FORECASTING VALUE-AT-RISK WITH NOVEL WAVELET BASED GARCH-EVT MODEL

Abstract

In this study, wavelet based GARCH-Extreme Value Theory (EVT) is proposed to model financial return series to forecast daily value-at-risk. Wavelets based GARCH-EVT is hybrid model combining the wavelet analysis and EVT. Proposed model contains three stages. In first stage, return series is decomposed into wavelet series and approximation series by applying the maximal overlap discrete wavelet transform. Second stage, detrended return series and approximation series are obtained by using wavelet series and scaling series. GARCH model is fitted to each obtained series to forecast daily volatility. Final stage, EVT is used to estimate quantile estimation of standardized residuals of GARCH model obtained for detrended return series and daily VaR value is forecasted by using volatility forecasts and quantile estimation. Daily VaR forecasting accuracy of proposed hybrid model is compared with the GARCH models specified under heavy-tailed distributions and GARCH-EVT model. Empirical findings show that wavelet based GARCH-EVT model is outperformed at high quantiles according to backtesting results.

Keywords: GARCH models, volatility forecasting, combining forecasts, decomposition, financial markets

JEL Classification: G32, G17, G15

1. INTRODUCTION

Most of the Value at Risk (VaR) models assume that financial return series are normally distributed. Modeling VaR with normality assumption, without considering the big and unpredictable losses, gives underestimate VaR forecasts. For this reason Extreme Value Theory (EVT) is good candidate to model tail of distribution that contains the extreme losses. McNeil and Frey (2000), Gencay et al. (2003), Gilli and Kellezi (2006), Onour (2010), Singh et al. (2013), Soltane et al. (2012), Chan and Gray (2006), Karmakar (2013) and Altun and Tatlidil (2015) have evaluated the performance of EVT in measuring the financial risk and also investigated tail behavior of financial returns series. Venkataraman (1997), Zangari (1996), Lee et al. (2008), Angelidis et al. (2004) have evaluated the performance of GARCH models under heavy-tailed distributions, such as student-t, mixture normal, generalized error distribution, skewed generalized error distribution, to forecast daily VaR. As a result of these studies, due to financial return series exhibit skewness and excess kurtosis, leptokurtic distributions are able to produce better daily VaR forecasts.

Wavelet analysis is a new tool in the field of applied mathematics. Fundamentals of the wavelet theory are provided by Daubechies (1992), Chui (1992) and Graps (2005). Wavelet analysis provides the opportunity to make semi-parametric estimations of highly complex structures without knowing the underlying functional form. In recent years, wavelet analysis is used to model financial returns series. Wavelet analysis decomposes financial return series into different scales that represent the low and high frequency sequences. Chi and Kai-jian (2006), Lai et al. (2006), Samia et al. (2009), Tan et al. (2010) and Cifter (2011) have evaluated performance of wavelet theory combining with ARMA-GARCH model in financial forecasting. Chi and Kai-jian (2006) proposed wavelet based value at risk model. Financial return series is decomposed into different scales and wavelet coefficients are de-noised according to the threshold selection rules. GARCH model is fitted de-noised return series to forecast daily VaR values. Lai et al. (2006) and Samia et al. (2009) decomposed the financial return series using wavelet analysis. GARCH model is fitted to decomposed time series and conditional volatility is modeled as a mixture of GARCH processes at each scale to forecast daily VaR values. Tan et al. (2010) proposed the price forecasting method based on wavelet transform combined with ARIMA-GARCH models. Wavelet analysis is used to decompose electricity price series into one approximation series and some detail series. GARCH model is fitted each detail series and ARIMA-GARCH model is fitted one approximation series. Price prediction is obtained by composing the forecasted values. Cifter (2011) proposed the wavelet based extreme value theory for univariate value at risk estimation. First stage, wavelet coefficients are used as threshold in Generalized Pareto distribution, in second stage EVT is applied with wavelet based thresholds.

In this paper, wavelet based GARCH-EVT model is proposed to forecast daily VaR based on wavelet transform combined with GARCH-EVT model. By applying wavelet transform, return series are decomposed into sub return series at different scales. Detrended return series is obtained by wavelet series and scaling series is used to add trend effects to proposed model. GARCH models are fitted to detrended return series and scaling series to forecast daily VaR with weighted volatility forecasts. The performance of proposed hybrid model is compared with the GARCH-normal, GARCH-student-t, GARCH-generalized error distribution, GARCH-skewed generalized error distribution and GARCH-EVT models for BIST-100 stock exchange index. The aim of this study is to evaluate the performance of wavelet based VaR model for daily-VaR forecast and also to show how distribution assumption made for residuals in GARCH model affects the daily-VaR forecasts. Backtesting methodology is used to compare model performance.

The rest of the paper organized as follows: Section II presents the VaR and EVT comprehensively. Section III presents GARCH models based on different distribution assumptions. Section IV presents the wavelet theory and wavelet based GARCH-EVT model. Section V presents empirical findings, model comparisons and final section presents the conclusion of study.

2. VALUE AT RISK AND EXTREME VALUE THEORY

Value-at-Risk (VaR) is defined as the largest possible loss of financial assets in a particular of time under a confidence level. VaR can be simply defined as follows:

$$VaR_{\alpha} = F^{-1}(1-\alpha)$$

where F is the distribution function of financial losses, F^{-1} denotes the inverse of F and α is the quantile at which VaR is calculated. EVT is used to model tail behavior of loss distribution and extreme events in financial time series. Modeling the extreme events with EVT, Peaks over Threshold (POT) methodology is used. POT method focuses on the distribution of exceedances over a threshold. F_u, which is the conditional excess distribution can be defined as follows:

$$F_u(y) = P(x-u \le y / x > u), \quad 0 \le y \le x_F - u \tag{1}$$

where X is a random variable, denotes the financial losses, u is a threshold, $y = x - u$ are the excesses, called as extreme losses, $x_F \leq \infty$ is the right endpoint of F which is the distribution function of X . F_u can be written in terms of F as follows,

$$F_u(y) = \frac{\Pr\{x - u \leq y, x > u\}}{\Pr(x > u)} = \frac{F(y+u) - F(u)}{1 - F(u)} = \frac{F(x) - F(u)}{1 - F(u)} \quad (2)$$

A theorem by Balkema and de Haan (1974) and Pickands (1975) indicates that, for sufficiently high threshold, the excess distribution function F_u, can be approximated by Generalized Pareto Distribution (GPD):

$$F_u(y) \approx G_{\xi,\sigma}(y), \quad \text{u} \to \infty$$

$$G_{\xi,\sigma}(y) = \begin{cases} 1 - (1 + \xi \dfrac{x - \mu}{\sigma})^{-1/\xi} & , \ \xi \neq 0 \\ 1 - e^{-y/\sigma} & , \ \xi = 0 \end{cases} \quad (3)$$

ξ is shape parameter, μ is the location parameter and σ is the scale parameter for GPD.

$F(x)$ can be isolated from Equation (2) and written as follows,

$$F(x) = (1 - F(u))F_u(y) + F(u) \quad (4)$$

$F_u(y)$ and $F(u)$ are replaced respectively by GPD and $(n - N_u)/n$, n is the total number of observations and N_u is the number of observations above the threshold. $\hat{F}(x)$ can be obtained as follows:

$$\hat{F}(x) = \frac{N_u}{n}(1 - (1 + \frac{\hat{\xi}}{\hat{\sigma}}(x - u))^{-1/\hat{\xi}} + (1 - \frac{N_u}{n}) = 1 - \frac{N_u}{n}(1 + \frac{\hat{\xi}}{\hat{\sigma}}(x - u))^{-1/\hat{\xi}} \quad (5)$$

VaR_p can be obtained by inverting Equation (6) for a given probability,

$$VaR_p = u + \frac{\hat{\sigma}}{\hat{\xi}}[(\frac{n}{N_u}p)^{-\hat{\xi}} - 1] \quad (6)$$

3. GARCH MODELS IN VAR ESTIMATION

Garch-normal model

Let $R_t = \ln\left(S_t / S_{t-1}\right) \times 100$ denotes the daily returns of the assets on time t and S_t represents the closed prices of the assets. Engle (1982) introduced the ARCH(q) model and expressed the conditional variance as a linear function of the past q squared residuals. Bollerslev (1986) proposed a generalization of the ARCH model, GARCH(1,1) model with normal error distribution can be written as follows:

$$
\begin{aligned}
&R_t = \mu + e_t \\
&e_t = \varepsilon_t \sigma_t, \quad \varepsilon_t \sim i.i.d.N(0,1) \\
&\sigma_t^2 = \omega + \alpha e_{t-1}^2 + \beta \sigma_{t-1}^2
\end{aligned} \tag{7}
$$

where respectively, μ and σ_t^2 are the conditional mean and variance. To ensure the stationarity condition and positive variance below equations must be hold.

$\alpha + \beta < 1, \alpha > 0, \beta > 0$ and $\omega > 0$

Log-likelihood function of GARCH-normal model under normality assumption can be written as:

$$
L(\psi) = -0.5\left(T \ln 2\pi + \sum_{t=1}^{T} \ln \sigma_t^2 + \sum_{t=1}^{T} \varepsilon_t^2 \right) \tag{8}
$$

According to GARCH-N model, one-day-ahead VaR forecast can be calculated as:

$$
VaR_{t+1} = \mu + F_\alpha(\varepsilon_t).\hat{\sigma}_t \tag{9}
$$

where $F_\alpha(\varepsilon_t)$ is the left quantile of standard normal distribution at α level.

Garch-student-t model

Bollerslev (1986, 1987) proposed the standardized student-t distribution with $\upsilon > 2$ degree of freedom. Student's-t is symmetric distribution and for $\upsilon > 4$, conditional kurtosis greater than 3, which exceeds the normal value. Under this specification, log-likelihood function, for a sample of T observations, can be written as follows:

$$
L(\psi) = T\left[\ln\Gamma\left(\frac{\upsilon+1}{2}\right) - \ln\Gamma\left(\frac{\upsilon}{2}\right) - \frac{1}{2}\ln\left[\pi(\upsilon-2)\right] \right] - \frac{1}{2}\sum_{t=1}^{T}\left[\ln\sigma_t^2 + (1+\upsilon)\ln\left(1 + \frac{\varepsilon_t^2}{\upsilon-2}\right) \right] \tag{10}
$$

where $\Gamma(\upsilon)$ is the gamma function and υ is the thickness parameter of the distribution tails. The one-day-ahead VaR forecast based on student-t distribution can be calculated as follows:

$$
VaR_{t+1} = \mu + F_\alpha(\varepsilon_t).\hat{\sigma}_t
$$

where $F_\alpha(\varepsilon_t)$ is the left quantile of the student-t distribution at α level.

Garch-GED model

In order to model the excess kurtosis observed asset prices, assumption on ε_t can be relaxed. Nelson (1991) proposed the generalized error distribution GED instead of assuming ε_t is normally distributed. Under this specification, log-likelihood function for GED distributed ε_t :

$$L(\psi)=\sum_{t=1}^{T}\left[\ln\left(\frac{\upsilon}{2}\right)-\frac{1}{2}\left|\frac{\varepsilon_t}{\lambda}\right|^{\upsilon}-(1+\upsilon^{-1})\ln(2)-\ln\Gamma\left(\frac{1}{2}\right)-\frac{1}{2}\ln\left(\sigma_t^2\right)\right] \tag{11}$$

where υ is the tail-thickness parameter and

$$\lambda=\left(\frac{\Gamma\left(\frac{1}{\upsilon}\right)}{2^{\frac{2}{\upsilon}}\Gamma\left(\frac{3}{\upsilon}\right)}\right)^{\frac{1}{2}} \tag{12}$$

where $\Gamma(.)$ is the gamma function. The Gaussian distribution is a special case of GED distribution when $\upsilon=2$. If $\upsilon<2$, GED has fatter tails than Gaussian distribution. According to Nelson (1991) specification, log-likelihood function can be written as follows:

$$L(\psi)=\sum_{t=1}^{T}\left[\ln\left(\frac{\upsilon}{\lambda}\right)-\frac{1}{2}\left|\frac{\varepsilon_t}{\lambda}\right|^{\upsilon}-\left(1+\upsilon^{-1}\right)\ln(2)-\ln\Gamma\left(\frac{1}{\upsilon}\right)-\frac{1}{2}\ln(\sigma_t^2)\right] \tag{13}$$

According to GARCH(1,1) model, $\sigma_t^2=\omega+\alpha e_{t-1}^2+\beta\sigma_{t-1}^2$ in the above equation. Parameters of the GARCH(1,1) model can be obtained by the numerical maximization procedure. The one-day-ahead VaR forecast based on GED distribution can be calculated as follows:

$$VaR_{t+1}=\mu+F_{\alpha}(\varepsilon_t,\upsilon).\hat{\sigma}_t$$

where $F_{\alpha}(\varepsilon_t,\upsilon)$ is the left quantile of GED distribution at α level.

Garch-SGED model

Lee et al. (2008) used the SGED distribution which provides a flexible distribution for modeling the empirical distribution of financial data. Probability density function of standardized SGED distribution can be written as follows:

$$f(\varepsilon_t)=C\exp\left(-\frac{\left|\varepsilon_t+\delta\right|^{\kappa}}{\left[1+sign(\varepsilon_t+\delta)\lambda\right]^{\kappa}\theta^{\kappa}}\right) \tag{14}$$

where

$$C=\frac{\kappa}{2\theta}\Gamma\left(\frac{1}{\kappa}\right)^{-1}, \theta=\Gamma\left(\frac{1}{\kappa}\right)^{0.5}\Gamma\left(\frac{3}{\kappa}\right)^{0.5}S(\lambda)^{-1}$$
$$S(\lambda)=\sqrt{1+3\lambda^2-4A^2\lambda^2}, \delta=\frac{2\lambda A}{S(\lambda)} \tag{15}$$
$$A=\Gamma\left(\frac{2}{\kappa}\right)\Gamma\left(\frac{1}{\kappa}\right)^{-0.5}\Gamma\left(\frac{3}{\kappa}\right)^{-0.5}$$

where κ is the shape parameter with constraint $\kappa>0$, λ is skewness parameter with $-1<\lambda<1$. SGED distribution turns out to be the standard normal distribution when $\kappa=2$ and $\lambda=0$. Log-likelihood function of GARCH-SGED model can be written as follows:

$$L(\psi)=-\frac{\left|R_t-\mu/\sigma_t+\delta\right|^{\kappa}}{\left[1+sign\left(R_t-\mu/\sigma_t+\delta\right)\lambda\right]^{\kappa}\theta^{\kappa}} \tag{16}$$

where ψ is the parameter vector. The one-day-ahead VaR forecast based on SGED distribution can be calculated as follows:

$$VaR_{t+1}=\mu+F_{\alpha}(\varepsilon_t,\kappa,\lambda).\hat{\sigma}_t$$

where $F_{\alpha}(\varepsilon_t,\kappa,\lambda)$ is the left quantile of SGED distribution at α level.

Garch-EVT model

McNeil and Frey (2000) proposed a two-stage model known as GARCH-EVT. Proposed model can be summarized as follow:

1. First step, GARCH (1,1) model is fitted to the return series by pseudo maximum likelihood estimation (PML) and gives the residuals for step-2 and also 1 day ahead predictions of μ_{t+1} and σ_{t+1}.
2. Second step, EVT-POT method is applied to the residuals of GARCH model. The most important point of this method is selection of threshold u. Using the parameter estimation of EVT-POT method and also predictions of μ_{t+1} and σ_{t+1}, VaR_{t+1} can be calculated easily.

The one-day-ahead VaR forecast based on GARCH-EVT model can be calculated as follows:

$$VaR_{t+1}=\mu+F_{\alpha}(\varepsilon_t;\xi,\sigma).\hat{\sigma}_t$$

where $F_{\alpha}(\varepsilon_t;\xi,\sigma)$ is obtained by the POT estimation procedure.

4. WAVELET ANALYSIS AND WAVELET BASED GARCH-EVT MODEL

Wavelet analysis, in contrast to Fourier analysis, gives insight in local behavior, whereas Fourier analysis gives insight in global behavior. The Fourier transform processes time-series by transforming the signal from the time domain into the frequency domain. However local effects are only visible in the time domain and not in the frequency domain. Wavelet analysis makes use of a fully scalable window, which is shifted along the signal in order to capture local behavior in the time domain (Hamburger 2003).

The use of wavelet analysis enables the analysis of non-stationary data, localization in time and time-scale decomposition, which proved to be useful in the analysis of economic and financial data (Ramsey 1999). A wavelet $\psi(t)$ is function of time t that satisfy admissibility condition,

$$C_{\psi} = \int_0^{\infty} \frac{|\varphi(f)|}{f} \mathrm{d}f < \infty \tag{17}$$

where $\varphi(f)$ is the Fourier transform of wavelet $\psi(t)$ in the frequency domain. Different wavelet families are available to capable of adapting to and accentuating certain data characteristics such as Haar wavelet, Daubechies wavelet, Symlets wavelet and Coiflets wavelet. Wavelet analysis is able to perform a process of separation, which is referred to as wavelet transform. There are two available wavelet transform: Discrete Wavelet Transform (DWT) and Continuous Wavelet Transform (CWT). Since most of the time series have finite number of values, DWT is used in finance and economics applications . Discrete wavelets are defined as:

$$\begin{aligned} \phi_{j,k} &= 2^{j/2}\phi\left(2^{j}t-k\right) \\ \psi_{j,k} &= 2^{j/2}\psi\left(2^{j}t-k\right) \end{aligned} \tag{18}$$

where $\phi_{j,k}$ and $\psi_{j,k}$ respectively represent the scaling signals and wavelets. Due to DWT is based on the power of two, length of the signal is need to be an integer power of two. To overcome this shortage, Maximal Overlap Discrete Wavelet Transform (MODWT) is used instead of DWT. MODWT can handle any sample size and wavelet variance estimator of MODWT is asymptotically more efficient than the estimator based on DWT. MODWT, similar to DWT, is a linear filtering operation that transforms a series into coefficients related to variations over a set of scales. The MODWT is suitable for multi-resolution analysis (MRA) and in contrast to DWT, MODWT is well-defined for all sample sizes N (Cornish and Percival 2006). In MODWT wavelet coefficients, $\hat{w}_{j,k}$, and scaling coefficients, $\hat{v}_{j,k}$, are obtained as follows:

$$\begin{aligned} \hat{w}_{j,t} &= \sum_{l=0}^{L_j-1} \hat{h}_{j,l} X_{t-l \bmod N} \\ \hat{v}_{j,t} &= \sum_{l=0}^{L_j-1} \hat{g}_{j,l} X_{t-l \bmod N} \end{aligned} \tag{19}$$

where $\hat{h}_{j,l}$ and $\hat{g}_{j,l}$ are respectively wavelet and scaling filters. The largest decomposition level $j = 1, 2, .., J$ is commonly determined such that $J \leq \log_2 N$.

In this study, we proposed wavelet based GARCH-EVT model. Proposed model can be summarized as follows:

Step 1. Implementing MODWT, financial returns series are decomposed into sub-return series at different scales j. W_j is the decomposed series by applying wavelet function and V_j is the decomposed series by applying the scaling function at scale $j = 1, 2, ..., J$. $f_w(t)$ represents detrended return series and $f_{V_j}(t)$ represents the level-j approximation of the original return series. $f_w(t)$ and $f_{V_j}(t)$ are respectively defined as follows:

$$f_w(t) = \sum_{j=1}^{J} W_j(t)$$
$$f_{V_j} = V_j \tag{20}$$

where $f_w(t)$ contains wavelet information of original returns series and f_{V_J} captures the trend of original return series at first decomposition level. $f_w(t)$ and f_{V_j} respectively refer to as detrended return series and approximation series.

Step 2. Second step, benchmark model GARCH(1,1) is fitted to $f_w(t)$ and f_{V_j} returns series by pseudo maximum likelihood estimation and gives the one-day-ahead forecasts of $\mu_{w,t+1}$, $\mu_{v,t+1}$, $\sigma_{w,t+1}$ and $\sigma_{v,t+1}$.

Step 3. Third step, EVT-POT method is applied to the standardized residuals of GARCH(1,1) model obtained for $f_w(t)$ wavelet based returns series. Threshold value of GPD is determined according to the 90th quantile value of standardized residuals. Using forecasts of $\mu_{w,t+1}$, $\mu_{v,t+1}$, $\sigma_{w,t+1}$ and $\sigma_{v,t+1}$, VaR_{t+1} can be calculated as follows:

$$VaR_{t+1} = \left(\mu_{w,t+1} + \mu_{v,t+1}\right) + F_{(\alpha,\varepsilon_t,\xi,\sigma)}\left(\sigma_{w,t+1}w_1 + \sigma_{vt+1}w_2\right)$$

where $F_{(\alpha,\varepsilon_t,\xi,\sigma)}$ is the corresponding quantile (95th, 97.5th or 99th) obtained by the POT estimation procedure. w_1 and w_2 are weights of forecasted volatilities. Optimal values of w_1 and w_2 are obtained according to the following optimization problem:

$$\begin{aligned} &Max\ \left(\sigma_{w,t+1}w_1 + \sigma_{v,t+1}w_2\right) \\ &w_1 + w_2 = 1 \\ &w_1 > 0.10 \\ &w_2 > 0.10 \end{aligned} \tag{21}$$

Objective function maximizes the forecasted volatility under the given constraints. Last two constraints avoid to occur zero weights for both returns series and also take account both trend and wavelet information into model at least 10 percent.

5. DATA AND EMPIRICAL FINDINGS

5.1. Data

Due to the unpredictable events and also extreme movements were occurred in Turkey stock exchange in recent years, ISE-100 index is selected to analyze the performance of wavelet based GARCH-EVT model on real data. S&P-500 and Nikkei-225 are also selected to compare the results of ISE-100 index with other financial markets. ISE-100, S&P-500 and Nikkei-225 respectively cover 1404, 1402 and 1367 daily observations from October 10, 2010 to July 31, 2015. Table 1 reports the descriptive statistics, unit root test results and ARCH-LM test results for all indexes.

Table 1. Descriptive statistics, ADF and ARCH-LM test results for ISE-100 index

ISE-100		S&P-500		Nikkei-225	
N. of Obs.	1404	N. of Obs	1402	N. of Obs	1367
Minimum	-0.1100	Minimum	-0.0689	Minimum	-0.1058
Maximum	0.0700	Maximum	0.0463	Maximum	0.0494
Mean	0.0003	Mean	0.0004	Mean	0.0005
Median	0.0010	Median	0.0007	Median	0.0007
Std. Deviation	0.0149	Std. Deviation	0.0099	Std. Deviation	0.0134
Skewness	-0.5850	Skewness	-0.4598	Skewness	-0.7155
Kurtosis	7.011	Kurtosis	7.606	Kurtosis	7.489
Jarque-Bera (JB)	1021.508 (0)	*Jarque-Bera (JB)*	1289.041 (0)	*Jarque-Bera (JB)*	1264.658 (0)
ADF test		ADF test		ADF test	
D-F = -10.784	p-value = 0.01	D-F = -11.1706	p-value = 0.01	D-F = -11.4653	p-value = 0.01
ARCH-LM test		ARCH-LM test		ARCH-LM test	
LM(2)	36.801 (0)	LM(2)	272.8072 (0)	LM(2)	95.1465 (0)
LM(5)	72.396 (0)	LM(5)	295.8739 (0)	LM(5)	108.9578 (0)
LM(10)	89.465 (0)	LM(10)	327.310 (0)	LM(10)	121.4966 (0)

*p values are shown in brackets

According to Table 1, mean is closed to 0 for all market returns. ISE-100 and Nikkei-225 indexes have bigger losses and exhibit higher volatility than S&P-500 index. Skewness and kurtosis are significantly different from the 0 and 3 for normal distribution and also *JB* test statistics are greater than the critical value at %5 level and *p-value* is 0. Therefore, log-returns of all indexes have the non-normal characteristics, excess kurtosis and fat tails. According to ADF test, log-returns of all indexes have not contain unit root and ARCH-LM test indicates that ARCH effects exist for all indexes.

5.2. Empirical Findings

Table 2. represents the parameter estimation of GARCH(1,1)-normal, GARCH(1,1)-student's-t, GARCH(1,1)-GED, GARCH(1,1)-SGED, GARCH(1,1)-EVT, W-GARCH(1,1)-EVT models for three indexes. Threshold value of GPD is determined with respect to 90^{th} quantile of the standardized residuals. To implement W-GARCH-EVT model Haar wavelet is selected as wavelet family because of its simplicity. Conditional variance parameters are highly significant and $\omega > 0, \alpha \geq, \beta \geq 0$ and $\alpha + \beta < 1$ conditions are hold to ensure the positive variance and stationarity condition.

Table 2. Parameter estimates of GARCH(1,1) model for three indexes

Parameter	ISE-100	S&P-500	Nikkei-225
Normal Distribution			
ω	0.000014	0.000004	0.000007
	(0.000004)	(0.000001)	(0.000002)
α	0.115529	0.132985	0.108763
	(0.020228)	(0.020696)	(0.018799)
β	0.825528	0.827618	0.854564
	(0.029143)	(0.02302)	(0.02422)
LL	3986.3	4679.2	4038.64
Student's t Distribution			
ω	0.000009	0.000004	0.000006
	(0.000004)	(0.000001)	(0.000002)
α	0.076194	0.143488	0.088476
	(0.020749)	(0.027886)	(0.020138)
β	0.881708	0.824626	0.877095
	(0.034727)	(0.028338)	(0.026974)
ν	6.506	5.57	9.927
	(1.068)	(0.918575)	(2.368)
LL	4023.6	4706.06	4053.66
Generalized Error Distribution			
ω	0.000012	0.000004	0.000007
	(0.000004)	(0.000001)	(0.000002)
α	0.094931	0.13897	0.098565
	(0.022041)	(0.026437)	(0.020936)
β	0.852554	0.822653	0.866207
	(0.034544)	(0.028625)	(0.027381)
ν	1.364	1.294	1.493
	(0.067691)	(0.068933)	(0.082621)
LL	4018.1	4712.22	4053.08
Skewed Generalized Error Distribution			
ω	0.000011	0.000003	0.000006
	(0.000004)	(0.000001)	(0.000002)
α	0.086264	0.135761	0.095597
	(0.020819)	(0.025208)	(0.019794)
β	0.863473	0.826555	0.869869
	(0.034415)	(0.02787)	(0.025828)
λ	0.868651	0.901101	0.905353
	(0.031729)	(0.029434)	(0.034725)
κ	1.384	1.345	1.571
	(0.069421)	(0.072513)	(0.092393)
LL	4026.42	4716.77	4056.62

Generalized Pareto Distribution			
ω	0.0000139	0.00000368	0.000007
	(0.000004)	(0.000001)	(0.0000011)
α	0.115576	0.132963	0.108763
	(0.020228)	(0.020696)	(0.018799)
β	0.82518	0.82737	0.854199
	(0.029143)	(0.02302)	(0.02422)
ξ	0.09963839	-0.3058944	0.04754548
σ	0.61521458	0.8580823	0.57497616
LL	3986.29	4679.2	4038.64
W-GARCH-EVT			
ω_1	0.0000161	0.0000142	0.0000131
	(0.000003)	0.000004	0.000012
α_1	0.050337	0.050075	0.051526
	(0.009735)	0.002681	0.005474
β_1	0.900047	0.900017	0.900268
	(0.023219)	0.004412	0.011681
ω_2	0.00002	0.0000121	0.000018
	(0.000005)	0.000002	0.000002
α_2	0.373369	0.273073	0.396863
	(0.059769)	0.039011	0.05238
β_2	0.483288	0.688103	0.425862
	(0.0734)	0.057467	0.044284
ξ	0.1192443	-0.1228539	0.1073522
σ	0.5569976	0.7820656	0.5476745
LL_1	4532.53	4382.46	3781.6
LL_2	3998.68	5240.83	4575.84

Standard errors are presented in parentheses.

5.3. Backtesting Results

To compare the forecasting ability of these models in terms of VaR forecasts, backtesting methodology is used. Kupiec (1995) proposed a Likelihood Ratio (LR) test for evaluating the model accuracy. The LR test statistic can be written as follows:

$$LR = -2\ln\left[\frac{p^{n_1}(1-p)^{n_0}}{\hat{\pi}^{n_1}(1-\pi)^{n_0}}\right] \sim \chi_1^2 \qquad (22)$$

where $\hat{\pi} = n_1/(n_0 + n_1)$ is the maximum likelihood estimation of p, n_1 represents the total violations and n_0 represents the total non-violations forecasts. Under the null hypothesis $(H_0 : p = \hat{\pi})$, LR statistics follows a chi-square distribution with one degree of freedom. Root mean square error (RMSE) also can be used to evaluate the overestimate or underestimate positions of the models for VaR forecasts. RMSE can be calculated as

$$RMSE = \sqrt{\frac{\sum_{t=1}^{T}\left(VaR_t - r_t\right)^2}{n}} \tag{23}$$

where VaR_t and r_t respectively represent the VaR forecast and return values for time t. In this study RMSE is calculated only for negative log-returns. Left tail of the distribution which contains the financial losses is considered.

Rolling window estimation procedure is used to evaluate the out of sample performance of models. Window length is differently determined for all indexes to evaluate out of sample performance of models with equal forecast period which contains 398 daily observations. Window lengths of equity indexes are determined as follows: 1006 for ISE-100, 1004 for S&P-500 and 969 for Nikkei-225 indexes.

Table 3. represents the backtesting results for ISE-100 index. According to backstesting results, GARCH-SGED, GARCH-EVT and W-GARCH-EVT models have the same violations at %95 and % 99 confidence levels. According to the LR-uc statistic value, there is no difference between these three models for both confidence levels. RMSE can be used in this case to decide which model is better than others.

Table 3. Out of sample performance of models according to bactesting results for ISE-100 index

		%95 confidence level		
ISE-100	**Number of Forecasts**	**Expected Violation**	**Observed Violation**	**LR-uc**
GARCH-normal	398	20	22	0.226 (0.635)
GARCH-student's t	398	20	23	0.485 (0.486)
GARCH-GED	398	20	20	0.001 (0.982)
GARCH-SGED	398	20	16	0.86 (0.354)
GARCH-EVT	398	20	16	0.86 (0.354)
W-GARCH-EVT	398	20	16	0.86 (0.354)
		%99 confidence level		
ISE-100	**Number of Forecasts**	**Expected Violation**	**Observed Violation**	**LR-uc**
GARCH-normal	398	4	9	4.711 (0.03)
GARCH-student's t	398	4	5	0.244 (0.621)
GARCH-GED	398	4	7	1.888 (0.169)
GARCH-SGED	398	4	4	0 (0.992)
GARCH-EVT	398	4	4	0 (0.992)
W-GARCH-EVT	398	4	4	0 (0.992)

RMSE values of the three models are given in Table 4. W-GARCH-EVT has minimum RMSE value for % 99 confidence levels and GARCH-SGED has minimum RMSE value for %95 confidence level. W-GARCH-EVT is outperformed at higher confidence level according to both RMSE and LR-uc statistic values for ISE-100 index.

Table 4. RMSE values of the models for ISE-100 index

Models (ISE-100)	RMSE
GARCH-SGED(0.01)	0.238029567
GARCH-SGED(0.05)	0.126702128*
GARCH-EVT(0.01)	0.263801298
GARCH-EVT(0.05)	0.139166049
W-GARCH-EVT (0.01)	0.236073666**
W-GARCH-EVT (0.05)	0.129426265

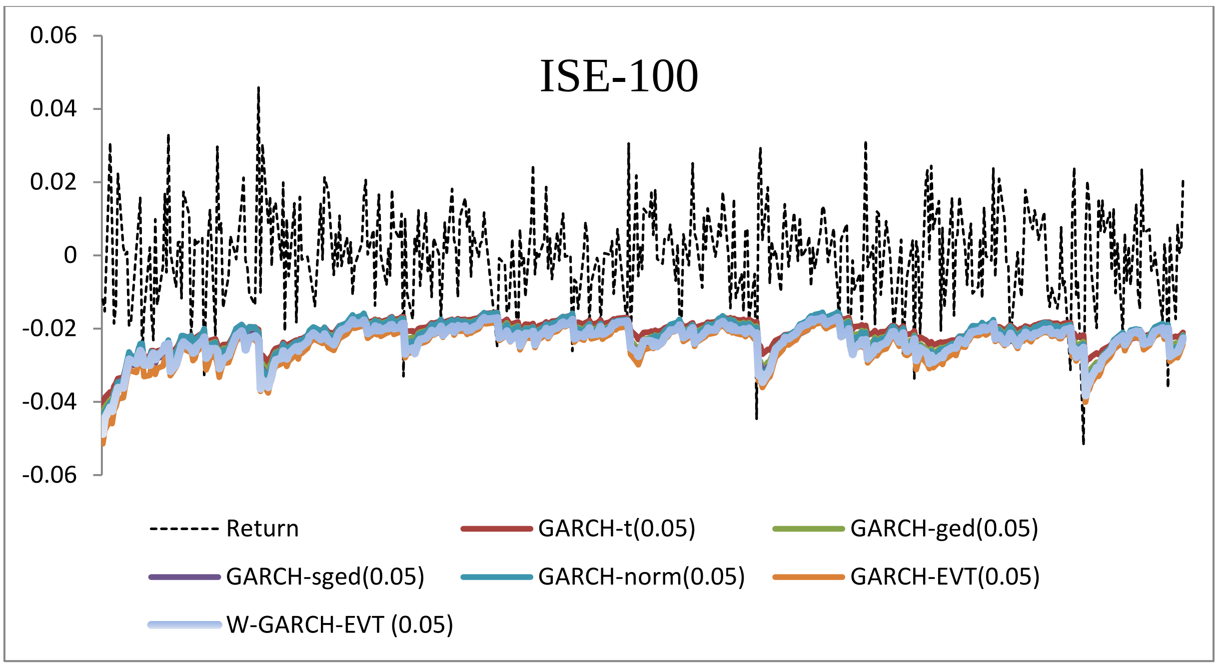

Figure 1. Daily VaR forecasts for ISE-100 index at %95 confidence level

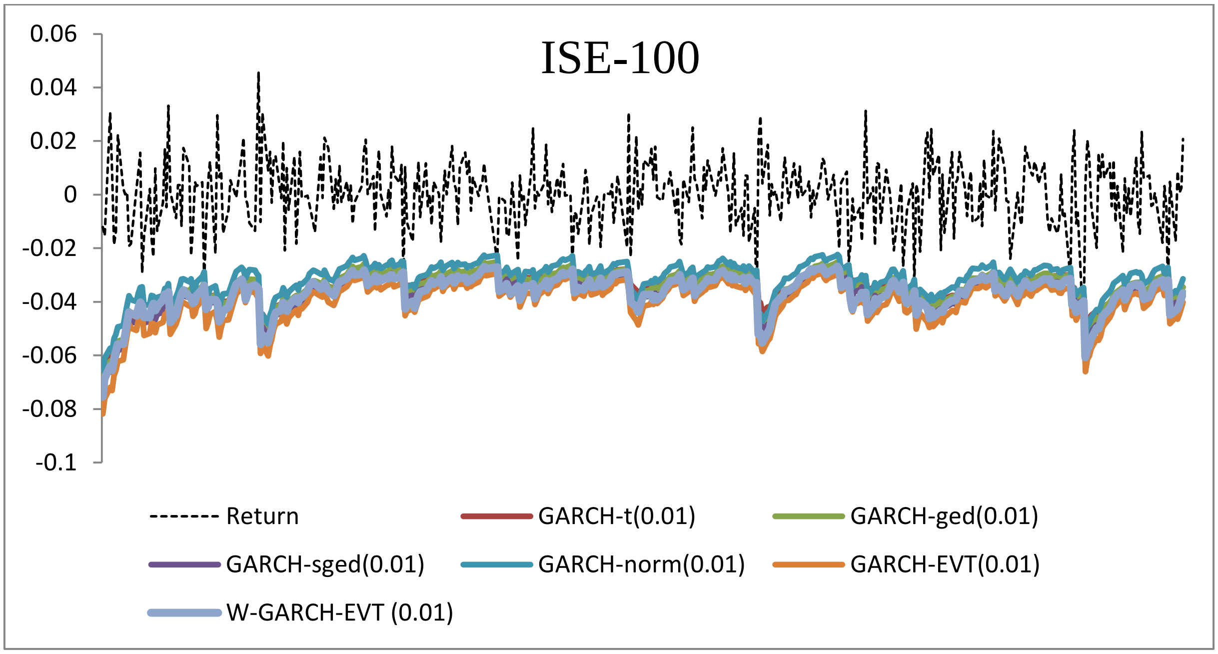

Figure 2. Daily VaR forecasts for ISE-100 index at %99 confidence level

Figure 1 and Figure 2 represent the daily VaR forecasts for ISE-100 index at % 95 and %99 confidence levels. As it seems in Figure 1 and Figure 2, Due to GARCH-EVT model gives big response to the changing volatility, GARCH-EVT model is overestimation position in most instances according to W-GARCH-EVT.

Table 5. represents the backtesting results for S&P-500 index. According to backstesting results, despite the fact that GARCH-EVT has the minimum violation value for both confidence levels, due to the smallest LR-uc statistic value indicates the best performed model, W-GARCH-EVT model has the best predictive performance for both confidence levels. GARCH-EVT model produces the overestimation VaR forecasts.

Table 5. Out of sample performance of models according to bactesting results for S&P-500 index

%95 confidence level				
S&P-500	**Number of Forecasts**	**Expected Violation**	**Observed Violation**	**LR-uc**
GARCH-normal	398	20	26	1.803 (0.179)
GARCH-student's t	398	20	27	2.41 (0.121)
GARCH-GED	398	20	26	1.803 (0.179)
GARCH-SGED	398	20	23	0.485 (0.486)
GARCH-EVT	398	20	16	0.859 (0.353)
W-GARCH-EVT	398	20	21	0.0629 (0.8019)
%99 confidence level				
S&P-500	**Number of Forecasts**	**Expected Violation**	**Observed Violation**	**LR-uc**
GARCH-normal	398	4	9	4.711 (0.03)
GARCH-student's t	398	4	6	0.896 (0.344)
GARCH-GED	398	4	6	0.896 (0.344)
GARCH-SGED	398	4	4	0 (0.992)

GARCH-EVT	398	4	2	1.217 (0.2698)
W-GARCH-EVT	398	4	4	0 (0.992)

RMSE is calculated only for best performed models. W-GARCH-EVT has minimum RMSE value for % 99 confidence levels and GARCH-SGED has minimum RMSE value for %95 confidence level. W-GARCH-EVT is outperformed at high confidence level according to both RMSE and LR-uc statistic values for S&P-500 index.

Table 6. RMSE values of models for S&P-500 index

Models	RMSE
GARCH-SGED(0.01)	0.13208733
GARCH-SGED(0.05)	0.07224983*
GARCH-EVT(0.01)	0.15851284
GARCH-EVT(0.05)	0.08849093
W-GARCH-EVT(0.01)	0.13001309**
W-GARCH-EVT(0.05)	0.07364195

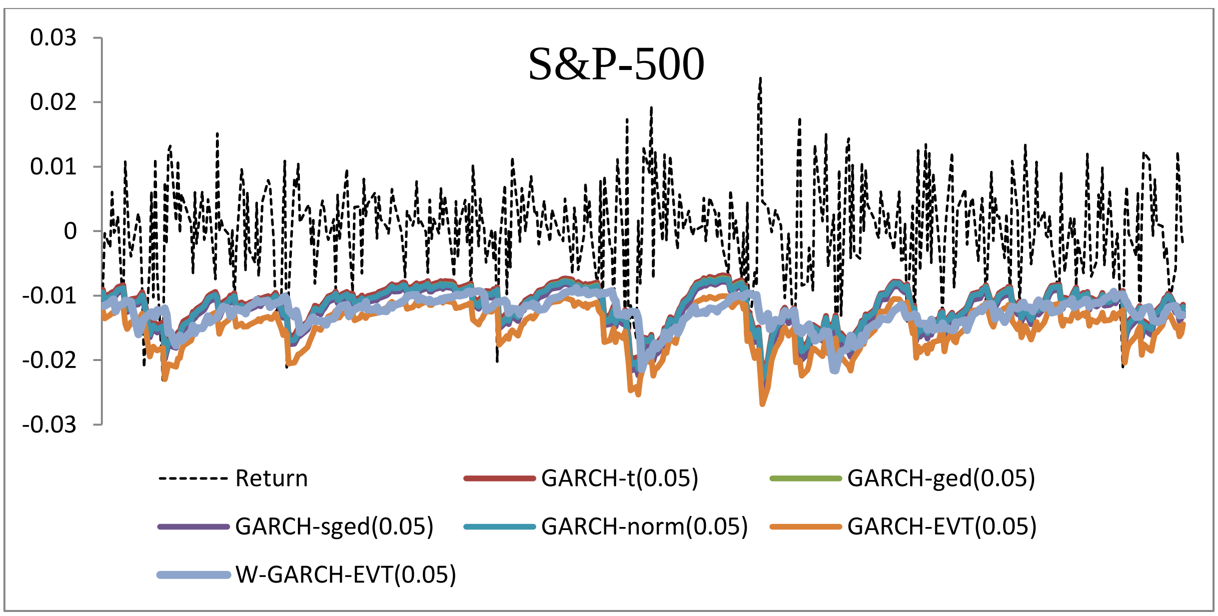

Figure 3. Daily VaR forecasts for S&P-500 index at %95 confidence level

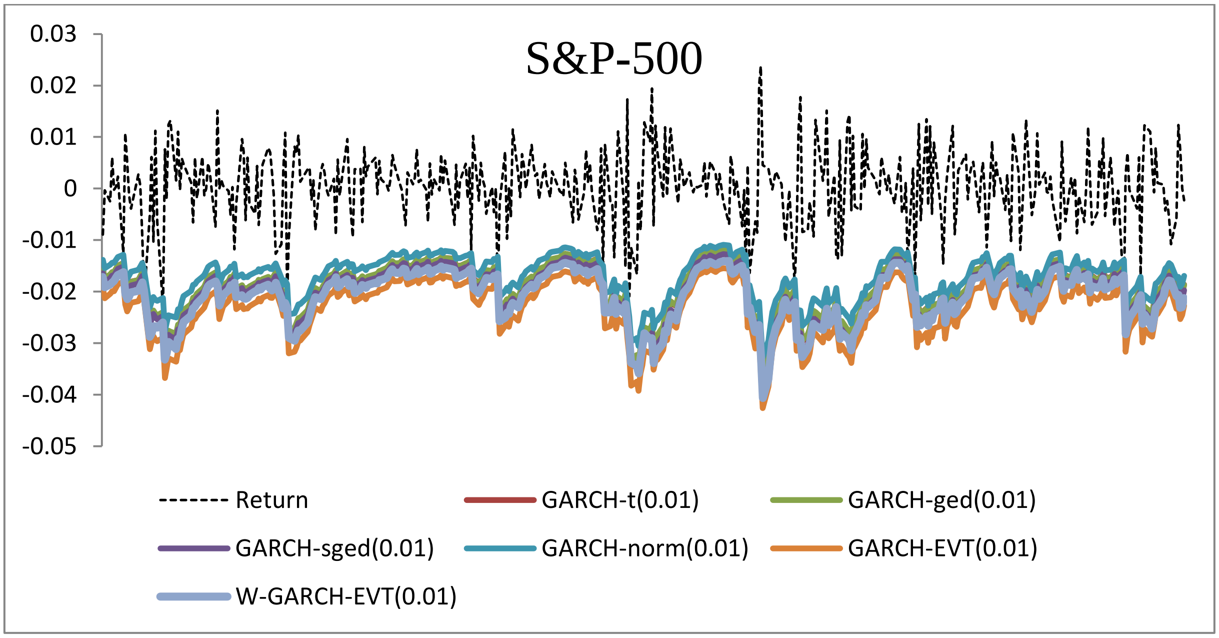

Figure 4. Daily VaR forecasts for S&P-500 index at %99 confidence level

Figure 3 and Figure 4 represent the daily VaR forecasts for S&P-500 index at % 95 and %99 confidence levels. The results of S&P-500 index are very similar to results of ISE-100 index. GARCH-EVT model gives bigger response than W-GARCH-EVT to changing volatility that causes to occur overestimation VaR forecasts.

Table 7. represents the backtesting results for Nikkei-225 index. Backtesting results show that GARCH-EVT and W-GARCH-EVT models are outperformed according to LR-uc statistic values for both confidence levels. Due to the smallest LR-uc value indicates the best performed model, W-GARCH-EVT model is better predictive performance than GARCH-EVT at %99 confidence level.

Table 7. Out of sample performance of models according to backtesting results for Nikkei-225 index

		%95 confidence level		
Nikkei-225	**Number of Forecasts**	**Expected Violation**	**Observed Violation**	**LR-uc**
GARCH-normal	398	20	23	0.485 (0.486)
GARCH-student's t	398	20	23	0.485 (0.486)
GARCH-GED	398	20	23	0.485 (0.486)
GARCH-SGED	398	20	21	0.063 (0.802)
GARCH-EVT	398	20	19	0.0434 (0.834)
W-GARCH-EVT	398	20	21	0.063 (0.802)
		%99 confidence level		
Nikkei-225	**Number of Forecasts**	**Expected Violation**	**Observed Violation**	**LR-uc**
GARCH-normal	398	4	12	10.611 (0.001)
GARCH-student's t	398	4	10	6.479 (0.011)

GARCH-GED	398	4	9	4.711 (0.03)
GARCH-SGED	398	4	7	1.888 (0.169)
GARCH-EVT	398	4	3	0.266 (0.605)
W-GARCH-EVT	398	4	4	0 (0.992)

Table 8 indicates the RMSE values of best performed models. Although GARCH-EVT model has smaller LR-uc value than W-GARCH-EVT model at % 95 confidence level, RMSE value indicates that W-GARCH-EVT model is performed better than GARCH-EVT for both confidence level.

Table 8. RMSE values of models for S&P-500 index

Models	RMSE
GARCH-EVT(0.01)	0.203544522
GARCH-EVT(0.05)	0.117238362
W-GARCH-EVT (0.01)	0.200064962**
W-GARCH-EVT (0.05)	0.111200969*

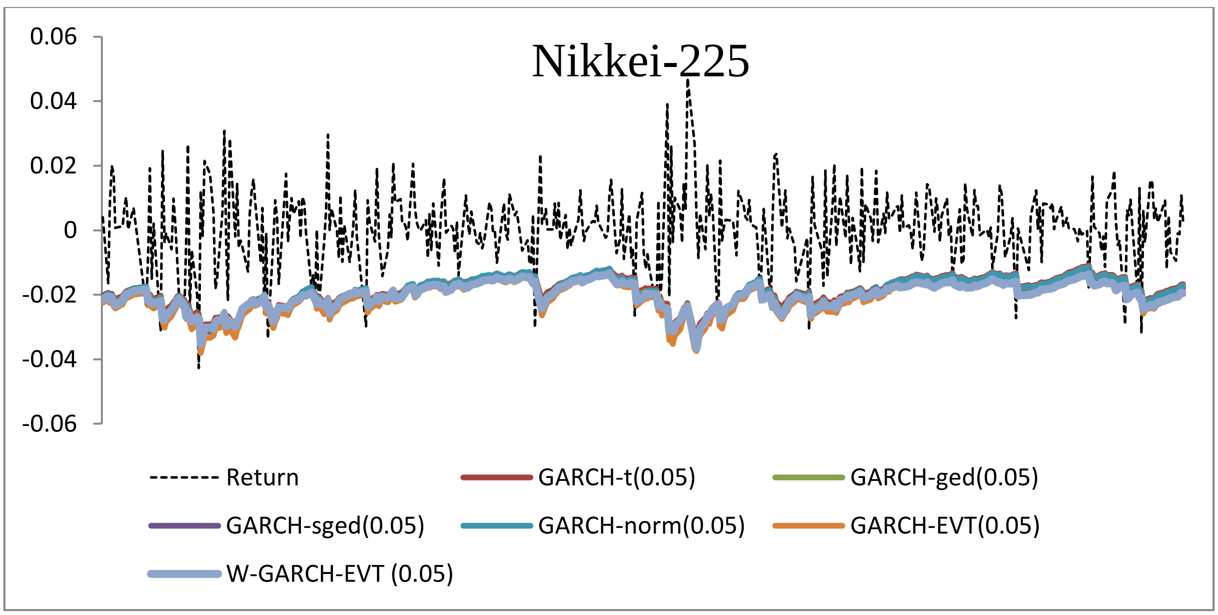

Figure 5. Daily VaR forecasts for Nikkei-225 index at %95 confidence level

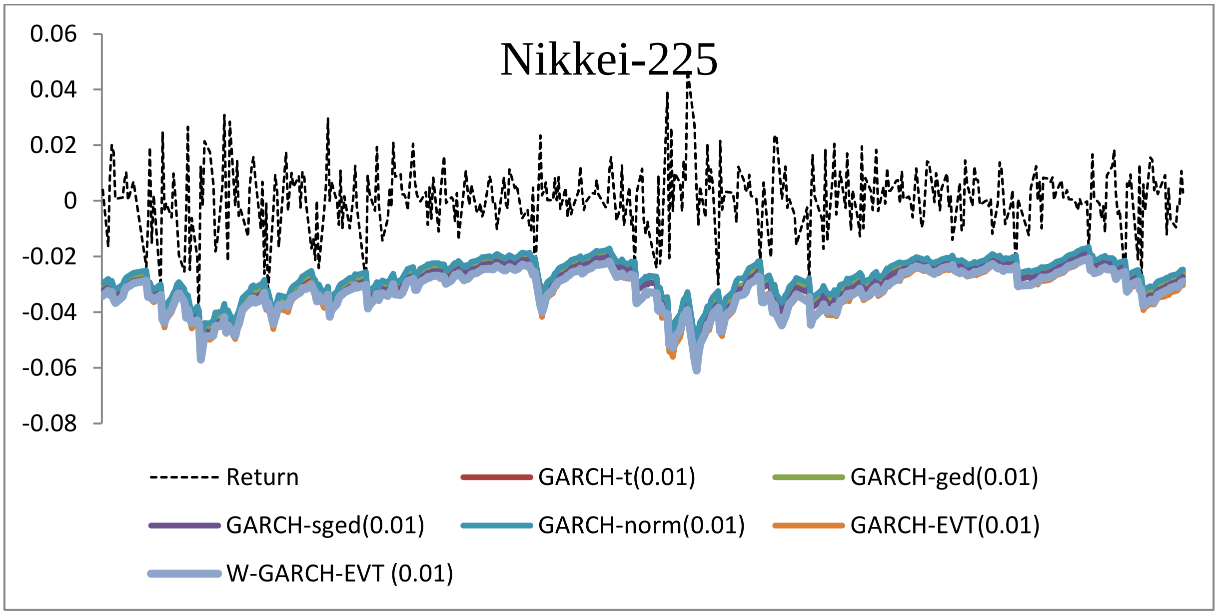

Figure 6. Daily VaR forecasts for Nikkei-225 index at %99 confidence level

Figure 5 and Figure 6 represent the daily VaR forecasts for Nikkei-225 index at % 95 and %99 confidence levels. The results of Nikkei-225 index are very similar to results of ISE-100 and also S&P-500 indexes. GARCH-EVT model produce overestimation VaR forecasts at high quantile.

6. CONCLUSION

This paper applies wavelet analysis to value-at-risk forecasts in ISE-100, S&P-500 and Nikkei-225 indexes and improves forecasting accuracy at higher confidence levels. The models based on past volatility rather than the extreme observations cannot be able to capture unpredictable and extreme losses. Proposed hybrid model combines the EVT and wavelet analysis to capture the extreme movements in financial markets. The contribution of this paper can be summarized as follows. Firstly W-GARCH-EVT has demonstrated its capability to improve the reliability of VaR forecasts at high confidence levels for three financial markets. Secondly, considering the overestimation problem of GARCH-EVT model, W-GARCH-EVT model produces reliable VaR forecasts and finds a solution to overestimation problems of VaR models. For this reasons, wavelet based GARCH-EVT model can be used to forecast market value-at-risk by financial institutions.

7. REFERENCES

Altun, E., & Tatlidil, H. (2015). A Comparison of Extreme Value Theory with Heavy-tailed Distributions in Modeling Daily VAR. *Journal of Finance and Investment Analysis*, 4(2), 69-83

Angelidis, T., Benos, A., & Degiannakis, S. (2004). The use of GARCH models in VaR estimation. *Statistical methodology*, *1*(1), 105-128.

Balkema, A. A., & De Haan, L. (1974). Residual life time at great age. *The Annals of probability*, 792-804.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of econometrics*, *31*(3), 307-327.

Bollerslev, T. (1987). A conditionally heteroskedastic time series model for speculative prices and rates of return. *The review of economics and statistics*, 542-547.

Chan, K. F., & Gray, P. (2006). Using extreme value theory to measure value-at-risk for daily electricity spot prices. *International Journal of Forecasting*,*22*(2), 283-300.

Chi, X., & Kai-jian, H. (2006, October). Wavelet denoised value at risk estimate. In *Management Science and Engineering, 2006. ICMSE'06. 2006 International Conference on* (pp. 1552-1557). IEEE.

Chui, C. K. (1992). An Introduction to Wavelets,(1992). *Philadelphia*, *SIAM*, *38*.

Cifter, A. (2011). Value-at-risk estimation with wavelet-based extreme value theory: Evidence from emerging markets. *Physica A: Statistical Mechanics and its Applications*, *390*(12), 2356-2367.

Cornish, C. R., Bretherton, C. S., & Percival, D. B. (2006). Maximal overlap wavelet statistical analysis with application to atmospheric turbulence.*Boundary-Layer Meteorology*, *119*(2), 339-374.

Daubechies, I. (1992). *Ten lectures on wavelets* (Vol. 61, pp. 198-202). Philadelphia: Society for industrial and applied mathematics.

Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica: Journal of the Econometric Society*, 987-1007.

Gencay, R., & Selcuk, F. (2004). Extreme value theory and Value-at-Risk: Relative performance in emerging markets. *International Journal of Forecasting*, *20*(2), 287-303.

Gilli, M. (2006). An application of extreme value theory for measuring financial risk. *Computational Economics*, *27*(2-3), 207-228.

Graps, A. (1995). An introduction to wavelets. *Computational Science & Engineering, IEEE*, *2*(2), 50-61.

Hamburger, Y. (2003). *Wavelet-based Value at Risk Estimation: A Multiresolution Approach*. Erasmus Universiteit.

Karmakar, M. (2013). Estimation of tail-related risk measures in the Indian stock market: An extreme value approach. *Review of Financial Economics,22*(3), 79-85.

Kupiec, P. H. (1995). Techniques for verifying the accuracy of risk measurement models. *THE J. OF DERIVATIVES*, 3(2).

Lai, K. K., He, K., Xie, C., & Chen, S. (2006, July). Market risk measurement for crude oil: a wavelet based var approach. In *Neural Networks, 2006. IJCNN'06. International Joint Conference on* (pp. 2129-2136). IEEE.

Lee, M. C., Su, J. B., & Liu, H. C. (2008). Value-at-risk in US stock indices with skewed generalized error distribution. *Applied Financial Economics Letters*, *4*(6), 425-431.

Lee, M. C., Su, J. B., & Liu, H. C. (2008). Value-at-risk in US stock indices with skewed generalized error distribution. *Applied Financial Economics Letters*, *4*(6), 425-431.

McNeil, A. J., & Frey, R. (2000). Estimation of tail-related risk measures for heteroscedastic financial time series: an extreme value approach. *Journal of empirical finance*, *7*(3), 271-300.

Nelson, D. B. (1991). Conditional heteroskedasticity in asset returns: A new approach. *Econometrica: Journal of the Econometric Society*, 347-370.

Onour, I. A. (2010). Extreme risk and fat-tails distribution model: empirical analysis. *Journal of Money, Investment and Banking*, (13).

Pickands III, J. (1975). Statistical inference using extreme order statistics. *the Annals of Statistics*, 119-131.

Ramsey, J. B. (1999). The contribution of wavelets to the analysis of economic and financial data. *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, *357*(1760), 2593-2606.

Samia, M., Dalenda, M., & Saoussen, A. (2009). Accuracy and Conservatism of VaR Models: A Wavelet Decomposed VaR Approach Versus Standard ARMA-GARCH Method. *International Journal of Economics and Finance*, *1*(2), p174.

Singh, A. K., Allen, D. E., & Robert, P. J. (2013). Extreme market risk and extreme value theory. *Mathematics and computers in simulation*, *94*, 310-328.

Soltane, H. B., Karaa, A., & Bellalah, M. (2012). Conditional VaR Using GARCH-EVT Approach: Forecasting Volatility in Tunisian Financial Market.*Journal of Computations & Modelling*, *2*(2), 95-115.

Tan, Z., Zhang, J., Wang, J., & Xu, J. (2010). Day-ahead electricity price forecasting using wavelet transform combined with ARIMA and GARCH models. *Applied Energy*, *87*(11), 3606-3610.

Venkataraman, S. (1997). Value at risk for a mixture of normal distributions: the use of quasi-Bayesian estimation techniques. *Economic Perspectives-Federal Reserve Bank of Chicago*, *21*, 2-13.

Zangari, P. (1996). An improved methodology for measuring VaR. *RiskMetrics Monitor*, *2*(1), 7-25.