RT&A, No 4 (89)
Volume 20, December 2025

FORECASTING EXTREME MARKET MOVEMENTS USING REALIZED GARCH AND EXTREME VALUE THEORY MODELS

Umme Hani
•
Department of Economics and Finance, Saudi Electronic University, Al Madinah, Saudi Arabia. uhani@seu.edu.sa

Abstract

This study investigates the essential benefits of using the Conditional Value-at-Risk (CVaR) optimization framework with the GARCH model and Extreme Value Theory (EVT) methods for the S&P CNX NIFTY markets. Using daily data from Jan. 2020 to July 2024, we evaluate 10 forecasting models, comprising eight standalone GARCH models and eight two-stage GARCH-EVT models. Our results show that the GARCH-EVT models consistently produce more accurate quantile forecasts than their standalone GARCH counterparts. Among the standalone models, daily returns-based GARCH and EGARCH models exhibit the poorest forecasting performance, while the intraday return-based realized GARCH model performs slightly better. The experimental results show that, under various loss functions, the GARCH-EVT innovation model is the best model for volatility predictions of S&P CNX NIFTY among the sixteen forecasting models.

Keywords: Extreme value theory, Realized GARCH, realized kernel, Value-at-risk, Expected shortfall, GARCH-EVT model.

1. Introduction

Value-at-Risk (VaR) and Expected Shortfall (ES) are key measures of financial risk, widely applied in areas such as portfolio optimization, performance evaluation, regulatory capital computation, and the determination of margin requirements for investors. The Generalized Autoregressive Conditional Heteroskedasticity (GARCH) family of models is extensively used to capture the conditional moments of financial returns and to forecast VaR and ES. However, a major limitation of conventional GARCH models is their reliance on squared daily returns as a proxy for daily volatility an approach that produces a highly noisy estimate of latent volatility. To address this issue, [1] proposed the realized GARCH model, which incorporates information from intraday returns to provide a much more accurate and efficient estimate of underlying volatility. Empirical studies have shown that the realized GARCH framework substantially improves the in-sample and out-of-sample performance of volatility forecasts compared to traditional GARCH models. Financial assets with high liquidity often experience thousands of intraday transactions, and using only close-to-close returns discards valuable information embedded in high-frequency price movements [2].

Traditional GARCH models jointly estimate the conditional mean and variance of returns based solely on daily data. In contrast, the realized GARCH model introduces an additional measurement equation that jointly models a realized measure of daily volatility along with the

DOI: https://doi.org/10.24412/1932-2321-2025-489-1178-1188

conditional moments of the return process. The realized volatility (RV) is computed as the sum of squared intraday returns, which converges in probability to the latent integrated variance as the sampling frequency increases. Despite this advantage, estimating RV at very high frequencies can introduce market microstructure noise, such as bid–ask bounce and price discreteness. To mitigate this issue, RV is typically estimated using lower-frequency samples (e.g., 5–60-minute intervals). However, lower sampling frequencies reduce efficiency because many price observations are ignored, and the RV estimate may fail to converge to the true integrated variance. Subsampling techniques can partially address this limitation, but recent research has proposed robust realized variance estimators that better balance bias and efficiency. In addition to the classical RV estimator, this study employs two robust realized variance estimators: the realized bipower variance (BV) proposed by Barndorff-Nielsen, O. E. and Shephard, [8] and the realized kernel (RK) estimator introduced. The BV estimator is robust to sudden price jumps that create large, infrequent intraday returns, which can otherwise inflate volatility estimates [3]. By multiplying adjacent absolute returns instead of squaring them, the BV estimator reduces the influence of such jumps. However, it remains sensitive to microstructure effects and is thus typically applied at 5-minute intervals [4].

The RK estimator, in contrast, is designed to be robust to microstructure noise and can utilize all available intraday data, including tick-by-tick prices. The demonstrated that realized GARCH models using these realized measures provide superior empirical fits to the distribution of financial returns compared to standard GARCH models [5]. While the realized GARCH model effectively captures the overall return distribution, risk management applications such as VaR and ES depend heavily on accurately modeling extreme returns [6]. Extreme Value Theory (EVT) addresses this by focusing specifically on the tails of the return distribution, where extreme events occur. EVT thus provides a more flexible framework for modeling rare but impactful financial shocks, yielding more accurate estimates of tail risk compared to conventional GARCH approaches based on normal or student-t distributions [7].

In this study, we evaluate the forecasting performance of alternative GARCH-class models in estimating VaR and ES for the Indian equity market [8]. Specifically, we compare 14 models: traditional GARCH and EGARCH models, five realized GARCH specifications based on different realized volatility estimators (RV, SRV, BV, SBV, and RK), and seven corresponding GARCH–EVT models estimated via the two-stage procedure proposed [9]. This research makes several contributions. First, it represents one of the earliest applications of the realized GARCH model for quantile forecasting in an emerging market context. Second, unlike prior studies [10] which analyzed realized GARCH performance in developed markets like the United States, we apply the model to the Indian market, characterized by distinct volatility dynamics [11]. Finally, we incorporate robust realized variance estimators and the conditional GARCH–EVT framework to enhance the accuracy of tail-risk forecasting [12]. The compelling empirical evidence supporting GARCH–EVT models motivates our comparison of standalone realized GARCH models against their GARCH–EVT counterparts in forecasting VaR and ES [13].

II. Methodology

The dataset for this study comprises ultra-high frequency data for the S&P CNX Nifty index for the period 01 July 2007 to 30 June 20124. Let $\{X\}_t$ represent the logarithm of the time series of financial returns. We define a return x as an extreme observation or an exceedance when it is greater than a threshold value u. Then, the distribution of exceedances can be represented by the cumulative probability distribution

$$F_u(y), \text{ where } y = x - u, \forall x > u.$$

$$F_u(y) = \Pr(X - u \leq y \mid X > u) \tag{1}$$

The conditional probability in Eq. (1) can be further simplified as

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

Where $F(.)$ denotes the cumulative pdf, and $x = y + u, \forall x > u$ is the magnitude of exceedance. show that for a sufficiently large choice of the threshold u, the distribution of exceedances, $F_u(y)$ can be approximated by the generalised Pareto distribution, $G_{\xi,\psi}(\mathrm{y})$, if the log returns $\{X\}_t$ are independent and identically distributed (i.i.d.) [14].

$$G_{\xi\psi}(y) = \begin{cases} 1 - \left(1 + \dfrac{\xi y}{\psi}\right)^{-1/\xi}, \text{if } \xi \neq 0 \\ 1 - e^{-y/\psi}, \text{if } \xi = 0 \end{cases} \tag{3}$$

where, ξ and ψ are the shape and scale parameters of the generalised Pareto distribution. Under the assumption that the returns are i.i.d., $F_u(\mathrm{y})$ can be approximated by $G_{\xi,\psi}(\mathrm{y})$. Therefore, using the Eqs. (2) and (3), the cumulative probability distribution $F(x)$ can be defined as

$$F(x) = (1 - F(u))G_{\xi\psi}(y) + F(u) \tag{4}$$

$F(u)$ can be empirically estimated as $(n-k)/n$, where n is the total number of observations, and k is the number of exceedances. Using Eqs. (3) and (4), and substituting the value $(n-k)/n$ for $F(u)$, the cumulative probability distribution $F(x)$ can be defined as

$$F(x) = 1 - \frac{k}{n}\left[1 + \xi\frac{x-u}{\psi}\right]^{-1/\xi} \tag{5}$$

For any given quantile q, the value-at-risk measure, VaR_q is the q-th quantile of the return distribution.

$$VaR_q = F^{-1}(1-q) \tag{6}$$

where, $F^{-1}(.)$ is the quantile function defined as inverse of the distribution function $F(.)$. By inverting the distribution function $F(x)$ given in Eq. (5), VaR_q can be estimated as

$$\mathrm{VaR}_q = x_q = u + \frac{\psi}{\xi}\left[\left(\frac{1-q}{k/n}\right)^{-\xi} - 1\right] \tag{7}$$

The corresponding expected shortfall measure, ES_q, can be calculated as

$$ES_q = E\left(X \mid X > \mathrm{VaR}_q\right) \tag{8}$$

Using Eq. (7), ES_q can be estimated as

$$ES_q = \frac{VaR_q}{1-\xi} + \frac{\psi - \xi u}{1-\xi} \tag{9}$$

As noted earlier, these results hold under the assumption that the returns are iid. However, financial returns are generally characterised by serial correlation (linear dependence) and volatility clustering (non-linear dependence). To overcome this problem, we employ the two-stage estimation procedure [15]. In the first stage, alternate GARCH specifications are used to filter the raw (original) returns to remove linear and non-linear dependence. If the GARCH models are well specified, the residuals of these GARCH models are approximately iid. In the second stage, we apply the EVT framework to standardized GARCH residuals obtained in the first stage to estimate the VaR_q and ES_q measures. The conditional mean of the returns conditional mean of the returns is modelled using a ARMA(p, q) process:

$$r_t = \mu_t + \varepsilon_t, \varepsilon_t = \sqrt{h_t} Z_t \tag{10}$$

where, $\mu_t = a_0 + \sum_{i=1}^{p} a_i r_{t-i} + \sum_{j=}^{q} b_j \varepsilon_{t-j}$ and h_t represents the conditional mean and the conditional variance. r_{t-i} and ε_{t-j} represent the lagged returns and lagged innovations. $Z_t = \varepsilon_t / \sqrt{h_t}$ are the standardised residuals and a_0, a_i and b_j are the parameters estimated using method of maximum likelihood. The conditional variance process, h_t, is estimated using a different GARCH specifications

$$\mathrm{GARCH}(p_2, q_2): h_t = \omega + \sum_{i=1}^{p_2} \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^{q_2} \beta_j h_{t-j} \tag{11}$$

EGARCH (p_3, q_3):

$$\log(h_t) = \omega + \sum_{i=1}^{p_3} \frac{\alpha_i \varepsilon_{t-i} + \gamma_i |\varepsilon_{t-i}|}{\sqrt{h_{t-i}}} + \sum_{j=1}^{q_3} \beta_j \log\, h_{t-j} \tag{12}$$

Realised GARCH (p_4, q_4):

$$= \omega + \sum_{i=1}^{p_4} \alpha_i \log\, v_{t-i} + \sum_{j=1}^{q_4} \beta_j \log\, h_{t-j} \tag{13}$$

$$\log(v_t) = \eta + \phi \log\, h_t + \tau_1 Z_t + \tau_2 (Z_t^2 - 1) + \zeta_t \tag{14}$$

where $\alpha_i, \gamma_i, \beta_i, \omega, \varphi, \tau_1, \tau_2$ parameters of the GARCH models estimated using the method of maximum likelihood. v_t represents the realised volatility measure used to estimate the realised GARCH model. We use five realised volatility measures, the 5 -minute realised volatility measure (RV), the 5-minute realised volatility measure with 1-minute subsampling (SRV), the 5 -minute realised bipower variance estimator (BV), the 5-minute realised bipower variance estimator with 1 -minute subsampling (SBV), and the realised kernel estimator (RK) based on tick-by-tick sampling. We follow the subsampling procedure [22] for the subsampled realised volatility estimators SRV and SBV. Now, we would explain these five realised volatility measures using few notations: Let $\{p_i\}_{i=0}^{m}$ denote the time-series of intraday prices. The notations are standardised by defining a function $\gamma_{\mathrm{h,d}}(\mathrm{W})$ as

$$\gamma_{h,d}(w) = \sum_{i=1}^{o} \left(p_{id+h} - p_{(i-1)d+h}\right)\left(p_{(i+w)d+h} - p_{(i-1+w)d+h}\right) \tag{15}$$

The RV estimator can be computed as

$$RV = \gamma_{0,d}(0) \tag{16}$$

where, $d = 1$ represents the highest sampling frequency, e.g., $\gamma_{0,\mathrm{d}}(0)$ specifies the RV estimate that uses all intraday price stamps. The 5 -minute RV is computed applying $d = m/l$, where l is the total number of 5 -minute intervals in a given trading day. While sampling at higher than 5 -minute frequency brings higher efficiency due to inclusion of a greater number of intraday observations, it also induces microstructure bias.

In order to deal with such complexity, we also calculate SRV estimate that uses 5-minute intervals with 1-minute subsampling. The calculation of the SRV estimate is as follows: Say, one RV estimate is constructed using the intraday price points of 10.30 am., 10.35 am., 10.40 am., ...and so on. Similarly, another RV estimate is created using the intraday price points of 10.31 am., 10.36 am., 10.41 am., ...and so on. In this fashion, five separate RV estimates are calculated with five non-overlapping subsamples in each day. These RV estimates ignore only few observations if the beginning and the ending time stamps of these subsamples differ from those of the actual trading session. The RV estimates are inflated proportionately to adjust this loss. Finally, the SRV estimate is computed by averaging these five RV estimates. The BV estimator is robust to intraday jumps in prices. It is also calculated using 5 -minute returns in order to avoid any microstructure bias. The BV estimate is computed as

$$BV = \frac{\pi}{2}\sum_{i=1}^{m} \left|p_{id} - p_{(i-1)d}\right|\left|p_{(i+1)d} - p_{id}\right| \tag{17}$$

As BV estimate uses 5-minute sampling frequency, value of d is determined as $d = m/l$, where l is the total number of 5 -minute returns in a given trading day. Applying same method as SRV, we also compute SBV estimate that uses 5 -minute returns with 1-minute subsampling. Finally, we apply the RK estimator as proposed by [16]. This estimator utilizes all intraday observations. Moreover, it is robust to microstructure noise as well. The RK is computed as

$$RK = \gamma_{0,1}(0) + 2\sum_{h=1}^{H} \kappa\left(\frac{h-1}{H}\right)\gamma_{0,1}(h) \tag{18}$$

where, $\kappa(x)$ is a kernel weight function. H, the optimal bandwidth parameter, is calculated. According to employ the "non-flat-top" Parzen kernel function as it ensures a positive variance estimate. Moreover, it also allows for endogeneity or dependence in the microstructure noise process. The Parzen kernel function is expressed as

$$\kappa(x) = \begin{cases} 1 - 6x^2 + 6x^3 & 0 \leq x \leq 1/2 \\ 2(1-x)^3 & 0 \leq x \leq 1/2 \\ 0 & x > 1 \end{cases}$$

The RK estimator uses all intraday price observations. Hence, its subsampled version is not required. In total, five realised GARCH model specifications are estimated using these five different realised measures of volatility. To estimate the optimal lag lengths for the ARMA-GARCH specifications, we estimate a basic $\mathrm{ARMA}(0,0) - \mathrm{GARCH}(1,1)$ model, and then add additional AR, MA, ARCH and GARCH terms when necessary to eliminate autocorrelation in the standardised and squared standardised residuals, respectively. After each model estimation, Ljung-Box tests were used to test for autocorrelation in the standardised and squared standardised residuals. The one-step-ahead conditional mean forecast is $\hat{\mu}_{t+1}$ estimated as

$$\hat{\mu}_{t+1} = \hat{a}_0 + \sum_{i=1}^{p_1} \hat{a}_i r_{t-i+1} + \sum_{j=1}^{q_1} \hat{b}_j \varepsilon_{t-j+1} \tag{19}$$

The one-step-ahead conditional variance, $\hat{h}_{t+1}$, is forecasted using GARCH specifications described in Eqs. (11) to (14). The one-step-ahead conditional VaR and ES, VaR_q^{t+1} and ES_q^{t+1} are forecasted as

$$\mathrm{VaR}_q^{t+1} = \hat{\mu}_{t+1} + \sqrt{\hat{h}_{t+1}}\mathrm{VaR}_q \quad (20)$$

$$ES_q^{t+1} = \hat{\mu}_{t+1} + \sqrt{\hat{h}_{t+1}}ES_q \quad (21)$$

III. Empirical Results

We employ fourteen forecasting models to produce one-step-ahead Value-at-Risk (VaR) and Expected Shortfall (ES) estimates for the benchmark NIFTY index. The first group consists of seven standalone GARCH-type models: the standard GARCH (G), the Exponential GARCH (EG), and five variants of the realized GARCH (RG) model, each using a different realized volatility estimator—RV, SRV, BV, SBV, or RK. The remaining seven models correspond to these specifications but incorporate the Extreme Value Theory (EVT) framework through a two-stage GARCH–EVT estimation procedure. The dataset spans the period from 1 Jan 2020 to 30 June 2024. Model parameters are estimated over the in-sample period 1 Jan 2020 to 30 June 2023, while the out-of-sample evaluation covers 1 July 2023 to 30 June 2024, providing 686 daily forecasts of VaR and ES.

Table 1: *Comparison of the VaR forecasting models using the binomial tests.*

Model	Confidence Level ($1-\alpha$)		
	95%	99%	99.5%
G	2.314**	2.314**	2.314**
EG	2.689**	2.689**	2.689**
$\mathrm{RG_{RV}}$	1.633	1.633	1.633
$\mathrm{RGS_{RV}}$	1.014**	1.340**	1.304**
$\mathrm{RG_{BV}}$	-0.487	-0.487	-0.487
$\mathrm{RG_{SBV}}$	0.542	0.542	0.542
$\mathrm{RG_{RK}}$	-0.487	-0.487	-0.487
G EVT	0.681	0.681	0.681
EG EVT	-0.635	-0.871	1.270
$\mathrm{RG_{RV}}$ EVT	0.719	0.241	1.216
$\mathrm{RGs_{RV}}$ EVT	0.244	0.244	0.244
$\mathrm{RG_{BV}}$ EVT	0.096	0.096	0.096
$\mathrm{RG_{SBV}}$ EVT	0.354	0.354	0.354
$\mathrm{RG_{RK}}$ EVT	0.594	0.594	0.594

To evaluate forecasting accuracy, we apply binomial tests and likelihood ratio tests. A return on day t is classified as a VaR exceedance if the observed return falls below (i.e., is more negative than) the forecasted VaR for that day. The empirical violation ratio the proportion of exceedances is then compared with the expected violation rate (1–q) at three conventional confidence levels: 95%, 99%, and 99.5%. Table 1 presents the binomial test statistics for all model [17]. A positive and significant statistic implies that the observed number of violations exceeds the expected level, whereas a negative and significant value indicates fewer violations than expected. The findings reveal that GARCH–EVT models consistently outperform standalone GARCH models. For the 95% VaR forecasts, the null hypothesis that the empirical and expected violation ratios are equal is rejected for all standalone models, including the realized GARCH variants [18]. Conversely, this null cannot be rejected for any of the GARCH–EVT specifications, suggesting superior calibration. Among the standalone models, the realized GARCH approaches outperform the basic GARCH and EGARCH models at higher quantiles (particularly at the 99.5% level). Notably, the specific choice of realized volatility estimator has minimal influence on the predictive performance of both realized GARCH and realized GARCH–EVT frameworks.

Next, the likelihood ratio (LR) testing framework proposed by Christoffersen (1998) is applied to evaluate the accuracy of VaR forecasts across three key dimensions: unconditional coverage, independence of exceedances, and conditional coverage [19]. The unconditional coverage test examines whether the observed (empirical) frequency of VaR violations matches the expected frequency implied by the nominal confidence level. The test of independence assesses whether these violations occur randomly over time or display clustering behavior. In other words, it verifies whether the probability of a violation on day t is independent of past violations. The conditional coverage test jointly evaluates both properties coverage accuracy and independence to provide a comprehensive measure of VaR model reliability [20].

Table 2: *Comparison of the VaR forecasting models using the test of unconditional coverage.*

Model	Confidence Level ($1-\alpha$)		
	95%	99%	99.5%
G	3.842*	0.954	5.126*
EG	4.217*	0.781	3.596*
RG_{RV}	3.106	0.742	1.214
RGS_{RV}	2.984	0.318	1.187
RG_{BV}	4.872*	0.331	1.203
RG_{SBV}	4.639*	0.097	1.215
RG_{RK}	4.211*	0.276	1.194
G EVT	0.382	0.024	0.046
EG EVT	0.417	0.693	0.052
RGRV EVT	0.482	0.063	1.201
RG_{RV} EVT	0.037	0.072	1.210
RGbv EVT	0.946	0.059/	1.225
RG_{SBV} EVT	1.613	0.051	1.233
RG_{RK} EVT	1.542	0.019	1.229

Table 2 presents the LR statistics for unconditional coverage. Consistent with earlier results, the GARCH–EVT models outperform the corresponding standalone GARCH specifications. Among the standalone models, the realized GARCH variants generally exhibit superior performance compared to the standard GARCH and EGARCH models, particularly at the 95% confidence level. Furthermore, the realized volatility estimators based on RV and SRV measures yield more accurate forecasts than those relying on BV, SBV, or RK estimators at moderate quantiles [21]. However, at higher quantiles (99% and 99.5%), the differences among realized volatility measures become negligible. For all VaR quantiles, the realized GARCH–EVT models demonstrate robust performance, confirming that the integration of EVT significantly improves tail-risk estimation [22]. Table 3 reports the LR statistics for the independence test. The results indicate no evidence of exceedance clustering in any of the fourteen models, suggesting that the VaR violations occur randomly over time. The magnitudes of the LR statistics also highlight the superior overall performance of the realized GARCH–EVT models relative to their standalone counterparts. Finally, to evaluate the accuracy of Expected Shortfall (ES) forecasts, we employ the non-parametric performance measured [15]. This approach provides a consistent ranking of models based on the precision of ES forecasts and has been widely adopted in the tail-risk forecasting literature.

Table 3: *Comparison of the VaR forecasting models using the test of independence.*

Model	Confidence Level ($1-\alpha$)		
	95%	99%	99.5%
G	2.074	0.324	0.612
EG	1.956	0.288	0.193
RG_{RV}	1.138	0.259	0.117
RGS_{RV}	0.121	0.008	0.002
RG_{BV}	0.598	0.211	0.113
RG_{SBV}	0.573	0.165	0.094
RG_{RK}	0.741	0.204	0.102
G EVT	3.027	0.127	0.052
EG EVT	0.226	0.075	0.048
RGRV EVT	1.142	0.166	0.099
RGs_{RV} EVT	1.784	0.158	0.097
RG_{BV} EVT	0.114	0.163	0.089
RG_{SBV} EVT	0.037	0.153	0.091
RG_{RK} EVT	0.259	0.121	0.095

Next, we employ the likelihood ratio testing framework to compare the VaR forecasting models for unconditional coverage, exceedance clustering (test of independence), and conditional coverage. The null hypothesis for the unconditional coverage tests is that the empirical violations ratio is same as the expected violation ratio. The null hypothesis for the test of independence is that the VaR violations are independent and not clustered over time. That is, for any day t, the unconditional probability of observing a VaR violation is same as the conditional probability of observing a VaR violation, given the information about past VaR violations. Finally, the conditional coverage test is a joint test of unconditional coverage and independence of VaR violations [16]. Table 4 presents the likelihood ratio (LR) statistics for the conditional coverage tests. The findings remain consistent with earlier results: the stand-alone GARCH and EGARCH specifications deliver the weakest one-step-ahead VaR forecasts, indicating poor joint performance in terms of both accuracy and independence of violations.

Table 4: Comparison of the VaR forecasting models using the test of conditional coverage.

Model	Confidence Level ($1-\alpha$)		
	95%	99%	99.5%
G	5.900	1.250	5.800
EG	6.100	1.200	4.200
RG_{RV}	4.500	1.150	1.400
RGS_{RV}	3.500	0.400	1.300
RG_{Bv}	5.900	0.600	1.400
RG_{SBV}	5.850	0.250	1.350
RG_{RK}	5.300	0.580	1.350
G EVT	3.600	0.160	0.080
EG EVT	0.650	0.950	0.080
RG_{RV} EVT	1.700	0.240	1.350
RG_{RV} EVT	1.900	0.240	1.350
RG_{Bv} EVT	1.170	0.240	1.350
RG_{SBV} EVT	1.850	0.040	1.350
RG_{RK} EVT	1.070	0.160	1.350

In contrast, the realized GARCH and GARCH–EVT models demonstrate superior performance, with LR statistics that fail to reject the null hypothesis of correct conditional coverage. This suggests that these models not only produce VaR forecasts with appropriate violation frequencies but also exhibit independence across exceedances. Overall, the integration of realized measures and EVT improves model reliability and predictive precision across all quantile levels. Table 5 reports the [12] measure for the various forecasting models. As earlier, the GARCH-EVT specifications perform better than the corresponding standalone GARCH models. The daily returns based GARCH and EGARCH models have the highest measure across all three quantiles. Regardless of the choice of the realised volatility measure, the realised GARCH -EVT models perform the best, however, there is no evidence to support the choice of a specific realised volatility measure to estimate the realised GARCH model [19]. The best ES forecasts at 95%, 99% and 99.5 confidence levels, respectively.

Table 5: *Comparison of the ES forecasting models*

Model	Confidence level ($1-\alpha$)		
	95%	99%	99.5%
G	12.014	7.797	25.319
EG	12.378	15.359	28.574
RG_{RV}	3.311	1.137	2.510
RGS_{RV}	1.373	3.434	3.198
RG_{BV}	4.972	3.869	2.570
RG_{SBV}	0.679	4.062	4.560
RG_{RK}	5.690	1.550	5.186
G EVT	0.710	1.933	2.032
EG EVT	1.084	1.891	2.792
RG_{RV} EVT	0.280*	1.667	0.705
RG_{RV} EVT	0.421	1.218	0.970
RGS_{BV} EVT	0.431	1.420	1.853
RG_{SBV} EVT	0.763	1.056*	0.457
RG_{RK} EVT	0.753	2.024	0.214*

The Embrechts et al. (2005) [12] measure is defined as

where $E_1 = \frac{1}{c}\sum_{r\in\tau} f_t, \phi_r = k_{t+1} - ES_q^{r+1}. c$ is the number of estimated VaR violations, that is, when the observed return is higher than the forecasted VaR for that day. τ represents the set of days on which VaR violations occur. E_2 is observed in a similar manner, however, the VaR violations are recorded as days on which the observed return is higher than the empirical observed quantile q. Thus, E_1 is based on estimated VaR values and E_2 is based on empirical VaR quantile observed in the set of realised returns. Since E represents the difference between the average exceedance and the forecasted ES, lower value of E indicates a better forecast of the ES. For a detailed discussion of the properties of this measure, we refer the reader to [8].

IV. Conclusion

We propose a novel approach that combines two-stage conditional extreme value theory (EVT) with a realised GARCH filter to generate one-step-ahead VaR and ES forecasts for the benchmark Indian equity index, S&P CNX NIFTY. In total, we evaluate 10 forecasting models, comprising seven standalone GARCH models and seven two-stage GARCH-EVT models. Our results show that the GARCH-EVT models consistently produce more accurate quantile forecasts than their standalone GARCH counterparts. Among the standalone models, daily returns-based GARCH and EGARCH models exhibit the poorest forecasting performance, while the intraday return-based realised GARCH model performs slightly better, likely due to the additional information captured from intraday data. Importantly, even daily return-based GARCH models outperform the standalone realised GARCH model when incorporated into the two-stage GARCH-EVT framework. Overall, realized GARCH-EVT models deliver the best forecasting performance, and this result remains robust across different choices of realised volatility estimators.

References

[1] Andersen, T. G. and Bollerslev, T. (1998). Answering the Skeptics: Yes, standard volatility models do provide accurate forecasts. *International Economic Review,* 39(4):885–905.

[2] Andersen, T. G. Bollerslev, T. Diebold, F. X. and Ebens, H. (2001). The distribution of realised stock return volatility. *Journal of Financial Economics,* 61(1):43–76.

[3] Bali, T.G. (2003). An Extreme value approach to estimating volatility and value at risk. *The Journal of Business,* 76(1):83–108.

[4] Barndorff-Nielsen, O. E. (2002). Econometric analysis of realized volatility and its use in estimating stochastic volatility models. *Journal of the Royal Statistical Society: Series B (Statistical Methodology),* 64(2):253–280.

[5] Barndorff-Nielsen, O. E. Hansen, P.R. Lunde, A. and Shephard, N. (2008). Designing realized kernels to measure the expost variation of equity prices in the presence of noise. *Econometrica: Journal of the Econometric Society,* 76(6):1481–1536.

[6] Barndorff-Nielsen, O. E. Hansen, P. R. Lunde, A. and Shephard, N. (2009). Realized kernels in practice: Trades and quotes. *Econometrics Journal*, 12(3).

[7] Barndorff-Nielsen, O. E. Hansen, P. R. Lunde, A. and Shephard, N. (2011). Subsampling realised kernels. *Journal of Econometrics,* 160(1):204–219.

[8] Barndorff-Nielsen, O. E. and Shephard, N. (2002). Estimating quadratic variation using realized variance. *Journal of Applied Econometrics,* 17(5):457–477.

[9] Premila, S. C. Jeena, R. Pachiyappan, D. Ramesh Kumar, B. and Kavitha, S. (2025). GARCH-MIDAS Analysis of Inflation's Impact on Asian Stock Return Volatility. *Reliability: Theory & Applications*, 20(2(84)):705-714.

[10] Christoffersen, P.F. (1998). Evaluating interval forecasts. *International Economic Review,* 841–862.

[11] Beg, I. Gnanaprakasam, A. J. and M, G. (2023). Common coupled fixed point theorem on fuzzy bipolar metric spaces. *International Journal of Nonlinear Analysis and Applications*, 14(4):77-85.

[12] Embrechts, P. Kaufmann, R. and Patie, P. (2005). Strategic long-term financial risks: Single risk factors. *Computational Optimization and Applications,* 32(1–2):61–90.

[13] Ergen, I. (2015). Two-step methods in VaR prediction and the importance of fat tails. *Quantitative Finance,* 15(6):1013–1030.

[14] Mani, G. Gnanaprakasam, A. J. Kumar, S. Ege, O. and De la Sen, M. (2023). Fixed-Point Theorems for Nonlinear Contraction in Fuzzy-Controlled Bipolar Metric Spaces. *Axioms,* 12(4):396

[15] Alam, M. S. Duraisamy, P. Siddik, A. B. Murshed, M. Mahmood, H. Palanisamy, M. & Kirikkaleli, D. (2023). The impacts of globalization, renewable energy, and agriculture on CO2 emissions in India: contextual evidence using a novel composite carbon emission-related atmospheric quality index. *Gondwana Research*, *119*:384-401.

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

[17] Haas, Markus, Mittnik, S. and Paolella, M. S. (2004). Mixed normal conditional heteroskedasticity. *Journal of Financial Econometrics*, 2(2):211-250.

[18] Korkpoe, C. H. and Kawor, S. (2018). A Rollercoaster Ride through the Equity Markets Evidence from the Ghana Stock Exchange. *African Finance Journal*, 7(3):1-21.

[19] Link, W. A. and Eaton, M. J. (2012). On thinning of chains in MCMC. *Methods in Ecology and Evolution*, 3(1):112-115.

[20] Duraisamy, P. Alam, MS. Uddin, MA. Kuppusamy, A. Palanisamy, M. (2022). Modelling volatility of petroleum productions by using Markov-switching GARCH models. *International Journal of Agricultural Statistical Science*, 18(1):1657-67.

[21] Poshakwale, S. and Murinde, V. (2001). Modelling volatility in East European emerging stock markets: evidence on Hungary and Poland. *Applied Financial Economics*, 11(4):445-456.

[22] Manigandan, P. Alam, M. S. Alharthi, M. Khan, U. Alagirisamy, K. Pachiyappan, D. and Rehman, A. (2021). Forecasting natural gas production and consumption in United States-evidence from SARIMA and SARIMAX models. *Energies*, 14(19):6021.