MODELLING MULTIVARIATE VOLATILITIES: AN AD HOC METHOD

MINGJIN WANG AND QIWEI YAO
Department of Statistics, London School of Economics and Political Science
Houghton Street, London, WC2A 2AE, UK
E-mail: m.wang1@lse.ac.uk and q.yao@lse.ac.uk

Department of Business Statistics and Econometrics
Guanghua School of Management, Peking University
100871, Beijing, China

The goal of the paper is two-fold. We first survey the available methods for modelling multivariate volatility processes. We then propose a new and simple method with numerical illustration.

1 Introduction

Volatility plays an important role in controlling and forecasting risks in various financial operations. For a univariate return series, volatility is often represented in terms of conditional variances or conditional standard deviations. Many statistical models have been developed for modelling univariate conditional variance processes. While univariate descriptions are useful and important, problems of risk assessment, asset allocation, hedging in futures markets and options pricing require a multivariate framework, since high volatilities are often observed in the same time periods across different assets. Statistically this boils down to model time-varying conditional variance and covariance matrices of a vector-valued time series. Section 2 below lists some existing statistical models for multivariate volatility processes. We refer to Bauwens, Laurent and Rombouts (2005) for a more detailed survey on this topic. We propose a new and ad hoc method with numerical illustration in section 3. We concludes in section 4 with a brief summary.

2 Existing methods

Let $\mathbf{x}_t = (x_{1,t}, \cdots, x_{d,t})^\tau$ be a $d \times 1$ return series of d assets. Let $\mathcal{F}_t$ be the σ-algebra generated by $\{\mathbf{x}_k, k \leq t\}$, which represents the information set at time t. We assume

$$E(\mathbf{x}_t|\mathcal{F}_{t-1}) = 0, \qquad \mathrm{Var}(\mathbf{x}_t|\mathcal{F}_{t-1}) = \mathbf{\Sigma}_t = \big(\sigma_{ij,t}\big). \tag{1}$$

The goal is to model the conditional variance-covariance matrix $\mathbf{\Sigma}_t$ which is a $d \times d$ non-negative definite matrix. Different models for $\mathbf{\Sigma}_t$ have been proposed over the last two decades. We review some of the models drawing major attraction in the literatures below, and refer to Bauwens, Laurent and Rombouts (2005) for a more extensive survey.

2.1 The BEKK GARCH models

One of the most general forms, proposed by Engle and Kroner (1995), is the BEKK representation of a multivariate GARCH(p, q) process

$$\mathbf{\Sigma}_t = \mathbf{C}_0'\mathbf{C}_0 + \sum_{k=1}^{K}\sum_{i=1}^{q} \mathbf{A}_{ik}\mathbf{x}_{t-i}\mathbf{x}_{t-i}'\mathbf{A}_{ik}' + \sum_{k=1}^{K}\sum_{j=1}^{p} \mathbf{B}_{jk}\mathbf{\Sigma}_{t-j}\mathbf{B}_{jk}', \tag{2}$$

where $\mathbf{C}_0, \mathbf{A}_{ik}, \mathbf{B}_{jk}$ are $d \times d$ matrices and $\mathbf{C}_0$ is upper triangular.

Although the form of the above model is quite general especially when K is reasonably large, it suffers from the problems due to overparametrization. See Engle and Kroner (1995) for more discuss on the identification problem of this model.

Similar to univariate GARCH models, the standard estimation method for the BEKK model (2) is the quasi-maximum likelihood estimation (qMLE) facilitated by assuming $\mathbf{x}_t|\mathcal{F}_{t-1} \sim N(\mathbf{0}, \mathbf{\Sigma_t})$. The consistency and the asymptotic normality of the qMLE have been established by Comte and Lieberman (2003). Note that even for moderately large d, the qMLE is a solution of a high-dimensional nonlinear optimization problem. Therefore in practice some approximate and iterative estimation methods are often more efficient.

2.2 Factor and orthogonal models

In order to reduce the number of parameters in modelling multivariate volatilities, different types of decompositions for $\mathbf{\Sigma}_t$ are often employed in model-specifications. For instance, writing $\mathbf{\Sigma}_t$ as the sum of a time-varying part (usually with reduced rank) and a homoscedastic part, Engle, Ng and Rothschild (1990) proposed a factor multivariate GARCH model as follows:

$$\mathbf{\Sigma}_t = \mathbf{\Omega} + \sum_{k=1}^{K} \mathbf{g}_k\mathbf{g}_k'(\sum_{i=1}^{q} \alpha_{ik}\mathbf{f}_k'\mathbf{x}_{t-i}\mathbf{x}_{t-i}'\mathbf{f}_k + \sum_{j=1}^{p} \beta_{jk}\mathbf{f}_k'\mathbf{\Sigma}_{t-j}\mathbf{f}_k). \tag{3}$$

where α_{ik}, β_{jk} are non-negative constants, $\mathbf{\Omega}$ is a time-invariant non-negative definite constant matrix, and $\mathbf{g}_k, \mathbf{f}_k$ are $d \times 1$ constant vectors satisfying the constraints $\sum_{\ell=1}^{d} f_{k\ell} = 1$ for $k = 1, \cdots, K$ and $\mathbf{f}_k'\mathbf{g}_i = 0$ for $k \neq i$, and 1 for $k = i$.

Model (3) is called a Factor-GARCH($p, q; k$) model. The K linear combinations $\theta_{k,t} = \mathbf{f}_k'\mathbf{x}_t,\ k = 1, 2, \cdots, K$, represent K common factors of which the conditional variances are specified as K different univariate GARCH(p, q) models. It is easy to see that (3) is a special case of the BEKK model if we put $\mathbf{A}_{ik} = \sqrt{\alpha_{ik}}\mathbf{f}_k\mathbf{g}_k'$ and $\mathbf{B}_{jk} = \sqrt{\beta_{jk}}\mathbf{f}_k\mathbf{g}_k'$.

Factor-GARCH models can be estimated with qMLE method. In practice, the factor representation portfolios $\theta_{k,t}$ are usually set as known and a two-stage estimation scheme can then be invoked. See Lin (1992) for other estimation procedures.

Orthogonal GARCH (hereafter "O-GARCH") model, which is based on the principal components of $\mathbf{x}_t$, can virtually be viewed as a special case of Factor-GARCH model (Alexander and Chibumba 1997). Let the unconditional covariance matrix of $\mathbf{x}_t$ be $\mathbf{\Sigma}$. Based on the eigen-decomposition $\mathbf{\Sigma} = \mathbf{W}\mathbf{\Lambda}\mathbf{W}'$, where $\mathbf{W}'\mathbf{W} = \mathbf{I}_d$ and $\mathbf{\Lambda}$ is a diagonal matrix. O-GARCH specification first fits the conditional variance of each principal component $\boldsymbol{\zeta}_t \equiv (\zeta_{1,t}, \cdots, \zeta_{d,t})' = \mathbf{W}'\mathbf{x}_t$ with a univariate GARCH model:

$$\zeta_{i,t}|\mathcal{F}_{t-1} \sim N(0, \lambda_{i,t}),$$
$$\lambda_{i,t} = \omega_i + \sum_{u=1}^{q_i} \alpha_{iu}\zeta_{i,t-u}^2 + \sum_{v=1}^{p_i} \beta_{iv}\lambda_{i,t-v},$$

and then take $\mathbf{\Sigma}_t = \mathbf{W}\mathbf{\Lambda}_t\mathbf{W}'$ as the conditional variance matrix of $\mathbf{x}_t$, where $\mathbf{\Lambda}_t = \mathrm{diag}(\lambda_{1,t}, \cdots, \lambda_{d,t})$. This effectively assumes that the principal components are also *conditionally* uncorrelated.

Obviously O-GARCH model is easy to fit in practice even when d is large or very large. However, it treats unconditionally uncorrelated principal components as conditionally correlated as well, which is typically untrue. This may lead to nonsensical or even wrong results. See Fan, Wang and Yao (2004).

Recently, Fan, Wang and Yao (2004) proposed to model multivariate volatilities in terms of a decomposition based on the so-called conditionally uncorrelated components (CUC) of $\mathbf{x}_t$. It overcomes the aforementioned shortcoming of the O-GARCH models.

2.3 Conditional correlation models

It always holds that

$$\mathbf{\Sigma}_t = \mathbf{D}_t\mathbf{\Gamma}_t\mathbf{D}_t, \tag{4}$$

where $\mathbf{D}_t = \mathrm{diag}(\sqrt{\sigma_{11,t}}, \cdots, \sqrt{\sigma_{dd,t}})$, and $\mathbf{\Gamma}_t$ is the conditional correlation matrix $\mathbf{x}_t$.

Assuming that $\mathbf{\Gamma}_t$ does not change over time t and modelling each $x_{j,t}$ with a univariate GARCH model, Bollerslev (1990) proposed a constant conditional correlation (CCC) framework which simplified the estimation and inference procedures substantially. However, it is questionable if the time-invariant conditional correlation is a realistic assumption in practice.

The dynamic conditional correlation (DCC) model of Engle (2002) computes the time changing conditional correlation matrix from the standardized residuals series

$$\mathbf{\Gamma}_t = \text{diag}\{\mathbf{Q}_t\}^{-1/2}\mathbf{Q}_t\text{diag}\{\mathbf{Q}_t\}^{-1/2}, \tag{5}$$

where

$$\mathbf{Q}_t = \mathbf{S}(1-\theta_1-\theta_2) + \theta_1(\boldsymbol{\xi}_{t-1}\boldsymbol{\xi}'_{t-1}) + \theta_2\mathbf{Q}_{t-1}, \tag{6}$$

and $\xi_{k,t}$ are the standardized residuals obtained from the raw residuals $x_{k,t}/\{\hat{\sigma}_{kk,t}\}^{1/2}$, and $\mathbf{S}$ is the sample covariance matrix of $\{\boldsymbol{\xi}_t\}_{t=1}^n$.

A slightly different formulation was suggested by Tse and Tsui (2002):

$$\mathbf{\Gamma}_t = \mathbf{\Gamma}(1-\theta_1-\theta_2) + \theta_1\mathbf{\Gamma}_{t-1} + \theta_2\mathbf{\Psi}_{t-1} \tag{7}$$

to fit the correlation process. Here $\mathbf{\Gamma} = \{\rho_{ij}\}$ is a time-invariant $d \times d$ positive definite parametric matrix with unit diagonal elements and $\mathbf{\Psi}_{t-1}$ is, for example, the sample correlation matrix of $\{\boldsymbol{\xi}_t\}_{t-M}^{t-1}$. This specification is called varying correlation multivariate GARCH model or simply VC-MGARCH model.

Although qMLE method is available in principle for all these conditional correlation models, some two-stage estimation schemes have been developed to increase the computational efficiency, and have apparently been used more often in practice.

3 A new ad hoc method

3.1 Method

Note in (1), $\sigma_{ii,t} = \text{Var}(x_{i,t}|\mathcal{F}_{t-1})$. We may model $\sigma_{i,t}^2 \equiv \sigma_{ii,t}$ using any appropriate univariate volatility models based on univariate time series $\{x_{ti}\}$.

To model the off-diagonal elements $\sigma_{ij,t}$ with $i < j$, put

$$y_{ij,t} = (x_{i,t} + x_{j,t})/2. \tag{8}$$

We may model its conditional variance $\omega_{ij,t} = \text{Var}(y_{ij,t}|\mathcal{F}_{t-1})$ again by a simple univariate model. Note that for $1 \le i < j \le d$,

$$\sigma_{ij,t} = 2\omega_{ij,t} - \frac{\sigma_{i,t}^2 + \sigma_{j,t}^2}{2}. \tag{9}$$

Hence once we have derived univariate volatility models for each component $x_{i,t}$ and the combined series $y_{ij,t}$, the models for the conditional covariances is implied by (9) above.

In practice, we may use simple GARCH(1,1) models for modelling both $\sigma_{i,t}^2$ and $\omega_{ij,t}$, namely

$$\sigma_{i,t}^2 = \alpha_i + \beta_i x_{i,t-1}^2 + \gamma_i \sigma_{i,t-1}^2, \tag{10}$$

$$\omega_{ij,t} = \alpha_{ij} + \beta_{ij} y_{ij,t-1}^2 + \gamma_{ij} \omega_{ij,t-1}. \tag{11}$$

It is clear that the above proposal overcomes the difficulties due to over-parametrization, and can be implemented in a computationally efficient manner since all the components of $\mathbf{\Sigma}_t$ are practically fitted separately. Furthermore, we have the flexibility in choosing appropriate univariate models for $\sigma_{i,t}^2$ and $\omega_{ij,t}$, which may be GARCH, stochastic volatility models, semiparametric or nonparametric volatililty models, or some empirical methods such as rolling exponential smoothing. However the simplicity in both the structure and the feasibility does come with a price unfortunately. First the implied estimator for the conditional variance $\mathbf{\Sigma}_t$ may not necessarily be a non-negative definite matrix. (A quick remedy may be to shrink the negative eigenvalues of the estimated $\mathbf{\Sigma}_t$ to 0). Furthermore, the approach suffers from a kind of innate inconsistency in model specification. For example, under the GARCH(1,1) specification of (10) and (11), the conditional variance of a portfolio $\mathbf{a}'\mathbf{x}_t$ is not necessarily GARCH(1,1). Also note that we may define $y_{ij,t}$ differently from the form (8), still $\sigma_{ij,t}$ may be uniquely determined by $\sigma_{i,t}^2, \sigma_{j,t}^2$ and $\omega_{ij,t}$. However the estimator for $\sigma_{ij,t}$ implied may be different.

3.2 Numerical illustration

We illustrate the new method with two real data sets. The first one consists of the daily log returns (in percentages) of two exchange rate series, namely, the Deutsche mark (D) and the Japanese yen (J) versus U.S. dollar. It covers the period of 3 January 1990 — 23 June 1998, for a total of 2131 observations. The data was downloaded from the website of the Federal Reserve Bank of New York and has been analyzed by Tse and Tsui (2002) using the VC-MGARCH model. See Figures 1(a) & (b) for the time series plots of these two series. The second data set contains the four indices from Asian stock markets, i.e. the Hang Seng index of Hong Kong (HS), the Japan Nikkei 225 index (JN), the Shanghai Composite index (SH) and the Taiwan Weighted index (TW). Daily close prices adjusted for dividends and splits are obtained directly from the website of Yahoo!Finance. We applied log-difference transformation to

Table 1. **Summary Statistics of the Two Data Sets**

	D	J	HS	JN	SH	TW
Mean	0.0025	-0.0023	-0.0193	-0.0388	0.0101	-0.0411
Stdev	0.6746	0.6750	2.0974	1.6868	1.5109	1.9432
Min	-2.8963	-4.5228	-14.7346	-9.0145	-8.7277	-9.9360
Max	3.1030	3.2269	20.2083	8.8876	8.8491	9.7871
Skewness	0.0197	-0.5065	0.6226	0.0107	0.1881	-0.0199
Kurtosis	4.7731	6.5508	14.9947	5.2678	8.2629	5.2284
J-B	279.29	1210.62	9131.42	322.97	1748.09	311.91
$Q_1(10)$	13.9104	16.0221	14.8263	6.5265	8.2794	20.9968
$Q_1(20)$	21.8039	27.4127	30.1055	10.7677	16.9162	38.4476
$Q_2(10)$	287.0760	83.8406	226.8334	74.5163	131.2148	88.4256
$Q_2(20)$	460.5919	111.9082	242.1270	100.1453	192.3026	103.3381

Note: J-B stands for the Jarque-Bera statistics. $Q_1(k)$ *and* $Q_2(k)$ *represents the Ljung-Box portmanteau statistics of the original and squared return series, respectively.*

convert them into continuously compounded returns. Adjustment was also made to account for the effect of different holidays of these four markets. The data consist of 1507 observations covering the period of 1 August 1997 — 31 July 2004. The time series plots for the second data set are omitted to save the space.

Descriptive statistics for all the six series are reported in Table 1. All the series are leptokurtic and the nulls of normal distribution can be rejected based on the Jarque-Bera test for all series. The Ljung-Box portmanteau statistics of the two exchange rates series suggest that there exists no significant evidence for the autocorrelation structures in both the series. We extract the mean values from these two series and focus our attention to their covariance matrix modelling. For the Asian market data, the portmanteau statistics reveal some autocorrelation structure in HS and TW series. Accordingly we fit an AR(5) model for each of these four series first. The analysis reported below was conducted with the filtered series.

For the first data set, a univariate GARCH(1,1) model is fitted to D and J, respectively, using qMLE method subject to the "variance targeting" constraint in the sense that the long run variance is just the sample variance (see Engle (2002)). In order to obtain an estimator for conditional correlation between D and J, another univariate GARCH(1,1) model is fitted to (D+J)/2 using the same method. Table 2 presents the estimated parameters. Standard errors are omitted to save space. See Figure 1(c) and (d) for the fitted

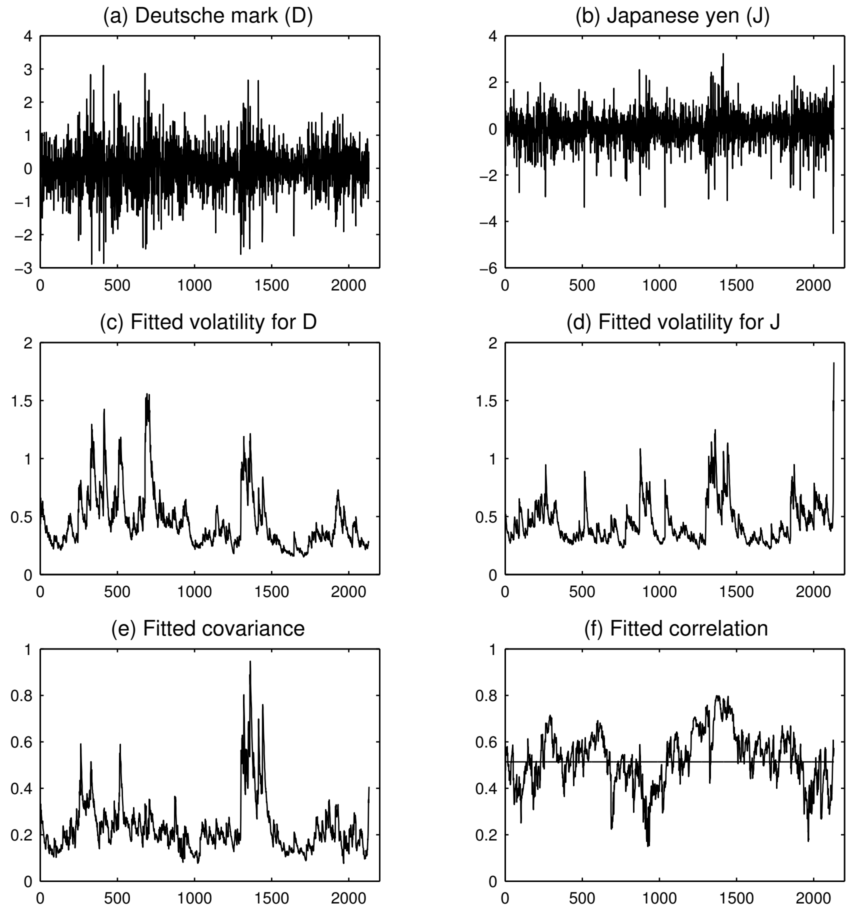

Figure 1. *Time series plot of daily log-return (in percentage) of (a) Deutsche mark (D) and (b) Japanese yen (J) versus US dollar; the fitted volatility of the return series of (c) Deutsche mark and (d)Japanese yen; and the fitted (e) conditional covariance and (f) conditional correlation between D and J using the ad hoc method.*

volatility for these two return series and (e) and (f) for the fitted covariance and correlation between D and J, respectively. A horizontal line in Figure 1(f) is drawn to show the level of unconditional correlation between these two series.

It is interesting to compare the fitted conditional correlation in Figure 1(f)

Table 2. **Estimation and Diagnostic Checking Results**

	α	β	γ	$Q(10)$	$Q(20)$
D	0.0061	0.0509	0.9357	11.5085	15.9470
J	0.0108	0.0439	0.9325	3.6654	10.7579
(D,J)	0.0066	0.0432	0.9376	15.8036	23.5034
HS	0.0731	0.0946	0.8887	7.9466	14.3098
JN	0.2203	0.0777	0.8447	5.0772	9.1901
SH	0.1056	0.1322	0.8214	7.3288	20.1829
TW	0.3347	0.0862	0.8242	5.8506	15.4980
(HS,JN)	0.1228	0.0860	0.8674	5.0718	16.5563
(HS,SH)	0.0562	0.0956	0.8737	12.4900	23.6990
(HS,TW)	0.1952	0.0902	0.8385	7.0120	13.5878
(JN,SH)	0.0916	0.0673	0.8655	9.1555	12.9979
(JN,TW)	0.2163	0.0559	0.8423	2.1990	4.5946
(SH,TW)	0.1785	0.1015	0.7867	7.4842	19.4355

with those in Figure 4 of Tse and Tsui (2002). The later was obtained using BEKK model and VC-MGARCH model, both of them needed an intensive searching method to maximize the corresponding likelihood functions. The magnitudes and the time-varying patterns in these two figures are very similar. This suggests that our ad hoc method is as capable as those more sophisticated models in representing dynamic correlation structure at least for this data set. Furthermore, the fitted conditional correlation process always stays between -1 and 1. Hence the corresponding conditional covariance is automatically a non-negative definite matrix.

To further check the possible misspecification of the fitted model, we use the Ljung-Box Q portmanteau statistics of the cross-product of the standardized error series.

More specifically, we use the $Q(k)$ statistics of $\hat{u}_{i,t}^2 - 1, t = 1, 2, \cdots, T$ to check adequacy of the volatility model for the i-th series and the $Q(k)$ statistics of $\hat{u}_{i,t}\hat{u}_{j,t} - \hat{\rho}_{ij,t}, t = 1, 2, \cdots, T$ to check the correlation modelling between i-th and j-th series, where $\hat{u}_{i,t} = x_{i,t}/\hat{\sigma}_{i,t}$ is the standardized residuals. χ_k^2 is selected as a null reference distribution[a]. The two columns on the right in Table 2 list the values of $Q(10)$ and $Q(20)$. Apparently, at any conventional level of significance, there is no evidence to indicate the remnant autocorrelation structure in the residuals. This confirms quantitatively that

[a] Although there is no rigorous theory for such a test so far, the simulation study in Tse and Tsui (1999) suggests that it indeed provides a reasonable test with good size and power.

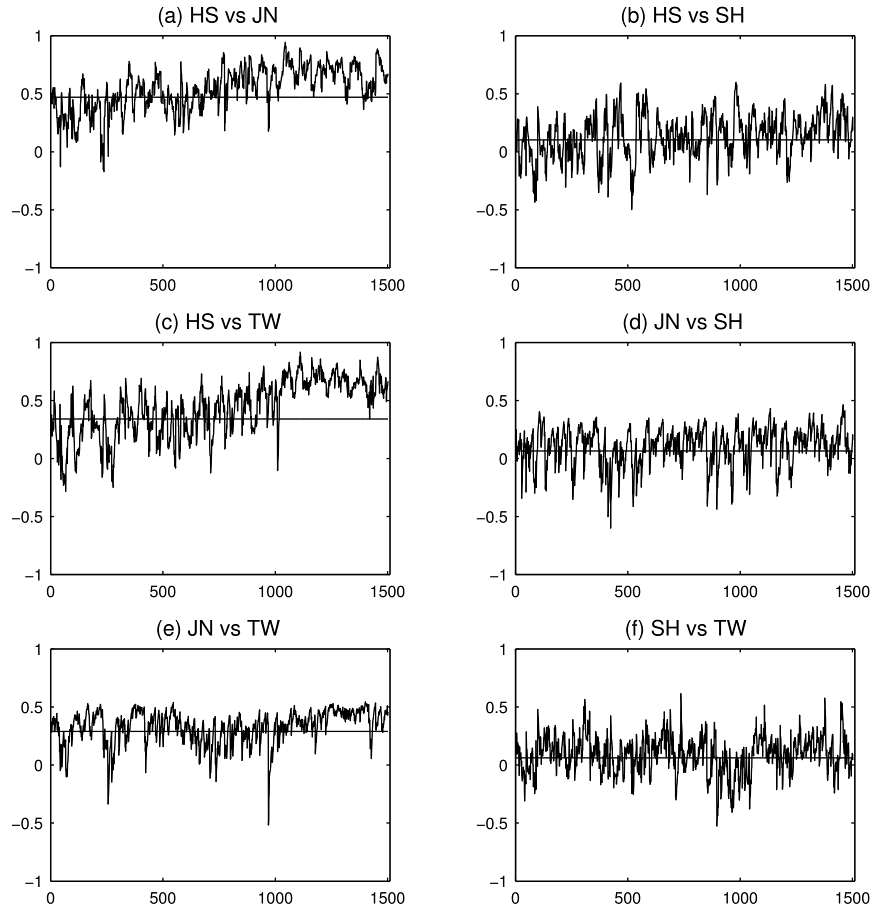

Figure 2. *The fitted conditional correlations between (a) HS and JN, (b) HS and SH, (c) HS and Tw, (d) JN and SH, (e) JN and TW, (f) SH and TW for the Asian Stock Market data using the ad hoc method.*

the new method works well for such a data set.

The intractability of estimating a high-dimensional volatility model is a notorious fact in modelling multivariate volatility processes. For instance, for a four-dimension BEKK model it requires to solve an optimization problem with at least 42(=10+16+16) parameters. However, our ad hoc method can handle such a situation in a pretty easy manner. As a matter of fact, we need to estimate 10 univariate GARCH(1,1) models only, and the six conditional covariance can be derived according to (9). Table 2 lists the estimated coefficients for the Asian market data set. The six fitted conditional correlation series are plotted in Figure 2, where the horizontal line in each panel is the corresponding unconditional correlation. Furthermore, the values of the Q portmanteau statistic in the two very-right columns of Table 2 suggest the adequacy of the fitting. Note the (global) unconditional correlations are pretty close to 0 in Figures 2(b), (d) and (f), it seems reasonable to observe some negative conditional correlations in those plots. Note that the estimated conditional variances are not guaranteed to be non-negative definite. We calculate the eigenvalues for each fitted covariance matrix and the negative values only occur at the smallest eigenvalues of 23 points over the whole 1502(=1507-5) observations.

4 Conclusion

After reviewing some of the major multivariate volatility models, we put forward a new ad hoc method to model the conditional covariance process. Numerical results based on two real data sets suggest that a practically meaningful fitting may be obtained in a computationally efficient manner from applying the proposed new method. Therefore it might be worthwhile to investigate the theoretically properties of this ad hoc method more thoroughly.

Acknowledgements

The research was partially supported by an EPSRC research grant. Mingjin Wang was also supported by the Natural Science Foundation of China (grant 70201007).

References

1. Alexander, C., and Chibumba, A. (1997). Multivariate orthogonal factor GARCH. University of Sussex, Mimeo.
2. Bauwens, L., Laurent, S. and Rombouts, J.V.K. (2005). Multivariate GARCH models: a survey. *Journal of Applied Econometrics*, forthcoming.
3. Bollerslev, T. (1990). Modelling the coherence in short-run nominal exchange rates: a multivariate generalized ARCH model. *Reveiw of Economics and Statistics*, **72**, 498-505.

4. Comte, F. and Lieberman, O. (2003). Asymptotic theory for multivariate GARCH processes. *Journal of Multivariate Analysis*. **84**, 61-84.
5. Engle, R. F., Ng, V.K. and Rothschild, M. (1990). Asset pricing with a factor-ARCH covariance structure: empirical estimates from treasury bills. *Journal of Econometrics*, **45**, 213-238.
6. Engle, R.F. (2002). Dynamic conditional correlation — a simple class of multivariate GARCH models. *Journal of Business of Economic Statistics*, **20**, 339-350.
7. Engle, R.F. and Kroner, K.F. (1995). Multivariave simultaneous generalised ARCH. *Econometric Theory*, **11**, 122-150.
8. Fan, J., Wang, M. and Yao, Q. (2004). Modelling multivariate volatilities via conditional uncorrelated components. *Under preparation*.
9. Lin, W. (1992). Alternative estimators for factor GARCH models - a Monte Carlo comparison. *Journal of Applied Econometrics*, **7**, 259-279.
10. Magnus, J.R. and Neudecker, H. (1999). *Matrix Differential Calculus with Applications in Statistics and Econometrics*. (Revised edition) Wiley, New York.
11. Tse, Y. K. and Tsui, A. (1999). A note on diagnosing multivariate conditional heteroscedasticity models. *Journal of Time Series Analysis*, **20**, 679-691.
12. Tse, Y. K. and Tsui, A. (2002). A multivariate GARCH modle with time-varying correlations. *Journal of Business and Economic Statistics*, **20**, 351-362.