Comparative Assessment of Value-at-Risk (VaR) Estimation Methods

Applied Econometrics II

Mr. Philippe De Peretti

Master 1 Econometrics and Statistics

Academic Year 2024–2025

Ali AHMAT SOULEYMAN
Victoria VIDAL - - EUDIER
Romane CRISMAT LE DIRACH

Contents

Abstract

Each model's ability to capture volatility dynamics and predict extreme losses is assessed using Kupiec and Christoffersen backtesting procedures. The GARCH-t model demonstrates the best performance, with strong reactivity to market shocks and robust VaR forecasts, especially at the 99% confidence level. The EWMA model, although based on a Gaussian assumption, shows satisfactory results and practical implementation advantages. In contrast, the Kalman filter model fails to provide accurate risk estimates, revealing significant underestimation and serial dependence in violations.

The study emphasizes the importance of combining statistical rigor and practical relevance in financial risk modeling. The results suggest that while GARCH-t remains the most effective under turbulent conditions, EWMA offers a viable and simple alternative for stable market environments.

This paper focuses on Value-at-Risk (VaR), a widely used measure for quantifying financial risk, although it can be complex to interpret when financial returns deviate from the normal distribution.

We begin by defining the fundamental properties of a risk measure, followed by an empirical evaluation of VaR using backtesting. Our study estimates VaR through three distinct conditional variance models: ARCH-GARCH, EWMA, and SVM (via the Kalman filter).

After computing the VaR estimates, the thesis evaluates their relevance using the Kupiec and Christoffersen tests, the latter being an extension of the former. The empirical analysis is conducted on a self-constructed equally weighted portfolio comprising six major European stock indices over the period 1996–2009.

Introduction

In the context of growing financial market uncertainties, the sharp evaluation of loss risk has become a key requirement for investors, regulators, and financial institutions. Among the most widely used tools for risk measurement, Value-at-Risk (VaR) stands as a benchmark, estimating, at a given confidence level, the maximum potential loss for a portfolio over a specified time horizon. However, VaR is highly sensitive to the quality of the model used to estimate conditional return volatility.

Several studies have explored comparative evaluations of volatility models for VaR computation. Bucevska (2013) highlighted the efficacy of GARCH models in capturing risk dynamics in emerging markets like Macedonia. Similarly, Obeng (2021) demonstrated the relevance of GARCH in the volatile environment of cryptocurrencies. Restrepo (2012) proposed a multivariate extension to better capture asset dependencies. Alternatively, the stochastic Kalman filter approach, as noted by Berardi et al. (2002), offers another estimation path, albeit more technically demanding. Galdi and Pereira (2007) compared several methodologies, concluding that model performance varies with asset class and economic regime.

This thesis compares the effectiveness of three widely applied VaR estimation techniques:

- the GARCH(1,1) model with Student's t-distributed innovations,
- a simplified stochastic volatility model estimated via the Kalman filter,
- and the Exponentially Weighted Moving Average (EWMA) approach.

Our goal is to evaluate these models on an equally weighted portfolio composed of six major European indices (CAC40, DAX, FTSE100, SMI, IBEX35, MDAXI), covering a period that includes significant market turbulence such as the 2008 financial crisis. Model accuracy is tested using rigorous backtesting techniques: the Kupiec test for exception frequency and the Christoffersen test for independence of violations.

The study proceeds in two stages :

1. The **theoretical phase** presents and contrasts the selected VaR estimation methods, discussing parametric and semi-parametric structures, responsiveness to market shifts, ability to capture extreme events, and theoretical limitations.
2. The **empirical phase** applies these methods to real financial data using SAS software. Custom scripts are developed for each model to compute the VaR series. The models are then evaluated using the aforementioned backtesting tools, and additional robustness checks are performed to examine sensitivity under varying assumptions.

The empirical findings reveal that the GARCH-t model yields the most statistically accurate VaR forecasts, particularly during high-volatility periods. The EWMA approach remains a practical and reliable solution under stable conditions. Meanwhile, although conceptually rich, the Kalman filter model exhibits weaknesses in backtesting performance, limiting its practical reliability in the tested context.

1. Literature Review

1.1 Value-at-Risk : Conceptual Framework and History

Value-at-Risk (VaR) is the most commonly applied method of market risk assessment, under which a firm estimates the maximum loss possible on its portfolio within specified confidence limits over a given time horizon. The 1990s popularize VaR— more particularly through J.P. Morgan's RiskMetrics project— and with it coming to be accepted as the standard within the financial industry, it has since been adopted into international prudential recommendations, most notably in accords Basel I and Basel II to legitimize application of VaR as a methodology for determining capital requirements by financial institutions. This brought about further advancement towards standardizing market risk measurement worldwide with VaR.

A wide application of VaR in different economic and geographic contexts has been demonstrated by various studies. Jorion (2011) points to its extensive proliferation within the finance field. Empirical research has tested its use in Southeast Asia (Cheong et al., 2011), the European Union (Iglesias, 2015), Latin America (Ozun and Cifter, 2007), Nordic countries (Jobayed, 2017), South Africa (Naradh et al., 2021).

In the period 2000–2012, VaR was analyzed by Iglesias (2015) based on the leading stock indices of eleven EU member countries. GJR-GARCH models (Glosten et al., 1993) were used to model returns. These belong to the family of GARCH (Generalized Autoregressive Conditional Heteroskedasticity), which enables one to model the conditional volatility of financial returns. The GJR-GARCH model particularly introduces an asymmetry effect whereby one can differentiate the impact of positive and negative shocks on volatility—a very important consideration in times of financial turbulence.

Ozun and Cifter (2007), for their part, combined a dynamic conditional copula with the EWMA (Exponentially Weighted Moving Average) approach—a method that assigns greater weight to more recent observations to better capture changes in volatility. This combination enabled a more refined assessment of asset dependence and market dynamics in Latin American countries.

Subbiah and Fabozzi (2016) tested different VaR methods in predicting extreme losses in financial markets of Asia, regarding which method gives better forecasts. Their results point to big differences in the models, especially on how well they adapt to the volatility and asymmetry of the Asian markets.

Choi and Min (2011) compared implied methods which use the prices of derivative instruments like options with unconditional methods based purely on historical return data. They made this comparison to determine the best methodologies to predict market risk in Asia.

Finally, Desheng and Chatpailin (2019) applied two methods of estimating VaR to BRICS markets: the delta-normal method—based on the assumption that returns are normally distributed and a linear approximation of the changes in prices—and historical method which applies past returns directly to construct an empirical distribution of potential losses. Backtesting techniques were used to evaluate the accuracy of these estimations. The frequency of exceedances was measured by Kupiec's test, and Christoffersen's test assessed their independence over time.

1.2 Classical Parametric Models

The parametric approach usually assumes that financial returns are distributed according to the normal distribution. In other words, it is assumed that a Gaussian distribution with a constant mean and variance, and perfect symmetry around the mean describes the behavior of returns. This assumption allows an easy analytical calculation of Value-at-Risk especially by using the variance-covariance method in which standard statistics of the portfolio are used to calculate VaR.

This structure fits rather well in placid market conditions, where swings are moderate, assets are loosely linked, and the macro environment is stable. These scenarios happen mostly in very diversified portfolios during periods of growth - or in some developed markets with high volumes and strong regulation.

Many empirical studies have proved that financial market returns do not follow a normal distribution, particularly during crises or when uncertainty is high, and thus concerning our portfolio evolution among time, its distributions is showing heavier tails than the ones of data following a normal distributions. A very commonly observed phenomenon is leptokurtosis that refers to the presence of fat tails in the return distribution indicating that extreme events happen with a probability larger than that predicted by the normal distribution. Other common departures are skewness, meaning that returns are asymmetrically distributed and that losses dominate gains of equal magnitude. Such dynamics become much more pronounced in periods of financial distress as was the case in the subprime crisis August 2008 or COVID-19 crisis March 2020.

Another point is that in reality, the assumption of constant variance is typically violated: financial returns volatility varies over time—calm and turbulent dynamic states. Such instability renders static parametric approaches less meaningful since they cannot capture the effect of leverage whereby volatility reacts more strongly to negative shocks than to positive ones as well as the conditional heteroskedasticity—variance depending on its previous values and previous shocks.

Thus, classical parametric models will be simple and easy to implement but relevant only in the conditions of stability in market dynamics. In more robust settings, dynamic models from the GARCH family introduced by Tim Bollerslev in 1986 or empirical approaches based on historical data will suit better.

1.3 Conditional Volatility Models: EWMA, GARCH

To better reflect the swings in money markets, a few conditional waltz of volatility have been suggested. Old school constant-variance methods do not allow for such flux; modern ones let volatility change with time and situation by situation based on recent news and bygone happenings. In this way, they can pre-empt periods when uncertainty is very or not so high—a critical component in the proper evaluation of where a potential loss might land regarding a portfolio.

The EWMA model came into J.P. Morgan's RiskMetrics project during the mid-1990s. It is believed that a principle under which recent performance should carry more weight than older observations in determining present volatility.

Though easy to implement and run, the EWMA model has some limitations. It doesn't explicitly model dynamic relationships between shocks and variance; neither does it account for asymmetries in returns observed quite often in practice.

To deal with these drawbacks, the GARCH family of models—Generalized Autoregressive Conditional Heteroskedasticity—came into being. The GARCH model proposed by Tim Bollerslev in 1986 is basically an improvement over the ARCH model originally introduced by Robert Engle in 1982. These models give conditional variance as a cobination of lagged variance and the lagged squared shocks. They capture volatility persistence—that is, the very important feature whereby after a shock, markets remain volatile for several periods.

In the standard GARCH model, positive and negative volatility shocks are assumed to be symmetric. However, most of the times negative news is more informative than positive news hence it imparts greater volatility. The leverage effect can be captured by applying the GJR-GARCH model-Glosten, Jagannathan, and Runkle adding a term that raises the response of variance to negative returns.

These conditional volatility models are typically used in the estimation of Value-at-Risk since they articulate risk a bit more accurately, especially in periods of crisis. Together with distributions that articulate financial returns better—such as the Student's t-distribution or asymmetric distributions—they produce more accurate estimates of VaR, particularly in the tails of the distribution where extreme losses occur.

Theoretical Part: Foundations and Methods for VaR Estimation

1. Definitions and Role of VaR

1.1 Basic Principles

Value-at-Risk (VaR) measures the maximum loss that a portfolio of money assets can sustain within specified confidence limits and time periods. (Value-at-Risk is used to assess the market risk associated with a portfolio. A portfolio is defined here as a set of positions, each linked to an underlying asset.) It answers the following question: what is the maximum loss I am likely to incur, with a probability of $x\,\%$, over a defined period of time?

Value-at-Risk (VaR) of a portfolio P at a confidence level π over a horizon Δt, can be defined mathematically as :

$$\Pr[\Delta P_t \leq -\text{VaR}_t] = \pi \tag{1}$$

Here, the probability that the loss exceeds the VaR is $1 - \pi$. In other words, if F is the return distribution function of P,

$$\text{VaR}_t = F^{-1}(\pi) \tag{2}$$

Economically, it may be said that VaR_t represents the maximum loss that can be incurred over the time horizon Δt with a probability of $1 - \pi$. Intuitively, Value-at-Risk represents the maximum loss that can be incurred on a portfolio, with a certain level of confidence, and over a given time horizon. For example, a VaR of £1 million at 99% implies that the loss will exceed this amount over the period in question with only 1% probability.

This indicator is defined by three essential dimensions:

- A confidence level, often set at 95% or 99%;
- A time horizon, which can be one day, ten days or more, according to the needs of the analysis;
- A potential loss expressed in absolute value (euros, dollars, etc.).

Market risk exposure is measured and controlled at financial institutions through use of VaR, which has now become an easy-to-adopt measure among these institutions. This is because the definition is based on assumptions about how returns are distributed—an issue that will be explored in much greater detail in the following sections.

1.2 Regulatory Framework

International banking regulations have a deep impact on the practice of risk management. Mostly they take effect through changes that increase Value-at-Risk's dominant role in capital requirement calculation. These changes flow from three major international accords originating with the Basel Committee on Banking Supervision and designed to enhance global financial stability.

Basel I (1988) was the first international effort on bank safety with a capital ratio of 8% on risk assets. Only in 1996 was a big change to include market risk in capital needs. This allowed use of firms' internal risk models, including VaR, but they must be checked by regulators.

Basel II (2004) adopted this framework and affirmed VaR to be the benchmark method for the assessment of market risk in Pillar 1. It also extended the application of internal models to credit risk and brought operational risk into the computation of regulatory capital for the first time. The accord is based on three pillars: minimum requirements, enhanced supervision, and market discipline via increased transparency.

Basel III (2010), after the 2007–2008 financial crisis, deals with making the banking system more resilient. It makes capital requirement quality and quantity increase, alongside new liquidity and leverage ratios. Market risk continues to be measured by VaR but Expected Shortfall is slowly taking over, as it is supposed to be more relevant for extreme risks.

2. Classical Methods of VaR Estimation

2.1 Parametric Approach (Variance-Covariance)

The parametric approach is thought to be a very simple way or an analytical VaR or even a normal VaR. It assumes that financial returns follow normal distribution and volatility can be estimated dependably from previous data.

In such cases, a close expression for Value-at-Risk can be derived assumed as the product of quantile normal distribution standard deviation of returns and risk exposure This method enables a rapid Value at Risk calculation without simulation and is, therefore, most applicable to portfolios comprising linear securities or when returns are assumed relatively symmetrical. It is based on the formula for variance-covariance which uses the covariance matrix of assets and their weightings in the portfolio.

This method was popularized via the RiskMetrics model by J.P. Morgan. in the 1990s. This framework assumes normal returns and presents a dynamic calculation of the volatility of returns using an exponentially weighted moving average. The EWMA model is described later in detail in a dedicated section.

Should financial returns prove to follow a normal distribution, a more operational expression for VaR can be derived. The simple form is also termed parametric VaR. If returns follow a normal distribution, then VaR can be written as:

$$\mathrm{VaR}_t = z_{1-\pi} \cdot \sigma \cdot \text{exposure} \tag{3}$$

where:

- $z_{1-\pi}$ is the quantile of the standard normal distribution corresponding to the confidence level (e.g. $z_{0.95} = 1.65$);
- σ is the standard deviation of the portfolio returns over the horizon under consideration;
- exposure refers to the market value of the position.

Although this approach is particularly simple to implement and inexpensive in terms of computational resources, it is based on strong assumptions, which can be summarised as follows:

- **Normality of returns:** financial returns are assumed to have a normal distribution, i.e. $r_t \sim \mathcal{N}(\mu, \sigma^2)$. This particularly implies that:

$$E[r_t] = \mu,\ \forall t, \quad \text{and} \quad \mathrm{Cov}(r_t, r_{t-k}) = 0,\ \forall k \neq 0.$$

  In simpler words, no autocorrelation between returns and they are identically distributed in each period.

- **Independence and homogeneity:** the sequence of returns (r_t) is assumed to be independent and equally spread out, which is expressed as $r_t \overset{\text{i.i.d.}}{\sim} \mathcal{N}(\mu, \sigma^2)$.

- **Parameter stability:** the moments (mean and variance) are constant over time, i.e. $\mu_t = \mu$ and $\sigma_t^2 = \sigma^2$, $\forall t$.

The scraped content talks about the use of assumptions in money calculations noting that while they can make things easier their truth is often doubted especially in times of money worries. The paper tries to closely look at these assumptions through comparing parts. To show how the method works in real life two instances are given one based on old model guesses and another linked to the setup of the RiskMetrics model.

Application Example of VaR Using the RiskMetrics Model

Consider a firm with a position of 140 million German marks (DEM), using USD as its reference currency. To assess market risk, we convert this to USD using the exchange rate:

$$\text{Exposure} = \frac{140\,\text{million}}{1.40} = 100\,\text{million USD}$$

Assuming daily volatility of DEM/USD is 0.565%, and using the 95% quantile of the standard normal distribution ($z = 1.65$), the one-day VaR is:

$$\text{VaR}_t = 1.65 \times 0.00565 \times 100\,000\,000 = 932\,250\,\text{USD}$$

This means there is a 5% chance that the loss will exceed $932,250 in one day.

Now consider a second example using the RiskMetrics framework for a $500 million portfolio. The expected return is zero, and daily volatility is $\sigma = 0.0321$. The 5% quantile return is:

$$\hat{r} = -1.65 \times 0.0321 = -0.0520$$

Thus, the adverse scenario value becomes:

$$V_1 = 500 \times e^{-0.0520} \approx 474.2\,\text{million}$$

Leading to:

$$\text{VaR} = 500 - 474.2 = 25.8\,\text{million USD}$$

This implies a 5% probability of losing more than $25.8 million in one day. While VaR provides a useful single-value risk measure, it does not capture the full shape of the loss distribution—particularly extreme events—which is why it should be complemented by other risk indicators.

2.2 Historical Simulation

2.2.1 General Principle

Historical simulation has wide applications for estimating Value-at-Risk. It relies on the direct use of time series of daily variations in market risk factors. It does not need any assumptions about the shape of the return distribution, unlike parametric approaches.

These returns are observed over a representative period and applied to the portfolio as if they were the returns on the current positions. By repeating this process for every historical observation, we get a series of scenarios that mimic possible losses and gains, thus helping us build an empirical distribution of portfolio variations.

2.2.2 Methodology

A typical approach involves four main steps:

1. **Collection of historical data**
   The historical returns of the assets that make up the portfolio are gathered over a specific period, typically several hundred days.
2. **Portfolio Variation Simulation**
   Every historical observation represents a scenario where the observed returns are applied to the positions held. A potential variation of the portfolio's value is then inferred.
3. **Empirical Distribution Construction**
   The simulated variations are sorted so as to build an empirical distribution of possible losses.
4. **Determining VaR**
   Identifying the quantile that relates to the desired confidence level leads to estimating the maximum potential loss.

2.2.3 Accuracy of Quantile Estimation

The estimation of Value-at-Risk by historical simulation is based on a sample of past returns. A formula for evaluating the standard error associated with an empirical quantile q calculated from a sample of size n was proposed by Kendall and Stuart (1972):

$$\text{Standard error} = \frac{1}{f(x)} \cdot \sqrt{\frac{q(1-q)}{n}} \tag{10}$$

where:

- $f(x)$ is the probability density function around the quantile x,
- q is the target quantile level,

- n is the number of observations.

Therefore, the greater the number of observations, the more stable the quantile estimate. High density around the target quantile also improves the accuracy of the estimate.

On the flip side, when the quantile falls in a low-density region, uncertainty begins to grow. These considerations are especially important for high confidence levels in VaR estimation, such as 99% or 99.9%, where data is sparse in the tail of the distribution. This might justify the use of complementary methods—such as the extreme value theory—to enhance precision in those critical regions.

2.3 Monte Carlo Simulation

2.3.1 Theoretical Framework

Monte Carlo simulation is a random method very good at estimating Value-at-Risk (VaR) when traditional value approaches do not work. Unlike parametric and historical methods that say there are straight relationships between the value of the portfolio and risk factors, Monte Carlo lets you make for very hard, not linear connections.

By making many random scenarios from given or thought of market variable distribution, it allows the real building of possible loss distributions-very good for when you have those types of portfolios that include derivatives, options, or structured products and where there are big effects because things are not linear and have convexity.

2.3.2 Methodology

The Monte Carlo simulation is implemented based on systematic steps as listed below:

1. Assess the mark-to-market value of the current portfolio using available market data.
2. Generate a random sample of realizations from a multivariate probability distribution that describes potential future changes in market factors.
3. Apply these simulated values to estimate market conditions at the end of the period.
4. Reassess the portfolio with this new data.
5. Calculate the simulated loss as the difference between the initial value of the portfolio and the final value.
6. Repeat the previous steps a large number of times to get an empirical distribution of variations in the portfolio.

The VaR results from this estimation as the α-quantile of that distribution; for example, if 5,000 simulations are run, the 99% confidence level VaR will be the 50^{th} worst loss observed.

Sometimes, a variation called *partial simulation* can be applied to reduce the computational cost. It implies that the portfolio is not recomputed at each iteration but the changes are linearized locally around a reference point.

2.3.3 Copula Extension of Monte Carlo Simulation

The copula version of Monte Carlo simulation allows for modeling complex dependency structures and non-normal marginal distributions, thereby liberating the approach from the rather constraining assumption of normal returns. Through copulas, for instance, Gaussian copulas the dependence between financial assets can be simulated more appropriately, especially under market stress. This methodology involves generating correlated samples and converting them into the required marginal distributions that would enhance the simulation ability to mirror extreme events and co-movements in financial markets.

3. Volatility Models

3.1 Volatility Estimation for Risk Measurement

Volatility estimation is an integral part of financial risk measurement because it represents the uncertainty of the returns at a particular time. This value then feeds into the calculation of possible losses over a specified time horizon and confidence level, as is done by Value-at-Risk (VaR).

Volatility models do not forecast VaR directly—they underpin it. Among the most commonly applied are the Exponentially Weighted Moving Average (EWMA) and conditional heteroscedasticity models such as ARCH and GARCH. Each assumes a particular structure for how variance evolves over time.

3.2 Volatility Estimation Using Simple Moving Averages

An approach to volatility estimation by using simple moving averages is the most basic method. It is the arithmetic mean of the squares of past returns over a fixed-size rolling window. The method assumes temporal homogeneity of returns by giving equal weight to each observation within the analysis period.

For a window of size T, the variance is calculated using the following formula:

$$\sigma^2 = \frac{1}{T}\sum_{t=1}^{T}(r_t - \bar{r})^2 \tag{11}$$

where r_t is the return at time t and $\bar{r}$ is the average return over the period under consideration.

3.3 EWMA Model

The EWMA model (*Exponentially Weighted Moving Average*) is a special case of GARCH models. It implies parameters alpha and beta are set to 1. The latter implies future volatility will always be a constant of current estimated volatility, and it ignores any impact on long run variance such that we have the relation: While traditional volatility forecasting method relies on moving averages that are equals, in other words fixed, EWMA method is an improvement of this approach. It sets the dynamic feature of volatility as an exponential moving average of historical observations.

The EWMA model (Exponentially Weighted Moving Average) is a method for estimating conditional volatility in financial time series. It relies on an exponentially decreasing weighting of past observations, assigning greater weight to recent returns while preserving memory of past movements.

This model is particularly useful in financial environments where volatility evolves over time, which cannot be captured under a constant variance assumption. Unlike the simple moving average (SMA), the EWMA model introduces a dynamic mechanism for updating the variance.

$$\sigma_t^2 = \lambda\sigma_{t-1}^2 + (1-\lambda)r_{t-1}^2 \tag{12}$$

where :

- σ_t^2 is the conditional variance estimated at time t,
- r_{t-1} is the return observed at time $t-1$,
- $\lambda \in (0,1)$ is the smoothing (or memory) parameter, determining the weight given to recent observations.

This is a weighted average of the variance that has been estimated already and the squared most recent return. Since this recursive scheme means that just the latest values for variance and return are required to update the estimate, there is no need to keep all past data. This makes the algorithm particularly efficient. Every old observation gets a weight of λ^{i-1}; therefore, weights fade away with an exponential factor as we move back in time. This property captures the idea that for predicting future volatility, recent events are much more informative than events from way back in time.

The choice of the λ parameter has a strong influence on the model's behavior: Each past observation is thus weighted by a factor λ^{i-1}, which implies that the weights decrease exponentially over time. This feature reflects the assumption that recent events are more informative for anticipating future volatility than older ones. The value of λ significantly impacts the dynamics of the model:

When λ is near 1, say 0.94 as employed in the RiskMetrics model, the model keeps a lengthy memory of past data and generates smoother volatility estimates. Conversely, a lower value like 0.7 makes the model more responsive to adjust more rapidly to recent market changes.

4. From ARCH to GARCH

4.1 General Principle

The ARCH model assumes that the conditional variance at time t depends only on past shocks. Its formulation is based on an autoregressive structure:

$$\varepsilon_t^2 = \alpha_0 + \alpha_1\varepsilon_{t-1}^2 + \alpha_2\varepsilon_{t-2}^2 + \cdots + \alpha_q\varepsilon_{t-q}^2 + v_t$$

where ε_t is the estimated residual from the model $y_t = \alpha_0 + \Phi y_{t-1} + \varepsilon_t$ and v_t is white noise. This specification has practical limitations, as it relies directly on past squared residuals.

A multiplicative form can be introduced to enable joint estimation of $\{y_t\}$ and the conditional variance:

$$\varepsilon_t = v_t \sqrt{\alpha_0 + \sum_{\mu=1}^{q} \alpha_\mu \varepsilon_{t-\mu}^2} \tag{15}$$

Subject to $\alpha_0 > 0$, $\alpha_\mu \geq 0$, and $0 \leq \sum \alpha_\mu \leq 1$, the expression of the long-run conditional variance becomes:

$$\sigma_t^2 = \frac{\alpha_0}{1 - \sum \alpha_\mu} \tag{16}$$

ensuring its positiveness and preventing explosive behavior.

4.2.1 Formulation of the GARCH Model

The GARCH model generalizes the ARCH model by incorporating not only past innovations but also lagged values of the conditional variance. It is expressed as:

$$\varepsilon_t = v_t \sqrt{h_t}, \quad \text{where} \quad E[v_t^2] = 1$$

and

$$h_t = \alpha_0 + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^{p} \beta_j h_{t-j} \tag{17}$$

where:

- h_t is the conditional variance at time t,
- $\alpha_0 > 0$, $\alpha_i \geq 0$, $\beta_j \geq 0$,

The condition $\sum \alpha_i + \sum \beta_j < 1$ ensures the stationarity of the process. Thus, the GARCH model can be seen as an ARMA process applied to the conditional variance. These models are conditionally heteroskedastic with a constant unconditional variance.

4.2.2 The GARCH(1,1) Model

The most commonly used variant in practice is the GARCH(1,1) model due to its simplicity and empirical robustness. It is defined as:

$$h_t = \omega + \alpha \varepsilon_{t-1}^2 + \beta h_{t-1} \tag{18}$$

This model imposes the constraint $\alpha + \beta < 1$ to guarantee stationarity. The closer the sum $\alpha + \beta$ is to 1, the slower the mean reversion of volatility toward its long-term average. Conversely, a smaller sum implies a quicker adaptation to new market information.

4.2.3 Link EWMA/GARCH models

The EWMA model may be viewed as a particular case of the GARCH(1,1) model. If we set $\omega = 0$, $\alpha = 1 - \lambda$, and $\beta = \lambda$, we obtain the structure of the exponential weighting model. Therefore, the GARCH(1,1) model generalizes the EWMA model by introducing a constant term and allowing empirical estimation of the parameters from the data.

In practice, when the value of the parameter ω is zero, the GARCH(1,1) behaves similarly to an EWMA model. However, if ω is negative, the model becomes unstable, which justifies preferring an EWMA specification in such cases. This structural connection between the two approaches helps understand the relevance of choosing between them based on the empirical characteristics of the time series.

In most theoretical applications, the GARCH(1,1) model is considered more powerful than the EWMA model because it incorporates mean reversion in variance—a feature that better reflects empirical behavior in financial markets. However, when $\omega = 0$, the GARCH(1,1) essentially reduces to an EWMA. The choice between the two depends on several factors: desired stability, acceptable model complexity, required accuracy, and the nature of volatility in the data.

4.2.4 Specification and Estimation

For the operational use of a GARCH model, the errors ε_t must be specified with a conditional distribution. In practice, several distributions can be considered:

1. Normal distribution,
2. Student's t-distribution with degrees of freedom ν,
3. Generalized error distributions.

The parameters of the GARCH model are usually estimated using the maximum likelihood method.

Estimation Assuming Normality
For a GARCH(1,1) model with T observations and normally distributed errors, the log-likelihood function is:

$$\log L = -\frac{1}{2}\log(2\pi) - \sum_{t=1}^{T}\left[\frac{1}{2}\log(h_t) + \frac{1}{2}\frac{(y_t - \alpha_0 - \varphi y_{t-1})^2}{h_t}\right] \tag{18}$$

Estimation Under the Student's t Hypothesis
With a Student's t-distribution having ν degrees of freedom, the log-likelihood is:

$$\log L = \frac{1}{2}\log\left(\frac{\pi(\nu-2)}{\nu}\right) - \log\left(\frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)}\right) - \sum_{t=1}^{T}\left[\log(h_t) + \frac{\nu+1}{2}\log\left(1 + \frac{(y_t - \alpha_0 - \varphi y_{t-1})^2}{h_t(\nu-2)}\right)\right] \tag{19}$$

where $\Gamma(x)$ denotes the gamma function, defined by:

$$\Gamma(x) = \int_0^\infty y^{x-1} e^{-y} \, dy \tag{20}$$

The choice of error distribution directly influences the robustness and accuracy of the estimates, particularly in the presence of extreme values.

5. The Kalman Filter

The Kalman filter allows us to extract the state variable we are interested in — the volatility — in order to model the returns of our portfolio with a stochastic volatility model.

The idea behind this filter is to update the model's parameters using a "forecast-upgrade" system: at each step, the model uses the measurement equation and the prediction error to adjust the state equation.

This makes it possible to recover the component we are looking for, even if it is latent and unobservable. The Kalman gain is a variable that tells us how much we should trust the new observation compared to our model prediction.

In other words, it adjusts how much the estimate of the volatility should change after we observe new data. If the observation is very noisy, the Kalman gain will be small, so the model will rely more on the past estimate. If the observation seems reliable, the Kalman gain will be higher, and the new data will have more influence on the updated volatility estimate.

Kalman filter is indeed an algorithms that recursively update an estimate of the state and find the innovations driving a stochastic process given a sequence of observations. It is called a stand-space model, used at first for rocket trajectories. We use Kalman filter in our study to modélise non linear evolution of the variance and to estimate stochastic variance of our portfolio.

In the Black and Scholes model stock price volatility is assumed to be constant, however studies have shown that return volatility is conditional to the information set available at the computation time. This introduces the use of the Kalman filter in order to estimate the stochastic volatility of our portfolio and to properly estimate our VAR.

5.1 State-Space Representation

The Kalman filter is based on a linear state-space model that describes the evolution of a hidden system state over time. The system dynamics are given by:

$$x_{k+1} = Ax_k + w_k$$

where:

- x_k is the hidden state vector at time k,
- A is the state transition matrix,
- w_k is the process noise, assumed to be Gaussian with zero mean and covariance matrix Q.

The measurement model, which links the state to the observed measurement z_k, is:

$$z_k = Hx_k + v_k$$

where:

- H is the observation matrix,
- v_k is the measurement noise, also assumed to be Gaussian with zero mean and covariance R.

5.2 Recursive Filtering: Prediction and Update

At each time step, the Kalman filter operates in two main stages:

Prediction Step

$$\hat{x}_k^- = A\hat{x}_{k-1}, \quad P_k^- = AP_{k-1}A^\top + Q$$

where $\hat{x}_k^-$ is the predicted state estimate and P_k^- is the predicted error covariance.

Update Step

Using the measurement z_k, the filter updates the predicted values:

$$K_k = P_k^- H^\top (HP_k^- H^\top + R)^{-1}$$

$$\hat{x}_k = \hat{x}_k^- + K_k(z_k - H\hat{x}_k^-)$$

$$P_k = (I - K_k H)P_k^-$$

where K_k is the Kalman gain, determining how much the prediction should be corrected using the measurement.

5.3 Recursive Cycle

After the update, the corrected state estimate and its covariance are projected to the next time step:

$$\hat{x}_{k+1}^- = A\hat{x}_k, \quad P_{k+1}^- = AP_kA^\top + Q$$

5.4 Statistical Interpretation

The Kalman filter minimizes the mean squared error (MSE) of the state estimates under Gaussian noise assumptions; it is equivalent to a recursive least squares estimator. It can also be viewed within a maximum likelihood framework, providing strong statistical justification for its use.

Additionally, the merit function associated with the Kalman filter has the same mathematical structure as a χ^2 test statistic, underlining its significant utility in problems involving data adjustment and the tracking of noisy observations.

Empirical Part: Application of Estimation Methods

1. Portfolio Construction and Descriptive Analysis

In this study, we construct an equally weighted portfolio based on six major European stock indices: the CAC40 (France), the DAX (Germany), the SMI (Switzerland), the FTSE100 (United Kingdom), the IBEX35 (Spain), and the MDAXI (Germany). Daily price data were imported via SAS macros from Excel files. A sorting and renaming step was used to harmonize the price columns.

To ensure the statistical integrity of the series, all observations containing missing values were removed, reducing the sample to 3,119 valid observations covering the period from March 1, 1996, to December 30, 2008.

The transformation of prices into logarithmic returns was then applied to each index:

$$r_t = \ln\left(\frac{P_t}{P_{t-1}}\right)$$

where r_t is the return at time t, and P_t is the price at time t.

The portfolio return is then obtained by simple averaging (equal weighting) of the six individual returns:

$$r_t^{\text{portfolio}} = \frac{1}{6}\sum_{i=1}^{6} r_t^i$$

This approach reflects a non-optimized portfolio, representing a diversified investor with no preconceived expectations about the future performance of the indices.

1.1 Temporal Evolution of the Portfolio's Log Return

The resulting series exhibits clear visual heteroscedasticity: volatility remains relatively stable until 2007, after which there is a sharp increase in variance during 2008, consistent with the global financial crisis. This behavior highlights the necessity of applying conditional volatility models in subsequent analysis.

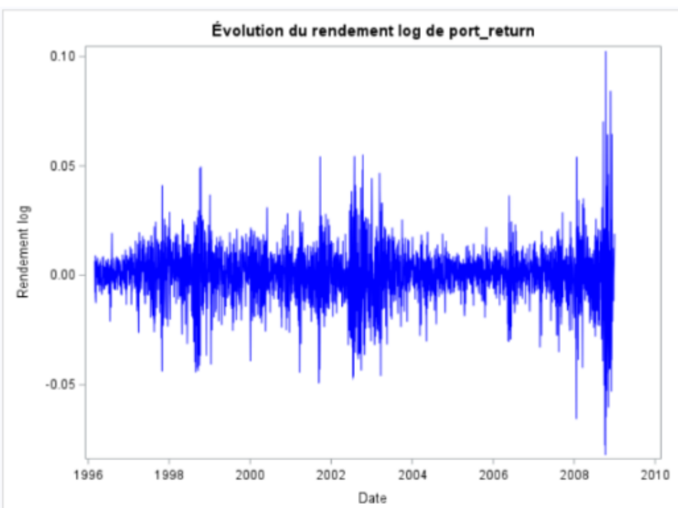

Figure 1: Temporal evolution of the log returns for the portfolio

2. Exploratory Analysis of Returns

Preliminary statistical analysis of portfolio returns was conducted using the `UNIVARIATE` procedure in SAS.

The average return is low and close to zero (0.018%), which is expected for daily returns. The standard deviation is 1.276%, reflecting moderate daily volatility, albeit with variation over time. The negative skewness (-0.165) indicates a slight leftward asymmetry, while the high kurtosis (5.74) suggests the presence of fat tails and frequent extreme events—contradicting the assumption of normality.

Empirical quantiles show wide dispersion: the worst 5% of returns fall below -2.02%, while the best 5% exceed $+1.81\%$. Extremes are particularly pronounced, with a minimum return of -8.2% and a maximum of $+10.3\%$.

The empirical distribution, when compared to the fitted normal curve, confirms these findings: it is more peaked in the center and exhibits significantly heavier tails. This justifies the consideration of alternative distribution laws, such as the Student-t, in Value-at-Risk models.

Finally, normality tests (Kolmogorov–Smirnov, Anderson–Darling, and Cramér–von Mises) strongly reject the normality hypothesis at the 1% level (p-value < 0.01), as illustrated. This statistical evidence highlights that methods based on normal distribution assumptions are ill-suited for modeling extreme financial risk.

3. Autocorrelation and Temporal Dependence Tests

To examine temporal dependence within the portfolio return series, we analyzed the autocorrelation structure of the raw returns and their squared values. This step is crucial for diagnosing the potential presence of conditional heteroscedasticity—commonly referred to as the ARCH effect—which is a key consideration in the selection of GARCH-type models.

Significant autocorrelation in squared returns, even when raw returns show little or no autocorrelation, typically indicates volatility clustering and time-varying conditional variance. These patterns justify the use of models that capture the dynamic behavior of volatility over time.

The simple autocorrelation (ACF) and partial autocorrelation (PACF) plots for the portfolio's log returns indicate no significant autocorrelation across all lags (up to 20). This property is typical of daily financial time series and suggests that the returns themselves are poorly predictable using standard ARMA models.

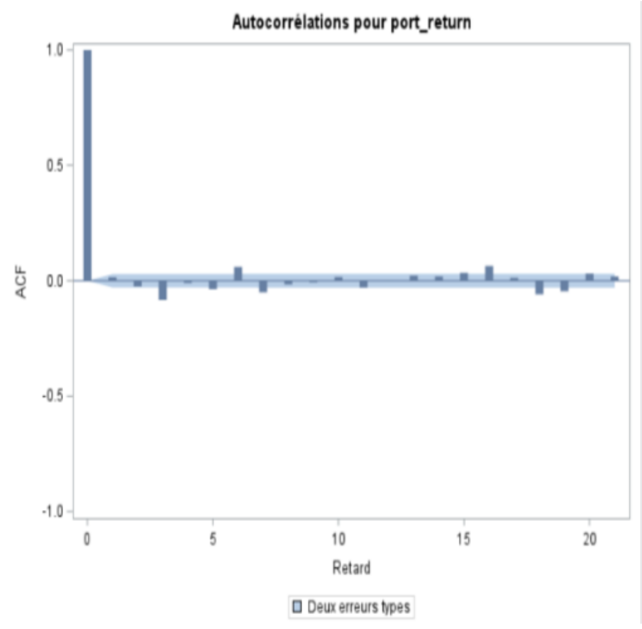

Figure 2: Autocorrelation of Returns

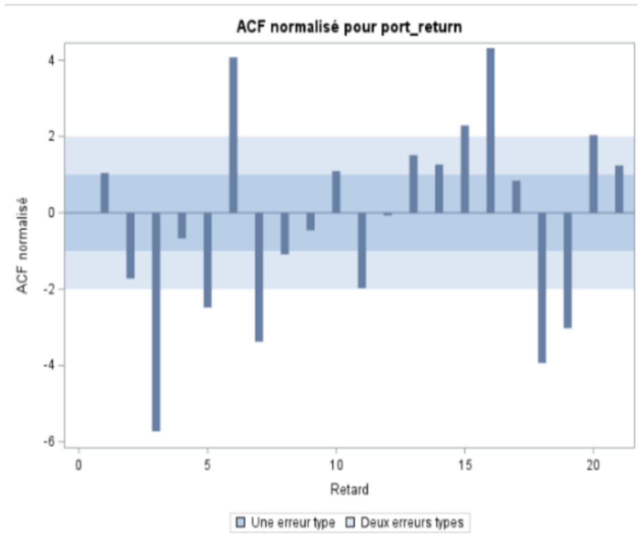

Figure 3: Autocorrelation of Squared Returns

3.1 Analysis of the Squares of Returns

In contrast, the squared returns, r_t^2, exhibit a significant autocorrelation structure across multiple lags. Figure 3 highlights a persistent second-order dependence, which is characteristic of conditional volatility. This behavior supports the presence of ARCH effects and motivates the application of GARCH-type models in the volatility modeling process.

La procédure ARIMA

Nom de la variable = port_sq	
Moyenne des séries de travail	0.000163
Ecart-type	0.000452
Nombre d'observations	3119

Vérification de l'autocorrélation pour le bruit blanc

Jusqu'au retard	Khi-2	DDL	Pr > khi-2	Autocorrélations					
6	1576.16	6	<.0001	0.254	0.324	0.295	0.264	0.368	0.206
12	2586.36	12	<.0001	0.234	0.184	0.249	0.263	0.206	0.246
18	3407.97	18	<.0001	0.203	0.179	0.210	0.240	0.175	0.238

Figure 4: Ljung–Box test results on squared returns using the ARIMA procedure

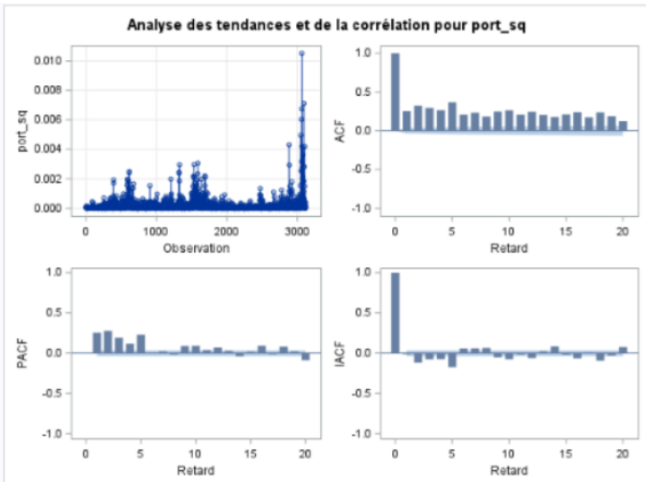

Figure 5: Trend and autocorrelation analysis of squared portfolio returns

The `ARIMA` procedure applied to the squared returns confirms this diagnosis. The Ljung–Box (portmanteau) tests show that the residual autocorrelations are significant up to 18 lags (p-value < 0.0001 for all lags tested). This suggests that the series exhibits a residual ARCH effect, which is a strong indicator of non-constant variance.

The Engle test (ARCH LM test), implemented using the `AUTOREG` procedure, confirms the presence of a significant ARCH effect. Rejection of the null hypothesis of homoscedasticity leads us to favor conditional variance models, such as GARCH.

3.2 Interim Conclusion

The results obtained strongly support the hypothesis of a variable volatility generating process—characterized by independence of raw returns but strong dependence in squared returns. This empirical behavior justifies the use of a GARCH(1,1) model, which will be estimated and tested in the following section.

4. Estimation of the GARCH(1,1) Model with Student's *t*-Distribution

Following the conditional heteroscedasticity diagnostics, a GARCH(1,1) model was estimated on the series of daily portfolio returns. This model captures both the ARCH effect (recent shocks on variance) and the GARCH effect (volatility persistence), according to the following specification.

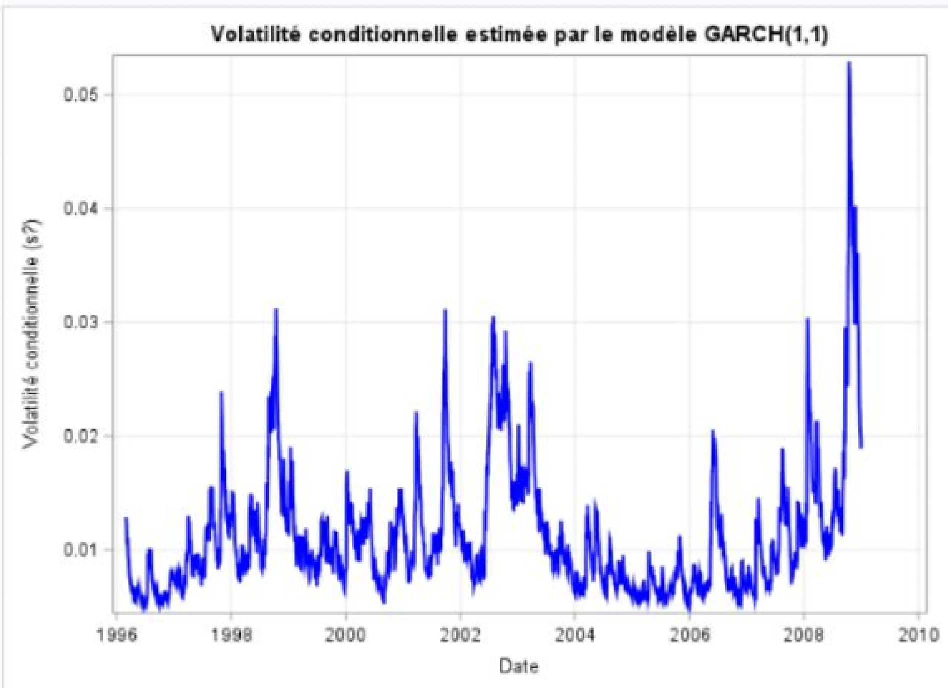

Figure 6: Conditional volatility estimated using a GARCH(1,1) model with Student's t-distributed errors, highlighting volatility clustering and persistence in daily portfolio returns

The choice of a Student's *t*-distribution for the innovations accounts for the previously observed leptokurtosis, as it places greater weight on extreme events than the normal distribution.

The `AUTOREG` procedure in SAS was used to estimate this model. The estimated coefficients α, β, ω, and the degrees of freedom $\nu = 11.76$ confirm that the return distribution has fat tails, but its statistical moments remain well-defined.

There is a sharp increase in volatility during periods of crisis (end of 2008), which validates the model's ability to capture the phenomenon of volatility clustering a well-known characteristic of financial time series.

5. Calculation of Conditional Value-at-Risk

The GARCH(1,1) model with Student's t-distributed innovations can be used to produce a dynamic estimate of Value-at-Risk (VaR), taking into account both conditional heteroscedasticity and the empirical distributional characteristics of financial returns.

The 95% and 99% quantiles were computed using the critical values of the Student-t distribution with $\nu = 11.76$ degrees of freedom. For each day t, the conditional VaR is calculated as follows :

$$\text{VaR}_{\alpha,t} = z_\alpha^{(t)} \cdot \hat{\sigma}_t$$
$$\text{where} \quad \hat{\sigma}_t = \sqrt{h_t} \quad \text{is the estimated conditional standard deviation}$$

5.1 Analyse graphique de la VaR

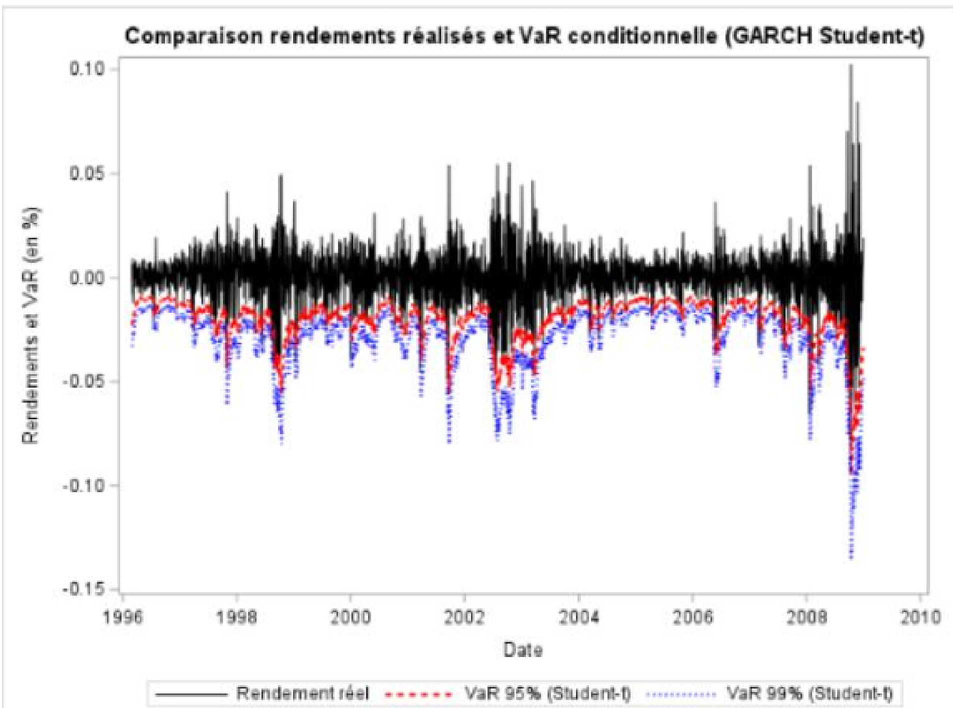

Figure 7: Comparison of actual returns and conditional VaR (GARCH Student-t)

This figure illustrates the comparison between actual returns and the two levels of conditional Value-at-Risk. It shows that VaR increases significantly during periods of stress (1998, 2002, 2008), reflecting the adaptability of the GARCH model to volatility shocks.

The VaR envelope appears to effectively capture extreme risk dynamics. However, a few exceedances appear during the 2008 crisis, requiring rigorous statistical validation.

6. Backtesting of VaR Estimated by the GARCH(1,1) Student-*t* Model

To evaluate the performance of the conditional Value-at-Risk estimated by the GARCH(1,1) model with Student-*t* innovations, we perform a series of backtesting procedures. This process enables us to assess whether the estimated VaR adequately captures extreme losses observed in the daily returns of the portfolio.

Two key dimensions are considered:

1. The frequency of exceptions (hedge backtesting, Kupiec test)
2. The independence of exceptions over time (Christoffersen test)

6.1 Kupiec Test (Proportion of Failures, POF) – Frequency of Exceptions

The Kupiec test (1995), also known as the Proportion of Failures (POF) test, evaluates whether the observed frequency of VaR violations aligns with the expected frequency implied by the chosen confidence level.

It is applied here to the conditional VaR estimates from the GARCH(1,1) model with Student-*t* innovations, using the 95% and 99% confidence levels.

Test de Kupiec – VaR conditionnelle à 95 % et 99 % (GARCH Student-t)

Niveau de VaR	Nb Exceptions	Nb Total	Taux observé	Statistique de Kupiec	P-value (H0 : VaR correcte)
VaR 95%	130	3119	0.041680	4.80635	0.02836
VaR 99%	33	3119	0.010580	0.10412	0.74694

Figure 8: Kupiec Test Results

The 99% VaR is validated, with an exception frequency consistent with the expected level (p-value = 0.74694). However, the 95% VaR is rejected at the 5% significance level, suggesting an underestimation of risk at this confidence level.

This result may be explained by the asymmetry of extreme losses or by a heavier-than-expected tail distribution during crisis periods.

6.2 Christoffersen Test – Independence of Exceptions

The Christoffersen test (1998) complements the analysis by evaluating whether VaR exceptions are independent over time, i.e., whether they occur randomly or tend to cluster together.

This test is crucial because a clustering of exceptions would indicate that the model fails to capture changes in volatility dynamics appropriately, thus violating the assumption of correct VaR calibration over time.

Test de Christoffersen – Indépendance des exceptions									
	t00	t01	t10	t11	pi0	pi1	Stat_Christoffersen	P_value	Indice
VaR 95%	2860	128	128	2	0.042838	0.0153846	3.042564	0.0811077	1
VaR 99%	3052	33	33	0	0.0106969	1E-10	0.7060102	0.4007715	2

Figure 9: Christoffersen Test Results

Both p-values are greater than 5%, which does not lead to a rejection of the null hypothesis of temporal independence of exceptions. Thus, the violations of the Student-t GARCH VaR are well dispersed over time—an expected property of a well-specified risk model.

In conclusion, the GARCH(1,1) model with Student-t innovations effectively captures the conditional volatility dynamics observed in the European equity portfolio. The 99% VaR estimated from this model passes the backtesting tests (Kupiec and Christoffersen), making it a robust tool for extreme risk management.

However, the 95% VaR appears to be slightly undercalibrated, emphasizing the importance of selecting an appropriate confidence level when assessing financial risk.

7. Estimation of Conditional Volatility Using a Kalman Filter

7.1 Stochastic variance modeling

In this section, we implement the estimation of the conditional variance of portfolio returns through a stochastic volatility (SV) model using the Kalman filter.

This approach is based on a state space representation, in which the logarithm of the variance follows an autoregressive process of order 1 (AR(1)):

$$y_t = \log(r_t^2) = h_t + \varepsilon_t, \quad \varepsilon_t \sim \mathcal{N}(0, R)$$
$$h_t = \mu + \phi(h_{t-1} - \mu) + \eta_t, \quad \eta_t \sim \mathcal{N}(0, Q)$$

The Kalman filter allows the latent state h_t to be estimated sequentially from observations y_t, by combining a forecast based on the system dynamics (process noise Q) with information from the new observation (observation noise R).

The parameters Q and R are estimated via maximum likelihood using the `PROC IML` procedure in SAS.

7.2 Estimation using sliding windows

To capture temporal variations in volatility, the model is applied using sliding windows of 500 days. This technique produces local and adaptive estimates of the conditional variance.

The figure below illustrates the evolution of conditional variance over the first three windows. This graphical representation highlights the model's ability to adapt to volatility regime changes, particularly around the financial crisis, while producing smoother dynamics compared to traditional GARCH models.

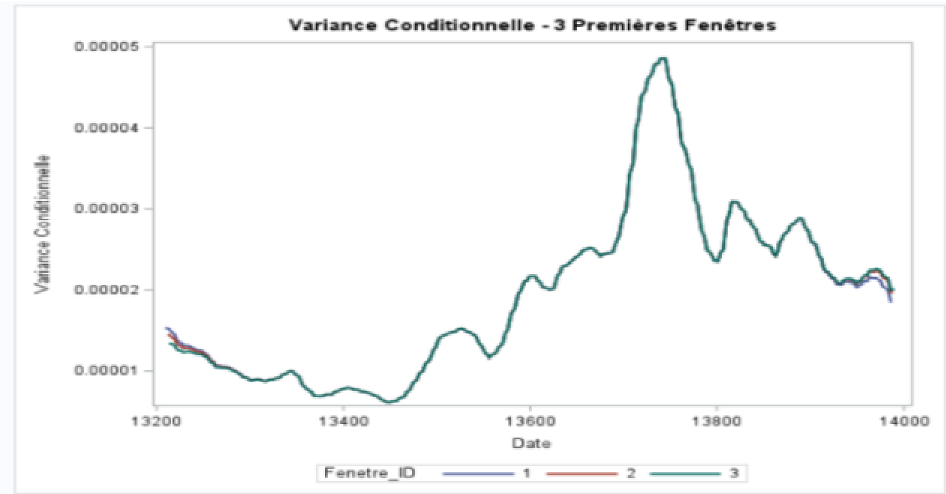

Figure 10: Enter Caption

7.3 Descriptive Statistics of Estimates

The outputs below present the descriptive statistics of the simplified stochastic volatility (SV) model estimated via the Kalman filter.

The mean of the conditional variance is approximately 3.97×10^{-4}, while the maximum reaches nearly 5.44×10^{-4}, reflecting the sharp increase in volatility during periods of financial stress.

Statistiques du Modèle SV Simplifié

La procédure MEANS

Variable	N	Moyenne	Ec-type	Minimum	Maximum
sigma2_t	1310000	0.000039793	0.000044126	3.230283E-6	0.000547216
VaR_95	1310000	0.0093848	0.0044279	0.0029566	0.0384809
VaR_99	1310000	0.0132699	0.0062609	0.0041805	0.0544113

Figure 11: Enter Caption

The conditional Value-at-Risk (VaR) is then computed at each date using the following formula:

$$\text{VaR}_{\alpha,t}^{\text{Kalman}} = z_\alpha \cdot \hat{\sigma}_t$$

where $z_\alpha^{(t)}$ is the critical value of the distribution corresponding to the chosen confidence level α, and h_t is the estimated conditional variance from the Kalman filter.

7.4 Backtesting of VaR (Kalman)

To evaluate the reliability of the estimated VaR values, we apply two standard backtesting procedures:

- The **Kupiec test**, which assesses whether the frequency of VaR exceptions matches the expected rate.
- The **Christoffersen test**, which evaluates the temporal independence of those exceptions.

The results of these tests are summarized in the table below.

Test de Kupiec – VaR conditionnelle à 95 % et 99 % (Filtre de Kalman)

Niveau de VaR	Nb Exceptions	Nb Total	Taux observé	Statistique de Kupiec	P-value (H0 : VaR correcte)
VaR 95%	221422	1310000	0.16902	247955.40	0
VaR 99%	128533	1310000	0.09812	366756.59	0

The Kupiec test strongly rejects the hypothesis of a correctly calibrated VaR. The observed frequency of exceptions is significantly higher than the theoretical thresholds (16.9% versus 5% for the 95% VaR), suggesting a severe underestimation of extreme risk.

Test de Christoffersen – Indépendance des violations (Kalman)

	N00	N01	N10	N11	Prob(viol|non-viol)	Prob(viol|viol)	Stat_Christoffersen	P_value	Indice
VaR 95% (Kalman)	1086543	2034	2034	219388	0.0018685	0.9908139	1137609.4	0	1
VaR 99% (Kalman)	1179081	2385	2385	126148	0.0020187	0.9814445	782716.25	0	2

The very high statistical values and near-zero p-values also indicate that the exceptions are clustered, and therefore that Kalman's VaR does not capture the temporal independence of the violations.

8. VaR estimation using the exponential weighted moving average method (EWMA)

EWMA model allows us an approach where weights aren't equal and decrease exponentially according to time. As we already have used the GARCH method to extract the conditional volatility of the series we have already justified the EWMA approach for our VaR estimations. However the Augmented Dickey Fuller test to check for the presence of an unitary root, the ARCH test effects and the Ljung Box for the conditional variance are required to start the EWMA approach. The choice of modeling our series with an EWMA model appears relevant, first because of the exponentially decreasing weights over time that allow to model the persistence of volatility, which is relevant in our case with the presence of ARCH effects. Also, since our objective is to estimate the VaR, the reactivity of volatility that the EWMA model provides is an advantage. Finally, because of its large use in the financial industry, notably theorized within RiskMetrics by JP Morgan, the EWMA approach is efficient for detecting volatility changes, The EWMA model is estimated using the smoothing parameter lambda = 0.94 (cf RiskMetrics). The recursive calculation of the model's formula allows us to extract the EWMA variance and the EWMA standard deviation (EWMA volatility). Although the EWMA model does not rely on any assumption regarding the distribution of returns, our calculated VaRs rely on the estimation of the conditional volatility from the EWMA model combined with the assumption of conditional normality, which allows us to use the quantiles of the normal distribution to obtain our estimated VaR95 and VaR99.

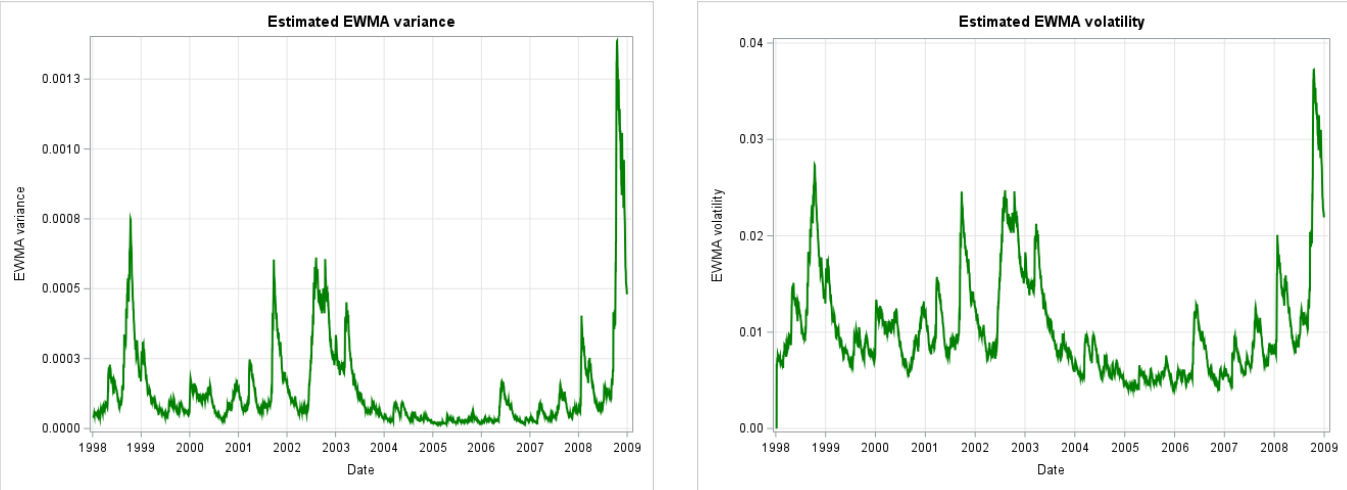

Figure 12: Estimated variance and volatility with exponential weighted moving average method along time

We observe (Figure 4.2) significant losses in our portfolio returns during the periods between 1998–1999, 2002–2004 and also between 2008–2009, which correspond to clusters of high volatility and periods of crises (Russian crisis, dotcom bubble, financial crisis). Regarding the dotcom bubble period, we note two peaks in 2000 and early 2008 that are exceptions to the estimated VaR95% and VaR99%, however the estimated VaRs over the period overall follow quite well the evolution of our portfolio returns.

Concerning the validation of our EWMA modeling, that is, the quality of our Value At Risk estimation, we conduct the Kupiec and Christofferson tests. We observe a total of 143 exceptions for the 95% VaR, which gives a rate of 5.25%, and 30 exceptions for the 99% VaR, which gives a rate of 1.10%, over a total of 2,722 observations. These observed exception rates are close to the theoretically expected levels.

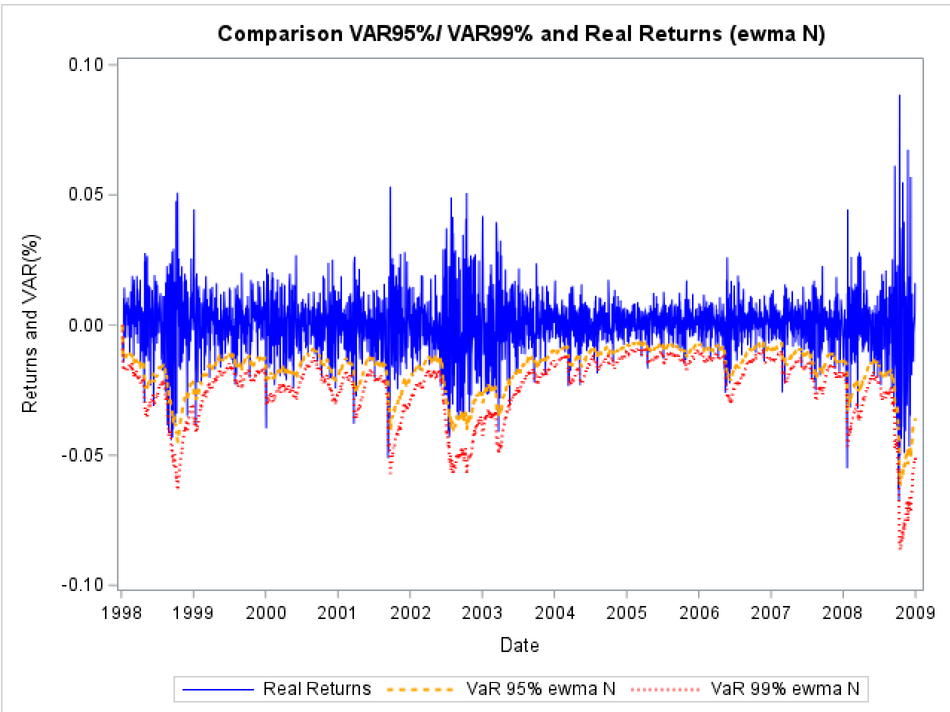

Figure 13: Estimated variance and volatility with exponential weighted moving average method along time

Kupiec test – Conditionnal VaR at 95 % and at 99 % (Normal EWMA)

VaR level	Number of Exceptions	Total number	Observed rate	Kupiec's Statistics	P-value (H0 : the VaR is accurate)
VaR 95%	143	2722	0.052535	0.36248	0.54713
VaR 99%	30	2722	0.011021	0.27759	0.59828

Test de Christoffersen – Indépendance des exceptions

	t00	t01	t10	t11	pi0	pi1	Stat_Christoffersen	P_value	Indice
VaR 95%	2443	135	136	7	0.0523662	0.048951	0.0325547	0.8568155	1
VaR 99%	2663	28	29	1	0.0104051	0.0333333	0.9521439	0.3291742	2

Figure 14: Estimated variance and volatility with exponential weighted moving average method along time

Regarding the p-values, we obtain with the Kupiec test a p-value of 54.71% for the 95% VaR and 59.82% for the 99% VaR. We do not reject the null hypothesis, so our EWMA model provides a correct estimation of the extreme loss risk according to the unconditional coverage test, and our VaR99% and VaR95% are properly estimated (valid).

Finally, we perform the Christofferson test, and we obtain a p-value of 85.68% for the 95% VaR and 32.92% for the 99% VaR. Thus, we do not reject the null hypothesis of temporal independence of exceptions. The exceptions appear to be random and occur randomly over time, in other words, our model correctly captures the dynamics of market risk. It is therefore relevant, and the observed exception rates are consistent with the specified risk levels (95% and 99%). We can thus conclude that our EWMA modeling is reliable for both the 95% and 99% VaR estimations.

Conclusion of the Backtesting of the VaR Estimated by the Kalman Filter

Limitations of the VaR model based on the stochastic volatility approach estimated by the Kalman filter are important in risk management. The model underestimates risk, as shown by high exception rates and clustering of violations (breaches of independence and correct calibration). While it can adequately capture average volatility, it cannot adequately model tail risk and volatility shocks. This implies that combining it with more adaptive models such as GARCH with Student's t or enhancing it with non-linear or Bayesian frameworks would offer a more robust risk assessment.

9 Comparison of GARCH, Kalman, and EWMA Models: Empirical Validation

Here is a set of comparisons between three volatility modeling approaches applied to an equally weighted portfolio comprising six leading European indices: CAC40, DAX, SMI, FTSE100, IBEX35, and MDAXI, over the period extending from 1996 through 2009. These approaches include the GARCH(1,1) model with Student's t-distribution, the Exponentially Weighted Moving Average (EWMA), and a simplified stochastic volatility model estimated via the Kalman filter.

- **GARCH(1,1) with Student's t-distribution** is the best model for conditional risk estimation. It captures volatility clustering and fat-tailed behavior in a reliable manner, thus providing a good representation of extreme market fluctuations. Its violation frequency at the 99% VaR level accords very well with theoretical expectations (p-value = 0.75). While at the 95% level there is a slight excess of exceptions, in that instance also the model passes backtesting tests and retains superiority in high-risk environments.
- **EWMA** is a simpler, less reactive model and meets the statistical requirements for conditional VaR estimation when the normality assumption about the distribution of returns is considered reasonable. Backtesting results using the Kupiec and Christoffersen tests are quite satisfactory.

However, during periods of extreme market stress, fat tails need to be taken into consideration hence this method would lack reliability.

- **Kalman filter** model, as applied in this study, systematically fails empirical validation. Exception frequencies were far higher than expected and the violations were strongly clustered. These results confirmed by near-zero p-values in both Kupiec and Christoffersen tests highlight the model's structural inability to account for tail risk as well as temporal dependence. It may smooth average volatility well but fails to adapt to sharp regime shifts and stress periods.

To sum up, the models that have heavy tails in their parametrization—like GARCH(1,1) with Student's t-distribution—are appropriate to be best accepted in environments of volatility for financial risk management. Under moderate conditions of the market, EWMA can be considered as a simpler model and quite robust. The Kalman filter as linear Gaussian is inflexible for modeling tail risk and would need extensions to compete effectively in practical risk management contexts—either through nonlinear dynamics, fat-tailed distributions, or Bayesian estimation.

General Conclusion

In this thesis, a few conditional volatility models and their estimation accuracy as well as predictive performance on an actual portfolio comprising the major European indices CAC40, DAX, SMI, FTSE100, IBEX35, MDAXI were assessed. The endeavor sought to model return volatility and subsequently assess the soundness of Value-at-Risk (VaR) forecasts for market risk management. The three models compared were:

GARCH(1,1) with Student-t innovations that can effectively capture fat tails and market shocks;

A simple **Kalman-filtered stochastic volatility** model, trying to follow hidden variance changes;

The **EWMA** model, liked in practice for its ease and quickness.

The empirical results indicate that the GARCH-t model gives the best accurate and robust VaR estimates, especially at the 99% level. The EWMA model performs well in validation tests and as a practical reliable tool due to normality assumption. On the other hand, the Kalman model misses extreme risk since backtests revealed both underestimation and violation clustering.

Methodologically, the paper highlights the role of thorough calibration and backtesting aided by maximum likelihood estimation and tests such as those by Kupiec and Christoffersen. Equally important is extremist model dynamics—not just average dynamics—in expressing reality.

Future research may try more flexible extensions of the Kalman model (e.g., Bayesian, Monte Carlo methods), hybrid approaches, or add macroeconomic variables to enrich volatility modeling.

In the end, the findings support using empirically validated models that balance theoretical rigor with practical reliability to manage market risk effectively.

References

Auffie, D., & Pan, J. (1997). An overview of value at risk. *Journal of Derivatives*, Spring, 7-48.

Forecasting stochastic Volatility using the Kalman filter: An Application to Canadian Interest Rates and Price-Earnings Ratio.

Doowoo Nam Nam-gu. (2012). The estimation and decomposition of value-at-risk for non-normal portfolio returns.

Escanciano & Olmo (2010). Backtesting Parametric Value-at-Risk With Estimation Risk.

J. Danielsson, Casper G. De Vries. (2000). Value-at-Risk and Extreme Returns. *Annales d'Économie et de Statistique*, No 60.

Jorion, P. (2006). *Value at risk: The new benchmark for managing financial risk.* New York: McGraw-Hill.

M. Bohdalová, M. Greguš. (2015). Estimating value-at-risk based on non-normal distributions. CBU International Conference on Innovation, Technology Transfer and Education, Prague, Czech Republic.

Tsay, R. S. (2010). *Analysis of financial time series.* New Jersey: John Wiley & Sons.

Galdi & Pereira (2008). Value at Risk (VaR) Using Volatility Forecasting Models EWMA, GARCH and Stochastic Volatility.

María Isabel Restrepo E. (2012). Estimating portfolio value at risk with GARCH and MGARCH models.

Vesna Bucevska. (2013). An empirical evaluation of GARCH models in value-at-risk estimation: Evidence from the Macedonian Stock Exchange. Faculty of Economics, University "Ss. Cyril and Methodius", Skopje.

Enterprise Risk Management. Master M2 EGR, Master in Economics of Global Risks, TSE.

Box, G. E. P., Jenkins, G. M., Reinsel, G. C., & Ljung, G. M. (2015). *Time series analysis: Forecasting and control* (5th ed.).

Lux, Härdle & Lessmann. (2020). Data driven value-at-risk forecasting using a SVR-GARCH-KDE hybrid.

Mária Bohdalová, Michal Greguš. (2013). Copula based VaR approach for European stocks portfolio. CBU International Conference on Integration and Innovation in Science and Education.

J.P. Morgan. (1996). RiskMetrics Technical Document.

Shaik & Padmakumari. (2022). Value-at-risk (VaR) estimation and backtesting during COVID-19: Empirical analysis based on BRICS and US stock markets.

Kuchynka, Alexandr. (2008). Volatility extraction using the Kalman filter. Charles University, Institute of Economic Studies (IES).

Lemoine, Matthieu & Pelgrin, Florian. (2003). Introduction aux modèles espace état et au filtre de Kalman.

Estimating Value at Risk with the Kalman Filter. (2003). A. Berardi University of Verona, S. Corradin, European Central Bank