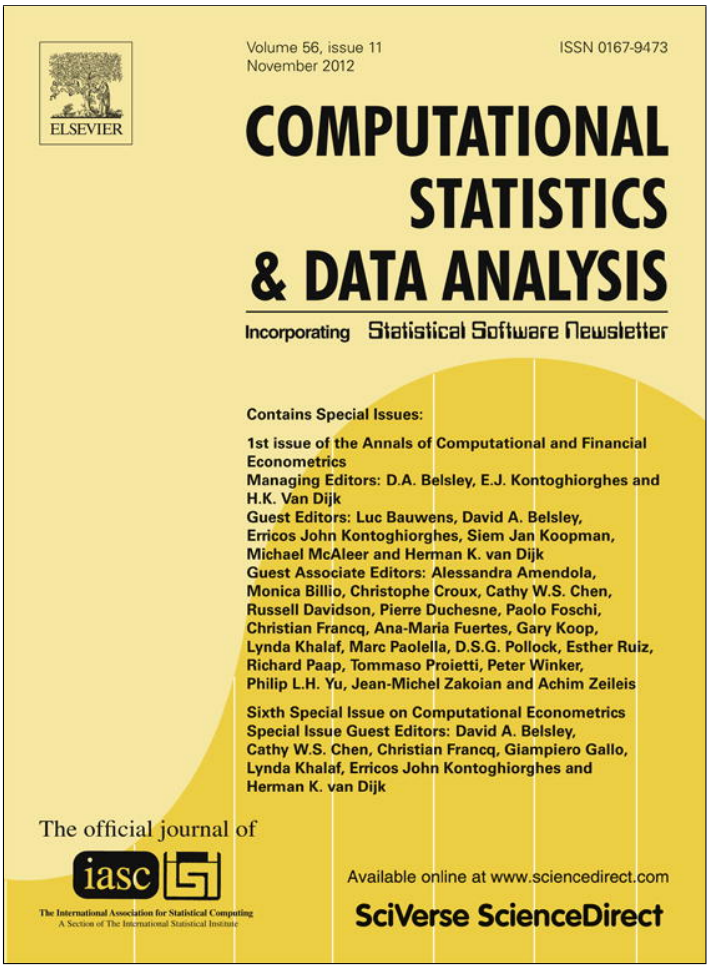

Computational Statistics and Data Analysis 56 (2012) 3006–3019

Contents lists available at SciVerse ScienceDirect

Computational Statistics and Data Analysis

journal homepage: www.elsevier.com/locate/csda

A Bayesian conditional autoregressive geometric process model for range data

J.S.K. Chan [a,*], C.P.Y. Lam [a], P.L.H. Yu [b], S.T.B. Choy [c], C.W.S. Chen [d]

[a] School of Mathematics and Statistics, The University of Sydney, Australia
[b] Department of Statistics and Actuarial Science, The University of Hong Kong, Hong Kong
[c] Discipline of Operations Management and Econometrics, The University of Sydney, Australia
[d] Graduate Institute of Statistics and Actuarial Science, Feng Chia University, Taiwan

ARTICLE INFO

Article history:
Available online 31 January 2011

Keywords:
Geometric process
Range data
CARR model
Bayesian analysis
WinBUGS

ABSTRACT

Extreme value theories indicate that the range is an efficient estimator of local volatility in financial time series. A geometric process (GP) framework that incorporates the conditional autoregressive range (CARR)-type mean function is presented for range data. The proposed model, called the conditional autoregressive geometric process range (CARGPR) model, allows for flexible trend patterns, threshold effects, leverage effects, and long-memory dynamics in financial time series. For robustness considerations, a log-t distribution is adopted. Model implementation can be easily done using the WinBUGS package. A simulation study shows that model parameters are estimated with high accuracy. In the empirical study on the range data of an Australian stock market index, the CARGPR model outperforms the CARR model in both in-sample estimation and out-of-sample forecast.

1. Introduction

Volatility has become a standard risk measure in financial markets. Accurate forecasting of volatility is important but difficult because financial time series often exhibit time-varying volatility and volatility clustering. They are periods of elevated volatility interspersed among more tranquil periods. Two main classes of models are derived to capture the dynamics of the volatility precisely: they are the generalized autoregressive conditional heteroskedasticity (GARCH) models (Bollerslev, 1986) and the stochastic volatility (SV) models (Hull and White, 1987). Essentially, the GARCH and SV models are return-based models as they are constructed using the data of closing prices, neglecting all intra-day price movement. Recent research has proposed using daily ranges to construct estimates of daily return volatility since daily ranges are known to be more efficient measures of return volatility (see Parkinson, 1980; Andersen and Bollerslev, 1998; Alizadeh et al., 2002), than daily returns. Chou (2005) proposed a range-based model called the conditional autoregressive range (CARR) model that described the dynamics of the conditional mean of the range. Later, Chen et al. (2008) allowed for exogenous threshold variables to fully examine asymmetric range effects in a threshold CARR model.

Most financial time series models do not account for trend movement explicitly. This paper proposes using the geometric process (GP) model of Lam (1988) to capture the trend movement in financial time series. The model contains a ratio parameter a which discounts a monotone process to a renewal process (RP) (Feller, 1949) with a constant mean μ. The two components, the mean μ and ratio a, allow separately the effects on the underlying RP and the effects on the strength and direction of trend movement. Moreover, as the ratio a affects both the mean and the variance of a GP, the model can

* Corresponding address: School of Mathematics and Statistics, The University of Sydney, NSW 2006, Australia. Tel.: +61 2 93514873; fax: +61 2 90369267.
E-mail addresses: jchan@maths.usyd.edu.au, jenniferskchan@gmail.com (J.S.K. Chan).

0167-9473/$ – see front matter
doi:10.1016/j.csda.2011.01.006

capture heteroskedasticity. Lastly, with the inherent geometric structure, forecasting using the GP model is simple and straightforward.

Following the idea of Chou (2005), this paper extends the modelling strategy of the GP model to the dynamic CARR model for range data to obtain a simple yet highly efficient model for capturing the dynamics of the volatility. In particular, the mean of the RP is assigned a CARR-type mean function, and the extended model is called the CARGPR model. By incorporating lagged returns in the mean function, the model can capture leverage effects or volatility asymmetry, which refers to the negative return sequences associated with an increase in volatility of the stock returns. The strength of the CARGPR model lies in its flexibility to adapt the dynamics of volatility using the CARR-type mean function and the trend movement specified by the ratio parameter in the GP model. The proposed model is further extended to accommodate a model shift after some time points called thresholds, and is distinguished from the regime switching model, where the changes occur when the outcomes exceed certain threshold levels. Parameter estimation in threshold autoregressive (TAR) models is usually performed in two approaches: the classical likelihood approach (Tong and Lim, 1980; Tong, 1990) and Bayesian approach (Geweke and Terui, 1993; Chen and Lee, 1995). In this paper, we adopt the Bayesian approach using Markov chain Monte Carlo (MCMC) algorithms, and we apply the Metropolis–Hastings algorithm to estimate the threshold time jointly with other model parameters. A variety of model structures and error distributions can be considered to provide a tailor-made analysis (Chiu and Wang, 2006). For robustness considerations, a heavy-tailed distribution such as Student's t-distribution is considered, and it is expressed in the scale mixture representation to allow a simpler Gibbs sampler for model implementation and to enable outlier diagnosis (Choy and Chan, 2008).

This paper is structured as follows. Section 2 introduces the CARGPR model with various extensions. Section 3 describes the Bayesian computational methods for statistical inference. Section 4 presents a simulation result to illustrate the performance of the CARGPR model. In Section 5, CARGPR models are fitted to the intra-day range data of the All Ordinaries (AORD) index for the Australian stock market. Finally, the paper is concluded in Section 6. The full conditional distributions for the Gibbs sampling algorithm are given in the Appendix.

2. Model development

2.1. The GP model

Lam (1988) first proposed modelling a monotone trend directly by a monotone process called a geometric process (GP). Let $X_1, X_2, \ldots$ be a set of positive random variables. If there exists a positive real number a, called the ratio, such that $\{Y_t = a^{t-1}X_t, t = 1, 2, \ldots\}$ after discounting by a forms a renewal process (RP) (Feller, 1949), then $\{X_t, t = 1, 2, \ldots\}$ is called a GP. The stationary RP $\{Y_t\}$ with a constant mean $E(Y_t) = \mu$ constitutes a special case of the linear model when $a = 1$. Hence the GP model is in fact a generalized model that allows trends when a is non-unit. Let the mean and variance of $\{Y_t\}$ be

$$E(Y_t) = \mu \quad \text{and} \quad \text{Var}(Y_t) = \sigma^2,$$

respectively. Then the mean and variance of $\{X_t\}$ are given by

$$E(X_t) = \mu/a^{t-1} \quad \text{and} \quad \text{Var}(X_t) = \sigma^2/a^{2(t-1)}, \tag{1}$$

respectively. This original GP model with a constant mean μ and a constant ratio a is very restrictive in applications, and these variables are replaced by a time-dependent mean μ_t and a time-dependent ratio a_t, respectively, in this paper.

By adopting some lifetime distributions to $\{Y_t\}$, the models can be implemented using a parametric approach. Chan et al. (2004) investigated statistical inference for a GP model with a gamma distribution and Lam and Chan (1998) considered a lognormal distribution. In our preliminary study, we found that the lognormal distribution consistently gives a better fit than the gamma distribution (see Section 5.4 for details). As many financial data are heavy tailed, the lognormal distribution is further replaced by the log-t distribution to achieve a robust analysis. To facilitate efficient Bayesian MCMC computation and outlier diagnostics, the t-distribution is expressed as a scale mixture of normal (SMN) distributions. Andrews and Mallows (1974) studied the class of SMN distributions and Choy and Chan (2008) investigated different scale mixture distributions. Such scale mixture formulation for t-distribution has been successfully employed for volatility model description in Chen et al. (2010). Student's t-distribution with location μ, scale σ, and number of degrees of freedom ν has the following SMN representation:

$$t_\nu(y|\mu, \sigma) = \int_0^\infty N\left(y \left| \mu, \frac{\sigma^2}{\lambda}\right.\right) G\left(\lambda \left| \frac{\nu}{2}, \frac{\nu}{2}\right.\right) \mathrm{d}\lambda,$$

which can be expressed hierarchically as

$$Y|\mu, \sigma, \lambda \sim N\left(\mu, \frac{\sigma^2}{\lambda}\right) \quad \text{and} \quad \lambda \sim G\left(\frac{\nu}{2}, \frac{\nu}{2}\right),$$

where $G(\alpha, \gamma)$ denotes the gamma distribution with mean α/γ.

In the GP model, we assume that $\ln Y_t \sim t_\nu(\upsilon_t, \tau^2)$ or $\ln Y_t|\lambda_t \sim N\left(\upsilon_t, \frac{\tau^2}{\lambda_t}\right)$ by conditioning on λ_t. Hence, $X_t = Y_t/a^{t-1}|\lambda_t \sim \text{LN}\left(\upsilon_t - \ln(a_t^{t-1}), \frac{\tau^2}{\lambda_t}\right)$, with the mean and variance given by

$$E(X_t) = \frac{\mu_t}{a_t^{t-1}} = \exp\left[\upsilon_t - \ln(a_t^{t-1}) + \frac{\tau^2}{2\lambda_t}\right] \tag{2}$$

and

$$\text{Var}(X_t) = \frac{\sigma_t^2}{a_t^{2(t-1)}} = \exp\left\{2[\upsilon_t - \ln(a_t^{t-1})] + \frac{\tau^2}{\lambda_t}\right\}\left[\exp\left(\frac{\tau^2}{\lambda_t}\right) - 1\right], \tag{3}$$

respectively. A lognormal distribution can be obtained as a special case when $\lambda_t = 1$.

2.2. The CARGPR model

Let P_t be the price of an asset measured at discrete time intervals (e.g., daily or weekly). The observed range is defined as

$$X_t = [\ln(\max P_t) - \ln(\min P_t)] \times 100, \tag{4}$$

where max (min) is the highest (lowest) price over the time interval. Parkinson (1980) showed that the range of any distribution is proportional to its standard deviation. Hence X_t is an estimator of σ_t for an asset price observed at finer intervals, for example, every 5 min during the trading hours of a day. To specify a dynamic structure in the mean function that describes the persistence of market shocks to the range of prices, Chou (2005) proposed the following CARR(p, q) model for X_t:

$$\begin{aligned} X_t &= \mu_t \epsilon_t, \\ \mu_t &= \beta_0 + \sum_{j=1}^{p} \beta_{1j}\mu_{t-j} + \sum_{j=1}^{q} \beta_{2j} X_{t-j}, \\ \epsilon_t|\Im_{t-1} &\sim f(.|\Im_{t-1}), \end{aligned} \tag{5}$$

where $\Im_{t-1}$ is the set of information up to time $t-1$ and $f(.|\Im_{t-1})$ is the conditional distribution for the errors ϵ_t with unit mean. The stationary condition for the process is

$$C = \sum_{j=1}^{p} \beta_{1j} + \sum_{j=1}^{q} \beta_{2j} < 1, \tag{6}$$

where C determines the persistence of range shocks and the unconditional (long-term) mean of X_t is $\beta_0/(1-C)$. Chou (2005) assumed that $X_t \sim W(\psi_t, \alpha)$, where ψ_t and α are the scale and shape parameters, respectively, $\psi_t = \mu_t/\Gamma\left(1 + \frac{1}{\alpha}\right)$, and $\Gamma(\cdot)$ is a gamma function. The mean and variance are given by μ_t and

$$\sigma_t^2 = \mu_t^2\left[\frac{\Gamma\left(1 + \frac{2}{\alpha}\right)}{\Gamma^2\left(1 + \frac{1}{\alpha}\right)} - 1\right], \tag{7}$$

respectively.

However, this CARR model does not allow for trend movement explicitly. To remedy this, we introduce the GP model and equate the mean function (5) to υ_t in (2) as

$$\upsilon_t = \beta_{\mu 0} + \sum_{j=1}^{p} \beta_{\mu 1j}\upsilon_{t-j} + \sum_{j=1}^{q} \beta_{\mu 2j} \ln(Y_{t-j}). \tag{8}$$

The extended model combining the modelling approaches of the GP and CARR techniques is called the conditional autoregressive geometric process range (CARGPR(p, q)) model.

2.3. The CARGPR model with covariate effects

As the daily range may evolve over time subject to certain external effects, exogenous variables Z_{tj} should be incorporated into the conditional mean function μ_t of the CARGPR model via υ_t as

$$\upsilon_t = \beta_{\mu 0} + \sum_{j=1}^{p} \beta_{\mu 1j}\upsilon_{t-j} + \sum_{j=1}^{q} \beta_{\mu 2j} \ln(Y_{t-j}) + \sum_{j=1}^{r} \beta_{\mu 3j} Z_{tj}. \tag{9}$$

Chou (2005) suggested the use of lagged return, trading volume, and seasonal factors because a negative relationship is often found between the range and the lagged return, suggesting a *leverage* effect that a decrease in return leads to higher volatility and, as expected, a positive relationship is often present between the range and the trading volume.

If we set $p = q = r = 1$ and drop the redundant subscript j in β, (9) becomes

$$\upsilon_t = \beta_{\mu 0} + \beta_{\mu 1}\upsilon_{t-1} + \beta_{\mu 2}\ln(y_{t-1}) + \beta_{\mu 3} z_t, \tag{10}$$

for $t = 2, \ldots, n$, and $\upsilon_1 = \beta_{\mu 0} + \beta_{\mu 3} z_t$ for $t = 1$. This function can be rewritten as

$$\upsilon_t = \beta_{\mu 0} \sum_{i=1}^{t} \beta_{\mu 1}^{i-1} + \beta_{\mu 2} \sum_{i=2}^{t} \beta_{\mu 1}^{i-2} \ln(y_{t-i+1}) + \beta_{\mu 3} \sum_{i=1}^{t} \beta_{\mu 1}^{i-1} z_{t-i+1}, \tag{11}$$

showing the complexity of parameter $\beta_{\mu 1}$ in υ_t. Note that the stationary constraint $C < 1$ in (6) does not apply to (10) with a log link function, as shown in (2). However, we find that the sum of parameters in υ_t is less than 1 for most of the models reported in Table 3 in the empirical study. When $a = 1$, X_t, which is just Y_t, is neither increasing nor decreasing, and the stationary constraint in (6) does not apply too.

On the other hand, the CARGPR model can be extended to allow for multiple trends to describe different stages of development, the growing stage ($a < 1$), stabilizing stage ($a = 1$), and declining stage ($a > 1$), for a certain event. In this case, the constant ratio a in (2) and (3) is replaced by a time-dependent ratio function log linked to a function of covariates. For example, the ratio function in the empirical study is

$$a_t = \exp(\beta_{a0} + \beta_{a1} \ln t). \tag{12}$$

2.4. The CARGPR model with threshold effects

Particularly when the time series is long, some structural changes may occur so that model shifts should be accommodated at some time points T called the turning points. Chan et al. (2006) extended the GP model to the threshold GP (TGP) model by fitting a separate GP to each stage, growing, stabilizing, and declining, of the development of an epidemic. Suppose that there are M GPs, $\text{GP}_m = \{X_t : T_m \le t < T_{m+1}\}$, $m = 1, \ldots, M$, with the turning points T_m ($T_1 = 1$) which mark the times of the model shifts. We assume that $X_t|\lambda_t \sim \text{LN}\left(\upsilon_{tm} - \ln(a_m^{t-T_m}), \frac{\tau_m^2}{\lambda_t}\right)$ for $T_m \le t < T_{m+1}$, where

$$\upsilon_{tm} = \beta_{\mu 0m} + \sum_{j=1}^{p} \beta_{\mu 1jm}\upsilon_{t-j,m} + \sum_{j=1}^{q} \beta_{\mu 2jm} \ln(Y_{t-j}) + \sum_{j=1}^{r} \beta_{\mu 3jm} Z_{tj}. \tag{13}$$

The mean and variance for X_t become

$$E(X_t) = \frac{\mu_{tm}}{a_m^{t-T_m}} = \exp\left[\upsilon_{tm} - \ln(a_m^{t-T_m}) + \frac{\tau_m^2}{2\lambda_t}\right] \tag{14}$$

and

$$\text{Var}(X_t) = \frac{\sigma_m^2}{a_m^{2(t-T_m)}} = \exp\left\{2[\upsilon_{tm} - \ln(a_m^{t-T_m})] + \frac{\tau_m^2}{\lambda_t}\right\}\left[\exp\left(\frac{\tau_m^2}{\lambda_t}\right) - 1\right], \tag{15}$$

respectively.

Tiwari et al. (2005) estimated the number of turning points using different model selection criteria. In applications, the number of turning points M and the range from which each turning point T_m is sampled are determined by examining the empirical time series. In general, the best model among models with $M = 1, 2, 3, \ldots$ thresholds can be selected based on some model selection criterion such as the Bayesian Information Criterion (BIC) and the Deviance Information Criterion (DIC) (Section 5.3).

3. Bayesian inference

The log-likelihood function and its derivatives as required in the classical likelihood approach are difficult to evaluate because υ_t in (11) is a complicated function of $\beta_{\mu 1}$. On the other hand, the Bayesian approach using MCMC techniques converts an optimization problem into a sampling problem, by simulation of a single or block of model parameters iteratively, conditional on other parameters and the data. The Gibbs sampling algorithm (Smith and Roberts, 1993; Gilks et al., 1996) and Metropolis–Hastings algorithm (Hastings, 1970; Metropolis et al., 1953) are the most popular MCMC techniques that produce samples from the intractable posterior distributions. For readers who are less familiar with Bayesian computation techniques, we recommend using the WinBUGS (Bayesian analysis Using Gibbs Sampling) package. See Spiegelhalter et al. (2004). The WinBUGS codes for the CARGPR model can be obtained from the authors upon request.

In the simulation and empirical studies, different CARGPR models are compared, and vague and non-informative priors are assigned to the model parameters. The Bayesian hierarchy for the CARGPR models (Models 1–4) is

$$\text{Data: } X_t \sim \text{LN}\left(\upsilon_t - \ln(a^{t-1}), \frac{\tau^2}{\lambda_t}\right).$$

$$\text{Priors: } a \sim U(0.95, 1.05), \qquad \beta_{\mu ij} \sim N(0, \sigma_\beta^2), \qquad \tau^2 \sim IG(\alpha_\tau, \gamma_\tau),$$

$$\lambda_t \sim G\left(\frac{\nu}{2}, \frac{\nu}{2}\right), \qquad \nu \sim G(\alpha_\nu, \gamma_\nu)I(1, 30),$$

where $I(a, b)$ indicates a truncated distribution with support (a, b) and $\lambda_t = 1$ for Model 1. With a ratio function a_t (Model 4), the priors are $\beta_{ai} \sim N(0, \sigma_\beta^2),\ i = 0, 1$. The Bayesian hierarchy for the threshold CARGPR model (Model 5) is

$$\text{Data: } X_t \sim \text{LN}\left(\upsilon_t - \ln(a_m^{t-T_m}), \frac{\tau_m^2}{\lambda_t}\right) I(T_m \le t < T_{m+1}).$$

$$\text{Priors: } a_m \sim U(0.95, 1.05), \qquad \beta_{\mu ijm} \sim N(0, \sigma_\beta^2), \qquad \tau_m^2 \sim IG(\alpha_\tau, \gamma_\tau),$$

$$\lambda_t \sim G\left(\frac{\nu_m}{2}, \frac{\nu_m}{2}\right), \qquad \nu_m \sim G(\alpha_\nu, \gamma_\nu)I(1, 30), \qquad T_m \sim U(c_m, d_m),$$

where T_m is assigned a discrete uniform prior on the range $[c_m, d_m]$. The full conditional distributions for the parameters in Model 5 are derived and reported in the Appendix to facilitate the MCMC sampler. Lastly, the Bayesian hierarchy for the CARR model with covariate using the Weibull distribution (Models 6 ($\alpha = 1$) and 7) is

$$\text{Data: } X_t \sim W\left(\mu_t/\Gamma\left(1 + \frac{1}{\alpha}\right), \alpha\right)$$

$$\text{Priors: } \alpha \sim G(c, d), \qquad \beta_{\mu 0} \sim N(0, \sigma_\beta^2), \qquad \beta_{\mu 1} \sim U(0, 1), \qquad \beta_{\mu 2} \sim U(0, 1 - \beta_{\mu 1}), \qquad \beta_{\mu 3} \sim N(0, \sigma_\beta^2).$$

The hyperparameter σ_β^2 is set to be very large whereas $\alpha_\tau, \gamma_\tau, \alpha_\nu, \gamma_\nu, c$, and d are set to zero for non-informative priors. In the Gibbs sampling scheme, a single Markov chain is run for 7000 iterations, discarding the initial 5000 iterations as the burn-in period to ensure convergence of parameter estimates. Convergence is also carefully checked by the history and autocorrelation function (ACF) plots. Simulated values from the Gibbs sampler after the burn-in period are taken to mimic a random sample of size 2000 from the joint posterior distribution for posterior inference. Parameter estimates are given by the posterior means or medians. To check if the posterior samples of 2000 iterations are sufficient, longer chains of 5000 iterations after burn-in are run for Models 1 and 2, and they give estimates similar to those from 2000 iterations. Moreover, the ACFs and history plots show that the posterior samples are quite uncorrelated. The computation time depends on the complexity of the model and the power of computer, and it is around 4 h using a Core 2 Duo 2 GHz PC for fitting the CARGPR models in the empirical study.

4. Simulation study

In this simulation study, we compare the model performance for models fitted to data of different sizes (small or medium) and adopted different data distributions (lognormal or log-t) and trend patterns (increasing or decreasing). We simulate $N = 100$ data sets; each contains $n = 200$ or $n = 700$ observations. Two models using lognormal (LN) and log-t (LT) distributions are considered, and each model adopts two sets of parameters with decreasing (set 1) and increasing (set 2) trends. Table 1 reports the mean and standard deviation (SD) of the parameter estimates over $N = 100$ replications as given by

$$\hat{\theta} = \frac{1}{N}\sum_{j=1}^{N}\hat{\theta}_j \quad \text{and} \quad \text{SD} = \left[\frac{1}{N-1}\sum_{j=1}^{N}(\hat{\theta}_j - \hat{\theta})^2\right]^{1/2},$$

respectively, where $\hat{\theta}_j$ is the posterior mean of θ in the j-th replication. The performance of the proposed models is further evaluated via three criteria: the absolute percentage bias (APB), root mean square error (RMS) and coverage percentage (CP), defined as

$$\text{APB} = \left|\frac{\hat{\theta} - \theta}{\theta}\right|,$$

$$\text{RMS} = \left[\frac{1}{N}\sum_{j=1}^{N}(\hat{\theta}_j - \theta)^2\right]^{1/2},$$

$$\text{CP} = \frac{100}{N}\sum_{j=1}^{N} I[\theta \in (\hat{\theta}_{j,0.025}, \hat{\theta}_{j,0.975})],$$

Table 1
Parameter estimates, their standard deviation, absolute percentage bias, root mean square error and coverage percentage in the simulation study.

Dist.		Set	a	$\beta_{\mu 0}$	$\beta_{\mu 1}$	$\beta_{\mu 2}$	ν	σ^2	Set	a	$\beta_{\mu 0}$	$\beta_{\mu 1}$	$\beta_{\mu 2}$	ν	σ^2
$n = 200$															
LT	θ	1	1.00100	1.000	−0.200	0.030	5.000	0.500	2	0.99800	−0.020	0.700	0.200	5.000	1.000
	$\hat{\theta}$		1.00091	0.995	−0.207	0.021	8.043	0.545		0.99790	−0.063	0.484	0.222	9.036	1.095
	SD		0.00116	0.216	0.241	0.065	3.877	0.098		0.00415	0.185	0.214	0.055	3.968	0.168
	APB		0.00009	0.005	0.035	0.311	0.609	0.089		0.00010	2.165	0.309	0.109	0.807	0.095
	RMS		0.00115	0.215	0.240	0.066	4.91	0.107		0.00413	0.189	0.304	0.059	5.65	0.192
	CP		85	98	98	95	92	93		91	89	80	95	91	95
$n = 700$															
LN	θ	1	1.00100	1.000	−0.200	0.030	–	0.500	2	0.99800	−0.020	0.700	0.200	–	1.000
	$\hat{\theta}$		1.00100	1.005	−0.212	0.031	–	0.501		0.99802	−0.023	0.659	0.212	–	0.999
	SD		0.00000	0.255	0.280	0.038	–	0.028		0.00056	0.032	0.060	0.026	–	0.055
	APB		0.00000	0.030	0.138	0.202	–	0.005		0.00002	0.128	0.059	0.062	–	0.001
	RMS		0.00000	0.411	0.327	0.034	–	0.031		0.00056	0.031	0.072	0.029	–	0.055
	CP		100	89	89	92	–	94		94	92	88	96	–	95
$n = 700$															
LT	θ	1	1.00100	1.000	−0.200	0.030	5.000	0.500	2	0.99800	−0.020	0.700	0.200	5.000	1.000
	$\hat{\theta}$		1.00100	1.011	−0.206	0.022	5.769	0.517		0.99800	−0.025	0.669	0.208	5.600	1.034
	SD		0.00000	0.269	0.301	0.030	1.359	0.045		0.00065	0.031	0.046	0.027	1.224	0.088
	APB		0.00000	0.202	0.729	0.230	0.114	0.011		0.00000	0.271	0.044	0.041	0.120	0.034
	RMS		0.00000	0.462	0.351	0.035	1.32	0.043		0.00065	0.031	0.055	0.028	1.358	0.094
	CP		100	85	88	95	94	96		94	97	93	93	94	94

Table 2
Summary statistics for the AORD stock market daily range data.

	Range X_t	Ln range $\ln(X_t)$	Return Z_t	Absolute return $\lvert Z_t \rvert$
Mean	1.4311	0.0723	−0.0471	1.0509
SD	0.9908	0.2649	1.4823	1.0457
Kurtosis	7.0649	−0.2866	3.8305	8.3178
Skewness	2.1410	0.2131	−0.5032	2.3368
Minimum	0.2568	−0.5904	−8.5536	0.0000
Maximum	8.0839	0.9076	5.3601	8.5536
Box–Ljung, Q_{12}	2416	2838	20.22[a]	721.7
Cramér–von Mises, W	5.647	0.110[a]	1.337	6.374
Jarque–Bera, JB	2170	8	499	2894

[a] p-value > 0.05. All other p-values are less than 0.02.

respectively, where $(\hat{\theta}_{j,0.025}, \hat{\theta}_{j,0.975})$ is the 95% credible interval of θ in the j-th replication and $I(E)$ is an indicator function for the event E. Models with smaller SD, APB and RMS and with CP closer to 95 are preferred.

From Table 1, the parameter estimates are close to their true values except $\hat{\nu}$ (both sets) and $\hat{\beta}_{\mu 1}$ (set 2) when $n = 200$. The complex function of $\beta_{\mu 1}$ in the mean function (11) explains the difficulty of estimating $\beta_{\mu 1}$ precisely. As for the number of degrees of freedom ν, it is well known that the shape of the t-distribution is rather insensitive to moderate to large numbers of degrees of freedom. However, both estimates improve substantially when the sample size increases to $n = 700$. The CP ranges from 80% to 100% for all CARGPR models, showing satisfactory coverage. There is no obvious difference in model performance between models showing a decreasing (set 1) or increasing (set 2) trend, nor between models adopting a lognormal or log-t distribution. Generally speaking, the results in the simulation study are satisfactory when $n = 200$ and are excellent when $n = 700$.

5. Empirical study

We analyze the intra-day high–low prices X_t defined in (4) from the All Ordinaries (AORD) index for the Australian stock market from 1 May 2006 to 30 April 2009 ($n = 763$) obtained from the website. As suggested in Chou (2005), the lag-one daily log return $Z_{t-1} = [\ln(P_{c,t-1}) - \ln(P_{c,t-2})] \times 100$, where $P_{c,t}$ is the closing price on day t, is taken as a covariate to allow for the leverage effect. Moreover, $|Z_t|$ is chosen to be the proxy of X_t in assessing the forecasting performance of the models. Summary statistics and three test statistics for X_t, $\ln X_t$, Z_t and $|Z_t|$ are reported in Table 2. The first statistic, Box–Ljung Q_{12}, tests the overall randomness of a time series based on 12 lagged autocorrelations. The second and third statistics, the Cramér–von Mises W and Jarque–Bera JB, test for normality in the data. W compares the empirical distribution with the hypothesized distribution while JB measures the departure from normality based on the sample kurtosis and skewness. From Table 2, all tests are significant, showing non-randomness and non-normality, except that Z_t is random and $\ln X_t$ is normal, confirming the lognormal assumption for X_t.

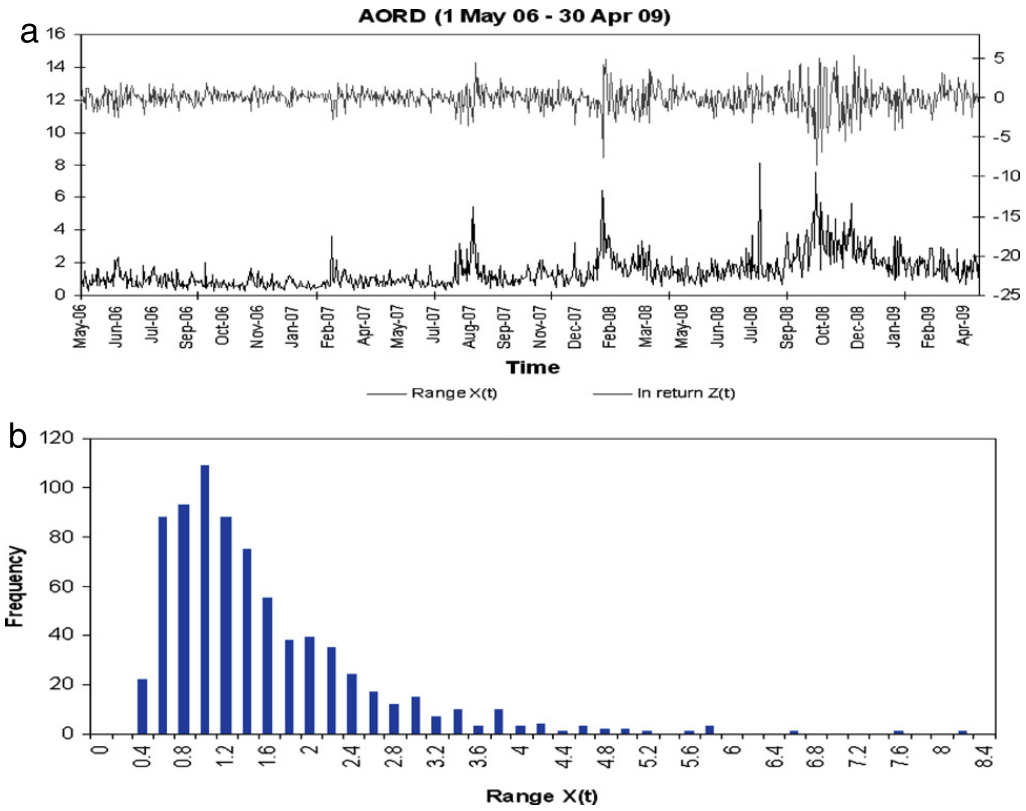

Fig. 1. (a) Observed daily range X_t of the AORD stock market price and (b) histogram of X_t.

Furthermore, time series plots of X_t and Z_{t-1} and the histogram of X_t are presented in Fig. 1. The summary statistics and histogram in Fig. 1(b) show substantial kurtosis and skewness effects in X_t. Moreover, the correlation between X_t and Z_t ($\hat{\rho}_{X,Z} = -0.241$) and their plots in Fig. 1(a) show that the leverage effect is present in the data, because the daily range is high or the price is volatile when the return is low, particularly during the period of the global financial tsunami which started in October 2008.

5.1. Model selection

The basic CARGPR(1, 1) model with lognormal (Model 1) and log-t distributions (Model 2) are first utilized, and Model 2 is preferred according to the DIC because the heavier tails of the log-t distribution can accommodate outliers. The BIC is slightly larger due to the rather heavy penalty for an additional parameter. Hence, the log-t distribution is adopted in all subsequent CARGPR models. By setting $a = 1$, the trend movement is not modelled, similar to the CARR model (Model 2.1), but it adopts a log-t instead of Weibull distribution with a log link function and uses the Bayesian approach in parameter estimation.

To describe different levels of persistence, the CARGPR(1, 2) and CARGPR(2, 1) models are considered. However, the CARGPR(2, 1) model has a technical problem in the implementation and β_{22} in the CARGPR(1, 2) model (Model 3) is insignificant, showing that the basic CARGPR(1, 1) model describes the market persistence effect well. Moreover both the BIC and the DIC of Model 3 show no improvement either. Hence the basic CARGPR(1, 1) model is adopted hereafter.

To allow for the leverage effect, Z_t is added to υ_t as an exogenous variable. Moreover, Models 1–3 are restricted to monotone trend data. Fig. 1(a) shows that the monotone increasing trend applies only till the global financial tsunami in October 2008, and decreases thereafter. To allow a flexible trend movement, a ratio function a_t in (12) is adopted in Model 4 and a threshold time effect in Model 5. Moreover, we set $M = 2$ and the range for sampling T_2 to be [610, 630], which covers the period from 19 September 2008 to 17 October 2008 for Model 5.

Lastly, the CARR models in (5) with the covariate Z_t in the mean μ_t using exponential (Model 6) and Weibull (Model 7) distributions and the Bayesian approach are also fitted for model comparison. Table 3 reports the posterior mean and the posterior standard error (in italics) of the model parameters, together with two model assessment criteria for Models 1–7.

5.2. Model assessment

To compare Models 1–7, the Bayes factor, BIC, and DIC (Spiegelhalter et al., 2002) are often used in Bayesian analysis. However, the former is often commented on as being too difficult to calculate, especially for models that involve many

Table 3
Parameter estimates, standard errors in italics, BIC and DIC for the AORD daily range data.

Model	Dist.	Type	T	$\beta_{\mu 0}$	$\beta_{\mu 11}$	$\beta_{\mu 21}$	$\beta_{\mu 22}$	$\beta_{\mu 31}$	a or β_{a0}	β_{a1}	τ^2	ν or α	BIC	DIC
CARGPR														
M1	LN	(1, 1)	–	−0.0234	0.7644	0.1879	–	–	0.9983	–	0.1762	–	1121	1098
			–	*0.0118*	*0.0356*	*0.0253*	–	–	*0.0005*	–	*0.0090*	–		
M2	LT	(1, 1)	–	−0.0192	0.7878	0.1735	–	–	0.9984	–	0.1568	19.26	1125	1095
			–	*0.0095*	*0.0263*	*0.0202*	–	–	*0.0005*	–	*0.0103*	*5.93*		
M2.1	LT	(1, 1)	–	0.0027	0.8080	0.1776	–	–	–	–	0.1570	17.96	1128	1102
($a = 1$)			–	*0.0031*	*0.0199*	*0.0182*	–	–	–	–	*0.0104*	*5.97*		
M3	LT	(1, 2)	–	−0.0222	0.7806	0.1661	0.0110	–	0.9983	–	0.1566	20.56	1132	1097
			–	*0.0103*	*0.0328*	*0.0415*	*0.0521*	–	*0.0006*	–	*0.0104*	*5.40*		
M4	LT	(1, 1)	–	−0.0129	0.8269	0.1124	–	−0.0536	0.0038	−0.0008	0.1441	18.63	1074	1047
			–	*0.0061*	*0.0259*	*0.0197*	–	*0.0067*	*0.0005*	*0.0001*	*0.0094*	*5.91*		
M5	LT	(1, 1)	1.00	−0.0225	0.8437	0.0972	–	−0.0765	0.9988	–	0.1490	21.56	1101	1037
			–	*0.0084*	*0.0311*	*0.0233*	–	*0.0089*	*0.0003*	–	*0.0096*	*4.86*		
	LT	(1, 1)	622	1.0530	0.1960	−0.0734	–	−0.0300	1.0070	–	0.1144	18.44		
			5.27	*0.3328*	*0.2279*	*0.0875*	–	*0.0135*	*0.0007*	–	*0.0153*	*6.29*		
CARR														
M6	Exp	(1, 1)	–	0.1571	0.6075	0.2820	–	−0.1261	–	–	–	–	1949	1930
			–	*0.0410*	*0.0733*	*0.0615*	–	*0.0365*	–	–	–	–		
M7	Wei	(1, 1)	–	0.1883	0.5858	0.2766	–	−0.1378	–	–	–	2.1860	1339	1316
			–	*0.0156*	*0.0286*	*0.0262*	–	*0.0181*	–	–	–	*0.0521*		

random effects, large numbers of unknowns, or improper priors (Ntzoufras, 2009). Alternately, the BIC and DIC defined as

$$\text{BIC} = -2\ln f(\mathbf{y}|\boldsymbol{\theta}) + p\ln n, \qquad (16)$$
$$\text{and} \quad \text{DIC} = \overline{D(\boldsymbol{\theta})} + p_D,$$

respectively, are adopted to approximate the Bayes factor. Both criteria contain two components: a measure of model fit and a penalty for model complexity, where $f(\mathbf{y}|\boldsymbol{\theta})$ is the likelihood function, $\overline{D(\boldsymbol{\theta})} = E_{\theta|y}[D(\boldsymbol{\theta})]$ is the posterior expectation of deviance, and p_D is the effective number of parameters defined as the difference between the posterior mean of deviance and the deviance evaluated at the posterior mean of parameters; that is,

$$p_D = E_{\theta|y}(D(\boldsymbol{\theta})) - D(E_{\theta|y}(\boldsymbol{\theta})) = \overline{D(\boldsymbol{\theta})} - D(\bar{\boldsymbol{\theta}}).$$

Clearly, the model with the smallest BIC and/or DIC values is preferred. The BIC and DIC values for Models 1–7 are presented in Table 3.

Moreover, five more measures, namely the root mean squared error (RMS), mean absolute error (MAE), coverage percentage (CP), width of the 95% confidence interval (CI) for $E(X_t)$ (CI(EX)) and width of the 95% CI for X_t (CI(X)), are defined as

$$\text{RMS}_{ih} = \left[\frac{1}{n_h}\sum_{t=1}^{n_h}(MR_{t+s_h,i} - \hat{X}_{t+s_h})^2\right]^{1/2},$$
$$\text{MAE}_{ih} = \frac{1}{n_h}\sum_{t=1}^{n_h}|MR_{t+s_h,i} - \hat{X}_{t+s_h}|,$$
$$\text{CP}_h = \frac{1}{n_h}\sum_{t=1}^{n_h} I(X_{t+s_h} \in (\text{CI}_{X_{t+s_h},\text{low}}, \text{CI}_{X_{t+s_h},\text{up}})),$$
$$\text{CI}(EX)_h = \frac{1}{n_h}\sum_{t=1}^{n_h}(\text{CI}_{E(X_{t+s_h}),\text{up}} - \text{CI}_{E(X_{t+s_h}),\text{low}}),$$
$$\text{CI}(X)_h = \frac{1}{n_h}\sum_{t=1}^{n_h}(\text{CI}_{X_{t+s_h},\text{up}} - \text{CI}_{X_{t+s_h},\text{low}}),$$

where $h = 0$ indicates the in-sample estimation with start $s_0 = 0$, $h = 1$ indicates the out-of-sample forecast with start $s_1 = n_1 = 763$, $i = 1$ indicates the measure of range $MR_{t,1} = X_t$, and $i = 2$ indicates $MR_{t,2} = |Z_t|$ as a proxy of X_t (Chou, 2005). The standardized variables when $X_t \sim \text{LN}(\omega_t, \varsigma_t)$, where $\omega_t = \upsilon_t - \ln a_t^{t-1}$ and $\varsigma_t^2 = \frac{\tau^2}{\lambda_t}$, and when $X_t \sim W(\psi_t, \alpha)$, where $\psi_t = \mu_t/\Gamma\left(1 + \frac{1}{\alpha}\right)$, are

$$S_{\text{LT},t} = \frac{\ln(X_t) - \omega_t}{\varsigma_t} \sim N(0, 1) \quad \text{and} \quad S_{W,t} = (X_t/\psi_t)^\alpha \sim \text{Exp}(1), \qquad (17)$$

Table 4
In-sample and out-of-sample model assessment for Models 4–7.

	In-sample model-fit criteria ($n = 763$)										Out-of-sample forecasting criteria ($n_1 = 50$)						
	RMS_1	MAE_1	RMS_2	MAE_2	CP	CI(EX)	CI(X)	Q_{12}	W	JB	RMS_1	MAE_1	RMS_2	MAE_2	CP	CI(EX)	CI(X)
M4	0.684	0.455	1.001	0.750	0.965	0.189	2.161	12.1	0.32	14.1	0.859	0.712	1.277	1.095	0.820	1.194	2.961
M5	*0.649*	*0.442*	*0.953*	*0.736*	*0.963*	0.284	*2.069*	*5.91*	*0.11*[a]	*6.40*	*0.521*	*0.383*	*0.800*	*0.659*	*0.940*	*0.419*	*1.764*
M6	0.700	0.485	0.994	0.763	0.996	0.362	5.265	21.2	23.7	–	0.577	0.664	0.878	0.851	1.000	2.359	4.618
M7	0.707	0.490	0.988	0.759	0.972	*0.165*	2.616	19.4	1.73	–	0.604	0.483	0.922	0.770	0.980	1.382	2.472

[a] p-value > 0.05, while the p-values for other test statistics are all <0.05.

respectively. Hence the corresponding 95% CIs ($\mathrm{CI}_{X_t,\mathrm{low}}$, $\mathrm{CI}_{X_t,\mathrm{up}}$) for X_t are

$$\left(\exp(\omega_t + \Phi^{-1}(0.025)\varsigma_t), \exp(\omega_t + \Phi^{-1}(0.975)\varsigma_t)\right), \tag{18}$$

and

$$\left([-\ln(0.975)]^{1/\alpha}\psi_t, [-\ln(0.025)]^{1/\alpha}\psi_t\right), \tag{19}$$

respectively, where $\Phi(\cdot)$ is the standard normal distribution function. On the other hand, the CIs for $E(X_t)$ are obtained from the 2.5 and 97.5 percentiles of the posterior sample for $E(X_t)$, where $E(X_t)$ is given by (2) and (5), respectively, for the log-t and Weibull distributions. The first three criteria measure the accuracy of the model while the last two measure the precision of the CIs. Models with smaller values of these criteria except CP_h are preferred. For CP_h, it should be close to 95%. Table 4 reports these measures together with the three test statistics Q_{12}, W and JB. The standardized variables $S_{\mathrm{LT},t}$ and $S_{E,t}$ are used to test the log-t and Weibull data distributions, respectively.

5.3. Numerical results

The results in Table 3 show that the parameter estimates are qualitatively consistent across the models. In particular, the ratio a for Models 1–3 is less than 1 and significant, showing a general monotone increasing trend. Since both $\beta_{\mu 11}, \beta_{\mu 21} > 0$ and are significant in all models, a persistence effect is present in the data. Moreover, τ^2 decreases across Models 1–5, showing an increase in model robustness while the number of degrees of freedom is around 20, indicating a moderate tail effect. Since Model 2 shows a better model fit than Model 2.1 ($a = 1$) according to both the BIC and DIC (1125 versus 1128 for the BIC and 1095 versus 1102 for the DIC), after allowing for model complexity, the superiority of the CARGPR model in allowing trend movement is clear.

For Models 4 and 5, as $\beta_{\mu 31}$ is significant and negative, a leverage effect is present in the data. Parameters β_{a0} and β_{a1} in Model 4 show that a_t changes from greater than 1 to less than 1, indicating a mild and short decreasing trend followed by an increasing trend thereafter. Moreover, the substantially larger DIC for Models 1 and 4 with a gamma distribution (1223 and 1167 for the BIC and 1200 and 1130 for the DIC) supports the assertion in Section 2.1 that lognormal and log-t distributions give better fits than a gamma distribution. For Model 5, the ratios a_1 and a_2 show an increasing trend before 7 October 2008 ($t = 622$) and a decreasing trend thereafter. The market volatility increases sharply from September 2008 to the maximum on 7 October 2008, during which the market price dropped continuously. The trends of the mean, $E(X_t)$ in (2) and (14), variance, $\mathrm{Var}(X_t)$ in (3) and (15), and the ratio $1/a_t^{t-1}$ for Models 4 and 5, as plotted in Fig. 2(a) and Fig. 3(a), respectively, show that the mean and variance capture the volatility clustering well. The 95% CIs for X_t in (18) are displayed in Fig. 2(b) and Fig. 3(b).

The coverage percentages (CPs), 96.5% and 96.3%, are reasonably close to 95%. Because there is no obvious volatility clustering after 7 October 2008, no outliers are detected, and $\beta_{\mu 112}$ and $\beta_{\mu 212}$ are both insignificant in Model 5, leading to a rather smooth trend after 7 October 2008. Obviously the significance of $\beta_{\mu 11}$ and $\beta_{\mu 21}$ for other models with a monotone increasing trend is due to the volatility clustering before 7 October 2008. Moreover, even though there is no stationary constraint for Y_t, the sum of parameters in the mean function υ_t in (9) is less than 1 for all the models reported in Table 3. This stationary condition is only violated by the second-stage model in Model 5.

The results for Models 6 and 7 are qualitatively the same as for Models 1–5. The shape parameter α is estimated to be 2.186 for the Weibull distribution. Using both the BIC and DIC, the CARR model using an exponential distribution (Model 6) is far from satisfactory but the model using the Weibull distribution (Model 7) is still no better than any of the CARGPR models (Models 1–5), because the CARGPR models accommodate the trend effect and adopt the more robust log-t distribution. The trends of the mean and variance for Model 7 are plotted in Fig. 4(a) and the 95% CI for X_t in (19) is plotted in Fig. 4(b). Again, the mean and variance capture the volatility clustering well, but the lower bound of the CI is very close to zero, revealing the characteristic of the Weibull distribution with higher density around zero when α is small.

Tests using Q_{12}, W and JB show that all the standardized residuals S_{it}, $i = LT, W$ in (17) are non-random and do not follow the hypothesized distribution except Model 5. Fig. 5 displays the histograms of $S_{\mathrm{LT},t}$ for Models 4 and 5 and $S_{W,t}$ for Model 7 superimposed on their hypothesized density functions. Again, the distribution of $S_{\mathrm{LT},t}$ from Model 5 is closest to the hypothesized standard normal density. Hence Model 5 is preferable to the other models.

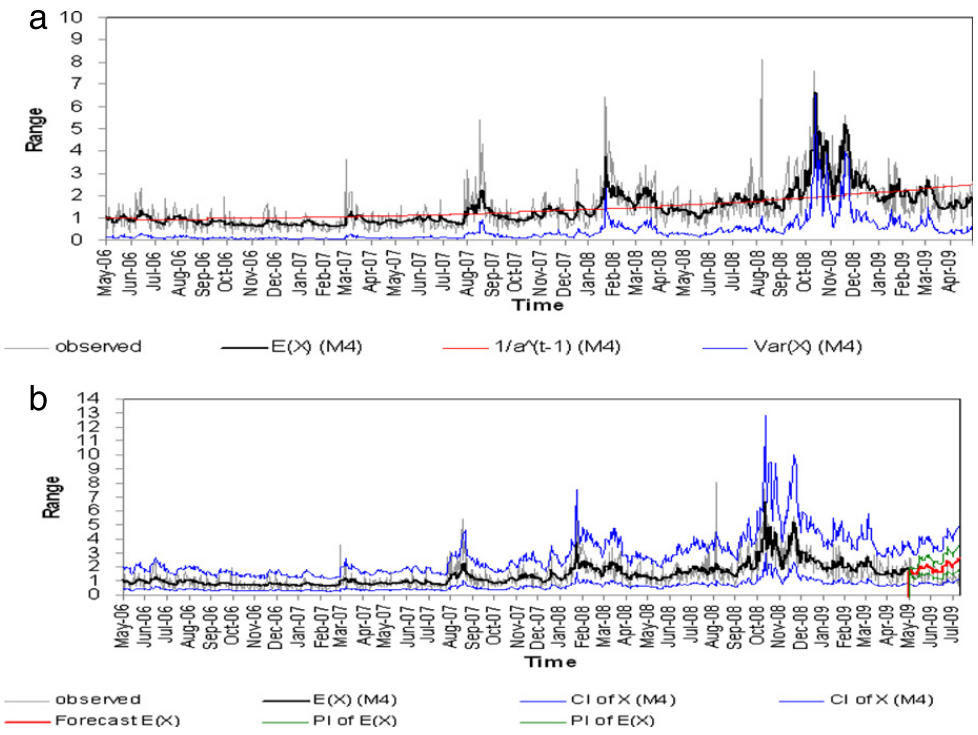

Fig. 2. (a) Trends of the mean, variance, and ratio. (b) Observed, expected, 95% CI of X_t, and 95% PI of the forecast $E(X_t)$ using Model 4.

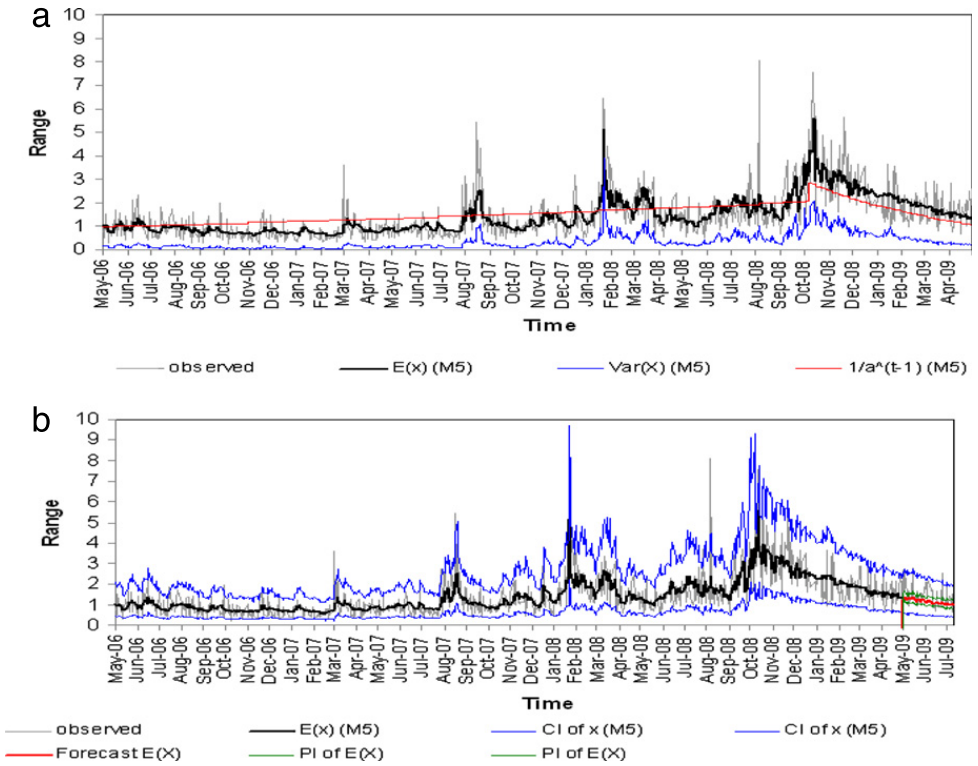

Fig. 3. (a) Trends of the mean, variance, and ratio. (b) Observed, expected, 95% CI of X_t, and 95% PI of the forecast $E(X_t)$ using Model 5.

Tables 3 and 4 show that Model 5 outperforms Models 4 and 7 across all in-sample model assessment criteria, BIC and DIC. The only two exceptions are the shorter CI(EX) for Model 7 and the slightly lower BIC for Model 4. The latter is due to the heavy penalty term in the BIC ($-2\ln f(\mathbf{y}|\boldsymbol{\theta})$ in (16) is 1021 and 1001, respectively, for Models 4 and 5). Fig. 6 compares the mean and 95% CI of X_t between Models 5 and 7. The CI estimate is clearly shorter for Model 5, giving more precise fitted values.

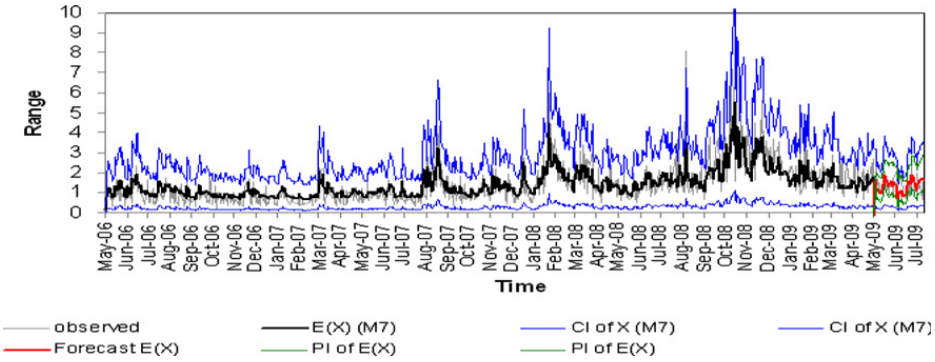

Fig. 4. (a) Trends of the mean and variance. (b) Observed, expected, 95% CI of X_t, and 95% PI of the forecast $E(X_t)$ using Model 7.

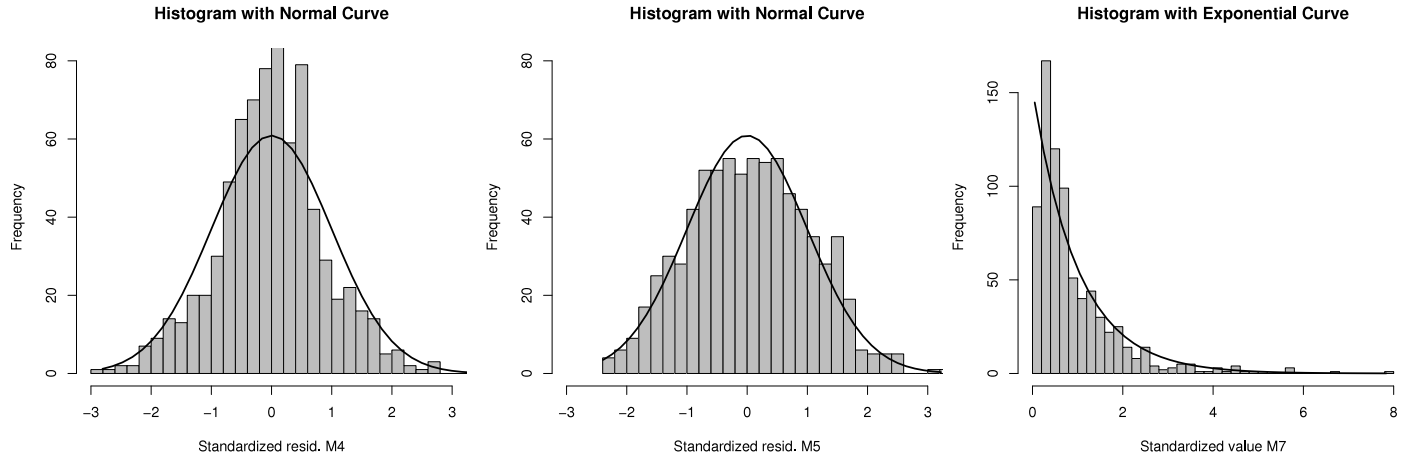

Fig. 5. Comparison of observed and hypothesized distributions for standardized X_t between Models 4, 5, and 7.

5.4. Forecasting

Models 4, 5, and 7 are used to forecast $n_1 = 50$ daily ranges from 1 May 2009 to 10 July 2009. These values are labelled as $x_{764}, \ldots, x_{813}$. The joint predictive distribution is given by

$$f(x_{764}, \ldots, x_{813}|\mathbf{x}) = \int \prod_{t=764}^{813} f_{\mathrm{LT}}\left(x|\upsilon_t - \ln(a_2^{t-T_2}), \tau^2, \alpha\right) f(\boldsymbol{\theta}|\mathbf{x})\mathrm{d}\boldsymbol{\theta},$$

where $\mathbf{x}$ denotes the vector of 763 observed daily ranges, $\boldsymbol{\theta}$ is the vector of model parameters, and $f_{\mathrm{LT}}(x|b, c, d)$ is the density function of the log-t distribution with location b, scale c, and number of degrees of freedom d. Given a set of parameter values, $\boldsymbol{\theta}^{(k)}$, at the k-th iteration of the Gibbs sampling output, a set of predicted values can be simulated successively from

$$x_t|x_{t-1}, \boldsymbol{\theta}^{(k)} \sim LT\left(\upsilon_t^{(k)} - \ln[(a_2^{(m)})^{t-T_2^{(k)}}], (\tau^{2(k)})^2, \alpha^{(k)}\right),$$

where $t = 764, \ldots, 813$. The random variate generation from the log-t distribution can be done via its scale mixture of normal representation. The posterior means and the corresponding 95% Bayesian prediction intervals of the predicted daily ranges are displayed in Fig. 2(b), Fig. 3(b), and Fig. 4(b) for Models 4, 5, and 7, respectively, and are also given in Fig. 7 for clarity.

In general, the forecasting error increases across the forecast period due to the accumulated uncertainty. However, Model 5 has a substantially lower forecasting error, and hence a much shorter prediction interval, because of the fitted decreasing trend and the insignificance of $\beta_{\mu 112}$ and $\beta_{\mu 212}$, leading to a relatively constant υ_t in the expression of $E(X_t)$. Since the observed X_t shows a gentle decline during the forecast period, the forecast using Model 4 with a fitted increasing trend is less satisfactory. Model 5 still outperforms Model 7 across all out-of-sample forecasting criteria in Table 4.

5.5. Outlier diagnostic

It is well known that Student's t-distribution provides a robust inference by downweighting the distorting effects of outliers. Expressing the t-distribution as an SMN distribution, Choy and Smith (1997) was the first to propose performing outlier diagnostics using the scale mixture variable λ in the SMN representation. An outlier is associated with a large value

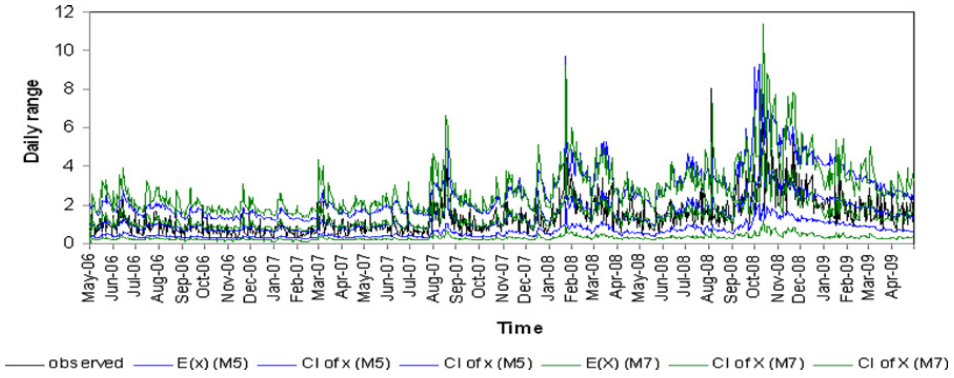

Fig. 6. Comparison of the expected and 95% CI of X between Models 5 and 7.

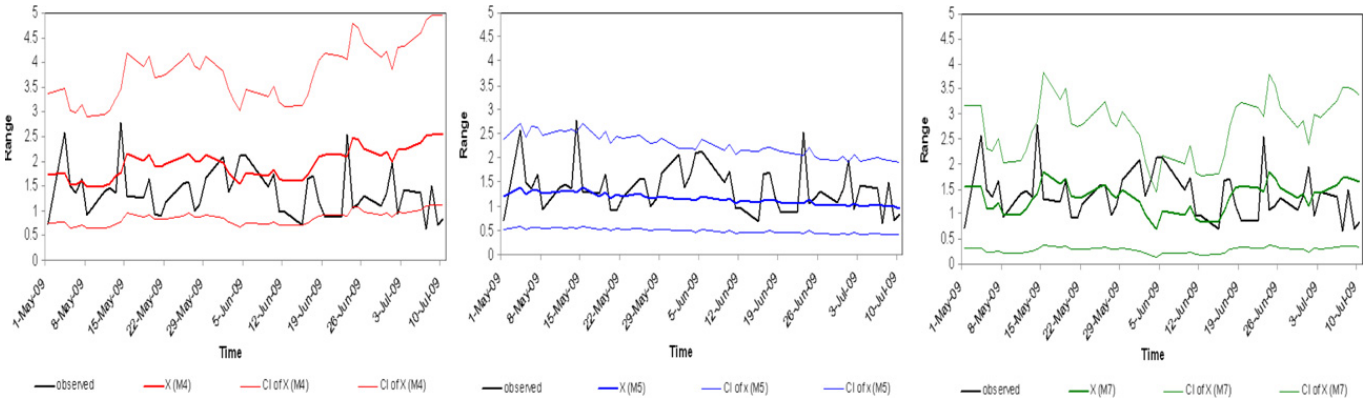

Fig. 7. Forecast range and 95% CI of the forecast range using Models 4, 5, and 7 respectively.

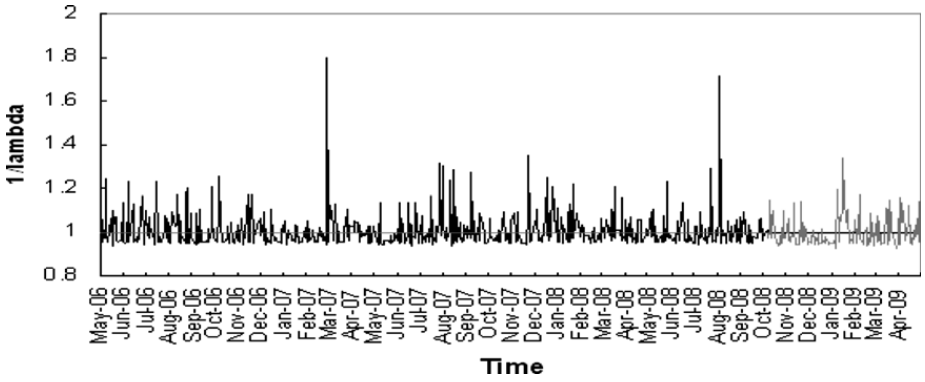

Fig. 8. Reciprocal of lambda $1/\lambda_t$ in outlier diagnostic using Model 5.

of $1/\lambda$ which inflates the variance of the corresponding normal distribution to accommodate the outlier. Therefore, the extremeness of observations is closely associated with the magnitude of λ.

Fig. 8 plots the reciprocal $1/\lambda_t$ in Model 5 across time. From the figure, two outliers, on 28 February 2007 and 5 August 2008, are detected, because their variances are inflated nearly twice as much as the variances at other time points. Table 5 reports the reciprocal $1/\lambda_t$, the observed value X_t, the mean $E(X_t)$, and the 95% CI of X_t for the two outliers. As the CIs do not contain X_t, the daily ranges on 28 February 2007 and 5 August 2008 are indeed outlying.

6. Conclusions

This paper extends the GP model to the CARGPR model for range data to describe the persistence dynamics in the mean function μ_t. The CARR-type range model is simpler than the GARCH and SV models, but yet it was shown to provide a superior volatility forecast. The performance of the proposed CARGPR model was shown to exceed that of the CARR-type models in four aspects: the accommodation of trend movement using an explicit ratio parameter or function, the adoption of heavy-tailed distributions such as the log-t distribution to describe different tail behavior, the use of the Bayesian approach via the Bayesian software WinBUGS to simplify the model implementation for non-experts, and, lastly, the representation of

Table 5
Summary information for the outliers in Model 5.

t	Date	$1/\lambda_t$	X_t	$E(X_t)$	95% CI of X_t
214	February 28 2007	1.8021	3.5846	0.7378	(0.2455, 1.8676)
577	August 5 2008	1.7176	8.0839	1.8200	(0.6091, 4.4166)

the t-distribution in an SMN representation to facilitate the MCMC algorithm in the Bayesian simulation and enable outlier detection. The simulation study shows that the CARGPR model provides highly accurate parameter estimates, particularly when the sample size is large. In the empirical study using the AORD daily range data, the CARGPR model achieves a better model fit and provides a sharper volatility forecast, confirming the superiority of the CARGPR model. Range data is sensitive to outliers. For the CARGPR model, the product of the random jump indicator and the random jump size may be added to υ_t to capture the spikes in a highly volatile financial times series. We believe that this will be a promising extension for the CARGPR model. On the other hand, the choice of volatility measure from high-frequency data is of utmost importance and should not be limited to the range measure. Perhaps the "realized volatility" of Fleming et al. (2003) may capture the dynamics of intra-day price movement better, provide more precise estimates of volatility, and hence offer a better choice than the range data for the CARGPR model. This is a worthwhile direction for further research. Moreover, an extension to a multivariate CARGPR model that simultaneously models the mean and the ratio is also a promising issue for future study.

Acknowledgements

The authors would like to thank the editor, an associate editor, and two anonymous referees for their constructive comments that helped to substantially improve the quality of the paper. Part of the work of Jennifer S.K. Chan, undertaken during a research visit to Feng Chia University (FCU), was funded by a grant from National Science Council. Cathy Chen is supported by NSC of Taiwan grant NSC96-2118-M-035-002-MY3.

Appendix

The full conditional distributions for the parameters in Model 5 are derived to facilitate the Gibbs sampling algorithm. Define $\mathbf{x} = (x_1, \ldots, x_n)$; $\boldsymbol{\beta}_m = (\beta_{\mu 0m}, \beta_{\mu 1m}, \beta_{\mu 2m}, \beta_{\mu 3m})$ (drop the redundant subscript j), $m = 1, 2$ indicate the GP model before and after the threshold $T_2 (T_1 = 1)$ respectively; $(s_1, e_1) = (1, T_2 - 1)$ and $(s_2, e_2) = (T_2, n)$ indicate the start and end times of the two GPs respectively; $\boldsymbol{\lambda} = (\lambda_1, \ldots, \lambda_n)$ and $\boldsymbol{\lambda}_{-t} = (\lambda_1, \ldots, \lambda_{t-1}, \lambda_{t+1}, \ldots, \lambda_n)$. The Gibbs sampler draws realizations iteratively from the following conditional distributions:

$$
\begin{aligned}
&f\left(\ln(a_m)|\boldsymbol{\beta}_m, \tau_m^2, \boldsymbol{\lambda}, \nu_m, T_2, \mathbf{x}\right) = N\left(-\frac{\sum_{t=s_m+1}^{e_m}(t-T_m)\left[\ln(x_t)-\upsilon_t\right]\tau_m^2}{\sum_{t=s_m+1}^{e_m}\lambda_t(t-T_m)^2}, \frac{\tau_m^2}{\sum_{t=s_m+1}^{e_m}\lambda_t(t-T_m)^2}\right) I(\ln(0.95), \ln(1.05))\\
&f\left(\boldsymbol{\beta}_m|a_m, \tau_m^2, \boldsymbol{\lambda}, \nu_m, T_2, \mathbf{x}\right) \propto \prod_{t=s_m}^{e_m} N\left(\ln(x_t)\,\middle|\, \upsilon_t - (t-T_m)\ln(a_m), \frac{\tau_m^2}{\lambda_t}\right)\\
&f\left(\tau_m^2|a_m, \boldsymbol{\beta}_m, \boldsymbol{\lambda}, \nu_m, T_2, \mathbf{x}\right) = IG\left(\frac{n}{2}, \frac{1}{2}\sum_{t=1}^{n}\lambda_t\left[\ln(x_t)-\upsilon_t+(t-T_m)\ln(a_m)\right]^2\right)\\
&f\left(\lambda_t|a_{\mathrm{m}}, \boldsymbol{\beta}_{\mathrm{m}}, \boldsymbol{\lambda}_{-t}, \nu_{\mathrm{m}}, T_2, \mathbf{x}\right) = IG\left(\frac{\nu_{\mathrm{m}}+1}{2}, \frac{\nu_{\mathrm{m}}}{2}+\frac{1}{2\tau_{\mathrm{m}}^2}\left[\ln(x_t)-\upsilon_t+(t-T_{\mathrm{m}})\ln(a_{\mathrm{m}})\right]^2\right) \qquad (20)\\
&f\left(\nu_m|a_m, \boldsymbol{\beta}_m, \boldsymbol{\lambda}, T_2, \mathbf{x}\right) \propto \prod_{t=s_m}^{e_m}\frac{1}{\nu_m}G\left(\lambda_t\,\middle|\,\frac{\nu_m}{2}, \frac{\nu_m}{2}\right)\\
&f\left(T_2|a_m, \boldsymbol{\beta}_m, \boldsymbol{\lambda}_m, \nu_m, \mathbf{x}\right) \propto \text{Multinomial}\,(\pi_{610}, \ldots, \pi_{630}),
\end{aligned}
$$

where $\pi_k = \dfrac{\prod_{t=1}^{k-1} f_{\mathrm{LN}}\left(x_t|\upsilon_t-\ln(a_1^{t-1}), \frac{\tau_1^2}{\lambda_t}\right)\prod_{t=k}^{n} f_{\mathrm{LN}}\left(x_t|\upsilon_t-\ln(a_2^{t-k}), \frac{\tau_2^2}{\lambda_t}\right)}{\sum_{k'=610}^{630}\left[\prod_{t=1}^{k'-1} f_{\mathrm{LN}}\left(x_t|\upsilon_t-\ln(a_1^{t-1}), \frac{\tau_1^2}{\lambda_t}\right)\prod_{t=k'}^{n} f_{\mathrm{LN}}\left(x_t|\upsilon_t-\ln(a_2^{t-k'}), \frac{\tau_2^2}{\lambda_t}\right)\right]}$, $k = 610, \ldots, 630$, $m = 1, 2$, $\mathrm{m} = 1I(t < T_2) + 2I(t \geq T_2)$ in (20), and υ_t is given by (13). The algorithm of Robert (1995) can be used to simulate the random variate $\ln(a_m)$ from a truncated normal distribution. The conditional distributions of $\boldsymbol{\beta}_m$ and ν_m are non-standard, and random variate generations from these full conditional distributions can be performed using Metropolis–Hastings algorithms.

References

Alizadeh, S., Brandt, M.W., Diebold, F.X., 2002. Range-based estimation of stochastic volatility models or exchange rate dynamics are more interesting than you think. Journal of Finance 57, 1047–1092.

Andersen, T., Bollerslev, T., 1998. Answering the skeptics: yes, standard volatility models do provide accurate forecasts. International Economic Review 39, 885–905.

Andrews, D.F., Mallows, C.L., 1974. Scale mixtures of normal distributions. Journal of the Royal Statistical Society. Series B 36, 99–102.

Bollerslev, T., 1986. Generalized autoregressive conditional heteroskedasticity. Journal of Econometrics 31, 307–328.

Chan, J.S.K., Lam, Y., Leung, D.Y.P., 2004. Statistical inference for geometric processes with gamma distributions. Computational Statistics & Data Analysis 47, 565–581.

Chan, J.S.K., Yu, P.L.H., Lam, Y., Ho, A.P.K., 2006. Modelling SARS data using threshold geometric process. Statistics in Medicine 25, 1826–1839.

Chen, C.W.S., Gerlach, R., Choy, S.T.B., Lin, C., 2010. Estimation and inference for exponential smooth transition nonlinear volatility models. Journal of Statistical Planning and Inference 140, 719–733.

Chen, C.W.S., Gerlach, R., Lin, E.M.H., 2008. Volatility forecast using threshold heteroskedastic models of the intra-day range. Computational Statistics & Data Analysis 52, 2990–3010 On Statistical & Computational Methods in Finance.

Chen, C.W.S., Lee, J.C., 1995. Bayesian inference of threshold autoregressive models. Journal of Time Series Analysis 16, 483–492.

Chiu, H.C., Wang, D., 2006. Using conditional autoregressive range model to forecast volatility of the stock indices. In: Proceedings of Joint Conference on Information Science 2006. Atlantis Press.

Chou, R., 2005. Forecasting financial volatilities with extreme values: the conditional autoregressive range (CARR) model. Journal of Money, Credit and Banking 37, 561–582.

Choy, S.T.B., Chan, J.S.K., 2008. Scale mixtures distributions in statistical modelling. Australian and New Zealand Journal of Statistics 50, 135–146.

Choy, S.T.B., Smith, A.F.M., 1997. Hierarchical models with scale mixtures of normal distribution. TEST 6, 205–221.

Feller, W., 1949. Fluctuation theory of recurrent events. Transactions of the American Mathematical Society 67, 98–119.

Fleming, J., Kirby, C., Ostdiek, B., 2003. The economic value of volatility timing using realized volatility. Journal of Financial Economics 67, 473–509.

Geweke, J., Terui, N., 1993. Bayesian threshold autoregressive models for nonlinear time series. Journal of Time Series Analysis 14, 441–454.

Gilks, W.R., Richardson, S., Spiegelhalter, D.J., 1996. Markov Chain Monte Carlo in Practice. Chapman and Hall, UK.

Hastings, W.K., 1970. Monte Carlo sampling methods using Markov chains and their applications. Biometrika 57, 97–109.

Hull, J., White, A., 1987. The pricing of options on assets with stochastic volatility. Journal of Finance 42, 281–300.

Lam, Y., 1988. Geometric process and replacement problem. Acta Mathematicae Applicatae Sinica 4, 366–377.

Lam, Y., Chan, J.S.K., 1998. Statistical inference for geometric processes with lognormal distribution. Computational Statistics & Data Analysis 27, 99–112.

Metropolis, N., Rosenbluth, A.W., Rosenbluth, M.N., Teller, A.H., 1953. Equations of state calculations by fast computing machines. Journal of Chemical Physics 21, 1087–1091.

Ntzoufras, I., 2009. Bayesian Modeling Using WinBUGS. Wiley, New Jersey, pp. 389–390.

Parkinson, M., 1980. The extreme value method for estimating the variance of the rate of return. Journal of Business 53, 61–65.

Robert, C.P., 1995. Simulation of truncated normal variables. Statistics and Computing 5, 121–125.

Smith, A.F.M., Roberts, G.O., 1993. Bayesian computation via the Gibbs sampler and related Markov chain Monte Carlo methods. Journal of the Royal Statistical Society. Series B 55, 3–23.

Spiegelhalter, D., Best, N.G., Carlin, B.P., van der Linde, A., 2002. Bayesian measures of model complexity and fit. Journal of the Royal Statistical Society. Series B 64, 583–639.

Spiegelhalter, D., Thomas, A., Best, N.G., Lunn, D., 2004. Bayesian inference using Gibbs sampling for Window version (WinBUGS) version 1.4.1. The University of Cambridge. www.mrc-bsu.cam.ac.uk/bugs/welcome.shtml.

Tiwari, R.C., Cronin, K.A., Davis, W., Feuer, E.J., 2005. Bayesian model selection for joint point regression with application to age-adjusted cancer rates. Applied Statistics 54, 919–939.

Tong, H., 1990. Nonlinear Time Series: A Dynamic System Approach. Oxford University Press, Oxford, UK.

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