

PORTFOLIO CONSTRUCTION USING NON-GAUSSIAN VAR ESTIMATION AND GSEV APPROACH

A Project Report

submitted by

**UNNIKRISHNAN NAMBIAR K
(Roll No. 17NA3FP14)**

*in partial fulfilment of the requirements
for the award of the degree of*

**BACHELOR OF TECHNOLOGY (HONOURS) AND
MASTER OF TECHNOLOGY**



**VINOD GUPTA SCHOOL OF MANAGEMENT
INDIAN INSTITUTE OF TECHNOLOGY KHARAGPUR**

APRIL 2021

DECLARATION

I hereby declare that this report represents my ideas in my own words and where others' ideas or words have been included, I have adequately cited and referenced the original sources. I also declare that I have adhered to all principles of academic honesty and integrity and have not misrepresented or fabricated or falsified any idea/ data/ fact/ source in my submission. I understand that any violation of the above will be cause for disciplinary action by the Institute and can also evoke penal action from the sources which have thus not been properly cited or from whom proper permission has not been taken when needed.

Unnikrishnan Nambiar K

Roll No. 17NA3FP14

Department of Ocean Engg. And Naval Arch.

Indian Institute of Technology Kharagpur

Date: 21 April 2021

CERTIFICATE OF APPROVAL

This is to certify that the project report titled “**PORTFOLIO CONSTRUCTION USING NON-GAUSSIAN VAR ESTIMATION AND GSEV APPROACH**” submitted by **UNNIKRISHNAN NAMBIAR K** to Indian Institute of Technology Kharagpur for the award of the degree of **Bachelor of Technology (Honours) and Master of Technology** is a bona fide record of research work carried out by him under my supervision. The contents of this report, in full or in parts, have not been submitted to any other Institute or University for the award of any degree or diploma.

Dr. Sujoy Bhattacharya

Associate Professor

Vinod Gupta School of Management

Indian Institute of Technology Kharagpur

West Bengal – 721302

Place: Kharagpur

Date: 21-04-2021

ACKNOWLEDGEMENT

ABSTRACT

In this implementation, we have made use of a multivariate analysis of equity portfolio returns that accounts for some of the most studied and discussed traits of financial data: fat tails, positive and negative skew, stochastic clustering of volatility, and asymmetric dependence with the aim of improving the model accuracy of quantitative portfolio risk measures. We have implemented the EGARCH heteroskedasticity model to account for volatility clustering and leverage effect, the skewed-t distribution to capture fat tails and asymmetric skew, and the skewed-t copula implemented in python to model asymmetric dependence. This model, known as the GSEV approach was further implemented on BRIC equity index data. We have further back tested this model, validating portfolios with lower downside risk than alternative approaches, namely the Markowitz framework. We further cement that investors can substantially curb the underestimation of risk by taking non-Gaussian models into account.

Keywords: GSEV, Stochastic volatility, Tail-risk, Skewed-t copula, Portfolio optimisation, BRIC Economies, Volatility Clustering, EGARCH processes.

CHAPTER 1

INTRODUCTION

Since the advent of the Markowitz framework, experts have debated and shown that financial data stray away from the idealised normal distribution. True financial asset return distributions have fat tails compared to the prescribed normal distribution (leptokurtosis) and more extreme negative values than positive values (negative skew) [Mandelbrot, 1963].

Their volatilities tend to cluster (heteroskedasticity) and are not constants as the primitive models suggest. Asset returns have asymmetric tail dependency, with correlations rising during periods of market stress [Karolyi and Stulz, 1996; Longin and Solnik, 2001; Alcock and Satchell, 2018]. This was pertinent in understanding the 2008 financial crisis where the tail risk of CDO portfolios were significantly underestimated [Watts, 2016]. Thus, the mean–variance model, according to experts, can lead to risk mismanagement and be harmful to investor health. Despite this, the mean–variance approach remains very much in use in the financial sector [Fabozzi, 2007] due to its simplicity.

Chen and Fan (2006) formalised a set of models in which a multivariate time series' conditional mean and conditional variance are parametrically defined, but the multivariate distribution of the standardised invention is specified semi-parametrically as a parametric copula evaluated at non-parametric marginal densities. The family of such models were introduced as “semiparametric copula-based multivariate dynamic models (SCOMYD)”.

The students-t copula was used in this implementation to model asymmetric dependency patterns in high dimensions. We expand on and extend previous work by proposing a risk management and portfolio development approach that takes into account all the four characteristic traits mentioned above and can be used even when there exist a large number of assets.

The idea of Value at Risk (VaR) is used to explain the above risk management concepts. By applying our proposed approach on BRIC equity index data, we compare it to existing frameworks. We show that, when compared to competing Markowitz model, the proposed method generates a substantial increase in investor utility certainty, with important consequences for market practice. This report proceeds as follows. We include a brief overview of the VaR estimation literature in Chapter 2 as well as a literature overview of alternative non-Gaussian approaches that have been used for portfolio construction. In Chapter 3, we develop a scalable model (GSEV) that accounts for the four stylized facts listed earlier, as well as discuss the dataset used in the implementation. In Chapter 4, we examine the empirical results of our evaluation methodologies, benchmark methods, and data, as well as the models' VaR forecasting efficiency. Chapter 5 wraps up the study and addresses the findings' practical consequences.

CHAPTER 2

LITERATURE SURVEY

VaR estimation methods have been heavily criticised, despite their widespread usage in the financial sector. VaR has been argued to be incoherent as a risk indicator [Artzner, 1997]. Furthermore, VaR ignores the fact that, depending on market conditions, the same dollar loss will result in dramatically different economic outcomes [At-Sahalia, 2000].

Models that account for stochastic models of uncertainty, on the other hand, consistently outperform static models [Berkowitz and O'Brien, 2002; McAleer and Da Veiga, 2008; Skoglund et al., 2010; Moreira and Muir, 2017]. Models that incorporate non-Gaussian developments have also been shown to do better than pure parametric Gaussian models [McAleer and Da Veiga, 2008]. According to surveys, the most used approaches at commercial banks are relatively simple historical simulation and its alternative, the filtered historical simulation approach [Christoffersen, 2009; Pérignon and Smith, 2008; Gurrola-Perez and Murphy, 2015].

Several studies have attempted to improve investor utility by incorporating non-Gaussian characteristics into portfolio construction methods. According to Patton (2004), using skewed-t marginals and a rotated Gumbel copula to account for asymmetry and skewness results in a slight increase in realised usefulness over the usual 1/N strategy. For investors with moderate to high levels of risk aversion, Jondeau and Rockinger (2005) combined a Taylor series approach with DCC-GARCH to find that the approach substantially outperforms mean-variance Markowitz portfolios.

It has been shown that full-scale optimisation based on S-shaped and bi-linear utility improves investor utility [Adler and Kritzman, 2007]. With the implementation of CRRA utility to model portfolio weights as a function of asset characteristics joint asset returns must be modelled [Aït-Sahalia and Brandt, 2001]. To account for asymmetric dependency, Gaussian marginals in conjunction with the Clayton copula were used and a large increase in average returns as well as stronger returns in bearish markets were found [Alcock and Hatherley, 2009].

Models for all four stylised facts had previously been developed, but they were focused on generalising the Archimedean copula method to higher dimensions ($n > 2$), which necessitated substantial and restrictive assumptions. [Viebig and Poddig, 2010]. Another interesting model was the Taylor Series approach to account for higher moments, which showed that unless sophisticated shrinkage estimators of higher moments were used, the system did not increase investor welfare [Martellini and Ziemann, 2010]. Furthermore, during studies based on the 2008 financial crisis, the implementation of the truncated Lévy-flight distribution to catch skew and heavy-tails found that the strategy outperformed mean-variance slightly [Xiong and Idzorek, 2011].

CHAPTER 3

METHODOLOGY

3.1 THE GSEV APPROACH

In this section, we demonstrate the advantages of a risk management approach that can accommodate fat tails, bidirectional skew, stochastic volatility clustering, and asymmetric tail dependence. This model builds on previous work that employs GARCH models and copulas to model the same [Nystrom and Skoglund, 2002; Ghorbel and Trabelsi, 2009; Viebig and Poddig, 2010]. Copulas are useful because they allow you to distinguish the dependency structure from the univariate densities when building portfolios. The estimation of constant and dynamic conditional correlation (CCC and DCC) multivariate GARCH models [Bollerslev, 1990; Engle, 2002] is another way to model this.

For addressing stochastic volatility and the leverage effect, we use a univariate exponential GARCH (EGARCH) procedure, the non-centred student's t distribution to model the heavy-tails in the GARCH residuals, and the skewed-t copula to allow for asymmetric tail dependence in our approach. This is represented as the “GSEV” approach. The fitting of the GARCH processes, the estimation of the univariate marginals, and the modelling of dependency structure are the three parts of the estimation problem. Filtering with a GARCH method produces an iid series, which makes fitting a parametric distribution easier. The estimation process is also simplified and accelerated with minimal efficiency loss.

With Student-t innovations, we use Nelson (1991)'s exponentially weighted GARCH model. We suit piecewise distributions to the univariate GARCH residuals, as McNeil and Frey (2000) and Nystrom and Skoglund (2002) did. We used maximum likelihood estimation to suit non centered-t distributions and a Gaussian kernel. The skewed-t copula is used to model the dependency structure of the GARCH residuals. We are interested in the skewed-t copula since it can handle tail dependence, asymmetric dependence between upper and lower tails, and heterogeneous dependence throughout asset pairs. It is also scalable in large dimensions. The following measures are used to approximate and simulate the model:

1. Estimate an AR (1)-EGARCH (1,1) model for each asset i , and obtain parameter estimates, conditional variances, and residuals.
2. Calculate the standardised residuals, $\epsilon_{i,t} = \varepsilon_{i,t} / \sigma_{i,t}$.
3. Fit the univariate distributions to, $\epsilon_{i,t}$.
4. Fit univariate skewed-t-distributions to each vector of residuals from step 2 by MLE.
5. Estimate the covariance matrix Σ^\wedge given by:

$$\left(\frac{v-2}{v} \right) \left(\text{cov}(X) - \frac{\hat{\beta} \hat{\beta}' (2v^2)}{(v-2)^2 (v-4)} \right)$$

where v and β are parameters of the skewed-t distributions estimated in Step 4. $\text{Cov}(X)$ is the sample covariance matrix.

6. Draw N independent d-dimensional vectors from the multivariate Gaussian distribution defined by:

$$Z \sim N(0, \Sigma^{\wedge})$$

7. Draw N independent random numbers from the inverse gamma distribution defined by:

$$W \sim IG\left(\frac{v}{2}, \frac{v}{2}\right)$$

8. Substitute Z and W into the following function to yield multivariate skewed t variables:

$$X = \mu + \beta W + \sqrt{W}Z$$

9. Approximate the cumulative distribution functions for the univariate marginals by numerical integration.
10. Transform X , based on Step 8, into uniformly distributed variables using the univariate cumulative distribution functions.
11. Convert Usk to innovations, Isk , using the inverse cumulative distribution functions for the univariate piecewise distributions developed in Step 3.
12. Simulate from an AR (1)-EGARCH (1,1) model for each asset i using innovations, I , and parameters $\theta_i E$.

This algorithm is rapid and can be applied to problems with large dimensions.

3.2 DATA

Using a dataset of major stock index prices from BRIC economies, we compared our proposed model to portfolio construction approaches typically used by managers or that are well-known in the literature. Our data set encompass the period of 2014-2017 and is comprised of Brazilian (BVSP), Russian (MOEX), Indian (NIFTY), and Chinese (SSE) Indices. All our return series are daily.

The main aim behind considering the BRIC nations for this study is to analyse the characteristics of volatility clustering for the emerging markets. In the last few years, Brazil, Russia, China, India, and south Africa have shown consistent return thus including those countries for this study is relevant and suitable. Many institutional and private investors consider emerging countries as smart alternative investment avenues.

CHAPTER 4

RESULTS

The average overview statistics for the properties in the dataset are shown in Table 1. To search for normality, we used the Jarque–Bera tests on each index. For each asset, normality was rejected. We have also conducted the Augmented Dickey Fuller test for the indices and the results show that there was strong evidence to believe in the absence of unit roots. These are shown in Table 2. The absolute value of returns has pronounced autocorrelation, which is consistent with volatility clustering. The ACF and PACF plots are shown in Fig1.

Table 1. Dataset Summary Description

INDEX	BRAZIL (BVSP)	RUSSIA (MOEX)	INDIA (NIFTY)	CHINA (SSE)
Annualized Return	0.009	0.017	0.009	0.005
std	0.018	0.019	0.012	0.015
min	-0.148	-0.091	-0.130	-0.103
25%	-0.008	-0.009	-0.005	-0.006
50%	0.001	0.000	0.001	0.001
75%	0.010	0.012	0.007	0.007
max	0.139	0.133	0.088	0.106
Skewness	-0.50	0.26	-0.92	-0.63
Kurtosis	11.68	3.12	14.64	7.60
Jarque Bera	9912.37	719.42	15701.69	4275.23
P (Normal)	0.00	0.00	0.00	0.00

Table 2. Augmented Dickey Fuller Test

INDEX	BRAZIL (BVSP)	RUSSIA (MOEX)	INDIA (NIFTY)	CHINA (SSE)
ADF Statistic	-10.01	-16.14	-8.76	-13.73
p Value	0.00	0.00	0.00	0.00
AIC Auto lag	17.00	5.00	23.00	7.00
Critical 1%	-3.43	-3.43	-3.43	-3.43
Critical 5%	-2.86	-2.86	-2.86	-2.86
Critical 10%	-2.57	-2.57	-2.57	-2.57

Table 3. VaR forecast evaluation: at 99.5% confidence.

$\beta = 99.5\%$	EW Markowitz	GSEV
Violations	0.0431	0.0048
UC: p-value	0.00	0.90
Consecutive viol	0.19	0.06
SI: p-value	0.00	0.03
CC: p-value	0.00	0.12

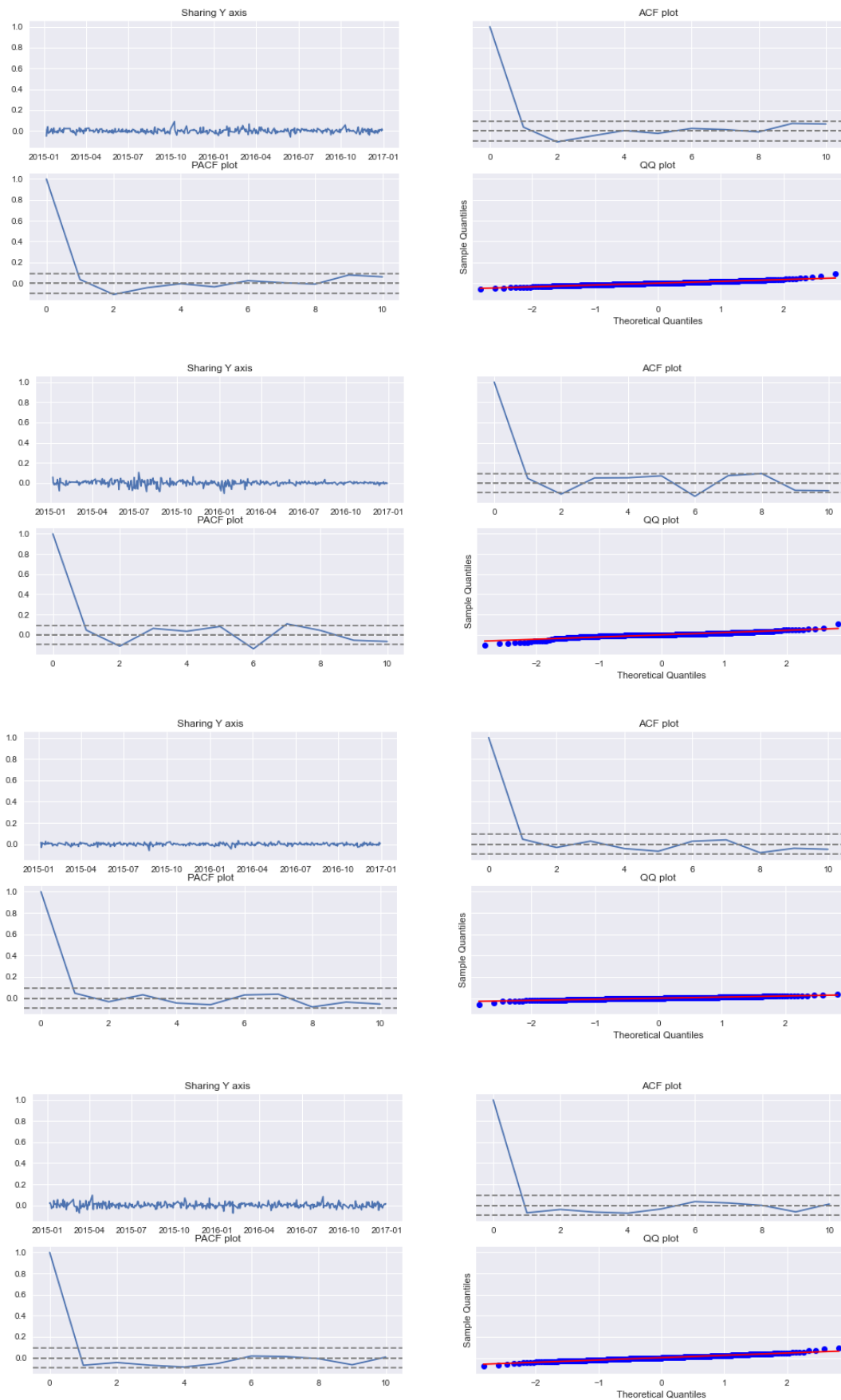


Fig.1. ACF, PACF plots BVSP, SSE, NIFTY, MOEX

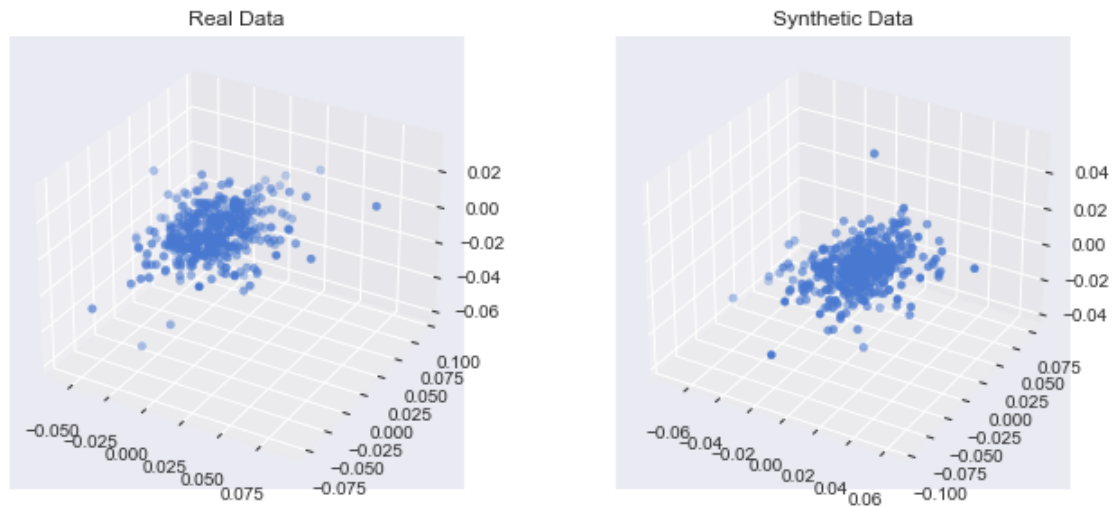


Fig.2. 3D comparison plot of real data and synthetic data from GSEV based on BSVP, MOEX, NIFTY data



Fig.3. Back test for EW Model based on Markowitz portfolio.



Fig.4. Back test for Minimum Variance Model based on GSEV portfolio.

It is apparent from Table 2 that there exists significant autocorrelation and the absence of unit roots. Fig 2 shows the synthetic data generated from the GSEV approach and the real data in a 3D plot. Hence, we can utilise this to predict better using the GSEV algorithm. The portfolio performance of the models is summarised in Figures 3 and 4. The out-of-sample CVaRs of the Markowitz EW model are significantly higher than the target. The GSEV approach, on the other hand, produces CVaRs that are similar to the target level producing the smallest risk forecast error. The model also consistently produces a lower maximum drawdown, as measured by the peak to trough return, a widely used metric by practitioners. Since the copula imposes a consistent dependency structure across asset pairs, the approach's output is likely to deteriorate as the number of assets and complexity grows. Table 3 shows the relative success of the two models.

CHAPTER 5

DISCUSSION

This implementation aimed to strengthen Gaussian approaches that systematically underestimate tail risk, in line with previous work in the field. Any strategy that quantifies tail risk using the Gaussian distribution poses a major risk to investor health. Conditional volatility approaches tend to be an important component of a sensible VaR estimation model since they reduce the proportion of consecutive breaches. The GSEV approach produces more reliable VaR forecasts than the GGEV approach, which employs a symmetrical dependency structure, suggesting the significance of accounting for asymmetric dependence. We believe that these findings are applicable to other risk control procedures, but we'll leave that to further study.

We conclude that the GSEV method outperforms mean–variance in our out-of-sample portfolio rebalancing study. In comparison to the mean–variance method, historical simulation produces a small increase in economic value. The added benefit of the GGEV and GSEV approaches is important. The use of a conditional volatility approach in combination with a heavy-tailed distribution is a popular unifying factor across both of these approaches.

The GSEV approach has the potential to provide significant value addition while minimising risk. The GSEV method also has the lowest maximum drawdown (MDD) values and the most reliable portfolio VaR estimates out of study. This demonstrates how important it is to account for asymmetric dependency. This study has a number of consequences for fund managers. Managers can use a conditional volatility model to quantify fat tail risk and optimise investment portfolios if asset return volatility is stochastic and clustered. Exponentially weighted covariance estimators, GARCH, stochastic volatility models, and realised volatility using HF data are some of the options suggested. In the case of HF results, noise from market microstructures must be taken into account. [Ait-Sahalia, 2010].

Investors can use a heavy-tailed distribution in the form of skewed-t distributions or other models if conditional returns are non-Gaussian. Finally, when calculating VaR and evaluating portfolio weights, investors should consider asymmetric dependency. The skewed-t copula seems to be a solid and underplayed choice. The dispersion of our multivariate approach's risk estimates is a potential field for future research. This method is time-consuming to implement because it necessitates the calculation of a large number of parameters. In this context, it would be fascinating to put Ait-Sahalia and Brandt's more frugal portfolio selection strategy into practise (2001).