Politecnico di Milano

FINANCIAL ENGINEERING

A.Y. 2023/2024

Filtered Historical Simulation

Final Project RM1

Authors:

Andrea Meschieri
Silvia Vittoria Idda

Table of Contents

Summary

Value at Risk (VaR) is a statistical measure used to quantify the potential maximum loss risk of an investment portfolio within a given confidence interval and over a given time horizon. Starting from the standard Historical Simulation approach, we developed a family of filter models including Short-Term Volatility filters (STV) and Long-Term Volatility filters (LTV). Moreover, we implemented two generalizations of the volatility filter. The first generalization tries to handle the covariance between each risk factor in the multi-dimensional case. The second is a class of filters that computes and updates the moments of the returns time series. We applied these models to real data and studied the impact of filtering on some properties of the risk factors distribution. Then, we introduced two back-testing methods and used them to assess the performance of our models.

Overall, our study demonstrates that the introduction of volatility filters and their generalizations can significantly enhance the accuracy and reliability of VaR estimates.

1 ARCH test

The ARCH test is an important tool in econometrics and finance to detect heteroscedasticity, which is crucial for accurate risk assessment, especially in Value at Risk (VaR) calculations. When dealing with homoscedasticity, parametric models can be effectively used to estimate the parameters of the distribution, such as the mean (μ) and variance (Σ). This allows a relatively simple calculation of VaR since losses, if returns X_t are Gaussian, can be represented by the following relationship:

$$\frac{L(X_t)}{V_t} \sim \mathcal{N}\left(-w_t \cdot \mu, w_t \cdot \Sigma w_t\right)$$

However, if the ARCH test indicates the presence of heteroscedasticity in the data set, it means that the variability of the errors or residuals in the regression model is not constant for all observations. In such cases, the use of parametric models may not be appropriate, and a nonparametric approach becomes necessary for accurate risk estimation.

In a nonparametric approach, instead of relying on specific distributional assumptions, methods that directly estimate the empirical distribution of the data or techniques such as bootstrapping are used to generate distributions to provide more robust estimates of risk.

Running this test with the built in function of matlab *archtest* our log-returns do not turn out to be homoschedastic, so we can only use a non parametric approach.

2 VaR implementation

In order to calculate the Value at Risk of a portfolio we need an estimate of the distribution of the losses over a fixed risk time horizon. In our case we estimated a one day VaR using a non parametric approach, looking at the historical time series of the stocks present in our portfolio. Moreover, in order to make the calculations more tractable, we used the following assumptions:

- Frozen portfolio (the composition of the portfolio stays constant over the risk horizon)
- Linear portfolio (the losses are a linear function of the risk factors)

Under these assumptions the losses over a 1 day horizon can be represented in this way:

$$L(X_t) = -V(t)(w \cdot X_t)$$

where $V(t)$ is the value of the portfolio at time t, X_t is the vector of the risk factors and w is the vector of weights (or sensitivities) related to each factor. Considering only stock prices as potential risk factors we can characterize the vectors X_t, w and $V(t)$:

$$X_i = ln(\frac{S_i(t+\delta)}{S_i(t)})$$

$$w_i = \frac{n_i S_i(t)}{V(t)} \quad and \quad V(t) = \sum_{j=1}^{n} n_j S_j(t)$$

where $S_i(t)$ is the i^{th} stock price at time t and n_i is the corresponding number of stocks.

2.1 Historical Simulation

The implementation of a historical simulation without filtering is very straightforward. After describing the losses at time t as a function of the vector of risk factors we can consider the empirical distribution generated by the time series $\{L(X_i)\}_{i=t}^{i=t-N}$. Since the value at risk is defined as $VaR_{1-\alpha} = \inf\{l \in \mathbb{R} : \Pr[L > l] \leq \alpha\}$, we ordered the losses $L(X_i)$ in ascending order and selected the quantile of order $1-\alpha$, meaning the $L(X_q)$ such that $q = \lfloor N(1-\alpha) \rfloor$. We assumed $\alpha = (1 - ConfidenceInterval)$ and N is the number of past observations.

2.2 General volatility filtering method

We now introduce a family of filtering methods that try to capture the dynamics of the volatility of the risk factors. The idea is to calculate for each day i an estimate of the volatility of each risk factor. These volatility estimates allow us to "normalize" the impact of each past observation and better fit the dynamics of the losses. Each volatility is calculated as an Exponentially Weighted Moving Average (EWMA), meaning:

$$\sigma_n^\lambda(t,j)^2 = Var_n^\lambda(X_j(t)) = \lambda \sigma_n^\lambda(t-1,j)^2 + (1-\lambda) X_j(t-1)^2$$

where $\sigma_n^\lambda(t-n,j)^2$ is the sample variance of $\{X_j(t-i)\}_{i=1}^{i=n}$ and λ is a chosen decay factor. The general volatility filter is map depending on four parameters, the decay factors λ and μ and the time windows P and Q.

$$\mathcal{F}(\lambda, \mu, P, Q) : \{X_j(t-i)\}_{i=1}^{N} \mapsto \{\widehat{X}_j(t-i)\}_{i=1}^{N}$$

where:

$$\widehat{X}_j(t-i) = \frac{\sigma_Q^\mu(t,j)}{\sigma_P^\lambda(t-i,j)} X_j(t-i)$$

In the filter definition σ_Q^μ is called the "revol" volatility while σ_P^λ is called the "devol" volatility.

2.3 Long Term Filtering (LTV) and Short Term Filtering (STV)

After introducing the general framework we focus on two specific filtering methods: long term and short term filtering. In the former approach we have a constant devol volatility, which means setting our filter parameter $\lambda = 1$. Moreover, we conventionally set the time windows $P = 500, \quad Q = 500$ and the revol parameter $\mu = 0.97$. This means that our filter takes this form:

$$\mathcal{F}_{LTV} = \mathcal{F}(1, 0.97, 500, 500)$$

When it comes to the short filter parameters we need to introduce a scaling factor $\lambda = 0.97$ in order to weigh each devol volatility. This more refined volatility re-scaling is supposed to increase how reactive the estimate is to short term volatility changes. The filter form is the following:

$$\mathcal{F}_{STV} = \mathcal{F}(0.97, 0.97, 500, 500)$$

3 Application

We perform our analysis to 1.500 daily returns from two futures contracts, BRENT and West Texas Intermediate ('WTI'), all starting at the end of February 2014.

3.1 Volatility estimate

The volatility for each portfolio is calculated as follow:

$$\sigma_n^\lambda(t, j) = \sqrt{\lambda \sigma_n^\lambda(t-1, j)^2 + (1-\lambda) X_j(t-1)^2}$$

To ensure robust volatility estimates, the first 500 days of data are used to calibrate the unweighted long-term volatility estimates and the EWMA volatilities. Consequently, our analysis focuses on the subsequent 1.000 days, providing us with a volatility vector of 1.000 elements for each futures contract.

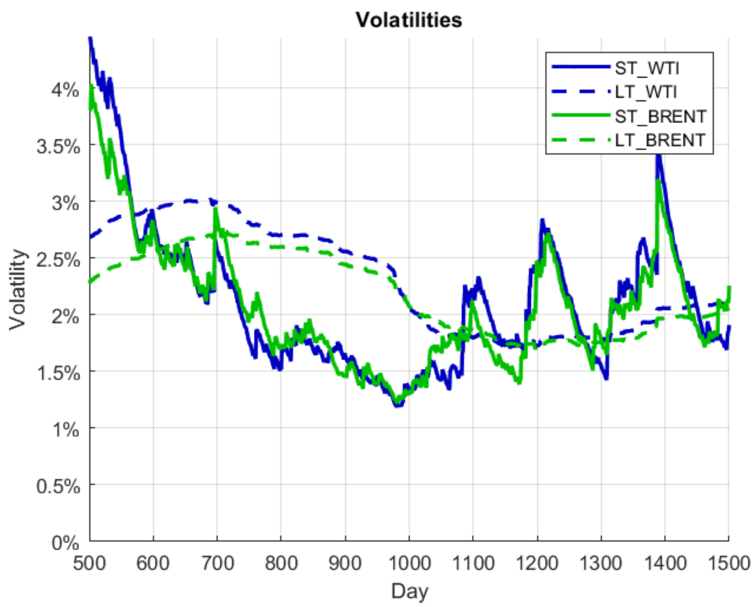


Figure 1: An illustration of short and long term volatility estimates for the two data series made by the STV and LTV models

The EWMA approach effectively captures the dynamic nature of volatility, reflecting periods of heightened and subdued market risk. This time-varying aspect of volatility is represented through a short-term filter factor that adjusts in response to recent return shocks. The variability in conditional volatility directly influences VaR estimates, highlighting the significance of adaptive models in risk management.

3.2 Filter factors

For the WTI data, we plotted the filter factor over a 1000-day time period.

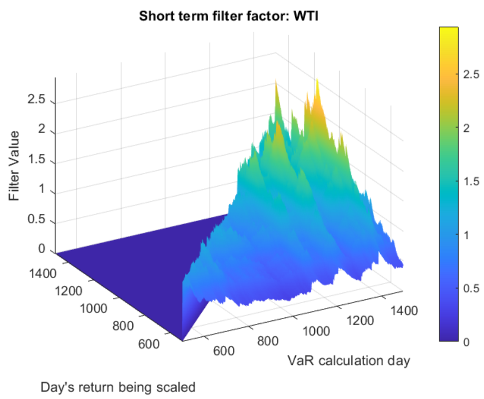


Figure 2: An illustration of short term filter factors for WTI

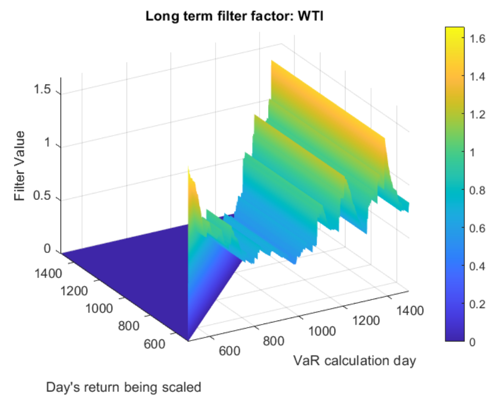


Figure 3: An illustration of long term filter factors for WTI

The plot of the long-term filter factors remains constant over time since it applies the same filter factor to all returns. In contrast, the short-term filter factor is more sensitive to changes in volatility. Figure 2 reflects the characteristics of Figure 1. For instance, between days 600 and 1100, the short-term volatility (STV) line is mostly below the long-term volatility (LTV) line, indicating that the filter factor is less than one during this period.

3.3 VaR calculation

To compute the VaR, we consider a low-variance portfolio, i.e., a portfolio composed by a long-short position in WTI vs Brent, where the number of WTI stocks is equal to 1 and the number of Brent stocks is equal to -1. Under these conditions, we calculate the portfolio value and the VaR using three different methods:

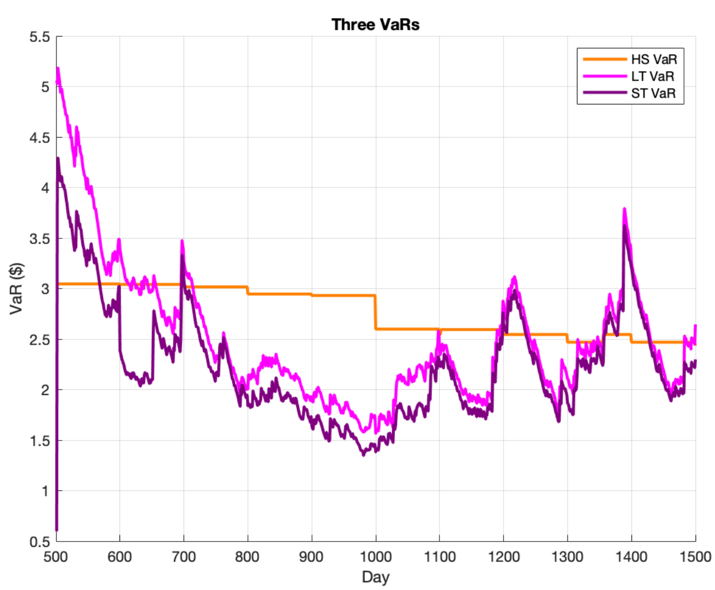


Figure 4: An illustration of ordinary, short and long term filtered VaR estimates for a low variance portfolio

The conditional volatility trends depicted in Figure 1 directly influence the estimation of Value-at-Risk

(VaR) within the filtered models. For instance, the two observed spikes in Exponentially Weighted Moving Average (EWMA) volatility for WTI from days 1200 to 1300 and 1300 to 1450 are mirrored by corresponding spikes in both Filtered Historical Simulation VaRs during the same periods. Additionally, it is evident that the VaR estimates show a relatively low sensitivity to the choice of the 'devol' volatility measure, as both Long-Term (LT) and Short-Term (ST) VaRs generally exhibit similar trajectories over time.

3.4 The Impact of Scaling on Skewness and Kurtosis

Filtering modifies various properties of the return distribution such as skewness and kurtosis. For a process $x_j(i)$ that follows a normal distribution $\mathcal{N}(0, \sigma)$, the LTV filtered returns will also be normally distributed, with a standard deviation equal to the most recent updated estimate $\sigma_j(t)$. Consequently, the percentiles of the distribution will be rescaled accordingly.

The impact of scaling is more complex for non-normal returns with a variable filter factor, as this introduces additional variability into the returns. This variability can result in more pronounced changes in the higher moments of the distribution. Insight into these effects can be obtained by comparing the descriptive statistics of the untreated data to those of the scaled data:

	Unfiltered Data		**Long Term Scaled Data**	
	WTI	BRENT	WTI	BRENT
Mean	-5.52e-04	-4.99e-04	-4.57e-04	-5.27e-04
Standard Deviation	0.0230	0.0214	0.0191	0.0226
Skewness	0.1999	0.2928	0.1999	0.2928
Kurtosis	6.0993	5.475	6.099	5.475

Table 1: The first four moments of the unfiltered and long term filtered risk factor returns

	Unfiltered Data		**Short Term Scaled Data**	
	WTI	BRENT	WTI	BRENT
Mean	-5.52e-04	-4.99e-04	1.25e-04	3.26e-04
Standard Deviation	0.0230	0.0214	0.0192	0.0227
Skewness	0.1999	0.2928	-0.2006	-0.1454
Kurtosis	6.0993	5.4747	5.8898	4.2559

Table 2: The first four moments of the unfiltered and short term filtered risk factor returns

For the long-term scaled data, as expected, the skewness and kurtosis of the scaled returns are identical to those of the original data. This indicates that the long-term scaling process maintains the shape of the distribution of returns. In contrast, for the short-term scaling approach, skewness and kurtosis are not preserved. The short-term scaling process created a data set that reacts to changes in short-term volatility, but it did so at the cost of transforming the distribution of returns in a less transparent way.

In general, scaling does not preserve the relationship between the standard deviation and the percentiles of the distribution. In addition, the effects of rescaling tend to be more pronounced as the decay parameter decreases, meaning that more weight is given to recent observations:

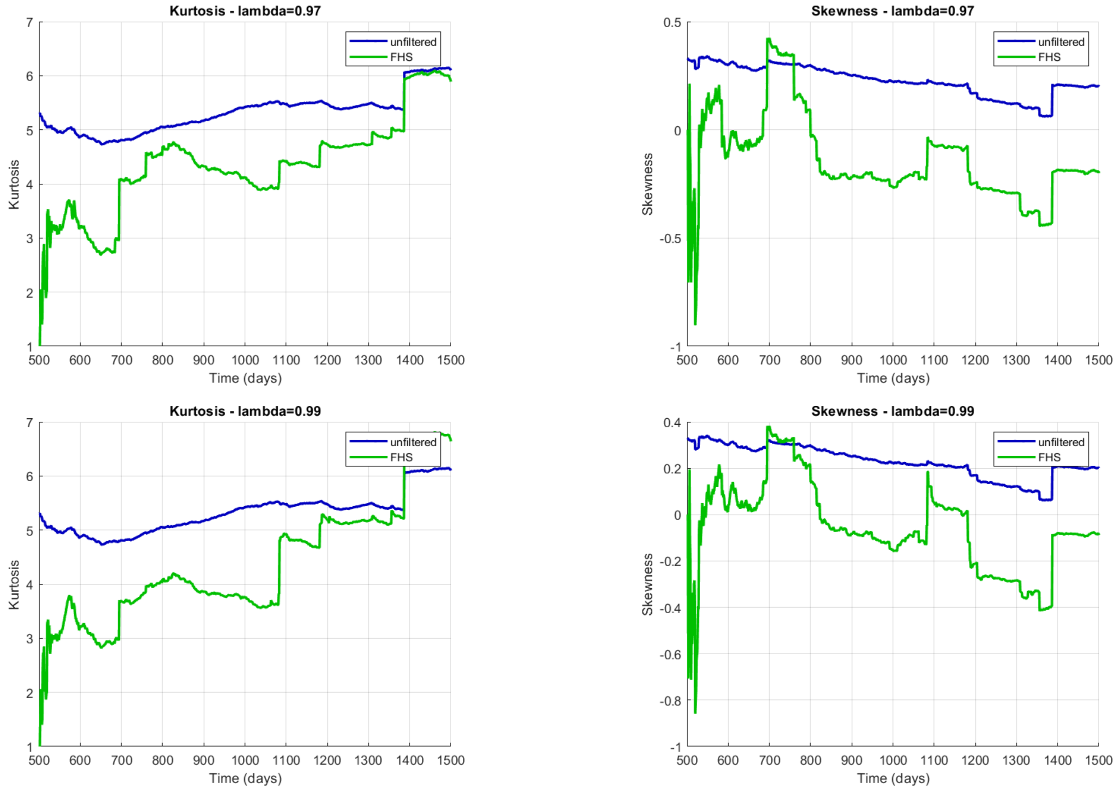
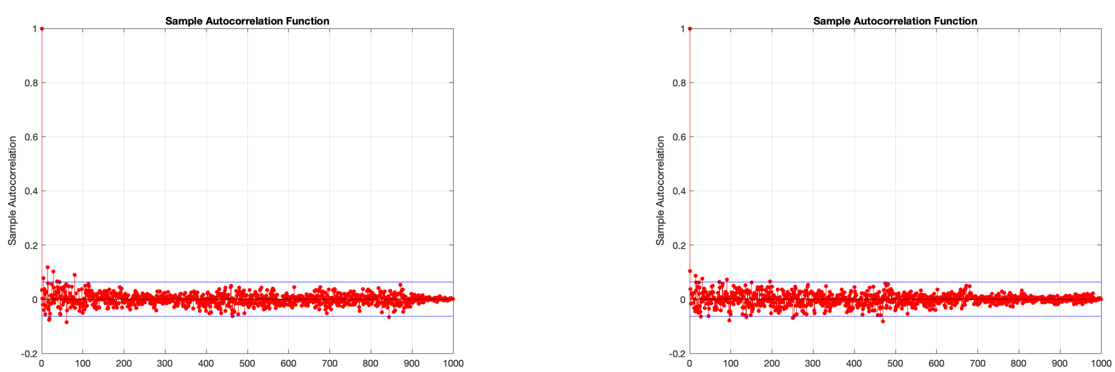


Figure 5: An illustration of the effect of volatility scaling on kurtosis and skewness estimates when using two different FHS decay factors

3.5 Autocorrelation

The rescaling process has a statistically significant impact on the autocorrelation structure. Autocorrelation is a statistical measure that represents the degree of similarity between a signal and a lagged version of it. Through the built in function of matlab *autocorr*, we can observe the autocorrelation of our risk factors:



(a) WTI

(b) BRENT

Figure 6: An illustration of the effect of volatility scaling on the autocorrelation

On the y-axis, we have the autocorrelation values for each lag, i.e., the number of lag samples. A value of 1 at lag zero is expected, since a signal is perfectly correlated with itself without any lag. The horizontal blue lines indicate the confidence limits for autocorrelation. If the autocorrelation values are within the confidence limits (5% in this case), they can be considered not significantly different from zero. Autocorrelation values outside the confidence interval indicate strong autocorrelation, positive or negative, at those specific lags.

4 Correlation Updating

4.1 Covariance Filtering

In this section we introduce the Covariance Filtering approach (CF) as an evolution of the standard volatility filtering methods. The objective is to capture in our Var estimate not only the variance of the risk factors but also their potential correlation. This can be done by extending the definition of the filter to include the covariance matrices of the risk factors. Let's first define the weighted (EMWA) covariance as:

$$Cov_n^\lambda(X_i(t), X_j(t)) = \lambda Cov_n^\lambda(X_i(t-1), X_j(t-1)) + (1-\lambda)X_i(t-1)X_j(t-1)$$

where $Cov_n^\lambda(X_i(t-n), X_j(t-n))$ is the sample covariance between $\{X_i(t-k)\}_{k=1}^{k=n}$ and $\{X_j(t-k)\}_{k=1}^{k=n}$, and λ is a chosen decay factor. Then it's possible to reformulate the usual devol and revol volatilities as matrices as follows:

$$\left(\Sigma_n^\lambda(t)^2\right)_{i,j} = Cov_n^\lambda(X_i(t), X_j(t))$$

The filter definition becomes:

$$\mathcal{F}_{COV}(\lambda, \mu, P, Q) :\mapsto \{\widehat{\overline{X}}(t-i)\}_{i=1}^N$$

$$\widehat{\overline{X}}(t-i) = \Sigma_Q^\mu(t) \circ \Sigma_P^\lambda(t-i)^{(-)}\overline{X}(t-i)$$

where: $\Sigma_Q^\mu(t)$ is the revol matrix and $\Sigma_P^\lambda(t-i)^{(-)}$ is the pseudo-inverse (Moore-Penrose) of the devol matrix.

4.2 Moment based filtering

The idea of filtering past samples using a volatility re-scaling is a nice way to incorporate the variability of the distribution of the risk factors in our model. However, this approach can present some drawbacks. For example, we lose the meaning of how such re-scaling affects the initial distribution. This can be understood in the following example: given a stationary historical sample $\{X(t-i)\}_{i=1}^N \sim N(\mu, \sigma^2)$, let's multiply it by a factor λ. Then the filtered X's will be distributed as $N(\mu\lambda, \sigma^2\lambda^2)$. This means that not only the volatility but also the mean can be affected by the filter. This can have unpredictable effects especially when considering multiple factors which might be negatively correlated with the losses. Moreover, lets not forget that usually our empirical sample $\{X(t-i)\}_{i=1}^N$ is not stationary. This means that using our historical simulation we are actually sampling from different distributions. However, we can make hypothesis about the dynamics of the underlying stochastic process for X.

4.3 The moment hypothesis

For instance one can say that the first k moments of the distribution of $\overline{X(t)}$ can be correctly estimated looking at the EWMA estimators over the last P (chosen time window) samples. This would mean for example that:

$$E[\overline{X(t)}] = \hat{\mu}(t)_P^\lambda = \hat{m}_1(t)_P^\lambda \quad k=1$$

for the mean and

$$Cov[\overline{X(t)}] = \hat{\Sigma}_P^\lambda(t)^2 = \hat{m}_2(t)_P^\lambda \quad k = 2$$

for the variance-covariance matrix. With this hypothesis in mind we can construct a family of filters that transforms the first k moments of the past samples and updates them to be equal to the ones of the current distribution of $\overline{X(t)}$.

4.4 The moment algorithm

The achieve this objective define the filtered $\hat{X}$ as a polynomial function of X.

$$\hat{X} = F(\alpha_0, ..., \alpha_{k-1})(X) = X + \sum_{i=0}^{k-1} \alpha_i X^i$$

Then we set the first k moments of the filtered $\hat{X}$ equal to the ones estimated in the current time window $[t - P, t]$:

$$E[\hat{X}^j] = E[(X + \sum_{i=0}^{k-1} \alpha_i X^i)^j] = \hat{m}_j(t)_P^\lambda \quad \forall j \in \{1, ..., k\}$$

We can express each $E[\hat{X}^j]$ as a linear combination of $\{E[X], ..., E[X^{2k-2}]\}$ which can be estimated by calculating the EWMA estimators for the moments of X. Therefore we end up with a system of k non-linear equations with k free parameters $\{\alpha_0, ..., \alpha_{k-1}\}$. For the case with $k = 2$ we can solve the system analytically and come up with the following filter:

$$\hat{\overline{X}}(t-i) = \hat{\Sigma}_P^\lambda(t) \circ (\hat{\Sigma}_P^\lambda(t-i)^{(-)}(\overline{X}(t-i) - \hat{\mu}(t-i)_P^\lambda)) + \hat{\mu}(t)_P^\lambda$$

We can see how see how similar it is to the classical Covariance Filtering method, the only difference being the mean "de-trending" and "re-trending" in addition to the volatility one.

5 Backtesting

Backtesting for Value at Risk (VaR) is an essential process to evaluate the effectiveness of the models used. VaR backtesting involves comparing VaR estimates derived from a model with actual observed losses over time. This allows us to assess whether the model can provide accurate estimates of loss risk.

To perform backtesting, we need to calculate the following:

- Number of exceptions: the number of times observed losses exceed the VaR estimated by the model.
- Likelihood Ratio: a statistic used to compare the likelihood of the tested model with that of the reference model.

  $$\text{LR} = -2 \ln \frac{L(x|\text{null hypothesis})}{L(x|\text{model to be tested})}$$

  If H0 is correct it follows a chi-square distribution
- P-value: the probability of observing a test statistic value more extreme than the one observed, assuming the null hypothesis is true.

  $$\text{p-value} = 1 - \Phi_{\chi^2}(LR)$$

  If the p-value is less than the significance level, it can be concluded that the model is not acceptable.

There are several VaR backtesting methods, such as the Kupiec and Christoffersen models, each offering different approaches to evaluating the accuracy and reliability of VaR estimates.

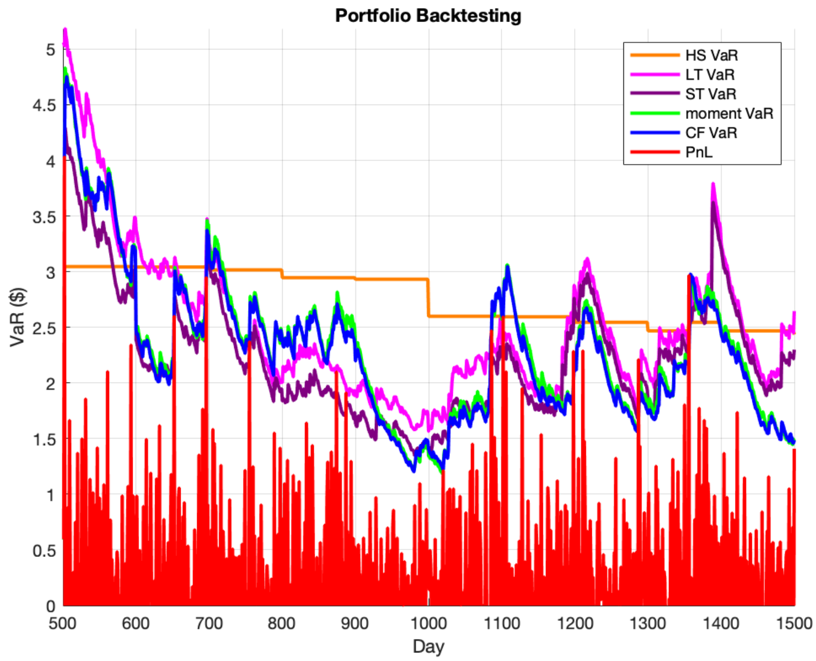


Figure 7: An illustration of model performance for WTI vs. Brent spread portfolio

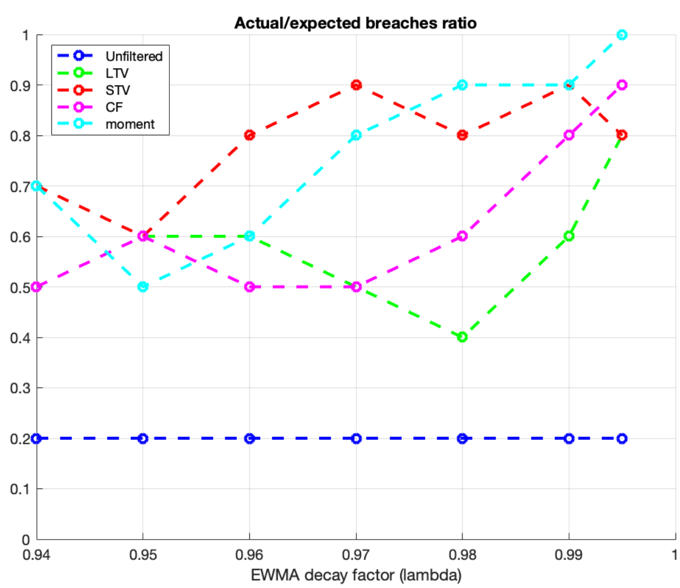


Figure 8: An illustration of the average ratios between observed and expected breaches for different models and for different values of the parameter λ

5.1 Kupiec method

The Kupiec model assumes the null hypothesis that VaR violations follow a Bernoulli distribution with a violation probability equal to the confidence level specified for the VaR itself:

$$L(x|\alpha) = \binom{N}{x} \alpha^x (1-\alpha)^{N-x}$$

where x is the number of exception, and N the number of observation. The testing model instead takes $\alpha = \hat{\alpha} = \frac{x}{N}$.

The likelihood ratio approximately follows a chi-square distribution with a 1 degree of freedom. If the calculated likelihood ratio exceeds 6.63, we can conclude that the model is not acceptable at the 99% confidence level.

Below the result of the backtesting performance of three 99% VaR models using 1.000 days historical data for a low variance portfolio:

	HS	**LTV**	**STV**	**CF**	**moment**
# exceptions	2	5	9	5	8
LR	9.6267	3.0937	0.1045	3.0937	0.4337
p-value	0.0019	0.0786	0.7465	0.0786	0.5102

Table 3: A summary of the backtesting performance

Since we have 1000 days VaR the value of expected exceptions is given by: $\alpha \cdot \text{Noss} = 10$. So, based on our results, we can conclude that the worst method is HS. It has a low number of exceptions, the likelihood ratio (LR) value is above the critical level, indicating a significant deviation from the model predictions, as confirmed by the low p-value. Both the LTV and CF models exhibit good values across the board —number of exceptions, p-value, and LR— underscoring their robustness and reliability as viable methods. The STV and moment model have the lowest LR value, indicating an outstanding theoretical fit. Moreover, they have a number of exceptions similar to the expected one, so they are the best models among those analyzed.

5.2 Christoffersen method

Christoffersen's model, built on the foundation of an elementary Markov chain for exceptions, underscores the significance of consecutive exceptions in subsequent days within financial markets.

t-1	**t**
yes(1)	yes(1)
no(0)	no(0)

The null hypotesis is:

$$L(x|\alpha) = \frac{N!}{N_{00}!\,N_{01}!\,N_{10}!\,N_{11}!} \alpha^{N_{01}+N_{11}} (1-\alpha)^{N_{00}+N_{10}}$$

The model to be tested is:

$$L(x|\hat{\alpha}_{ij}) = \frac{N!}{N_{00}!\,N_{01}!\,N_{10}!\,N_{11}!} (1-\hat{\alpha}_{01})^{N_{00}} \hat{\alpha}_{01}^{N_{01}} (1-\hat{\alpha}_{11})^{N_{10}} \hat{\alpha}_{11}^{N_{11}}$$

where:

$$\begin{cases} N_{00} = \text{no exception in } t-1 \text{ and no exception in } t \\ N_{01} = \text{no exception in } t-1 \text{ and exception in } t \\ N_{10} = \text{exception in } t-1 \text{ and no exception in } t \\ N_{11} = \text{exception in } t-1 \text{ and exception in } t \end{cases} \qquad \begin{cases} \hat{\alpha}_{01} = \frac{N01}{N_{01}+N_{00}} \\ \hat{\alpha}_{00} = 1-\hat{\alpha}_{01} \\ \hat{\alpha}_{11} = \frac{N11}{N_{11}+N_{10}} \\ \hat{\alpha}_{10} = 1-\hat{\alpha}_{11} \end{cases}$$

The likelihood ratio approximately adheres to a chi-square distribution with 2 degrees of freedom. Consequently, the critical value, marking the threshold not to be exceeeded, stands at 9.21.
However, when applied to our portfolio, there are no consecutive days where losses exceed the Value at Risk (VaR). This absence of sequential exceptional events makes the model ineffective in providing actionable insights or forecasts for our specific portfolio dynamics.