Enhancing Extreme Risk Assessment in Green Bonds:

A Monte Carlo Simulation Approach

Master's Thesis

Masoud Morvarid

June 2024

1st supervisor: Prof. Dr. Kees Koedijk

2nd supervisor: Dr. Labrini Zarpala

Utrecht School of Economics

Banking and Finance

Universiteit Utrecht

Contents

1- Introduction

Green bonds (GBs) are an essential tool to raise capital and invest in projects that generate benefits for the environment. Typically, these bonds finance projects such as renewable energy, green buildings, clean transportation, or sustainable land use, among others. Since the first green bond was issued by the World Bank in 2007 and the European Investment Bank in 2008, there has been substantial growth in the green bond market, with issuance worth $939 billion in 2023, up 3% on the same period last year. It's not a record – that was in 2021 when issuance reached $1.1 trillion. The green bond market has grown significantly with over 50 countries issuing green bonds. However, questions are being raised regarding the risk of investing in GBs compared to conventional bonds. As the urgency to combat climate change grows, effective risk management strategies are crucial, especially in green markets. Nevertheless, given the series of extreme events that have occurred in the global financial market since the early 1980s, such as the Latin American Debt Crisis, the Dot Com Bubble, and the recent COVID-19 pandemic-induced market crash. Consequently, assessing risk in extreme conditions would be crucial in this case. While current methodologies offer an understanding of extreme risk modeling for GBs, there is a gap in using Monte Carlo simulations to develop a risk assessment tool to efficiently capture extreme behavior without compromising the accuracy of the model. This thesis seeks to build a Monte Carlo framework focused on the tail behavior of GB returns. By applying a semi-parametric methodology to model the distribution tails and resulting detailed tails not just normal distributions. This would enhance the simulation's ability to model extreme risks accurately.

The research considers both the traditional and the modified Geometric Brownian Motion model to simulate extreme events. At first, I will consider customizing the Monte Carlo simulation for extreme events, incorporating a non-Gaussian distribution such as the student t-distribution with heavier tails and incorporating time-varying volatility. Moreover, I study the effect of Montecarlo with a particle filtering approach. The thesis's findings will have advanced the role of Monte Carlo Simulation in modeling extreme risk, providing better simulation to capture extreme events like COVID-19. By employing various modeling techniques and indices, the study provides insights into the extreme risk landscape in green markets and offers valuable implications for policymakers, and investors. Monte Carlo with t-distribution and volatility assumption by GARCH in both indexes led to a robust estimation of VaR and ES and can beat other methods in this study.

2- Literature review

Green bonds (GBs) have become increasingly popular in bond markets as they are utilized by authorities for environmentally friendly projects. GBs play a crucial role in mitigating environmental risks by offering policy instruments that incentivize sustainable investments and activities. These bonds, along with futures markets, provide mechanisms to share risks, create incentives for private enforcement of environmental contracts, and channel capital towards green projects, ultimately reducing pollution and promoting sustainability(Smith, 2022). Despite the advantages of GBs in promoting green investments, concerns have been raised regarding the risks associated with investing in them compared to conventional bonds. Research has focused on analyzing volatility dynamics, spillover effects on conventional bonds, and factors affecting the relationship between green and conventional bonds. Studies indicate that while conventional bonds may offer higher returns on average, the differences can be explained by bond properties. GBs have shown positive impacts on various aspects like announcement returns, environmental and operating performance, and attracting long-term investors. While GBs demonstrate superior financial convenience, issues related to risk assessment through measures like Value-at-Risk (VaR) need to be explored further.(Tsoukala & Tsiotas, 2021)

Value at risk (VaR) is a statistical measure that outlines the potential financial loss a company, portfolio, or position could face within a certain timeframe. It was first introduced by JP Morgan in 1994. For example, if a portfolio has a daily VaR of $1 million at the 99% confidence level, it means there's only a 1% chance that the portfolio's losses will exceed $1 million on any given day under normal market conditions. VaR has become a fundamental component of risk management, offering a straightforward way to estimate potential losses over a specified period at a given confidence level. Calculating VaR, despite its apparent simplicity, is quite complex. There are three primary approaches for calculating:

1. Parametric: It assumes that asset returns follow a roughly normal distribution.

2. Non-parametric: This approach doesn't rely on parameter assumptions and includes historical simulation, which uses past data to forecast potential future losses.

3. Semi-parametric: This approach blends elements of both non-parametric and parametric methods. Notable semi-parametric methods include the Filtered Historical Simulation (Barone - Adesi et al., 1999), the CaViaR method by (Engle & Manganelli, 2004), and methods based on the Extreme Value Theory.

Empirical studies, such as those referenced by (Abad et al., 2014), suggest that methods based on the Extreme Value Theory and the Filtered Historical Simulation are particularly effective for forecasting VaR. The Block Maxima and Peaks-Over-Threshold methods are two prominent EVT-based modeling approaches, employing the GEV and GPD limit distributions, respectively(Gumbel, 1958). The Block Maxima method involves segmenting data into blocks and using the maximum loss from each to fit a Generalized Extreme Value (GEV) distribution. This approach is validated by the Fisher–Tippett–Gnedenko theorem, which states that GEVD is the only potential asymptotic distribution for block maxima extreme values. (McNeil, 1998), presenting it as a three-parameter model.

$$H_{\xi,\mu,\psi}(m) = \begin{cases} \exp\left(-\left(1+\xi\frac{m-\mu}{\psi}\right)^{-1/\xi}\right) & \text{if } \xi \neq 0 \text{ and } 1+\xi\frac{m-\mu}{\psi} > 0, \\ \exp\left(-\exp\left(-\frac{m-\mu}{\psi}\right)\right) & \text{if } \xi = 0. \end{cases}$$

On the other hand, the Peaks-Over-Threshold (POT) method selects data points exceeding a predefined threshold, fitting them to a Generalized Pareto Distribution (GPD). Initially proposed by (Balkema & De Haan, 1974) and later refined by in 1975, this method focuses solely on the tail of the distribution, potentially offering a more tailored analysis. However, the choice of the threshold is crucial, as it significantly influences the parameter estimates of GPD and subsequently the final risk estimates, particularly in smaller samples. This decision represents a trade-off between bias and the variance of the estimators. The density function of GDP is:

$$G_{\xi,u,\psi}(x) = \begin{cases} 1-\left(1+\frac{\xi}{\psi}(x-u)\right)^{-1/\xi}, & \text{if } \xi \neq 0, \\ 1-\exp\left(-\frac{(x-u)}{\psi}\right), & \text{if } \xi = 0. \end{cases}$$

(McNeil & Frey, 2000) introduced a novel approach that integrates EVT with volatility models, termed conditional EVT, further expanding the toolkit for risk assessment in financial modeling. Except EVT, the Parametric approach can be refined like the VaR-x methodology presented by (Huisman et al., 1998) offers an enhanced approach to estimating VaR by accounting for the fat tails in asset return distributions. This method employs the Student's t-distribution, which has a flexible shape parameter (degrees of freedom) that can adjust for the tail thickness of the distribution. The tail index, derived from extreme value theory, is used to set the degrees of freedom, thereby directly incorporating the empirical tail behavior of asset returns into the VaR estimation.

In a recent study on this subject,(Zhuo et al., 2023) designed an extreme risk model for green bonds (GB) and the clean energy sector. They examined the extreme risk associated with GB and Global Clean Energy (GCE) indexes from 2012 to 2020, employing Value at Risk (VaR) and Expected Shortfall (ES) as risk assessment indicators. The researchers trained their models on data from periods with lower extreme risk events to validate their ability to predict potential extreme risks, notably during the COVID-19 episode. They excluded the ongoing Russia-Ukraine war from their analysis due to its unpredictable duration and impact, which complicates market recovery analysis post-extreme events. During the COVID-19 pandemic, the GCE index showed significant volatility, with VaR and ES indices varying between 2 to 10%. Conversely, the GB index remained relatively stable, with VaR and ES figures between 1-3%. To address this gap, they developed traditional rolling time window models, GARCH models with different parameter estimates, and semiparametric models for GCE and GB to measure extreme risk. The GAS, GARCH-FZ, and hybrid models leverage (Patton et al., 2019) statistical decision theory, incorporating semiparametric models for dynamic ES and VaR to improve the accuracy and robustness of their predictions. Semi-parametric models were employed to enhance the precision and stability of VaR and ES estimates for GBs and the GCE index. Identifying breakpoints is also deemed crucial for accurately understanding and modeling financial time series, though it adds complexity to the model creation process. They emphasized the importance of employing advanced statistical theory and Brownian motion to improve the prediction of extreme risks.

2-1- Monte Carlo

The simplest Monte Carlo (MC) method for estimating VaR over a one-day horizon at a specific significance level involves simulating N draws from the distribution of returns for the next day. The 99% VaR is determined by identifying the N/100th element after sorting the simulated one-day returns. MC is an algorithmic technique in finance that leverages randomness to forecast a multitude of potential outcomes in complex, uncertain systems. It generates random values from a probability distribution and performs calculations on these values to estimate outcomes that might be challenging or impossible to determine through conventional algebraic or numerical methods.

In finance, MC simulation is crucial for modeling asset prices, especially under the assumption that daily asset returns follow a normal distribution. By creating a probability distribution based on historical data for an asset's expected return and volatility, MC can produce numerous potential future price paths. While a single simulation provides limited information, aggregating multiple simulations offers a robust model for potential stock price movements, reflecting the widely accepted random walk hypothesis in financial theory.(Shonkwiler, 2013)

(Jascha, 2015) employed MC simulation with an exponentially weighted moving average to enhance the modeling process. To simulate the unpredictable nature of stock prices, volatility is combined with a Wiener process, a continuous-time stochastic process derived from Brownian motion that resembles a random walk. Geometric Brownian Motion

(GBM) is a preferred model for asset price simulation, defined by a differential equation where the incremental change in asset price depends on a deterministic drift component and a random volatility component. Practically, simulating GBM requires iterative computation, integrating both the deterministic drift and the random shocks induced by volatility over various iterations and time frames. This process produces a distribution of potential future asset prices, assisting investors and analysts in evaluating the likelihood of different price outcomes.

While normal, and lognormal are commonly used in financial modeling, other distributions like Student's t, skewed Student's t, generalized t, and generalized error distribution (GED) are also applied. Despite the mathematical elegance of modeling financial prices or logarithms following Brownian motion, researchers (Fama, 1963) have noted challenges in fitting real financial data to these models, advocating for the use of non-Gaussian or skewed distributions to better model return distributions.

2-2- Improvements in Montecarlo

Despite some criticisms of the statistical methods underpinning these findings, the normal distribution continues to be of practical interest specially when we want to focus on mean of sample. (De Domenico et al., 2023) have proposed various alternative distributions to better align with observed data, including:

alternative distributions are proposed to better fit the empirical data:

Truncated Lévy Distribution**:** It is used to model asset price fluctuations, particularly effective in an intermediate range, and decays exponentially beyond it. This distribution can capture the heavy tails observed in financial data but includes an exponential cut-off to ensure a realistic decay in extreme events.

Student's t-Distribution**:** This distribution accommodates heavy tails and has a shape parameter (degrees of freedom) that influences its tail behavior. In the context of high-frequency returns, it can replicate the observed power-law behavior of extreme movements.

q-Gaussian Distribution: It is a generalization of the normal distribution and is an important concept in the field of non-extensive statistical mechanics, a theory developed by Constantino Tsallis and useful for derivative pricing in fat tail conditions.

Modified Weibull Distribution**:** This is proposed to describe the behavior of empirical log returns, especially in the tails. It features a stretched exponential behavior for large variations, offering a different approach to capturing tail risk in financial data.

The Student's t-distribution is a popular model for asset returns due to its flexibility over the normal distribution, particularly in capturing heavier tails. Its degrees of freedom parameter, when reduced, results in fatter tails, thus offering a closer fit to the empirical distribution of asset returns. In this thesis I want to focus on this alternative and discuss it more on methodology section.

While GBM assumes a continuous path for price movements with fluctuations modeled through a normal distribution of returns, it fails to account for the abrupt and sizable price changes seen in real markets. Recognizing these limitations, (Merton, 1976) introduced the model, an extension of the GBM model that incorporates jumps to reflect sudden significant movements in stock prices. The Merton Jump Diffusion (JD) model introduces a jump component to the traditional GBM framework. This enhancement allows the model to incorporate sudden and significant changes in stock prices, reflecting more realistic market behavior. The JD model modifies the GBM equation by adding a jump term and applying weights to the mean and standard deviation offers a more precise prediction model for short-term periods (1 month) compared to using non-weighted parameters. The weighted method in this study is not an accurate method but sequential MC is new method that can be useful in this case it makes the model more complicated. The revised version of the Sequential Monte Carlo (SMC) framework, as developed by (Neslihanoglu & Date, 2019), is a groundbreaking enhancement in the efficiency of estimating extreme quantiles, especially relevant for dynamic systems and time series analysis. The effectiveness of this updated SMC method is proven in its application to the Cox–Ingersoll–Ross (CIR) model, showcasing its potential in financial risk management by adeptly identifying extreme quantiles of interest rates. Unlike conventional MC methods that rely on fixed distributions, the modified SMC adapts its estimations to meet specific requirements, particularly focusing on tail quantiles.

3- Methodology

Monte Carlo has a similarity to historical simulation. The main difference is that the simulation uses instead of observed changes in the market factors.

The data of GBs and GCE will be downloaded from FactSet as time series data for the study period. I will calculate daily log returns using these daily indexes. Then use the first n-day daily return to calculate the mean standard deviation as normal parameters for the first attempt. After that, I use them to generate 10000 synthetic returns scenarios for the next day and compare 1-day VaR with 95, 99, and 99.9% confidence intervals with actual data

I can use the latest n-day means and standard deviation for the next day repeat it for the coming days and compare them together.

The distinctive method that I want to test in this thesis is how to relax normal simplification in stochastic terms and use non-gaussian distribution for generating synthetic returns. MC is a time-consuming approach but it is straightforward and needs to be modified in time of rising extreme events.

3-1- Simulating with normal GBM

B_t is Brownian motion (a Wiener process) , dB_t are increments of a Wiener process that is a random term in the equation and never becomes negative. The price of an asset moving with a drift and a shock. The is a random amount of volatility that acts on the stock's price.

$$dS_t = \mu S_t d_t + \sigma_v S_t dB_t$$

$$dB_t \sim N(0, dt)$$

$$\ln S_t = \left[\mu - \frac{1}{2}\sigma^2\right] t + \sigma B_t$$

Firstly, I will model typical GBM under normal distribution assumption. The simulations depicted were generated by the following algorithm. Inputs are: starting price $S0$, period of the study T in years, volatility σ in per square root year, and drift μ in per year

Here is the algorithm:

inputs: S_0, T, μ, σ
$\Delta t = 1/365.0$ ▹ 1 day time increments in years
$n = T/\Delta t$ ▹ number of Δt steps in time T
for $t = 1, \dots, n$
$Z_t \sim N(0,1)$
$\Delta S_t = S_{t-1}\left(\mu\Delta t + \sigma\sqrt{\Delta t}Z_t\right)$
$S_t = S_{t-1} + \Delta S_t;$
endfor
▹ the last S_t is S_T

Before introducing my suggestions for enhancement of MC, I should declare that because there is no mean reverting term in GBM and each iteration is independent of the former one, in the long run, it can generate extreme synthetic situation however in reality returns of assets is a time series that has an autocorrelation that can revert back the values to its mean but GBM has not such term. We need this modification for short-term analysis.

3-2- Simulating GBM with fat-tail distribution

For studying with fat tail distribution it has a single parameter known as degree of freedom(DOF), v in the following equation. The t probability density function is:

$$f_\theta(y_t) = \frac{1}{\sqrt{\pi(v-2)}} \frac{\Gamma\left(\frac{v+1}{2}\right)}{\Gamma\left(\frac{v}{2}\right)} e^{-\frac{y_t}{2}} \left(1 + \frac{y_t^2 e^{-x_t}}{v-2}\right)^{-\frac{v+1}{2}}$$

$$\Gamma(z) = \int_0^\infty t^{z-1} e^{-t} dt$$

As DOF tends to infinity the t distribution tends to the standard normal density. I use the (Bailey, 1994) method for sampling from the t-distribution, the following algorithm:

repeat
$\mathrm{U} \sim \mathrm{U}(0,1); \mathrm{U} = 2\mathrm{U} - 1;$ uniform on -1 to 1
$\mathrm{V} \sim \mathrm{U}(0,1); \mathrm{V} = 2\mathrm{V} - 1;$ a point in the sqr.
until $\mathrm{W} = \mathrm{U}^2 + \mathrm{V}^2 <= 1$
$\mathrm{C} = \mathrm{U}^2/\mathrm{W}; \mathrm{R} = \nu\left(W^{-2/\nu} - 1\right)$
$\mathrm{T} = \sqrt{RC}$
if ($\mathrm{U} \sim \mathrm{U}(0,1) < .5$)
return T
else
return $-\mathrm{T}$

3-3- Time-varying volatility

The second fundamental change we make to the Monte Carlo methodology is that volatility is not constant over time. Volatility has its volatility, and highly volatile periods cluster together. In other words, if we have a very volatile period in the market, such as covid outbreak period, the subsequent period will not automatically normalize it. As I intend to generate more extreme events using the simulation, I decided it was reasonable to incorporate it. There are numerous methodologies to model this characteristic, technically referred to as heteroskedasticity. The one used in this thesis is GARCH.

My idea is that the algorithm I mentioned above doesn't change, I just need to update volatility in each time

$$\sigma_t^2 = \alpha_0 + \alpha_1 \epsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2$$

This is the GARCH(1,1) model that can be used to forecast volatility in trained time I will define the coefficients and then after each iteration, I need to update volatility. The new method for Monte Carlo is proposed by the RiskMetrics group, and the paper about that is work presented this method. No paper verifies this method, so when relaxing the normal distribution in GBM, I work on time-varying volatility in GBM and check whether it has meaningful results or not. Time-varying volatility makes the SDE of GBM unsolvable with analytical methods but numerical methods still may be useful. The last one is the t-distribution and time-varying volatility to check the combination.

3-4- Particle filtering Algorithm

Particle filtering, also known as Sequential Monte Carlo (SMC) methods, is particularly useful for non-linear and non-Gaussian state-space models. Here, I outline the methodology used, drawing on the framework established by(Kitagawa & Sato, n.d.). The particle filtering algorithm approximates the filtering and smoothing density functions using a large number of particles, which can be considered independent realizations from the distributions. The process involves several steps:

Initialization:

Generate N initial particles $\left\{X_0^{(i)}\right\}_{i=1}^{N}$ from the initial state distribution $p(X_0)$.

Prediction:

For each particle $X_{t-1}^{(i)}$, generate a predicted state $X_t^{(i)}$ using the state transition model F_t and system noise V_t :

$$X_t^{(i)} = F_t\left(X_{t-1}^{(i)}, V_t^{(i)}\right)$$

This yields N predicted particles $\left\{X_t^{(i)}\right\}_{i=1}^{N}$.

Update (Weighting):

Compute the importance weight $w_t^{(i)}$ for each predicted particle based on the likelihood of the observed data Y_t :

$$w_t^{(i)} = p\left(Y_t \mid X_t^{(i)}\right)$$

Normalize the weights:

$$\widetilde{w}_t^{(i)} = \frac{w_t^{(i)}}{\sum_{j=1}^{N} w_t^{(j)}}$$

Resampling:

Resample N particles from the current set $\left\{X_t^{(i)}, \widetilde{w}_t^{(i)}\right\}_{i=1}^{N}$ according to their normalized weights. This step helps to focus computational resources on the most promising particles and discard those with low weights.

Iteration:

Repeat the prediction, update, and resampling steps for each time step t in the data series.

Significance test of the result

I will simulate it in a daily interval for a forecasting period but the number of iterations is a variable that can be started from example from 1000 to 100000

Convergence check: I want to ensure that adding more iterations does not significantly change the outcome. If results keep fluctuating with more iterations, it might indicate that the simulation has not yet converged and more iterations are needed.

Backtesting plays a crucial role in validating VaR and ES models. Various loss functions, such as Kupiec and Lopez’s loss function are utilized for this purpose. Kupiec’s POF test is an unconditional test method used for backtesting VaR models to determine their accuracy in predicting financial losses over time. This test assesses the VaR model's effectiveness by counting the number of violations and comparing them against predetermined confidence levels. It serves as a benchmark to evaluate the performance of a VaR model, particularly in detecting where the model under risk. Lopez's loss function is a density forecast test that focuses on the distance between observed returns and forecasted VaR values when a breach occurs. (Zhang & Nadarajah, 2018a).

4- Empirical study

4-1- Data collection

I selected the S&P Global Clean Energy and S&P Green Bond Index as our core data sources .The S&P Green Bond Index is designed to track the global green bond market. S&P GB index includes bonds that are labeled "green" by the Climate Bonds Initiative (CBI) and meet the specific criteria outlined in the Eligibility Criteria and Sub-Index Rules. The S&P Global Clean Energy Index is designed to measure the performance of companies in global clean energy-related businesses from both developed and emerging markets, with a target constituent count of 100 .

The study period: The dataset covers a period of 10 years, from 2012 to 2021. Daily return of S&P GCE and S&P GB from January 1, 2012, to December 31, 2021. The sample period is from 1 January 2012 to 31 December 2016 and the forecast time is from 1 January 2017 to 31 December 2021. This period includes the significant market event of the COVID-19 pandemic. The indices are represented in Figure 1, showing the time series data for the specified period. All data was collected from FactSet.

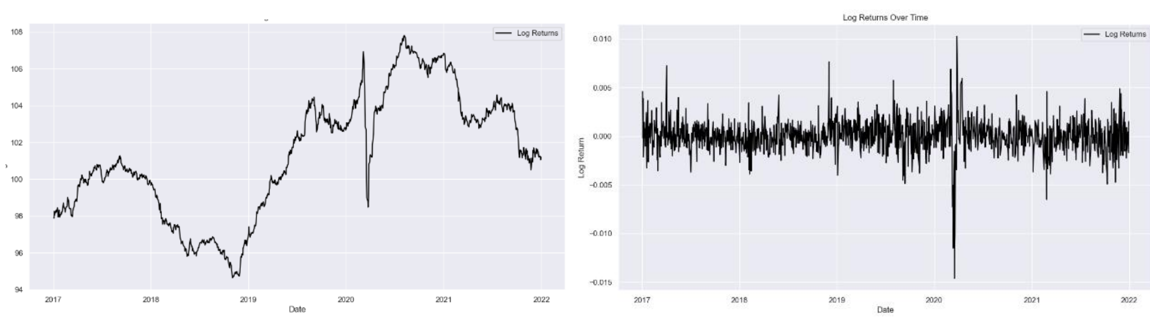

Figure1- Green Bond (GB) index times series from 2017 to 2022(left) – Log return time series(right)

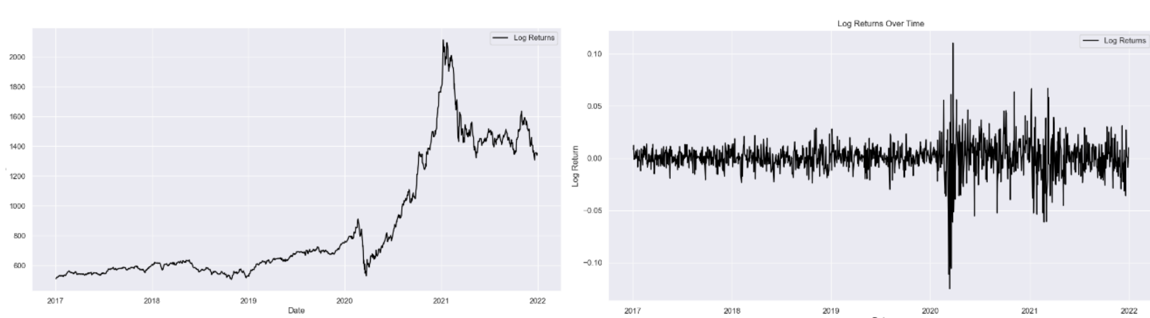

Figure2- The Global Clean Energy (GCE) index times series from 2017 to 2022(left) – Log return time series(right)

The summary statistics for each index are provided in Table 1. This includes the mean, standard deviation, t-distribution parameters (degrees of freedom, location, and scale), skewness, and kurtosis.

	GB	**GCE**
Mean	0.000	0.000
Standard Deviation	0.002	0.012
T-Distribution DF	3.0	6.5
T-Distribution Loc	0.000	0.000
T-Distribution Scale	0.001	0.010
Skewness	-0.514	-0.133
Kurtosis	8.142	1.025

Table1 – statistical summary for daily return distribution

It is necessary to test the non-normality hypothesis, which is more realistic for financial data compared to a normal distribution. Figures 2 and 3 show the leptokurtic shape of log-return distribution for the GB and GCE indices, respectively. These figures highlight the presence of fat tails in the data, which normal distributions fail to capture accurately.

Given the presence of leptokurtic distributions, I replaced the normal distribution with a t-distribution to better capture tail events. The t-distribution is a more realistic assumption for modeling the extreme values observed in financial markets. The red line in fig3,4 shows the normal distribution, If both sets of quantiles came from the same distribution, we should see the points forming a roughly straight line.

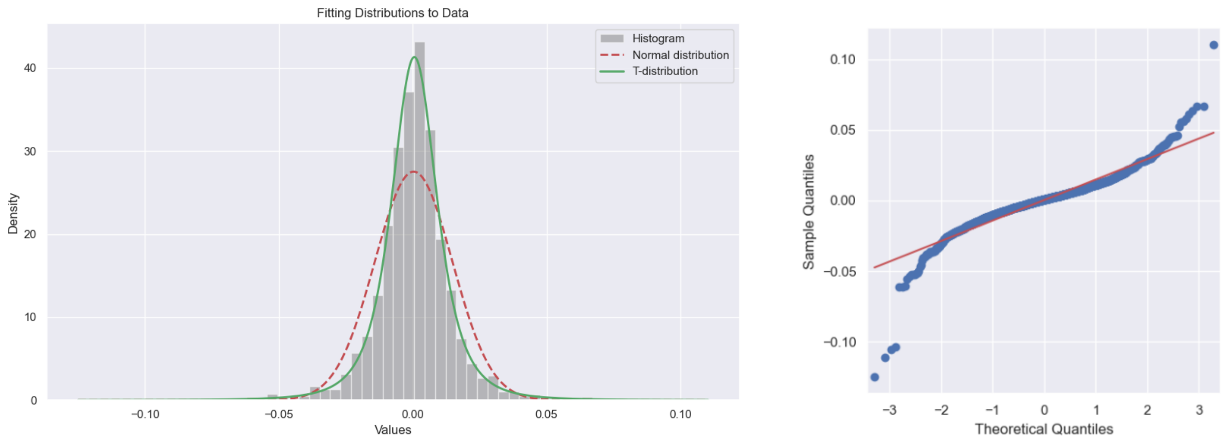

Figure3- Global Clean Energy index – log return distribution in-sample period(left) – QQ plot of distribution (right)

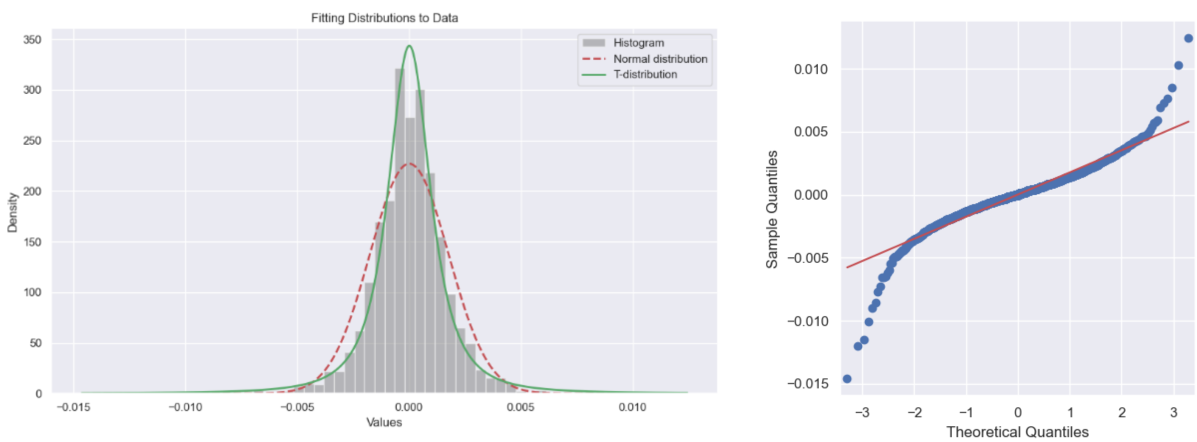

Figure4- Green Bond index – log return distribution in-sample period(left) – QQ plot of distribution (right)

4-2- GBM with Student's t-distribution

The GBM model, which incorporates both a deterministic drift term and a stochastic random shock term, is widely used to simulate the future price paths of financial assets. As discussed in methodology, GBM is a model for modeling asset value and use Montecarlo as a common way for solving this stochastic Markov process. In this part, I focus on random shock generation. the method builds a distribution of potential future values. simulation is based on a chosen probability distribution, like the normal distribution or any other distribution. Since I want to capture tail events I have chosen a t-distribution with a fixed degree of freedom of 3 was used for the random shock. Maximum Likelihood Estimation (MLE) methods are employed to estimate the parameters of the model, including the location parameter, scale parameter, and degrees of freedom (DOF). As indicated in Table

1, the DOF is a crucial parameter in this study and is highly sensitive to the initial guess in the MLE method. During the estimation process, it was observed that the DOF frequently fell within the range of 3 to 6. A fixed value of 3 was selected for the DOF to maintain a conservative approach.

VaR is calculated by identifying the worst-case losses at a specific confidence level, such as the 5th percentile. To estimate the 1-day VaR and ES at a 95% confidence level, Monte Carlo simulations were employed. For each step, 100,000 random trials were generated to capture the potential price movements. In volatile markets, the accuracy of simulations tends to decrease over longer time horizons. The model assumptions were recalibrated using both daily and 14-day time horizons to address this. This recalibration aimed to assess the effectiveness of the model under different temporal conditions. The study aims to provide insights into the trade-offs between responsiveness and accuracy in VaR and ES estimation by comparing the results from these two recalibration frequencies. (Denoted by 1-d and 14-d)

4-3- Forecasting volatility

For the volatility, we can assume that this is the same as the standard deviation of historical return in a specific rolling window. My basic assumption is that it equals to one-year standard deviation. However, it can be refined to more realistic volatility closer to the market condition. It seems that it has a contradicts to GBM approach with the random walk assumption incorporated with an autoregressive assumption of GARCH volatility but when I look at the initial guess for volatility in conventional GBM we need to assume volatility based on the historical volatility of the market or any other way. So this is acceptable to estimate volatility of market based on GARCH and then apply random shock to it. In (A New Monte Carlo Simulation Methodology) presented this approach.

Volatility in the out-of-sample one-step ahead recursive window forecast is used to consider the GARCH:

$$E[\hat{\sigma}_t^2 - \mathrm{V}] = (\hat{\alpha}_0 + \hat{\beta}_1)\, \boldsymbol{E}[\hat{\sigma}_{n+t-1}^2 - V]$$

At time T, all historical values of $\hat{\varepsilon}_T^2$ and $\hat{\sigma}_T^2$ are known. In a one-step-ahead forecast, $\hat{\varepsilon}_T$ is observable as it is the most recently estimated residual in the GARCH model. To forecast volatility for the n-step-ahead period (T+n), we need the $\hat{\varepsilon}_{T+\mathrm{n}}$ that is utilized in the forecast. This process continues up to the last observation in the out-of-sample period. There are ways to update parameters: fixed rolling and recursive. The fixed rolling employs a fixed window of data to re-estimate the parameters, while the recursive window uses an expanding window. Predictions from the previous period are used to forecast the next period, and the data window for parameter re-estimation expands as the model forecasts the next day. There is no agreement on which one is better except for the computation time in the long period. In figures 5,6 volatility in the out-of-sample period from 2017 to 2021 is illustrated for both indices.

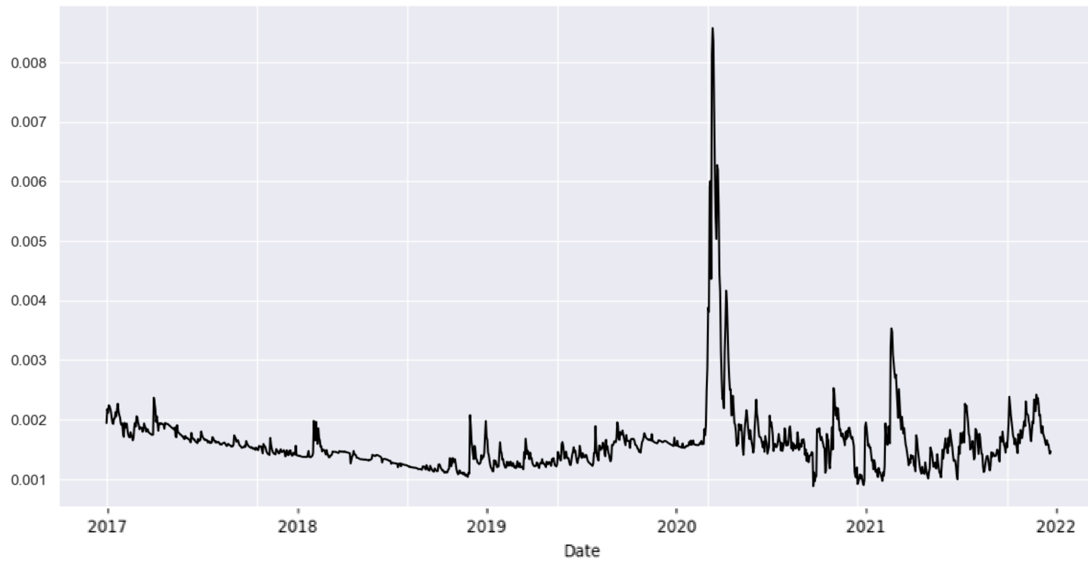

Figure 5 - GB rolling volatility prediction out-of-sample

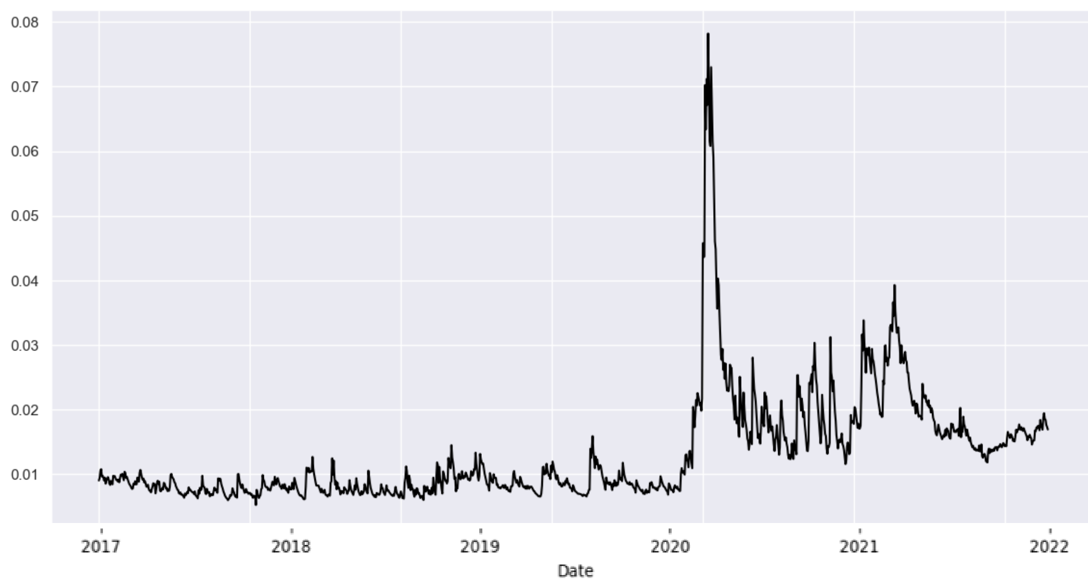

Figure 6 - GCE rolling volatility prediction out-of-sample

4-4- VaR estimation

In this study, I developed and compared 10 different methods to evaluate the impact of incorporating GARCH and t-distribution within a Monte Carlo simulation framework. For each day in the forecast period, I conducted simulations to generate potential future price paths based on the estimated volatility. Each day, I reported the 5th percentile of the simulated return distribution, representing the VaR. Additionally, the ES was calculated to provide a more comprehensive risk measure, capturing the expected loss in the worst-case scenarios. To enhance the accuracy and responsiveness of the Monte Carlo simulations, I recalibrated the model parameters at two different frequencies: daily (1-day recalibration) and bi-weekly (14-day recalibration). bi-weekly recalibration updates the model parameters every 14 trading days. This less frequent recalibration smooths out short-term market noise and focuses on medium-term trends. While this approach may not capture rapid market changes as quickly as daily recalibration, it reduces the computational burden and the potential for overfitting to short-term fluctuation, I will discuss the particle filtering approach to overcome overfitting.

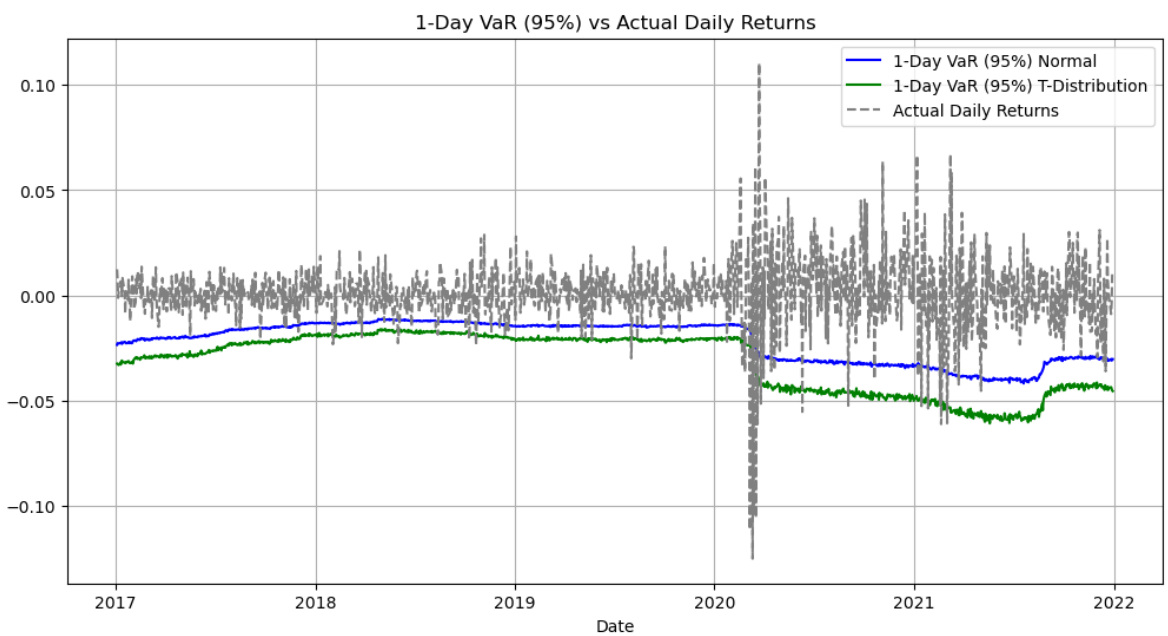

Figure 7- 1-day VaR 95% CL out of sample period with normal and student’s t and fixed volatility for GCE

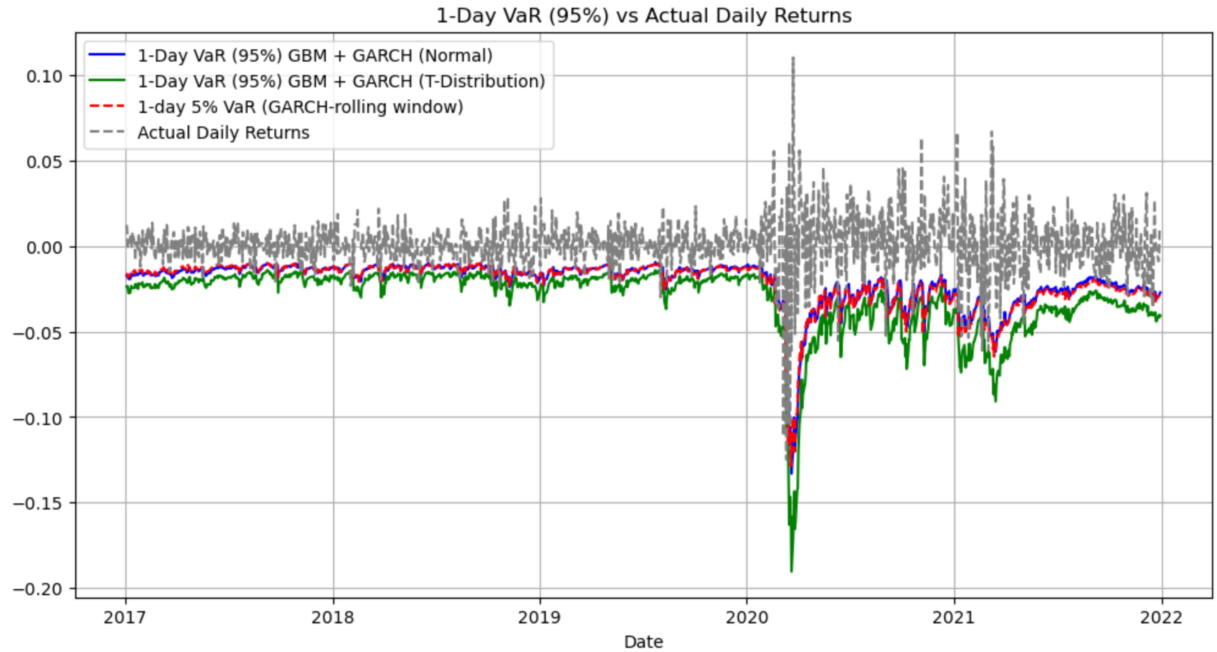

Figure 8- 1-day VaR 95% CL out of sample period with normal and student's t and volatility with GARCH for GCE

VaR in Montecarlo simulation with a rolling window approach responds to shocks ineffectively, resulting in a delayed dampening of the shock effect. In contrast, Montecarlo with the GARCH(1,1) model with the same rolling window immediately responds to fluctuations and dampens the effect more promptly.

4-5- Monte Carlo with Particle filtering

Particle filtering, also known as Sequential Monte Carlo (SMC) methods, provides a framework for estimating the subsequent distribution of state variables in dynamic systems. This approach is particularly useful for financial applications(Nkemnole & Abass, 2019). The particle filter method sequentially updates and resamples particles representing possible future states of the system, incorporating observed data to refine the predictions.

In this study, I employed Monte Carlo simulations integrated with particle filtering instead of the conventional MC approach. The particle filter adjusts the weights of the particles based on observed data and then resamples to generate a new set of particles according to their weights of deviation from the observed price. This method enhances the ability to capture extreme events and handle the intrinsic uncertainty of financial markets. After the prediction of simulated price, two steps were added before the next step generation

- *Weighting:* Adjust the weights of the particles based on observed data. The weighting step is essential for updating the particle filter based on new observed data. Here's a detailed explanation of the weighting approach used in this study: For each particle, calculate the likelihood of the observed price given the predicted price of that particle. This likelihood is computed using a probability density function (PDF) that describes the distribution of the predicted prices. Common choices include the normal distribution and the t-distribution, depending on the characteristics of the financial data.

The weight of each particle is proportional to this likelihood. Mathematically, the weight w_i of particle i is given by:

$$w_i = \text{PDF}\ (\ \text{observed\_price}\ \mid\ \text{predicted\_price}\ _i\ )$$

The PDF used can be the normal distribution $\mathcal{N}(\mu, \sigma)$ or the t-distribution $t(\mu, \sigma, \nu)$ on the model's assumptions about the distribution of returns.

After calculating the weights, they are normalized to sum up to 1. This normalization ensures that the weights represent a valid probability distribution.

$$w_i = \frac{w_i}{\sum_{j=1}^{n_{\text{particles}}} w_j}$$

- *Resampling:* Generate a new set of particles by sampling from the current particles, with replacement, according to their weights. Based on the normalized weights, particles are resampled with replacement. Particles with higher weights have a higher probability of being selected multiple times, while particles with lower weights are less likely to be chosen. This process effectively discards particles that poorly predict the observed price and focuses on those that align well with the observed data.

The new set of particles replaces the old set, and each particle is assigned an equal weight after resampling. This equal weighting ensures that the resampling step starts afresh without bias from the previous weights. This method emphasizes the intrinsic uncertainty of the market. Compared to simple recalibration with observed prices, particle filtering offers significant advantages:

Uncertainty Handling: Particle filtering explicitly accounts for the uncertainty inherent in financial markets by propagating a diverse set of particles through the model. Each particle represents a possible future price path, allowing the model to capture a wide range of outcomes, including extreme events in the tails of the distribution.

Avoiding Overfitting: Simple resampling with exact market prices can lead to overfitting, where the model becomes too closely aligned with past data and loses its predictive power.

Figure 10 and Figure 11 illustrate the predicted bands using the particle filtering approach. The graphs show that the prediction bands become wider when using particle filtering compared to conventional methods. This widening of the bands indicates an increased level of uncertainty in the predictions, which is a significant feature of the particle filtering approach.

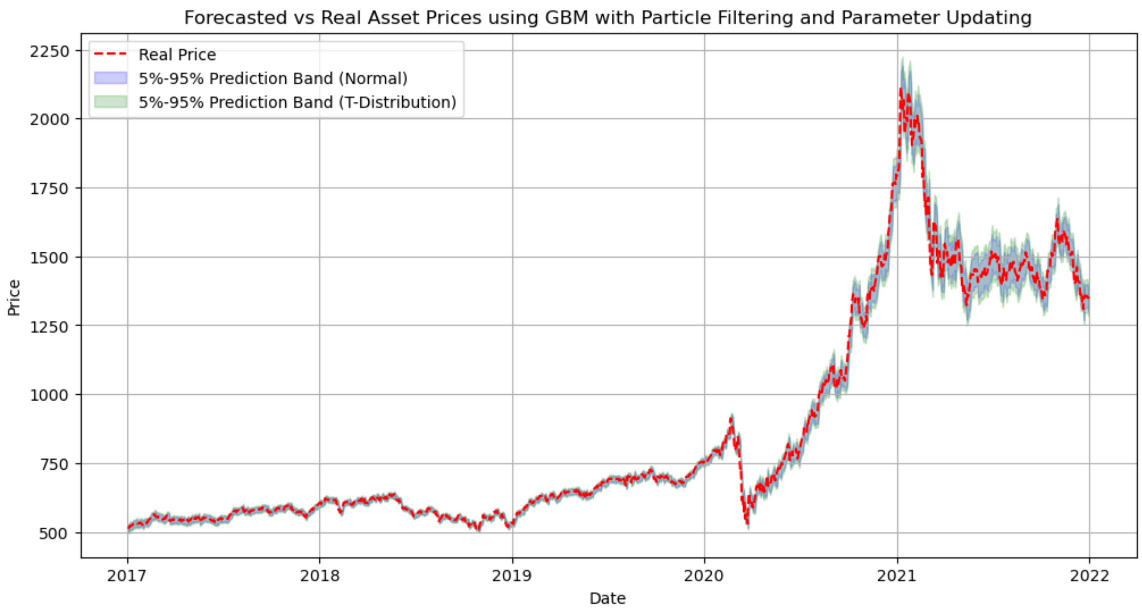

Figure 9- Simple Monte Carlo Simulation - 5th to 95th Percentile Prediction Band for GCE with Daily Updated Parameters

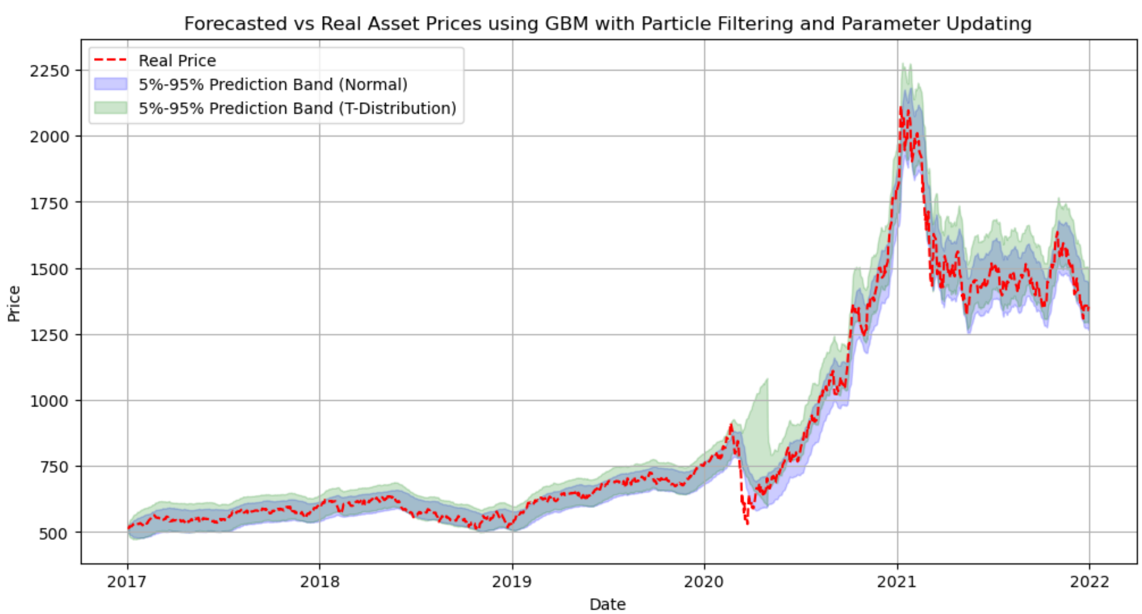

Figure 10- Monte Carlo Simulation with Particle Filtering - 5th to 95th Percentile Prediction Band for GCE with Daily Updated Parameters

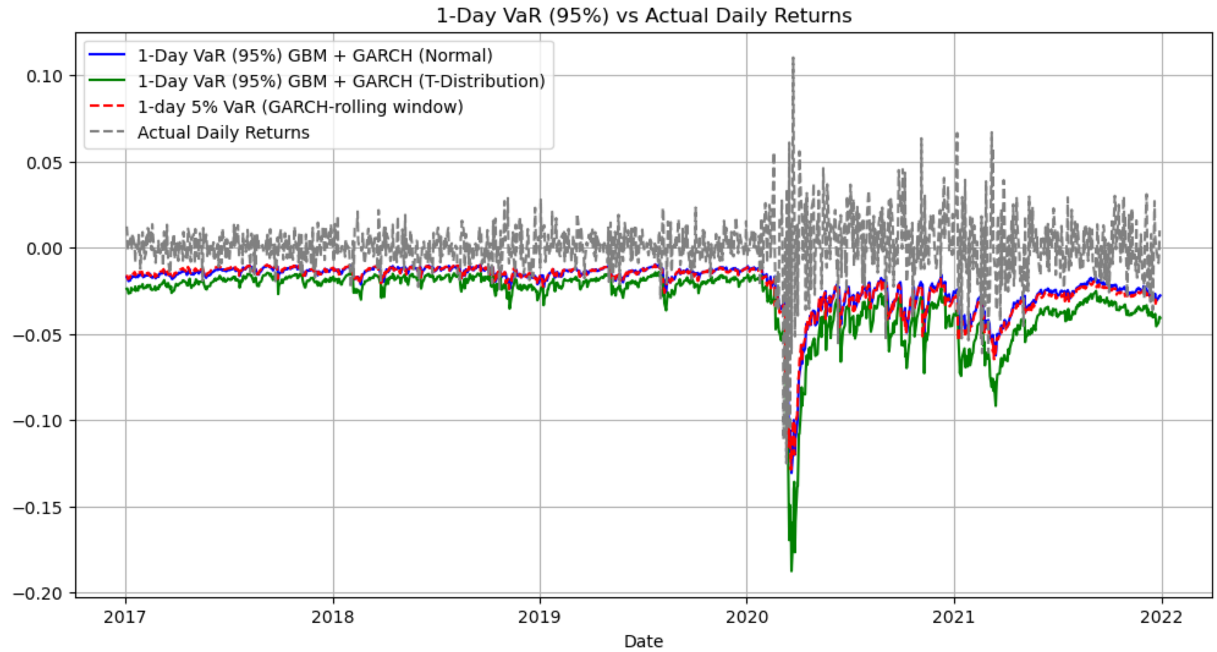

Figure 11- 1-day VaR 95% CL out of sample period with particle filtering Monte Carlo approach

5- Analysis and Results

The analysis employs various statistical methods to model extreme risks and assesses the accuracy and reliability of these models using backtesting techniques. The objective is to identify which methods provide the most accurate predictions for extreme events in the financial markets represented by these indices. While the entire period is initially presented, the focus then shifts to the period from 2020 to the end of 2021, capturing the impact of the COVID-19 outbreaks and their recovery post-event prolonged to the end of 2021.

5-1- Kupiec POF test

Backtesting is a set of procedures used to verify the actual losses observed are matched with what I expected. The Kupiec-POF test represents the widely used test for assessing the reliability of these risk models in process of backtesting. If the model passes the Kupiec test, it means the number of exceptions is consistent with the expected number based on the confidence level. The test is concerned whether the reported VaR is violated than α percent of time(Zhang & Nadarajah, 2018b). The test statistic is

$$POF = 2\ln\left[\left(\frac{1-\hat{\alpha}}{1-\alpha}\right)^{n-I(\alpha)}\left(\frac{\hat{\alpha}}{\alpha}\right)^{I(\alpha)}\right]$$

$$\hat{\alpha} = \frac{1}{n}I(\alpha)$$

$$I(\alpha) = \sum_{t=1}^{n} I_t(\alpha),$$

The test statistic value is obtained from the chi-square distribution with 1 degree of freedom. Where n denotes the number of observations, the test statistic indicates that if the proportion of VaR violations is exactly equal to 100α percent, the POF test value will be zero, showing no evidence of inadequacy in the underlying VaR model. However, as the proportion of VaR violations deviates from α percent, the POF test statistic increases, providing growing evidence that the proposed VaR model either consistently understates or overstates the portfolio's actual risk level. As show in table 2, For a 95% confidence level, we expect approximately 25.2 exceedances out of the total sample size.

This is a very simple test that shows us when I forecast the volatility term with GARCH model (t-distribution) it makes the model more conservative and overestimate the VaR. normal distribution in Montecarlo simulation with GARCH volatility forecasting can be consistent. Particle filtering slightly changed the number of exceedances and made the model in this period a bit conservative comparing one method. The Simple GARCH(1,1) model is known for its reliability in estimating the volatility of assets, especially assets that exhibit high levels of volatility, such as GCE. This model is particularly effective because it captures time-varying volatility by considering both past returns and past forecast errors. However, in the case of GB, the GARCH(1,1) model faced challenges. Specifically, during extreme conditions, the exceedance levels can violate the established thresholds but the level of exceedance is lower than the historical approach.

Method	Num of Exceedances (25.2 expected 95%)		Kupiec Test Statistic		Kupiec Test p-value	
	GCE	GB	GCE	GB	GCE	GB
GBM-RW-n-1d	28	20	0.317	1.212	**0.574**	**0.271**
GBM-RW-t-1d	19	13	1.749	7.499	**0.186**	0.006
GBM-RW-n-14d	30	23	0.909	0.208	**0.340**	**0.648**
GBM-RW-t-14d	19	13	1.749	7.499	**0.186**	0.006
GBM-RW-n-1d-SMC	29	23	0.576	0.208	**0.448**	**0.648**
GBM-RW-t-1d-SMC	19	13	1.749	7.499	**0.186**	0.006
GBM-garch-n-1d	12	30	8.954	0.909	0.003	**0.340**
GBM-garch-t-1d	2	13	37.372	7.499	0.000	0.006
GARCH (1,1)	12	27	8.954	0.132	0.003	**0.716**
GBM-garch-n-14d	20	30	1.212	0.909	**0.271**	**0.340**
GBM-garch-t-14d	4	17	28.601	3.156	0.000	0.076
GBM-garch-n-1d-SMC	12	31	8.954	1.313	0.003	**0.252**
GBM-garch-t-1d-SMC	1	12	43.149	8.954	0.000	0.003
RW – 1 year	37	31	5.115	1.313	0.024	**0.252**

Table 2 - Kupiec POF test

5-2- Lopez magnitude loss function

Lopez (1998, 1999) introduced a method to measure the distance between observed returns and the forecasted VaR (α) when a violation occurs. loss function accounts for the total number of violations and the squared distance from the corresponding VaR. By utilizing this loss function, we gain a better insights into the performance of their VaR models and assess their ability to capture and estimate market risk accurately. Results in ascending order are shown in table 3. GCE is more volatile than GB so particle filtering helps slightly to better match the actual return.

Lopez function(Zhang & Nadarajah, 2018b):

$$L_t(x_t, \mathrm{VaR}_t\ (\alpha)) = \begin{cases} 1 + [x_t - \mathrm{VaR}_t\ (\alpha)]^2, & x_t \leq \mathrm{VaR}_t \\ 0, & x_t > \mathrm{VaR}_t \end{cases}$$

$$\hat{L} = \frac{1}{n}\sum_{t=1}^{n} L_t(x_t, \mathrm{VaR}_t\ (\alpha))$$

	GB	GCE
GBM-GARCH-t-1d-SMC	0.014	0.001
GBM-GARCH-t-1d	0.014	0.002
GBM-RW-t-1d-SMC	0.017	0.022
GBM-RW-t-1d	0.017	0.024
GBM-RW-t-14d	0.019	0.032
GBM-GARCH-t-14d	0.021	0.003
GBM-RW-n-1d	0.038	0.051
GBM-RW-n-1d-SMC	0.040	0.051
GBM-RW-n-14d	0.041	0.052
GARCH (1,1)	0.042	0.025
GBM-GARCH-n-14d	0.046	0.033
GBM-GARCH-n-1d	0.048	0.027
GBM-GARCH-n-1d-SMC	0.048	0.026
RW-1y	0.053	0.057

Table 3 – Lopez loss function ($\alpha = 0.05$)

The loss function can reveal the inefficiency of certain methods, such as GARCH (1,1) and all methods that assume a normal distribution. Even when the number of exceedances is similar across these methods, the loss function highlights how far the predicted returns deviate from the actual returns. The lowest loss values are observed for the GBM-GARCH-t-1d-SMC and GBM-GARCH-t-1d models (0.014), indicating that these models perform well in capturing the return dynamics of GB. t-distribution-based models are effective for both. However, incorporating GARCH in GCE needs to be estimated with the t-distribution to achieve optimal performance and reliability.

5-3 Unconditional Test by Acerbi and Szekely

Acerbi showed that ES can in practice be jointly elicited with VaR. The test by (Acerbi & Szekely, 2014) is designed to evaluate the accuracy of risk models in predicting extreme losses in a financial portfolio. The test focuses on the Expected Shortfall. This test evaluates whether the ES predictions from a model match the observed extreme losses in a given dataset. Returns simulated under the null hypothesis that the risk model accurately predicts ES and VaR.

The unconditional test is sensitive to the magnitude and number of exceedances in the unconditional test, losses are scaled by the expected number of exceedances, given a significance level (α) and estimation window (T). The null and alternative hypotheses are defined as:

$H_0: ES_{\alpha,t}^{uF} = ES_{\alpha,t}, \forall t$

The observed Expected Shortfall (ES) is consistent with the Expected Shortfall predicted by the risk model.

$H_1: ES_{\alpha,t}^{uF} \geq ES_{\alpha,t} \forall t$

The observed Expected Shortfall (ES) is not consistent with the Expected Shortfall predicted by the risk model.

The Acerbi-Szekely method provides a rigorous approach to backtesting ES by comparing the observed ES with a benchmark distribution derived from simulations to estimate the distribution of the ES statistic under the null hypothesis. This involves generating a large number of simulated return paths and calculating the ES for each path. In table 4, you can find the result with 5% significance level all models with rolling window estimation of volatility even along with t-distribution reject the null hypothesis and are inconsistent in this study.

	p-value GB	p-value GCE
GBM-RW-t-14d	0.0004	0.0028
GBM-RW-t-14d-SMC	0.0007	0.0017
GBM-RW-t-1d-SMC	0.0034	0.0173
GBM-RW-t-1d	0.0036	0.0113
GBM-GARCH-t-14d	0.1478	0.2373
GBM-RW-n-14d	0.2226	0.2528
GBM-RW-n-1d-SMC	0.4922	0.5681
GBM-RW-n-1d	0.638	0.6189
RW – 1Y	0.9115	0.8534
GBM-GARCH-n-14d	0.9188	0.9221
GBM-GARCH-t-1d	0.9909	0.7434
GBM-GARCH-t-1d-SMC	0.9953	0.8875
GBM-GARCH-n-1d-SMC	1	1
GBM-GARCH-n-1d	1	1
GARCH (1,1)	1	0.9997

Table 4 - Acerbi and Szekely test ($\alpha = 0.05$)

Incorporating these tests in Table 5 shows five methods that pass the test and significantly estimate VaR.

	GB			GCE		
	POF	Loss	Acerbi	POF	Loss	Acerbi
GBM-RW-n-1d	20	7	0.638	28	11	0.6189
GBM-RW-t-1d	13	4	0.0036	19	5	0.0113
GBM-RW-n-14d	23	9	0.2226	30	13	0.2528
GBM-RW-t-14d	13	5	0.0004	19	6	0.0028
GBM-RW-n-1d-SMC	23	8	0.4922	29	12	0.5681
GBM-RW-t-1d-SMC	13	3	0.0034	19	4	0.0173
GBM-garch-n-1d	30	12	1	12	9	1
GBM-garch-t-1d	13	2	0.9909	2	2	0.7434
GARCH(1,1)	27	10	1	12	7	0.9997
GBM-garch-n-14d	30	11	0.9188	20	10	0.9221
GBM-garch-t-14d	17	6	0.1478	4	3	0.2373
GBM-garch-n-1d-SMC	31	13	1	12	8	1
GBM-garch-t-1d-SMC	12	1	0.9953	1	1	0.8875
RW 1-y	31	14	0.9115	37	14	0.8534

Table 5 – Aggregating all tests – loss function numbered in order from 1-14

6- Discussion and conclusion

In this study, I developed a framework to assess extreme risks associated with S&P Green Bonds (GB) and Global Clean Energy (GCE) using Monte Carlo simulations enhanced with advanced statistical models. By applying Value at Risk (VaR) and Expected Shortfall (ES) as key risk indicators, I evaluated the performance and stability of these measures with three statistical tests.

The VaR with t-distribution and GARCH for GB fluctuated between 0.1% and 1.7%, while for GCE it ranged from 1.6% to 15.2%. Although I did not investigate the specific causes of these shocks, it is evident that the GCE market was significantly impacted by the COVID-19 outbreak.

The Monte Carlo simulation approach was used with several assumptions. Based on the Efficient Market Hypothesis (EMH), market prices evolve according to a random walk and are unpredictable. The Monte Carlo simulation utilizes Geometric Brownian Motion (GBM) to generate random future states. I modified two conventional assumptions regarding the shock and volatility terms in GBM. Then, I applied particle filtering to Monte Carlo simulations to capture extreme risk events. By shifting from a normal to a t-distribution, I aimed to fit a more realistic distribution of returns. This adjustment decreased the exceedance level to match the 95% confidence level, reducing it from 28 in two years with a normal distribution to 19 with a t-distribution in GBs. Additionally, I incorporated the volatility term, another impactful factor in GBM, by fitting a GARCH model to sample data. This further decreased the exceedance level from 28 in two years to 12. However, incorporating both modifications led to an overestimation of VaR with 95% CL. On the other hand, particle filtering did not significantly affect the results. The Acerbi test also highlighted the importance of forecasting volatility with GARCH, which outperformed all other models. When examining all models, particle filtering Monte Carlo with GARCH and t-distribution provided the most accurate results based on exceedance and loss function tests, with only one exceedance in GCE and 12 in GB, and the lowest loss value. However, in the Acerbi test, the simple Monte Carlo method produced more consistent results. Considering the methods from various perspectives, in terms of both VaR and ES, the simple Monte Carlo with GARCH and t-distribution also demonstrated more consistent performance.

The time horizon for simulation is another factor examined in this research. Parameters need to be updated frequently to reflect changing market conditions. During normal conditions, price movements remain within a predictable range, allowing simulations to rely on initial parameters. However, there is no evidence that in a volatile market, parameters need frequent updating. This is why modeling with a two-week updating frequency underestimates risk compared to daily updates. It is not the goal of this research to find the optimal time horizon, but this can be studied in future research.

If the ultimate goal is to develop a framework for modeling extreme risk, these models should be tested on other asset classes to enhance Monte Carlo simulation accuracy. Additionally, exploring other types of distributions could provide further improvements. Future research could benefit from using Generalized Geometric Brownian Motion (gGBM), which can incorporate memory effects and capture more complex asset dynamics. This approach addresses GBM's limitations in capturing sudden jumps in asset prices and its assumption that price changes are independent and identically distributed. the crucial role of gGBM in handling structural changes. These models are usually used for option pricing and require a strong knowledge of stochastic calculus. Furthermore, future research should focus on external factors such as macroeconomic indicators and geopolitical events that influence market behavior.

References

Abad, P., Benito, S., & López, C. (2014). A comprehensive review of Value at Risk methodologies. *The Spanish Review of Financial Economics*, *12*(1), 15–32.

Acerbi, C., & Szekely, B. (2014). *BACKTESTING EXPECTED SHORTFALL Introducing three model-independent, non-parametric back-test methodologies for Expected Shortfall*.

Bailey, R. W. (1994). Polar generation of random variates with the t-distribution. *Mathematics of Computation*, *62*(206), 779–781.

Balkema, A. A., & De Haan, L. (1974). *Limit laws for order statistics*.

Barone-Adesi, G., Giannopoulos, K., & Vosper, L. (1999). VaR without correlations for portfolios of derivative securities. *Journal of Futures Markets*, *19*(5), 583–602.

De Domenico, F., Livan, G., Montagna, G., & Nicrosini, O. (2023). Modeling and simulation of financial returns under non-Gaussian distributions. *Physica A: Statistical Mechanics and Its Applications*, *622*, 128886.

Engle, R. F., & Manganelli, S. (2004). CAViaR: Conditional autoregressive value at risk by regression quantiles. *Journal of Business & Economic Statistics*, *22*(4), 367–381.

Fama, E. F. (1963). Mandelbrot and the stable Paretian hypothesis. *The Journal of Business*, *36*(4), 420–429.

Gumbel, E. J. (1958). *Statistics of extremes*. Columbia university press.

Huisman, R., Koedijk, K. G., & Pownall, R. A. (1998). VaR-x: Fat tails in financial risk management. *Journal of Risk*, *1*(1), 47–61.

Kitagawa, G., & Sato, S. (n.d.). *Monte Carlo Smoothing and Self-Organising State-Space Model*.

McNeil, A. J. (1998). *Calculating quantile risk measures for financial return series using extreme value theory*. ETH Zurich.

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*(3–4), 271–300.

Merton, R. C. (1976). Option pricing when underlying stock returns are discontinuous. *Journal of Financial Economics*, *3*(1–2), 125–144.

Neslihanoglu, S., & Date, P. (2019). A modified sequential Monte Carlo procedure for the efficient recursive estimation of extreme quantiles. *Journal of Forecasting*, *38*(5), 390–399.

Nkemnole, B., & Abass, O. (2019). Estimation of geometric Brownian motion model with a t-distribution–based particle filter. *Journal of Economic and Financial Sciences*, *12*(1). https://doi.org/10.4102/jef.v12i1.159

Patton, A. J., Ziegel, J. F., & Chen, R. (2019). Dynamic semiparametric models for expected shortfall (and Value-at-Risk). *Journal of Econometrics*, *211*(2), 388–413. https://doi.org/10.1016/j.jeconom.2018.10.008

Shonkwiler, R. W. (2013). *Finance with Monte Carlo*. Springer.

Smith, V. K. (2022). Chapter 23: Can environmental bonds manage policy-induced risks? In *The Economics of Environmental Risk* (pp. 332–365). Edward Elgar Publishing. https://doi.org/10.4337/9781035301614.00033

Tsoukala, A., & Tsiotas, G. (2021). Assessing green bond risk: an empirical investigation. *Green Finance*, *3*(2), 222–252. https://doi.org/10.3934/GF.2021012

Zhang, Y., & Nadarajah, S. (2018a). A review of backtesting for value at risk. In *Communications in Statistics - Theory and Methods* (Vol. 47, Issue 15, pp. 3616–3639). Taylor and Francis Inc. https://doi.org/10.1080/03610926.2017.1361984

Zhuo, J., Ren, X., & Duan, K. (2023). Modelling extreme risks for green bond and clean energy. *Environmental Science and Pollution Research*. https://doi.org/10.1007/s11356-023-27071-5