, we apply the likelihood ratio test statistic of Kupiec (1995), which we compute as

$$LR_i = 2\left\{\log[CR_i^{\gamma_i}(1 - CR_i)^{T-\gamma_i}] - log[\alpha^{\gamma_i}(1-\alpha)^{T-\gamma_i}]\right\}, \tag{27}$$

where $\gamma_i = \sum_t^T I_{it}$. LR_i follows a $\mathcal{X}^2(1)$ distribution. A similar approach has been applied by Bauwens et al. (2006) and Creal et al. (2011).

Results of the 95% Value-at-Risk tests on the portfolios are shown in Table 7. The reported values are the statistics of Kupiec to test the null hypothesis $H_0 : CR_i = \alpha$ of a correct coverage. The p values are reported in parentheses.

We observe that for almost all portfolios and all combinations of time series, the Cholesky models perform better than the traditional ones; in particular, the full models have better coverage, that is, the p-values are higher overall. The only exception is the scalar CH-BEKK model, which reports the worst performance because it always rejects the null hypothesis. Moreover, considering the mixed portfolio, the standard DCC model presents good results that are sometimes better than those of the proposed models. However, we can affirm that the Kupiec test shows the satisfactory performance of the Cholesky models, particularly in their full specifications.

Conclusions

Generally, multivariate GARCH models are subject to two problems. On the one hand, the requirement of the positive definiteness of the covariance matrix poses some difficulties. On the other hand, the number of parameters to estimate in the full specification of the models increases drastically with the inclusion of additional time series.

To overcome the requirement of positive definiteness, we proposed two extensions of the BEKK and DCC models by applying two spherical transformations to the Cholesky factorization of the covariance and correlation matrices. With the application of the parameterization of the covariance matrix, we introduce the Cholesky-BEKK model. In contrast, by employing the parameterization of the correlation matrix, we introduce the Cholesky-DCC model. These transformations allow us to define the updating functions of the models without constraints on the parameters. Moreover, they permit a reduction in the total number of parameters to be estimated when considering the full specification because, by construction, both transformations result in triangular matrices of angles. Thus, fewer parameters are required to be estimated when maximizing the log-likelihood.

Moreover, we applied the proposed full and scalar models to WTI and Brent crude oil returns, along with different exchange rate returns, and compared their results with those of the standard BEKK and DCC models. The application shows that the proposed models can capture the dynamics of covariances and correlations. Moreover, we performed the Kupiec test on the Value-at-Risk of three portfolio combinations and confirm the satisfactory performance of the Cholesky models in their full specification.

Finally, we observe that the proposed models, when applied to any time series, could be useful for practitioners and policymakers as they might be able to estimate and predict the conditional volatility of the time series without imposing the positive definiteness of the covariance or correlation matrices. In this way, they can control or mitigate risky positions or enforce regulations to avoid potential misconduct in markets, such as advisors or brokers recommending unsuitable products to their clients, which can be dangerous when markets are under pressure.

Abbreviations

AUD	Australian dollar
BEKK	Baba, engle, kraft, and kroner
CAD	Canadian dollar
CCC	Constant conditional correlation
CH-BEKK	Cholesky-BEKK
CH-DCC	Cholesky-DCC
DCC	Dynamic conditional correlation
EUR	Euro
GARCH	Generalized autoregressive conditional heteroskedasticity
GBP	British pound
JPY	Japanese yen
MGARCH	Multivariate generalized autoregressive conditional heteroskedasticity
MXN	Mexican peso
NOK	Norwegian crown
WTI	West Texas instrument
VAR	Vector autoregressive
VaR	Value-at-risk
VEC	Operator that transforms a matrix into vector

Acknowledgements
The authors would like to thank the anonymous reviewers for their comments and suggestions. Luca Vincenzo Ballestra, Graziella Pacelli, and Riccardo De Blasis are members of the Gruppo Nazionale Calcolo Scientifico-Istituto Nazionale di Alta Matematica (GNCS-INdAM).

Author contributions
All authors contributed equally to this work. All authors read and approved the final manuscript.

Funding
Not applicable.

Availability of data and material
The datasets generated and/or analyzed during the current study are available in the Thomson Reuters Refinitive Datascope and US Energy Information Administration databases.

Declarations

Competing interests
The author declare that they have no competing interests.

Received: 19 January 2023 Accepted: 11 November 2024
Published online: 12 January 2025

References

Bauwens L, Laurent S, Rombouts JVK (2006) Multivariate GARCH models: a survey. J Appl Economet 21(1):79–109
Bernardi M, Catania L (2018) Portfolio optimisation under flexible dynamic dependence modelling. J Empir Financ 48:1–18
Bollerslev T (1986) Generalized autoregressive conditional heteroskedasticity. J Econom 31(3):307–327
Bollerslev T (1990) Modelling the coherence in short-run nominal exchange rates: a multivariate generalized arch model. Rev Econ Stat 72(3):498–505
Bollerslev T, Engle RF, Wooldridge JM (1988) A capital asset pricing model with time-varying covariances. J Polit Econ 96(1):116–131
Caporin M, McAleer M (2012) Do we really need both bekk and dcc? A tale of two multivariate GARCH models. J Econom Surv 26(4):736–751
Creal D, Koopman SJ, Lucas A (2011) A dynamic multivariate heavy-tailed model for time-varying volatilities and correlations. J Bus Econom Stat 29(4):552–563
Ding Z, Engle RF (2001) 'Large Scale Conditional Covariance Matrix Modeling, Estimation and Testing', NYU Working Paper No. S-DRP-01-07
Engle R (2002) Dynamic conditional correlation. J Bus Econom Stat 20(3):339–350
Engle RF, Kroner KF (1995) Multivariate simultaneous generalized arch. Economet Theor 11(1):122–150
Granger CWJ (1969) Investigating causal relations by econometric models and cross-spectral methods. Econometrica 37(3):424–438
Karolyi GA (1995) A multivariate GARCH model of international transmissions of stock returns and volatility: the case of the United States and Canada. J Bus Econom Stat 13(1):11–25
Kearney C, Patton AJ (2000) Multivariate GARCH modeling of exchange rate volatility transmission in the European monetary system. Financ Rev 35(1):29–48
Kou G, Pamucar D, Dinçer H, Yüksel S (2023) From risks to rewards: a comprehensive guide to sustainable investment decisions in renewable energy using a hybrid facial expression-based fuzzy decision-making approach. Appl Soft Comput 142:110365
Kou G, Dinçer H, Yüksel S (2024) Pattern recognition of financial innovation life cycle for renewable energy investments with integer code series and multiple technology S-curves based on Q-ROF DEMATEL. Financ Innov 10(1):53
Kou G, Pamucar D, Yüksel S, Dinçer H, Deveci M, Umar M (2024) Evaluation of multidimensional carbon neutrality policies in transportation using a novel quantum picture fuzzy rough modeling. IEEE Trans Eng Manage 71:7681–7700
Kroner KF, Ng VK (1998) Modeling asymmetric comovements of asset returns. Rev Financ Stud 11(4):817–844
Kupiec PH (1995) Techniques for verifying the accuracy of risk measurement models. J Deriv 3(2):73–84
Longin F, Solnik B (2001) Extreme correlation of international equity markets. J Financ 56(2):649–676
Pedeli X, Fokianos K, Pourahmadi M (2015) Two Cholesky-log-GARCH models for multivariate volatilities. Stat Model 15(3):233–255
Pinheiro JC, Bates DM (1996) Unconstrained parametrizations for variance-covariance matrices. Stat Comput 6(3):289–296
Pourahmadi M, Wang X (2015) Distribution of random correlation matrices: hyperspherical parameterization of the Cholesky factor. Stat Probab Lett 106:5–12
Rapisarda F, Brigo D, Mercurio F (2007) Parameterizing correlations: a geometric interpretation. IMA J Manag Math 18(1):55–73
Rebonato R, Jackel P (1999) The most general methodology for creating a valid correlation matrix for risk management and option pricing purposes. J Risk 2(2):17–27
Tse YK, Tsui AKC (2002) A multivariate generalized autoregressive conditional heteroscedasticity model with time-varying correlations. J Bus Econom Stat 20(3):351–362
Vrontos ID, Dellaportas P, Politis DN (2003) A full-factor multivariate GARCH model. Economet J 6(2):312–334
Zhang W, Leng C, Tang CY (2015) A joint modelling approach for longitudinal studies. J R Stat Soc : Series B (Stat Methodol) 77(1):219–238

Publisher's Note