Instituto de Ciências Matemáticas e de Computação

ISSN - 0103-2569

Bayesian Estimation and Prediction of Stochastic Volatility Models via INLA

Mauricio Zevallos and Ricardo S. Ehlers

№ 386

RELATÓRIOS TÉCNICOS DO ICMC

São Carlos
Outubro 2012

Bayesian Estimation and Prediction of Stochastic Volatility Models via INLA

Mauricio Zevallos[a] and Ricardo S. Ehlers[b*]

[a] Campinas State University, Brazil

[b] São Paulo State University, Brazil

Abstract

In this paper we assess Bayesian estimation and prediction using integrated Laplace approximation (INLA) on a stochastic volatility model. This was performed through a Monte Carlo study with 1000 simulated time series. To evaluate the estimation method, two criteria were considered: the bias and square root of the mean square error (smse). The criteria used for prediction are the one step ahead forecast of volatility and the one day Value at Risk (VaR). The main findings are that the INLA approximations are fairly accurate and relatively robust to the choice of prior distribution on the persistence parameter. Additionally, VaR estimates are computed and compared for three financial time series returns indexes.

Keywords: Stochastic volatility models, quasi-maximum likelihood, INLA, Bayesian methods.

1 Introduction

Stochastic volatility (SV) models have been applied with success in order to model the time-varying volatility present in financial time series. To estimate these models several estimation methods have been proposed in the literature, quasi-maximum likelihood methods (Harvey et al. 1994), generalized method

*Corresponding author. Email: ehlers@icmc.usp.br

of moments (Andersen and Sorensen 1996) and Markov Chain Monte Carlo Methods (MCMC) (Jacquier et al. 1994), to name a few. For an account of recent developments in the estimation of SV models see Broto and Ruiz (2004) and Shephard and Andersen (2009) and the references therein.

In particular, MCMC methods are considered one of the most efficient estimation methods. However, its implementation can be very computationally demanding and it requires a training to assess the convergence of the chains. This could be very hard for practitioners interested in fast answers and available computational programs.

In order to overcome the problems associated with MCMC methods, Rue et al. (2009) have recently proposed a very fast, non-sampling based, Bayesian methodology named Integrated Nested Laplace Approximations (INLA). The INLA program is free for downloading and many examples of applications in several fields have appeared in the recent literature. In particular, in Martino et al. (2010) the use of this method is illustrated in the analysis of two financial time series. However, to the best of our knowledge, no study on the performance of this method in terms of estimation and prediction on SV models has yet been made.

The first objective of this paper is to assess the efficiency of INLA methodology for estimating SV models in terms of parameter estimates and predicted volatilities on simulated time series. Second, as a practitioner-oriented point of view, we evaluated the method by estimating the value-at-risk (VaR) of three financial time series returns indexes. VaR estimates are important criteria in risk management. Moreover, in these applications we compare the performance of INLA method with another fast estimation method, the Quasi-Maximum Likelihood Estimation (QMLE) of Harvey et al. (1994). Even though the QMLE method provides an approximated filter for the volatility, it is widely used in applications because it is fast and is implemented in commercial software. The focus of the paper is on the practical side of the methods. The results show that Bayesian methods using the INLA package provide accurate approximations to the model parameters in small computing times. In particular, we observe that the VaR values obtained using INLA react very well to the up and down movements in the return series. We also investigate the sensitivity of parameters and volatility estimates to the choice of prior distributions.

The rest of the paper is organised as follows. In section 2 the methodology is briefly presented. This is assessed through simulated time series in section 3 and through the statistical analysis of real-life time series in section 4.

Finally, some final comments are given in section 5.

2 Methodology

Let r_t the continuosly compounded return on an asset between times $t-1$ and t, that is $r_t = \ln(P_t) - \ln(P_{t-1})$, where P_t is the price at time t. The stochastic volatility model introduced by Taylor (1986) is defined as,

$$\begin{aligned} r_t &= \exp(h_t/2)\varepsilon_t, \quad \varepsilon_t \sim NID(0,1) &(1)\\ h_t &= \gamma + \phi h_{t-1} + \eta_t, \quad \eta_t \sim NID(0,\sigma_\eta^2) &(2) \end{aligned}$$

where η_t is independent of η_s for all t and s. In this model h_t is an unnobserved (latent) variable as well as $\sigma_t = \exp(h_t/2)$, which is the *volatility* of the asset.

The model parameters are $(\gamma, \phi, \sigma_\eta^2)$ or equivalently $\varphi = (\gamma, \phi, \tau)$ where $\tau = 1/\sigma_\eta^2$, $\phi \in (0,1)$ is the persistence parameter and σ_η^2 is the volatility in the log-volatility equation.

Following the Bayesian paradigm, we need to complete the model specification by specifying the prior distributions of the parameters in φ. Here we adopt a reparameterization that maps each restricted parameter space onto the real line. This is important computationally since the algorithm to be used for estimation involves optimization steps which work better in this unrestricted space. For the log-volatility precision we simply take the logarithm and assume that $\log(\sigma_\eta^{-2}) \sim logGamma(a,b)$ so that the prior mean and variance of σ_η^{-2} are a/b and a/b^2 while for the persistence parameter we take a logit type transformation and assume that

$$\kappa = \text{logit}\left(\frac{\phi+1}{2}\right) = \log(1+\phi) - \log(1-\phi) \sim N(c, d^{-1})$$

with reverse map given by $\phi = (e^\kappa - 1)/(e^\kappa + 1)$. In the applications of this paper, we choose the mean c and the precision parameter d so that the corresponding prior distribution for ϕ is roughly uniform in (0,1). The hyperparameters a and b are chosen so that the log-precision has a relatively vague prior distribution. Finally, the common mean parameter γ is assigned a vague normal prior distribution with mean zero and a large variance. The basis for inference on model parameters and prediction of future volatilities is provided by the posterior marginal distributions of γ, ϕ, τ and h_t given $r_1, \ldots, r_n$.

The model in equations (1) and (2) can be rewritten as $r_t|h_t \sim N(0, e^{h_t})$ and $h_t = \gamma + f_t$ with $f_t|f_{t-1}, \ldots, f_1, \tau, \phi \sim N(\phi f_{t-1}, 1/\tau)$. Now, defining $\boldsymbol{x} = (\gamma, h_1, \ldots, h_n)$ and $\boldsymbol{\theta} = (\phi, \tau)$ then $\boldsymbol{x}|\boldsymbol{\theta} \sim N(\mathbf{0}, \boldsymbol{Q}^{-1}(\boldsymbol{\theta}))$ so that $\boldsymbol{x}$ is a Gaussian Markov random field (GMRF) with a sparse precision matrix $\boldsymbol{Q}(\boldsymbol{\theta})$. The posterior marginal distributions are obtained from the joint posterior distribution of $(\boldsymbol{x}, \boldsymbol{\theta})$ which is not available in closed form. However, the numerical methods developed in Rue et al. (2009) (Integrated nested Laplace approximations, INLA) can be applied for approximate inference in this class of models. The INLA approach provides very accurate approximations to the marginal posterior densities of h_t, γ, ϕ and τ in small computing times.

Going into the details of the approximations and associated numerical issues is not the focus of the present paper and the interested readers are referred to Rue et al. (2009) and Martino et al. (2010).

All the computations in this paper were implemented using the open-source statistical software language and environment `R` (R Development Core Team 2006). In particular, we used the add-on package `INLA` (Martino and Rue 2009) which is freely available and can be downloaded from the website `www.r-inla.org`.

3 Simulations

In this section, we perform a Monte Carlo study to evaluate the methodology described in the previous section. We generated m =1000 replications of 1000 observations plus 5 from the SV model (1)-(2). The adopted setup for the true parameter values is the same used in Jacquier et al. (1994).

The performance of the INLA method was assessed in terms of the quality of parameter estimates and the quality of the volatility predictions. Thus, for each generated sample, the estimation of parameters is made based on the first 1000 observations. Then, using these estimates we calculated the predictions of the volatilities and compared these values with the actual data, the last 5 generated observations.

Bayesian estimation using the INLA methods was carried out using the functions in the package `INLA` developed in `R`. The true parameter values were considered as initial values in the estimation process and the prior distributions are as described in Section 2 with hyperparameters $a = 1$, $b = 0.1$, $c = 2.2$ and $d = 1/1.5$. These choices lead to prior distributions for the log-precision and persistence which are relatively vague and roughly uniform in (0,1) respectively.

Finally, in order to get some assessment of the robustness of these simulations we repeated all procedures setting $c = 2.5$ and $d = 1/2$ which corresponds to a roughly uniform prior distribution for ϕ in $(0.5, 0.999)$. We also analised this interval because financial time series returns usually exibit high persistence (ϕ) values.

3.1 Parameter estimates

Let $\hat{\theta}^{(i)}$ the estimate of parameter θ for the i-th replication, $i = 1, \ldots, m$. To evaluate the estimation method, two criteria were considered: the bias and square root of the mean square error (smse), which are defined as,

$$bias = \frac{1}{m}\sum_{i=1}^{m} \hat{\theta}^{(i)} - \theta, \tag{3}$$

$$smse^2 = \frac{1}{m}\sum_{i=1}^{m} \left\{\hat{\theta}^{(i)} - \theta\right\}^2. \tag{4}$$

The results obtained with $m = 1000$ replications are reported in Table 1. The main findings are,

(i) The bias for ϕ is very small. For fixed ϕ the worst results occur for low σ_η, that is for the following (ϕ, σ_η) cases: (0.90, 0.135), (0.95, 0.0964) and (0.98, 0.0614). But even in those situations, the relative bias is inferior to 10%. The bias of ϕ is negative except for the $(\phi, \sigma_\eta) = (0.90, 0.6750)$ case.

(ii) INLA severely overestimates σ_η with large biases for the cases listed in (i).

(iii) INLA methods present large SMSE values for $\hat{\phi}$ and $\hat{\sigma}_\eta$ in the cases listed in (i).

(iv) When assigning a roughly uniform prior for ϕ in (0.5, 0.999) insted of in (0, 0.999) the changes in both bias and smse are not dramatic. In fact, the values of smse rarelly change for γ, ϕ and σ and the bias for estimation of ϕ tend to reduce slightly.

Table 1 around here

3.2 Predicted volatilities

Let $\hat{\sigma}_n^{(i)}[k]$ be the estimated predicted volatility at time $n+k$ based on data $\{y_1, \dots, y_n\}$ and $\sigma_{n+k}^{(i)}$ the true volatility for the i-th replication, $i = 1, \dots, m$. To assess the quality of out-of-sample volatility predictions three criteria are considered: the bias, the relative absolute error (rel.ad) and the square root of the mean square error (smse).

$$
\begin{aligned}
bias &= \frac{1}{m} \sum_{i=1}^{m} \hat{\sigma}_n^{(i)}[k] - \sigma_{n+k}^{(i)}, & (5)\\
rel.ad &= \frac{1}{m} \sum_{i=1}^{m} \left| \frac{\hat{\sigma}_n^{(i)}[k] - \sigma_{n+k}^{(i)}}{\sigma_{n+k}^{(i)}} \right|, & (6)\\
smse^2 &= \frac{1}{m} \sum_{i=1}^{m} \left\{ \hat{\sigma}_n^{(i)}[k] - \sigma_{n+k}^{(i)} \right\}^2 . & (7)
\end{aligned}
$$

The results obtained by applying the above criteria for $m = 1000$ replications with $k = 1, 2, 3, 4, 5$ are given in Table 2. When a roughly uniform prior distribution in $(0.5, 0.999)$ was assigned, the results were practically the same and are therefore omitted. The main findings are,

(i) Underestimation of the predicted volatility.

(ii) For each combination of parameters, the bias is almost the same for each k, the prediction horizon, excepting for (ϕ, σ_η) cases (0.90, 0.675), (0.95, 0.4835) and (0.98, 0.30) where the bias increases. That is, for fixed ϕ the worst cases occur for large σ_η. This was expected.

(iii) For each ϕ we observe the smaller values of rel.ad for the lowest σ. On the contrary, the bigger values of rel.ad are observed for larger σ. These values are large for large k.

(iv) SMSE increases with k and also increases with σ for all values of k.

Table 2 about here

4 Applications

In this section, the described methodology is applied to real financial time series data. Specifically, we estimated the value at risk (VaR) of three stock market indexes: the SP500 of USA, the FTSE100 of UK and the NIKKEI225 of Japan. The time series under study are the daily continuously compounded returns in percentage, as defined at the beginning of Section 2.

The time series returns (Figures 1, 2 and 3) are calculated based on the closing quotations from 2 January 2003 to 31 October 2007. Thus we have 1216 returns for the SP500, 1222 returns for the FTSE100 and 1188 returns for the NIKKEI225. The SP500 time series was analized by Martino et al. (2010). These time series were obtained from the website `http://finance.yahoo.com/`.

Figure 1 about here

Figure 2 about here

Figure 3 about here

For each time series, we estimated SV models using INLA considering the following three different distributions for the errors ε_t in (1), the Gaussian, the Student-t with ν degrees of freedom and the Normal Inverse Gaussian (NIG) (Barndorf-Nielsen 1997) distributions. For the degrees of freedom parameter in the Student t distribution we take the transformation $\nu^* = \log(\nu - 2)$ and assume that $\nu^* \sim N(0, \sigma^2_{\nu^*})$. The NIG distribution has a skewness parameter β and a shape parameter ψ. These are assigned the following prior distributions $\beta \sim N(0, 10)$ and $\phi^* = \log(\psi - 1) \sim N(1, 0.5)$ which are the INLA default specification.

The estimated posterior means and standard deviations for each parameter are shown in Table 3. We can observe high persistence estimates ($\hat{\phi}$). In addition, we obtained moderate values of ν the degrees of freedom in the t Student distribution, indicating not too heavy tails.

For each time series we also estimated the one day 99% VaR for the last 252 observations (one stock market year approximately). To reproduce a real scenario, the VaR estimates were calculated using the one-step ahead predicted volatility. Thus given the observations $r_1, \ldots, r_n$ we estimated the parameters and then calculated VaR_{n+1}. In consequence, we estimated the model 252 times.

In the Bayesian approach, the VaR calculation is based on the one-step ahead predictive posterior distribution of the returns. This is given by,

$$p(r_{n+1}|\boldsymbol{r}) = \int p(r_{n+1}|h_{n+1})p(h_{n+1}|\boldsymbol{r})dh_{n+1}, \tag{8}$$

which is not available analytically and not directly computed by the INLA program. Let $\hat{h}_n[1]$ the posterior mean of h_{n+1} which is estimated by INLA. We then calculated the Bayesian VaR as

$$VaR_{n+1} = q \exp(\hat{h}_n[1]/2). \tag{9}$$

where q is the 99% quantile of the standard normal, t or NIG distribution. If the error distribution is t or NIG the hyperparameters are fixed at the estimated values of ν, β and ψ given in Table 3.

Figures 4, 5 and 6 show the last 252 returns and the VaR estimates for the INLA method. In 252 observations we expected $100/252 = 0.397\%$ observations below the VaR. When using INLA estimates with Gaussian errors we obtained 11 for SP500, 8 for FTSE100 and 8 for NIKKEI225. However, when using a NIG distribution we obtained better results, obtaining 6, 6 and 3 observations outside the VaR limits, respectively. The VaR estimates follow very well the volatility in the market and reacts well to extreme down movements (large negative return values). Additionally, we have included in Figures 4-6 the estimated VaR for the QMLE method of Harvey et al. (1994) (assuming Gaussian errors). We note that the INLA estimates provide a better reaction to the ups and downs while the VaR computed from QMLE are sometimes unnecessarily large. Besides, the VaR values based on INLA estimates for the three error distributions are very similar thus indicating some sort of robustness in terms of choice of these distributions.

Figure 4 about here

Figure 5 about here

Figure 6 about here

5 Conclusions

In this paper we evaluated Bayesian methods to estimate the parameters in a stochastic volatility model. We employed the Bayesian approach using the numerical methods known as Integrated Laplace approximations (INLA) to obtain the approximations to the posterior marginals of interest.

The approximations for parameters and predicted volatilities appear to be accurate in terms of bias, mean square errors and relative absolute errors. In terms of parameter estimation, the simulation results were relatively robust to the choice of prior on the persistence ϕ. Of course, as in any Monte Carlo study, our results are limited to our particular selection of sample sizes, prior distributions, etc. However, we note that these are typical choices in most financial application and we hope that our findings are useful to the practitioners.

Acknowledgements

This research was partially supported by FAPESP and Laboratório EPI-FISMA for the first author. We would like to thank Marcio Laurini for the discussion during this work.

References

Andersen, T. and B. Sorensen (1996). GMM estimation of a stochastic volatility model: A Monte Carlo study. *Journal of Business and Economic Statistics*.

Barndorf-Nielsen, O. E. (1997). Normal inverse Gaussian distributions and stochastic volatility modelling. *Scandinavian Journal of Statistics 24*, 1–13.

Broto, C. and E. Ruiz (2004). Estimation methods for stochastic volatility models: A survey. *Journal of Economic Surveys 18*, 613–649.

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

Jacquier, E., N. G. Polson, and P. E. Rossi (1994). Bayesian analysis of stochastic volatility models (with discussion). *Journal of Business and Economic Statistics 12*, 371–418.

Martino, S., K. Aas, O. Lindqvist, L. Neef, and H. Rue (2010). Estimating stochastic volatility models using integrated nested laplace approximations. *The European Journal of Finance (doi:10.1080/1351847X.2010.495475)*, 1–17.

Martino, S. and H. Rue (2009). Implementing approximate Bayesian inference for latent Gaussian models using integrated Laplace approximations: a manual for `inla` program. Technical report, Norwegian University for Science and Technology, Trondheim (Available from http://www.math.ntnu.no/hrue/GMRFLib).

R Development Core Team (2006). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. ISBN 3-900051-07-0.

Rue, H., S. Martino, and N. Chopin (2009). Approximate Bayesian inference for latent Gaussian models by using integrated nested Laplace approximations (with discussion. *Journal of the Royal Statistical Society: Series B 71*(2), 319–392.

Shephard, N. and T. G. Andersen (2009). Stochastic volatility: Origins and overview. In *Handbook of Financial Time Series*, pp. 233–254. Springer.

Taylor, S. (1986). *Modelling Financial Time Series*. Wiley.

Table 1: Bias and square root of the mean squared error using INLA.[a] estimates with restriction $\hat{\phi} \in (0, 0.999)$, [b] estimates with restriction $\hat{\phi} \in (0.5, 0.999)$.

Parameters			$\hat{\gamma}$		$\hat{\phi}$		$\hat{\sigma}$	
γ	ϕ	σ	**Bias**	**SMSE**	**Bias**	**SMSE**	**Bias**	**SMSE**
-0.7061	0.90	0.1350	[a] -0.0020	0.0661	-0.0961	0.1431	0.1568	0.1659
			[b]-0.0007	0.0662	-0.0782	0.1322	0.1558	0.1652
-0.7360	0.90	0.3629	0.0093	0.1220	-0.0075	0.0302	0.4259	0.4335
			0.0102	0.1221	-0.0053	0.0296	0.4259	0.4335
-0.8210	0.90	0.6750	-0.0009	0.2234	0.0022	0.0192	0.8323	0.8410
			-0.0003	0.2234	0.0032	0.0194	0.8342	0.8431
-0.3530	0.95	0.0964	-0.0049	0.0744	-0.0842	0.1276	0.2017	0.2089
			-0.0035	0.0744	-0.0676	0.1132	0.2001	0.2073
-0.3680	0.95	0.2600	0.0038	0.1692	-0.0082	0.0205	0.5228	0.5312
			0.0050	0.1693	-0.0061	0.0195	0.5250	0.5336
-0.4106	0.95	0.4835	-0.0110	0.3143	-0.0044	0.0146	1.0000	1.0130
			-0.0103	0.3144	-0.0032	0.0143	1.0097	1.0234
-0.1412	0.98	0.0614	-0.0067	0.1087	-0.0739	0.1196	0.2379	0.2445
			-0.0052	0.1086	-0.0606	0.1196	0.2365	0.2432
-0.1472	0.98	0.1657	-0.0013	0.2675	-0.0110	0.0174	0.5746	0.5862
			-0.0001	0.2674	-0.0090	0.0157	0.5830	0.5954
-0.1642	0.98	0.3082	0.0147	0.4682	-0.0064	0.0111	1.0975	1.1157
			0.0154	0.4684	-0.0051	0.0104	1.1240	1.1441

Table 2: Bias, mean absolute deviation and square root of the mean squared error for predictions evaluation using INLA.

Parameters		**Measure**	**Steps ahead**				
ϕ	σ		**1**	**2**	**3**	**4**	**5**
0.90	0.135	bias	-0.0150	-0.0129	-0.0110	-0.0100	-0.0110
		rel.ad	0.1149	0.1160	0.1182	0.1213	0.1229
		smse	0.1451	0.1482	0.1494	0.1534	0.1554
0.90	0.363	bias	-0.0760	-0.0852	-0.0832	-0.0869	-0.0923
		rel.ad	0.2617	0.2826	0.2962	0.3086	0.3206
		smse	0.4001	0.4254	0.4268	0.4505	0.4611
0.90	0.675	bias	-0.1938	-0.2193	-0.2444	-0.2525	-0.2716
		rel.ad	0.4399	0.5118	0.5592	0.6116	0.6549
		smse	0.8676	0.9688	1.0293	1.0725	1.1092
0.95	0.0964	bias	-0.0071	-0.0092	-0.0078	-0.0064	-0.0074
		rel.ad	0.1167	0.1175	0.1207	0.1202	0.1210
		smse	0.1467	0.1484	0.1523	0.1515	0.1526
0.95	0.26	bias	-0.0604	-0.0677	-0.0677	-0.0677	-0.0620
		rel.ad	0.2221	0.2324	0.2490	0.2620	0.2726
		smse	0.3286	0.3437	0.3596	0.3766	0.3882
0.95	0.4835	bias	-0.1185	-0.1342	-0.1348	-0.1744	-0.1841
		rel.ad	0.3612	0.4024	0.4612	0.4849	0.5182
		smse	0.7418	0.7995	0.8610	0.9678	1.1294
0.98	0.0614	bias	-0.0138	-0.0145	-0.0125	-0.0130	-0.0138
		rel.ad	0.1021	0.1032	0.1051	0.1055	0.1070
		smse	0.1320	0.1332	0.1350	0.1362	0.1381
0.98	0.1657	bias	-0.0498	-0.0513	-0.0520	-0.0454	-0.0499
		rel.ad	0.1878	0.1970	0.2110	0.2186	0.2299
		smse	0.2839	0.2950	0.3104	0.3156	0.3356
0.98	0.308	bias	-0.1047	-0.1077	-0.1152	-0.1285	-0.1448
		rel.ad	0.2799	0.3050	0.3307	0.3452	0.3681
		smse	0.6464	0.7182	0.7479	0.7992	0.8697

Table 3: Posterior means and standard deviations (in parentheses) for the estimates of the selected time series using INLA.

Series	Model	γ	ϕ	$\tau = 1/\sigma^2$	ν	β	ψ
SP500	Normal	-0.57 (0.18)	0.98 (0.01)	3.28 (0.98)			
	Student t	-0.51 (0.20)	0.99 (0.01)	3.60 (1.19)	10.75 (2.60)		
	NIG	-0.53 (0.20)	0.98 (0.01)	3.70 (1.24)		-0.26 (0.16)	2.48 (0.70)
FTSE	Normal	-0.60 (0.21)	0.98 (0.01)	1.92 (0.49)			
	Student t	-0.54 (0.23)	0.98 (0.01)	2.03 (0.58)	14.70 (3.71)		
	NIG	-0.57 (0.22)	0.98 (0.01)	1.93 (0.53)		-0.25 (0.28)	5.40 (2.28)
NIKKEI	Normal	0.15 (0.16)	0.97 (0.01)	2.87 (0.77)			
	Student t	0.21 (0.17)	0.98 (0.01)	3.28 (1.00)	11.50 (2.88)		
	NIG	0.19 (0.17)	0.98 (0.01)	3.42 (1.05)		-0.43 (0.20)	2.67 (0.63)

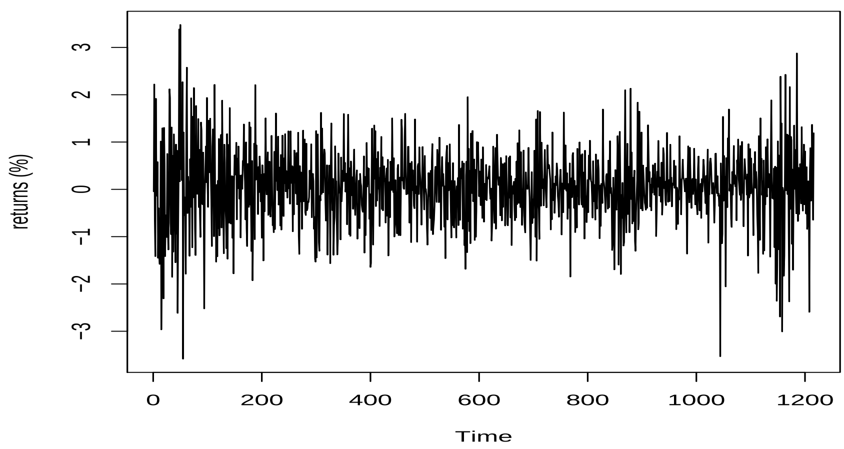

Figure 1: SP500 time series returns.

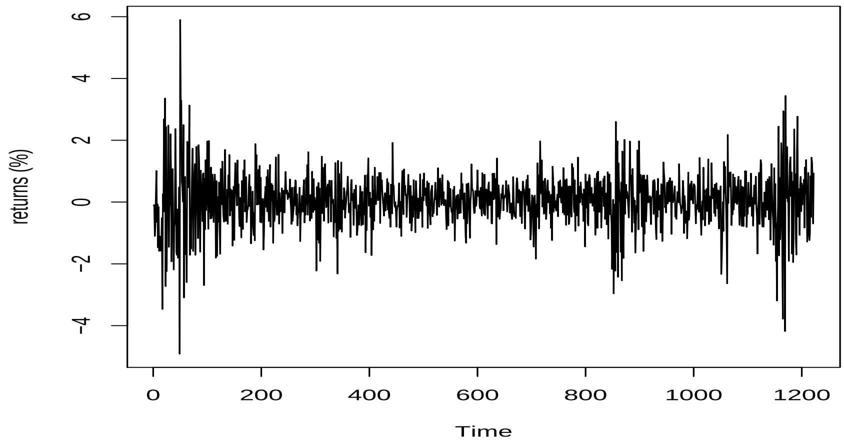

Figure 2: FTSE100 time series returns.

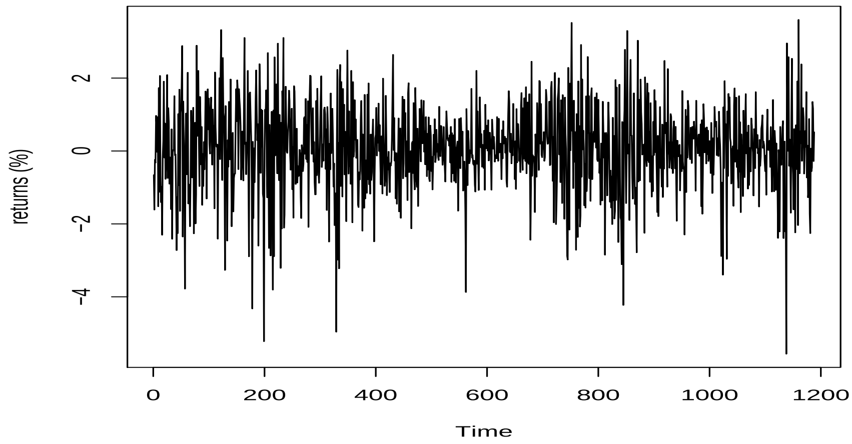

Figure 3: NIKKEI225 time series returns.

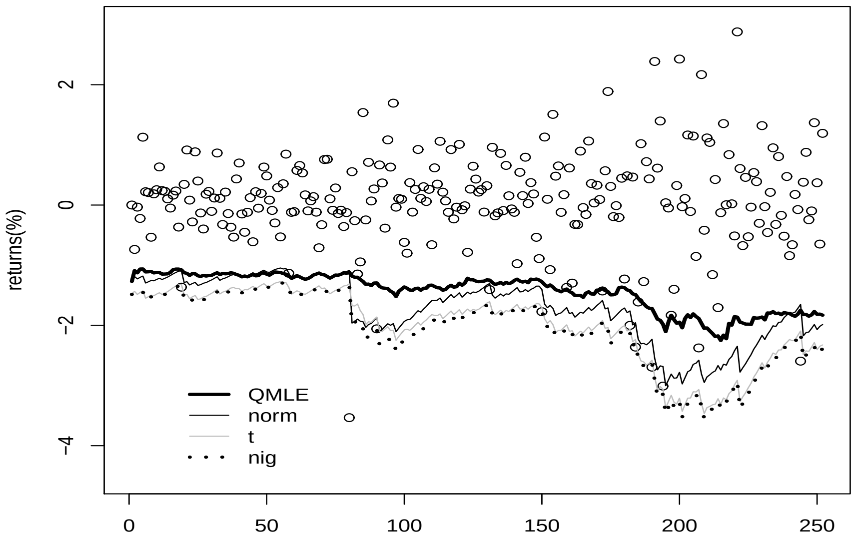

Figure 4: 99% Value at risk of SP500

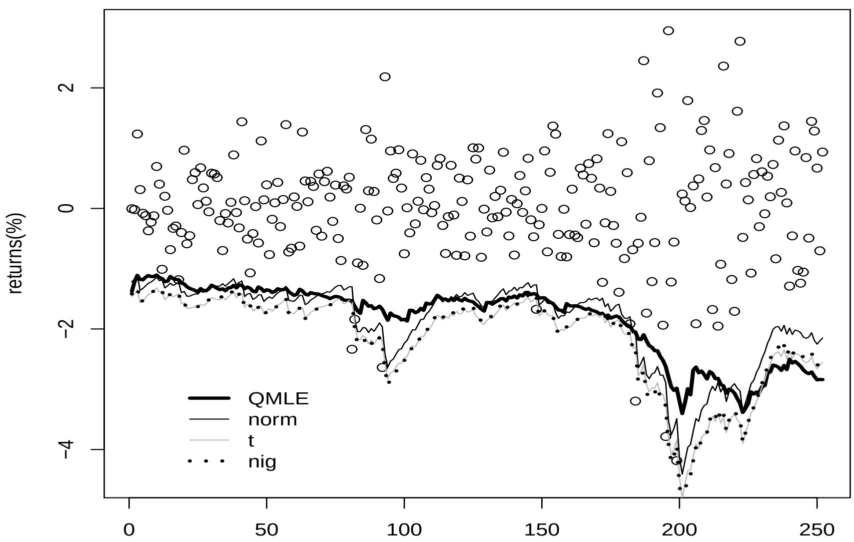

Figure 5: 99% Value at risk of FTSE100

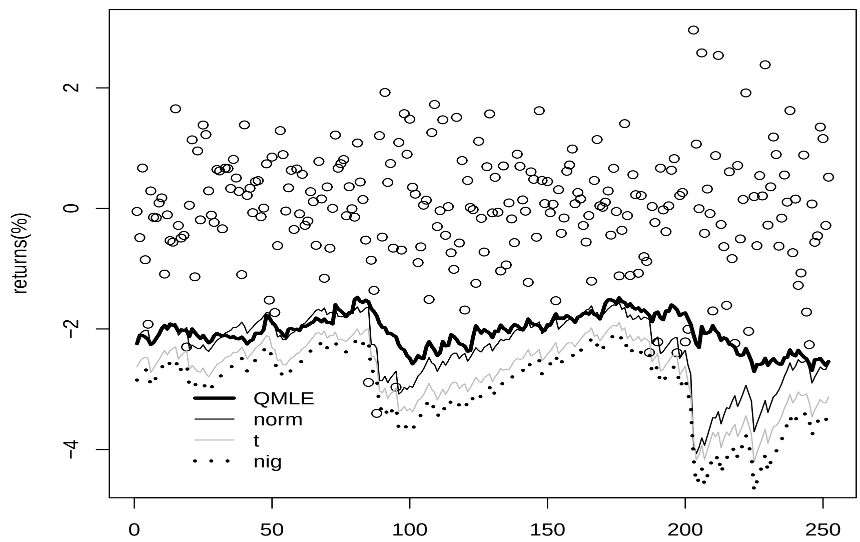

Figure 6: 99% Value at risk of NIKKEI225