IEEE Second International Conference on Data Stream Mining & Processing
August 21-25, 2018, Lviv, Ukraine

The Methods Bayesian Analysis of the Threshold Stochastic Volatility Model

Peter Bidyuk
National Technical University of Ukraine "Ighor Sikorsky Kyiv Polytechnic Institute"
Kiev, Ukraine
pbidyke@gmail.com

Aleksandr Gozhyj
Petro Mohyla Black Sea National University
Nikolaev, Ukraine
alex.gozhyj@gmail.com

Zdislaw Szymanski
Splolecna Akademia Nauk
Warsaw, Poland
zszymanski@san.edu.pl

Irina Kalinina
Petro Mohyla Black Sea National University
Nikolaev, Ukraine
irina.kalinina1612@gmail.com

Volodymyr Beglytsia
Petro Mohyla Black Sea National University
Nikolaev, Ukraine
science@chmnu.edu.ua

***Abstract*—The paper considers the Bayesian analysis of the threshold stochastic volatility models. Studies of methods for analyzing stochastic volatility and improving models of stochastic volatility significantly improve the quality of forecast models and their estimates. Bayesian inference is performed by tailoring Markov chain Monte Carlo (MCMC) or sequential Monte Carlo (SMC) schemes that take into account the specific characteristics of models. The results of applying the method demonstrated in models heteroscedastic non-stationary processes.**

Keywords—Bayesian analysis; stochastic volatility; Threshold stohastic volatility model

I. INTRODUCTION

With the current economic instability, trading on stock markets carries a high risk. Therefore, the study of volatility, a statistical measure of stock prices, becomes instrumental and often indispensable. Currently, the study of volatility has become the basis of the financial economics, and one of the main tools of financial analysis and modeling, in particular. The ground of these studies lies in various probabilistic and statistical volatility models. Statistical volatility models are widely used in various financial tasks, such as the estimation of the standard deviation of the market returns, risk assessment, evaluation of the financial instruments, etc. There are dozens of volatility measurement methods, ranging from technical indicators such as the average true range (ATR), historical volatility, stochastic volatility of various types, standard deviation, etc. In addition to the financial analysis, conditional variance models are widely used in medical and technical diagnostic systems, risk assessment and management, social studies, etc. The study of stochastic volatility analysis methods and the improvement of the models` structure can substantially improve the quality of their forecasting and estimates. Therefore, this research is aimed at the investigation of the method of Bayesian analysis of the threshold stochastic volatility model.

II. VOLATILITY MODELING WAYS

The stochastic volatility model is based on the conditional heteroscedastic model. The conditional heteroscedasticity model (ARCH) was developed by Robert F. Engle [1] to create a model of inflation in the UK. This model was later used for stock prices and exchange rates modeling [2]. The further development of ARCH is the generalized autoregressive conditional heteroscedasticity (GARCH) described in the investigation [3,4], which is still actively used for volatility forecasting [5, 6]. Models like GARCH allow recreating the phenomenon of volatility clustering (GARCH-effect). Parameters of ARCH / GARCH models are most often evaluated by the maximum likelihood estimation. One of the main disadvantages of the GARCH model is that the model memory is "not long enough" since its autocorrelation function (ACF) is characterized by the exponential decay. When the sum of coefficients of the model $\alpha + \beta$ is close to 1, the GARCH model degenerates into a non-stationary process, named an integrated generalized autoregressive conditionally heteroscedastic (IGARCH) process [2]. However, the latter model implies the dependence of volatility on the initial conditions which does not disappear within an infinite planning horizon, and therefore cannot be claimed to be an adequate reflection of the reality. An alternative approach is to use stochastic processes or models whose theoretical properties imply the presence of the so-called "long" memory.

III. MODELS OF STOCHASTIC VOLATILITY

The main idea behind the stochastic volatility models is to increase the number of randomness sources. In conditional heteroscedasticity models, there is only one source of randomness, and the variation of the process is assumed to be dependent on its previous realizations in one form or another. An alternative way of modeling is to provide the price dynamics in the form of a simple model, like a differential equation. However, the volatility σ in this equation is rather a separate stochastic process, than a parameter. Consequently, there are two independent sources of randomness. The first stochastic volatility model was suggested by [11]. In particular, it assumed that logarithmic volatility is a process AR (1):

$$\begin{aligned} r_t &= \mu\sigma_t\varepsilon_t \\ \ln\sigma_t^2 &= \varphi\ln\sigma_{t-1}^2 + \nu_t \end{aligned} \qquad (1)$$

where μ is a positive constant, the inclusion of which allows to remove the free term from the equation for the sake of volatility, and ϕ is an autoregressive parameter that determines the memory in its relation to volatility. The main properties of autoregressive stochastic volatility models (ARSV) are investigated in [12,13].

The stochastic volatility model can be represented as follows:

$$x(k) = \psi_0 + \psi_1 x(k-1) + y(k) \tag{2}$$

$$y(k) = \sqrt{h(k)u(k)},\ u(k) \approx N(0,1) \tag{3}$$

$$\log h(k+1) = \alpha + \phi \log h(k) + \eta(k),\ \ \eta(k) \approx N(0,\sigma^2) \tag{4}$$

where $x(k)$ is the statistical time series on which the model is based; $u(k)$ and $\eta(k)$ are independent of white noise stochastic processes. In some formulations, it is assumed that ψ_0 and ψ_1 are zero. The AR process (1) with the innovation $y(k)$ in the time series, as determined by the equation (2), explains the possibility of autoregression in the process $x(k)$.

Threshold stochastic volatility model. Dispersion of the incomes tends to increase with the decrease of share prices. Such dispersion behavior can be described using a constant correlation coefficient ρ between $u(k)$ and $\eta(k)$, keeping all other assumptions unchanged. In the initial model defined by the equations (2) – (4), it is zero. Numerous empirical studies have shown that the coefficients ρ are negative in the assumption that negative income is associated with a positive dispersion.

This investigation suggests a new approach to fixing the time series dispersion asymmetry. Since it has been established that the dispersion tends to grow under the influence of bad news (disappointing global forecasts), then it is likely that the dynamics of autoregression in the equation (4) is determined by the income in a previous period of time. There is a hypothesis that the amount of income is dependent on the prior income (income sign). This kind of income asymmetry can also be taken into account, summing up the equation (2) to the piece-linear structure. Thus, it will be more natural to consider the threshold nonlinear structures than the linear autoregressive processes represented by the equations (2) – (4).

Let`s define a set of Bernoulli random variables as follows:

$$s(k) = \begin{cases} 0, & if \quad x(k-1) < 0, \\ 1, & if \quad x(k-2) \geq 0. \end{cases} \tag{5}$$

The threshold stochastic volatility model takes the following form:

$$x(k) = \psi_{0s(k)} + \psi_{1s(k)} x(k-1) + y(k) \tag{6}$$

$$y(k) = \sqrt{h(k)\,u(k)}, \qquad u(k) \sim N(0,1) \tag{7}$$

$$\log h(k+1) = \alpha_{s_{k+1}} + \phi_{s_{k+1}} \log h(k) + \eta(k), \quad \eta(k) \sim N(0,\sigma^2) \tag{8}$$

Like in the initial formulation, $u(k)$ and $\eta(k)$ are stochastically independent. At a time $k-1$, when there is an unexpected fall in prices due to the disappointing news, $x(k-1) < 0$ and $s(k) = 0$. On the contrary, if there is good news at the time $k-1$, then $x(k-1) > 0$ and $s(k) = 1$. Therefore, the value $s(k)$ is determined by a magnitude sign $x(k-1)$. In the threshold stochastic volatility model, the values of the parameters ψ_0, ψ_1, α and ϕ switch between these two modes, which corresponds to the asset prices increases and falls.

In the symmetric case, the two sets of parameters are identical. In particular, if $\phi_0 = \phi_1$, taking into account that $\alpha_0 \geq \alpha_1$, it follows that the dispersion will be higher when the prior income is negative, than when it is positive.

The generalized model ϕ_0 may differ from ϕ_1. Indeed, the coefficient ϕ_{s_t} is describing the magnitude of the impact of prior income on the current dispersion. If ϕ_0 is more than ϕ_1, then the dispersion in previous periods will have a greater impact on the current dispersion after the price fall then after its rise. It is expected that in a similar hypothetical market situation, more time will be needed to "handle" the negative information that is contained in the previous dispersion data. This kind of asymmetry has not yet been sufficiently described in the researches related to the stochastic volatility analysis.

IV. Bayesian Analysis Of The Threshold Stochastic Volatility Model

In the standard Bayesian conclusion, marginal posterior distributions of unknown parameters are used. However, in many cases, the common posterior distribution or even marginal posterior distribution do not have closed-form solutions. It is also quite difficult to obtain the model values from the desirable posterior distribution.

Monte Carlo method-based approaches for Markov chains are Markov update algorithms aimed at obtaining a sample from the common posterior distribution. A separate case is Gibbs sampling. This method based on the Monte Carlo procedure is close to the approach based on the generation and reproduction of the data samples.

Let`s make a sample selection of the unified distribution $F(\omega_1,...,\omega_m)$ where $\omega = (\omega_1,...\omega_m)^T$ is a vector of unknown parameters or hidden variables. With the known initial values $[\omega_1^{(0)},...,\omega_m^{(0)}]$, the algorithm gives an estimate of the value $\omega_1^{(i)}$ with $F(\omega_1 \mid \omega_2^{(i-1)},...,\omega_m^{(i-1)})$, value $\omega_2^{(i)}$ with $F(\omega_2 \mid \omega_1^{(i)}, \omega_3^{(i-1)}...,\omega_m^{(i-1)})$, and value $\omega_m^{(i)}$ with $F(\omega_m \mid \omega_1^{(i-1)},..., \omega_{m-1}^{(i-1)})$, considering that $i = 1,...,M+N$. Under non-rigid conditions, the vector of parameters

$[\omega_1^{(i)},...,\omega_m^{(i)}]$ coincides with the distribution to the shared distribution $F(\omega_1,...,\omega_m)$ with $i \to \infty$.

Usually, the first M of the transitive iterations are skipped, and the last N of the iterations are taken for an approximate sample selection, dependent on $F(\omega_1,...,\omega_m)$. The density of the probability distribution can be performed in two different ways.

The first approach is the traditional assessment of core density distribution. The second approach is shown by the formula:

$$f(\omega_j) = \frac{1}{N}\sum_{i=M+1}^{M+N} f(\omega_j \mid \omega_{-j}^{(i)}) \tag{9}$$

where ω_{-j} means all the parameters except ω_j. The point estimates of any function ω, e.g. $g(\omega)$, can also be found using the Gibbs variable selection. One of the commonly used approaches is to use the posterior average, i.e.

$$\frac{1}{N}\sum_{i=M+1}^{M+N} g(\omega^{(i)}) \tag{10}$$

The Gibbs sampler (sampling) can be used if it is possible to make a model with all conditional distributions. Gibbs specification and other models based on the Markov chain Monte Carlo (MCMC) methods are investigated in [15,16].

Let`s make a sample selection and calculate the distribution parameters following the given MCMC algorithm:

(1) Calculate $\rho(k)$, $k = 1,...,n$.

(2) Calculate σ^2 following the Metropolis step model (random samples).

(3) Calculate $h(k), k = 1,...,n$, using multi-stage discretization.

(4) Calculate $(\psi_0, \delta, \psi_1, c)^T$ using its multidimensional normal complete conditional distribution.

(5) Calculate x_0 using its normal full conditional distribution.

(6) Calculate $(\alpha, \gamma, \phi, d)^T$ using its multidimensional normal complete conditional distribution.

The completion of the algorithm iteration using the MCMC method.

Dispersion smoothing and forecasting. To implement the procedure of the dispersion smoothing and forecasting, the Gibbs sample selection is used. After performing the iteration required for Gibbs sampling, one can get the approximation (approximate sample) from the common posterior distribution $f(x_0, H_n, \theta \mid X_n)$, marked as $(x_0^{(i)}, h_1^{(i)},....,h_n^{(i)}, \theta^{(i)}), i = M+1,...,M+N$. Smoothed estimates $h(k),(k = 1,..,n)$ are the h_t estimates calculated from the marginal posterior distribution $f(h(k) \mid X_n)$. The natural choice is the marginal posterior expectation, $E(h(k) \mid X_n)$ which can be estimated as a sample mean:

$$\frac{1}{N}\sum_{i=M+1}^{M+N} h(k)^{(i)} \tag{11}$$

To predict the future dispersion, using the currently available information, it is necessary to generate the samples from $f(h_{n+j} \mid X_n)$ with $j > 0$. This can be effectively done using the composition method. Thus, when j = 1 it can be written:

$$\begin{aligned} &f(h_{n+1} \mid X_n) = \\ &= \int f(h_{n+1} \mid h_n, \theta, X_n) f(h_n, \theta \mid X_n) d(h_n, \theta) = \\ &= \int f(h_{n+1} \mid h_n, \theta, s_{n+1}) f(h_n, \theta \mid X_n) d(h_n, \theta). \end{aligned} \tag{12}$$

Therefore, $h_{n+1}^{(2)}, i = M+1,...,M+N$ as an approximate sample of $f(h_{n+1} \mid X_n)$, is modeled using the log-normal distribution density $f(h_{n+1} \mid h_n^{(i)}, \theta^{(i)}, s_{n+1})$. Using this sample, the estimates of the distribution density as the h_{n+1} point estimates can be formed. This approach is generalized to calculate a multi-year forecast.

It should be taken into account, that $x(n+1)^{(i)}$, $i = M+1,...,M+N$ calculated during the step 1, are model values with $f(x(n+1) \mid X_n)$. If it is necessary to have an extreme and percentile p-th forecast, let`s say, with p = 1 for estimating the Value-at-Risk (VaR) value on financial markets, then the sample $x(n+1)^{(i)}$ will provide the choice from the p-th empirical percentile. Obviously, as soon as the $h_{n+j}^{(i)}$ from the distribution $f(h_{n+j} \mid X_n)$ is known, the value of the multi-year VaR forecast can be also calculated.

V. Example Of Stochastic Volatility Models Usage

As an example of the stochastic volatility usage, the following model can be presented. A model of the stochastic volatility can be used to make a formal description of the mental dispersion on the market in case of financial risk estimation. That can be done in the following way.

$$\begin{cases} y(k) \mid h(k) = e^{\frac{1}{2}h(k)}\varepsilon(k), \quad \varepsilon(k) \sim N(0,1) \\ h(k) \mid h(k-1), \mu, \phi, \tau^2 = \mu + \phi(h(k-1) - \mu) + \\ + \upsilon(k), \quad \upsilon(k) \sim N(0, \tau^2) \end{cases} \tag{13}$$

To calculate potential losses, the Value-at-Risk (VaR) method can be used. It is based on the estimates of the exchange rate volatility, calculated on the basis of the reciprocal stochastic volatility model. In general, a different kind of model, which complies with the suitability and adequacy of the process, can be used. To investigate the volatility estimate methodology and possible VaR loses,

other selections, generated with the MCMC procedure, can be used. In this case, the posterior VaR mean value can be calculated by the equation:

$$E\left[VaR_{\pi}(N+1)\,|\,\mathbf{r}\right]=\frac{1}{M}\sum_{l=1}^{M}VaR_{\pi}^{(l)}(k) \qquad (14)$$

where N is a number of values, which were actually used from the amount M generated through the MCMC procedure; $\mathbf{r}$ represents the available measurements of the key variable of the investigated process; $VaR^{(l)}(N+1)$ is the value of $VaR_{\pi}(k)$ with $l-$th iteration of the MCMC procedure, which is calculated by the equation:

$$VaR_{\pi}^{(l)}(k)=\left(h^{(l)}(k)\right)^{1/2}\xi_{\pi}^{(l)}(k) \qquad (15)$$

where $\xi_{\pi}^{(l)}(k)$ is the quantile of the generated distribution

$$\hat{f}^{(l)}[\xi(k)]=\sum_{s=1}^{z^{*(l)}}\frac{n_s^{(l)}}{N}\Phi\left[\xi(k)\,|\,\mu_s^{(l)},\,\sigma_s^{2\,(l)}\right]$$

The calculated sample of values $VaR_{\pi}^{(l)}(k),\ l=1,2,...,M$ gives an opportunity to find the median value and the Bayesian tolerance intervals using the quantile of the distribution.

This example shows how a nonexistent (heteroscedastic) process is generated with the following model:

$$\begin{aligned}h(k)&=\alpha_0+\alpha_1 r^2(k-1)+\alpha_2 r^2(k-2)+\\&+\beta_1 h(k-1)+\beta_2 h(k-2)+\xi(k)=\\&=0{,}032+0{,}23\,r^2(k-1)-0{,}095\,r^2(k-2)+\\&+0{,}67\,h(k-1)-0{,}15\,h(k-2)+\xi(k)\end{aligned}.$$

To describe the key variable, the following equation is used:

$$y(k)\,|\,h(k)=\exp(-0{,}5h(k))\,\varepsilon(k),\quad \varepsilon(k)\sim N(\bar{\varepsilon},1).$$

The generation of an innovative process (random variables) was performed using this combination of normal distributions:

$$\xi(k)\sim 0{,}75\ N(0{,}11;\ 0{,}55)+0{,}25\ N(-0{,}95;\ 2{,}25).$$

According to the MCMC algorithm, the sequences of the pseudorandomized numbers in the overall amount of values (in 20000 algorithm iterations) was generated to estimate the parameters of the model. The first 10,000 values were not further examined because they refer to the transitional stage of the estimation process, during which the chosen data generating method is configured. It means that 10000 values were actually used. Considering that $\bar{\varepsilon}=E[\varepsilon(k)]=\alpha_0^{-1/2}E[\xi(k)]$, the posterior distribution for $\bar{\varepsilon}$ can be calculated as follows:

$$\bar{\varepsilon}^{(l)}=\frac{1}{(\hat{\alpha}^{(l)})^{1/2}}\sum_{s=1}^{p}\mu_s^{(l)},\quad l=1,...,M$$

where p is the number of distributions used to generate the mixture.

The results of imitational modeling of the non-stationary heteroscedastic process and the estimation of the mathematical model parameters are given in the Table I. It contains posterior average estimates of the model parameters calculated using the proposed method. For purposes of comparison, it also includes the values of estimates calculated using a simple Gaussian distribution of the innovative random process.

TABLE I. RESULTS OF IMITATIONAL MODELING OF THE NON-STATIONARY HETEROSCEDASTIC PROCESS

№ п/п	Parameter and its value		Suggested method	Gaussian distribution
	Parameter	*True value*	*Posterior average*	*Posterior average*
1	α_0	0,032	0,0297 (7,2%)	0,0355 (10,94%)
2	α_1	0,230	0,245 (6,5%)	0,258 (12,21%)
3	α_2	-0,095	-0,0877 (7,7%)	-0,0998 (5,05%)
4	β_1	0,670	0,6581 (1,8%)	0,749 (11,79%)
5	β_2	-0,150	-0,163 (8,7%)	-0,132 (12,01%)
6	$\bar{\varepsilon}$	0,0095	0,0078 (17,9%)	0,0114 (20,0%)
7	Average error of estimate %		8,30%	12,01%

The percentage in parentheses indicates the average estimate errors related to the exact values of the given model. It is evident that the parameter estimates, calculated according to the suggested method, are much closer to the true values. Thus, the average estimate error in percentages is 8.3% and 12.01% accordingly. I.e. the estimate errors decreased 1.5 times approximately. Consequently, the alternative method, chosen for the comparison, allows getting the estimates, close to the true values of the used model.

VI. CONCLUSION

Investigations related to the probabilistic-statistical volatility modeling are highly important due to the necessity of the forecasts estimates quality improvement and the decisions taken on their basis. Therefore, a particular consideration is given to the method of Bayesian analysis of non-stationary (heteroscedastic) processes, which are widely distributed in various spheres of human life. The development of the volatility estimation methods based on the Bayesian analysis allows to significantly improve the quality of forecasts and their estimation.

The usage of various modifications of Monte Carlo method-based approaches for Markov chains makes it

possible to correctly solve the issues of mathematical model parameters estimation within the complex structures, provided that there are random influences with arbitrary distributions. Further investigation of the Bayesian analysis may be aimed at improving the methods of estimating the parameters of various probabilistic and statistical volatility models with the help of adaptive estimation schemes; expanding the criteria basis for analyzing the quality of intermediate and final results; building specialized decision support systems for the analysis of nonlinear non-stationary processes in order to take substantiated financial and economic decisions. In particular, this applies to the current systems of risk management, analysis of the price formation processes on stock exchanges, investments and economical diagnostic.

REFERENCES

[1] R. Engle, " Autoregressive conditional heteroscedasticity with estimates of variance of united kingdom inflation," Econometrica. vol. 50. pp. 987–1008. 1982.

[2] R. Engle, and T. Bollerslev, "Modelling the persistence of conditional variances," Econometric Reviews. vol. 5, no. 1, pp. 1–50. 1986.

[3] M. Asai, M. McAleer, and Jun. Yu, "Multivariate stochastic volatility: areview". Econometric Reviews, vol. 25, pp.145-75, 2006.

[4] M. Asai, and M. McAleer, "The structure of dynamic correlations in multi variate stochastic volatility models," Journal of Econometrics, vol.150, pp.182-192, 2009.

[5] T. Bollerslev, "Generalized autoregressive conditional heteroskedasticity". Journal of Econometrics. vol. 31, no. 3. pp. 307–327, 1986.

[6] T. Bollerslev, "A conditionally heteroskedastic time series model forspeculative prices and rates of return ," The Review of Economics and Statistics. vol. 69, no. 3. pp. 542–547, 1987.

[7] E. Jacquier, N. G. Polson, and P. E. Rossi. "Bayesian analysis of stochastic volatility models," Journal of Business and Economic Statistics, 20, pp.69-87, 1994.

[8] E. Jacquier, N. G. Polson, and P. E. Rossi. "Bayesian analysis of stochastic volatility models with fat-tails and correlated errors," Journal of Econometrics, vol.122, pp.185-212, 2004.

[9] N. Shephard, Stochastic Volatility: Selected Readings. Oxford: University Press. 2005.

[10] N. Shephard, and T. A. Andersen, Stochastic Volatility: Origins and Overview. Handbook of Financial Time Series. New York: Springer. 2009.

[11] T. Bollerslev, R. Chou, K. Kroner, "Arch modeling in finance : A review of the theory and empirical evidence," Journal of Econometrics. vol. 52, no. 1-2, pp. 5–59, 1992.

[12] T. Andersen, "Stochastic autoregressive volatility: a framework for volatility modeling, "Mathematical Finance. vol. 4, pp. 75– 102, 1994.

[13] S. Taylor, "Modeling stochastic volatility: A review and comparative study," Mathematical Finance.Vol. 4, no. 2, pp. 183–204, 1994.

[14] Capobianco, E. State-spaces to chastic volatility models: areview of estimation algorithms. Applied Stochastic Models and Data Analysis. vol. 12, pp. 265–279,1996.

[15] T. Andersen, T. Bollerslev, S. Lange, "Forecasting financial market volatility: Sample frequency visa-vis forecast horizon," Journal of Empirical Finance. vol. 6, no. 5, pp. 457–477,1999.

[16] S. Babichev, V. Lytvynenko, M. Korobchynskyi, and M.A. Taiff, "Objective clustering inductive technology of gene expression sequences features," Communications in Computer and Information Science, 716, pp. 359-372, 2017