Article:

Iqbal, F. and Triantafyllopoulos, K. orcid.org/0000-0002-4144-4092 (2019) Bayesian inference of multivariate rotated GARCH models with skew returns. Communications in Statistics - Simulation and Computation. ISSN 0361-0918

https://doi.org/10.1080/03610918.2019.1620272

Bayesian inference of multivariate rotated GARCH models with skew returns

Farhat Iqbal and Kostas Triantafyllopoulos
Department of Statistics, University of Balochistan, Quetta, Pakistan
and School of Mathematics and Statistics,
Hicks Building, University of Sheffield, S3 7RH, Sheffield, UK
email: kostas@sheffield.ac.uk, Tel: +44 114 222 3741

May 10, 2019

Abstract

Bayesian inference is proposed for volatility models, targeting financial returns, which exhibit high kurtosis and slight skewness. Rotated GARCH models are considered which can accommodate the multivariate standard normal, Student t, generalised error distributions and their skewed versions. Inference on the model parameters and prediction of future volatilities and cross-correlations are addressed by Markov chain Monte Carlo inference. Bivariate simulated data is used to assess the performance of the method, while two sets of real data are used for illustration: the first is a trivariate data set of financial stock indices and the second is a higher dimensional data set for which a portfolio allocation is performed.

Keywords: Volatility, skew returns, GARCH, BEKK, rotated BEKK, multivariate time series, portfolio allocation.

1 Introduction

The time-varying dynamics of conditional covariances of asset returns play a crucial role for asset pricing, portfolio allocation and risk management and hence their modelling and forecasting has gained considerable attention for the past three decades. Multivariate generalised autoregressive conditional heteroscedastic (MGARCH) models have been routinely used to study and examine the relationship between the volatilities and co-volatilities of multivariate financial time series.

A wide range of MGARCH models have been proposed in the literature to accommodate time-varying multivariate volatility but the number of parameters increases rapidly as the the dimension of the returns grows; this is widely known as the *curse of dimensionality*, see e.g. Bauwens et al. (2006) and Silvennoinen and Teräsvirta (2009), among others. Bollerslev et al. (1988) first introduced the half-vec (*vech*) form for the conditional covariance matrices and its special case is the popular Baba-Engle-Kraft-Kroner (BEKK) model of Engle and Kroner (1995). Though the BEKK model provides rich dynamics of conditional covariances, the estimation of this model demands heavy computations. The diagonal BEKK model, where parameter-matrices are assumed diagonal, provides some simplification over the full BEKK model. Several models have been proposed in the literature based on transformations of the returns (van der Weide, 2002; Fan et al., 2008; Boswijk and van der Weide, 2011). Noureldin et al. (2014) proposed the rotated BEKK

(RBEKK) model that utilises the BEKK parametrisation using covariance targeting and aiming at higher dimensional data by exploiting returns rotation. With respect to estimation, maximum likelihood is usually adopted to carry out inference of multivariate GARCH-type models. On the other hand, the Bayesian paradigm is well suited for MGARCH models and provides certain advantages in comparison to the llikelihood-based inference. Several papers have adopted Bayesian estimation, in particular Markov chain Monte Carlo (MCMC) inference, see e.g. Vrontos et al. (2003), Osiewalski and Pipien (2004) and the review of Virbickaite et al. (2015).

The main contribution of the article is to develop a Bayesian approach for the estimation of the parameters of the RBEKK models allowing for heavy-tailedness and asymmetry in the distribution of the returns. Under a Bayesian framework, a block-sampling MCMC scheme is proposed for the estimation of the rotated volatility covariance matrix. Results of Monte Carlo simulations show that the method accurately estimates the volatilities and correlations of multivariate returns. The model is first evaluated via a bivariate simulated data set; for this data set the true simulated volatilities, correlations and model parameters are always within the in-sample credible intervals, which are remarkably narrow. Consequently, an empirical study is considered for a trivariate data set consisting of Frankfurt (Dax), Paris (Cac40) and Tokyo (Nikkei) stock indices. Finally, a higher dimensional data set is analysed consisting of eight shares from Dow Jones industrial average index, for which an asset allocation is performed and its evaluation is conducted using cumulative portfolio returns. The assessment and performance of the model suggest very accurate in-sample and volatility forecasting providing clearly improved performance in comparison to maximum likelihood estimation. The utility of this paper is to offer MCMC inference for the volatility of financial returns, which exhibit heavy tails and asymmetry; the covariance targeting approach allows the estimation of higher dimensional data sets.

The structure of the remainder of the article is as follows. In Section 2, we discuss the specification of the MGARCH model and the rotated BEKK models. Section 3 describes in detail the proposed Bayesian inference and forecasting for the RBEKK model. Results of simulation studies are presented in Section 4 and application to the real data sets are discussed in Section 5. Finally, the paper concludes with closing comments.

2 Multivariate GARCH models

Consider a general multivariate GARCH model

$$\mathbf{r}_t = \mathbf{H}_t^{1/2} \boldsymbol{\varepsilon}_t, \tag{1}$$

where $\mathbf{r}_t = (r_{1t}, \cdots, r_{Kt})'$, $t = 1, 2, \ldots, T$, is a K-dimensional daily asset returns, $\boldsymbol{\varepsilon}_t$ is a K-dimensional i.i.d process with mean zero and identity covariance matrix. This model assumes a zero-mean $\mathrm{E}[\mathbf{r}_t|\boldsymbol{\Omega}_{t-1}] = \mathbf{0}$ and conditional covariance matrix $\mathrm{E}[\mathbf{r}_t\mathbf{r}_t'|\boldsymbol{\Omega}_{t-1}] = \mathbf{H}_t$, with elements $H_{ij,t}, i = j = 1,\ldots,K$, where $\boldsymbol{\Omega}_{t-1}$ denotes the information set at time $t-1$ and $\mathrm{E}[\cdot]$ denotes expectation. A parametrisation for the conditional covariance matrix $\mathbf{H}_t$ completes the multivariate GARCH model.

One of the most widely used models for the conditional covariances adopts the BEKK specification

$$\mathbf{H}_t = \mathbf{C}\mathbf{C}' + \mathbf{A}\mathbf{r}_{t-1}\mathbf{r}_{t-1}'\mathbf{A}' + \mathbf{B}\mathbf{H}_{t-1}\mathbf{B}', \tag{2}$$

where $\mathbf{C}, \mathbf{A}$ and $\mathbf{B}$ are $K \times K$ square matrices with $\mathbf{C}$ being a positive definite symmetric parameter matrix. The fully parameterised model includes $2.5K^2 + 0.5K$ parameters and only feasible for small value of K.

Noureldin et al. (2014) propose a rotated version of the above model, the rotated BEKK (RBEKK) model that utilises the BEKK parametrisation using covariance targeting. More specifically, the model is fitted using the rotated returns

$$\tilde{\mathbf{r}}_t = \mathbf{H}^{*-1/2}\mathbf{r}_t = \mathbf{P}\boldsymbol{\Lambda}^{-1/2}\mathbf{P}'\mathbf{r}_t, \tag{3}$$

where $\mathbf{H}^* = \mathbf{P}\boldsymbol{\Lambda}\mathbf{P}'$ is the unconditional covariance of $\mathbf{r}_t$ and the matrices of eigenvectors $\mathbf{P}$ and eigenvalue $\boldsymbol{\Lambda}$ are obtained using spectral decomposition. Now, the unconditional covariance matrix of rotated returns $\tilde{\mathbf{r}}_t$ is $\mathrm{Var}[\tilde{\mathbf{r}}_t] = \mathbf{I}_K$, where $\mathbf{I}_K$ denotes the K-dimensional identity matrix. Using a covariance-targeting the conditional covariance matrix of $\tilde{\mathbf{r}}_t$ is defined as

$$\mathbf{G}_t = (\mathbf{I}_K - \tilde{\mathbf{A}}\tilde{\mathbf{A}}' - \tilde{\mathbf{B}}\tilde{\mathbf{B}}') + \tilde{\mathbf{A}}\tilde{\mathbf{r}}_{t-1}\tilde{\mathbf{r}}'_{t-1}\tilde{\mathbf{A}}' + \tilde{\mathbf{B}}\mathbf{G}_{t-1}\tilde{\mathbf{B}}', \tag{4}$$

with $\mathbf{G}_0 = \mathbf{I}_K$ and assume $(\mathbf{I}_K - \tilde{\mathbf{A}}\tilde{\mathbf{A}}' - \tilde{\mathbf{B}}\tilde{\mathbf{B}}') \geq 0$ on the sense of being positive semidefinite. The RBEKK model is the restricted BEKK model with parameters $\mathbf{A}^* = \mathbf{H}^{*1/2}\tilde{\mathbf{A}}\mathbf{H}^{*-1/2}$, $\mathbf{B}^* = \mathbf{H}^{*1/2}\tilde{\mathbf{B}}\mathbf{H}^{*-1/2}$ and $\mathbf{C}^* = \mathbf{H}^{*1/2}(\mathbf{I}_K - \tilde{\mathbf{A}}\tilde{\mathbf{A}}' - \tilde{\mathbf{B}}\tilde{\mathbf{B}}')\mathbf{H}^{*1/2}$. A high-order lag structure or asymmetric term can also be introduced in (4).

Noureldin et al. (2014) study three different specifications of RBEKK models called scalar RBEKK (S-RBEKK), diagonal RBEKK (D-RBEKK) and common persistence RBEKK (CP-RBEKK). Here, we discuss the scalar and diagonal specifications. The scalar specification assumes $\tilde{\mathbf{A}} = \alpha^{1/2}\mathbf{I}_K$ and $\tilde{\mathbf{B}} = \beta^{1/2}\mathbf{I}_K$. The (i,j)th element of $\mathbf{G}_t$ is given by

$$g_{ij,t} = (1 - \alpha - \beta)\mathrm{I}_{(i=j)} + \alpha\tilde{r}_{i,t-1}\tilde{r}_{j,t-1} + \beta g_{ij,t-1}, \qquad i,j = 1,\ldots,K,$$

where $\mathrm{I}_{(\cdot)}$ is the indicator function. Note that in this specification, all the elements of $\mathbf{G}_t$ have the same dynamic parameters. It is assumed that $\alpha, \beta \geq 0$ and $\alpha + \beta < 1$ is required for covariance stationarity. The total number of parameters is 2 in the scalar specification.

In the diagonal specification, $\tilde{\mathbf{A}}$ and $\tilde{\mathbf{B}}$ are diagonal with elements $\alpha_{ii}^{1/2}$ and $\beta_{ii}^{1/2}$, $i = 1\ldots,K$, i.e., $\tilde{\mathbf{A}} = \mathrm{diag}\{\alpha_{ii}^{1/2}\}$ and $\tilde{\mathbf{B}} = \mathrm{diag}\{\beta_{ii}^{1/2}\}$. This specification implies

$$g_{ij,t} = (1 - \alpha_{ii}^{1/2}\alpha_{jj}^{1/2} - \beta_{ii}^{1/2}\beta_{jj}^{1/2})\mathrm{I}_{(i=j)} + \alpha_{ii}^{1/2}\alpha_{jj}^{1/2}\tilde{r}_{i,t-1}\tilde{r}_{j,t-1} + \beta_{ii}^{1/2}\beta_{jj}^{1/2}g_{ij,t-1}, \qquad i, j = 1, \ldots, K.$$

Assuming $\alpha_{ii}^{1/2} > 0$ and $\beta_{ii}^{1/2} > 0$ along with $\alpha_{ii} + \beta_{ii} < 1$ ensure that $(\mathbf{I}_K - \tilde{\mathbf{A}}\tilde{\mathbf{A}}' - \tilde{\mathbf{B}}\tilde{\mathbf{B}}')$ is positive definite. The total number of parameters to be estimated in the D-RBEKK model is $2K$.

It is noted that the BEKK and REBEKK models have the same number of parameters. The diagonal RBEKK model implies a full BEKK model for returns. Besides, fitting a diagonal RBEKK model implies rather rich dynamics for the unrotated returns (Noureldin et al. 2014, p. 18). The transformation of the raw returns enables the fitting of flexible multivariate models to the rotated returns using covariance targeting (Noureldin et al., 2014). Since the diagonal BEKK model implies a full BEKK model for unrotated returns, fitting this model provides rich dynamics with smaller number of parameters. This may be attractive for modelling both the volatilities and correlations in large dimensions as only $2K$ dynamic parameters need to be estimated in the diagonal specification.

3 Bayesian inference for RBEKK Models

3.1 Preliminaries

Estimation of the RBEKK models involves a two-step estimation procedure. In the first step, a method of moment estimator is used to obtain an estimate of the rotated volatility matrix $\mathbf{H}^*$:

$$\widehat{\mathbf{H}}^* = T^{-1} \sum_{t=1}^{T} \mathbf{r}_t \mathbf{r}_t'.$$

By noting that $\mathbf{H}^* = \mathbf{P}\boldsymbol{\Lambda}\mathbf{P}'$ (see equation (3)), this estimator can be decomposed into $\widehat{\mathbf{P}}$ and $\widehat{\boldsymbol{\Lambda}}$ to obtain the rotated returns $\tilde{\mathbf{r}}_t = \widehat{\mathbf{P}}\widehat{\boldsymbol{\Lambda}}^{-1/2}\widehat{\mathbf{P}}'\mathbf{r}_t$, $t = 1, 2, \ldots, T$.

In the second step, we adopt a Bayesian approach to estimate the parameters of RBEKK models with skewed and heavy-tailed distributions for the errors by constructing MCMC algorithms. The conditional likelihood function for model (1) can be written as

$$L(\boldsymbol{\theta}; \mathbf{r}) = \prod_{t=1}^{T} |\mathbf{H}_t|^{-1/2} p_\varepsilon(\mathbf{H}_t^{-1/2}\mathbf{r}_t), \tag{5}$$

where $\mathbf{r} = (\mathbf{r}_1', \ldots, \mathbf{r}_T')'$ is the sample of returns and p_ε is the joint density function for $\boldsymbol{\varepsilon}_t$, and $\boldsymbol{\theta}$ is the vector of unknown parameters in the model. The likelihood function for (3) can be defined in a similar manner.

It is noted that MCMC is proposed only in the second step (estimation on the rotated conditional covariance matrix). This enables the study of uncertainty around the volatility of the rotated returns, but it does not provide uncertainty around the the matrix transformation in order to obtain the rotated returns. Joint Bayesian estimation of the conditional returns and the unconditional returns might be possible, but it is not explored in this paper any further. Markov chain Monte Carlo estimation of the unconditional returns in Step 1 would negate the advantage of the reduction of parameters in the model. One possibility to move forward this idea is to estimate the unconditional covariance volatility matrix in the first step by employing particle filters (Creal, 2012) and then to adopt the proposed MCMC approach for the estimation of the conditional returns in the second step.

3.2 Asymmetric error distribution

The standardised multivariate normal or Student t distributions are often used for the error distribution; the latter being capable to describe heavy tailed returns, which are typically observed in finance. However, as it is common for financial returns to exhibit asymmetry a suitable asymmetric error distribution may be considered. Bauwens and Laurent (2005) describe a method of constructing a multivariate skew distribution from a symmetric one. They show that the multivariate skew densities can be written as

$$s(\mathbf{x}|\boldsymbol{\gamma}) = 2^K \left(\prod_{i=1}^{K} \frac{\gamma_i}{1+\gamma_i^2} \right) f(\mathbf{x}^*),$$

where $f(\cdot)$ is a symmetric multivariate density, $\mathbf{x}^* = (x_1^*, \ldots, x_K^*)'$, $\gamma_i > 0$ is the shape parameter, $x_i^* = x_i \gamma_i^{I_i}, i = 1, \ldots, K$, with $I_i = -1$, if $x_i \geq 0$, and 1 otherwise, being the indicator function. Note that $\gamma_i = 1$ yields the symmetric distribution, $\gamma_i > 1 (< 1)$ indicates right (left) skewness and $\gamma_i^2 = Pr(x_i \geq 0)/Pr(x_i < 0)$ (Fernandez and Steel, 1998). The first two moments of x_i^* are given by

$$m_i = M_{i,1}\left(\gamma_i - \frac{1}{\gamma_i}\right), \tag{6}$$

$$s_i = \left(M_{i,2} - M_{i,1}^2\right)\left(\gamma_i^2 + \frac{1}{\gamma_i^2}\right) + 2M_{i,1}^2 - M_{i,2}, \tag{7}$$

with

$$M_{i,r} = \int_0^\infty 2u^r f_i(u) du.$$

The log-likelihood function for multivariate skew-normal distribution is

$$\begin{aligned} L(\boldsymbol{\theta}) = -\frac{1}{2}\sum_{t=1}^{T}\left[\log|\mathbf{H}_t| + \sum_{i=1}^{K}\left(s_i \sum_{j=1}^{K} p_{ij,t} r_{j,t} + m_i\right)^2 \gamma_i^{2I_i}\right] \\ + T\left[\sum_{i=1}^{K}(\log\gamma_i + \log s_i) - log(1+\gamma_i^2)\right] + \frac{TK}{2}[\log(2) - \log(\pi)], \end{aligned} \tag{8}$$

where $p_{ij,t}$ corresponds to the j^{th} element of the i^{th} row of the inverse Cholesky factor of the matrix $\mathbf{H}_t$.

The multivariate generalised error distribution (GED) can also be used as an alternative heavy tailed distribution. For the GED, the tail parameter δ needs to be estimated along with other parameters. The multivariate standard normal is obtained for $\delta = 2$ and the value $\delta < 2 (> 2)$ leads to thicker (thinner) tails than the standard normal (see Fioruci et al., 2014 for further discussion on univariate and multivariate skew distributions). For expressions of the likelihood function of models with skew t and GED return distributions the reader is referred to Braione and Scholtes (2016) and to references therein.

3.3 Markov chain Monte Carlo estimation

Let $\boldsymbol{\theta}$ denote the set of all unknown parameters which includes the parameters of RBEKK models, $(\alpha^{1/2}, \beta^{1/2})$ in case of the scalar specification and $(\alpha_{11}^{1/2}, \ldots, \alpha_{KK}^{1/2}, \beta_{11}^{1/2}, \ldots, \beta_{KK}^{1/2})$ in case of the diagonal, the shape parameters for each returns $(\gamma_1, \ldots, \gamma_K)$ and the tail parameter ν or δ when using the multivariate Student t or GED, respectively.

In a Bayesian framework (Chib and Greenberg, 1995) prior distributions for all parameters of interest need to be specified. These are assumed to be *a priori* independent and normally distributed truncated to the intervals they are defined. We focus our attention on the diagonal specification as this provides rich dynamics compared to the scalar specification. For $\alpha_{ii}^{1/2}$ and $\beta_{ii}^{1/2}, i = 1, \ldots, K$, we assume $\alpha_{ii}^{1/2} \sim N(\mu_{\alpha_{ii}^{1/2}}, \sigma^2_{\alpha_{ii}^{1/2}}) I_{(0<\alpha_{ii}^{1/2}<1)}$ and $\beta_{ii}^{1/2} \sim N(\mu_{\beta_{ii}^{1/2}}, \sigma^2_{\beta_{ii}^{1/2}}) I_{(0<\beta_{ii}^{1/2}<1)}$. A prior distribution for the tail parameter (the degrees of freedom) is $\nu \sim N(\mu_\nu, \sigma^2_\nu) I_{(\nu>2)}$, for the multivariate Student t or $\delta \sim N(\mu_\delta, \sigma^2_\delta) I_{(\delta>0)}$, for multivariate GED. The choice of prior distribution for skewness parameter is adopted from Fioruci et al. (2014), that is $\gamma_i \sim$

$N(0, 0.64^{-1})I_{(\gamma_i>0)}$. The values for hyper parameters are specified as $\mu_{\alpha_{ii}} = \mu_{\beta_{ii}} = \mu_\nu = \mu_\delta = 0$ and $\sigma^2_{\alpha_{ii}} = \sigma^2_{\beta_{ii}} = \sigma^2_\nu = \sigma^2_\delta = 100$. The relatively large variances reflect on a weakly informative prior specification; it is possible to set up a hierarchical prior structure by specifying an inverse prior distribution for each variance, but there was little benefit in the estimation and this approach was not adopted to avoid extra computation cost.

A block Metropolis-Hastings algorithm is constructed to sample from the posterior distribution $p(\boldsymbol{\theta}|\mathbf{r}_t)$ where all the parameters are updated as a block. More specifically, a random walk Metropolis algorithm is used where at each iteration a new vector from a multivariate normal distribution centred around the current vector with a variance-covariance proposal matrix is generated. The proposal matrix is calculated from a pilot tuning that is carried out by running one-dimensional random walk Metropolis updates with univariate normal candidate distributions whose variances are calibrated to obtain good acceptance rates. The main steps of the algorithm are:

1. Initialize the parameter vector $\boldsymbol{\theta}^0$ and set $n = 0$.
2. Draw a sample $\boldsymbol{\theta}^{(n+1)}$ from the distribution of $\boldsymbol{\theta}|\mathbf{r}_t$.
3. Set $n = n + 1$ and go to step 2, until $n = N$, for a large N.

In step 2 we sample from the conditional posterior distribution of $\boldsymbol{\theta}$ whose kernel is given by

$$\kappa(\boldsymbol{\theta}|\mathbf{r}_t) = p(\boldsymbol{\theta})L(\boldsymbol{\theta}|\mathbf{r}_t), \tag{9}$$

where $p(\boldsymbol{\theta})$ is the prior probability of $\boldsymbol{\theta}$. The random walk Metropolis-Hastings method is used for this purpose as described in the following two steps:

First, we generate a candidate vector $\tilde{\boldsymbol{\theta}}$ from the multivariate normal distribution $N(\boldsymbol{\theta}^{(n)}, c\widehat{\boldsymbol{\Sigma}}_{\boldsymbol{\theta}})$, where c is a constant and $\widehat{\boldsymbol{\Sigma}}_{\boldsymbol{\theta}}$ is the covariance matrix calculated from a pilot tuning. Let

$$\tau_{\boldsymbol{\theta}}^{(n)} = \min\{1, (\kappa(\tilde{\boldsymbol{\theta}}|\mathbf{r}_t))/(\kappa(\boldsymbol{\theta}^{(n)}|\mathbf{r}_t))\},$$

where $\kappa(\tilde{\boldsymbol{\theta}}|\mathbf{r}_t)$ is given in eq. (9).

Then, define

$$\boldsymbol{\theta}^{(n+1)} = \begin{cases} \tilde{\boldsymbol{\theta}}, & \text{with probability} \quad \tau_{\boldsymbol{\theta}}^{(n)} \\ \boldsymbol{\theta}^{(n)} & \text{with probability} \quad 1 - \tau_{\boldsymbol{\theta}}^{(n)} \end{cases}$$

The constant c is used to tune the acceptance rate (usually lying between 0.2 and 0.5) to achieve fast convergence. This builds an irreducible and aperiodic Markov chain in the parameter space $\boldsymbol{\theta}^{(0)}, \boldsymbol{\theta}^{(1)}, \ldots, \boldsymbol{\theta}^{(N)}$. For large N, $\boldsymbol{\theta}^{(n)}$ tends in distribution to a random variable with density $p(\boldsymbol{\theta}|\mathbf{r}_t)$.

4 Simulation studies

In this section, we study the performance of the proposed MCMC sampler for the RBEKK model through Monte Carlo simulations. We performed two simulation studies on the bivariate diagonal RBEKK model. Samples of sizes $T = 2000$ and $T = 5000$ were simulated from this model. The true parameter vector for the bivariate RBEKK model (equations (3) and (4)) with diagonal specification is set to $\boldsymbol{\theta} = (\alpha_{11}^{1/2} = \sqrt{0.05}, \alpha_{22}^{1/2} = \sqrt{0.10}, \beta_{11}^{1/2} = \sqrt{0.90}, \beta_{22}^{1/2} = \sqrt{0.80})$. Note that the diagonal specification assumes that $\tilde{\mathbf{A}} = \text{diag}\{\alpha_{ii}^{1/2}\}$ and $\tilde{\mathbf{B}} = \text{diag}\{\beta_{ii}^{1/2}\}$, for $i = 1, 2$. For the sake of brevity, results for errors from the multivariate skew t distribution, $ST(\nu, \boldsymbol{\gamma})$, with $\nu = 8$ are only reported. For the shape parameter, we selected negatively skewed errors with $\boldsymbol{\gamma} = (0.8, 0.8)'$. Such a setting generates heavy tailed and skewed returns that are commonly observed in financial time series.

The proposed MCMC algorithm is then used to estimate the parameters of the diagonal RBEKK model for each simulated series. The MCMC algorithm is run using $N = 20000$ iterations discarding the initial 10000 iterations as burn-in samples. Geweke convergence diagnostic (Geweke, 1992) were used to check the convergence of the Markov chains. The MCMC chains provide good mixing performance and fast convergence.

Table 1 shows posterior means, standard deviations and 95% credible intervals of the model parameters for the two simulated series. Observe the accuracy of the estimation and note that the 95% credible intervals always include the true parameters. The posterior standard deviations become smaller as the sample size increases and the length of the credible intervals decrease reflecting upon precision. Besides the availability of point estimates, Bayesian estimation routinely provides parameter uncertainty, e.g. via the

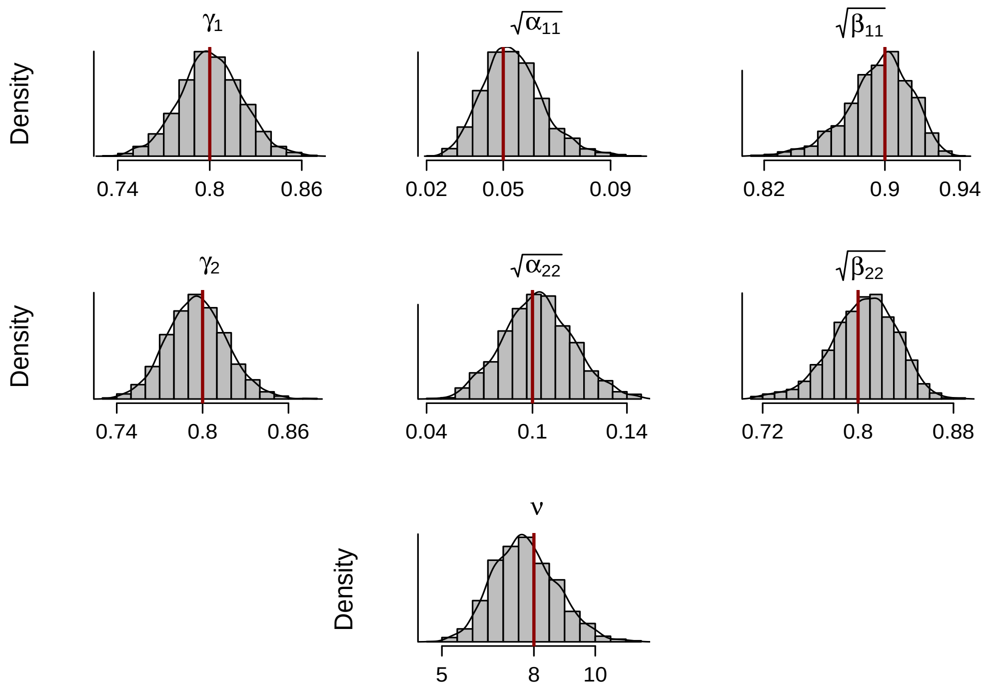

Figure 1: Histograms and density of the posterior MCMC samples of model parameters of the first simulated series with $T = 2000$. Vertical line represents the true value of the parameter.

empirical posterior densities. This parameter uncertainty may be introduced in the estimation of volatilities, correlations, value-at-risk (VaR), portfolio selection, and so on. The histograms and densities of the posterior samples of each parameter for the first simulated series with sample size $T = 2000$ is shown in Figure 1. The histograms for large sample size $T = 5000$ (not reported) exhibit higher degree of symmetry and smaller variance as compared to those of size $T = 2000$.

The in-sample volatilities and correlations are also estimated and pre-

Table 1: Estimates of the model parameters with multivariate skew t errors.

	T = 2000			T = 5000		
Parameter	Mean	SD	95% CI	Mean	SD	95% CI
$\gamma_1 = 0.8000$	0.8006	0.0203	[0.7607, 0.8412]	0.8112	0.0160	[0.7808, 0.8432]
$\sqrt{\alpha_{11}} = 0.2236$	0.2410	0.0196	[0.2041, 0.2808]	0.2280	0.0145	[0.1998, 0.2573]
$\sqrt{\beta_{11}} = 0.9487$	0.9354	0.0122	[0.9072, 0.9557]	0.9402	0.0090	[0.9207, 0.9559]
$\gamma_2 = 0.8000$	0.8367	0.0203	[0.7984, 0.8775]	0.8320	0.0171	[0.7917, 0.8749]
$\sqrt{\alpha_{22}} = 0.3162$	0.2980	0.0202	[0.2582, 0.3384]	0.3055	0.0182	[0.2704, 0.3417]
$\sqrt{\beta_{22}} = 0.8944$	0.8961	0.0157	[0.8626, 0.9238]	0.8843	0.0159	[0.8498, 0.9129]
$\nu = 8$	7.1044	0.3994	[6.5762, 8.2296]	8.0223	0.4212	[7.2345, 8.8977]

sented in Figure 2. The estimated in-sample volatilities, $\widehat{H}_{ii,t}$, for $i = 1, 2$, and the in-sample correlations, $\widehat{R}_{12,t}$, for last 1000 observations, $t = 4901, \ldots,$ 5000 for the simulated series with $T = 5000$ are presented. True values for volatilities and correlations along with 95% credible intervals are plotted and it can be seen that MCMC provides accurate estimates of volatilities and correlations. It can also be noted that the true values for volatilities and correlations are always included in the credible intervals. The point estimates and credible intervals for the one-step-ahead volatilities $\widehat{H}_{ii,T+1}$ and correlations $\widehat{R}_{12,T+1}$ can easily be obtained. These one-step-ahead point predictions (using mean) along with the corresponding predictive intervals and true values are also shown in Figure 2.

Finally, Figure 3 shows plots of volatilities and co-volatilities estimated under the assumption of skew student-t and standard normal errors against true values for a random sample of 100 observations when errors are generated from skew student-t distribution. The plots show that using the skew

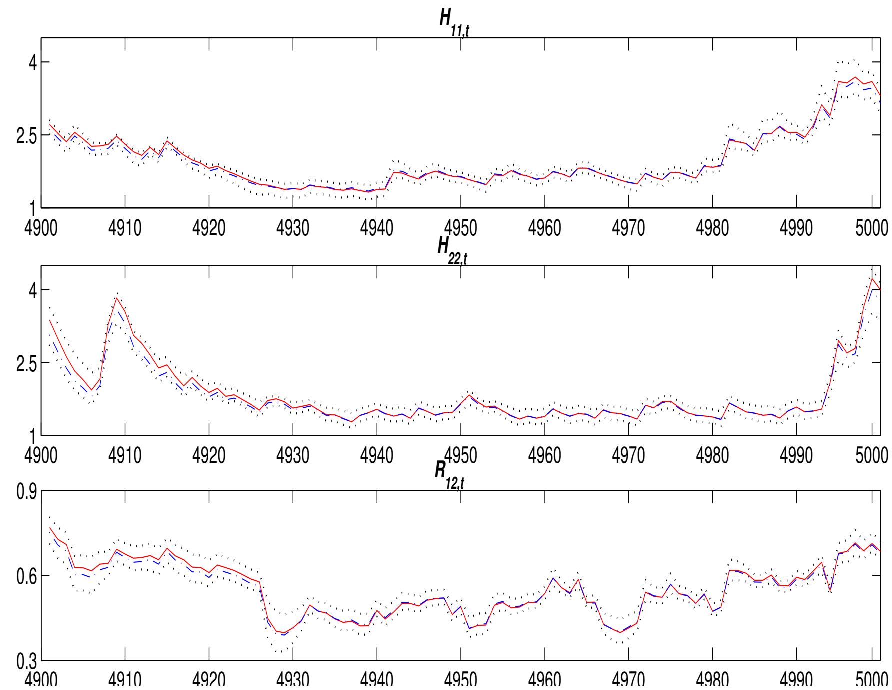

Figure 2: True (solid) and Bayesian (dashed-dot) estimates and 95% intervals (dotted) for volatilities $H_{ii,t}$, for $i = 1, 2$, and correlations $R_{12,t}$, for the last 1000 observations for the simulated series of sample size T= 5000.

t model provides an improved performance from the normal model. The skew t model provides estimates remarkably close to the simulated volatilities, while at some points of time the difference between the two estimates is quite pronounced, e.g. $H_{11,t}$ at $t = 20$. In-sample maximum likelihood estimates (MLEs) provided similar results, and hence they were not reported here. However, we favour here the MCMC approach as it provides more information of uncertainty quantification of the parameters subject to estimation.

Simulations with higher-dimensional systems in which $K = 10$ and $K =$

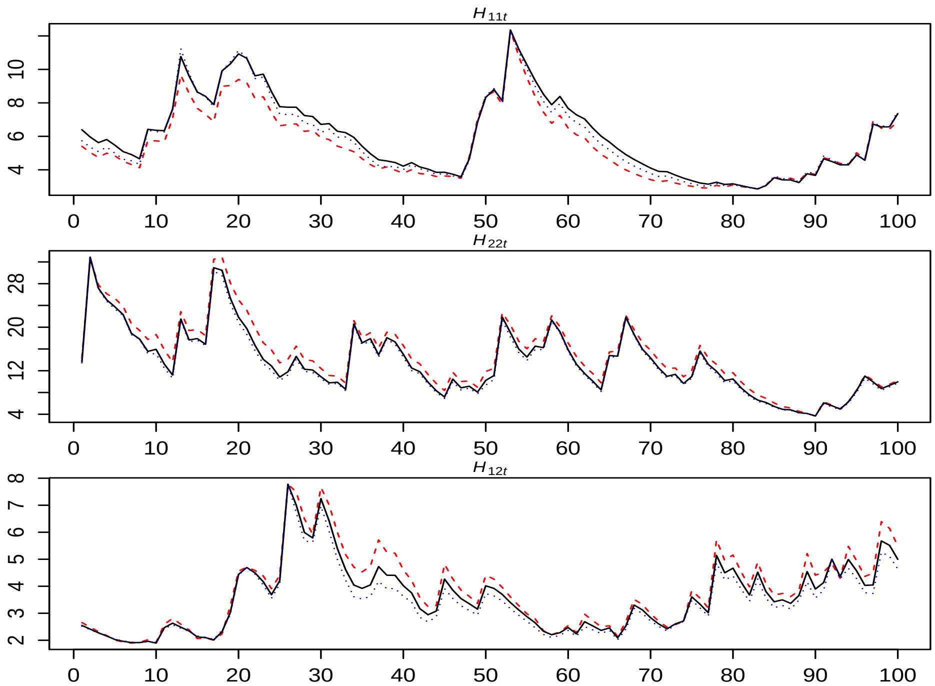

Figure 3: Plots of Bayesian estimates of volatilities and co-volatilities estimated using skew student-t (dotted) and standard normal (dashed) errors against true (solid) values for a random 100 observations from the simulated series of sample size T= 2000 when errors are generated from skew student-t distribution.

20 were also performed; for each simulated data set we have used 20000 iterations for the MCMC. The accuracy of the estimates was not significantly affected by large dimensions, though greater computational cost was required to obtain the estimates, in particular regarding $K = 20$. Table 2 represents the computational time taken by an Intel Core i7 laptop with 16GB RAM for the estimation of RBEKK models. The code is written in the programming language R (`https://www.r-project.org`) with few routines in C++ language. Use of cluster and parallel computing are expected to improve computational time.

Table 2: Computational time (in minutes) taken by various RBEKK models for $T = 1000$.

	Normal	Student-t	GED	skew-Normal	skew-t	skew-GED
$K = 2$	0.66	0.29	0.51	1.30	0.40	0.39
$K = 5$	1.71	1.21	2.36	2.86	0.68	1.00
$K = 10$	6.53	6.83	12.76	18.67	11.94	19.13
$K = 20$	50.61	52.11	58.78	88.76	82.23	94.98

5 Empirical studies

In this section, we illustrate the proposed Bayesian approach using two real data sets.

5.1 DAX, CAC40 and NIKKEI indices

The first data set consists of the daily closing prices of the stock market indices in Frankfurt (DAX), Paris (CAC40), and Tokyo (NIKKEI) from January 04, 2007 to October 31, 2018, a total of 2828 observations. This time period includes the global financial crisis of 2007-2008. The multivariate time series of de-meaned returns $r_{i,t}$ are defined as $100\times\{\log P_{i,t}-\log P_{i,t-1}\}$, where $P_{i,t}$ is the daily closing price for stock i on day t. The skewness coefficients for log-return series of DAX, CAC and NIKKEI are $(0.1181, 0.1111, -0.4940)$ whereas kurtosis are $(10.3499, 10.6778, 10.8184)$. As we can see, all three log-return series have heavy-tails than the normal. The DAX and CAC40 log-returns are slightly positively skewed whereas the NIKKEI log-returns are found highly negatively skewed. Hence, we fit the diagonal RBEKK model with heavy-tailed and skewed distributions for this data set.

A Metropolis-Hastings algorithm with all the parameters updated as a block is adopted for the MCMC updates. A total of 20000 iterations are conducted with initial 10000 discarded in the burn-in stage. The simulated Markov chains were then checked for convergence and good mixing. Visual inspection of the marginal trace plots, density estimates and autocorrelation plots along with formal tests showed good convergence of the Markov chains. The autocorrelation and trace plots for the skewness parameter is shown in Figure 4. The trace plots seem to be stationary; this is evidenced visually by both the trace plots and the ACF plots. However, there seems to be some autocorrelation present. To minimise the effect of autocorrelation, thinning of 5 iterations is applied. Furthermore, we have initiated several chains and all had a similar effect. In addition to that running the chain for longer than 20000 iterations did not seem to have an effect and hence we do not report on longer chains. For this type of data we consider convergence as reasonable, although we do note some autocorrelation effect.

The deviance information criterion (DIC) and log marginal likelihood (LogML) are used to compare various models for the error terms (see e.g. Spiegelhalter et al., 2002, for a discussion). These include normal, skew normal, Student t, skew t, GED and skew GED. Note that smaller values of DIC and LogML are desired for a favourable model. The DIC is given by $\text{DIC} = 2\text{E}[D(\boldsymbol{\theta}_M)] - D(\text{E}[\boldsymbol{\theta}_M])$ where $D(\cdot)$ is the deviance function defined as minus twice the log-likelihood function and $\boldsymbol{\theta}_M$ is the vector of parameters in

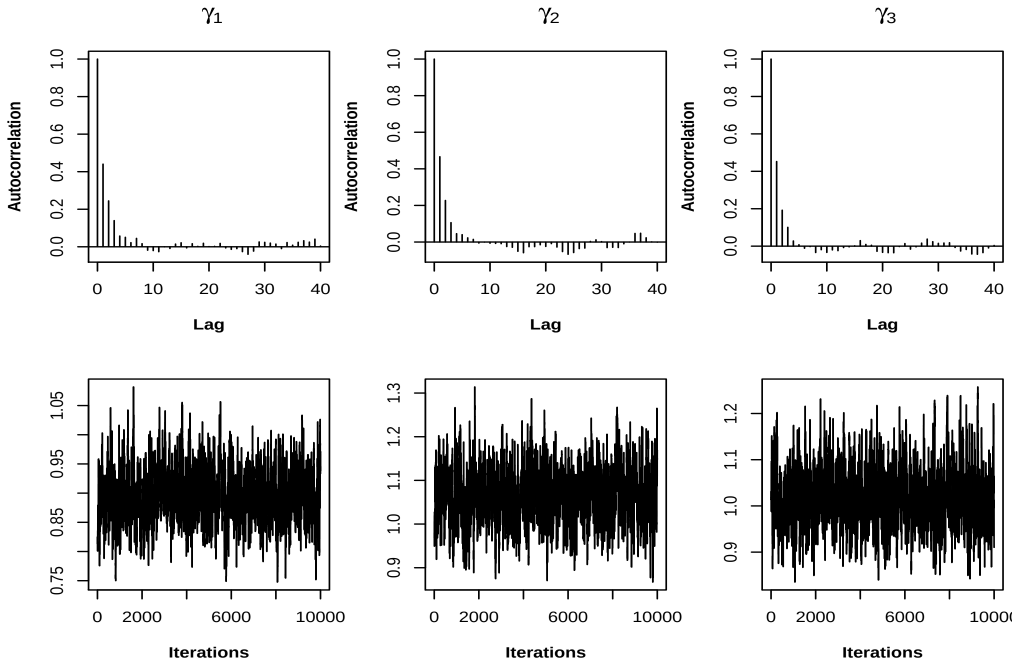

Figure 4: Autocorrelation and trace plots for the skewness parameter.

model M. To check whether the inclusion of skewness substantially improves the model fit, we calculate the weights associated with each DIC. The DIC weights are obtained as

$$w_M \propto \exp\left(-\frac{\mathrm{DIC}_M - \mathrm{DIC}_B}{2}\right),$$

where DIC_B is the value associated with the 'best' model. The DIC weights are then normalised to sum to 1. When the difference of DIC_M from DIC_B is large (i.e. model M does poorly compared to the full model B) then the weight w_M is small, while when $\mathrm{DIC}_M \approx \mathrm{DIC}_B$, then the weight gets close to 1. We also compute the log marginal likelihood (LogML) from the posterior distribution using non-parametric self-normalised importance sampling (Neddermeyer, 2009).

Table 3 below shows the DIC and LogML values for the diagonal RBEKK model, with six different innovation distributions, along with their weights. It can be seen that models with heavier tails exhibit a better behaviour than the normal distribution. Also note that skew distributions provide lower DIC and LogML values as compared to their symmetric versions. Moreover, the multivariate skew t model is found to provide the best fit among the competing models. The DIC weight of this model is also found very large and far from all other models. The log marginal likelihoods of the RBEKK model with Student t and skew t distributions are, respectively, -11127.84 and -11115.34. This implies a Bayes factor of 2.7×10^5 in favor of the RBEKK model with skew t distribution, indicating overwhelming evidence of the latter. These findings indicate that incorporating the heavy tails with asymmetry in the error distribution provides a better fit for RBEKK model and consequently more precise volatilities and correlation estimates. Hence, in the sequel we report results for the multivariate skew t model only.

Table 4 presents the summary of the MCMC estimates of the diagonal RBEKK model using a multivariate skew t distribution for the returns. Posterior means and standard deviations along with 95% credible intervals are displayed. The p-values of the convergence diagnostic (CD) of Geweke (1992) are also presented. For the Dax and NIKKEI log-returns, 95% credible intervals for the skewness parameter do not include 1 and hence confirms

Table 3: DIC and LogML values and weights for various Diagonal RBEKK models.

model	DIC	weight	LogML
Normal	22754.05	0.0000	-11393.97
Student t	22219.65	0.0001	-11127.84
GED	22322.69	0.0000	-11184.51
skew Normal	22708.01	0.0000	-11374.83
skew t	22196.37	0.9999	-11115.34
skew GED	22306.35	0.0000	-11177.78

asymmetry in these series whereas estimates for skewness for CAC40 series are not found significant. The estimate of the tail parameter (ν) indicates the appropriateness of heavy tail of the distribution of the returns.

5.2 Portfolio allocation

In this section we discuss asset allocation, which is one of the main utilities and target applications of volatility estimation, in particularly regarding medium to high dimensional financial time series. Since parameter uncertainty largely affects the optimal asset allocation (Jorion, 1986), the Bayesian paradigm can offer an ideal estimation approach (see Kang, 2011, and Jacquier and Polson, 2013). We consider the global minimum variance (GMV) portfolio, with time-varying covariance matrices, which minimises the portfolio variance. Multivariate GARCH models were first used by Cecchetti et al. (1988) for optimal portfolio allocation and since then many studies have shown that using GARCH-type models reduce the portfolio risk (see Rossi and Zucca, 2002 and Yang and Allen, 2005, among others).

Table 4: MCMC estimates of the Diagonal RBEKK model with multivariate Skew t errors.

		Mean	SD	95% interval	CD
DAX	γ_1	0.9011	0.0214	[0.8595, 0.9441]	0.6841
	α_{11}	0.0525	0.0061	[0.0412, 0.0657]	0.6518
	β_{11}	0.9328	0.0084	[0.9145, 0.9476]	0.9749
CAC40	γ_2	0.9647	0.0254	[0.9175, 1.0157]	0.7397
	α_{22}	0.1108	0.0184	[0.0774, 0.1490]	0.9490
	β_{22}	0.8574	0.0252	[0.8034, 0.9025]	0.7353
NIKKEI	γ_3	0.9441	0.0223	[0.8988, 0.9862]	0.5103
	α_{33}	0.0594	0.0088	[0.0438, 0.0790]	0.5515
	β_{33}	0.9226	0.0123	[0.8947, 0.9432]	0.5809
	ν	6.2094	0.3277	[5.6353, 6.8884]	0.9953

The posterior means are computed by averaging the simulated draws. SD is the standard deviation. The 95% intervals are calculated using the 2.5th and 97.5th percentiles of the simulated draws. The p-values of convergence diagnostic statistic proposed by Geweke (1992) are reported under CD.

The one-step-ahead conditional covariance matrix $\widehat{\mathbf{H}}_{T+1}$ is used to solve the portfolio allocation problem (see Yilmaz, 2011). The optimal portfolio weights for time $T+1$ are obtained by solving the following optimization problem:

$$\mathbf{w}^* = \arg \min_{\mathbf{w}:\mathbf{w}'\mathbf{1}_K=1} \text{Var}[\mathbf{r}_T^*],$$

where $\mathbf{w} = (w_1, \ldots, w_K)'$ is the weight vector, $\mathbf{1}_K$ is a K-vector of ones, $\mathbf{r}_T^* = \mathbf{w}'\mathbf{r}_T$ is the portfolio return at time T and $\mathbf{r}_T$ is the vector of observed returns. Without imposing short-scale constraint, i.e., $w_i \geq 0, \forall i = 1, 2, \ldots, K$, the problem has the following analytical solution

$$\mathbf{w}^*_{T+1} = \frac{\widehat{\mathbf{H}}^{-1}_{T+1}\mathbf{1}_K}{\mathbf{1}'_K\widehat{\mathbf{H}}^{-1}_{T+1}\mathbf{1}_K}. \tag{10}$$

The proposed MCMC estimation enables us to approximate the posterior mean of the optimal portfolio weights as

$$\mathrm{E}[\mathbf{w}^*_{T+1}|\mathbf{r}_T] \approx \frac{1}{N}\sum_{n=1}^{N}\mathbf{w}^{*(n)}_{T+1},$$

where $\left\{\mathbf{w}^{*(n)}_{T+1}\right\}_{n=1}^{N}$ is a posterior sample of the vector of optimal portfolio weights for each value of one-step-ahead conditional covariance matrix $\left\{\widehat{\mathbf{H}}^{(n)}_{T+1}\right\}_{n=1}^{N}$ in the MCMC sample. In this way, we solve the allocation problem at every MCMC iteration and obtain the approximate posterior mean of the optimal portfolio weights. The approximate posterior credible intervals for $\mathbf{w}^*_{T+1}$ can also be obtained in by using the quantiles of the sample of optimal portfolio weights. Similarly, the optimal portfolio variance $\sigma^2_{\mathbf{w},T+1} = \mathbf{w}^{*'}_{T+1}\widehat{\mathbf{H}}_{T+1}\mathbf{w}^*_{T+1}$ can also be calculated and samples can be drawn from its posterior distribution using MCMC.

In our approach uncertainty around the conditional volatility covariance matrix $\left\{\widehat{\mathbf{H}}^{(n)}_{T+1}\right\}_{n=1}^{N}$ is passed onto uncertainty around the portfolio weights and hence one can obtain a predictive sample $\sigma^{2(n)}_{\mathbf{w},T+1}$ $(n = 1,\ldots,N)$ of the portfolio returns $\sigma^2_{\mathbf{w},T+1}$. Following a more traditional approach, an investor may obtain a point estimate of $\widehat{\mathbf{H}}_{T+1}$ (e.g. the mode of $\widehat{\mathbf{H}}^{(n)}_{T+1}$, $n = 1,\ldots,N$) and then solve the minimum portfolio problem once, hence obtain a single forecast value of $\sigma^2_{\mathbf{w},T+1}$. The two approaches should be equivalent for large N, but our approach benefits by providing uncertainty around the volatility of the portfolio returns, which may be vital information for the investor. The mode of the forecast sample $\sigma^{2(n)}_{\mathbf{w},T+1}$ $(n = 1, \ldots, N)$ should provide a point forecast of $\sigma^2_{\mathbf{w},T+1}$, if this is needed.

Bayesian estimates of conditional correlations when estimated with both skew t and Normal distributions for DAX, CAC40 and NIKKEI are presented in Figure 5. We observe that the correlation estimates of DAX and CAC40 largely agree over the two models: skew t and normal. However, for the correlation of DAX and NIKKEI and CAC40 and NIKKEI there appear to be some notable differences between the two models. We do not know the true correlations in the empirical studies, but the correlation estimates using the normal model appear to be more variable and less consistent for certain periods of time.

The first 2700 observations of DAX, CAC40 and NIKKEI data set are used for the estimation of the parameters of the RBEKK model with the skew t distribution and the remaining 128 observations (roughly six months data) are left for the out-of-sample forecasts. The forecast returns are obtained using a rolling window of size 2700 and the model is re-estimated each time a new observation vector is obtained (for the last 128 time points). This approach is appealing to practitioners and manages to capture the returns-dynamics better than in studies which the returns are estimated once from historical data, see e.g. Aguilar and West (2000). Figure 6

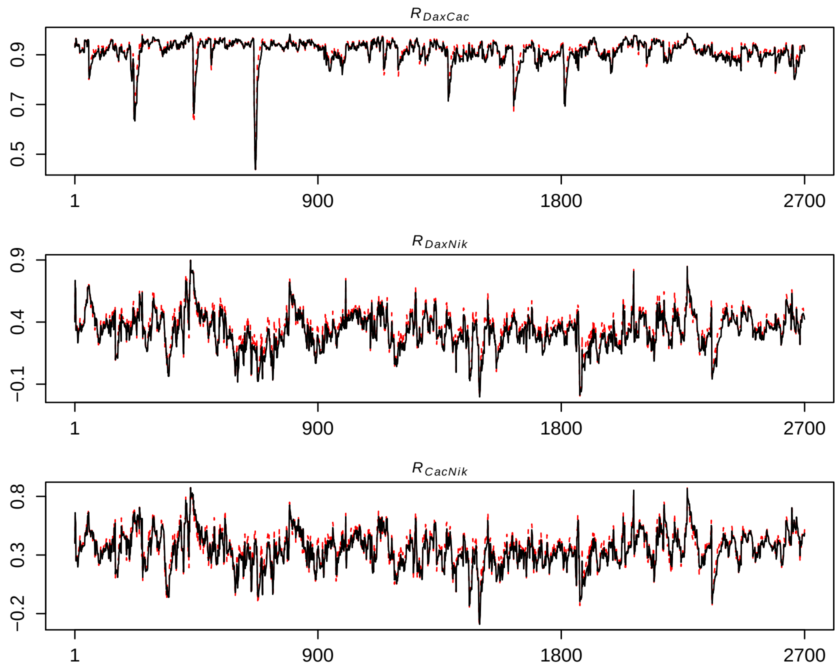

Figure 5: Bayesian estimates of conditional correlations for DAX, CAX40 and NIKKEI index; Skew-t (solid) and Normal (dashed).

shows the log-returns of three stock indices for the out-of-sample period of $t = 2701, \ldots, 2828$.

The optimal portfolio weights for the out-of-sample period are estimated. Figure 7 depicts the dynamics of the estimated portfolio weights and variances along with their corresponding 95% credible intervals. The results of the first two portfolio weights $w^*_{i,T+1}, i = 1, 2$ are reported as the third weight can be obtained from these two.

Finally, to illustrate the proposed MCMC inference on a higher-dimensional data set, we use daily closing prices of eight stocks from the Dow Jones industrial average (DJIA) index. The data set of 2242 observations is obtained from Yahoo!Finance for a sample period from January 02, 2001 to December 31, 2009. Noureldin et al. (2014) used the same data with ten stocks for the maximum likelihood estimation of RBEKK models assuming Gaussian distributions for the returns. In parallel to the earlier analysis, we split the data into two parts; the first 2100 observations are used for the estimation of the parameters of the model while the remaining 142 observation are used for out-of-sample forecasting. The diagonal RBEKK model is estimated using both Bayesian and MLE methods assuming a skew t distribution for the returns. The skew t distribution is chosen as its DIC value was found the lowest among other competing distributions. One-day-ahead forecasts of conditional covariance matrices are obtained and the GMV optimal portfolio without the short-scale constraint is constructed using both methods.

Figure 8 shows the cumulative portfolio returns estimated using the MCMC and MLE methods. For comparison purposes, cumulative portfolio returns using equal weights are also plotted in that figure. We remark that both cumulative returns (under the Bayesian and MLE estimation) outperform the equal weight allocation, which is very volatile and has a poor performance in the beginning and in the end of the sampling period. MCMC provides consistently better cumulative returns in comparison to the MLE, which is struggling to provide positive portfolio returns for about the first 60 trading days. If the assumed model is the true model, then the MLE is expected to at least as good as MCMC. We remark that the MLE algorithm took shorter time to run than MCMC, due to the large number of iterations required for the chains to reach convergence. The difference is more pronounced in higher dimensions. However, it is well known that maximum likelihood estimation for medium to high dimensional data can suffer from local maxima; the simulation study in Section 4 shows that the MCMC has the edge in precision. Moreover, MCMC benefits by its capability of uncertainty analysis around the volatility estimates, the portfolio wights, the cumulative returns and associated expected risk and value-at-risk.

Figure 9 shows Bayesian and maximum likelihood estimates of portfolio variances for eight stocks of DJIA index. We remark that the variance performance of the portfolio between MCMC and MLE is similar (as both methods minimise the variance of the GMV portfolio). However, there are periods of time in which the portfolio variance under the skew t model is systematically smaller than the variance using the normal model, see e.g. $t = 40 - 60$ in

Figure 9.

6 Conclusions

In this article we have proposed Markov chain Monte Carlo estimation for a class of multivariate GARCH models for the estimation of volatility of financial returns. The proposed model benefits from covariance targeting and allows a parsimonious yet flexible treatment of asymmetry and heavy tails of the returns. In the core of the methodology there is a blocked Metropolis algorithm, which is shown to be efficient for small and medium dimensions in a simulation study. Volatility estimation on several assets and a portfolio study demonstrate the capabilities of the methodology. The proposed Bayesian estimation offers parameter uncertainty quantification, for example posterior summaries of volatilities, co-valatilities and dynamic correlation are provided. This is a useful consideration in assessing the uncertainty on portfolio returns, risk management and value-at-risk. As one referee has pointed out the model specification (1) proposed in Noureldin et al. (2014) can be generalised by replacing $\mathbf{H}_t^{1/2}$ by a more general invertible matrix $\mathbf{Z}$. Depending upon the structure of $\mathbf{Z}$, the MCMC inference proposed in this paper may be extended to this interesting case.

The approach of covariance targeting may be applied to other GARCH-type models, such as the dynamic conditional correlation of Engle (2002), although further research should be conducted towards this direction. The combination of covariance targeting and MCMC inference is a promising approach, which we aim to explore in the near future for other multivariate volatility models.

Acknowledgements: The first author would like to thank Higher Education Commission Pakistan for funding this research at the University of Sheffield, UK.

References

[1] Aguilar O, West M (2000) Bayesian dynamic factor models and portfolio allocation. Journal of Business and Economic Statistics 18: 338–357.

[2] Bauwens L, Laurent S (2005) A new class of multivariate skew densities, with application to generalized autoregressive conditional heteroscedasticity models. Journal of Business & Economic Statistics 23: 346–354.

[3] Bauwens L, Laurent S, Rombouts JVK (2006) Multivariate GARCH models: A survey. Journal of Applied Economics 21: 79–109.

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

[5] Boswijk PH, van der Weide R (2011) Method of moments estimation of Go-GARCH models. Journal of Econometrics 163(1): 118–126.

[6] Braione M, Scholtes N (2016) Forecasting value-at-risk under different distributional assumptions. Econometrics, 4: 1–27.

[7] Cecchetti SG, Cumby RE, Figlewski S (1998) Estimation of the optimal futures hedge. Review of Economics and Statistics 70: 623–630.

[8] Chib S, Greenberg E (1995) Understanding the Metropolis-Hastings algorithm. The American Statistician 49: 327–335.

[9] Creal D (2012) A survey of sequential Monte Carlo methods for economics and finance. Econometric Reviews 31: 245–296.

[10] Engle RF (2002) Dynamic conditional correlation: a simple class of multivariate generalized autoregressive conditional heteroskedasticity models. Journal of Business Economics and Statistics 20: 339–350.

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

[12] Fan J, Wang M, Yao Q (2008) Modelling multivariate volatilities via conditionally uncorrelated components. Journal of the Royal Statistical Society: Series B 70: 679–702.

[13] Fernandez C, Steel M (1988) On Bayesian modelling of fat tails and skewness. Journal of the American Statistical Association 93: 359–371.

[14] Fioruci JA, Ehlers RS, Filho, MGA (2014) Bayesian multivariate GARCH models with dynamic correlations and asymmetric error distributions. Journal of Applied Statistics 41: 320–331.

[15] Geweke, J (1992) Evaluating the accuracy of sampling-based approaches to calculating posterior moments. In J.M.Bernardo, J.O.Berger, A.P. Dawid, & A.F.M. Smith (Eds.) Bayesian Statistics (Vol.4). Oxford: Clarendon Press.

[16] Jacquier E, Polson NG (2013) Asset allocation in finance: a Bayesian prespective. In: Demian P, Dellaportas P, Polson NG, Stephen DA (Eds.), Bayesian Theory and Applications, first ed. Oxford University Press, Oxford, pp. 501–516 (Chapter 25).

[17] Jorion P (1986) BayesStein estimation for portfolio analysis. Journal of Financial and Quantitative Analysis 21: 279–292.

[18] Kang L (2011) Asset allocation in a Bayesian copula-GARCH framework: an application to the passive funds versus active funds problem. Journal of Asset Management 12: 45–66.

[19] Neddermeyer, J (2009) Computationally efficient nonparametric importance sampling. Journal of the American Statistical Association 104: 788–802.

[20] Noureldin D, Shephard N, Sheppard K (2014) Multivariate rotated ARCH models. Journal of Econometrics 179: 16–30.

[21] Osiewalski J, Pipien M (2004) Bayesian comparison of bivariate ARCH-type models for the main exchange rates in Poland. Journal of Econometrics 123: 371–391.

[22] Rossi E, Zucca C (2002) Hedging interest rate risk with multivariate GARCH. Applied Financial Economics 12: 241–251.

[23] Silvennoinen A, Teräsvirta T (2009) Multivariate GARCH modes. In T.G. Andersen, R.R. Davis, J.P. Kreiss, and T. Mikosch (Eds.). Handbook of Financial Time Series 201–209. Springer-Verlag.

[24] Spiegelhalter DJ, Best NG, Carlin BP, Lindeven der A (2002) Bayesian measure of model complexity and fir (with discussion). Journal of the Royal Statistical Society: Series B 64: 1–34.

[25] van der Weide R (2002) Go-GARCH: a multivariate generalized orthogonal GARCH model. Journal of Applied Econometrics 17: 549–564.

[26] Virbickaite A, Ausín MC, Galeano P (2015) Bayesian inference methods for univariate and multivariate GARCH models: a survey. Journal of Economic Surveys 29: 76–96.

[27] Vrontos ID, Dellaportas P, Politis DN (2003) A full-factor multivariate GARCH model. Econometrics Journal 6: 312–334.

[28] Yang W, Allen DE (2005) Multivariate GARCH hedge ratios and hedging effectiveness in Australian futures markets. Accounting and Finance 45: 301–321.

[29] Yilmaz T (2011) Improving portfolio optimization by DCC and DECO GARCH: evidence from Istanbul stock exchange. International Journal of Advanced Economics and Business Management 1: 81–92.

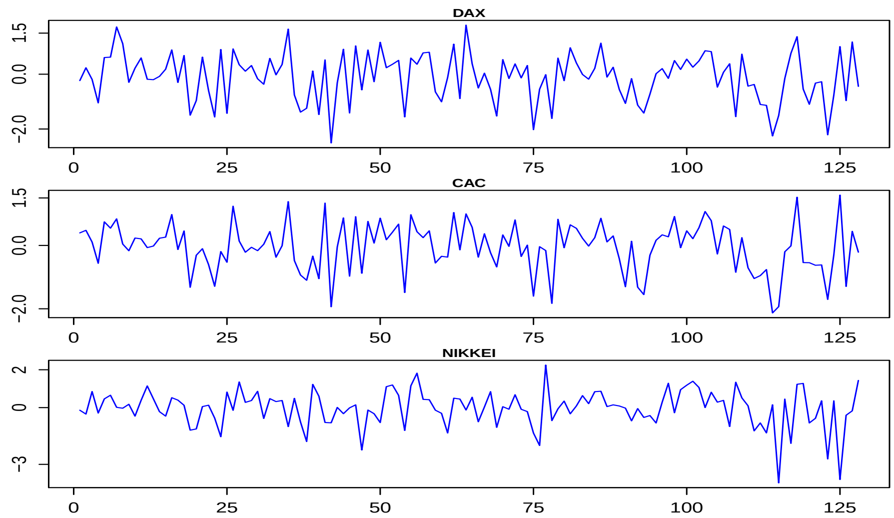

Figure 6: Log-returns (in %) of DAX, CAC40 and NIKKEI index for $t = 2701, \ldots, 2828$.

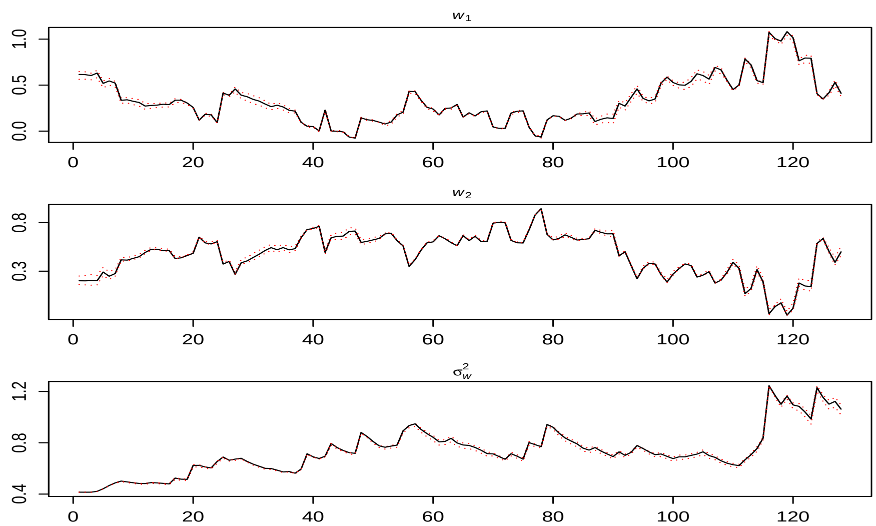

Figure 7: A sequence of portfolio weights and variances along with their corresponding 95% intervals for $t = 2701, \ldots, 2828$.

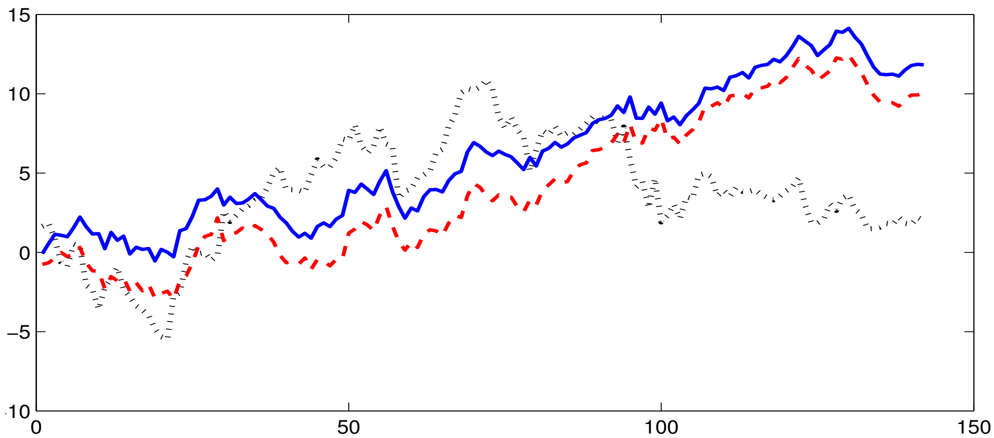

Figure 8: Out-of-sample cumulative portfolio returns for eight stocks of DJIA index; Bayesian (solid) and MLE (dashed) using the GMV portfolio. The equal weighted portfolio returns (dotted) are also displayed.

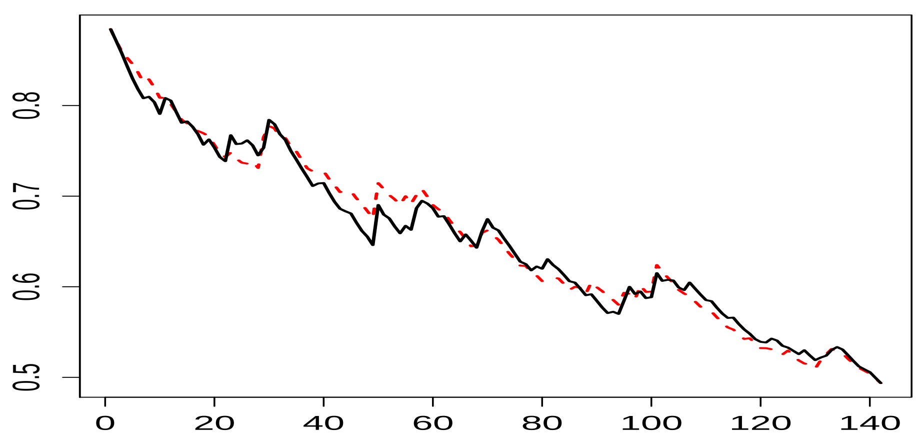

Figure 9: Out-of-sample portfolio variances for eights stocks of DJIA index; Bayesian (solid) and MLE (dashed).