Statistics and Its Interface Volume 13 (2020) 77–89

Modeling RCOV matrices with a generalized threshold conditional autoregressive Wishart model

Yan Cui*, Fukang Zhu, and Wai Keung Li

In this article, we propose a generalized threshold conditional autoregressive Wishart (GTCAW) model to analyze the dynamics of the realized covariance (RCOV) matrices. This model extends the idea of [29] to a threshold framework. It is believed that, as in many financial time series, the dynamic of RCOV matrices exhibits nonlinearity and may be better explained by a threshold type model. The noncentrality matrix and scale matrix of the Wishart distribution are piecewise linear driven by the lagged values of RCOV matrices and retain two different sources of dynamics. The GTCAW model guarantees the symmetry and positive definiteness of RCOV matrices, some simulation results on the maximum likelihood estimation are also given. Real data examples based on daily RCOV matrices present the nonlinear behavior in these time series and the usefulness of the proposed model.

AMS 2000 subject classifications: Primary 91B84; secondary 62M10.
Keywords and phrases: GTCAW, RCOV matrices, Threshold, Volatility, Wishart distribution.

1. INTRODUCTION

Volatility is a common measure of risk in financial time series and multivariate volatility modeling is also of vital importance in many investment decisions such as the calculation of derivative prices, portfolio optimization and risk management. The most popular volatility model among practitioners is the multivariate generalized autoregressive conditional heteroscedastic (MGARCH) model introduced by [5], which is widely accepted to describe the properties of financial market returns. Later, the multivariate stochastic volatility (MSV) model proposed by [15] provides an alternative to GARCH models in accounting for the time-varying and persistent volatility. Recent reviews on MGARCH and MSV models can be found in [4] and [2], respectively. In practice, the covariance of returns is unknown and is either treated as measurable given past observations in the case of MGARCH models or is assumed to be a latent quantity in the case of MSV models.

*Corresponding author.

Recently, an alternative approach of direct modeling covariance has attracted substantial interest. By using the high-frequency returns data, one can construct the realized covariance (RCOV) matrix as a precise estimate for the covariance matrix of low-frequency returns (see e.g., [1], [3]). RCOV matrix provides a more accurate measure of daily ex-post covariation that is observable. For instance, econometric forecasting gains were demonstrated in [13], [17] and [18] while improvements in portfolio choice can be found in [17] and [8].

Modeling RCOV matrices offers much improvements over conventional MGARCH and MSV models, but it should satisfy the requirement that each covariance matrix must be symmetric and positive definite at any time point in the process. One of the most flexible methods to deal with this requirement is to use the matrix-variate distributions with support restricted to the set of all symmetric and positive definite matrices. The pioneer time series models for modeling RCOV matrices are based on Wishart distributions. For example, [14] introduced the Wishart autoregressive (WAR) model which involves a Wishart distribution with a time-varying noncentrality matrix and a constant but nonzero scale matrix. Later, the conditional autoregressive Wishart (CAW) model was proposed by [13], which involves a central Wishart distribution with a time-varying scale matrix that has the BEKK-GARCH specification of [12]. [29] generalized WAR and CAW models, making the noncentrality and scale matrix of the Wishart distribution both time-varying in two different dynamic structures.

In financial econometrics, it is well known that nonlinearity is a salient feature of volatility. Therefore, nonlinear time series modeling is worthy paying more attention to. This class of models captures the dynamic behavior of time series including limit cycles, amplitude dependent frequencies and jump phenomena by switching the regimes. [25] proposed the self-exciting threshold autoregressive (SETAR) model to capture nonlinear phenomena in financial data, [20] introduced the double-threshold ARCH model and [7] extended it to the double-threshold GARCH model to handle the dynamics in the mean and variance process. During these two decades, the high-dimensional nature of economic and financial data leads researchers to multivariate threshold modeling. [26] first proposed the testing and modeling procedure on multivariate threshold autoregressive models and applied them to high-frequency financial data examples. [22] considered a regime switching dynamic correlation model with constant correlation within a regime but different correlations across regimes. The recent works of [19] proposed a threshold varying conditional correlation model to capture the asymmetric behavior of the mean and the variance in financial time series. For more properties and applications, see a thorough review on the development of the family of the threshold time series models in the finance literature by [10].

Motivated by multivariate threshold models, we propose a generalized threshold conditional autoregressive Wishart (GTCAW) model to analyze the dynamics of the RCOV matrices. This model extends the idea of [29] to a threshold framework. The noncentrality matrix and scale matrix of the Wishart distribution are both piecewise linear driven by the lagged values of RCOV matrices and retain two different sources of dynamics. The GTCAW model is also in accord with the requirement that each RCOV matrix guarantees the symmetry and positive definiteness as mentioned before. The parameters can be estimated by maximum likelihood estimation and the model is applied to empirical data examples to show its adequacy in terms of model fitting and forecasting.

The rest of the paper is organized as follows. In Section 2, we introduce the generalized threshold conditional autoregressive Wishart model. Simulation experiments are performed in Section 3. Section 4 reports the empirical data analysis by using the proposed model on daily RCOV matrices for modelling stocks from the New York Exchange. At last, we conclude in Section 5. Below, we largely follow the notations in [29].

2. GENERALIZED THRESHOLD CONDITIONAL AUTOREGRESSIVE WISHART MODEL

2.1 Model formulation

Let $\boldsymbol{Y}_t = (Y_{ij,t})$ be a stochastic, symmetric and positive definite RCOV matrix of n asset returns observed at time t $(t = 1, ..., T)$. The matrix $\boldsymbol{Y}_t$ given the past observations $\mathcal{F}_{t-1} = \{\boldsymbol{Y}_{t-1}, \boldsymbol{Y}_{t-2}, ...\}$ is assumed to follow a Wishart distribution:

$$\boldsymbol{Y}_t|\mathcal{F}_{t-1} \sim W_n(\nu, \boldsymbol{\Lambda}_t, \boldsymbol{\Sigma}_t), \tag{1}$$

where $\nu > n-1$ is the degrees of freedom (DF), $\boldsymbol{\Lambda}_t = (\Lambda_{ij,t})$ is the $n \times n$ symmetric and positive semi-definite noncentrality matrix and $\boldsymbol{\Sigma}_t = (\Sigma_{ij,t})$ is the $n \times n$ symmetric and positive definite scale matrix. The density function of $\boldsymbol{Y}_t$ (see [21], p.442) is given by:

(2)

$$f(\boldsymbol{Y}_t|\mathcal{F}_{t-1}) = 2^{-\nu n/2}\pi^{-n(n-1)/4}\left[\prod_{i=1}^{n}\Gamma\left(\frac{\nu+1-i}{2}\right)\right]^{-1} \times (\det \boldsymbol{\Sigma}_t)^{-\nu/2}(\det \boldsymbol{Y}_t)^{(\nu-n-1)/2} \times \exp\left\{-\frac{1}{2}\mathrm{tr}\left[\boldsymbol{\Sigma}_t^{-1}(\boldsymbol{Y}_t+\boldsymbol{\Lambda}_t)\right]\right\} \times {}_0F_1\left(\frac{\nu}{2};\frac{1}{4}\boldsymbol{\Sigma}_t^{-1}\boldsymbol{\Lambda}_t\boldsymbol{\Sigma}_t^{-1}\boldsymbol{Y}_t\right),$$

where $\Gamma(\cdot)$ is the gamma function and ${}_0F_1$ is the hypergeometric function of matrix argument. Let $l_0 < l_1 < \cdots < l_{s-1} < l_s$ be a partition of the real line, where $l_0 = -\infty$ and $l_s = \infty$. Let d be the delay parameter and r_{t-d} be a real-valued threshold variable. We assume that the matrices $\boldsymbol{\Lambda}_t$ and $\boldsymbol{\Sigma}_t$ are driven by the lagged values of $\boldsymbol{Y}_t$ in each regime $(j = 1, ..., s)$:

$$\boldsymbol{\Lambda}_t = \sum_{h=1}^{R_j}\boldsymbol{M}_h^{(j)}\boldsymbol{Y}_{t-h}(\boldsymbol{M}_h^{(j)})', \tag{3}$$

$$\boldsymbol{\Sigma}_t = \boldsymbol{C}^{(j)}(\boldsymbol{C}^{(j)})' + \sum_{i=1}^{P_j}\boldsymbol{B}_i^{(j)}\boldsymbol{\Sigma}_{t-i}(\boldsymbol{B}_i^{(j)})' + \sum_{k=1}^{Q_j}\boldsymbol{A}_k^{(j)}\boldsymbol{Y}_{t-k}(\boldsymbol{A}_k^{(j)})', \quad l_{j-1} < r_{t-d} \le l_j, \tag{4}$$

where $\boldsymbol{M}_h^{(j)}$ is an $n \times n$ parameter matrix, $\boldsymbol{C}^{(j)}$ is an $n \times n$ lower triangular matrix, $\boldsymbol{A}_k^{(j)} = \mathrm{diag}(\alpha_{1k}^{(j)}, ..., \alpha_{nk}^{(j)})$ and $\boldsymbol{B}_i^{(j)} = \mathrm{diag}(\beta_{1i}^{(j)}, ..., \beta_{ni}^{(j)})$ are $n \times n$ diagonal parameter matrices in the jth regime and subject to $\sum_{k=1}^{Q_j}(\alpha_{hk}^{(j)})^2 + \sum_{i=1}^{P_j}(\beta_{hi}^{(j)})^2 < 1$ for $h = 1, ..., n$. One can see that equation (3) accounts for the autoregressive property, equation (4) resembles the diagonal BEKK model of [12] and captures conditional heteroscedasticity. Besides, each RCOV matrix of the process is guaranteed to be symmetric and positive definite as long as the initial matrices $\boldsymbol{\Sigma}_0, \boldsymbol{\Sigma}_{-1}, ..., \boldsymbol{\Sigma}_{-P_j+1}$ are symmetric and positive definite. Then, we shall call the model defined by equations (1)-(4) the jth regime generalized threshold conditional autoregressive Wishart (GTCAW)(P_j, Q_j, R_j, s) model, $j = 1, ..., s$. We will assume $\boldsymbol{Y}_t$ to be stationary and ergodic with finite fourth order moments.

For simplicity, the same threshold structure of the autoregressive and conditional variance equations are considered. Thus, the GTCAW(P_j, Q_j, R_j, s) model will be simplified as GTCAW(p, q, r, s) if $P_j = p$, $Q_j = q$ and $R_j = r$ for any $j = 1, ..., s$. Similar to [12], we suppose that the main diagonal elements in $\boldsymbol{C}^{(j)}$ and the first diagonal element of each $\boldsymbol{A}_k^{(j)}, \boldsymbol{B}_i^{(j)}$ and $\boldsymbol{M}_h^{(j)}$ in the jth regime are restricted to be positive, which guarantees that the model is identifiable. It is obvious that the GTCAW(p, q, r, s) model is an extension of the generalized conditional autoregressive Wishart (GCAW) model in [29] for capturing nonlinearity. Since the GCAW model consists of the WAR model of [14] and the CAW model of [13], we know that if $\boldsymbol{M}_1^{(j)} = \cdots = \boldsymbol{M}_r^{(j)} = 0$ (i.e.,

$r=0$), the GTCAW(p,q,r,s) model becomes the threshold CAW model, denoted as TCAW(p,q,s). If $\boldsymbol{A}_1^{(j)}=\cdots=\boldsymbol{A}_q^{(j)}=\boldsymbol{B}_1^{(j)}=\cdots=\boldsymbol{B}_p^{(j)}=0$ (i.e., $p=q=0$), then the GTCAW(p,q,r,s) model reduces to the threshold WAR model, denoted as TWAR(r,s).

To define the standardized residuals for model diagnostics, first we derive the vector representation of the GTCAW(p,q,r,s) model. Denote vech$(\cdot)$ as the operator that transforms a matrix by stacking the lower triangular part including the diagonal of the matrix into a vector, and vec$(\cdot)$ as the operator that stacks all columns of a matrix into a vector. Let $\mathcal{L}_n$ and $\mathcal{D}_n$ denote the elimination and duplication matrix, respectively, such that for any symmetric $n\times n$ matrix $\boldsymbol{Y}$, $\text{vec}(\boldsymbol{Y})=\mathcal{D}_n(\boldsymbol{Y})$ and $\text{vech}(\boldsymbol{Y})=\mathcal{L}_n\text{vec}(\boldsymbol{Y})$. Define $\boldsymbol{y}_t=\text{vech}(\boldsymbol{Y}_t)$, $\boldsymbol{\lambda}_t=\text{vech}(\boldsymbol{\Lambda}_t)$, $\boldsymbol{\sigma}_t=\text{vech}(\boldsymbol{\Sigma}_t)$ and $\boldsymbol{c}^{(j)}=\text{vech}(\boldsymbol{C}^{(j)}(\boldsymbol{C}^{(j)})')$. Then the vector representations of equations (3) and (4) turn out to be

$$(5)\qquad \boldsymbol{\lambda}_t=\sum_{h=1}^{r}\mathcal{M}_h^{(j)}\boldsymbol{y}_{t-h},$$

$$(6)\qquad \boldsymbol{\sigma}_t=\boldsymbol{c}^{(j)}+\sum_{i=1}^{p}\mathcal{B}_i^{(j)}\boldsymbol{\sigma}_{t-i}+\sum_{k=1}^{q}\mathcal{A}_k^{(j)}\boldsymbol{y}_{t-k},\qquad l_{j-1}<r_{t-d}\le l_j,$$

where $\mathcal{M}_h^{(j)},\mathcal{A}_k^{(j)}$ and $\mathcal{B}_i^{(j)}$ are square matrices with dimension $n(n+1)/2$ such that

$$\mathcal{M}_h^{(j)}=\mathcal{L}_n(\boldsymbol{M}_h^{(j)}\otimes\boldsymbol{M}_h^{(j)})\mathcal{D}_n,$$
$$\mathcal{B}_i^{(j)}=\mathcal{L}_n(\boldsymbol{B}_i^{(j)}\otimes\boldsymbol{B}_i^{(j)})\mathcal{D}_n,$$
$$\mathcal{A}_k^{(j)}=\mathcal{L}_n(\boldsymbol{A}_k^{(j)}\otimes\boldsymbol{A}_k^{(j)})\mathcal{D}_n,$$

$j=1,...,s$ and $\otimes$ denotes the Kronecker product.

Next, we will give the conditional expectation for the jth ($j=1,...,s$) regime GTCAW (p,q,r,s) model as follows

$$(7)\qquad E(\boldsymbol{y}_t|\mathcal{F}_{t-1})=\boldsymbol{\lambda}_t+\nu\boldsymbol{\sigma}_t$$
$$=\nu\boldsymbol{c}^{(j)}+\sum_{i=1}^{p}\nu\mathcal{B}_i^{(j)}\boldsymbol{\sigma}_{t-i}+\sum_{k=1}^{\max(q,r)}(\nu\mathcal{A}_k^{(j)}+\mathcal{M}_k^{(j)})\boldsymbol{y}_{t-k},$$

$$(8)\qquad \text{Var}(\boldsymbol{y}_t|\mathcal{F}_{t-1})$$
$$=2\mathcal{D}_n^{+}[\nu(\boldsymbol{\Sigma}_t\otimes\boldsymbol{\Sigma}_t)+\boldsymbol{\Sigma}_t\otimes\boldsymbol{\Lambda}_t+\boldsymbol{\Lambda}_t\otimes\boldsymbol{\Sigma}_t](\mathcal{D}_n^{+})',\qquad l_{j-1}<r_{t-d}\le l_j,$$

with $\mathcal{A}_k^{(j)}=0$ for $k>q$ and $\mathcal{M}_k^{(j)}=0$ for $k>r$, where $\mathcal{D}_n^{+}=\mathcal{L}_n(\boldsymbol{I}_{n^2}+\mathcal{K}_{nn})/2$, $\boldsymbol{I}_{n^2}$ is an n^2-dimension identity matrix and $\mathcal{K}_{nn}$ is the commutation matrix such that for any $m\times n$ matrix X, $\text{vec}(X')=\mathcal{K}_{mn}\text{vec}(X)$, see also [29].

2.2 Parameter estimation

Recall the notation $\boldsymbol{c}^{(j)}=\text{vech}(\boldsymbol{C}^{(j)}(\boldsymbol{C}^{(j)})')$, $\boldsymbol{A}_k^{(j)}=\text{diag}(\alpha_{1k}^{(j)},...,\alpha_{nk}^{(j)})$ for $k=1,...,q$ and $\boldsymbol{B}_i^{(j)}=\text{diag}(\beta_{1i}^{(j)},...,\beta_{ni}^{(j)})$ for $i=1,...,p$. Let $\boldsymbol{\alpha}_k^{(j)}=(\alpha_{1k}^{(j)},...,\alpha_{nk}^{(j)})'$ and $\boldsymbol{\beta}_i^{(j)}=(\beta_{1i}^{(j)},...,\beta_{ni}^{(j)})'$ within the jth regime. For ease of presentation and without loss of generality, the case of $s=2$ is considered, thus there is only one threshold to be estimated. Then the parameter vector of GTCAW(p,q,r,s) turns out to be $\boldsymbol{\phi}=(\nu,(\boldsymbol{c}^{(j)})',(\boldsymbol{\alpha}_1^{(j)})',...,(\boldsymbol{\alpha}_q^{(j)})',(\boldsymbol{\beta}_1^{(j)})',...,(\boldsymbol{\beta}_p^{(j)})',\text{vec}(\boldsymbol{M}_1^{(j)})',...,\text{vec}(\boldsymbol{M}_r^{(j)})',l)'=(\boldsymbol{\theta},l)'$ ($j=1,2$), which can be estimated by maximum likelihood (ML) estimation. Denote $I_{1t}=I(r_{t-d}\le l)$ and $I_{2t}=I(r_{t-d}>l)$ where $I(\cdot)$ is the indicator function. Then by equation (2), one can find that the conditional log-likelihood function for the jth regime ($j=1,2$) is given by

$$(9)\qquad L_T(\boldsymbol{\phi})=\sum_{t=1}^{T}\Bigg\{-\frac{\nu}{2}\ln\det\boldsymbol{\Sigma}_t+\frac{\nu-n-1}{2}\ln\det\boldsymbol{Y}_t$$
$$-\frac{1}{2}\text{tr}\left[\boldsymbol{\Sigma}_t^{-1}(\boldsymbol{Y}_t+\boldsymbol{\Lambda}_t)\right]+\ln{}_0F_1\left(\frac{\nu}{2};\frac{1}{4}\boldsymbol{\Sigma}_t^{-1}\boldsymbol{\Lambda}_t\boldsymbol{\Sigma}_t^{-1}\boldsymbol{Y}_t\right)\Bigg\}I_{jt}$$
$$-T\times\left[\frac{\nu n}{2}\ln 2+\frac{n(n-1)}{4}\ln\pi+\sum_{i=1}^{n}\ln\Gamma\left(\frac{\nu+1-i}{2}\right)\right]I_{jt}$$

The ML-estimates $\widehat{\boldsymbol{\phi}}$ (except $\hat{l}$) can be obtained by maximizing the conditional log-likelihood function (9) numerically. Since $L_T(\boldsymbol{\phi})$ is not differentiable with respect to the threshold l, the maximization of the log-likelihood function can be done in the following two steps in practice:

1. For each l in the appropriate set of candidates $[l_L,l_U]$, find $\widehat{\boldsymbol{\theta}}_l$ such that
$$\widehat{\boldsymbol{\theta}}_l=\arg\max_{\boldsymbol{\theta}\in\Theta}(L_T(\boldsymbol{\theta},l)),$$
where Θ is the parameter space for $\boldsymbol{\theta}$.
2. The threshold is estimated by searching over all candidates
$$\hat{l}=\arg\max_{l\in[l_L,l_U]}(L_T(\widehat{\boldsymbol{\theta}}_l,l)),$$
and the final estimate for $\widehat{\boldsymbol{\phi}}$ is $(\widehat{\boldsymbol{\theta}}_{\hat{l}},\hat{l})'$.

Notice that to restrict $\nu>n-1$ and the main diagonal elements in $\boldsymbol{C}^{(j)}$ and the first diagonal element of each $\boldsymbol{A}_k^{(j)}$, $\boldsymbol{B}_i^{(j)}$ and $\boldsymbol{M}_h^{(j)}$ in the jth regime to be positive as mentioned before, we estimate $\sqrt{\nu-n+1}$, $\sqrt{C_{ll}^{(j)}}$, $\sqrt{A_{11,k}^{(j)}}$, $\sqrt{B_{11,i}^{(j)}}$ and $\sqrt{M_{11,h}^{(j)}}$ instead. As in the threshold time series literature, the estimator of l is expected to be super-consistent and consistency together with asymptotic normality hold for the other parameters in equations (3) and (4), see [9].

2.3 Model selection and diagnostic checking

Once the GTCAW(p,q,r,s) model has been estimated, we can use the Akaike information criterion (AIC) to choose the optimal order (p, q, r). The AIC of the GTCAW(p, q, r) model is given by

$$\text{AIC} = -2L_T(\widehat{\boldsymbol{\phi}}) + 2\dim(\boldsymbol{\phi}), \tag{10}$$

where $\widehat{\boldsymbol{\phi}}$ is the ML-estimates of $\boldsymbol{\phi}$, $L_T(\cdot)$ is the log-likelihood function given in equation (9) and $\dim(\boldsymbol{\phi}) = n^2(2r+1) + n(2p+2q+1)+2$ is the number of parameters.

For model diagnostic checking, we know that the GCAW(p, q, r) model (with the case $j = 1$) of [29] can be rewritten as the vector autoregressive moving average $(\max(p+r, q), p)$ model:

$$\begin{aligned}\boldsymbol{y}_t =& \nu \boldsymbol{c} + \sum_{l=1}^{\max(p+r,q)} \left(\nu \mathcal{A}_l + \mathcal{B}_l + \mathcal{M}_l - \sum_{h=1}^{l-1} \mathcal{B}_{l-h}\mathcal{M}_h\right) \boldsymbol{y}_{t-l} \\ &- \sum_{i=1}^{p} \mathcal{B}_i \boldsymbol{a}_{t-i} + \boldsymbol{a}_t,\end{aligned}$$

where $\boldsymbol{a}_t$ is a martingale difference subject to $E(\boldsymbol{a}_t) = 0$ and $E(\boldsymbol{a}_t\boldsymbol{a}_s) = 0$ for $s \neq t$. Following this idea, we define the standardized residual vector for our GTCAW(p, q, r, s) model as

$$\widehat{\boldsymbol{e}}_t = \text{Var}(\boldsymbol{y}_t|\mathcal{F}_{t-1})^{-1/2}[\boldsymbol{y}_t - E(\boldsymbol{y}_t|\mathcal{F}_{t-1})], \tag{11}$$

where the conditional expectation and conditional variance are given in equations (7) and (8) with $\boldsymbol{\phi} = \widehat{\boldsymbol{\phi}}$ respectively, and $\text{Var}(\boldsymbol{y}_t|\mathcal{F}_{t-1})^{-1/2}$ is the inverse of the Cholesky factor of matrix $\text{Var}(\boldsymbol{y}_t|\mathcal{F}_{t-1})$. As we all know, if the model fits the data correctly and captures the temporal dependence in elements of $\boldsymbol{y}_t$ adequately, the standardized residuals $\widehat{\boldsymbol{e}}_{ij,t}$ in the vector $\widehat{\boldsymbol{e}}_t$ should be approximately serially uncorrelated. Here, we choose the portmanteau test [6] and the multivariate portmanteau test [16] to check the adequacy of the fitted models. The corresponding portmanteau test statistics are defined as

$$\begin{aligned}Q_1 &= n^2 \sum_{k=1}^{m} \widehat{r}_{ij,k}^2/(n-k), \\ Q_2 &= n^2 \sum_{k=1}^{m} \text{tr}(\widehat{E}_k' \widehat{E}_0^{-1} \widehat{E}_k \widehat{E}_0^{-1})/(n-k),\end{aligned}$$

where $\widehat{r}_{ij,k} = \sum_{t=k+1}^{T} \widehat{\boldsymbol{e}}_{ij,t}\widehat{\boldsymbol{e}}_{ij,t-k} / \sum_{t=1}^{T} \widehat{\boldsymbol{e}}_{ij,t}^2$ is the kth autocorrelation of the residual component $\widehat{\boldsymbol{e}}_{ij,k}$, $\widehat{E}_k = \sum_{t=k+1}^{T} \widehat{\boldsymbol{e}}_t \widehat{\boldsymbol{e}}_{t-k}'$ is the kth residual autocovariance matrix and m is the number of lags to be chosen appropriately. The construction for these two test statistics are similar but different for the DF of the χ^2 distribution.

2.4 Forecasting

Given the parameter estimates, one can easily obtain the forecasts of $\boldsymbol{y}_t$. Here, we predict the one-day-ahead covariance matrix by the conditional expectation given past observations of $\boldsymbol{y}_t$. According to equation (7), it can be shown that

$$\begin{aligned}\widehat{\boldsymbol{y}}_t(1) &= E(\boldsymbol{y}_{t+1}|\mathcal{F}_t) \\ &= \nu \boldsymbol{c}^{(j)} + \sum_{i=1}^{p} \nu \mathcal{B}_i^{(j)} \boldsymbol{\sigma}_t + \sum_{k=1}^{\max(q,r)} (\nu \mathcal{A}_k^{(j)} + \mathcal{M}_k^{(j)}) \boldsymbol{y}_t, \\ &\qquad l_{j-1} < r_{t-d} \leq l_j.\end{aligned}$$

By calculating $\widehat{\boldsymbol{y}}_t(1)$, then we have the one-day-ahead forecasts of the RCOV matrices $\widehat{\boldsymbol{Y}}_t(1)$, which are obtained by substituting the elements of $\widehat{\boldsymbol{y}}_t(1)$ into a symmetric matrix.

3. SIMULATION STUDIES

For simulating the GTCAW model, first we assume that $n = 2$, $s = 2$ for simplicity, and the threshold structure of the autoregressive and conditional variance equations are the same for ease of presentation. Extension to other complex cases can be done in a straightforward manner. By the virtue of the SETAR model introduced by [25], we use the first diagonal element of the RCOV matrix series with delay parameter equal to one, i.e., $Y_{11,t-1}$ as the threshold variable.

Similar to the generation of the noncentral Wishart distribution of [29], the simulation procedure with threshold framework is as follows:

1. Give the initial values, i.e., $\boldsymbol{Y}_0 = \boldsymbol{\Sigma}_0 = \boldsymbol{I}_2$ and the threshold value l, determine the regime using the relationship between $Y_{11,t-1}$ and l, calculate $\boldsymbol{\Lambda}_t$ and $\boldsymbol{\Sigma}_t$ by equations (3) and (4).
2. Generate a $\nu \times n$ random matrix $\boldsymbol{Z}$ such that its elements are uncorrelated standard normal random numbers.
3. Obtain the Cholesky decompositions of $\boldsymbol{\Lambda}_t$ and $\boldsymbol{\Sigma}_t$ such that $\boldsymbol{\Lambda}_t = \widetilde{\boldsymbol{N}}_t' \widetilde{\boldsymbol{N}}_t$ and $\boldsymbol{\Sigma}_t = \boldsymbol{D}_t' \boldsymbol{D}_t$.
4. Construct the $\nu \times n$ mean matrix as
$$\boldsymbol{N}_t = \begin{pmatrix} \widetilde{\boldsymbol{N}}_t \\ \boldsymbol{O}_{\nu-n,n} \end{pmatrix},$$
where $\boldsymbol{O}_{\nu-n,n}$ is the $(\nu - n) \times n$ zero matrix.
5. Generate the RCOV matrix as
$$\boldsymbol{Y}_t = (\boldsymbol{N}_t + \boldsymbol{Z}\boldsymbol{D}_t)'(\boldsymbol{N}_t + \boldsymbol{Z}\boldsymbol{D}_t).$$

After constructing the simulated RCOV matrices, we apply the two steps estimation procedure as discussed in Section 2.2. Since the threshold is estimated by the grid search method over the set of candidates in $[l_L, l_U]$, the choice for the upper bound l_U as well as the lower bound l_L is necessary and of vital importance. A common strategy used in practice is to replace the bounds by some numbers determined based on the data (see [11], [28]). Specifically, we fix $\alpha_1 = 0.2$, $\alpha_2 = 0.8$ and find the empirical α_ith quantile for the sorted data $Y_{11,t-1}$, denoted as $q_i (i = 1, 2)$. Then the threshold is estimated by searching over all values of $Y_{11,t-1}$

Table 1. True values of the parameters used in simulation studies

Para.	Value		Para.	Value	
Regime	1			2	
$\boldsymbol{A}_1^{(1)}$	0.2069		$\boldsymbol{A}_1^{(2)}$	0.0621	
		−0.0567			−0.1606
$\boldsymbol{B}_1^{(1)}$	0.2461		$\boldsymbol{B}_1^{(2)}$	0.2077	
		−0.0145			−0.0926
$\boldsymbol{C}^{(1)}$	0.1060		$\boldsymbol{C}^{(2)}$	0.1065	
	0.2374	0.2978		−0.1294	0.1874
$\boldsymbol{M}_1^{(1)}$	0.1232	−0.2918	$\boldsymbol{M}_1^{(2)}$	0.2361	0.0565
	−0.2352	0.0378		−0.0880	−0.0140
ν	10		l	0.3	

Table 2. Bias and MSE for GTCAW(1,1,1,2) model as sample size $T = 200$

Para.	Bias		Para.	Bias	
Regime	1			2	
$\boldsymbol{A}_1^{(1)}$	−0.0082		$\boldsymbol{A}_1^{(2)}$	−0.0076	
		0.0305			0.0246
$\boldsymbol{B}_1^{(1)}$	−0.0127		$\boldsymbol{B}_1^{(2)}$	−0.0266	
		0.0004			0.0440
$\boldsymbol{C}^{(1)}$	−0.0057		$\boldsymbol{C}^{(2)}$	−0.0108	
	0.0055	−0.0495		−0.0018	−0.1050
$\boldsymbol{M}_1^{(1)}$	−0.0008	0.0063	$\boldsymbol{M}_1^{(2)}$	−0.0140	−0.0134
	0.1374	−0.0024		−0.0196	−0.0159
ν	0.2338		l	−0.0021	
Para.	MSE		Para.	MSE	
Regime	1			2	
$\boldsymbol{A}_1^{(1)}$	0.0013		$\boldsymbol{A}_1^{(2)}$	0.0015	
		0.0045			0.0028
$\boldsymbol{B}_1^{(1)}$	0.0567		$\boldsymbol{B}_1^{(2)}$	0.0345	
		0.0592			0.0948
$\boldsymbol{C}^{(1)}$	0.0003		$\boldsymbol{C}^{(2)}$	0.0004	
	0.0027	0.0098		0.0014	0.0181
$\boldsymbol{M}_1^{(1)}$	0.0069	0.0013	$\boldsymbol{M}_1^{(2)}$	0.0192	0.0041
	0.2196	0.0108		0.1077	0.0701
ν	0.3900		l	1.1344×10^{-5}	

in the interval $[q_1, q_2]$ with estimates having the maximum log-likelihood value. We obtain the parameter estimate $\widehat{\boldsymbol{\theta}}_l$ for each given threshold l, finally select the threshold $\hat{l}$ and the corresponding $\widehat{\boldsymbol{\phi}}_{\hat{l}}$ which maximizes the log-likelihood function.

Here, we conduct a simulation experiment on the GTCAW(1,1,1,2) model with $m = 100$ replications of 2 by 2 RCOV matrices and we choose the sample size $T = 200,\ 400$. With the known threshold variable, there are 24 parameters to be estimated including the DF ν and the threshold l. Each estimate is compared in terms of its bias and mean squared error (MSE) according to the following formulas:

$$\text{bias} = \frac{1}{m}\sum_{i=1}^{m}(\widehat{\vartheta}_i - \vartheta^0), \quad \text{MSE} = \frac{1}{m}\sum_{i=1}^{m}(\widehat{\vartheta}_i - \vartheta^0)^2,$$

Table 3. Bias and MSE for GTCAW(1,1,1,2) model as sample size $T = 400$

Para.	Bias		Para.	Bias	
Regime	1			2	
$\boldsymbol{A}_1^{(1)}$	−0.0029		$\boldsymbol{A}_1^{(2)}$	−0.0049	
		0.0240			0.0177
$\boldsymbol{B}_1^{(1)}$	0.0003		$\boldsymbol{B}_1^{(2)}$	−0.0231	
		−0.0074			0.0171
$\boldsymbol{C}^{(1)}$	−0.0046		$\boldsymbol{C}^{(2)}$	−0.0067	
	0.0123	−0.0342		−0.0052	−0.0740
$\boldsymbol{M}_1^{(1)}$	0.0013	0.0061	$\boldsymbol{M}_1^{(2)}$	−0.0330	−0.0016
	0.1180	0.0140		0.0009	−0.0382
ν	0.1247		l	$-9.3948 * 10^{-4}$	
Para.	MSE		Para.	MSE	
Regime	1			2	
$\boldsymbol{A}_1^{(1)}$	0.0005		$\boldsymbol{A}_1^{(2)}$	0.0010	
		0.0023			0.0013
$\boldsymbol{B}_1^{(1)}$	0.0413		$\boldsymbol{B}_1^{(2)}$	0.0225	
		0.0360			0.0543
$\boldsymbol{C}^{(1)}$	0.0002		$\boldsymbol{C}^{(2)}$	0.0002	
	0.0017	0.0040		0.0013	0.0083
$\boldsymbol{M}_1^{(1)}$	0.0044	0.0006	$\boldsymbol{M}_1^{(2)}$	0.0157	0.0032
	0.0916	0.0050		0.0923	0.0388
ν	0.1529		l	$1.8081 * 10^{-6}$	

where $\widehat{\vartheta}_i$ is the estimate of the true parameter ϑ^0 in the ith replication. The true values of the parameters are shown in Table 1 and the bias and MSE values of the estimates for $T = 200$ and 400 are presented in Tables 2 and 3, respectively. From them, we can find that the bias and MSE of the estimate in matrices both give reasonably small values. Furthermore, as the sample size increases, all the values of MSE gradually decrease. In particular, the mean estimated threshold is 0.2979 (0.2991) when $T = 200$ (400), very close to the true value 0.3. In addition, the MSE for the threshold parameter drops much more than 50% (from 1.1344×10^{-5} to 1.8081×10^{-6}), which indicates the super-consistency of the threshold estimator. And the obtained value of DF shows a little larger bias and MSE, it may be due to the fact that the convergence of DF in Wishart distribution is relatively slower. In general, all the estimates are fairly close to the true values.

4. REAL DATA ANALYSIS

We now apply the proposed GTCAW models to three examples of daily RCOV matrices for modeling stocks traded at the New York Stock Exchange (NYSE) in this section. It is well known that the stocks have very different dynamics in a bull or bear market. For example, the correlations between stock returns will become stronger in the down market. Therefore, we consider the piecewise linear approach, threshold, to feasibly and easily capture such phenomenon.

Table 4. Summary statistics for the realized variances and covariances

Stock	Mean $\times 10^{-5}$	Minimum $\times 10^{-5}$	Maximum $\times 10^{-4}$	SD $\times 10^{-5}$	Skewness	Kurtosis
Example (a)			Realized variance			
AIG	17.40	3.05	9.16	11.17	2.47	13.08
GE	5.89	0.91	1.57	3.02	1.01	3.59
			Realized covariance			
AIG-GE	3.86	−1.53	1.51	2.82	1.10	4.35
Example (b)			Realized variance			
BA	5.46	1.10	2.19	3.25	1.70	7.33
C	16.73	2.80	9.62	10.89	2.56	15.25
			Realized covariance			
BA-C	3.69	−1.30	1.75	3.13	1.66	6.30
Example (c)			Realized variance			
KO	2.69	0.42	1.10	1.73	1.81	7.22
MCD	3.08	0.42	2.68	2.57	4.34	33.46
WMT	3.23	0.48	1.61	2.34	2.53	11.61
			Realized covariance			
KO-MCD	0.94	−0.36	0.91	0.96	3.21	23.80
KO-WMT	0.89	−0.11	0.47	0.85	1.45	5.63
MCD-WMT	0.82	−0.69	0.44	0.80	1.40	5.50

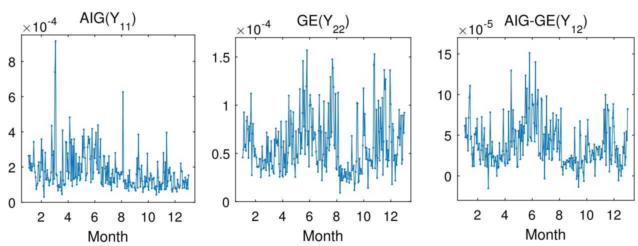

Figure 1. Time series of the realized variances and covariances for example (a).

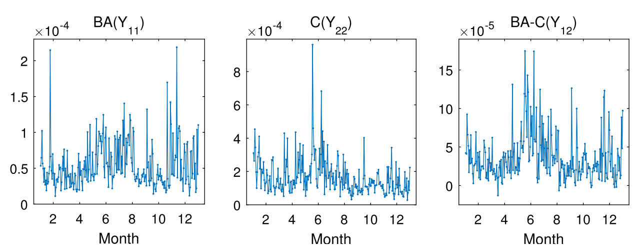

Figure 2. Time series of the realized variances and covariances for example (b).

4.1 Data

Here, we select seven stocks (after data cleaning) traded at NYSE: American International Group (AIG), General Electric (GE), Boeing (BA), Citigroup (C), Coca-Cola (KO), McDonald's (MCD) and Wal-Mart (WMT), starting at 3 January 2012 and ending on 31 December 2012, with totally 250 observations. The original data set was also considered to deal with high-dimensional RCOV matrices

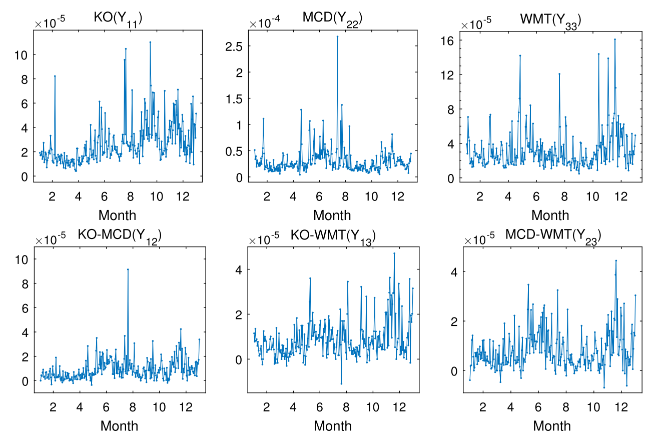

Figure 3. Time series of the realized variances and covariances for example (c).

by using 30 stocks in [23]. The above seven stocks are divided into three data set examples, that is (a) AIG and GE, (b) BA and C and (c) KO, MCD and WMT. For example (a), we just choose them by a random choice and they are both multi-national corporations that include financial service. The other two combinations of stocks have relative high correlations. BA and C are the components of the First National City Bank Financial Group. As for example (c), we pick up these three stocks not only because their brand familiarity, but also due to the fact that they are all from consumer goods industry and have cooperation mutually.

Then, we construct the daily RCOV matrices by using the threshold multiscale realized volatility matrix (MRVM)

Table 5. Goodness of fit measures for the GTCAW models of example (a)

Order	$\dim(\boldsymbol{\phi})$	Threshold	$L_T(\widehat{\boldsymbol{\phi}})$	AIC	p-value			
					$\widehat{e}_{11,k}$	$\widehat{e}_{12,k}$	$\widehat{e}_{22,k}$	$\widehat{E}_k$
Central Wishart distribution with threshold (GTCAW(0,0,0,s))								
(0,2)	8	1.5842	−238.0044	492.0089	0.06	0.00	0.00	0.00
TWAR(r,s) (GTCAW(0,0,r,s))								
(1,2)	16	1.5842	−228.4030	488.8060	0.07	0.00	0.00	0.00
(2,2)	24	1.5842	−212.6215	473.2430	0.13	0.01	0.04	0.01
TCAW(p,q,s) (GTCAW(p,q,0,s))								
(0,1,2)	12	1.3044	−210.7074	445.4148	0.23	0.01	0.03	0.28
(1,1,2)	16	1.5842	−198.7482	429.4964	**0.32**	**0.21**	**0.08**	**0.44**
(1,2,2)	20	1.5842	−197.6880	435.3760	**0.22**	**0.15**	**0.07**	**0.31**
(2,1,2)	20	1.5106	−187.1481	<u>408.2961</u>	**0.47**	**0.11**	**0.08**	**0.35**
(2,2,2)	24	1.5106	−185.0764	418.1528	**0.33**	**0.15**	**0.14**	**0.49**
(2,3,2)	28	1.5106	−185.0042	426.0084	**0.27**	**0.12**	**0.10**	**0.30**
(3,2,2)	28	1.5106	−184.1077	424.2155	**0.39**	**0.11**	**0.08**	**0.11**
(3,3,2)	32	1.5106	−184.0850	432.1700	**0.22**	**0.09**	**0.07**	**0.14**
GTCAW(p,q,r,s)								
(0,1,1,2)	20	1.8627	−207.3659	454.7317	0.18	0.00	0.01	0.10
(1,1,1,2)	24	1.5787	−197.1957	442.3915	**0.29**	**0.14**	**0.05**	**0.38**
(1,2,1,2)	28	1.5842	−195.9558	447.9117	**0.23**	**0.13**	**0.06**	**0.38**
(2,1,1,2)	28	1.5106	−186.3690	<u>418.7381</u>	**0.25**	**0.04**	**0.04**	**0.26**
(2,2,1,2)	32	1.5106	−184.5951	433.1902	**0.29**	**0.08**	**0.07**	**0.28**
(2,3,1,2)	36	1.5106	−184.2820	440.9639	**0.28**	**0.09**	**0.07**	**0.30**
(3,2,1,2)	36	1.5106	−183.7655	447.5310	**0.18**	**0.05**	**0.03**	**0.05**
(3,3,1,2)	40	1.5106	−183.9569	439.9139	**0.18**	**0.05**	**0.03**	**0.05**
(0,1,2,2)	28	1.8627	−199.4685	454.9370	0.19	0.01	0.05	0.13
(1,1,2,2)	32	1.5842	−195.0493	454.0985	**0.15**	**0.10**	**0.03**	**0.24**
(1,2,2,2)	36	1.5842	−194.4081	460.8162	**0.13**	**0.09**	**0.03**	**0.24**
(2,1,2,2)	36	1.5106	−181.6999	435.3998	**0.13**	**0.05**	**0.04**	**0.18**
(2,2,2,2)	40	1.5106	−180.3721	440.7443	**0.13**	**0.07**	**0.04**	**0.19**
(2,3,2,2)	44	1.5106	−180.2026	448.4052	**0.14**	**0.07**	**0.04**	**0.19**
(3,2,2,2)	44	1.5106	−179.8962	447.7924	**0.07**	**0.03**	**0.02**	**0.02**
(3,3,2,2)	48	1.5106	−179.7972	455.5943	**0.08**	**0.03**	**0.02**	**0.02**

Note: $\dim(\boldsymbol{\phi})$, the number of parameters; $L_T(\hat{\boldsymbol{\phi}})$, the optimized log-likelihood; AIC, Akaike information criterion; p-value, calculated by the portmanteau test using 20 lags of the standardized residuals.

estimator defined by [24]. Similar to [23], it generates three set-ups of time series with 250 matrices, whose sizes are 2 by 2, 2 by 2 and 3 by 3, respectively. Since the threshold MRVM estimator cannot guarantee positive definite property in finite samples, we have also checked that all the matrices in the data series are positive definite.

In addition, descriptive statistics for the realized variances and covariances of each example are presented in Table 4, and their time plots are shown in Figures 1-3, respectively. From the graphs, the realized variances and covariances show fluctuations and some time series have significant outliers during the year, which may lead to heavy-tailed cases. In fact, one can find that all realized variances and covariances have larger kurtosises than that of the normal distribution, exhibiting the heavy-tailed phenomena.

4.2 Model fitting

We now fit the GTCAW model to the full sample of daily RCOV matrices, with lag orders (p,q,r) ranging from (0,0,0) to (3,3,2). For simplicity, we also choose the regime $s=2$ and $Y_{11,t-1}$ as the threshold variable. Clearly, many other possibilities for the threshold variable exist, but our purpose here is to illustrate the usefulness of the GTCAW models in describing the nonlinear presence in the data. The current choice seems to be adequate for our purpose. Notice that the GTCAW(0,0,0,2) model represents the central Wishart distribution $W_n(\nu, 0, \boldsymbol{CC}')$ with a threshold. When $p=q=0$ and $r \neq 0$, it turns out to be the TWAR(r,2) model. If $r=0$, it reduces to the TCAW$(p,q,2)$ model as mentioned before. In each example, we report the estimated threshold value, the maximum values of the log-likelihood function given in (9), AIC values given in (10) and the results of diagnostic checking on the standardized residuals $\widehat{\boldsymbol{e}}_t$ with 20 lags as described in Section 2.3, which are shown in Tables 5-7, respectively. For comparison purposes, we also fit non-threshold GCAW model with the optimal order of each example to the RCOV matrices. The results can be seen in Table 8.

Table 6. Goodness of fit measures for the GTCAW models of example (b)

Order	dim(ϕ)	Threshold	$L_T(\widehat{\phi})$	AIC	p-value			
					$\widehat{e}_{11,k}$	$\widehat{e}_{12,k}$	$\widehat{e}_{22,k}$	$\widehat{E}_k$
Central Wishart distribution with threshold (GTCAW(0,0,0,s))								
(0,2)	8	0.4230	−229.0383	474.0766	0.03	0.00	0.00	0.00
TWAR(r,s) (GTCAW(0,0,r,s))								
(1,2)	16	0.4080	−206.9145	445.8291	0.02	0.03	0.00	0.00
(2,2)	24	0.4160	−195.7852	439.3703	0.06	0.00	0.02	0.06
TCAW(p,q,s) (GTCAW(p,q,0,s))								
(0,1,2)	12	0.4080	−195.0041	414.0082	0.04	0.00	0.04	0.00
(1,1,2)	16	0.3590	−183.4360	<u>394.8721</u>	**0.97**	**0.90**	**0.91**	**0.42**
(1,2,2)	20	0.4060	−178.9340	397.8681	0.05	0.01	0.22	0.01
(2,1,2)	20	0.4080	−179.8978	399.7957	**0.51**	**0.47**	**0.82**	**0.15**
(2,2,2)	24	0.4080	−175.2716	398.5433	0.08	0.02	0.20	0.01
(2,3,2)	28	0.4080	−174.2561	404.5122	0.05	0.01	0.15	0.00
(3,2,2)	28	0.4080	−173.2772	402.5545	0.04	0.01	0.12	0.01
(3,3,2)	32	0.4080	−172.3572	432.1700	0.04	0.01	0.12	0.00
GTCAW(p,q,r,s)								
(0,1,1,2)	20	0.4080	−193.0410	426.0821	0.03	0.00	0.04	0.04
(1,1,1,2)	24	0.3600	−180.7873	<u>405.5745</u>	**0.94**	**0.83**	**0.87**	**0.46**
(1,2,1,2)	28	0.4080	−177.1697	410.3394	0.04	0.01	0.18	0.02
(2,1,1,2)	28	0.4080	−178.1396	412.2792	**0.47**	**0.42**	**0.81**	**0.06**
(2,2,1,2)	32	0.4080	−173.6099	411.2198	0.04	0.01	0.19	0.00
(2,3,1,2)	36	0.4080	−172.9911	417.9821	0.03	0.01	0.17	0.00
(3,2,1,2)	36	0.4080	−171.6847	415.3693	0.02	0.01	0.11	0.00
(3,3,1,2)	40	0.4080	−170.8503	421.7006	0.02	0.00	0.11	0.05
(0,1,2,2)	28	0.4080	−180.9398	417.8795	0.04	0.01	0.37	0.06
(1,1,2,2)	32	0.4080	−178.4936	420.9872	0.02	0.00	0.18	0.00
(1,2,2,2)	36	0.4080	−176.1822	424.3644	0.02	0.00	0.17	0.00
(2,1,2,2)	36	0.4080	−173.8349	419.6698	0.01	0.00	0.18	0.00
(2,2,2,2)	40	0.4080	−171.6268	423.2536	0.02	0.01	0.17	0.00
(2,3,2,2)	44	0.4080	−170.9751	429.9503	0.02	0.01	0.16	0.00
(3,2,2,2)	44	0.4080	−169.6830	427.3659	0.01	0.00	0.10	0.00
(3,3,2,2)	48	0.4080	−168.8594	433.7188	0.01	0.00	0.09	0.00

Note: dim(ϕ), the number of parameters; $L_T(\hat{\phi})$, the optimized log-likelihood; AIC, Akaike information criterion; p-value, calculated by the portmanteau test using 20 lags of the standardized residuals.

From them, we have the following findings:

1. First, we consider the diagnostic checking for each model and all elements of the residual series passing the portmanteau tests at the 1% significance level are in bold. Notice that in example (c), due to the increase in the dimension of the RCOV matrices, we only consider the portmanteau test of [6]. In fact, as we have been looking at many tests at the same significance level, some of the cases that do not pass the diagnostic checking are probably due to the type one error. Then based on AIC values, we choose the best specification among the models that pass the diagnostic checking in each example. From Tables 5-7, one can find that the best fitting models are TCAW(2,1,2), TCAW(1,1,2) and GTCAW(3,2,1,2) models in examples (a)-(c), respectively, since they have the smallest AIC values. Besides, the best performance in GTCAW(p,q,r,s) cases for examples (a)-(b) are GTCAW(2,1,1,2) and GTCAW(1,1,1,2), respectively.
2. With the optimal order, we conduct the corresponding non-threshold GCAW models. From Table 8, one can find that many elements do not pass the diagnostic checking. In addition, despite the GCAW models have less parameters, they show much larger AIC values than their threshold counterparts. This leads to a demonstration on the nonlinear dynamic presence in the RCOV series and the usefulness of the threshold approach.
3. The ML-estimates of the preferred models for each example are given in Tables 9-11. One can find that some estimates show very low values and we also conduct a likelihood ratio test (LRT) for the significance of those estimates. More specifically, let the null hypothesis be $\vartheta_i = 0$, and we refit the model by setting those parameters with small estimates to zeros and then apply a LRT to the reduced model. In summary, the preferred models all accept the null hypothesis and we fix the corresponding quantities in Tables 9-11. Although these

Table 7. Goodness of fit measures for the GTCAW models of example (c)

Order	$\dim(\boldsymbol{\phi})$	Threshold	$L_T(\widehat{\boldsymbol{\phi}})$	AIC	p-value					
					$\widehat{\boldsymbol{e}}_{11,k}$	$\widehat{\boldsymbol{e}}_{12,k}$	$\widehat{\boldsymbol{e}}_{13,k}$	$\widehat{\boldsymbol{e}}_{22,k}$	$\widehat{\boldsymbol{e}}_{23,k}$	$\widehat{\boldsymbol{e}}_{33,k}$
Central Wishart distribution with threshold (GTCAW(0,0,0,s))										
(0,2)	14	0.1910	1481.6959	−2935.3918	0.05	0.11	0.00	0.00	0.00	0.00
TWAR(r,s) (GTCAW(0,0,r,s))										
(1,2)	32	0.1910	1548.4026	−3032.8051	0.16	0.19	0.04	0.02	0.00	0.00
(2,2)	50	0.1930	1564.6422	−3029.2846	0.42	0.25	0.08	0.02	0.00	0.00
TCAW(p,q,s) (GTCAW(p,q,0,s))										
(0,1,2)	20	0.1730	1580.4893	−3120.9785	0.08	0.00	0.00	0.00	0.00	0.20
(1,1,2)	26	0.1910	1605.3007	−3158.6015	0.24	0.01	0.00	1.00	0.00	0.21
(1,2,2)	32	0.1910	1602.9856	−3141.9712	0.22	0.08	0.00	1.00	0.74	0.05
(2,1,2)	32	0.1910	1614.3620	−3164.7240	0.19	0.01	0.00	1.00	0.51	0.15
(2,2,2)	38	0.1910	1600.3317	−3124.6633	0.27	0.08	0.00	1.00	0.18	0.04
(2,3,2)	44	0.1910	1612.9541	−3137.9083	0.09	0.13	0.00	1.00	0.01	0.03
(3,2,2)	44	0.2020	1606.4562	<u>−3130.9123</u>	**1.00**	**1.00**	**0.74**	**1.00**	**0.97**	**0.12**
(3,3,2)	50	0.1820	1615.3461	−3130.6921	0.03	0.07	0.00	0.02	0.00	0.02
GTCAW(p,q,r,s)										
(0,1,1,2)	38	0.1580	1596.1638	−3116.3277	0.09	0.00	0.04	0.00	0.00	0.21
(1,1,1,2)	44	0.1910	1620.5998	−3153.1997	0.21	0.02	0.00	1.00	1.00	0.10
(1,2,1,2)	50	0.1910	1626.2514	−3152.5027	0.44	0.01	0.01	1.00	0.00	0.09
(2,1,1,2)	50	0.1910	1620.7963	−3141.5926	0.13	0.01	0.00	1.00	0.99	0.06
(2,2,1,2)	56	0.1910	1626.5640	−3141.1279	0.31	0.01	0.00	1.00	0.00	0.05
(2,3,1,2)	62	0.1910	1629.2613	−3134.5226	0.62	0.02	0.00	1.00	0.00	0.03
(3,2,1,2)	62	0.2040	1626.6942	<u>−3139.3885</u>	**0.98**	**0.99**	**0.99**	**1.00**	**1.00**	**1.00**
(3,3,1,2)	68	0.1910	1632.8585	−3129.7171	0.53	0.02	0.00	1.00	0.11	0.02
(0,1,2,2)	56	0.1580	1613.3820	−3114.7641	0.32	0.01	0.02	0.00	0.00	0.18
(1,1,2,2)	62	0.1890	1626.1696	−3128.3391	0.35	0.06	0.01	0.00	0.00	0.06
(1,2,2,2)	68	0.1910	1633.1398	−3130.2796	0.34	0.02	0.01	1.00	0.00	0.04
(2,1,2,2)	68	0.1910	1634.2738	−3132.5476	0.29	0.02	0.00	1.00	0.00	0.03
(2,2,2,2)	74	0.1910	1633.2540	−3118.5079	0.22	0.01	0.01	1.00	0.00	0.02
(2,3,2,2)	80	0.1910	1637.8375	−3115.6750	0.55	0.02	0.00	1.00	0.00	0.01
(3,2,2,2)	80	0.1950	1629.7196	−3099.4392	**0.94**	**0.99**	**0.99**	**1.00**	**1.00**	**1.00**
(3,3,2,2)	86	0.1890	1640.2279	−3108.4558	0.11	0.04	0.00	0.00	0.00	0.00

Note: $\dim(\boldsymbol{\phi})$, the number of parameters; $L_T(\hat{\boldsymbol{\phi}})$, the optimized log-likelihood; AIC, Akaike information criterion; p-value, calculated by the portmanteau test using 20 lags of the standardized residuals.

Table 8. Model fitting results for the non-threshold GCAW models with the preferred orders

Example	(p,q,r)	$\dim(\boldsymbol{\phi})$	$L_T(\widehat{\boldsymbol{\phi}})$	AIC	$\widehat{\boldsymbol{e}}_{11,k}$	$\widehat{\boldsymbol{e}}_{12,k}$	$\widehat{\boldsymbol{e}}_{22,k}$	$\widehat{\boldsymbol{e}}_{13,k}$	$\widehat{\boldsymbol{e}}_{23,k}$	$\widehat{\boldsymbol{e}}_{33,k}$
(a)	(2,1,0)	10	−241.5557	503.1114	0.05	0.00	0.00	-	-	-
(b)	(1,1,0)	8	−212.0289	440.0579	0.00	0.00	0.02	-	-	-
(c)	(3,2,1)	31	1575.8346	−3089.6691	0.05	0.00	0.00	0.00	0.00	0.03

estimates are insignificant, the lag order (p,q,r) does not change and therefore the preferred models are still efficient.

4. By simply collecting the data series in each example, we count the numbers of $\boldsymbol{Y}_{11,t}$ in both regimes, that is (a) 135/115, (b) 80/170 and (c) 113/137, upper number refers to regime 1. This indicates that there are enough observations in every regime, thus, the threshold approach is feasible and meaningful.
5. The TWAR model is the worst in terms of both the diagnostic checking results and the AIC values among all models. One can see that all residual series of the TWAR models fail the multivariate portmanteau test. This indicates that the constant scale matrix cannot capture the dependence enough in RCOV matrices.
6. For the efficient models which pass the diagnostic checking in examples (a)-(b), the TCAW$(p,q,2)$ models yield the smaller AIC values than the corresponding GTCAW$(p,q,r,2)$ cases. The reason is mainly because the TCAW$(p,q,2)$ model has less parameters than its GTCAW counterpart. However, in the above three examples, we can see that for any fixed lag order (p,q), the

Table 9. ML-estimates for TCAW(2,1,2) and GTCAW(2,1,1,2) models of example (a)

Para.	Estimate		Para.	Estimate	
TCAW(2,1,2)(GTCAW(2,1,0,2))					
Regime	1			2	
$A_1^{(1)}$	0.0000		$A_1^{(2)}$	0.1531	
		−0.1577			0.1584
$B_1^{(1)}$	0.0293		$B_1^{(2)}$	0.0064	
		0.2410			0.0281
$B_2^{(1)}$	0.8928		$B_2^{(2)}$	0.0000	
		0.7140			0.0000
$C^{(1)}$	0.0966		$C^{(2)}$	0.3318	
	0.0657	0.0324		0.0729	0.1783
ν	11.8341		l	1.5106×10^{-4}	
GTCAW(2,1,1,2)					
Regime	1			2	
$A_1^{(1)}$	0.0000		$A_1^{(2)}$	0.1509	
		−0.1440			0.1451
$B_1^{(1)}$	0.1653		$B_1^{(2)}$	0.0786	
		0.2323			0.0261
$B_2^{(1)}$	0.9023		$B_2^{(2)}$	0.0001	
		0.7430			0.0000
$C^{(1)}$	0.0872		$C^{(2)}$	0.3258	
	0.0569	0.0301		0.0638	0.1778
$M_1^{(1)}$	0.0000	−0.2381	$M_1^{(2)}$	0.0000	−0.3296
	−0.0276	−0.2953		−0.0067	−0.2646
ν	11.8304		l	1.5106×10^{-4}	

Table 10. ML-estimates for TCAW(1,1,2) and GTCAW(1,1,1,2) models of example (b)

Para.	Estimate		Para.	Estimate	
TCAW(1,1,2)(GTCAW(1,1,0,2))					
Regime	1			2	
$A_1^{(1)}$	0.0000		$A_1^{(2)}$	0.1396	
		−0.1945			0.1713
$B_1^{(1)}$	0.0000		$B_1^{(2)}$	0.5961	
		−0.6285			0.6298
$C^{(1)}$	0.2032		$C^{(2)}$	0.1611	
	0.1284	0.1403		0.0918	0.2012
ν	10.2182		l	0.3590×10^{-4}	
GTCAW(1,1,1,2)					
Regime	1			2	
$A_1^{(1)}$	0.0000		$A_1^{(2)}$	0.1387	
		−0.1984			0.1675
$B_1^{(1)}$	0.0000		$B_1^{(2)}$	0.6189	
		−0.6137			0.6483
$C^{(1)}$	0.1872		$C^{(2)}$	0.1571	
	0.1355	0.1199		0.0833	0.1899
$M_1^{(1)}$	0.3633	−0.2213	$M_1^{(2)}$	0.0529	0.0140
	−0.3833	−0.1735		0.2882	0.0224
ν	10.2052		l	0.3600×10^{-4}	

GTCAW$(p, q, r, 2)$ models always have larger maximum log-likelihood values than the TCAW$(p, q, 2)$ models.

7. From the economic perspective, it seems very appealing to model financial dynamics in terms of a threshold model in which the expanding phase and contracting phase are governed by different regimes. With the above results, we conclude that our GTCAW models can fit RCOV matrices more efficiently.

4.3 Forecasting results

To further examine the performance of the GTCAW model, we also perform an out-of-sample forecasting exercise. Here, we compute the one-day-ahead forecast of $\boldsymbol{Y}_t$ in the preferred GTCAW models and their corresponding non-threshold GCAW counterparts for comparison.

The forecasting exercise is carried out along the following lines: we choose the moving window sample consisting of k days and forecast the next day, where $k = 235, ..., 249$. Each model is re-estimated and the one-day-ahead forecasts are computed based on the updated parameter estimates. Then, we compare the forecast errors during the 15 periods between the forecasts and the ex-post realization of the RCOV matrices $\boldsymbol{Y}_{t+1}$. The predictive accuracy is measured by the spectral norm (SN) and the Frobenius norm (FN) according to the following formulas:

$$\begin{aligned} SN &= \frac{1}{T_0}\sum_t ||\boldsymbol{Y}_{t+1} - \widehat{\boldsymbol{Y}}_t(1)||_2 \\ &= \frac{1}{T_0}\sum_t \left\{\lambda_{max}[(\boldsymbol{Y}_{t+1} - \widehat{\boldsymbol{Y}}_t(1))^H(\boldsymbol{Y}_{t+1} - \widehat{\boldsymbol{Y}}_t(1))]\right\}^{1/2}, \\ FN &= \frac{1}{T_0}\sum_t ||\boldsymbol{Y}_{t+1} - \widehat{\boldsymbol{Y}}_t(1)||_F \\ &= \frac{1}{T_0}\sum_t \left\{\sum_{i,j}\left[\boldsymbol{Y}_{ij,t+1} - \widehat{\boldsymbol{Y}}_{ij,t}(1)\right]^2\right\}^{1/2}, \end{aligned}$$

where T_0 is the number of forecast periods and equals to 15, λ_{max} represents the maximum eigenvalue of the matrix and H denotes the conjugate transpose for a matrix.

Table 12 reports the results of forecasting accuracy for the preferred GTCAW models of each example together with their non-threshold cases. In every example, one can see that the GTCAW model yields smaller predictive errors than its non-threshold counterpart. This indicates the necessity of including threshold in modeling the RCOV matrices. Besides, we also consider using the previous RCOV estimator as the predictor based on martingale theory for comparison. By calculating the above predictive accuracy SN and FN respectively for each example, we have the results ($\times 10^{-4}$), (a) 0.6233 and 0.6420, (b) 0.7985 and 0.8170 together with (c) 0.3531 and 0.3707. By contrast, we know that our GTCAW models can improve the forecasting accuracy.

5. CONCLUSION

The current paper extends the GCAW model [29] to its threshold framework for investigating the asymmetric nonlinear behavior of the RCOV matrices of asset returns. The

Table 11. ML-estimates for TCAW(3,2,2) and GTCAW(3,2,1,2) models of example (c)

Para.	Estimate			Para.	Estimate		
TCAW(3,2,2)(GTCAW(3,2,0,2))							
Regime		1				2	
	0.0000				0.1258		
$\boldsymbol{A}_1^{(1)}$		−0.2073		$\boldsymbol{A}_1^{(2)}$		0.1533	
			−0.1455				0.1911
	0.1422				0.0000		
$\boldsymbol{A}_2^{(1)}$		0.0534		$\boldsymbol{A}_1^{(2)}$		−0.0799	
			0.0986				0.0370
	0.0000				0.5423		
$\boldsymbol{B}_1^{(1)}$		0.0011		$\boldsymbol{B}_1^{(2)}$		0.3720	
			−0.4036				0.0559
	0.0011				0.0037		
$\boldsymbol{B}_2^{(1)}$		0.0003		$\boldsymbol{B}_2^{(2)}$		0.0010	
			0.0002				0.0021
	0.0403				0.3936		
$\boldsymbol{B}_3^{(1)}$		−0.0028		$\boldsymbol{B}_3^{(2)}$		0.6597	
			0.1906				0.3067
	0.1069				0.1037		
$\boldsymbol{C}^{(1)}$	0.0414	0.1021		$\boldsymbol{C}^{(2)}$	0.0292	0.0312	
	0.0390	0.0195	0.0940		0.0462	0.0638	0.0799
ν	12.1021			l	0.2020×10^{-4}		
GTCAW(3,2,1,2)							
Regime		1				2	
	0.0000				0.0000		
$\boldsymbol{A}_1^{(1)}$		−0.2157		$\boldsymbol{A}_1^{(2)}$		−0.1116	
			−0.0991				−0.1649
	0.1228				0.1165		
$\boldsymbol{A}_2^{(1)}$		0.0473		$\boldsymbol{A}_2^{(2)}$		0.0985	
			0.1316				−0.0169
	0.0000				0.0021		
$\boldsymbol{B}_1^{(1)}$		−0.0001		$\boldsymbol{B}_1^{(2)}$		0.0012	
			−0.0027				−0.0012
	0.0021				0.0032		
$\boldsymbol{B}_2^{(1)}$		0.0002		$\boldsymbol{B}_2^{(2)}$		0.0002	
			0.0004				0.0065
	0.0339				0.0000		
$\boldsymbol{B}_3^{(1)}$		0.0003		$\boldsymbol{B}_3^{(2)}$		0.7957	
			0.1918				0.4723
	0.0894				0.1308		
$\boldsymbol{C}^{(1)}$	0.0374	0.0956		$\boldsymbol{C}^{(2)}$	0.0320	0.0041	
	0.0383	0.0243	0.0959		0.0439	0.0924	0.0057
	0.6185	−0.0888	−0.0533		0.3528	−0.0637	0.1794
$\boldsymbol{M}_1^{(1)}$	0.1555	0.1056	0.0358	$\boldsymbol{M}_1^{(2)}$	0.0032	−0.2712	0.4120
	0.0311	−0.3085	0.4127		0.0326	0.0879	0.2490
ν	11.9890			l	0.2040×10^{-4}		

Table 12. Forecast errors for preferred GTCAW models and corresponding non-threshold cases

	GTCAW model				GCAW model			
Example	Order	dim($\boldsymbol{\phi}$)	SN$\times 10^{-4}$	FN$\times 10^{-4}$	Order	dim($\boldsymbol{\phi}$)	SN$\times 10^{-4}$	FN$\times 10^{-4}$
(a)	(2,1,0,2)	20	0.5806	0.6161	(2,1,0)	10	0.6606	0.7115
(b)	(1,1,0,2)	16	0.7142	0.7469	(1,1,0)	8	0.7597	0.7875
(c)	(3,2,1,2)	62	0.2068	0.2169	(3,2,1)	31	0.2233	0.2332

ML-estimation can be implemented by the closed form conditional density function and model checking can be derived by the conditional expectation and variance. We also study the finite sample performance of the ML-estimation using Monte Carlo simulation, which shows reasonable results. For empirical applications, the GTCAW models outperform their non-threshold counterparts in terms of fitting and forecasting by comparison.

For simplicity, we choose the regime $s = 2$, and impose a diagonal structure on the $\boldsymbol{A}_k$ and $\boldsymbol{B}_i$ in every regime in case that the curse of dimensionality problem appears to be less acute for the proposed model. We will discuss the more complex cases ($s > 2$) and study the asymptotic properties of the model in the future.

ACKNOWLEDGEMENTS

The authors thank two anonymous referees and the associate editor for their helpful and constructive comments. Li's work is supported partially by the Hong Kong General Research Fund grant 17303315. Zhu's work is supported by National Natural Science Foundation of China (Nos. 11871027, 11731015), Science and Technology Developing Plan of Jilin Province (No. 20170101057JC), and Cultivation Plan for Excellent Young Scholar Candidates of Jilin University.

Received 30 August 2018

REFERENCES

[1] Andersen, T.G., Bollerslev, T., Diebold, F.X. and Labys, P. (2003). Modeling and forecasting realized volatility. *Econometrica*, **71**, 579–625. MR1958138

[2] Asai, M., McAleer, M. and J, Yu. (2006). Multivariate stochastic volatility: a review. *Econometric Reviews*, **25**, 145–175. MR2256285

[3] Barndorff-Nielsen, O.E. and Shephard, N. (2004). Econometric analysis of realized covariation: high frequency based covariance, regression, and correlation in financial economics. *Econometrica*, **72**, 885–925. MR2051439

[4] Bauwens, L., Laurent, S. and Rombouts, J.V.K. (2006). Multivariate GARCH models: a survey. *Journal of Applied Econometrics*, **21**, 79–109. MR2225523

[5] Bollerslev, T., Engle, R.F. and Wooldridge, J.M. (1988). A capital asset pricing model with time-varying covariances. *Journal of Political Economy*, **96**, 116–131.

[6] Box, G.E.P. and Pierce, D.A. (1970). Distribution of the residual autocorrelations in autoregressive integrated moving average time series models. *Journal of the American Statistical Association*, **65**, 1509–1526. MR0273762

[7] Brooks, C. (2001). A double-threshold GARCH model for the French Franc/Deutschmark exchange rate. *Journal of Forecasting*, **20**, 135–143.

[8] Callot, L.A.F., Kock, A.B. and Medeiros, M.C. (2017). Modeling and forecasting large realized covariance matrices and portfolio choice. *Journal of Applied Econometrics*, **32**, 140–158. MR3611064

[9] Chan, K.S. (1993). Consistency and limiting distribution of the least squares estimator of a threshold autoregressive model. *Annals of Statistics*, **21**, 520–533. MR1212191

[10] Chen, C.W.S., So, M.K.P. and Liu, F.C. (2011). A review of threshold time series models in finance. *Statistics and Its Interface*, **2**, 167–181. MR2812813

[11] Cheng, X., Li, W.K., Yu, P.L.H., Zhou, X., Wang, C. and Lo, P.H. (2011). Modeling threshold conditional heteroscedasticity with regime-dependent skewness and kurtosis. *Computational Statistics and Data Analysis*, **55**, 2590–2604. MR2802338

[12] Engle, R.F. and Kroner, K.F. (1995). Multivariate simultaneous generalized ARCH. *Econometric Theory*, **11**, 122–150. MR1325104

[13] Golosnoy, V., Gribisch, B. and Liesenfeld, R. (2012). The conditional autoregressive Wishart model for multivariate stock market volatility. *Journal of Econometrics*, **167**, 211–223. MR2885447

[14] Gourieroux, C., Jasiak, J. and Sufana, R. (2009). The Wishart autoregressive process of multivariate stochastic volatility. *Journal of Econometrics*. **150**, 167–181. MR2535514

[15] Harvey, A., Ruiz, E. and Shephard, N. (1994). Multivariate stochastic variance models. *Review of Economic Studies*, **61**, 247–264.

[16] Hosking, J.K.M. (1980). The multivariate portmanteau statistic. *Journal of the American Statistical Association*, **75**, 602–607. MR0590689

[17] Jin, X. and Maheu, J.M. (2013). Modeling realized covariances and returns. *Journal of Financial Econometrics*, **11**, 335–369.

[18] Jin, X. and Maheu, J.M. (2016). Bayesian semiparametric modeling of realized covariance matrices. *Journal of Econometrics*, **192**, 19–39. MR3463662

[19] Kwan, W., Li, W.K. and Ng, K.W. (2010). A multivariate threshold varying conditional correlations model. *Econometric Reviews*, **29**, 20–38. MR2750495

[20] Li, C.W. and Li, W.K. (1996). On a double-threshold autoregressive heteroscedastic time series model. *Journal of Applied Econometrics*, **11**, 253–274.

[21] Muirhead, R.J. (1982). *Aspects of Multivariate Statistical Theory*. Wiley, New Jersey. MR0652932

[22] Pelletier, D. (2006). Regime switching for dynamic correlations. *Journal of Econometrics*, **131**, 445–473. MR2276007

[23] Shen, K., Yao, J. and Li, W.K. (2018). Forecasting high-dimensional realized volatility matrices using a factor model. *Quantitative Finance*, forthcoming.

[24] Tao, M., Wang, Y. and Chen, X. (2013). Fast convergence rates in estimating large volatility matrices using high-frequency financial data. *Econometric Theory*, **29**, 838–856. MR3092465

[25] Tong, H. and Lim, K.S. (1980). Threshold autoregression, limit cycles and cyclical data (with discussion). *Journal of the Royal Statistical Society: Series B*, **42**, 245–292.

[26] Tsay, R.S. (1998). Testing and modeling multivariate threshold models. *Journal of the American Statistical Association*, **93**, 1188–1202. MR1649212

[27] Tse, Y.K. and Tsui, A.K.C. (2002). A multivariate generalized autoregressive conditional heteroscedasticity model with time-varying correlations. *Journal of Business and Economic Statistics*, **20**, 351–362. MR1939906

[28] Wang, C., Liu, H., Yao, J.F., Davis, R.A. and Li, W.K. (2014). Self-excited threshold Poisson autoregression. *Journal of the American Statistical Association*, **109**, 777–787. MR3223749

[29] Yu, P.L., Li, W.K. and Ng, F.C. (2017). The generalized conditional autoregressive Wishart model for multivariate realized volatility. *Journal of Business and Economic Statistics*, **35**, 513–527. MR3716040

Yan Cui
School of Mathematics
Jilin University
Changchun
China
E-mail address: ycui0311@163.com

Fukang Zhu
School of Mathematics
Jilin University
Changchun
China
E-mail address: zfk8010@163.com

Wai Keung Li
Faculty of Liberal Arts and Social Sciences
Department of Mathematics and Information Technology
The Education University of Hong Kong
Hong Kong
China
E-mail address: waikeungli@eduhk.hk