ISSN : 0974 - 7435 Volume 10 Issue 23

2014

BioTechnology
An Indian Journal

FULL PAPER

BTAIJ, 10(23), 2014 [14526-14533]

Bayesian estimates of VaR using GM(1,1)-POT-VaR model

Wang Jin[1]*, Wang Ruiqing[2]
[1]College of Physics Engineering, Anyang Normal University, Anyang, (CHINA)
[2]School of Computer Engineering, Anyang Normal University, Anyang, (CHINA)
E-mail: aywx123@163.com

ABSTRACT

Electricity price connotes a grey system, due to uncertainty and incomplete information for partial external or inner parameters. A two-stage method based on grey system and extreme value theory is proposed to estimate the value-at-risk. In stage one, to capture the dependencies, seasonalities and volatility-clustering, a gray GM(1,1) model is utilized to filter electricity price series. In this way, an approximately independently and identically distributed residual series with better statistical properties is acquired. In stage two, a peaks over threshold method is adopted to explicitly model the tails of the residuals of GM(1,1) model, and accurate estimates of electricity market value-at-risk can be produced. For conquering the difficulty of lacking observed data over threshold, Bayesian estimation based on Markov Chain Monte Carlo simulation is used to estimate the parameters of peaks over threshold model. The empirical analysis shows that the proposed model can be rapidly reflect the most recent and relevant changes of electricity prices and produce accurate forecasts of value-at-risk at all confidence levels, and the computational cost is far less than the existing two-stage value-at-risk estimating models, further improving the ability of risk management for electricity market participants.

KEYWORDS

Bayesian estimation; Value-at-risk; Grey GM(1,1) model; Extreme value theory.

INTRODUCTION

Value-at-risk (VaR) is a risk management tool to quantify the level of risk exposure in advance and has become one of the most popular risk measurement tools in practice. With VaR as the risk measure, the purchasing risk of electric utility is calculated using a normal distribution based Delta model [1]. With assumption that the probability distribution of electricity price is normal, the impacts of different bidding strategies on the selling risk for generation companies has been analyzed based on Monte Carlo simulation, the results show that the minimum risk bidding strategy is the one based on marginal cost [2]. By introducing capacity sufficient rate and must-run rate as exogenous explanatory variables to depict the generators' market power and the supply-demand relationship, a GARCH model with Gaussian distribution innovations (N-GARCH) has been used to assess the price of volatility risk in electricity markets [3]. Considering that N-GARCH based VaR calculating model cannot effectively address the leptokurtosis and heavy-tailed phenomenon in the data of profit and loss, a re-sampling method based on a bias-correction step and the bootstrap has been developed, further improving the VaR forecasting accuracy of the N-GARCH model [4]. By utilizing Gram-Charlier series expansion of normal and student-t distribution to depict the residuals distribution of ARMAX-GARCH model, an estimating model of VaR has been proposed, showing that the model with Gram-Charlier series expansion of normal density function can rapidly reflect the recent and relevant changes of electricity prices and produce accurate forecasts of VaR at all confidence levels [5].

With GARCH-based model, the impacts of probability distribution assumption for residuals on VaR estimation accuracy are analyzed for normal, student-t, skewed student-t and general error distribution (GED), showing that the accuracy and stability of estimates of VaR are heavily dependent on the selection of probability distribution for innovations [6]. Extreme value theory (EVT) provides a firm theoretical foundation to study the asymptotical distribution of extreme value for order statistics, without assuming the probability distribution for the sample data. EVT allows extrapolation beyond the sample and can accurately describe the behavior of the tails of the real data. Rozario R. estimated the VaR of electricity market using a technique from extreme value theory known as peaks over thresholds (POT), showing that the estimated results perform well for moderate to very high confidence levels (95-90%), but struggle at higher levels (>99%) owing to the extreme clustering and other dependence evident in the data [7]. Bystrom Hans N.E. extended the classic unconditional EVT approach by first filtering the data via GARCH model to capture some of the dependencies in electricity return series, and thereafter applying ordinary EVT techniques. In this way the independently and identically distributed (IID) assumption behind the EVT-based tail-quantile estimator is less likely to be violated, and the better tail estimates can be acquired [8]. Up to now electric power energy cannot be stored economically and therefore the influencing factors such as load, climate, transmission network, installed capacity have an un-tempered effect on electricity prices. In particular, electricity price exhibits considerably richer structure than load curve and has the following characteristics: mean reversion, seasonalities, heteroscedasticities, lepkurtosises and extreme behavior with fast-reverting spikes. To obtain an approximately IID residual series with better statistical properties, an ARMAX-GARCH model with Gram-Charlier series expansion of normal density function and skewed student-t distribution over the error items is used to pre-filter the raw data to capture the dependences of electricity price series, further improving the effectiveness of the VaR estimates via POT model [9,10].

Although the approximately IID residual series can be acquired by using GARCH models to pre-filter the electricity price series, the high non-linearity for the GARCH models leads to very large computational costs and hinders the wide application in practice. Considering that the incomplete and uncertain information for the spot prices is in line with the characteristics of grey variables, a gray system and extreme value theory based two-stage model for estimating VaR is proposed in this paper (referred as GM(1,1)-POT-VaR). In stage one, to acquire the approximately IID residuals with better statistical properties, a gray GM(1,1) model is used to pre-filter the electricity price series. In stage two, an EVT based POT model is employed to explicitly deal with the right tail of the residuals of the GM(1,1), and accurate estimates of VaR in electricity market can be produced. There are several contributions. First, the paper proposes a model that has the potential to generate more accurate quantile estimates. The seasonalities, skewnesses and kurtosises are accommodated via an GM(1,1) specification. In forecasting VaR, EVT is applied to the residuals. Clearly, the proposed combination is a sophisticated approach to forecasting VaR. The second contribution is to acquire an approximately IID residual series with better statistical properties by using a GM(1,1) model. The effectiveness of the VaR estimates via POT model can be further improved. The third contribution is to compare the accuracy of VaR forecasts under the proposed model with a number of conventional approaches. Tail quantiles are estimated under each competing model and the frequency with which realized returns violate these estimates provides an initial measure of model success. The empirical analysis indicates that the GM(1,1)-POT-VaR model can rapidly reflect the most recent and relevant changes of electricity prices and can produce accurate forecasts of VaR at all significance levels. Moreover, the computational costs is far less than the proposed models [7-10], further improving the risk management ability of electricity market participants. These results suggest that the proposed approach is robust and therefore useful.

GRAY GM(1,1) MODEL

The grey model is a modelling method based on the concept of grey generating function and differential fitting, having the advantages that the predicted results can be tested and less original data are needed. Let the observed data series be $X^{(0)} = \{x^{(0)}(k)\}$ and the first-order accumulated generating operation series of $X^{(0)}$ is $X^{(1)} = \{x^{(1)}(k) | x^{(1)}(k) = \sum_{j=1}^{k} x^{(0)}(j)\}$, among them, k=1,2,…,n. Then, the dynamic process of $x^{(1)}(k)$ can be described by the following GM(1,1) model:

$$x^{(0)}(k)+az^{(1)}(k)=u. \qquad (1)$$

where, a and u are the model parameters to be estimated, $z_1^{(1)}(k)=\lambda x_1^{(1)}(k)+(1-\lambda)x_1^{(1)}(k-1)$ $(0\le\lambda\le 1)$ is the background value. In traditional GM(1,1) model the λ is usually taken to be a fixed value 0.5 [11]. Let $\hat{a}=[a,u]^T$, then the estimated values by least squares method is

$$\hat{a}=(B^TB)^{-1}B^TY_N. \qquad (2)$$

in which,

$$Y_N=[x^{(0)}(2),x^{(0)}(3),\dots,x^{(0)}(n)]^T,\ B=\begin{bmatrix}-\lambda x^{(1)}(2)-(1-\lambda)x^{(1)}(1) & 1\\ -\lambda x^{(1)}(3)-(1-\lambda)x^{(1)}(2) & 1\\ \vdots & \\ \lambda x^{(1)}(n)-(1-\lambda)x^{(1)}(n-1) & 1\end{bmatrix}$$

After calibrated $\hat{a}$, the solution to Eq. (1) with initial condition $x^{(1)}(1)=x^{(0)}(1)$ is

$$\hat{x}^{(1)}(k)=\left(x^{(0)}(1)-\frac{u}{a}\right)e^{-a(k-1)}+\frac{u}{a}. \qquad (3)$$

Using the first-order inverse accumulated generating operation of $\hat{x}^{(1)}(k)$, the modeling value $\hat{x}^{(0)}(k)$ can be derived from Eq. (3):

$$\hat{x}^{(0)}(k)=(1-e^{a})(x^{(0)}(1)-u/a)e^{-a(k-1)} \qquad (4)$$

With the operation of electricity market, the new data of electricity price continue to emerge. In order to utilize the rich information contained in the new observed values, the new-information grey model is used in this paper. That is, each new obtained value will be added to the tail of the series, at the same time, the first observed value will be removed from the series. The research has shown that new-information grey model have some advantages such as small data sets required, less computational complexity, objective and reliable forecasted results [11].

EXTREME VALUE THEORY

There exists strong temporal dependence in the electricity price series due to the specific features of electric power. It violates the underlying assumption that the data series to which EVT is applied should be a sequence of IID random variables. In this paper, a two-stage approach, provided by McNeil and Frey, is used to this problem. Firstly, the heteroscedasticities, skewnesses, lepkurtosises and seasonalities of electricity price series are filtered by the GM(1,1) model in section 2 to obtain a nearly IID normalized residual series. In stage two, the EVT framework is applied to the standardized residuals to better capture the heavy-tails and improve the accuracy of VaR estimation [12].

POT model

POT is to model the excess distribution for the IID sample data that exceed a high threshold. Given the distribution function $F_z(z)$ of a random variable Z, the distribution function of values of z above a certain threshold u, $F_u(y)$, is called the conditional excess distribution function and is defined as

$$F_u(y)=\text{Prob}(Z-u\le y\,|\,Z>u)=\frac{F_z(z)-F_z(u)}{1-F_z(u)},\ \forall\, 0\le y\le z_F-u, \qquad (5)$$

where Z is a random variable, u is a given threshold, $y=z-u$ are the excesses and $z_F\le\infty$ is the right endpoint of $F_z(z)$.

The theorem of Balkema-De Haan-Pickands states that for large u, the conditional excess distribution function $F_u(y)$ is well approximated by the generalized Pareto distribution (GPD) $G_{\xi,\beta}(y)$, which is defined as

$$G_{\xi,\beta}(y)=\begin{cases}1-\left(1+\dfrac{\xi}{\beta}y\right)^{1/\xi} & \forall\,\xi\ne 0\\ 1-e^{-y/\beta} & \forall\,\xi=0\end{cases} \qquad (6)$$

for $y\in[0,\infty)$ if $\xi\ge 0$ and $y\in[0,-\beta/\xi]$ if $\xi<0$. ξ is the shape parameter or tail index and $\beta>0$ is the scaling parameter. In general, we cannot fix an upper bound for financial losses and only distributions with shape parameter $\xi>0$ are suited to model fat-tailed distributions. Therefore, we will only discuss the situation of $\xi>0$ in the remainder of this paper.

If T is the total number of observations and T_u the number above the threshold u, the value of $F_z(u)$ can be well approximated by the estimate $(T-T_u)/T$ for sufficiently high u. Replacing $F_u(y)$ by the GPD and $F_z(u)$ by $(T-T_u)/T$, we obtain the estimate of $F_z(z)$ from Eq. (5)

$$\hat{F}_z(z) = 1 - \frac{T_u}{T}\left(1+\frac{\xi}{\beta}(z-u)\right)^{1/\xi} \tag{7}$$

for $z>u$. Inverting Eq. (7) for a given probability p ($p \geq 90\%$), the estimates of the p-th tail quantile for the sample distribution can be gotten,

$$\hat{F}_z^{-1}(p) = u + \frac{\beta}{\xi}\left(\left(Tp/T_u\right)^{-\xi} - 1\right), \tag{8}$$

which is valid for positive excesses, that is $z > u$.

A reasonable threshold u must be chosen to effectively estimate the values of parameters ξ and σ. So far, no automatic algorithm with satisfactory performance for choice of threshold u is available. A popular graphical tool for visually selecting u is the sample mean excess plot defined by the points $(u, e_n(u))$. Let $z_{(1)}>z_{(2)}>\ldots>z_{(T)}$ represent the IID order random variables, $e_n(u)$ can be calculated by

$$e_n(u) = \sum_{i=k}^{n} (z_{(i)} - u) \Big/ (n-k+1), \tag{9}$$

where $k=min\{i|z_{(i)}>u\}$, $n-k+1$ is the number of observations exceeding threshold u [13]. If the GPD provides a good description of the data $e_n(u)$ should be approximately linear in u. So we can select the value that locates at the beginning of the sample mean excess plot which is roughly linear as the suitable threshold. If the sample mean excess plot is upward sloping when $z \geq u$, the distribution of the observations exceeding threshold u is the GPD with positive ξ; If the sample mean excess plot is downward sloping when $z \geq u$, the observations exceeding threshold u follows a distribution with short tails; If the sample mean excess plot is horizontal when $z \geq u$, the distribution of the observations exceeding threshold u is the exponential distribution.

Bayesian estimates of model

Compared with Bayesian estimate, the disadvantages of maximum likelihood estimation (MLE) mainly include: (1) To solve constrained maximization problem by MLE, we face some challenging questions such as convergence and sensitivity of the estimated results to initial values selecting; (2) In practice we are usually interested in the highly nonlinear functions of the parameters, but not the parameters itself. When evaluating their confidence intervals the delta or bootstrap sampling methods have to be used, but these methods are quite time-consuming and difficult to be implemented; (3) It is difficult to prove if the asymptotically optimal conditions is satisfied. And in practice the large sample size is also difficult to be met especially in extreme value estimation; (4) With the regime-switching GARCH or mixture distribution model, it is difficult to determine the number of regimes and mixture distribution. Therefore, in this paper we utilize the Bayesian methods to calculate the parameters values of the POT model to avoid the increasing errors caused by insufficient sample data and the complexity of extreme value optimization by using MLE methods.

From Eq. (6), the likelihood function of the exceedances over the threshold, also called joint conditional density function, can be given by

$$L(y \mid \xi, \beta) = \frac{1}{\beta^{n_u}} \prod_{i=1}^{n_u} \left(1+\frac{\xi}{\beta} y_i\right)^{(1+\xi)/\xi}, \tag{10}$$

where, y_i is the sample observations of the exceedances $Y = \{z-u \mid z>u\}$.

To implement Bayesian estimation, a proper prior should be chosen. The Bayesian factors are very sensitive to the hyperparameters selection of the non-informative prior, particularly the non-informative prior will result in the instability of the posterior distribution and the convergence problem of Gibbs sampling, and therefore an informative prior is suggested [14]. With the assumption that ξ is independent of β, referring to the suggestion of selecting prior distribution in the literature, in this paper the following information prior

$$\xi \sim N(\mu, \sigma^2), \mu > 0, \sigma > 0 \qquad \beta \sim IG(a, b), a > 0, b > 0 \tag{11}$$

is specified, where the hyperparameters μ, σ, a, b can be evaluated by the moments estimation method based on the historical data. The joint posterior density of parameters ξ and β conditional on the exceedances is $\pi(\xi,\beta \mid y) \propto L(y \mid \xi,\beta)p(\xi)p(\beta)$, in which $p(\xi)$ and $p(\beta)$ are the marginal prior distribution for the parameters ξ and β. From the likelihood Eq. (10) and the prior distributions Eq. (11), we can obtain the posterior density distribution of ξ and β conditional on the exceedances:

$$\pi(\xi,\beta|y) \propto \exp\left(-\frac{b}{\beta}-\frac{(\xi-\mu)^2}{2\sigma^2}\right)\beta^{-(\alpha+1+n_u)}\prod_{i=1}^{n_u}\left(1+\frac{\xi}{\beta}y_i\right)^{(1+\xi)/\xi} \quad (12)$$

Under square loss, the Bayesian estimates of the parameters are equal to the posterior mean values, which can be acquired by computing the integral of the posterior distribution. From Eq. (12), it can be seen that the posterior density distribution is a complex two-dimensional nonstandard distribution, and it is difficult to calculate the mean values of the parameters directly. With the help of the Markov Chain Monte Carlo (MCMC) algorithm, a Bayesian calculating method which has developed rapidly in recent years and proved to be effective in practice, the parameters estimates can be derived by implementing computer simulation. The basic concept of MCMC algorithms is to draw from probability distributions based on constructing a Markov chain with stationary distribution $\pi(\xi,\beta|y)$, which collectively forms an approximation of the desired posterior. With these large samples at hand, it makes it possible to conduct various statistical inferences for the posterior density distribution, such as the posterior mean and variance [15].

When direct sampling is difficult, Gibbs sampling is the most widely used MCMC algorithm for obtaining a sequence of observations which approximately form the joint probability distribution of two or more random variables. Among others, the single-site Gibbs sampler is the most attractive because of involving only a single variable sampling and therefore we adopt the single-site Gibbs sampler to calculate the posterior mean values in this paper [15]. The full conditional distributions can be derived from the joint posterior presented in Eq. (12),

$$\begin{aligned}\pi(\xi|\beta,y) &\propto e^{-\frac{(\xi-\mu)^2}{2\sigma^2}}\prod_{i=1}^{n_u}\left(1+\frac{\xi}{\beta}y_i\right)^{\frac{1+\xi}{\xi}}\\ \pi(\beta|\xi,y) &\propto \beta^{-(a+1+n_u)}e^{-\frac{b}{\beta}}\prod_{i=1}^{n_u}\left(1+\frac{\xi}{\beta}y_i\right)^{\frac{1+\xi}{\xi}}\end{aligned} \quad (13)$$

Let $(\xi^{(0)},\beta^{(0)})$ and $(\xi^{(t-1)},\beta^{(t-1)})$ denote the arbitrary starting values and the estimated values of (ξ, β) in the $(t$-1)-th iteration respectively, the t-th iteration can be proceeded by the following steps: (1) Draw $\xi^{(t)}$ from $\pi(\xi|\beta^{(t-1)},y)$; (2) Draw $\beta^{(t)}$ from $\pi(\beta|\xi^{(t-1)},y)$. Now we have completed one iteration of the scheme by visiting each variable and then acquired the sampling sequence $((\xi^{(0)},\beta^{(0)}),(\xi^{(1)},\beta^{(1)}),\cdots(\xi^{(t)},\beta^{(t)}))$ that is a Markov chain and its stationary distribution is $\pi(\xi,\beta|y)$. To diagnose the convergence of the chain, the ergodic means from the well separated subsamples are calculated at a certain distance. If the ergodic means tend to be stability, the procedures of Gibbs sampling can be stopped and the posterior means can be estimated by the mean values of simulated samples.

What needs to be explained is that when drawing from the full conditional distributions, which can not be drawn directly, the acceptance-rejection sampler will be used [15]. If the probability density $p(x)$ can be transformed into $p(x)=c\cdot h(x)\cdot g(x)$, where $0<g(x)\leq 1$, $c\geq 1$ is a constant, $h(x)$ is a easily sampling probability density, the implementation of drawing from $p(x)$ can be proceeded as follows: (1) Draw u^* from uniform distribution $U(0,1)$ and y from $h(y)$; (2) If $u^*\leq g(y)$, set $x=y$, otherwise go to step (1).

When using the acceptance-rejection sampler to draw from the two full conditional distributions $\pi(\xi|\beta,y)$ and $\pi(\beta|\xi,y)$, the constant c is equal to 1, $h(x)$ are normal distribution for $\pi(\xi|\beta,y)$ and inverse gamma distribution for $\pi(\beta|\xi,y)$, $g(x)=\prod_{i=1}^{n_u}(1+y_i\xi/\beta)^{-(1+\xi)/\xi}$ also satisfies the conditions $0<g(x)\leq 1$.

ESTIMATION AND EVALUTION OF VAR

EVT aims to depict the behavior of extreme observations. The characteristics of electricity price data naturally lend itself to EVT analysis. For instance, electricity itself is non-storable. As such the balance between supply and demand must be maintained. This leads to an extremely turbulent market where spot prices can rise from average levels to many times within a very brief period. Large spot price movements expose market participants to significant market risk over short periods of time. In this situation risk managers are interested in a risk measure like VaR. The strong dependence in the sequence of electricity prices violates the underlying assumption that the data series to which EVT models are applied should be a sequence of IID random variables. In this paper, a two-stage approach, provided by McNeil and Frey [12], is used to this problem. Firstly, the dependences, skewnesses, lepkurtosises and seasonalities of electricity price series are filtered by a grey GM(1,1) model to obtain a nearly IID residual series $\{\varepsilon_t\}$. In stage two, the EVT framework is applied to the tails of the nearly IID residuals to better capture the heavy-tails and improve the accuracy of VaR estimation.

GM(1,1)-POT-VaR estimating model

VaR is one of the most intuitive and comprehensible risk measures. It is based on the standard statistical technology and has become an international popular risk measurement technology. Assuming normal market conditions and no trading in a given portfolio, VaR is defined as a threshold value such that the probability that the worst loss on the portfolio over a target horizon exceeds this value is the given level of probability. Mathematically, the VaR of the portfolio with a confidence interval p, VaR_p, is defined as [2]

$$VaR_p = \inf\{x \in R \mid \mathrm{Prob}(\Delta P \geq x) \leq 1 - p\}, \tag{14}$$

where Prob(·) denotes the portfolio probability distribution and ΔP the portfolio losses over the given holding period.

For a given time horizon t, suppose that the system demand for electricity is Q_t, the retail price to ultimate customers is P_0, the spot price is $p_t=\mathrm{E}(p_t|I_{t-1})+\varepsilon_t$, where E(·) is the conditional expectation operator, I_{t-1} the information set available at time t-1 and ε_t the random shock such that $\mathrm{E}(\varepsilon_t)=0$ and $\mathrm{E}(\varepsilon_t\varepsilon_s)=0$, $\forall t \neq s$. The trading losses of an electric utility over the target horizon t is

$$\Delta P_t = Q_t\left(\mathrm{E}(p_t \mid \mathrm{I}_{t-1}) + \varepsilon_t - P_0\right). \tag{15}$$

As the retail price, P_0, is a regulated price approved by electricity regulatory departments and the electric power demand, Q_t, can be accurately forecasted, Q_t and P_0 can be regarded as constant. Let $f_\varepsilon(\varepsilon_t \mid I_{t-1})$ denote the conditional probability density function of ε_t conditional on I_{t-1}. The VaR of an electric utility in the specified period t with the pre-assigned probability level p, denoted by $VaR_{p,t}$, is

$$1 - p = \mathrm{Prob}(\Delta P_t \geq VaR_{p,t}) = \int_{\frac{VaR_{p,t} - Q_t(\mathrm{E}(p_t|\mathrm{I}_{t-1}) - P_0)}{Q_t}}^{\infty} f_\varepsilon(x \mid \mathrm{I}_{t-1})dx \tag{16}$$

Now inverting Eq. (16) for the given probability p, we obtain

$$VaR_{p,t} = Q_t\left(\mathrm{E}(p_t \mid \mathrm{I}_{t-1}) - P_0 + F_\varepsilon^{-1}(p \mid \mathrm{I}_{t-1})\right), \tag{17}$$

where $F_\varepsilon(\cdot)$ is the conditional cumulative distribution function of ε_t, F_e^{-1} is the quantile function defined as the inverse of the distribution function F_ε.

The spot price presents the properties of incomplete and uncertain information. It is in line with the characteristics of grey variables, so we can estimate the expected values of the electricity spot price $\mathrm{E}(p_t|\mathrm{I}_{t-1})$ and the p-quantile $F_\varepsilon^{-1}(p \mid \mathrm{I}_{t-1})$ of the residual series ε_t by Eq. (4) and (8). Then we can calculate the VaR of an electric utility in the specified period t by Eq. (17).

Backtesting for VaR estimates

It is of crucial importance to assess the accuracy of VaR estimates. Backtesting or verification testing is the way that we verify whether forecasted losses are in line with actual losses. The most widely known backtesting method based on failure rates has been suggested by Kupiec [16]. Kupiec's test measures whether the number of violation exception is in line with the expected number for the chosen confidence interval. Under the null hypothesis that the VaR estimated model is correct at a pre-assigned confidence interval, the observed failure rate should act as an unbiased measure of the level of significance as sample size is increased. Denoting the number of times that the actual portfolio returns fall outside the VaR estimate as N and the total number of observations as T, the following likelihood ratio (LR)

$$LR = -2\log\left((1-c)^N c^{T-N}\right) + 2\log\left(\left(\frac{N}{T}\right)^N \left(1 - \frac{N}{T}\right)^{T-N}\right) \tag{18}$$

is asymptotically χ^2 (chi-squared) distributed with one degree of freedom. If the value of LR exceeds the critical value of the χ^2 distribution, the null hypothesis will be rejected and the model is deemed as inaccurate. On the contrary, the null hypothesis will be accepted and the model should be considered correct.

EMPIRICAL RESULTS

The PJM is organized as a day-ahead market. Participants submit their buying and selling bid curves for each of the next 24 hours. Then the market operator aggregates bids for each hour and determines market clearing prices and volumes for each hour of the following day. In this paper, a total of 1197 observations of average daily electricity spot prices in dollars per megawatt hour ($/MWh) and average daily loads in gigawatt (Gw) are employed to validate the performance of the VaR calculating model. The sample period begins on 1st June 2007 and ends on 9th September 2010.

Estimates of GM(1,1) model

To improve the filtering effects of GM (1,1) model, the data window length is set to 7 considering the significant weekly seasonality of the electricity price series. Table 1 illustrates the Ljung–Box Q statistics for the residuals and their square sequences. It is seen from Table 1, the Ljung–Box Q statistics of the square series of residuals are not significant at up to 24 lags, suggesting that no potential time-varying volatility exists in the residual series; the Ljung–Box Q statistics at 7 or 24 lags for the residual series are far less than the daily average electricity spot price series, indicting that there are still weak serial dependences, so we can conclude that the residual series is a stationary series with weakly serial correlation and without volatility clustering, meeting the prerequisite of EVT modeling [17].

TABLE 1: Ljung-Box test for residuals of GM(1,1) model

Statistics	Electricity prices ($/MWh)	Residuals ($/MWh)
Ljung-Box Q(6)	3868.28(0)	882.765(0)
Ljung-Box Q(24)	11348.94(0)	1235.612(0)
Ljung-Box $Q^2(6)$	3117.13(0)	314.055(0)
cLjung-Box $Q^2(24)$	7892.50(0)	456.305(0)

Estimates of GM(1,1)-POT-VaR model

To apply EVT, the threshold can be selected by the mean excess function or Hill plots. From a closer inspection of the mean excess function, we find that the sample mean excess plot is roughly linear when the value of the threshold u is fixed 6.295. After selecting the threshold u, the estimates of the shape and scale parameters, ξ and σ, can be determined by fitting the GPD to the residuals via Bayesian estimator. Inserting the estimates of ξ and σ into Eq. (8), the tail quantiles of the standardized residual series at a given confidence level c can be calculated. Table 2 reports the estimated results for tail index, scale parameter and tail quantiles. It can be seen that the ξ estimates is positive and statistically significant, indicating that the right tail of the distribution of standardized residuals is characterized by the Fréchet distribution. In Bayesian estimation, the estimated parameters are considered as random variables and it actually increases the uncertainty of asset return distribution, thus the VaRs calculated by Bayesian estimation are generally greater than by MLE.

TABLE 2: Estimates of GPD parameters and quantiles

Threshold	Shape parameter	Scale parameter	Confidence level	Tail quantile
Bayes	-0.11848	4.028975	95.0%	8.954089
			97.5%	11.42550
			99.0%	14.39604
MLE	-0.15951	4.029428	95.0%	8.91774
			97.5%	11.28731
			99.0%	14.04341

VaR estimates and backtesting

Without loss of generality, in this paper we assume that an electric utility has the obligation to serve 1MW of load 24 hours a day and the retail price has been frozen at a level equivalent to 0$/MWh. Substituting the calculated results at subsection 5.1 and 5.2 into Eq. (17), the VaR at each confidence level can be estimated. The Kupiec's test results for actual and forecasted losses are shown in Table 3. It can be seen from that the null hypotheses of ARMAX-GARCH-st-VaR [9], ARMAX-GARCHSK-VaR [10] and our proposed GM(1,1)-POT-VaR models cannot be rejected in each significance levels. Summarizing the results for the Kupiec's tests, the VaR predictions by these methods are insignificantly different from the proposed downfall probability, but because the GM(1,1)-POT-VaR model is easier to deal with and possesses the advantages of less computational costs, this further improves the risk management ability for electricity market participants to some extent.

TABLE 3: Backtests of estimated VaRs

Confidence level	Estimated model	Expected	Real	LR
95%	GARCHSK[10]	60	59	0.013
	GARCH-st[9]	60	60	0.000
	GM(1,1)-MLE	60	61	0.023
	GM(1,1)-Bayes	60	60	0.000
97.5%	GARCHSK[10]	30	28	0.013
	GARCH-st[9]	30	30	0.000
	GM(1,1)-MLE	30	32	0.144
	GM(1,1)-Bayes	30	29	0.030
99%	GARCHSK[10]	12	13	0.087
	GARCH-st[9]	12	13	0.087
	GM(1,1)-MLE	12	11	0.082
	GM(1,1)-Bayes	12	11	0.082

DISCUSSION AND CONCLUSION

The distinctive characteristics of electric energy which cannot be effectively stored through time and space and needs instantaneous balance of supply and demand make electricity price present highly volatility and occasional extreme movements of magnitudes rarely seen in markets for regular financial assets, thus volatility of price risk identification, evaluation and management in electricity market are more important than in financial markets. Considering various influencing factors on electricity prices and their pertinences, a gray system and extreme value theory based two-stage model for estimating VaR is proposed. In stage one, to capture the most important characteristics such as seasonalities, heteroscedasticities, skewnesses and lepkurtosises and to acquire the approximately IID residuals with better statistical properties, a gray GM(1,1) model is used to pre-filter the electricity price series. In stage two, an EVT based model is employed to explicitly deal with the right tail of the residuals of the GM(1,1)l, and accurate estimates of VaR in electricity market can be produced. The empirical analysis based on the historical data of the PJM electricity market indicates that the GM(1,1)-POT-VaR model can rapidly reflect the most recent and relevant changes of electricity prices and can produce accurate forecasts of VaR at all significance levels. Moreover, the computational costs is far less than the proposed models [9, 10], further improving the risk management ability of electricity market participants. These results present several potential implications for electricity market risk quantifications and hedging strategies.

REFERENCES

[1] M. Li, Q. Zhang, Y. SI; Power Syst. Protect. & Contr., 14, 12(2010).
[2] J. Liao, H. Jiang, J.C. Peng; Relay, 11, 30(2007).
[3] R.H. Huang, J. Zhang, L.Z. Zhang, Z.L. L; Proc. the CSEE, 19, 85(2009).
[4] C. Hartz, S. Mittnik, M. Paolella; Compu. Stat. & Data Analy., 51, 2295(2006).
[5] R.Q. Wang, H.F. Wang, X. Wang; Power Syst. Protect. & Contr., 40, 49 (2009).
[6] R.Q. Wang, F.X. Wang, X.J. Guo; J. Hainan Normal Univ., 25, 36(2012).
[7] http://www.sal.tkk.fi/publications/pdf-files/eleh07.pdf, (Nov 2014)
[8] H.N.E. Bystrom; Inter. Review Econ. & Fin., 14, 41(2005).
[9] R.Q. Wang, X. Wang, Y.Z. Li; East China Electric Power, 41, 1335(2013).
[10] R.Q. Wang, F.X. Wang, M.C. Xu; ARMAX-GARCHSK-EVT Model Based Risk Measure of Electricity Market, Paper presented at the 32nd Chinese Control Conference, Xi’an, China, 26-28, July(2013).
[11] L.L. Dong, D. Huang; Contr. Eng. China, 13, 252(2006).
[12] A.J. McNei, R. Frey; J. Empirical Fin., 7, 271(2000).
[13] M. Gilli, E. Kellezi; Compu. Economics, 27, 207(2006).
[14] J.G. Ibrahim, M.H. Chen, D. Sinha; Bayesian survival analysis, 1nd Ed., Berlin Heidelberg; New York (2001).
[15] S. Mao, J.L. Wang, X.H. Pu; Advanced Mathematical Statistics, 1nd Ed., Higher Education Press; Beijing (2004).
[16] P. Kupiec P; J. Derivatives, 39, 73(1975).
[17] S. Coles; An introduction to statistical modeling of extreme values, 1nd Ed., Springer-Verlag; London (2001).