JIRSS (2026)
Vol. 25, No. 2, pp 119-136
DOI: 10.22034/jirss.2026.2039785.1076

A time-varying cholesky garch model for dynamic dependence modeling

Ferdous Mohammadi Basatini[1], **Saeid Rezakhah**[2]

[1]Department of Mathematics and Statistics, Sho.C., Islamic Azad University, Shoushtar, Iran.
[2]Faculty of Mathematics and Computer Science, Amirkabir University of Technology, Tehran, Iran.

Received: 2024-08-29, Accepted: 2026-06-09, Published online: 2026-08-12

Abstract. To describe instantaneous dependencies, we consider time-varying coefficients in Cholesky GARCH (C-GARCH) model. The regression coefficients are updated through a Gaussian random-walk. These coefficients are estimated using the Kalman filter method based on the most recent information as it becomes available and inducing dynamic prior distribution of the coefficients. The simulation and real data studies reveal the superior performance of the proposed time-varying Cholesky GARCH (TC-GARCH) model compared to the previous models by capturing instantaneous dependencies.

Keywords. Cholesky decomposition, Dependencies, Gaussian random-walk, time-varying covariance matrix.

MSC: 37M10, 91B84.

1 Introduction

The study of dependencies in multivariate time series has always been of interest. In financial time series these dependencies can reflect the effects of different markets on each other, so evaluation of the instantaneous dependencies is of great importance in modelling multivariate financial time series. Many tasks of modern financial management, including portfolio selection, option pricing and risk assessment require modelling

Corresponding Author: Ferdous Mohammadi Basatini (fe.mohamadi2024@iau.ac.ir).
Saeid Rezakhah (rezakhah@aut.ac.ir).

and prediction instantaneous dependencies. Multivariate GARCH processes are established to model the instantaneous dependencies through time-varying covariance matrix $\mathbf{\Sigma}_t$. Maintaining positive-definiteness constraint of the time-varying covariance matrix $\mathbf{\Sigma}_t$ for all t is an important issue in multivariate GARCH models. Special attention is needed to maintain this constraint when the dimension of data is large. Cholesky GARCH (C-GARCH) model introduced by Dellaportas and Pourahmadi (2012) based on modified Cholesky decomposition (MCD) of covariance matrix inherently guarantees the positive-definiteness constraint of the time-varying covariance matrix $\mathbf{\Sigma}_t$. Pourahmadi (1999) proposed the modified Cholesky decomposition of covariance matrix that provides some statistically interpretable parametrization of the covariance matrix. The MCD method decomposes the covariance matrix parsimoniously by a unit lower triangular matrix with linear regression coefficients as off-diagonal entries and a diagonal matrix having variances on its diagonal. Dellaportas and Pourahmadi (2012) considered the least square estimators of the linear regression coefficients in the MCD. They used univariate GARCH processes to model variances of individual assets.
Assuming the constancy of the coefficients leads to misspecified models and potentially improper financial decisions. Hence, we need to apply models with time-varying coefficients to capture all the dynamic aspects of the financial time series. In this paper, we introduce a time-varying Cholesky GARCH (TC-GARCH) model with time-varying regression coefficients of ordered variables where these coefficients are updated using newly arriving information through Gaussian random-walk. The proposed model has the advantage that the regression parameters and the multivariate GARCH parameters are estimated separately. The time-varying regression coefficients are estimated by using the Kalman filter based on the most recent information as it becomes available and inducing dynamic prior distribution of the coefficients.
Simulation study and empirical application show competitive performance of the TC-GARCH model in comparison with the competing models in forecasting instantaneous dependencies. We rely on four loss functions to evaluate the ability of the TC-GARCH model in calculating instantaneous dependencies in comparison with different models. Three competing models are used where two of them are in the class of the C-GARCH models with constant dependency and third one is a dynamic correlation conditional GARCH (DCC-GARCH) model. Additionally different models are evaluated on the basis of how well they forecast one-step-ahead Value-at-risk (VaR).

A variety of multivariate extensions of univariate GARCH models, introduced by Bollerslev (1986), has been developed so far. VEC and BEKK models are known as the first ones, arose as a direct generalization of the univariate GARCH models which introduced by Bollerslev et al. (1988) and Engle and Kroner (1995), respectively. In these models, the number of parameters increases rapidly with the dimensionality and therefore the estimation of the parameters is computationally complex and expensive. Some other methods impose strong assumptions on the conditional correlation matrices. For instance, Bollerslev (1990) assumed constant conditional correlation for the GARCH model (CCC-GARCH), which may not be satisfied in real data. Engle (2002) and Tse and Tsui (2002) assumed dynamic conditional correlation for the GARCH model (DCC-GARCH), which is computationally expensive in high-dimensional cases. To overcome the curse of dimensionality, several methods have been proposed where the idea of orthogonal transformation has attracted much attention. For instance, Alexander (2001) proposed an orthogonal-GARCH (O-GARCH) model and used principal component analysis (PCA) to reduce the number of variables. Also Van der Weide (2002) presented a class of generalized orthogonal GARCH (Go-GARCH) models and Muriel Torrero (2014) did a study on tailed GO-GARCH models. Paolella et al. (2021) proposed a new robust orthogonal GARCH model for a multivariate set of non-Gaussian asset returns. In the class of C-GARCH models, Pedeli et al. (2015) studied Cholesky Log-GARCH (CL-GARCH) models. They considered the same structure as that in Dellaportas and Pourahmadi (2012) for lower triangular matrix of Cholesky decomposition, to describe the constant dependence structure, and Log-GARCH model to describe variances in diagonal matrices. Recently, some studies are done via Cholesky decomposition. For example, Darolles et al. (2018) presented a new C-GARCH model with time-varying slope coefficients. Using Cholesky decomposition, Näf et al. (2019) developed a new construction that generalizes the multivariate generalized hyperbolic distribution.

Kang et al. (2020) developed an order-averaged Cholesky-log-GARCH model to mitigate the variable-ordering issue inherent in modified Cholesky decomposition. Zheng and Ye (2024) developed Cholesky-based Generalized Autoregressive Score (GAS) models suitable for large-dimensional time-varying covariance matrices. Recently, several studies have incorporated regularization techniques into Cholesky-based GARCH models to address high-dimensional settings. Zhu et al. (2024) developed a sparse banded time-varying precision matrix estimator under the modified Cholesky decomposition framework using SCAD and group lasso penalties, while Kemper (2020) proposed sparse Cholesky-GARCH (CHAR) models based on penalization methods to reduce model complexity. Both studies demonstrated that introducing sparsity can improve estimation accuracy and scalability in high-dimensional applications.

This article is organized as follows: Section 2 is allocated to introducing the TC-GARCH model and its features. Section 3 is devoted to estimating the time-varying regression coefficients and multivariate GARCH parameters. The simulation experiment is studied in Section 4. The models evaluation criteria are presented in Section 5. By analyzing a real dataset of monthly stock returns of 12 US bluechips in Section 6, we show the competing performance of the TC-GARCH model. Conclusions are presented in the last section.

2 The model

Cholesky GARCH model is a class of multivariate GARCH models that present an ordering of variables which in turn are modeled by univariate GARCH model and describes dependencies through linear regressions of each variable with respect to the previous ones in the presented ordering. So one needs to detect some ordering between multivariate time series, say $\mathbf{Y}_t = (Y_{1t}, Y_{2t}, \ldots, Y_{pt})$, and regress each time series variable Y_{jt} with respect to the preceding variables at the same time, say $Y_{1t}, Y_{2t}, \ldots, Y_{(j-1)t}$.

Such ordering is studied by Basford and Tukey (1999) and Dellaportas and Pourahmadi (2004). Now assume that $\mathbf{Y}_t = (Y_{1t}, Y_{2t}, \ldots, Y_{pt})', t = 1, \ldots, n$ is a p-dimensional ordered Gaussian time series with mean zero and covariance matrix $\mathbf{\Sigma}_t$. Dellaportas and Pourahmadi (2012) and Pedeli et al. (2015) considered the linear regression of Y_{jt} on its predecessor variables, $Y_{1t}, Y_{2t}, \ldots, Y_{(j-1)t}$ as

$$Y_{jt} = \sum_{k=1}^{j-1} \phi_{jk} Y_{kt} + \epsilon_{jt}, \tag{2.1}$$

where ϕ_{jk}'s are regression coefficients, ϵ_{jt}'s are independent Gaussian variables with mean zero and conditional variances d^2_{jt} for $j = 1, \ldots, p$ and $\epsilon_{1t} = Y_{1t}$. They studied the case where d^2_{jt}'s follow GARCH and log-GARCH, respectively. Allowing the regression coefficients ϕ_{jk}'s to be time-varying, provides a more flexible model to describe the instantaneous dependencies.

2.1 Time-Varying Cholesky GARCH Model

We introduce the TC-GARCH model as the case where the coefficients in (2.1) are time-varying, denoted by ϕ_{jkt}. We study the case where these coefficients evolve through time according to Gaussian random-walk. The state space representation of the new model can be written as

$$\begin{aligned} Y_{jt} &= \mathbf{X}'_{jt} \boldsymbol{\phi}_{jt} + \epsilon_{jt}, \quad t = 1, 2, \ldots, n, \quad j = 2, \ldots, p, \\ \boldsymbol{\phi}_{jt} &= \mathbf{A}_j \boldsymbol{\phi}_{j(t-1)} + \mathbf{a}_{jt} \end{aligned} \tag{2.2}$$

where $\mathbf{X}_{jt} = (Y_{1t}, Y_{2t}, \ldots, Y_{(j-1)t})'$ and $\boldsymbol{\phi}_{jt} = (\phi_{j1t}, \ldots, \phi_{j(j-1)t})'$. The $\mathbf{A}_j$ is a $(j-1) \times (j-1)$ identity matrix, reflecting the assumption of random-walk dynamics for the time-varying Cholesky coefficients, and $\mathbf{a}_{jt} = (a_{j1t}, \ldots, a_{j(j-1)t})'$ are white noises at time t with some covariance matrix $\mathbf{Q}_j$, such that state innovations $\{a_{jt}\}$ are independent of the observation innovations $\{\epsilon_{jt}\}$ for all j and t.
The noise sequences ϵ_{jt}'s follow GARCH(r,q) models for $j = 1, \ldots, p$ as

$$\begin{aligned} \epsilon_{jt} &= d_{jt} z_t, \\ d^2_{jt} &= w_j + \sum_{i=1}^{r} \alpha_{ji} \epsilon^2_{j(t-i)} + \sum_{l=1}^{q} \beta_{jl} d^2_{j(t-l)}, \end{aligned} \tag{2.3}$$

where z_t's are independent and identically distributed (i.i.d) from $N(0, 1)$ distribution. In the matrix representation, (2.2) can be written as

$$\mathbf{T}_t \mathbf{Y}_t = \boldsymbol{\epsilon}_t,$$

where $\boldsymbol{\epsilon}_t = (\epsilon_{1t}, \epsilon_{2t}, \ldots, \epsilon_{pt})$ and $\mathbf{T}_t$ is the unit lower triangular matrix

$$\mathbf{T}_t = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ -\phi_{21t} & 1 & 0 & \cdots & 0 \\ -\phi_{31t} & -\phi_{32t} & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ -\phi_{p1t} & -\phi_{p2t} & \cdots & -\phi_{p(p-1)t} & 1 \end{pmatrix}.$$

Let $\boldsymbol{\Upsilon}_t$ be the information up to t, so

$$\mathbf{D}_t = \mathrm{Cov}(\epsilon_t|\boldsymbol{\Upsilon}_{t-1}) = \mathrm{Cov}(\mathbf{T}_t\mathbf{Y}_t|\boldsymbol{\Upsilon}_{t-1}) = \mathbf{T}_t\mathrm{Cov}(\mathbf{Y}_t|\boldsymbol{\Upsilon}_{t-1})\mathbf{T}_t' = \mathbf{T}_t\boldsymbol{\Sigma}_t\mathbf{T}_t',$$

or equivalently

$$\boldsymbol{\Sigma}_t = \mathbf{T}_t^{-1}\mathbf{D}_t\mathbf{T}_t'^{-1},$$

where $\mathbf{D}_t$ is diagonal matrix as $\mathbf{D}_t = diag(d_{1t}^2, \ldots, d_{pt}^2)$. The $\boldsymbol{\Sigma}_t$ is a symmetric positive-definite matrix for $t = 1, \ldots, n$. This is by the fact that a symmetric matrix $\boldsymbol{\Sigma}_t$ is positive definite if and only if there exists a unique unit lower triangular matrix $\mathbf{T}_t$, with 1's as diagonal entries, and a unique diagonal matrix $\mathbf{D}_t$ with positive diagonal entries such that $\mathbf{T}_t\boldsymbol{\Sigma}_t\mathbf{T}_t' = \mathbf{D}_t$ (Newton, 1988, page 539). Also, $\boldsymbol{\Sigma}_t$'s are invertible matrices for $t = 1, \ldots, n$. This can be verified by the fact that if λ_t is an eigenvalue of $\boldsymbol{\Sigma}_t$ and x_t the corresponding eigenvector, so $\boldsymbol{\Sigma}_t x_t = \lambda_t x_t$. Thus multiplying x_t' from the left, it leads to

$$x_t'\boldsymbol{\Sigma}_t x_t = \lambda_t x_t' x_t = \lambda_t \|x_t\|^2.$$

The left hand side is positive as $\boldsymbol{\Sigma}_t$ is positive definite and x_t is a non-zero vector as it is an eigenvector; so λ_t is positive. If the eigenvalues of a matrix are positive, the matrix is invertible (Hoffman and Kunze, 1971).

Ordering regression variables is an important step in MCD and is a drawback. For a set of p variables there are $p!$ choices. Basford and Tukey (1999) proposed a method for ordering variables based on a scatter plot matrix, nice in the sense that one brings the more correlated variables closer to the main diagonal, called "greedy close algorithm". This simple exploratory data analysis idea has led to the more powerful conceptual idea of banded sample covariance matrix estimation (Bickel and Gel, 2011; Bickel and Levina, 2008). In the context of the ordering variables in the regression, Dodge and Rousson (2000) presented an ordering for variables in setting regression, just for two variables, based on the asymmetric properties correlation coefficients. Dellaportas and Pourahmadi (2004) and also Dellaportas and Pourahmadi (2012) suggested using Akaike information criterion (AIC) or Bayesian information criterion (BIC) for ordering variables in MCD and Pedeli et al. (2015) used the same criterion. In this paper, we consider the overall ordering of regression variables using the BIC criterion. We follow the method of Dellaportas and Pourahmadi (2004) and also Dellaportas and Pourahmadi (2012) and Pedeli et al. (2015) in this order.

3 Estimation

There are two types of parameters that must be estimated, the time-varying regression coefficients ϕ_{jkt} and the GARCH parameters w_j, α_{ji} and β_{jl}. These two types of parameters will be estimated separately.

3.1 Estimation of the Time-Varying Regression Coefficients in TC-GARCH

For estimating $\boldsymbol{\phi}_{jt}$'s, we apply the Kalman filter method through a set of equations that allows updating the estimates of $\boldsymbol{\phi}_{jt}$ by arriving the new observation. For this, some inferences about the distribution of the state vector $\boldsymbol{\phi}_{jt}$ and its initial values are required.

Let $p(\boldsymbol{\phi}_{j0})$ be the probability distribution of state vector at time zero. The state equation (2.2) with such initial distributions determines the distribution of the state vectors $p(\boldsymbol{\phi}_{jt}|\boldsymbol{\phi}_{j(t-1)})$, for $t = 1, \ldots, n$ and $j = 2, \ldots, p$. We refer to these distributions as prior distributions of state vector at time t.

To specify the likelihood model of $p(y_{jt}|\boldsymbol{\phi}_{jt})$, by observing a new measurement y_{jt}, the prior distributions to be updated, such update information about $\boldsymbol{\phi}_{jt}$ expressed by the conditional distribution $p(\boldsymbol{\phi}_{jt}|y_{jt})$, called posterior distribution. Let the prior distribution of $\boldsymbol{\phi}_{j0}$ to be normal with mean vector $\mathbf{m}_{j0} = (m_{j10}, \ldots, m_{j(j-1)0})$ and covariance matrix $\mathbf{P}_{j(0)}$, also the errors ϵ_{jt} and $\mathbf{a}_{jt}$ in (2.2) are normality distributed. So

$$\begin{aligned} p(\boldsymbol{\phi}_{j0}) &= N(\boldsymbol{\phi}_{j0}|\mathbf{m}_{j0}, \mathbf{P_{j0}}), \\ p(y_{jt}|\boldsymbol{\phi}_{jt}) &= N(y_{jt}|\mathbf{X'}_{jt}\boldsymbol{\phi}_{jt}, d^2_{jt}), \\ p(\boldsymbol{\phi}_{jt}|\boldsymbol{\phi}_{j(t-1)}) &= N(\boldsymbol{\phi}_{jt}|\boldsymbol{\phi}_{j(t-1)}, \mathbf{Q}_j), \end{aligned} \tag{3.1}$$

where $\mathbf{Q_j}$ is covariance matrix of the multivariate random walk noises $\mathbf{a}_{jt} = (a_{j1t}, \ldots, a_{j(j-1)t})'$. As Särkkä (2013) explained in Section 2.1 that Bayesian inference provides the equations for computing the posterior distributions and point estimates for any model once the model specification has been set up. He followed the Gaussian approximation of posterior distribution which is introduced by Gelman et al. (2004), and concluded that

$$p(\boldsymbol{\phi}_{jt}|y_{j(1:t)}) \simeq N(\boldsymbol{\phi}_{jt}|\mathbf{m}_{jt}, \mathbf{P}_{jt}),$$

where $\mathbf{m}_{jt}$ and $\mathbf{P}_{jt}$ are mean and covariance matrix of the Gaussian approximation, respectively and can be computed by matching the first two moments of the posterior distribution. So

$$p(\boldsymbol{\phi}_{j(t-1)}|y_{j(1:t-1)}) \simeq N(\boldsymbol{\phi}_{j(t-1)}|\mathbf{m}_{j(t-1)}, \mathbf{P}_{j(t-1)}), \tag{3.2}$$

where $\mathbf{m}_{j(t-1)} = (m_{j1(t-1)}, \ldots, m_{j(j-1)(t-1)})$ and $\mathbf{P}_{j(t-1)}$ are mean and covariance matrix of the Gaussian approximation of the posterior distribution $p(\boldsymbol{\phi}_{j(t-1)}|y_{j(1:t-1)})$ at the time step $t-1$, respectively.

The joint distribution of $\boldsymbol{\phi}_{jt}$ and $\boldsymbol{\phi}_{j(t-1)}$ based on the behaviour of Markovian dynamic models can be written as

$$p(\boldsymbol{\phi}_{jt}, \boldsymbol{\phi}_{j(t-1)}|y_{j(1:t-1)}) = p(\boldsymbol{\phi}_{jt}|\boldsymbol{\phi}_{j(t-1)})p(\boldsymbol{\phi}_{j(t-1)}|y_{j(1:t-1)}).$$

The distribution of $\boldsymbol{\phi}_{jt}$ given the measurement history up to time step $t-1$ can be calculated by integrating over $\boldsymbol{\phi}_{j(t-1)}$

$$p(\boldsymbol{\phi}_{jt}|y_{j(1:t-1)}) = \int p(\boldsymbol{\phi}_{jt}|\boldsymbol{\phi}_{j(t-1)})p(\boldsymbol{\phi}_{j(t-1)}|y_{j(1:t-1)})\, d\boldsymbol{\phi}_{j(t-1)}.$$

As $p(\boldsymbol{\phi}_{jt}|\boldsymbol{\phi}_{j(t-1)})$ is Gaussian by (3.1) and $p(\boldsymbol{\phi}_{j(t-1)}|y_{j(1:t-1)})$ is approximated by (3.2), so

$$p(\boldsymbol{\phi}_{jt}|y_{j(1:t-1)}) \simeq N(\boldsymbol{\phi}_{jt}|\mathbf{m}^-_{jt}, \mathbf{P}^-_{j(t)}) \tag{3.3}$$

where mean and covariance matrix

$$\begin{aligned}\mathbf{m}^-_{jt} &= \mathbf{A}_j\mathbf{m}_{j(t-1)},\\ \mathbf{P}^-_{j(t)} &= \mathbf{A}_j\mathbf{P}_{j(t-1)}\mathbf{A}'_j + \mathbf{Q}_j.\end{aligned}$$

Having Gaussian approximation (3.3) as the prior distribution for the measurements, we obtain the approximation of the posterior distribution $\boldsymbol{\phi}_{jt}$

$$p(\boldsymbol{\phi}_{jt}|y_{j(1:t)}) \simeq N(\boldsymbol{\phi}_{jt}|\mathbf{m}_{jt}, \mathbf{P}_{j(t)}).$$

The mean vector and covariance matrix of the posterior distribution can be updated recursively using the Kalman filter (Harvey, 1989) as follows:

$$\begin{aligned}\mathbf{m}^-_{jt} &= \mathbf{A}_j\mathbf{m}_{j(t-1)}\\ \mathbf{P}^-_{j(t)} &= \mathbf{A}_j\mathbf{P}_{j(t-1)}\mathbf{A}'_j + \mathbf{Q}_j\\ \mathbf{m}_{jt} &= \mathbf{m}^-_{jt} + \mathbf{K}_{jt}\left(y_{jt} - \mathbf{X}'_{jt}\mathbf{m}^-_{jt}\right)\\ \mathbf{P}_{j(t)} &= \mathbf{P}^-_{j(t)} - \mathbf{K_{jt}}\mathbf{X}'_{jt}\mathbf{P}^-_{j(t)}\\ \mathbf{K}_{jt} &= \left[\mathbf{X}'_{jt}\mathbf{P}^-_{j(t)}\mathbf{X}_{jt} + d^2_{jt}\right]^{-1}\mathbf{P}^-_{j(t)}\mathbf{X}_{jt}.\end{aligned} \tag{3.4}$$

Considering the mean square error as the loss function, the $\mathbf{m}_{jt}$'s are the best estimators for the $\boldsymbol{\phi}_{jt}$'s in Bayesian concept (Papoulis and Pillai, 2002, Section 7.3). Suppose we have observed a time series up to $t-1$, then $\hat{\boldsymbol{\phi}}_{j(t-1)}$ is the mean square error estimator of $\boldsymbol{\phi}_{j(t-1)}$ based on information up to this time.

3.2 Estimation of the multivariate GARCH parameters in TC-GARCH

Assuming normality of the multivariate time series $\mathbf{Y}_t$ in (2.2) we have

$$\mathbf{T}_t\mathbf{Y}_t|\boldsymbol{\Upsilon}_{t-1} \sim N_p(0, \mathbf{D}_t) \tag{3.5}$$

and the log-likelihood function can be written as

$$
\begin{aligned}
L(\boldsymbol{\theta}) - K &= \sum_{t=1}^{n}(\log|\boldsymbol{\Sigma}_t| + \mathbf{Y}_t'\boldsymbol{\Sigma}_t^{-1}\mathbf{Y}_t) \\
&= \sum_{t=1}^{n}(\sum_{j=1}^{p}\log d_{jt}^2 + \mathbf{Y}_t'\mathbf{T}_t'\mathbf{D}_t^{-1}\mathbf{T}_t\mathbf{Y}_t) \\
&= \sum_{t=1}^{n}\sum_{j=1}^{p}(\log d_{jt}^2 + \frac{\epsilon_{jt}^2}{d_{jt}^2}) \\
&= \sum_{j=1}^{p}\left\{\sum_{t=1}^{n}\left(log d_{jt}^2 + \frac{\epsilon_{jt}^2}{d_{jt}^2}\right)\right\}
\end{aligned}
\tag{3.6}
$$

where K is some constant not depending on the parameters. We consider the case where d_{jt}^2's follow $GARCH(1,1)$ model as defined in (2.3). The last representation in (3.6) shows that the log-likelihood function can be viewed as a sum of p univariate Gaussian likelihoods for the mutually uncorrelated returns ϵ_{jt}, $j = 1, \dots, p$. It is the source of a rather unique and appealing property of the Cholesky GARCH models that is not available in any other multivariate GARCH models. Let $d_{jt}^2(\boldsymbol{\theta}_j)$ follows a GARCH (1,1) model and $\boldsymbol{\theta}_j = (w_j, \alpha_j, \beta_j)$, $j = 1, \dots, p$ denotes the parameter vector of the GARCH (1,1). Consider the $L(\boldsymbol{\theta}_j)$, $j = 1, \dots, p$ as

$$
L(\boldsymbol{\theta}_j) = \sum_{t=1}^{n}(\log d_{jt}^2(\boldsymbol{\theta}_j) + \frac{\epsilon_{jt}^2}{d_{jt}^2(\boldsymbol{\theta}_j)}).
$$

The derivatives of $L(\boldsymbol{\theta}_j)$ with respect to the parameters are determined as

$$
\frac{\partial(L\boldsymbol{\theta}_j)}{\partial\theta_{j(l)}} = \sum_{t=1}^{n}\frac{1}{2d_{jt}^2(\boldsymbol{\theta}_j)}\left(\frac{\epsilon_{jt}^2}{d_{jt}^2(\boldsymbol{\theta}_j)} - 1\right)\frac{\partial d_{jt}^2(\boldsymbol{\theta}_j)}{\partial\theta_{j(l)}}
$$

where $\theta_{j(l)}$ is the $l - th$ element of $\boldsymbol{\theta}_j$. Partial derivatives of $d_{jt}^2(\boldsymbol{\theta}_j)$ can obtain as

$$
\begin{aligned}
\frac{\partial d_{jt}^2(\boldsymbol{\theta}_j)}{\partial w_j} &= 1 + \beta_j\frac{\partial d_{j(t-1)}^2(\boldsymbol{\theta}_j)}{\partial w_j}, \\
\frac{\partial d_{jt}^2(\boldsymbol{\theta}_j)}{\partial \alpha_j} &= \epsilon_{j(t-1)}^2 + \beta_j\frac{\partial d_{j(t-1)}^2(\boldsymbol{\theta}_j)}{\partial \alpha_j}, \\
\frac{\partial d_{jt}^2(\boldsymbol{\theta}_j)}{\partial \beta_j} &= d_{j(t-1)}^2(\boldsymbol{\theta}_j) + \beta_j\frac{\partial d_{j(t-1)}^2(\boldsymbol{\theta}_j)}{\partial \beta_j}.
\end{aligned}
$$

Some numerical method such as quasi-Newton algorithm can be used to estimate parameters of the GARCH model.

To capture time-varying measurement uncertainty, the observation-error variance is modeled by a GARCH process, while the covariance matrix of the state innovations $\mathbf{Q}_j$ for $j = 2, \ldots, p$ is estimated via the Expectation-Maximization (EM) algorithm. This approach avoids imposing restrictive assumptions on the state-error covariance structure while retaining computational tractability in the state-space estimation framework.

The estimates obtained from the Kalman filter and the univariate GARCH models can be combined to construct the estimated time-varying covariance matrix. Let

$$\widehat{\mathbf{T}}_t = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ -\widehat{\phi}_{21t} & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ -\widehat{\phi}_{p1t} & -\widehat{\phi}_{p2t} & \cdots & 1 \end{pmatrix}$$

and $\widehat{\mathbf{D}}_t = \operatorname{diag}\left(\widehat{d^2_{1t}}, \ldots, \widehat{d^2_{pt}}\right)$. Then the estimated time-varying covariance matrix of the proposed TC-GARCH model is

$$\widehat{\boldsymbol{\Sigma}}_t = \widehat{\mathbf{T}}_t^{-1}\widehat{\mathbf{D}}_t\widehat{\mathbf{T}}_t^{\prime -1}.$$

Consequently, the estimated time-varying correlation matrix is obtained as

$$\widehat{\mathbf{R}}_t = \widehat{\mathbf{V}}_t^{-1/2}\widehat{\boldsymbol{\Sigma}}_t\widehat{\mathbf{V}}_t^{-1/2},$$

where $\widehat{\mathbf{V}}_t = \operatorname{diag}\left(\widehat{\Sigma}_{11,t}, \ldots, \widehat{\Sigma}_{pp,t}\right)$. The matrix $\widehat{\mathbf{R}}_t$ is used throughout the remainder of the paper to evaluate the ability of the proposed TC-GARCH model to capture dynamic instantaneous dependencies.

4 Simulation

We conduct a Monte Carlo simulation study to evaluate the finite-sample performance of the proposed time-varying Cholesky-GARCH model. Let $\boldsymbol{\epsilon}_t = (\epsilon_{1t}, \ldots, \epsilon_{pt})'$ denote the innovation vector with conditional covariance matrix $\mathbf{D}_t = \operatorname{diag}(d^2_{1t}, \ldots, d^2_{pt})$, where, for $j = 1, \ldots, p$, the innovation variances follow independent GARCH(1, 1) processes,

$$d^2_{jt} = \omega_j + \alpha_j\epsilon^2_{j,t-1} + \beta_j d^2_{j,t-1},$$

with $\omega_j > 0$, $\alpha_j \geq 0$, $\beta_j \geq 0$, and $\alpha_j + \beta_j < 1$. The time-varying Cholesky coefficients are generated from a random-walk state equation as in (2.2), such that $\mathbf{a}_t \sim \mathcal{N}(\mathbf{0}, q_{\text{var}}\mathbf{I}_m)$. Thus, q_{var} controls the degree of time variation in the dependence structure, with larger values producing more rapidly changing Cholesky coefficients. Using the coefficients in $\boldsymbol{\phi}_t$, we construct the lower triangular Cholesky matrix of matrix $\mathbf{T}_t$ in (6.1). The observation vector $\mathbf{Y}_t = (y_{1t}, \ldots, y_{pt})'$ is then generated according to $\mathbf{T}_t\mathbf{Y}_t = \boldsymbol{\epsilon}_t$, or equivalently,

$$\mathbf{Y}_t = \mathbf{T}_t^{-1}\boldsymbol{\epsilon}_t.$$

Consequently, the true time-varying covariance matrix is $\mathbf{\Sigma_t} = \mathbf{T}_t^{-1}\mathbf{D}_t\mathbf{T}_t'^{-1}$, and the corresponding true correlation matrix $\mathbf{R}_t$ is obtained from $\mathbf{\Sigma}_t$. We compare the proposed model with DCC-GARCH, Cholesky-GARCH (C-GARCH), Cholesky log-GARCH (CL-GARCH), and Time-varying Cholesky-GARCH (TC-GARCH). The C-GARCH and CL-GARCH models assume constant Cholesky coefficients estimated by ordinary least squares, whereas TC-GARCH allows the coefficients to vary over time and estimate them through the Kalman filtering framework. In the proposed TC-GARCH models, the observation-error variances is modeled using GARCH specifications, while the covariance matrix of the state innovations is estimated via the EM algorithm.

The performance of each method is evaluated by comparing the estimated time-varying correlation matrix $\widehat{\mathbf{R}}_t$ with the true correlation matrix $\mathbf{R}_t$. As a measure of accuracy, we employ the average Frobenius norm error

$$\text{FE} = \frac{1}{n}\sum_{t=1}^{n}\left\|\widehat{\mathbf{R}}_t - \mathbf{R}_t\right\|_F,$$

where $\|\cdot\|_F$ denotes the Frobenius norm. Smaller values of FE indicate more accurate recovery of the true dynamic dependence structure.

To investigate the effects of dimensionality, sample size, and temporal variation, we consider several simulation settings by varying the dimension p, the sample size n, and the state-innovation variance q_{var}. For each setting, the simulation is repeated 200 times, and the reported results are averaged across all replications.

Table 1 reports the average Frobenius norm error between the estimated and true time-varying correlation matrices under different combinations of sample size n, dimension p, and state innovation variance q_{var}. Lower values indicate more accurate estimation of the dynamic correlation structure. Several patterns emerge from the results. First, the estimation error generally increases with the dimension p, reflecting the growing difficulty of estimating high-dimensional correlation matrices. Second, larger values of q_{var} lead to substantially higher errors for all methods, indicating that rapidly changing dependence structures are more challenging to recover than slowly varying ones.

Among the competing approaches, DCC-GARCH consistently produces the largest estimation errors across all simulation settings. The Cholesky-based methods, namely C-GARCH and CL-GARCH, provide considerable improvements over DCC, particularly in moderate- and high-dimensional settings. The proposed TC-GARCH method further reduces the estimation error in nearly all scenarios, demonstrating the benefit of incorporating time-varying Cholesky coefficients through a state-space framework. The advantage of TC-GARCH is particularly evident when $q_{\text{var}} = 0.001$ and 0.01, where the dependence structure evolves relatively smoothly over time. Although the performance gap narrows when $q_{\text{var}} = 0.1$, the proposed method remains competitive and generally achieves the smallest error. Overall, the results suggest that explicitly modeling the temporal evolution of the Cholesky coefficients improves the estimation of dynamic correlation matrices, especially in low- and moderate-variation environments.

Table 1: Average Frobenius norm error under different simulation settings.

n	p	q_{var}	DCC-GARCH	C-GARCH	CL-GARCH	TC-GARCH
200	5	0.001	2.8355	0.3784	0.4285	0.3001
	5	0.010	2.8794	0.4790	0.5270	0.4181
	5	0.100	4.0012	1.7436	1.7756	1.7242
	10	0.001	4.5906	0.8579	0.9776	0.6963
	10	0.010	4.7444	1.0988	1.2117	0.9816
	10	0.100	8.0468	4.9027	4.9272	4.8773
	15	0.001	6.1670	1.3359	1.5210	1.0812
	15	0.010	6.4166	1.7440	1.9024	1.5701
	15	0.100	12.2539	8.7985	8.8085	8.8011
500	5	0.001	2.8221	0.2522	0.3295	0.1890
	5	0.010	2.9359	0.5182	0.5721	0.4818
	5	0.100	4.3418	2.1988	2.2168	2.1640
	10	0.001	4.5660	0.5697	0.7647	0.4274
	10	0.010	4.9376	1.2135	1.3352	1.1338
	10	0.100	8.9868	6.2096	6.1971	6.1908
	15	0.001	6.1628	0.9198	1.2278	0.6797
	15	0.010	6.8454	2.0236	2.2078	1.8791
	15	0.100	12.2958	9.1262	9.1167	9.1041
800	5	0.001	3.0278	0.5771	0.6279	0.5362
	5	0.010	2.8276	0.1937	0.2857	0.1473
	5	0.100	4.5951	2.3285	2.3461	2.2846
	10	0.001	4.5828	0.4701	0.7059	0.3581
	10	0.010	5.1876	1.4170	1.5308	1.3034
	10	0.100	9.2681	6.6192	6.6088	6.5673
	15	0.001	6.1414	0.7493	1.1219	0.5622
	15	0.010	7.1453	2.3918	2.5552	2.1714
	15	0.100	12.9390	10.2511	10.2453	10.2184

5 Evaluation criteria

To evaluate the estimation ability of the TC-GARCH model in comparison with to other models, we use four loss functions, two quadratic loss functions $\triangle_{1t}$ and $\triangle_{2t}$, the mean absolute error (MAE_t) and MSE_t. Also, to investigate the forecasting capability of our proposed TC-GARCH model, we use one-step-ahead VaR.

5.1 Loss functions

To measure the accuracy of the time-varying correlation matrix estimation $\hat{\mathbf{R}}_t$, we consider the moving blocks approach as a benchmark to calculate $\mathbf{R_t}$ (Lopes et al., 2012). The loss functions $\triangle_{1t}$ and $\triangle_{2t}$, MAE$_t$ and MSE$_t$ are obtained as

$$\triangle_{1t} = tr(\mathbf{R_t}^{-1}\widehat{\mathbf{R}}_t - \mathbf{I})^2,$$
$$\triangle_{2t} = tr(\widehat{\mathbf{R}}_t^{-1}\mathbf{R_t} - \mathbf{I})^2,$$
$$MAE_t = \frac{1}{p^2}\sum_{i=1}^{p}\sum_{j=1}^{p}|\hat{r}_{ijt} - r_{ijt}|,$$
$$MSE_t = \frac{1}{p^2}\sum_{i=1}^{p}\sum_{j=1}^{p}(\hat{r}_{ijt} - r_{ijt})^2,$$

where $\hat{r}_{ijt}$'s are defined as Section 4. Also $r_{ijt} = \frac{\sigma_{ijt}}{\sigma_{it}\sigma_{jt}}$ where σ_{ijt}, σ_{it}, and σ_{jt} are elements of the time-varying covariance matrix $\mathbf{\Sigma}_t$ obtained by moving blocks approach. For the comparison of different models we rely on the averages of the above measures over t. Selection of the block size q is based on criteria $\triangle_{it}, i = 1, 2$ with the moving blocks estimator serving as $\mathbf{R}_t$ and the time-varying correlation matrix evaluated by the TC-GARCH model and other models as $\hat{\mathbf{R}}_t$. More specifically, the average loss functions $\hat{\triangle}_i = \sum_{t=1}^{n} \triangle_{it}/n, i = 1, 2$ are calculated for several values of q. Let $q_1, \ldots, q_m, \ldots, q_M$ be the ordered series of q-values, the value q_m is that stabilizes the absolute difference $|\hat{\triangle}_i^{(q_m)} - \hat{\triangle}_i^{(q_{m-1})}|$ for two loss functions $\triangle_{1t}$ and $\triangle_{2t}$. According to this procedure, we chose $q = 41$ in our real data set.

5.2 VaR **forecasting**

To evaluate the forecasting performance of the TC-GARCH model in forecasting future returns, we consider the one-step-ahead VaR. A portfolio with weight w_i in stock market index i has return as

$$r_t = \mathbf{w}'\mathbf{Y}_t$$

where $\mathbf{Y}_t$ denotes the vector of stock market returns and $\mathbf{w}'$ is the weight vector of the any given indices.
The portfolio variances are as follows

$$\sigma_t^2 = \mathbf{w}'\mathbf{\Sigma}_t\mathbf{w}$$

where $\mathbf{\Sigma}_t$ is a conditional covariance to be obtained by multivariate GARCH models. The one-step-ahead VaR with probability η, VaR(η), is calculated as VaR$_t(\eta) = F^{-1}(\eta)\sigma_t$, and $F^{-1}(\eta)$ is the inverse distribution of the standardized returns (r_t/σ_t). Some likelihood ratio (LR) tests were done to appraise the precision of the VaR forecasts from competing models say, unconditional coverage test, independent test and conditional coverage test (Ardia, 2009; Valizadeh et al., 2021).

Unconditional coverage test

In order to test whether VaR forecasts supply the pre-specified probability, the unconditional coverage (UC) test is used. If the real loss exceeds the VaR forecasts, this is named an "exception", which is a Bernoulli random variable with probability ξ. The $H_0 : \xi = \eta$ is the null hypothesis of the UC test and the LR statistic of the unconditional coverage (LR_{UC}) is described as:

$$LR_{UC} = -2\log(\frac{\eta^n(1-\eta)^{T-n}}{\hat{\xi}^n(1-\hat{\xi})^{T-n}}).$$

Where T is the number of the forecasting samples, n is the number of the exceptions and $\hat{\xi} = \frac{n}{T}$ is the MLE of ξ under H_1. Then under H_0 the LR_{UC} is asymptotically distributed as a χ^2 random variable with one degree of freedom.

Independent test

The exceptions must be distributed over the entire sample period independently and do not make clusters. To check the clustering of the exceptions the independent (IND) test is used. The null hypothesis of the IND test is defined as the probability of an exception on time t is not influenced by what happened the time before. Hence, $H_0 : \xi_{10} = \xi_{00}$, and ξ_{ij} shows that the probability of an event i on time $t-1$ should be followed by an event j on time t when $i, j = 0, 1$. The LR statistic of the IND test (LR_{IND}) is organized as:

$$LR_{IND} = -2\log(\frac{\hat{\xi}^n(1-\hat{\xi})^{T-n}}{\hat{\xi}_{01}^{n_{01}}(1-\hat{\xi}_{01})^{n_{00}}\hat{\xi}_{11}^{n_{11}}(1-\hat{\xi}_{11})^{n_{10}}}),$$

where n_{ij} is the number of observations with value i followed by value j $(i, j = 0, 1)$, $\hat{\xi}_{01} = \frac{n_{01}}{n_{00}+n_{01}}$ and $\hat{\xi}_{11} = \frac{n_{11}}{n_{10}+n_{11}}$. Under H_0, the LR_{IND} is asymptotically distributed as a χ^2 random variable with one degree of freedom.

Conditional coverage test

The conditional coverage (CC) test covers the properties of both the UC and IND tests. The null hypothesis of the CC test checks both the exception cluster and the agreement of the exceptions with the VaR confidence level. The $H_0 : \xi_{01} = \xi_{11} = \eta$ is the null hypothesis of the CC test. The LR statistics of the CC test (LR_{CC}) is formed as:

$$LR_{CC} = -2\log(\frac{\eta^n(1-\eta)^{T-n}}{\hat{\xi}_{01}^{n_{01}}(1-\hat{\xi}_{01})^{n_{00}}\hat{\xi}_{11}^{n_{11}}(1-\hat{\xi}_{11})^{n_{10}}}).$$

Where under H_0, LR_{CC} is asymptotically distributed as a χ^2 random variable with two degrees of freedom.

6 Real data

We study a financial dataset consisting of n = 251 monthly stock returns of 12 different US bluechips. We show the competing performance of our proposed TC-GARCH

model in comparison with the C-GARCH, CL-GARCH, and DCC-GARCH models in calculating instantaneous dependencies. Also, we consider GARCH(1,1) in all the above models.
Note that to estimate the time-varying correlation matrix $\mathbf{R}_t$, we performed the following steps in order.

- Input time series and do ordering $Y_1, \ldots, Y_p$ based on BIC criterion.
- For any $t = 1, \ldots, n$, estimate $\boldsymbol{\phi}_{jt}$'s for $j = 2, \ldots, p$ based on Kalman filter at (3.4) and create matrix $\hat{\mathbf{T}}_t$ as the following

$$\hat{\mathbf{T}}_t = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ -\hat{\phi}_{21t} & 1 & 0 & \cdots & 0 \\ -\hat{\phi}_{31t} & -\hat{\phi}_{32t} & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ -\hat{\phi}_{p1t} & -\hat{\phi}_{p2t} & \cdots & -\hat{\phi}_{p(p-1)t} & 1 \end{pmatrix}. \tag{6.1}$$

- For any $t = 1, \ldots, n$ and $j = 2, \ldots, p$, using $\hat{\boldsymbol{\phi}}_{jt}$'s and equation (2.2), we obtain ϵ_{jt}'s as inputs GARCH models in (2.3) and estimate the GARCH parameters by (3.6). Then $\hat{d}^2_{jt}$'s are calculated and used as entries of diagonal matrix $\hat{\mathbf{D}}_t$.
- At each time t, we estimate time-varying matrix $\boldsymbol{\Sigma}_t$ as $\hat{\boldsymbol{\Sigma}}_t = \hat{\mathbf{T}}_t^{-1}\hat{\mathbf{D}}_t\hat{\mathbf{T}}_t'^{-1}$ and calculate time-varying correlation matrix $\hat{\mathbf{R}}_t$ based on $\hat{\boldsymbol{\Sigma}}_t$.

6.1 Monthly stock return of 12 US bluechips

The financial dataset consists of compound monthly log returns from a panel of 12 US bluechips. These 12 US bluechips include companies Oracle Corp. (ORCL), Apple Inc. (AAPL), Amgen Inc. (AMGN), Cisco Systems Inc. (CSCO), Intel Corp. (INTC), Schlumberger Ltd. (SLB), Microsoft Corp. (MSFT), Bank of America Corp. (BAC), Citigroup Inc. (C), Caterpillar Inc. (CAT), Dow Chemical Co. (DOW) and American Express Co. (AXP). The n = 251 stock returns were collected from January 1990 to December 2010. The BIC criterion is used to determine the order of variables. Table 2 presents the results of comparing the performance of different models. Comparing the results, the TC-GARCH model has the best performance and C-GARCH model has the worst performance in terms of the $\triangle_1$, MSE and MAE. Also, DCC-GARCH model outperforms C-GARCH and CL-GARCH models by presenting lower value of $\triangle_1$, $\triangle_2$, MSE and MAE.

To evaluate the forecast ability of the TC-GARCH model, the first 151 observations are used to estimate the parameters and the last 100 samples are used as out-of-sample. One-step-ahead VaR forecasts at the risk level of $\eta = 0.01, 0.05$ and 0.1 are calculated for the models and the accuracy tests are performed. The results of LR tests, the number of the expected exceptions (Ex.e) and empirical exceptions (Em.e) are summarized in Table 3. It is observed that the CL-GARCH model has a better performance in

Table 2: The averages of two loss functions ($\triangle_1$ and $\triangle_2$), MAE and MSE for DCC-GARCH, C-GARCH, CL-GARCH, and TC-GARCH models in stock return of 12 US bluechips

Method	$\triangle_1$	$\triangle_2$	MAE	MSE
DCC-GARCH	38.474	0.0193	0.1315	0.03093
C-GARCH	46.898	0.3704	0.19213	0.06503
CL-GARCH	48.979	0.2932	0.1848	0.0609
TC-GARCH	27.211	0.0429	0.1115	0.0244

Table 3: VaR forecasting results for DCC-GARCH, C-GARCH, CL-GARCH and TC-GARCH models for the stock returns of 12 US bluechips at $\eta = 0.01,\ 0.05,\ 0.10$.

Measure	Model	0.01	0.05	0.10
Ex.e	–	1	5	10
Em.e	DCC-GARCH	4	6	11
	C-GARCH	2	5	10
	CL-GARCH	3	5	9
	TC-GARCH	1	3	10
LR_{UC}	DCC-GARCH	5.182	0.1984*	0.1079*
	C-GARCH	0.7827*	0*	0*
	CL-GARCH	2.632*	0*	0.1145*
	TC-GARCH	0*	0.976*	0*
LR_{IND}	DCC-GARCH	8.6431	11.249	7.5443
	C-GARCH	5.696	6.401	8.8506
	CL-GARCH	3.686*	6.401	6.885
	TC-GARCH	0.0406*	7.505	8.8506
LR_{CC}	DCC-GARCH	13.825	11.4476	7.652
	C-GARCH	6.4788	6.401	8.8506
	CL-GARCH	6.318	6.401	7.00002
	TC-GARCH	0.0406*	8.48	8.8506

Notes: (1) At the 5% significance level the critical values of LR_{UC} and LR_{IND} are 3.84, and for LR_{CC} is 5.99.
(2) * indicates that the model passes the test at the 5% significance level.

comparison with the C-GARCH and DCC-GARCH models according to the number of tests passed. The C-GARCH model outperforms the DCC-GARCH model. Also the TC-GARCH model has the best performance and DCC-GARCH has the worst performance in terms of the number of tests passed.

7 Conclusion

The dependencies are one of the most important aspects of multivariate models. A time-varying version of the Cholesky GARCH model is proposed, say TC-GARCH which is capable of capturing the dynamic aspects of the dependencies in multivariate time series. The time-varying regression coefficients in the TC-GARCH model are updated through Gaussian random-walk and are estimated by the Kalman filter based on the most recent information as it becomes available. Also, this model inherently guarantees positive-definiteness constraint of the time-varying covariance matrix. The simulation and real data studies show the competitive performance of the TC-GARCH model in fitting and forecasting data.

While the proposed TC-GARCH model provides a flexible and computationally tractable framework for modeling dynamic dependencies, its application to high-dimensional time series remains challenging. In particular, the model relies on an ordering of the variables and requires the estimation of a number of time-varying Cholesky coefficients that grows quadratically with the dimension. Although the BIC criterion is used to determine the ordering, searching for a near-optimal ordering becomes computationally prohibitive as p increases due to the $p!$ possible permutations. Furthermore, the large number of dynamic parameters may increase estimation variability and computational cost, indicating that some form of sparsity is desirable in high-dimensional settings. Future work may therefore explore sparse or banded Cholesky structures to enhance both scalability and statistical efficiency.

In addition, extending the volatility component to more general GARCH specifications, such as FIGARCH or hyperbolic GARCH models, may provide further insight into the effects of long-memory volatility on instantaneous dependencies.

References

Alexander C. Orthogonal GARCH. Mastering Risk. 2001;2:21–38.

Ardia D. Bayesian estimation of a Markov-switching threshold asymmetric GARCH model with Student t innovations. Econometrics Journal. 2009;12(1):105–126.

Basford KE, Tukey JW. Graphical Analysis of Multiresponse Data Illustrated with a Plant Breeding Trial. London: Chapman and Hall/CRC Press; 1999.

Bickel PJ, Gel YR. Banded Regularization of Autocovariance Matrices in Application to Parameter Estimation and Forecasting of Time Series. Journal of the Royal Statistical Society: Series B (Statistical Methodology). 2011;73(5):711–728.

Bickel PJ, Levina E. Covariance Regularization by Thresholding. The Annals of Statistics. 2008;36(6):2577–2604.

Bollerslev T. Generalized autoregressive conditional heteroscedasticity. Journal of Econometrics. 1986;31(3):307–327.

Bollerslev T. Modelling the coherence in short-run nominal exchange rates: a multivariate generalized ARCH approach. The Review of Economics and Statistics. 1990;72(3):498–505.

Bollerslev T, Engle RF, Wooldridge JM. A capital asset pricing model with time-varying covariances. The Journal of Political Economy. 1988;96(1):116–131.

Darolles S, Francq C, Laurent S. Asymptotics of Cholesky GARCH models and time-varying conditional betas. Journal of Econometrics. 2018;204(2):223–247.

Dellaportas P, Pourahmadi M. Large Time-Varying Covariance Matrices with Applications to Finance. Department of Statistics, Athens University of Economics and Business; 2004.

Dellaportas P, Pourahmadi M. Cholesky–GARCH models with applications to finance. Statistics and Computing. 2012;22(4):849–855.

Dodge Y, Rousson V. Direction dependence in a regression line. Communications in Statistics - Theory and Methods. 2000;29(9):1957–1972.

Engle RF. Dynamic conditional correlation: a simple class of multivariate generalized autoregressive conditional heteroskedasticity models. Journal of Business and Economic Statistics. 2002;20(3):339–350.

Engle RF, Kroner KF. Multivariate simultaneous generalized ARCH. Econometric Theory. 1995;11(1):122–150.

Gelman A, Carlin JB, Stern HS, Rubin DB. Bayesian Data Analysis. 2 ed. Chapman and Hall; 2004.

Harvey AC. Forecasting, Structural Time Series Models and the Kalman Filter. Cambridge: Cambridge University Press; 1989.

Hoffman KM, Kunze R. Linear Algebra. Prentice-Hall; 1971.

Kang X, Deng X, Tsui KW, Pourahmadi M. On variable ordination of modified Cholesky decomposition for estimating time-varying covariance matrices. International Statistical Review. 2020;88(2):363–394.

Kemper JPFM. Sparse Cholesky-GARCH Models. Erasmus School of Economics, Erasmus University Rotterdam, (masterthesis). 2020;.

Lopes HF, McCulloch RE, Tsay RS. Cholesky Stochastic Volatility Models for High-Dimensional Time Series. (techreport). 2012;.

Muriel Torrero NO. A note on the tails of the GO-GARCH process. Stat Journal. 2014;3:23–30.

Näf J, Paolella MS, Polak P. Heterogeneous tail generalized COMFORT modeling via Cholesky decomposition. Journal of Multivariate Analysis. 2019;172:84–106.

Newton HJ. TIMESLAB: A Time Series Analysis Laboratory. Pacific Grove, CA: Wadsworth and Brooks/Cole; 1988.

Paolella MS, Polak P, Walker PS. A Non-Elliptical Orthogonal GARCH Model for Portfolio Selection under Transaction Costs. Journal of Banking and Finance. 2021;125:106–146.

Papoulis A, Pillai SU. Probability, Random Variables, and Stochastic Processes. 4 ed. McGraw-Hill Higher Education; 2002.

Pedeli X, Fokianos K, Pourahmadi M. Two Cholesky-Log-GARCH Models for Multivariate Volatilities. Statistical Modelling. 2015;15(3):233–255.

Pourahmadi M. Joint mean-covariance models with applications to longitudinal data: unconstrained parameterization. Biometrika. 1999;86(3):677–690.

Särkkä S. Bayesian Filtering and Smoothing. Cambridge University Press; 2013.

Tse YK, Tsui AKC. A multivariate GARCH model with time-varying correlations. Journal of Business and Economic Statistics. 2002;20(3):351–362.

Valizadeh T, Rezakhah S, Mohammadi Basatini F. On Time-Varying Amplitude HGARCH Model. International Journal of Finance and Economics. 2021;26:2538–2547.

Van der Weide R. GO-GARCH: A Multivariate Generalized Orthogonal GARCH Model. Journal of Applied Econometrics. 2002;17(5):549–564.

Zheng T, Ye S. Cholesky GAS models for large time-varying covariance matrices. Journal of Management Science and Engineering. 2024;9(1):115–142.

Zhu X, Chen Y, Hu J. Estimation of banded time-varying precision matrix based on SCAD and group lasso. Computational Statistics & Data Analysis. 2024;189:107849.