Forecasting Value-at-Risk Measures of Financial Time Series Using GARCH-EVT Theory

Vivek Kantamani (vk2389)

Introduction

Under the assumption of normally distributed and mutually independent log-returns, traditional mathematical finance techniques suggest that prices follow a log-normal geometric random walk (i.e. geometric Brownian motion with drift, in the continuous case).

Although distributions of log-returns typically exhibit characteristics that violate these underlying assumptions (namely, volatility clustering and heavy-tails), such financial models are the industry standard for characterizing distributions of prices and returns.[1]

In this setting, VaR is a statistic that quantifies the extent of potential losses under specific market conditions associated with an investment position over a specified time frame, commonly used by financial institutions as a risk-management measure to determine the probabilities associated with potential losses in their investment portfolios. In practice, VaR is used as a measure of downside investment risk (tail risk) - the risk of extreme market events.

In particular, VaR refers to the maximal possible loss which should not be exceeded for a specified significance level α during a specified time horizon T. Thus, VaR can be characterized as the quantile of the lower tail (the loss tail) of the return distribution, or simply the upper quantile of the distribution of losses.[2]

Let the loss function L_t specified for a fixed portfolio over a time horizon T, $t-1$ to t, be defined as

$$L_t = V_{t-1} - V_t = -R_t V_{t-1}$$

Where
i. V_t indicates the value of the value of the portfolio at time t, and
ii. $R_t = \frac{V_t - V_{t-1}}{V_{t-1}}$ specifies the return of the portfolio at time t.

Then VaR is given by a bound over the loss function L over the time horizon T for significance level α defined as[2]

$$P(L_t > VaR(\alpha, T)) = \alpha$$

VaR models are traditionally constructed using historical returns, the variance-covariance method (characterized by a normal approximation of historical returns), or Monte-Carlo simulation, i.e. the use of computational models (primarily, exponentially-weighted moving average models) to estimate volatility and simulate projected returns.[3,4]

Methods for evaluating VaR models are referred to as backtesting methods, which compare the true losses achieved on a portfolio to calculated VaR measures over a suitable period of time. The difficulties of backtesting VaR measures are manifold, including (but not limited to) sampling variation, misspecification of the return distribution, and inaccurate measurement of losses - the infrequency of tail events makes it particularly difficult to measure loss tail probabilities based on historical data.[5]

The difficulty associated with the specification and model-validation of VaR models is compounded when considering forecasting VaR measures. Recent research suggests that the shortcomings of historical and variance-covariance estimation of VaR measures can be improved by modeling the conditional volatility of financial time series using Generalized Autoregressive Heteroskedastic (GARCH) models. In particular, GARCH models of return distributions can be used to estimate conditional loss distributions, which can subsequently be used to estimate and forecast risk measures like VaR.[6]

With the introduction of more sophisticated conditional volatility estimates, GARCH-VaR models have the potential to improve estimation of risk measures due to their ability to account for time-dependent volatility criteria.[6]

This report aims to synthesize literature on prediction of VaR measures in GARCH-VaR models through a case study examining the implementation of VaR forecasting techniques using GARCH-EVT theory.
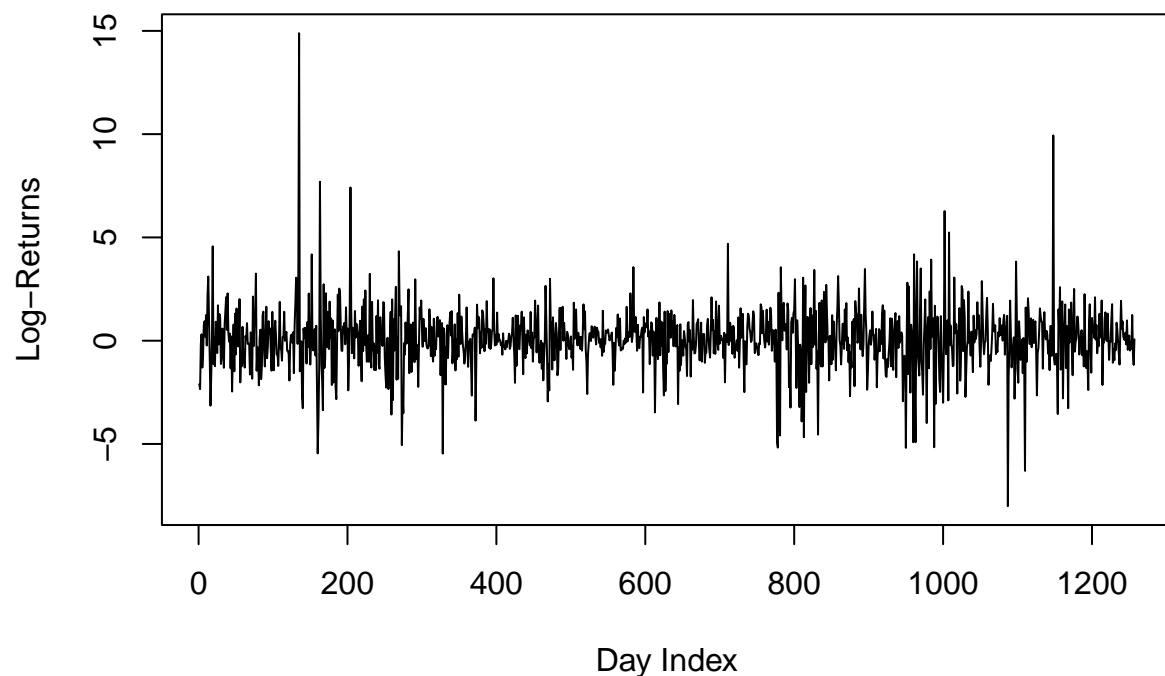
Exploratory Data Analysis

We will construct log-returns from daily adjusted closing prices for Alphabet Inc (`GOOG`) from 2015-01-01 to 2020-01-01.

## GOOG Adjusted Daily Closing Prices: 2015–01–01 to 2020–01–01
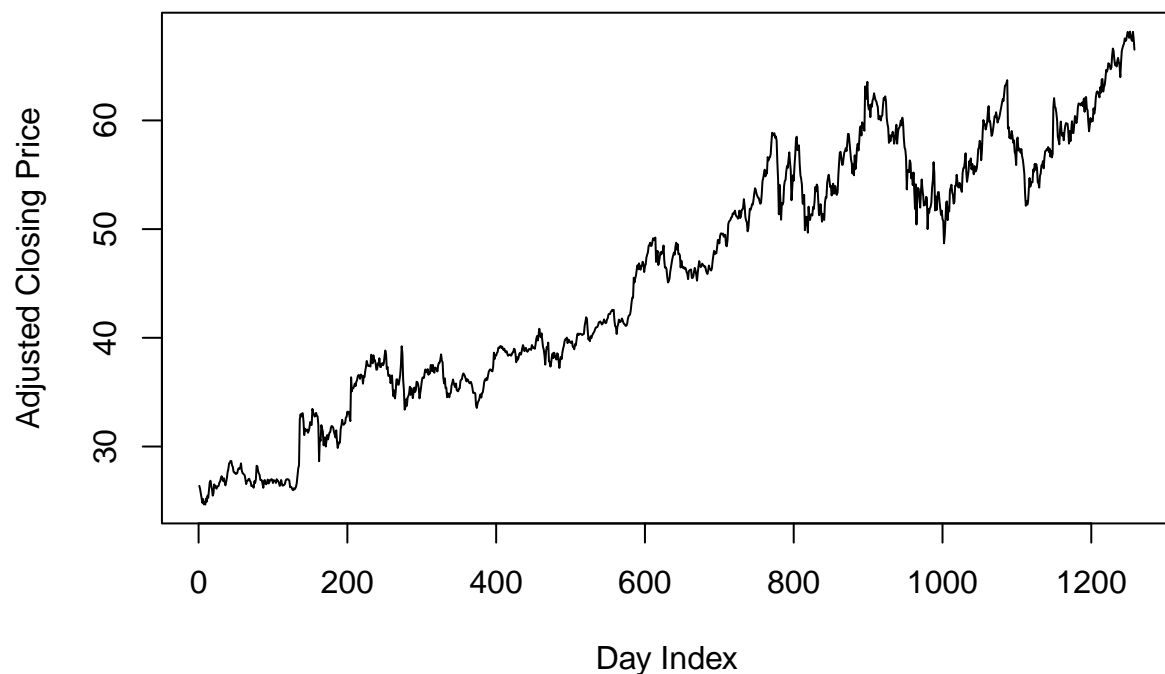
GOOG Log–Returns: 2015–01–01 to 2020–01–01

On the basis of the ACF and PACF plots of the log-returns, we note that there is no strong indication of AR(p) or MA(q) structure in our data. Furthermore, as indicated by the KPSS test ($\alpha = 0.05$), the time series of log-returns is stationary (likely due to the differencing necessary to construct log-returns from prices) which will facilitate forecasting methods.

Because the ARMA(p,q) model does not appear appropriate for our data (on the basis of exploratory data analysis), we will consider GARCH(r,s) models to fit our data. To this effect, we note that the Ljung-Box test indicates serial correlation ($\alpha = 0.05$) in the log-returns for both a moderate number (20) and large number (100) of lags.
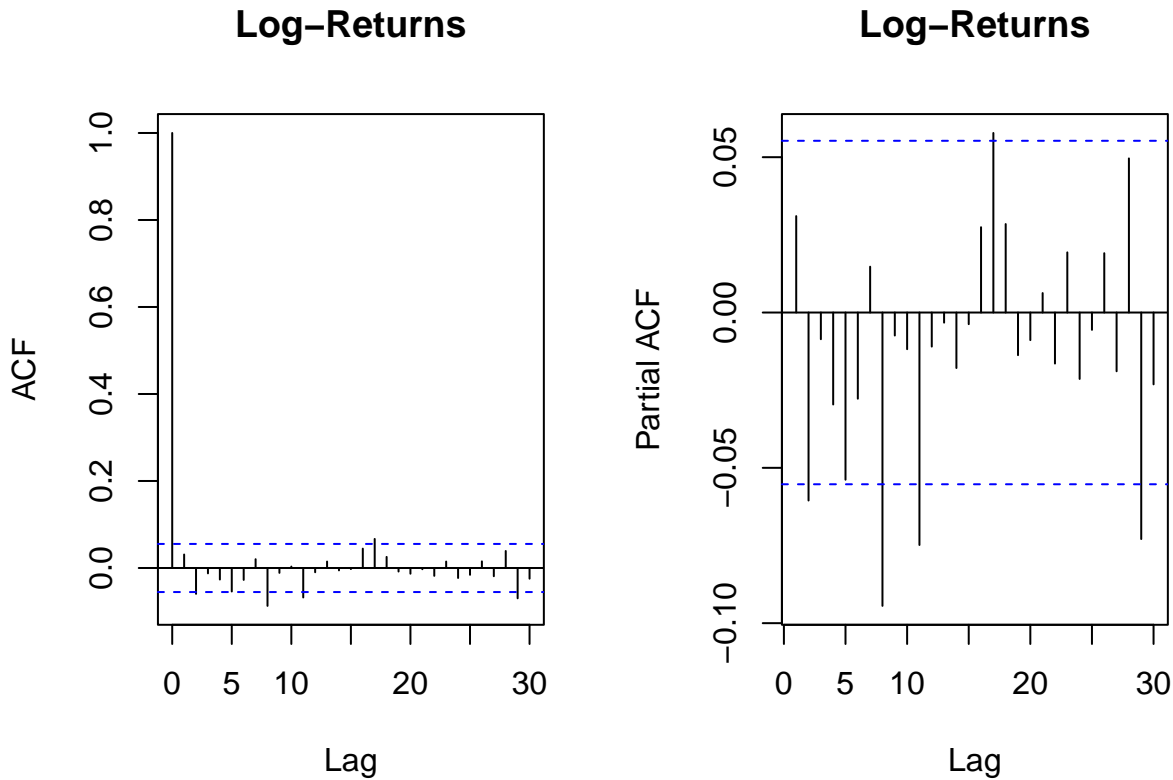
KPSS Test p-value	Result
0.1	Stationary

	Ljung-Box Test p-value	Result
20 Lags	0.0106	Serial Autocorrelation
100 Lags	0.005816	Serial Autocorrelation

Time Series Model Validation

Previous literature has suggested that univariate GARCH models can produce stable extrapolative estimates of VaR and volatility when the models are tuned with respect to volatility clustering and the assumption of non-normal conditional distributions. Furthermore, multivariate GARCH models seeking to improve forecasting results by accounting for the comovements between markets during periods of high-volatility have shown only moderate improvements relative to the increased computational intensity of multivariate modeling (i.e. "the curse of dimensionality").[6]

Let r_t denote the log-return of a portfolio between time periods $t-1$ and t, and let $\mathcal{F}_t$ denote the information filtration (a family of non-decreasing σ-algebras) associated with the distribution of returns.

We can define the distribution of log-returns as

$$r_t = \mu_t + a_t$$

Where
i. $E[r_t|\mathcal{F}_{t-1}] = \mu_t$ indicates the conditional mean of r_t.
ii. $\sigma_t^2 := Var(r_t|\mathcal{F}_{t-1}) = Var(a_t|\mathcal{F}_{t-1})$ indicates the conditional variance of r_t.

GARCH models are used to model the stochastic dynamics of conditional variances.
We say that σ_t^2 follows a GARCH(r,s) model if
i. $a_t = \sigma_t \varepsilon_t, \quad \varepsilon_t \sim IID(0,1)$, and
ii. $\sigma_t^2 = \alpha_0 + \sum_{i=1}^{r} \alpha_i a_{t-i}^2 + \sum_{j=1}^{s} \beta_j \sigma_{t-j}^2$
For $\alpha_0 > 0$, $\alpha_i \geq 0$ for $i = 1, \ldots, r$, and $\beta_j \geq 0$ for $j = 1, \ldots, s$.[7]

The literature suggests that low-dimensional model orders $r, s \in (0,1)$ are used in practice, with the GARCH(1,1) model being the most prevalent to model log-return distributions.[8]

We will primarily consider a univariate GARCH framework for our model, introducing tuning parameters for conditional volatility and accommodating a variety of distributions - normal, student, generalized error distribution - and their skewed counterparts. We will account for the possibility of modeling a conditional mean structure by broadening the framework to include ARMA-GARCH models (although exploratory data analysis indicates no significant ARMA(p,q) structure).

In order to assess goodness-of-fit among ARMA-GARCH models, model-selection will be predicated on the corrected Akaike Information Criteria (AICC), Schwarz's Bayesian Information Criterion (SBC) and the Hannan-Quinn Information Criterion (HQC) defined as follows:

$$AICC = -2\ln(L) + 2k + 2\frac{k(k+1)}{N-k-1}$$

$$SBC = -2\ln(L) + \ln(N)k$$

$$HQC = -2\ln(L) + 2\ln(\ln(N))k$$

Where
i. L is the value of the likelihood function evaluated at the parameter estimates.
ii. N is the number of observations.
iii. k is the number of estimated parameters.[9]

To facilitate parsimony, note that values between 0 and 2 will be considered for the model parameters (p, q, r, s) of the ARMA(p,q)-GARCH(r,s) model.

Across all goodness-of-fit measures, it is clear that the ARMA(0,0)-GARCH(1,0) model with a skewed normal error distribution provides the best fit. The lack of an ARMA component synthesizes with exploratory data analysis results.

When performing diagnostics on our GARCH(1,0) model, we note that
i. The residuals are normally distributed, validating the residual structure of skewed Normal (Shapiro-Wilk test, $\alpha = 0.05$).
ii. The model captures the serial auto-correlation of the data (Weighted Box-Ljung test, $\alpha = 0.05$).

	p-value	Result
Shapiro-Wilk Test	0	Normally Distributed Residuals
Weighted Ljung-Box Test	>0.3369	No Serial Autocorrelation (Standardized Residuals)

	Residual Distribution	Likelihood	AICC	SBC	HQC
ARMA(0,0)+GARCH(1,0)	Skewed Normal	2260.887	-13.444	-8.311	-1.174
ARMA(0,0)+GARCH(1,1)	Skewed Normal	2238.409	-11.417	-1.154	13.119
ARMA(0,0)+GARCH(2,0)	Skewed Normal	2255.635	-11.433	-1.169	13.104
ARMA(1,0)+GARCH(1,0)	Skewed Normal	2259.122	-11.436	-1.172	13.1
ARMA(0,1)+GARCH(1,0)	Skewed Normal	2259.06	-11.436	-1.172	13.101
ARMA(1,1)+GARCH(1,0)	Skewed Normal	2258.557	-9.426	5.964	27.374
ARMA(0,0)+GARCH(1,0)	Skewed t	2134.06	-13.328	-8.195	-1.059
ARMA(0,0)+GARCH(1,1)	Skewed t	2098.596	-11.288	-1.025	13.248
ARMA(0,0)+GARCH(2,0)	Skewed t	2127.991	-11.316	-1.053	13.22
ARMA(1,0)+GARCH(1,0)	Skewed t	2131.499	-11.32	-1.056	13.217
ARMA(0,1)+GARCH(1,0)	Skewed t	2131.502	-11.32	-1.056	13.217
ARMA(1,1)+GARCH(1,0)	Skewed t	2131.498	-9.31	6.08	27.49

Log–Returns and Fitted Conditional Mean
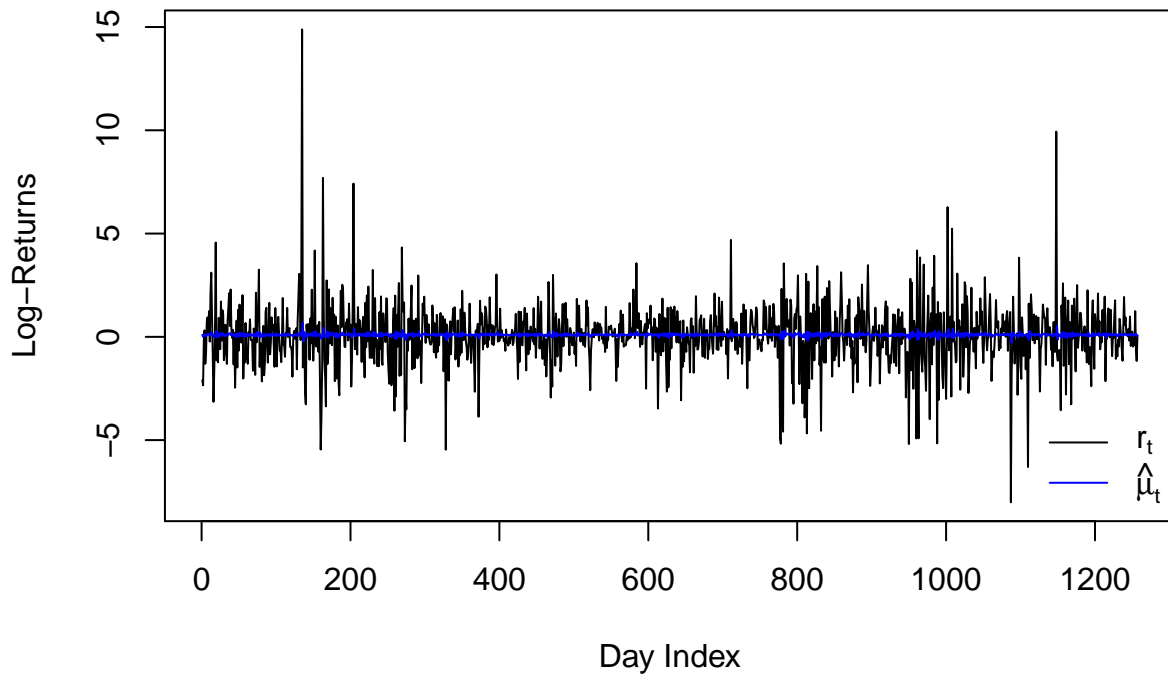
Forecasting

Assuming that historical returns on the portfolio are available, a fitted GARCH(r,s) model for the log-returns of a portfolio can naturally be used to forecast the conditional loss distributions of the model.

In particular, provided
i. Stationarity of the GARCH(r,s) model (for $\alpha_1 + \beta_1 < 1$), and
ii. The assumption that $\varepsilon_t \sim IID(0, \sigma^2)$,
The unconditional variance of a GARCH(r,s) model is given by

$$\theta = \frac{\alpha_0}{1 - \alpha_1 - \beta_1}$$

We may then write the one-step ahead forecast of the conditional variance as follows,

$$E[r_{t+1}^2 | \mathcal{F}_t] = \sigma_{t+1}^2 = \kappa\theta + (1 - \kappa)[(1 - \lambda)a_t^2 + \lambda\sigma_t^2]$$

For $\kappa, \lambda \in \mathbb{R}$.
This expression can be generalized to the h-step ahead forecast for conditional variance as follows,

$$E[\sigma_{t+h}^2 | \mathcal{F}_t] = \kappa\theta[1 + (1 - k) + \cdots + (1 - \kappa)^{n-2}] + (1 - \kappa)^{n-1}\sigma_{t+1}^2$$

Using
i. Maximum likelihood estimates for our parameters, $\kappa_{MLE}, \theta_{MLE}, \lambda_{MLE}$,
ii. The maximum likelihood estimate of unconditional variance as an estimate of σ_1^2, $\hat{\sigma}_1^2 = \theta_{MLE}$,
iii. The (historical) observed log-returns, $r_i, i = 1, \ldots, t$,
The h-step ahead forecast of the conditional variance (and conditional mean, if desired) of returns can be computed.[7]

It is worth noting that McNeil et al. show that the unconditional variance θ of a GARCH(r,s) model is a consistent estimator of the long-run conditional variance.[10]

$$E[\sigma_{t+h}^2 | \mathcal{F}_t] \xrightarrow{p} \theta \ \ as \ \ n \to \infty$$

These methods can be extended to estimate risk measures (quantiles) such as VaR (i.e., the conditional tail probability of extreme loss for a specified portfolio) by combining GARCH methods with extreme value theory (EVT).

Provided a fitted ARMA-GARCH model to a time series of log-returns of a portfolio, and estimated h-step ahead forecasts for conditional variance $E[\sigma_{t+h}^2 | \mathcal{F}_t]$ using the methods above, we may apply extreme valuate theory to fit the generalized Pareto distribution (GPD) to the tail of the fitted residuals to forecast the following risk measures,[7,10]
i. VaR: $\hat{VaR}_{\alpha,t} = \hat{\mu}_{t+1} + \hat{\sigma}_{t+1} q_\alpha(\varepsilon_{t+1}^{GPD})$
ii. Expected Shortfall: $E[\hat{S}]_{\alpha,t} = \hat{\mu}_{t+1} + \hat{\sigma}_{t+1} E[S]_\alpha(\varepsilon_{t+1}^{GPD})$

Although the GARCH(r,s) model assumes an $IID(0, \sigma^2)$ structure (e.g. $\overset{iid}{\sim} N$, $\overset{iid}{\sim} t$) on the residuals ε_t, McNeil et al. note that by the theory of quasi-maximum likelihood (QMLE), $\hat{\mu}_{t+1}$ and $\hat{\sigma}_{t+1}$ are still consistent estimators of μ_{t+1} and σ_{t+1} although the distributions of the residuals ε_t's may have been misspecified. In particular, they conclude that among a variety of numerical methods for estimation of VaR or expected shortfall, methods based on GARCH-EVT are most accurate.[10]

In order to apply these methods to our log-return model and construct a valid GARCH-EVT-VaR model, we must
i. Calculate the VaR time series using $\hat{VaR}_{\alpha,t} = \hat{\mu}_{t+1} + \hat{\sigma}_{t+1} q_\alpha(\varepsilon_{t+1}^{GPD})$, and
ii. Backtest the VaR estimates

To establish a statistical framework for VaR backtests, we will follow the recommendations outlined by the Federal Reserve Board of Washington D.C. in their 2005 review of backtesting procedures.

In particular, if we denote the loss on a portfolio over a time interval $T \in (t, t+1)$ as L_{t+1}, we define the indicator function of exceedances, which indicates whether a loss in excess of reported VaR has occurred, as[11]

$$\mathbb{I}_{t+1}(\alpha) = \begin{cases} 1 & if \ L_{t+1} \leq VaR(\alpha, T) \\ 0 & if \ L_{t+1} > VaR(\alpha, T) \end{cases}$$

To assess the accuracy of the VaR model, we will use the probability of failure (POF) test statistic popularized by Kupiec (1995):[12]

$$POF = 2\log\left(\left(\frac{1 - \hat{\alpha}}{1 - \alpha}\right)^{T - \mathbb{I}(\alpha)} \left(\frac{\hat{\alpha}}{\alpha}\right)^{\mathbb{I}(\alpha)}\right)$$

Where

i. $\hat{\alpha} = \frac{1}{T}\mathbb{I}(\alpha)$, and

ii. $\mathbb{I}(\alpha) = \sum_{t=1}^{T} \mathbb{I}_t(\alpha)$

Under the assumption that the VaR measures are accurate, the POF test statistic grows in value as the proportion of true VaR violations $\hat{\alpha}(100\%)$, differs from $\alpha(100\%)$. Thus the POF test statistic assesses whether the proposed VaR model systematically understates or overstates the underlying level of risk of the portfolio under the null assumptions of unconditional coverage of VaR estimate exceedances, and independence of the indicator sequence of exceedances, $\{\ldots, \mathbb{I}_{t-1}(\alpha), \mathbb{I}_t(\alpha)\}$.[11,12]

In order to forecast the one-step ahead VaR estimates, we will use our fitted GARCH(1,0) model in a rolling-window estimation procedure based on an out-of-sample window of one-tenth of our sample size (approximately 120 observations).

For the purposes of analyzing our model, we will use $\alpha_{VaR} = 0.01$ in constructing a VaR measure. Furthermore, because our returns our constructed from daily adjusted closing prices, we will assume a time horizon T of one day.

The Kupiec (1995) POF test indicates that the VaR model ($\alpha_{VaR} = 0.01$) accurately characterizes the underlying level of risk of the log-returns, demonstrating an appropriate number of exceedances and independence of the indicator sequence of exceedances ($\alpha = 0.05$).

POF Test (Kupiec, 1995) p-value	Result
0.05245	Correct Exceedances
0.0937	Correct Exceedances and Independent Indicators

Log−Returns and GARCH−EVT−VaR Model with VaR Forecasts and Simulated Bootstrap CI
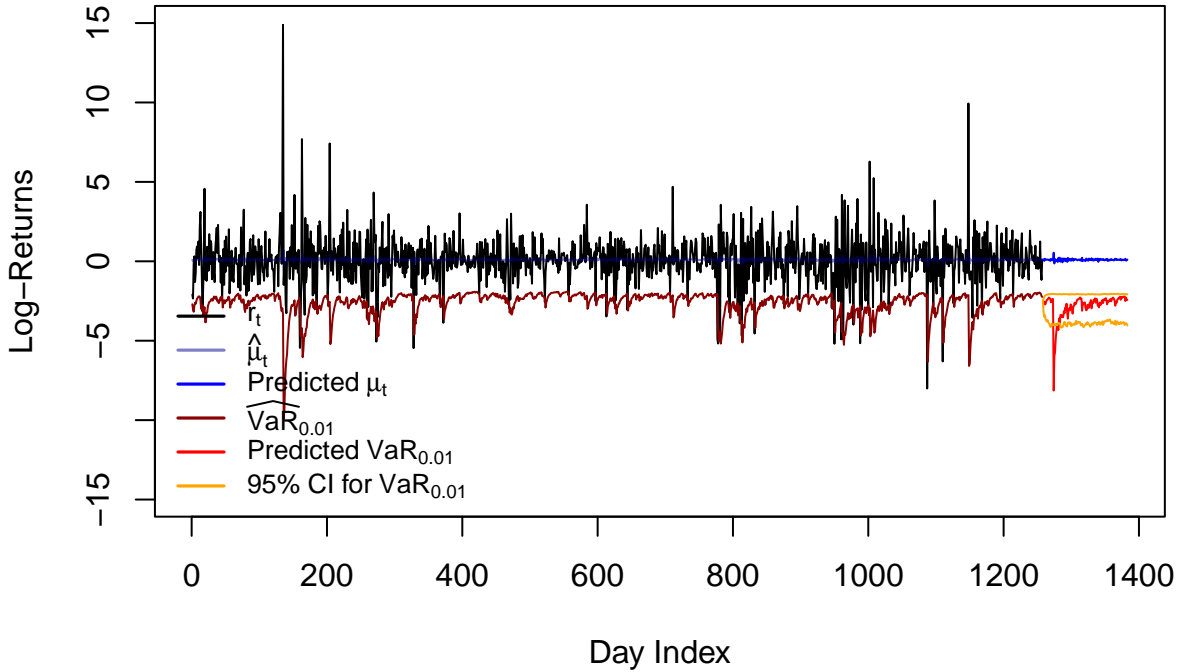
References

(1) Ruppert, David, and David S. Matteson. Statistics and Data Analysis for Financial Engineering. 2nd ed., Springer Texts in Statistics, 2015.

(2) El Barmi, Hammou. "Chapter 19: Risk Management." STAT GU4261 Statistical Methods in Finance.

(3) Kenton, Will. "Understanding Value at Risk (VAR) and How It's Computed." Investopedia, Investopedia, 23 Mar. 2023, https://www.investopedia.com/terms/v/var.asp.

(4) Harper, David R. "What Is Value at Risk (VAR) and How to Calculate It?" Investopedia, Investopedia, 29 Apr. 2023, https://www.investopedia.com/articles/04/092904.asp#:~:text=VAR%20is%20determined%20by%20three,and%20the%20Mon

(5) Zhang, Y., and S. Nadarajah. "A Review of Backtesting for Value at Risk." Communications in Statistics - Theory and Methods, vol. 47, no. 15, 2017, pp. 3616–3639., https://doi.org/10.1080/03610926.2017.1361984.

(6) Linssen, Thedo. "Value at Risk Forecasting." Erasmus University Rotterdam - Erasmus School of Economics, 2018.

(7) Haugh, Martin. "Risk Management and Time Series." IEOR E4602: Quantitative Risk Management. https://martin-haugh.github.io/teaching/qrm/.

(8) Braione, Manuela, and Nicolas Scholtes. "Forecasting Value-at-Risk under Different Distributional Assumptions." Econometrics, vol. 4, no. 4, 2016, p. 3., https://doi.org/10.3390/econometrics4010003.

(9) Ghani, I M, and H A Rahim. "Modeling and Forecasting of Volatility Using ARMA-GARCH: Case Study on Malaysia Natural Rubber Prices." IOP Conference Series: Materials Science and Engineering, vol. 548, no. 1, 2019, p. 012023., https://doi.org/10.1088/1757-899x/548/1/012023.

(10) McNeil, Alexander J., et al. Quantitative Risk Management Concepts, Techniques and Tools. Princeton University Press, 2015.

(11) Campbell, Sean D. "A Review of Backtesting and Backtesting Procedures." Finance and Economics Discussion Series, vol. 2005, no. 21, 2005, pp. 1–23., https://doi.org/10.17016/feds.2005.21.

(12) Kupiec, Paul H. "Techniques for Verifying the Accuracy of Risk Measurement Models." The Journal of Derivatives, vol. 3, no. 2, 1995, pp. 73–84., https://doi.org/10.3905/jod.1995.407942.