Efficient Value at Risk - Risk Factor Mapping

Philip Johansson Casserstål
Oscar Stenander Eklund

Master of Science in Engineering Mathematics Degree Project
Lunds Tekniska Högskola
Supervisor: Magnus Wiktorsson

June 2026

Abstract

This project examines whether a Taylor-based risk-factor mapping framework can accurately and efficiently estimate Value at Risk for option-heavy portfolios without the high computational cost of full revaluation. The method maps changes in portfolio value to a selected set of market risk factors, including equity spot returns, foreign exchange rates, interest-rate curves, and principal components of implied-volatility surfaces. Portfolio sensitivities, including spot gamma, second-order volatility terms, and spot-factor cross sensitivities, are calculated using bump-and-reprice finite differences.

The framework is evaluated in two dimensions. Firstly, the Taylor approximation is compared with full revaluation using historical data from 2022 to 2026 and randomized option-heavy portfolio tests. Secondly, several VaR scenario-generation models are tested, including EWMA, GARCH, GJR-GARCH and state-dependent extensions, in order to assess their ability to capture time-varying volatility, tail risk and violation clustering. The VaR forecasts are evaluated using exceedance backtesting, unconditional coverage tests and duration-based tests for violation independence.

The results show that the Taylor-based approach closely matches full-revaluation estimates, with low average relative errors and similar exceedance patterns during backtesting. Additional analysis of stock, foreign exchange, and bond instruments indicates that the remaining approximation errors are mainly driven by option nonlinearity. The best-performing VaR specifications produce violation rates close to the expected levels and show no evidence of violation clustering. Overall, the results suggest that the framework can significantly reduce scenario valuation cost while preserving the main risk signals of full revaluation.

Acknowledgments

We would like to express our deepest gratitude to our supervisor, Magnus Wiktorsson at LTH, for providing guidance and insights of great value. We also wish to thank Erik Lindström for examining this project.

Furthermore, we would like to thank Scila AB for their guidance and for allowing us to collaborate with them on this project.

Philip Johansson Casserstål & Oscar Stenander Eklund
Stockholm, June 2026

Contents

List of Figures

List of Tables

1 Introduction

1.1 Motivation

Calculating VaR for certain non-linear portfolios can become computationally expensive given portfolios of heavy sizes, especially when it comes to option-heavy portfolios where the value calculations can contain pricing frameworks with repeated option valuation across many shocked market states. Only attributing a portfolio with the help of full repricing, especially given the context of VaR, is not always the optimal strategy in financial risk management[Jor06, pp. 209–210]. A competitive alternative to using full repricing calculations in this context is to use what is called risk factor mapping. The motivation is that risk factor mapping is computationally much more efficient than full repricing since it often relies on an approximation formula instead of recalculating the exact value of a portfolio.

When the given portfolio has major non-linear properties, for example in the case of an option-heavy portfolio, the use of risk-factor-mapping can significantly lower the computational cost of the portfolio PnL calculation at the expense of computational accuracy. Instruments are mapped to sensitivities with respect to selected risk factors, such as spot, interest-rate tenors, and implied-volatility principal components. This approach aligns with the regulatory framework set out in the Fundamental Review of the Trading Book[Bas19]. In this project, this idea is extended by mapping the implied volatility surface to principal component risk factors, so that volatility-surface movements can be represented by a small number of PC shocks rather than by all volatility nodes individually.

1.2 Problem formulation

The central problem of this project can be formulated as follows:

Can a Taylor-based risk-factor mapping framework approximate VaR of option-heavy portfolios with sufficient accuracy relative to full revaluation, while reducing the computational cost of scenario valuation?

- How can option portfolios be mapped to a reduced set of risk factors, including spot, implied-volatility principal components, interest-rate tenors and FX rates?
- Which Taylor terms are necessary to obtain accurate P&L approximations for non-linear portfolios, including delta and gamma, second-order volatility terms and relevant cross-sensitivities?
- How accurately does the Taylor approximation reproduce full-revaluation P&L and VaR under the same Filtered Historical Simulation scenarios?

- How does the choice of volatility and covariance scenario model affect the VaR forecasts, in terms of exceedance rates and violation clustering?
- To what extent can the resulting VaR estimates be validated through backtesting against realized portfolio P&L over the 2022–2026 validation period?

2 Background

2.1 Empirical properties of financial returns

The viewpoint of many market analysts has been, over the years, and remains, to use an event-based approach in which one tries to rationalize a market movement by labeling it to a certain economic or political event. The result of more than half a century of empirical studies on financial time series has shown that rather random variation of price movement can be connected to non-trivial statistical properties.[Con01] It is important to list and explain relevant statistical properties and phenomena that directly affect asset price movements before starting the modeling part.

2.1.1 Asymmetry

Practice has shown that the market does not react symmetrically between positive and negative asset returns. For example, negative equity returns tend to be followed by increased volatility, which is greater than it would be in the case of a positive price movement of the same size. This phenomenon is commonly referred to as the leverage effect. In general, volatility tends to be negatively correlated with asset return.[Bou+01] In the case of derivatives and especially the option market, this is especially interesting since the option value is dependent on both the underlying asset and the implied volatility. The phenomenon is intuitively logical; when the stock market plummets, the uncertainty, risk premiums, and the demand for protected options rise heavily. As a result, the implied volatility also rises, especially for put options.

2.1.2 Volatility Clustering and Heavy Tails

When dealing with a large number of portfolios with a large number of assets, it is important to find a simplification for the model in question. A common simplification used is the assumption that asset returns are normally distributed and independent and identically distributed (i.i.d). Practice has shown that in reality, most financial return series do not adhere to this assumption. A key reason why this assumption is unjustified is because of the volatility clustering effects. While raw returns often exhibit little linear autocorrelation, absolute and squared returns tend to be positively auto-correlated, indicating persistence in volatility. There appears to be no simple formulae for scaling VaR when returns have volatility clustering effects in the normal assumption case.[Con01]

Volatility clustering is the financial markets concept that large absolute returns tend to be followed by further large absolute returns, while calm periods tend to be followed by calm periods [LMN15]. The issue is that, while the returns do not necessarily tend to be auto-correlated themselves, the absolute size of the returns tends to be dependent over

time. This means that the direction of the next day's return is difficult to predict, but the size of the return is easier. This clustering further complicates the VaR prediction for simpler linear models, which assume a normal i.i.d. Instead, a more sophisticated approach to solve this is to apply a GARCH model to simulate daily returns over the risk horizon. This is further explained in section 1.4.1.

Another difficulty that arises when making the i.i.d normal asset returns assumption is that, in reality, asset returns tend to exhibit extreme values more often than what the normal distribution predicts. The market makes an extreme move more often than what the simple Gaussian model would give. That means that asset returns assumed to be i.i.d normal would underestimate the probability of great losses. This is extra problematic when dealing with Value at Risk (VaR), since VaR aims to predict the loss distribution in the tails of the distribution. To solve the issue with the tails, one could incorporate the Student's t-distribution into the modeling, using an appropriate number of degrees of freedom for the sampled data[Con01]. Using the t-distribution generally improves the accuracy of the VaR calculations by incorporating heavier tails compared to the normal distribution. Combining a GARCH model with the student's t-distribution thus seems to be the best of both worlds since it minimizes errors arising from both volatility clustering and thinner tails.

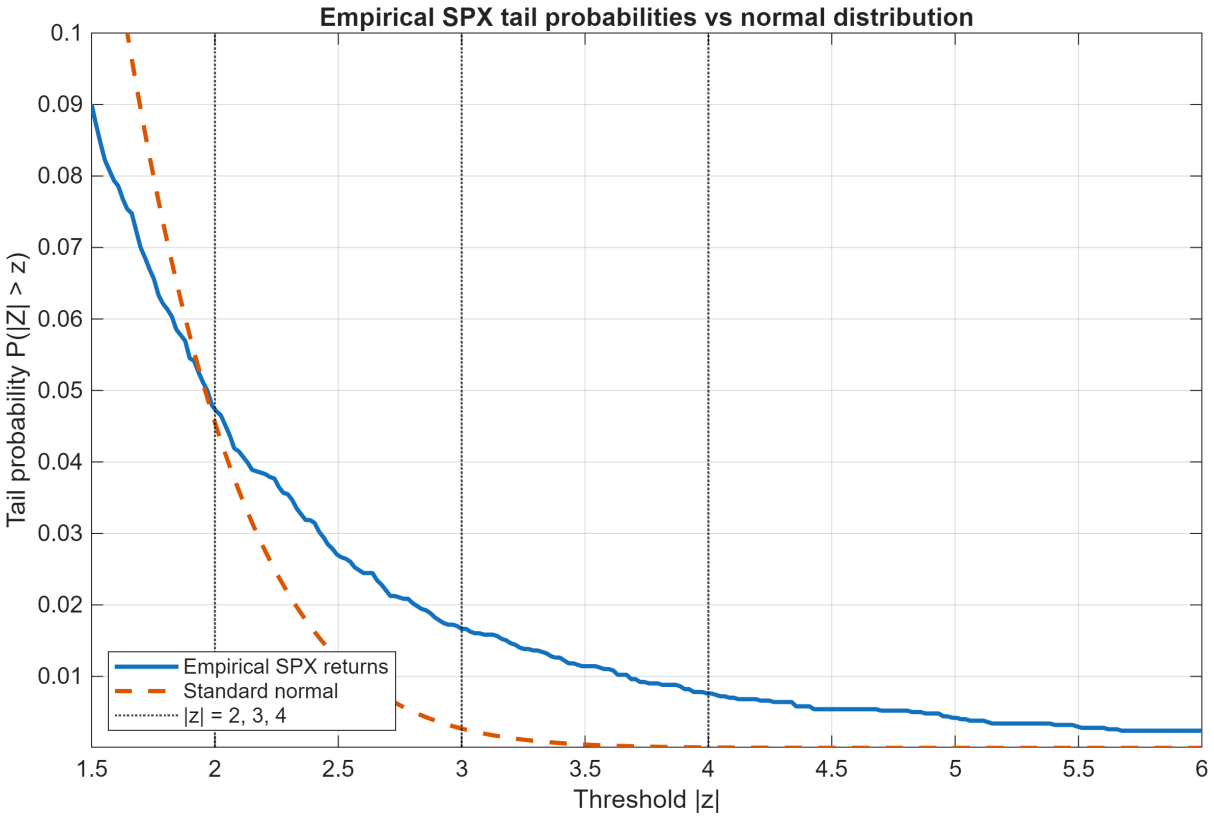


Figure 1: Comparison of tail behavior between the standard normal distribution and empirical data of the S&P-500 standardized daily log returns between the years 2006-2026.

2.2 Value at Risk

Financial risk is one that relates to possible losses owing to financial market activities. These losses can occur due to, for example, interest-rate movements or defaults on financial obligations. For financial institutions, the function is primarily to manage financial risks actively. It is necessary to assess, mitigate, or advise on financial risks. The recent growth of the risk management industry can be traced directly to the increased volatility of financial markets since the early 1970s. The common denominator across financial market crashes and events throughout history is that they are unpredictable. The rapidity of changes in price and volatility caused major financial losses before any real risk management was developed clearly. The question asked was: how do you limit potential losses while still allowing traders to take views on markets?[Jor06, Chapter 1]

Many conventional risk-measurements were since developed. One example is to establish stop-loss limits. However, this approach is still incomplete. It does not consider the volatility of the risk factors, or their correlation; these could vary across markets. Since sensitivity measures are not additive, they cannot be used to aggregate risks [Jor06, pp. 18–21]. This is where Value at Risk (VaR) comes in. VaR combines the price-yield relationship with the probability of an adverse market movement. Thus, VaR is a statistical risk measure of potential losses.

The definition of VaR is as follows: VaR summarizes the worst loss over a target horizon that will not be exceeded with a given level of confidence[Jor06, p. 22]. Let $L_{t+1} = -\Delta V_{t+1}$ denote the portfolio loss over the risk horizon. For confidence level α, the Value at Risk is defined as the α-quantile of the loss distribution,

$$\mathrm{VaR}_\alpha = F_L^{-1}(\alpha),$$

or equivalently,

$$\mathbb{P}(L_{t+1} > \mathrm{VaR}_\alpha) = 1 - \alpha.$$

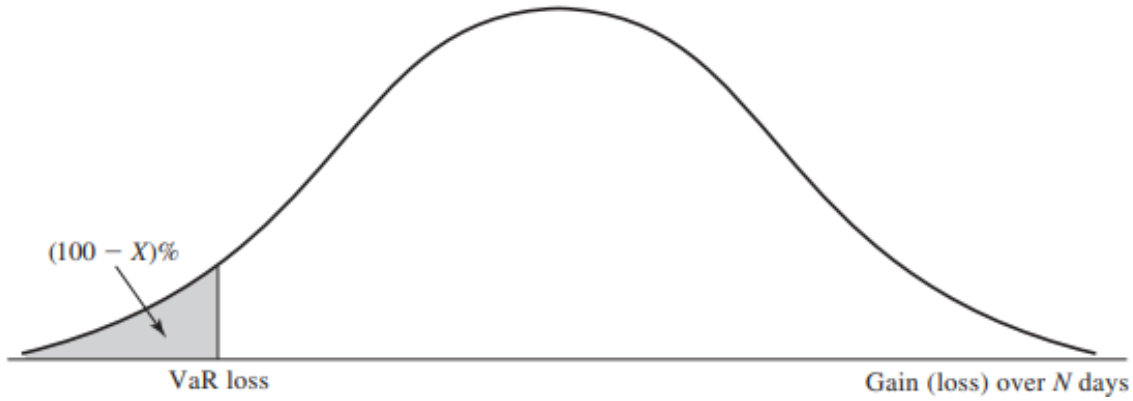


Figure 2: Figure shows a normal distribution with the VaR loss at a certain percentile prescribed

2.3 Risk Factor mapping

In practice, VaR measures are based on the risk factor mapping of the portfolio. A risk factor mapping is the construction of a model that relates the portfolio return to variations of its risk factors. In factor models, the coefficient parameters on the risk factor variations are called the portfolio's sensitivities to variations in the risk factors. These sensitivities quantify how small changes in each risk factor affect the portfolio's value. Consequently, the distribution of the portfolio P&L can be approximated by combining the sensitivities with the distribution of the risk factor changes[Ale09, pp. 25–26]. When the P&L of the portfolio is obtained, the Value at Risk can also be obtained at a certain confidence level and in a certain time horizon ahead. In practice, for portfolios containing derivative instruments, such as options, these sensitivities correspond to well-known risk measures such as delta, gamma, and vega. These quantities play a central role in approximating portfolio value changes and will be discussed in detail in subsequent sections.

The motivation for why one would map the portfolio to its risk factors, as opposed to mapping to its instruments, is that this provides an efficient framework for hedging certain risks and for capital allocation. Furthermore, many portfolios are simply too large to measure VaR by mapping to all of their instruments. Simply mapping to the instruments would mean that one would measure VaR by the level of each asset in a portfolio, which would require one to model the multivariate distribution for the return of each asset. Only a few simple portfolios are small enough to not require the aid of risk factor mapping. Which risk factors affect the final P&L of the portfolio directly depends on the financial assets that belong to the portfolio. For example, considering an equity portfolio which has positions in cash equity and index futures, one would typically consider risk factors such as:

- Market spot equities such as S&P 500, OMXS30 etc.

- Forex spot rates such as EURUSD, GBPUSD etc.
- Dividend yields in each market.
- Spot LIBOR rates, with maturity in this case being equal to the maturity of the domestic and foreign currencies.

For an option-heavy equity portfolio, one would consider risk factors such as implied volatilities, underlying asset prices, and the term structure of interest rates. The drawback of using risk factor mapping on a portfolio is that it comes with a source of model risk in the estimation of VaR:

- The choice of risk factors is subjective. Different risk managers might interpret the necessity of different risk factors for a portfolio.
- The risk factor sensitivities might have estimation errors.
- We only capture the systematic VaR, and ignore the specific risk of the portfolio.

[Ale09, p. 27]Having established the conceptual framework of risk factor mapping, the next step is to specify how portfolio sensitivities are computed for different types of financial instruments. In particular, for derivative portfolios, sensitivity-based approximations rely on Taylor expansions of the portfolio value function, which we introduce in the following sections.

2.4 Volatility Forecasting

The reason to forecast variances and covariances is to use them when forecasting a portfolio's P&L over a given time period. Often, volatility forecasting can be done by examining historical data, rather than having to use option pricing theory to obtain implied volatilities. Only using implied volatility comes with numerous different issues from a practical point of view. Implied volatilities are solely based on the expectations of a given option pricing model. Most pricing models assume that the standard deviation is constant, and therefore will be difficult to interpret, and will also lead to imprecise forecasts if the given pricing model is not correctly specified[Hul21, Chapters 15 & 23]. The evidence of which forecasting ability is superior of historical volatility over implied volatility is mixed, depending on the time series considered. In practice, combining the forecasting methods gives the best of both worlds.

2.4.1 Exponential Weighted Moving Average

The utilization of an Exponential Weighted Moving Average (EWMA) approach when capturing the dynamic features of volatility has been shown to be at least as simple but

yet an improvement to the Simple Moving Average model (SMA).[JP 96, Chapter 5]

The core idea of EWMA is to use an exponential moving average of historical observations, where the latest observations carry the highest weight in the volatility estimate. In contrast to an equal-weighted moving average, volatility reacts faster to shocks in the market since the latest data are more substantially weighted than the data in the past. Furthermore, following a shock, the volatility declines exponentially as the weight of the shock observation falls. In contrast, the use of a simple moving average leads to abrupt changes in the standard deviation once the shock falls out of the measurement sample. The general mathematical formula for forecasting volatility according to the EWMA model, based on T observations of data, is

$$\sigma^2 = (1-\lambda)\sum_{t=1}^{T}\lambda^{t-1}(r_t-\bar{r})^2.$$

The parameter $1 > \lambda > 0$ is known as the decay factor. This parameter determines the relative weights that are applied to the returns. One feature of the EWMA is that it can be written in a recursive form, which makes it easier to handle since the sum of returns no longer has to be considered. Furthermore, we can keep matters simple by setting the sample mean $\bar{r}$ to zero. This assumption is based on the assumption that one possesses a sufficiently large data set (the general assumption is an infinite data set). Given that we know the return observation the previous day, the general form of the EWMA can be written as:

$$\sigma^2_{t+1|t} = \lambda\sigma^2_{t|t-1} + (1-\lambda)\, r_t^2. \tag{1}$$

This form is particularly useful for volatility forecasting in different VaR models, for example, Filtered Historical Simulation. The EWMA model can be interpreted as a special case of a GARCH(1,1) model where the intercept term is set to zero and the persistence parameter equals the decay factor[Hul21, Chapter 23]. Consequently, the model assumes that volatility shocks decay exponentially through time, with the persistence determined by the parameter λ.

2.4.2 GARCH

The Generalized Auto-regressive Conditional Heteroskedasticity (GARCH) model is a more general version of the EWMA model. While the EWMA model can be seen as a restricted version of a GARCH model, the standard GARCH(1,1) model introduces a constant long-run average variance, allowing the process to be mean-reverting[Bol86]. The GARCH(1,1) model defines the conditional variance at time t as a function of the

previous period's residual and the previous period's variance:

$$\sigma_t^2 = \omega + \alpha r_{t-1}^2 + \beta \sigma_{t-1}^2 \tag{2}$$

where ω is the constant intercept, α represents the reaction of volatility to new market shocks and β represents the persistence of volatility. For the model to be stationary and the variance to remain positive, the constraints $\omega > 0$, $\alpha, \beta \geq 0$, and $\alpha + \beta < 1$ must be satisfied.

A GARCH model with a high α value corresponds to a model that changes quickly depending on shocks, while a high β value is more robust and makes slower adjustments. The EWMA model assumes $\alpha + \beta = 1$, while the GARCH model is more flexible.

2.4.3 GJR-GARCH

A significant limitation of the standard GARCH model is its symmetric response to shocks. It assumes that positive and negative returns of the same magnitude have the identical impact on future volatility. However, financial data has asymmetric properties, where negative shocks tend to increase volatility more than positive shocks. To address this asymmetry, Glosten, Jagannathan, and Runkle introduced the GJR-GARCH model [GJR93]. The conditional variance is expressed as:

$$\sigma_t^2 = \omega + (\alpha + \gamma I_{t-1}) r_{t-1}^2 + \beta \sigma_{t-1}^2 \tag{3}$$

where I_{t-1} is an indicator function that takes the value of 1 if the previous return r_{t-1} is negative, and 0 otherwise:

$$I_{t-1} = \begin{cases} 1 & \text{if } r_{t-1} < 0 \\ 0 & \text{if } r_{t-1} \geq 0 \end{cases}$$

With this structure, the impact of a positive shock is captured by α, while the impact of a negative shock is captured by $\alpha + \gamma$. If the leverage parameter $\gamma > 0$, it indicates that negative shocks provide a larger contribution to future volatility than positive shocks of the same magnitude. If $\gamma = 0$ we get the standard GARCH(1,1) again. Overall, the GJR-GARCH model provides an even more flexible option compared to the GARCH. However, it comes at the cost of needing to find a good value of yet another parameter.

2.5 Market Regimes and Covariance

Financial markets do not generally operate under a single, constant set of dynamics. Instead, they tend to alternate between phases, called market regimes or states. A regime can be characterized by a level of volatility, something earlier mentioned being

time-dependent. It is also characterized by correlation behavior between assets, and the tail heaviness of return distributions might also change. For example, we might have calm, low-volatility periods with small covariance between assets and turbulent periods with more movements and clearer correlations. In the context of Value at Risk prediction, a model that captures regime shifts in both volatilities and correlation has a better chance of producing good VaR estimates.

2.6 Filtered Historical Simulation

In the first generations of VaR models, back in the 1990s, the concept of Historical Simulation (HS) was introduced. The HS model simply used some history in the changes of the selected risk factors, for example, spanning the daily changes of certain risk factors back for example 1000 data points in history and reevaluating the portfolio solely based on the history. HS models were simple to construct and understand. Furthermore, they made no assumptions about the tail of the joint risk factor return distribution. Moreover, provided that a period of stress had occurred in the data window, the model would incorporate it into its risk estimate.[GM15]

However, empirical research shows that standard HS models are not well-posed due to a number of issues appearing with VaR estimates. Firstly, many risk factor returns are not well-described by processes with constant volatility. This means that risk factor returns are therefore not necessarily independent over time. Moreover, current market conditions contain some information about returns in the near future: if conditional volatility is elevated, then usually larger returns are expected than if it is not. These issues can be in some ways solved by introducing a family of models called Filtered Historical Simulation Value-at-Risk models. The core idea of these models is that they scale the historical data based on the current market conditions. Meaning, if the current market conditions are less volatile, they would scale the estimates to be damped down, and if they are more volatile, they would be turned up.[GM15].

In an unfiltered historical simulation model, each historical vector of risk-factor returns is treated as a possible scenario for the next trading day. Each of these possible repetitions of prior days generates P&L. These P&Ls are then gathered into a distribution, and a certain percentile of this distribution is estimated. In contrast, FHS models identify the volatilities of returns as a crucial property. For each day i, they calculate some estimate of the volatility of risk factor j on that day, $\sigma_j(i)$. All the historical returns in the data window are then devolatilized by dividing them by the relevant volatility estimate.

Consider the return of a certain risk factor:

$$r_t = \sigma_t \epsilon_t,$$

where σ_t denotes the conditional volatility of the return and ϵ_t represents a standardized innovation with mean zero and unit variance. The objective of the filtering procedure is to remove the time-varying volatility component from the historical returns, producing a sequence of approximately stationary residuals. Given a volatility estimate σ_t obtained from a volatility forecasting model such as EWMA, equation (1), the historical returns are standardized according to

$$\epsilon_t = \frac{r_t}{\sigma_t}.$$

The sequence $\{\epsilon_t\}_{t=1}^{T}$ represents the filtered residuals of the return process. Under the FHS framework, these residuals are assumed to be approximately independent and identically distributed. Future market scenarios are generated by sampling from the empirical distribution of the standardized residuals. Let

$$\epsilon^{(i)} \sim \{\epsilon_1, \epsilon_2, \ldots, \epsilon_T\}$$

denote a randomly selected residual from the historical sample. To generate simulated future returns, the sampled residuals are scaled by a forecast of the next-period volatility. Let $\sigma_{t+1|t}$ denote the forecasted conditional volatility obtained from the volatility model. The simulated return is then given by

$$r_{t+1}^{(i)} = \sigma_{t+1|t} \epsilon^{(i)}.$$

This procedure generates a set of possible future realizations of the risk factor returns that incorporate both the empirical distribution of historical shocks and the current estimate of market volatility. In practice, the portfolio is typically exposed to multiple risk factors[GM15]. Let

$$\mathbf{r}_t = (r_{1,t}, r_{2,t}, \ldots, r_{J,t})^\top$$

denote the vector of risk factor returns. The filtering procedure is applied component-wise, producing standardized residuals

$$\epsilon_{j,t} = \frac{r_{j,t}}{\sigma_{j,t}}, \qquad j = 1, \ldots, J.$$

The resulting residual vectors

$$\boldsymbol{\epsilon}_t = (\epsilon_{1,t}, \ldots, \epsilon_{J,t})^\top$$

form the basis for generating simulated scenarios of future risk factor returns. These simulated return scenarios are subsequently used as inputs in the risk factor mapping framework, where portfolio value changes are approximated using sensitivities with respect to the underlying risk factors.

2.7 Maximum Likelihood Estimation

Finding the parameters that describe a certain process for some volatility model might not always be trivial. One method of trying to find the parameters is to estimate the parameters using historical data for the process. Maximum Likelihood Estimation (MLE) is an estimation method whose goal is to find the parameters that are most likely to have generated the observed data [Jak21].

The principle of MLE is straightforward. Suppose we have a set of observations $x_1, x_2, \ldots, x_T$ assumed to come from a probability distribution with unknown parameters θ. The likelihood function $L(\theta)$ measures how probable the observed data are for a given choice of θ. MLE selects the value of θ that maximizes $L(\theta)$, i.e., the parameter set that makes the observed data most likely. In practice, it is often easier to maximize the log-likelihood $\ell(\theta) = \log L(\theta)$, because sums are easier to handle than products.

For illustration, consider the derivation of the log-likelihood function for a GARCH(1,1) model under the assumption that the returns are normally distributed with zero mean. Let r_t denote the observed return at time t, and let $\mathcal{F}_{t-1}$ represent the information set available at time $t-1$ (all past returns and estimated variances). The conditional variance σ_t^2 follows the GARCH(1,1) recursion

$$\sigma_t^2 = \omega + \alpha r_{t-1}^2 + \beta \sigma_{t-1}^2.$$

The innovation at time t is r_t itself because we assume a zero conditional mean. Under the normality assumption, the conditional density of r_t given $\mathcal{F}_{t-1}$ is

$$f(r_t \mid \mathcal{F}_{t-1}) = \frac{1}{\sqrt{2\pi\sigma_t^2}} \exp\left(-\frac{r_t^2}{2\sigma_t^2}\right).$$

The likelihood function for the entire sample of T observations is the product of the conditional densities:

$$L(\theta) = \prod_{t=1}^{T} f(r_t \mid \mathcal{F}_{t-1}),$$

where $\theta = (\omega, \alpha, \beta)$ is the vector of GARCH parameters. Taking logarithms yields the log-likelihood function:

$$\ell(\theta) = -\frac{1}{2}\sum_{t=1}^{T}\left(\log(2\pi) + \log(\sigma_t^2) + \frac{r_t^2}{\sigma_t^2}\right).$$

For the log-likelihood function regarding the GJR-GARCH, see A.

2.7.1 Optimization

After obtaining a log-likelihood function, the next step is to find the parameter values that maximize it. In other words finding the estimator

$$\hat{\theta} = \arg\max_{\theta} \ell(\theta).$$

For simple models, closed-form solutions exist. For GARCH-type models, however, the log-likelihood is a non-linear function and therefore numerical optimization methods are required.

There exist different optimization methods, and they differ in things like stability, speed, and computational heaviness. In this project, we use the Nelder–Mead algorithm. It is a derivative-free optimization method that is well-suited for problems where the objective function does not have a derivative or is expensive to calculate. [Die23]. It is a very robust method and can deal with constraints well. The algorithm maintains a simplex of $n + 1$ points in the n-dimensional parameter space and iteratively reflects, expands, contracts, or shrinks the simplex to move towards a minimum. The algorithm is terminated when the size of the simplex falls below a specified tolerance or when the change in the objective function is sufficiently small. Important to note that Nelder–Mead, along with most numerical optimizers, does not guarantee convergence to the global optimum. For settings and algorithm, see B.

2.8 Principal Component Analysis

Many financial portfolios contain a multitude of different assets, whose returns depend on and correlate with the same worldwide financial variables. When the number of financial assets is sufficiently large, this can lead to complications in modeling attempts since the high-dimensional statistical models are hard to apply. To simplify the task of modeling multiple asset returns, one can apply dimension reduction techniques. The goal is to search and find the underlying structure of the asset returns. In the eyes of risk management, the characteristics of this problem lean more towards reducing the number of risk factors to a sufficiently small amount where the remaining risk factors are no longer excessively

correlated. A suitable dimension-reduction technique in this case is to use Principal Components Analysis (PCA)[Tsa10, Chapter 9].

Given a k-dimensional random variable, $r = (r_1, ..., r_k)'$ with covariance matrix Σ_r, a *Principal Components Analysis* is concerned with using a few linear combinations of r_i to explain Σ_r.

PCA can be applied either to the covariance matrix Σ_r or to the correlation matrix ρ_r. Since the correlation matrix is the covariance matrix of the standardized vector

$$r^* = S^{-1}r,$$

where $S = \text{diag}(\sigma_1, \ldots, \sigma_k)$ contains the standard deviations of the components of r, it suffices to present the theory in terms of Σ_r.

Let $w_i = (w_{i1}, \ldots, w_{ik})^\top \in \mathbb{R}^k$ be a weight vector and consider the linear combination

$$y_i = w_i^\top r = \sum_{j=1}^{k} w_{ij} r_j.$$

If r represents returns of k assets, then y_i can be interpreted as the return of a portfolio with weights w_{ij}. Since multiplying w_i by a scalar does not change the portfolio composition up to scale, we impose the normalization constraint

$$w_i^\top w_i = 1.$$

Using standard properties of linear combinations, the variance and covariance of y_i are given by

$$\text{Var}(y_i) = w_i^\top \Sigma_r w_i, \tag{4}$$

$$\text{Cov}(y_i, y_j) = w_i^\top \Sigma_r w_j. \tag{5}$$

The objective of PCA is to construct orthogonal linear combinations of r that successively maximize variance.[Tsa10, pp. 405–408]

$$\text{Cov}(y_2, y_1) = 0.$$

In general, the ith principal component is

$$y_i = w_i^\top r$$

that maximizes $\mathrm{Var}(y_i)$ subject to

$$w_i^\top w_i = 1, \quad \mathrm{Cov}(y_i, y_j) = 0, \quad j = 1, \ldots, i-1.$$

Since Σ_r is symmetric and positive semi-definite, it admits a spectral decomposition. Let

$$(\lambda_1, e_1), \ldots, (\lambda_k, e_k)$$

denote the eigenvalue–eigenvector pairs of Σ_r, ordered such that

$$\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_k \geq 0.$$

Then the principal components are given by

$$y_i = e_i^\top r, \quad i = 1, \ldots, k.$$

Moreover,

$$\mathrm{Var}(y_i) = e_i^\top \Sigma_r e_i = \lambda_i, \tag{6}$$

$$\mathrm{Cov}(y_i, y_j) = e_i^\top \Sigma_r e_j = 0, \quad i \neq j. \tag{7}$$

Thus, the principal components are uncorrelated, and their variances are equal to the eigenvalues of Σ_r. The total variance of r satisfies

$$\mathrm{tr}(\Sigma_r) = \sum_{i=1}^{k} \lambda_i = \sum_{i=1}^{k} \mathrm{Var}(y_i).$$

Hence, the proportion of total variance explained by the ith principal component is

$$\frac{\lambda_i}{\sum_{j=1}^{k} \lambda_j},$$

and the cumulative proportion explained by the first m components is

$$\frac{\sum_{j=1}^{m} \lambda_j}{\sum_{j=1}^{k} \lambda_j}.$$

In practice, a $m \ll k$ means that the cumulative proportion of explained variance is sufficiently large, thereby achieving dimensionality reduction[Tsa10, pp. 409–412].

2.9 Options and Implied Volatility

An option is a financial derivative that gives the holder the right, but not the obligation, to buy (call option) or sell (put option) an underlying asset at a predetermined price on

or before a specified expiration date. The predetermined price is called the strike, and the expiration date is called the maturity date. The price of an option depends on several factors, including the current asset price, the strike price, time to expiration, the risk-free interest rate, and the volatility of the underlying asset. [Ale08]

Implied volatility is a forward-looking measure of the uncertainty of an asset's return. It answers the question: what constant volatility would the underlying asset need to have in order for a given option pricing model to produce the market price of the option? [Hul23] Unlike historical volatility, which is based on past returns, implied volatility is not directly observable; it must be inferred by inverting the pricing formula.

In practice, implied volatility is quoted for a finite set of strike prices and maturities, forming an implied volatility surface. Since the options in a portfolio generally have strikes and maturities that do not coincide exactly with the quoted nodes, an interpolation procedure is required to obtain the implied volatility used for pricing and risk management. The choice of interpolation method affects the smoothness of the surface and the absence of arbitrage opportunities.

2.10 Black–Scholes and Full Revaluation

The Black–Scholes model [BS73] is a European option pricing model. For a call option, the price is given by

$$C = S_0 N(d_1) - K e^{-rT} N(d_2),$$

with

$$d_1 = \frac{\ln(S_0/K) + (r + \sigma^2/2)T}{\sigma\sqrt{T}}, \qquad d_2 = d_1 - \sigma\sqrt{T},$$

where S_0 is the current spot price, K the strike price, T the time to maturity, r the risk-free rate, σ the implied volatility, and $N(\cdot)$ the cumulative distribution function of the standard normal distribution. For a corresponding put option, the price is given by

$$P = K e^{-rT} N(-d_2) - S_0 N(-d_1).$$

In a Monte Carlo simulation for Value at Risk, a straightforward way to obtain the P&L distribution of an option portfolio is to perform full revaluation. For each scenario, compute the required metrics, then evaluate the Black–Scholes formula for every option, and finally aggregate the portfolio value. This method is accurate but computationally expensive because it requires evaluating transcendental functions such as exponential, logarithm, and normal CDF for each option in every scenario. With a large number of options and many scenarios, this becomes extra relevant.

2.11 Greeks and Taylor Approximation

For hedging options, there exist sophisticated procedures that traders use. These procedures involve calculating measures such as delta, gamma, and vega. These are collectively referred to as the Greeks. The Greeks represent different aspects of risk in an option position. When calculating the Greeks, there is often an assumed option pricing model in the background; for example, the Black-Scholes-Merton model, where the volatility is set to the current implied volatility.

Delta is defined as the rate of change in the option price with respect to the price of the underlying asset. Assume a portfolio of options or other derivatives; then, in general, we have:

$$\Delta = \frac{\partial \prod}{\partial S},$$

where $\prod$ is the value of the portfolio and S is the price of the underlying asset. In the same way we define a portfolio's delta, we can define a portfolio's gamma. The gamma (Γ) of a portfolio of options is defined as the rate of change of the portfolio's delta with respect to the price of the underlying asset[Hul21, Chapter 19]. In the same mathematical notation as before, we can define:

$$\Gamma = \frac{\partial^2 \Pi}{\partial S^2}.$$

When Greek letters are calculated, the implied volatility is generally set to constant since the underlying model, such as Black-Scholes-Merton, assumes constant volatility for the underlying asset. Real-world practice proves that this is generally not the case - the volatility of an asset changes over time. The vega of an option, ν, is the rate of change in its value with respect to the volatility of its underlying asset. Mathematically,

$$\nu = \frac{\partial f}{\partial \sigma},$$

where f is the option price and σ is the volatility of the underlying asset, which is normally the option's implied volatility. Finally, when considering options, one must also take into account the interest rate as a risk factor. Rho, ρ, of an option measures the rate of change of its price with respect to the interest rate, r. This Greek can be written as:

$$\rho = \frac{\partial f}{\partial r},$$

where f is the standard notation for the option price, and r is the interest rate. It measures the sensitivity of the value of a portfolio to a change in the interest rate when everything else remains the same. When considering the Greek letters from a risk management perspective, they do not merely serve as hedging metrics; mathematically, they correspond to partial derivatives of the portfolio value function. Consequently, they appear naturally

in a Taylor expansion of the portfolio value with respect to its underlying risk factors. This provides a direct link between option sensitivities and approximations of portfolio P&L, which forms the basis of the delta–gamma VaR framework. This link can be directly derived below using the model of the portfolio value as a function of its underlying risk factors,

$$\Pi = \Pi(S, \sigma, t, r),$$

where S denotes the underlying asset price, σ the volatility, and t time. Consider small changes in the risk factors:

$$\Delta S, \quad \Delta\sigma, \quad \Delta t, \quad \Delta r.$$

A second-order Taylor expansion of the portfolio value yields

$$\Delta\Pi \approx \frac{\partial\Pi}{\partial S}\Delta S + \frac{\partial\Pi}{\partial\sigma}\Delta\sigma + \frac{\partial\Pi}{\partial t}\Delta t + \frac{\partial\Pi}{\partial r}\Delta r + \frac{1}{2}\frac{\partial^2\Pi}{\partial S^2}(\Delta S)^2 + \text{higher-order terms}.$$

Using the standard definitions of the Greeks,

$$\Delta = \frac{\partial\Pi}{\partial S}, \qquad \Gamma = \frac{\partial^2\Pi}{\partial S^2}, \qquad \nu = \frac{\partial\Pi}{\partial\sigma}, \qquad \Theta = \frac{\partial\Pi}{\partial t}, \qquad \rho = \frac{\partial\Pi}{\partial r},$$

the approximation can be written compactly as

$$\Delta\Pi \approx \Delta\,\Delta S + \nu\,\Delta\sigma + \Theta\,\Delta t + \rho\,\Delta r + \frac{1}{2}\Gamma(\Delta S)^2. \tag{8}$$

For sufficiently small changes in the risk factors, higher-order terms may be neglected. The first-order approximation corresponds to the delta approximation, while inclusion of the second-order term in the underlying price yields the delta–gamma approximation. This simple mathematical derivation provides the direct link between the change in P&L for a portfolio and the changes in P&L for its risk factors. This framework is widely used in market risk modeling, particularly in delta–gamma VaR[Ale09, pp. 250–253].

2.11.1 Delta-Gamma-Vega

As discussed earlier, the portfolio value can be modeled using the full Taylor approximation (8), which takes into account all the options' risk factors. In practice, this can be simplified into the delta-gamma-vega approximation, which retains only the terms

$$\Delta\Pi \approx \Delta\,\Delta S + \frac{1}{2}\Gamma(\Delta S)^2 + \nu\,\Delta\sigma,$$

while neglecting higher-order terms as well as the time and interest-rate components. The reasoning for this simplification is both mathematical and economic.

First, the Taylor expansion is intended to approximate small changes in risk factors over short VaR horizons (typically one to ten trading days). Over such short intervals, Δt is small, and its contribution through $\Theta \Delta t$ is generally minor compared to the components' contributions from ΔS and $\Delta \sigma$.

Second, for many equity portfolios, sensitivity to interest rate movements is typically small relative to price and volatility sensitivities. Moreover, short-horizon interest rate changes are generally modest compared to equity price fluctuations, reducing the contribution of $\rho \Delta r$ to overall P&L variability. Consequently, the term $\rho \Delta r$ contributes little to short-horizon P&L variability.

Third, having the gamma term captures the nonlinear effect in the underlying price, which has to be considered. The quadratic term $\frac{1}{2}\Gamma(\Delta S)^2$ introduces curvature into the P&L distribution and corrects an insufficient and purely linear delta approximation. Higher-order derivatives are theoretically present, but contribute terms of order $(\Delta S)^3$ or higher and are therefore negligible for sufficiently small risk factor changes.

The delta–gamma–vega specification represents a good balance between tractability and accuracy. It includes the linear exposure, but also the principal source of nonlinearity in the underlying asset price, and the first-order exposure to volatility risk. Furthermore, it avoids the instability and estimation noise associated with higher-order terms. For short risk horizons, this approximation provides a sufficiently accurate representation of portfolio P&L[Ale09, Chapter 7].

Extended Taylor Approximation

The standard delta–gamma–vega approximation can be extended by considering the portfolio value as a function of a vector of market risk factors rather than only the underlying price and a single volatility input. Let

$$\mathbf{X} = (X_1, \ldots, X_d)^\top$$

denote the relevant risk-factor vector. The portfolio value is then written as

$$\Pi = \Pi(\mathbf{X}).$$

A second-order multivariate Taylor expansion gives

$$\Delta\Pi \approx \nabla\Pi^\top \Delta\mathbf{X} + \frac{1}{2}\Delta\mathbf{X}^\top H_\Pi \Delta\mathbf{X},$$

where $\nabla\Pi$ is the vector of first-order sensitivities and H_Π is the Hessian matrix of second-order sensitivities. Written component-wise, this becomes

$$\Delta\Pi \approx \sum_{i=1}^{d} \frac{\partial \Pi}{\partial X_i}\Delta X_i + \frac{1}{2}\sum_{i=1}^{d}\sum_{j=1}^{d} \frac{\partial^2 \Pi}{\partial X_i \partial X_j}\Delta X_i \Delta X_j.$$

The diagonal elements of the Hessian represent curvature with respect to individual risk factors, while the off-diagonal elements represent cross sensitivities between different risk factors. An important task of the mapping is to understand the cross-risks[Ale09, Chapter 5]. That means understanding how different risk metrics or parameters influence each other. Dealing with these risks can be addressed in many different ways: for example, in static/dynamic hedging, global/local hedging, or the choice of model dynamics. It is important to remember that most of these risks cannot be hedged with market instruments, so it is important to critically examine their behavior and measure the associated risks.

These cross sensitivities are closely related to higher-order Greeks used in option risk management. For instance, vanna is the mixed sensitivity

$$\text{Vanna} = \frac{\partial^2 V}{\partial S \partial \sigma} = \frac{\partial \Delta}{\partial \sigma} = \frac{\partial \nu}{\partial S},$$

and measures the interaction between spot and volatility risk. Volga, also referred to as vomma, is given by

$$\text{Volga} = \frac{\partial^2 V}{\partial \sigma^2} = \frac{\partial \nu}{\partial \sigma},$$

and measures convexity with respect to volatility. The approximation includes the standard spot delta and gamma, first-order sensitivities to FX, interest-rate and volatility principal-component shocks, selected second-order sensitivities to non-spot risk factors, and cross sensitivities between the spot move and the non-spot risk factors. Schematically, the approximation can be written as

$$\Delta\Pi \approx \Delta_S \Delta S + \frac{1}{2}\Gamma_S (\Delta S)^2 + \Delta_{FX}\Delta FX + \sum_j a_j \Delta X_j + \frac{1}{2}\sum_j b_j (\Delta X_j)^2 + \sum_j c_j \Delta S \Delta X_j,$$

where X_j denotes non-spot risk factors such as interest-rate shocks and implied-volatility principal-component shocks. The coefficients a_j, b_j and c_j can be estimated by bump-and-reprice. This structure captures the most relevant nonlinearities and spot–factor interactions for the option portfolio, while avoiding the computational cost and numerical instability associated with estimating the full Hessian matrix.[PLA26]

The motivation for including cross terms is that option values are generally affected by simultaneous movements in several risk factors. In particular, equity option portfolios

are sensitive not only to changes in the underlying spot price and implied volatility separately, but also to their interaction. A negative equity shock is often accompanied by an increase in implied volatility, and the effect on the option value may therefore not be well captured by treating the spot and volatility movements independently.

In a multivariate Taylor expansion, such interaction effects appear as mixed second-order derivatives. For example, the term

$$\frac{\partial^2 \Pi}{\partial S \partial \sigma} \Delta S \Delta \sigma$$

captures the interaction between changes in the underlying price and changes in volatility. In option risk terminology, this type of sensitivity is closely related to vanna. Similarly, second-order volatility terms such as

$$\frac{1}{2} \frac{\partial^2 \Pi}{\partial \sigma^2} (\Delta \sigma)^2$$

are related to volga, or vomma, and capture convexity with respect to volatility.

Including such terms can make the Taylor approximation closer to full revaluation, provided that the sensitivities are estimated accurately and the scenario shocks are not too large. The benefit is that important nonlinear and interaction effects can be captured without repricing the full portfolio under every scenario. However, the drawback is that higher-order and cross sensitivities are more difficult to estimate robustly. They may introduce numerical noise, depend on the selected bump size, and increase the computational cost of the sensitivity calculation. Therefore, a practical implementation often uses a reduced set of second-order and cross terms rather than the full Hessian matrix.

2.12 Backtesting

Assessing the accuracy of a VaR model is not straightforward. A single day's outcome reveals little about the model's overall reliability. However, by examining a long sequence of daily outcomes and comparing them with the corresponding VaR forecasts, systematic patterns can emerge. Several statistical tests have been developed to evaluate VaR models. These tests typically focus on two essential properties:

1. Unconditional coverage = whether the number of violations matches the expected.
2. Independence = whether violations occur independently over time.

The following subsections describe the tests used in this project: the Kupiec test for unconditional coverage and the Christoffersen–Pelletier duration test for independence.

2.12.1 Kupiec Test

The Kupiec test [Kup95] evaluates whether the observed number of Value-at-Risk violations over a given time period is consistent with the expected violation rate. Under the null hypothesis, violations follow a binomial distribution with probability $p = 1 - \alpha$.

Let M denote the number of observations and x the number of observed violations. The observed violation frequency is given by $\hat{p} = x/M$.

The Kupiec test is based on a likelihood ratio statistic:

$$\mathrm{LR}_{\mathrm{POF}} = -2 \ln \left(\frac{(1-p)^{M-x} \, p^x}{(1-\hat{p})^{M-x} \, \hat{p}^x} \right).$$

Under the null hypothesis, the test statistic follows a chi-squared distribution with one degree of freedom:

$$\mathrm{LR}_{\mathrm{POF}} \sim \chi^2(1).$$

If the value of the test statistic is sufficiently large, the null hypothesis is rejected, indicating that the VaR model does not produce the correct proportion of violations.

2.12.2 Christoffersen–Pelletier Duration Test

The Christoffersen–Pelletier duration test [CP04] evaluates whether Value-at-Risk violations occur independently over time by examining the durations between consecutive violations.

Let D_i denote the number of days between two consecutive VaR violations. Under a correctly specified VaR model, violations occur independently with probability $p = 1-\alpha$, implying that durations should follow an exponential distribution. Christoffersen and Pelletier test is based on modeling the durations using a Weibull distribution,

$$f(D; a, b) = a^b b D^{\,b-1} \exp\Big[-(aD)^b\Big],$$

where $a > 0$ is a scale parameter and $b > 0$ is a shape parameter.

The key parameter is the Weibull shape parameter b. Under the null hypothesis of independently occurring violations, the Weibull distribution reduces to the exponential distribution, corresponding to $b = 1$.

Values of $b < 1$ indicate violation clustering, meaning that violations tend to occur closer together than expected under independence. Conversely, values of $b > 1$ suggest

that violations occur more regularly than expected. Therefore, the estimated value of b provides a measure of whether VaR violations are independently distributed over time, with deviations from $b = 1$ indicating potential model misspecification.

3 Method

3.1 Data Description

Our model is trained and validated on historical data. We choose a training period starting at 2006-05-01 and ending at 2021-12-31. The validation period starts at 2022-01-01 and ends at 2026-01-02. This means a total of fifteen 15 years and 8 months of training data and four years for validation. Using these dates makes the validation period long enough to be able to analyze results, while also giving the model access to a lot of training data. With these dates, the validation period includes both calm moments of the market as well as an extreme volatility market in most of 2022; this is particularly useful when evaluating our model in different types of markets. The training data also includes both calm periods of small market movements as well as some more intense periods such as the 2008 and 2020 market crashes.

For the validation, we choose to create a portfolio with real market-price data. Our model is constructed to be able to deal with different kinds of assets and not only options. However, for evaluation, we decided to construct a full option portfolio of SPX options. This includes a mix of European call and put options. The mix consists of two options bought each quarter. The maturities of the options start at each quarter as well and decrease according to the maturity date. Strike prices were chosen around the money in terms of the current price of the underlying index value. The moneyness will then move depending on underlying movements until the maturity date. Using this portfolio gives us the possibility of evaluating how our model deals with options specifically, which is the most interesting asset due to its nonlinear price properties. This portfolio also tests the model on different maturity and moneyness levels, which is something we want the model to be able to do. The final reason for using this portfolio structure is that the SPX index and SPX options are commonly used assets for research but also for trading. Therefore, the results can be compared with other academic work as well as being a realistic portfolio.

Most of the data used was collected from Bloomberg. This includes historical option prices, estimated implied volatilities for some maturities and strikes. Index values and OIS rates for some maturities.

3.2 Risk factor extraction

Keeping in mind that the idea of the mapping should be mainly based on non-linear instruments such as European options, we first need to consider the risk factors that affect option prices. Other instruments such as futures, forwards, and swaps were also considered. The common denominator for these is that they all depend on an underlying

asset. However, the degree of nonlinearity differs across instruments. Futures and forwards are often approximately linear in the underlying, while European options exhibit pronounced nonlinear exposure through delta, gamma, and volatility sensitivity. This is particularly interesting in a framework where P&L is directly approximated, rather than fully calculated using a pricing framework such as the Black-Scholes formula. At the end, we opted for using European call/put options for the portfolio, which is sufficient given the non-linearity of the derivative and that there exists a clear framework to calculate the "real" price when using full revaluation. To extend the whole pipeline, the work is also complemented by having different bond types, stocks, and Forex spots allowed in the given portfolio. This totals a more complete, useful, and realistic modeling framework where many common assets and derivatives can be VaR-modeled without issues.

Looking at the risk factors for these assets and starting with European call and put options, we can see that there are a few factors that directly drive the prices of these. Option prices are of course directly affected by the price of the underlying asset. Secondly, a single option inherits an implied volatility metric. Given that the framework should work for many different option assets and many options on the same underlying with different values of moneyness and time to maturity, it is highly necessary to not only look at single implied volatility metrics, but also at the whole volatility surface for the underlying asset in question. Thirdly, options factor in changes in the yield market. Directly from the Black-Scholes formula, it is evident that the prices are also driven by the risk-free rate. What specific yield rates and tenors to label the risk-free rate is dependent on the specific underlying in question. For SPX options, the risk-free rate is suitably linked to different OIS rate tenors. Lastly, one could also argue that option prices have a certain dependency on the time factor. However, for this work, time is not treated as a market risk factor. The one-day passage of time is deterministic and can either be accounted for through repricing at the next valuation date or omitted from the stochastic risk-factor shock vector. In this project, the stochastic VaR scenarios are driven by market risk factors: spot, FX, rates, and implied volatility.

Dividend-yield or forward-curve risk is also not modeled as a separate risk factor. Its effect is either implicitly embedded in the observed option marks and implied-volatility data or treated as outside the scope of the risk-factor set. This is a limitation for equity-index option portfolios, since changes in forward levels may affect option prices. Below we go through specifics on what risk factors were modeled and supported in the risk-factor mapping framework.

Table 1: Supported instrument types and associated risk factors in the framework

Instrument type	**Main market inputs**	**Extracted risk factors**	**Role in P&L model**
Stock position	Equity spot level	Spot log return	Linear exposure to the underlying asset
FX spot position	FX spot level	FX log return	Currency exposure and FX translation risk
Zero-coupon bond	Yield curve	Interest-rate tenor changes	Discounting and yield-curve sensitivity
Fixed-coupon bond	Yield curve and coupon schedule	Interest-rate tenor changes	Price sensitivity to curve movements
European option	Underlying spot, volatility surface, yield curve	Spot return, volatility-surface shocks, rate changes	Nonlinear exposure through delta, gamma, vega and rate sensitivity

The purpose of the extraction step is to transform observable market data into a stationary set of daily risk-factor changes. These changes form the input matrix for the volatility model and the filtered historical simulation procedure. The simulated shocks are later mapped back to market levels and used both in the Taylor approximation and in the full-revaluation benchmark. Collecting all extracted risk factors, the one-day risk-factor vector is written as

$$X_t = \left(r_t^S, r_t^{FX}, \Delta r_t^{(1)}, \ldots, \Delta r_t^{(K)}, y_{1,t}^{vol}, \ldots, y_{m,t}^{vol}\right)^\top,$$

where r_t^S denotes the equity spot log return, r_t^{FX} the FX log return, $\Delta r_t^{(k)}$ the absolute change in interest-rate tenor k, and $y_{j,t}^{vol}$ the j-th principal-component score of the implied-volatility surface changes.

A simulated spot shock x^S is applied multiplicatively as

$$S' = S\exp(x^S),$$

while an FX shock is applied analogously,

$$F' = F \exp(x^{FX}).$$

Interest-rate shocks are applied additively to each tenor,

$$r^{(k)\prime} = r^{(k)} + \Delta r^{(k)}.$$

For implied volatility, PCA shocks are first mapped back to node-level volatility changes, after which node volatilities are updated multiplicatively,

$$\sigma_i' = \sigma_i \exp(\Delta v_i).$$

3.2.1 Equity spot risk factors

Equity spot levels are included as risk factors because they directly affect both stock positions and equity options. For a stock position, the exposure to the spot level is approximately linear. For an option position, the exposure is nonlinear and enters through delta and gamma in the Taylor approximation. For an equity spot level S_t, the extracted risk factor is the daily log return,

$$r_t^S = \log\left(\frac{S_t}{S_{t-1}}\right).$$

Log returns are used because they represent relative price changes and are additive over time. In the empirical implementation, the main equity spot risk factor is the SPX index level.

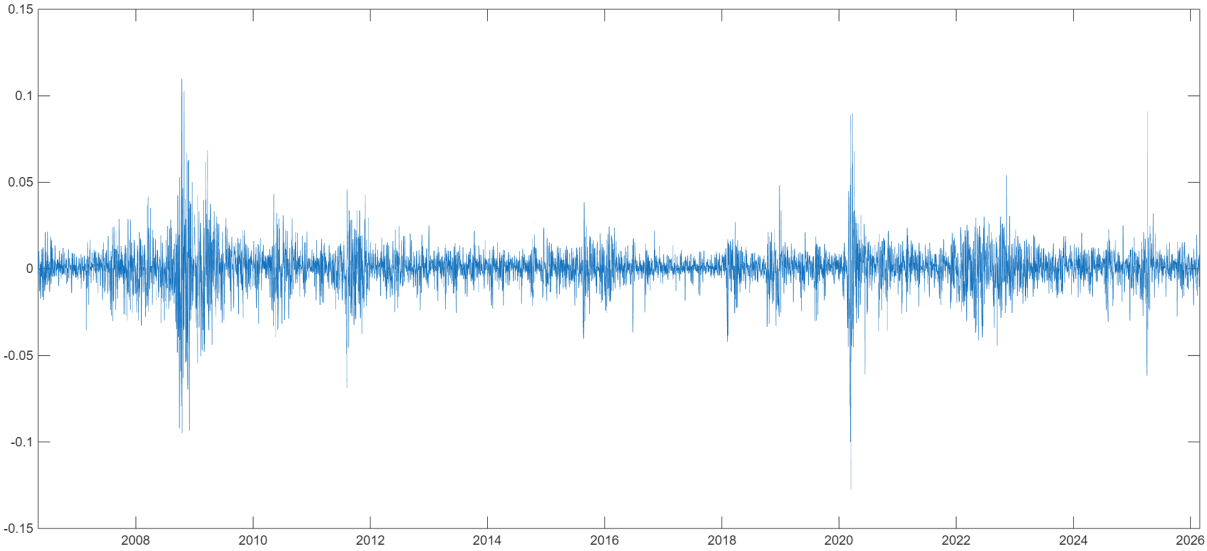


Figure 3: SPX index daily log returns used as the equity spot risk factor.

3.2.2 FX-spot risk factors

FX spot rates are included to support direct FX spot positions and to capture currency translation effects when instruments are denominated in a different currency than the reporting currency. Similar to equity spot levels, FX spot rates are transformed into log returns. For an FX rate F_t, the extracted risk factor is

$$r_t^{FX} = \log\left(\frac{F_t}{F_{t-1}}\right).$$

In the empirical implementation, the EURUSD exchange rate is included as the FX risk factor.

3.2.3 Interest-rate risk factors

Interest-rate risk factors are required for fixed-income instruments and for option pricing through discounting. Zero-coupon and fixed-coupon bonds are directly sensitive to changes in the yield curve, while European option prices depend on the risk-free rate used in the pricing model. The yield curve is represented by a finite set of tenor points. For each tenor k, the extracted risk factor is the absolute rate change,

$$\Delta r_t^{(k)} = r_t^{(k)} - r_{t-1}^{(k)}.$$

Absolute changes are used rather than log returns since interest rates can be close to zero and are naturally modeled in rate units. In the empirical implementation, the risk-free curve for SPX options is represented using USD OIS rates.

Table 2: Interest-rate risk-factor extraction

Item	**Specification**
Raw market input	Yield-curve tenor rates
Example in implementation	USD OIS curve
Included tenors	1m, 3m, 6m, 1y, 2y, 3y, 4y, 5y
Transformation	Absolute daily rate change $r_t^{(k)} - r_{t-1}^{(k)}$
Risk-factor role	Bond pricing, discounting and option rate sensitivity
Scenario interpretation	Shift in individual yield-curve tenor points

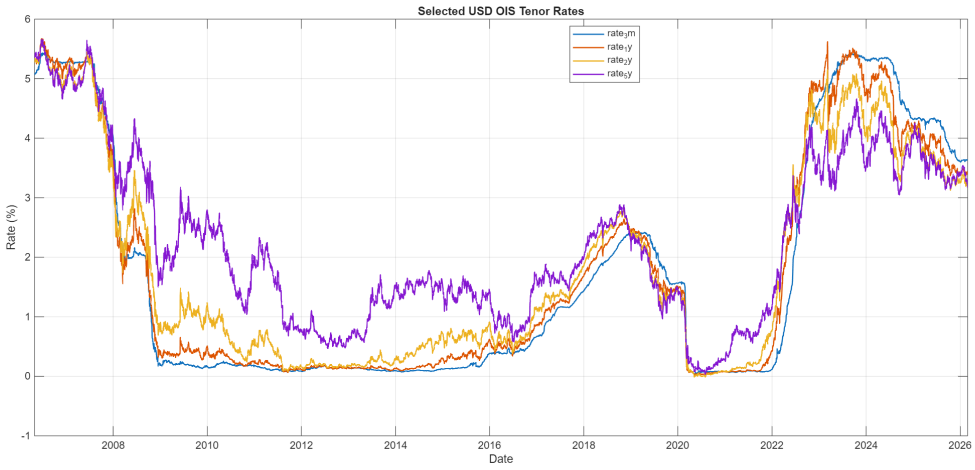


Figure 4: Selected USD OIS tenor levels used as raw inputs for interest-rate risk-factor extraction over the period 2006-2026.

3.2.4 Implied-volatility surface risk factors

Implied volatility is a central risk factor for option portfolios. Since we are particularly interested in portfolios containing a magnitude of options, both on the same underlying, but also on different underlying assets, which all have different strikes and maturities exposed to different implied volatilities, it is in our interest to model the whole implied-volatility surface as a risk factor. Single implied-volatility time series are simply not sufficient for the work; instead, we treat different strikes/maturities with the implied volatility time series as nodes on the surface. We choose a suitable number of these nodes at reasonable strikes/maturities and extract the time series of implied volatility metrics over the given time period. Doing it this way simplifies the mapping framework since each European option in any given portfolio can be directly interpolated and linked to the volatility surface. Let $\sigma_{t,i}$ denote the implied volatility at node i on day t. The extracted node-level risk factor is the daily log change,

$$\Delta v_{t,i} = \log\left(\frac{\sigma_{t,i}}{\sigma_{t-1,i}}\right).$$

The full set of node changes describes movements in the implied-volatility surface. Since this creates a high-dimensional risk-factor set, the volatility-surface changes are later reduced using principal component analysis. The resulting volatility principal components are used as risk factors in the scenario generation and Taylor approximation.

Table 3: SPX implied-volatility surface grid

Dimension	**Nodes**
Maturities	30d, 60d, 3m, 6m, 12m, 24m
Moneyness levels	80, 90, 95, 97.5, 100, 102.5, 105, 110, 120
Total number of volatility nodes	54
Node transformation	Daily log volatility changes
Dimension reduction	PCA on volatility-surface changes
Number of volatility PCs	4

The volatility surface is therefore represented by a finite vector of node volatilities. This node representation is suitable for risk-factor modeling, since historical changes can be computed for each node. However, individual options in the portfolio generally do not lie exactly on the selected grid. The node representation must therefore be combined with an interpolation operator when option prices are computed.

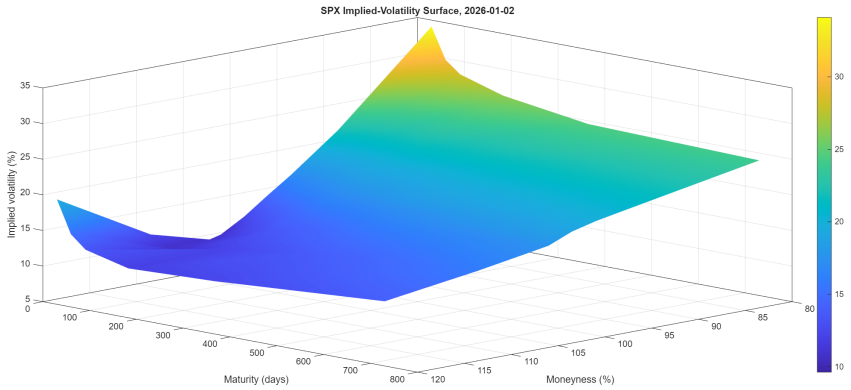


Figure 5: Example SPX implied-volatility surface across moneyness and maturity nodes.

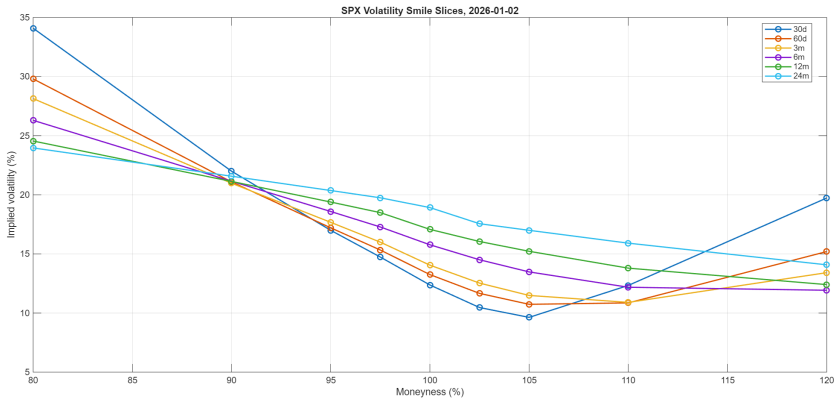


Figure 6: Different smile slices of the SPX implied volatility surface.

3.2.5 Principal Components Analysis on volatility surface

Having a multitude of implied volatility nodes on a grid is advantageous in the sense that it provides well-approximated interpolated option-specific implied volatility metrics. However, using all provided nodes as risk factors is often not optimal due to the high correlation between them. A market shock that affects one part of the surface often affects nearby maturities and moneyness levels as well. Thus, we do not need all volatility nodes to explain the presence of the surface. We are interested in reducing the high dimensionality of the volatility surface into a few interpretable, systematic factors. A viable option in this case is to use principal components analysis on the volatility surface, discussed in 2.8. We choose to reduce the surface to four components in this case due to the high number of used nodes. Normally, one could use three components instead, which would result in the classic level/skew/curve. Adding a fourth component would also add a certain maturity-dependent rotation, which would ideally push short-term implied volatility up while pulling long-term implied volatility down. Let

$$\Delta v_t = (\Delta v_{t,1}, \ldots, \Delta v_{t,N})^\top$$

denote the vector of node-level log-volatility changes on day t, where $N = 54$ in the SPX implementation. The node changes are defined as

$$\Delta v_{t,i} = \log\left(\frac{\sigma_{t,i}}{\sigma_{t-1,i}}\right).$$

Before PCA is applied, each node-change series is centered and standardized,

$$\widetilde{\Delta v}_{t,i} = \frac{\Delta v_{t,i} - \mu_i}{s_i},$$

where μ_i and s_i denote the historical mean and standard deviation of node i. Standardization is used so that the PCA is not dominated by nodes with larger absolute volatility variation. PCA is then applied to the standardized matrix of volatility-surface changes. If w_j denotes the j-th principal-component loading vector, the corresponding volatility principal-component score is

$$y_{j,t}^{vol} = w_j^\top \widetilde{\Delta v}_t.$$

The first m principal-component scores are retained and used as volatility risk factors in the scenario generation. In the empirical implementation, $m = 4$ components are used. The volatility-PC representation is also used when shocks are mapped back to the volatility surface. A simulated or bumped shock in principal component j is transformed back to node-level log-volatility changes through the PCA loadings. For a pure PC bump

of size h_j, the corresponding node-level change is

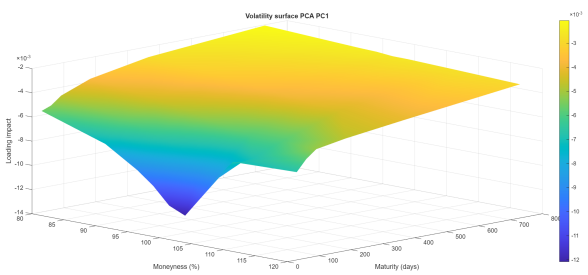
$$\Delta v_i = h_j w_{i,j} s_i,$$

where $w_{i,j}$ is the loading of node i on component j, and s_i is the standard deviation used in the PCA scaling. The shocked node volatility is then obtained,

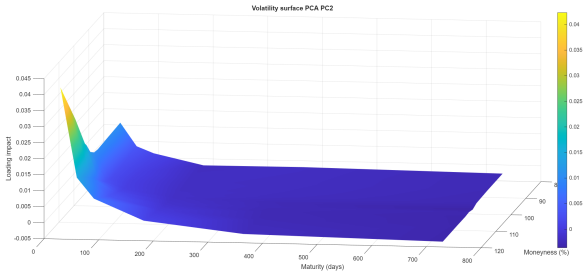
$$\sigma_i' = \sigma_i \exp(\Delta v_i).$$

Component	Dominant interpretation	Explained variance	Cumulative explained variance
PC1	Level shift	84.19%	84.19%
PC2	Moneyness/skew movement	3.124%	87.32%
PC3	Curvature/smile movement	2.856%	90.17%
PC4	Maturity-dependent rotation	1.583%	91.76%

Table 4: Explained variance of the retained SPX implied-volatility surface principal components.

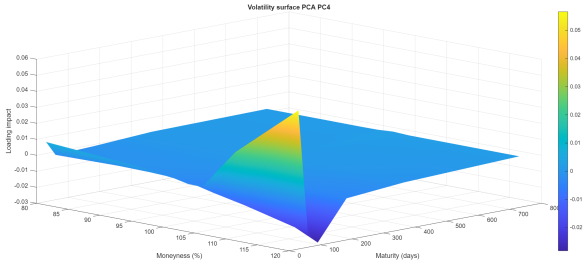


(a) PC1



(b) PC2

(c) PC3



(d) PC4

Figure 7: Loading surfaces of the four retained principal components of the implied-volatility surface. Each loading vector is reshaped back to the maturity–moneyness grid. Together, the components describe the dominant systematic movements of the volatility surface.

3.2.6 Volatility Surface Interpolation

Having only a discrete grid of implied volatility nodes with their corresponding PCA components is rarely enough in the context of risk factor mapping. The issue with this is that, in reality, the case is usually that portfolios inherit options that have combinations of moneyness and maturities which do not exist on the volatility grid. In this case, it would seem impossible to obtain the implied volatility for the given option[Reb04]. Thus, it would require an interpolation scheme to obtain the implied volatility. Let Σ_t denote the vector of implied-volatility nodes at time t. For an option with time to maturity τ and moneyness M, the implied volatility used in the pricing formula is written as

$$\sigma_{\text{opt},t} = \mathcal{I}(\Sigma_t; \tau, M),$$

where $\mathcal{I}$ denotes the volatility-surface interpolation operator. This operator maps the discrete volatility-surface representation to the option-specific implied volatility.

One simple interpolation approach is to use bilinear interpolation. The idea is extremely simple: every four nodes on the grid closest to the given unknown implied volatility node are assigned a weight. Depending on how close the option combination (M;T) is to the closest grid nodes, these are assigned larger or smaller weights. The implied volatility of the known nodes is distributed to the unknown IV node according to the weights. Bilinear interpolation is simple, transparent, and local, but it may generate surfaces that are less smooth across maturity and moneyness.

The second approach, used in the main empirical implementation, is cubic-spline interpolation. The procedure is applied in two steps. First, for each maturity slice, a one-dimensional interpolation is performed across the moneyness dimension. This gives an interpolated volatility value at the option's moneyness for each quoted maturity. Second, these interpolated values are themselves interpolated across the maturity dimension to obtain the final option-specific implied volatility. Schematically,

$$\sigma_{\text{opt},t} = \mathcal{I}_T \left(\mathcal{I}_M(\Sigma_t; M); \tau \right),$$

where $\mathcal{I}_M$ denotes interpolation in the moneyness direction and $\mathcal{I}_T$ denotes interpolation in the maturity direction. More specifically, the cubic spline method divides the grid into intervals between the moneyness nodes. Then, for each interval i, a unique third-degree polynomial function $S_i(M)$ is defined. For each interval, when you plug in the left node M_i into the equation, the result must equal that node's volatility:

$$S_i(M_i) = d_i = \sigma_i$$

When you plug in the right node (M_{i+1} into the same equation, this must equal the next volatility (σ_{i+1}). Then, to prevent the volatility from having too sharp and unnatural corners at the internal nodes, the slopes are forced to match exactly where the two intervals meet by having:

$$S_i'(M_{i+1}) = S_{i+1}'(M_{i+1}).$$

The same is done with the second derivative. The boundary conditions in our case are set to Natural Cubic Spline conditions, where we assume the curve straightens out at the edges: $S_1''(M_1) = 0$ and $S_{n-1}''(M_n) = 0$. Finally, this system of equations are solved, where we obtain the coefficients of the third degree polynomial and evaluates:

$$\sigma_{target} = a_i(M_{target} - M_i)^3 + b_i(M_{target} - M_i)^2 + c_i(M_{target} - M_i) + d_i.$$

For points outside the volatility grid, the queried maturity and moneyness are restricted to the boundary of the available grid. This avoids extrapolating the implied-volatility surface far outside the region supported by market data. The interpolated volatility is also floored at a small positive value to ensure that the pricing input remains well-defined.

The same interpolation operator is used consistently in the base valuation, bump-and-reprice sensitivity calculation and full-revaluation benchmark. This consistency is important because otherwise differences between Taylor approximation and full revaluation could be caused by different volatility interpolation assumptions rather than by the Taylor approximation itself.

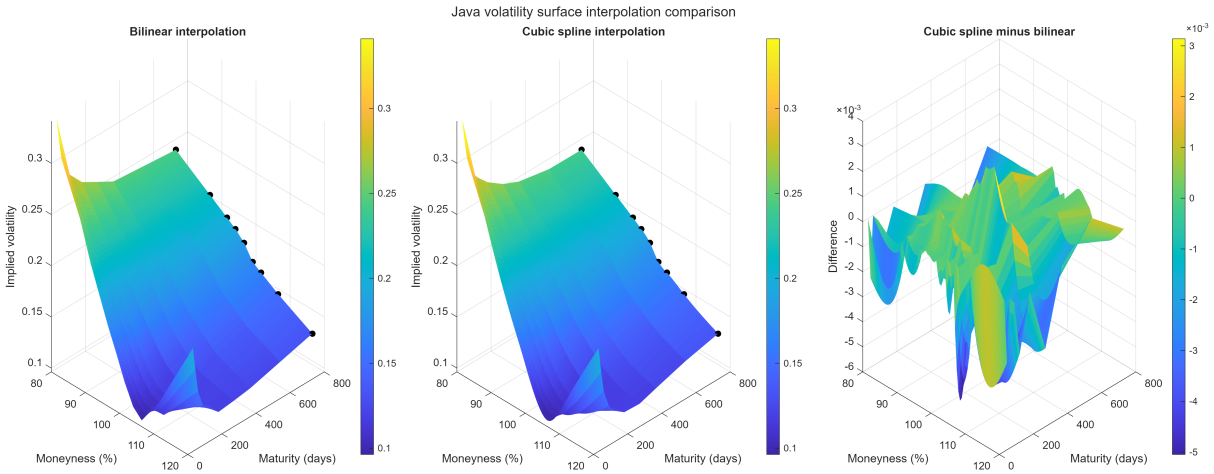


Figure 8: Comparison between bilinear and cubic-spline interpolation of the SPX implied-volatility surface. The node points represent observed volatility grid values. The cubic-spline interpolation produces a smoother surface, while the difference plot shows where the two interpolation methods deviate.

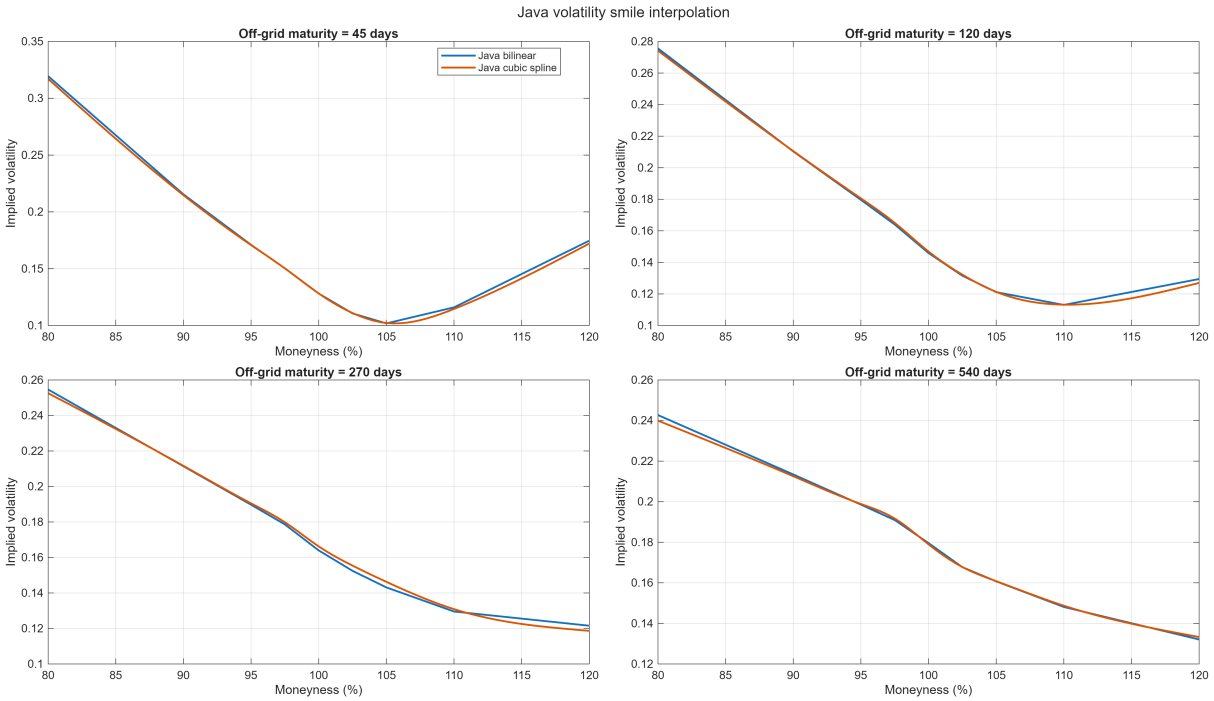


Figure 9: Interpolated implied-volatility smile slices at off-grid maturities. These slices illustrate how bilinear and cubic-spline interpolation behave when the option maturity does not coincide with a quoted volatility-surface maturity node.

Figure 8 compares bilinear and cubic-spline interpolation on the SPX implied-volatility surface. Both methods reproduce the observed volatility nodes, but they differ in how the surface is constructed between the nodes. Bilinear interpolation produces a piecewise linear surface, while cubic-spline interpolation produces a smoother surface across both moneyness and maturity.

This smoothness is particularly relevant for the bump-and-reprice procedure. Since sensitivities are estimated using finite differences, artificial kinks in the interpolation scheme may lead to unstable sensitivity estimates[Hul21, Chapter 19]. Cubic-spline interpolation is therefore used in the main implementation to obtain smoother option-specific implied volatilities. However, spline interpolation may introduce local overshooting near sparse regions or boundaries of the grid, which is why the interpolated volatilities are kept within reasonable positive bounds.

3.3 Volatility Modeling

We assume that daily log-returns have zero mean and follow either a Normal or a Student-t distribution, focusing solely on their conditional volatility. This no-drift assumption is a common short-term approximation for financial returns. The volatility is allowed to be time-varying and is updated daily. Four different methods for modeling conditional volatility were tested:

1. Exponentially Weighted Moving Average (EWMA) with fixed decay factor $\lambda = 0.94$.
2. Standard GARCH(1,1).
3. GJR-GARCH(1,1).
4. Self-switching two-state GJR-GARCH.

The EWMA is the most basic method and serves as a baseline volatility model. The standard GARCH and GJR-GARCH models estimate their parameters uniquely for each risk factor via maximum likelihood under the Student-t assumption, and the conditional variance is updated daily using the estimated recursion.

The self-switching GJR-GARCH method uses a two-state volatility approach with a smooth transition between regimes. First, a GARCH(1,1) model is estimated on the entire return series of a given risk factor. Its one-step-ahead conditional volatility v_t serves as a state indicator. The training period is split into low-volatility and high-volatility subsets based on whether v_t exceeds a threshold θ. The threshold is chosen as a quantile of the $\{v_t\}$ series over the training period. Separate GJR-GARCH(1,1) models are then estimated on each subset, yielding two parameter sets:

$$(\omega_L, \alpha_L, \gamma_L, \beta_L) \quad \text{and} \quad (\omega_H, \alpha_H, \gamma_H, \beta_H).$$

To avoid abrupt jumps when the state switches, a soft mixing weight w_t is defined as a linear function of v_t over an interval of width $2\theta\delta$ centered at θ, where δ is a relative transition width. Precisely,

$$w_t = \begin{cases} 0, & v_t \le \theta(1-\delta), \\ \dfrac{v_t - \theta(1-\delta)}{2\theta\delta}, & \theta(1-\delta) < v_t < \theta(1+\delta), \\ 1, & v_t \ge \theta(1+\delta). \end{cases}$$

Figure 10 illustrates this weighting function for $\theta = 1$ and $\delta = 0.2$. At each time t, the GJR parameters are linearly blended:

$$\begin{aligned} \omega_t &= (1-w_t)\omega_L + w_t\omega_H, \\ \alpha_t &= (1-w_t)\alpha_L + w_t\alpha_H, \\ \gamma_t &= (1-w_t)\gamma_L + w_t\gamma_H, \\ \beta_t &= (1-w_t)\beta_L + w_t\beta_H. \end{aligned}$$

The conditional variance for the next day is then computed using these blended parameters:

$$\sigma_{t+1}^2 = \omega_t + \Big(\alpha_t + \gamma_t \cdot I_{\{r_t<0\}}\Big) r_t^2 + \beta_t \, \sigma_t^2,$$

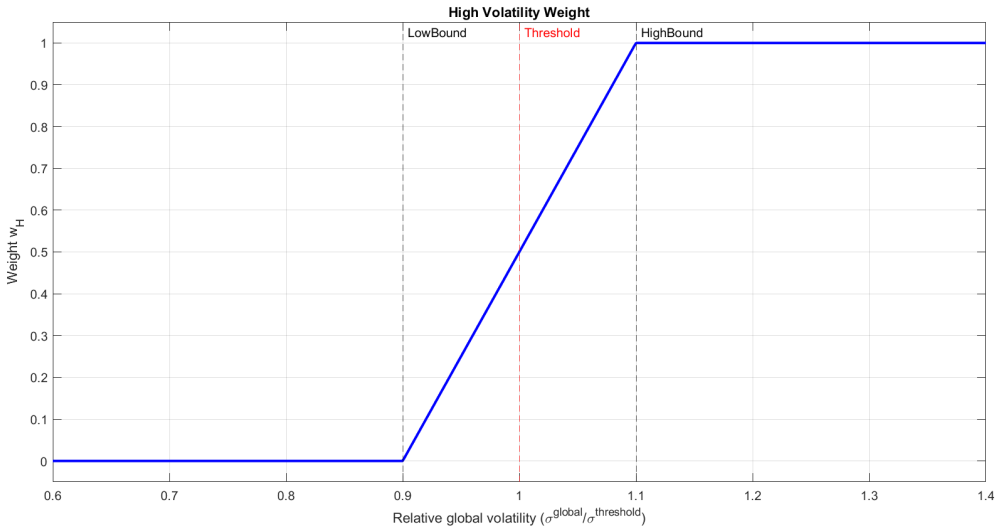


Figure 10: Soft transition weight function for the self-switching GARCH model with transition width $\delta = 0.2$ and threshold $\theta = 1$ (normalized). The weight increases linearly from 0 to 1 over the interval $[\theta(1-\delta),\ \theta(1+\delta)]$, providing a smooth blend between low-vol and high-vol GJR parameters.

Table 5 shows the four models along with their equations and parameters.

Table 5: Overview of volatility models tested

Model	**Volatility equation (for a single factor)**	**Parameters**
EWMA	$\sigma_t^2 = \lambda\sigma_{t-1}^2 + (1-\lambda)r_{t-1}^2$	λ (decay factor)
GARCH(1,1)	$\sigma_t^2 = \omega + \alpha r_{t-1}^2 + \beta\sigma_{t-1}^2$	ω, α, β
GJR-GARCH(1,1)	$\sigma_t^2 = \omega + (\alpha + \gamma I_{r_{t-1}<0})r_{t-1}^2 + \beta\sigma_{t-1}^2$	$\omega, \alpha, \gamma, \beta$
Self-switching GARCH (GJR)	Global GARCH determines state; low- and high-vol GJR parameters blended smoothly.	Two GJR sets, quantile, transition width δ

All methods except EWMA used maximum likelihood estimation to find the parameters for each volatility model for each risk factor. A Nelder-Mead optimizer was used as the optimization tool to minimize the negative log-likelihood function over the parameter space. All training data were used in the estimation of the parameters.

3.4 Covariance Modeling

After obtaining a time series of time-varying conditional volatilities for each risk factor using any of the volatility models described in the previous section, the standardized residuals

$$\epsilon_{j,t} = \frac{r_{j,t}}{\sigma_{j,t}}$$

are computed for all training days. These residuals are approximately independent and identically distributed with zero mean and unit variance, and form the basis for Filtered Historical Simulation, which is used to model the joint distribution of all risk factors simultaneously, explaining their covariance.

The number of risk factors can be large, depending on the portfolio's assets. Modeling their covariance is challenging because covariances may be time-varying and may exhibit complex non-linear patterns that are difficult to capture with parametric models such as covariance matrices or copulas. FHS offers a simple, non-parametric alternative: it resamples historical residual vectors as a whole, thereby preserving the entire empirical dependence structure, including tail dependencies and asymmetries, provided they have occurred in the past. Moreover, FHS is computationally light and easy to implement.

Three strategies based on FHS were used to generate scenarios for the residuals:

1. Standard FHS – samples residuals uniformly from all historical days, ignoring any market state.

2. Three-state hard sampling – classifies each historical day into one of three market states (STABLE, NORMAL, RISKY) and, for the current day, draws residuals only from the state that matches the current state.

3. Three-state soft sampling – similar to hard sampling, but instead of deterministically picking the current state, the state for each scenario is sampled probabilistically using a softmax function based on the current market metric.

3.4.1 State classification

The three states (Stable, normal, risky) are defined using a volatility-based metric. For each day, we compute

$$\text{metric}_t = \sum_j \frac{\sigma_{j,t}}{\text{median}(\sigma_j)},$$

where $\sigma_{j,t}$ is the conditional volatility of risk factor j at time t and $\text{median}(\sigma_j)$ is its median over the training period. This metric captures the current level of forecasted turbulence relative to typical conditions. Optionally, the metric can be smoothed by

a rolling average over the last k days to reduce noise and make state transitions more persistent.

The historical distribution of the metric is split into three regions using two quantile thresholds, for example 0.4 and 0.9. Accordingly,

- Days with metric $\leq$ lower threshold $\rightarrow$ `Stable`
- Days with lower $<$ metric $\leq$ upper threshold $\rightarrow$ `Normal`
- Days with metric $>$ upper threshold $\rightarrow$ `Risky`

The same quantile thresholds are then used to classify the current day's metric into a current state.

3.4.2 Softmax probabilities for soft sampling

When soft sampling is used, the state for each scenario is not deterministic; instead, a state is drawn randomly with probabilities that depend on the distance of the current metric to the thresholds. Let m be the current smoothed metric, L the lower threshold, and U the upper threshold. Define three centers:

$$c_{\text{STABLE}} = \frac{L}{2}, \qquad c_{\text{NORMAL}} = \frac{L+U}{2}, \qquad c_{\text{RISKY}} = U \times 1.5.$$

The unnormalized score for state k is

$$s_k = \exp\left(-\frac{(m-c_k)^2}{T}\right),$$

where T is a temperature parameter that controls the softness. The final probability for state k is then

$$p_k = \frac{s_k}{\sum_i s_i}.$$

A random number is drawn from a uniform distribution, and the corresponding state is selected according to these probabilities. The temperature T determines how sharply the probabilities change around the centers: a low temperature yields nearly deterministic assignment (hard switching), while a higher temperature gives a smoother mixture.

The soft sampling approach avoids abrupt changes in the bootstrap distribution when the market metric fluctuates near the threshold boundaries, leading to more stable VaR forecasts.

3.5 Risk-factor mapping and P&L calculations

Once the risk factor scenarios have been generated, the next natural step is to perform the mapping itself in order to obtain the portfolio P&L. The purpose of the risk-factor mapping is to avoid repricing the full portfolio under every simulated market state. Instead, the portfolio value is approximated locally around the current market state using sensitivities with respect to the selected risk factors. The theory behind obtaining P&L by using the reduced second-order multivariate Taylor expansion of the portfolio value function is previously discussed in section 2.12 and is of the full form:

$$\Delta\Pi \approx \Delta_S \Delta S + \frac{1}{2}\Gamma_S(\Delta S)^2 + \Delta_{FX}\Delta FX + \sum_j a_j \Delta X_j + \frac{1}{2}\sum_j b_j(\Delta X_j)^2 + \sum_j c_j \Delta S \Delta X_j,$$

where X_j denotes non-spot risk factors such as interest-rate shocks and implied-volatility principal-component shocks, and Π denotes the portfolio value. The risk factor shocks are of course already obtained, so the natural next step is to find the coefficients a_j, b_j and c_j. These coefficients can be seen as closely related to the actual Greeks of each risk factor; they are the actual sensitivities. Of course, because we use the extended version of the Taylor approximation, the coefficients a, b and c can not be directly referred to as "the Greeks". This version of the Taylor expansion is the reduced Hessian Matrix of the portfolio value function; reduced in the sense that the yield rate cross-terms are not presented.

The first two terms represent the standard delta-gamma approximation with respect to the underlying equity spot. The coefficient Δ_S measures the linear exposure to the underlying, while Γ_S captures the curvature of the option portfolio with respect to spot movements. The inclusion of Γ_S is an important addition for an option-heavy portfolio since it captures the non-linearity of the underlying asset price and option-price relationship. The term $\Delta_{FX}\Delta FX$ represents the first-order sensitivity to FX movements. This term is included to account for currency exposure and translation effects when instrument values are expressed in a reporting currency different from the instrument currency. In the current implementation, FX exposure is treated linearly.

The coefficients a_j are the sensitivities with respect to the implied volatility principal components as well as the sensitivities for each yield curve principal component. That means for $j = 1, ..., 4$, these are the sensitivities that directly corresponds to $IV_{PC_1}, ...IV_{PC_4}$. In theory, this can be closely compared to the vega Greeks since it is directly related to implied volatility. For $j = 5, ...7$, these coefficients are the sensitivities related to the principal components of the yield curves $r_{PC_1}, ...r_{PC_3}$.

The coefficients b_j represent the second-order sensitivities related to the quadratic properties of the implied volatility and yield curve principal components. These terms capture curvature effects that are not explained by the linear sensitivities alone. For volatility principal components, the b_j-terms are closely related to volga or vomma effects, since they measure convexity with respect to changes in implied volatility. Including these terms improves the approximation when volatility shocks are large or when the portfolio has nonlinear exposure to volatility.

The spot-cross terms $\sum_j c_j \Delta S \Delta X_j$ are provided in the formula in order to capture the interaction effects between simultaneous movements in risk factors. The most important example is the interaction between equity spot and implied volatility. In equity markets, negative spot returns are often accompanied by increases in implied volatility. If spot and volatility shocks were treated independently in the P&L approximation, this interaction effect could be missed. The spot-volatility cross terms are therefore closely related to vanna-type sensitivities. It is also provided in order to minimize the approximation error compared to full repricing.

In a complete second-order Taylor expansion, all pairwise cross terms between all risk factors would appear. However, estimating the full Hessian is computationally expensive and may introduce numerical noise, especially when the number of risk factors is large. For this reason, the implementation uses a reduced Hessian structure: it includes the most important second-order effects, such as spot gamma, individual non-spot curvature terms and spot–risk-factor cross terms, while excluding cross terms between different non-spot risk factors. This provides a compromise between accuracy and computational efficiency.

Term	Interpretation	Main purpose
$\Delta_S \Delta S$	Spot delta	Linear equity exposure
$\frac{1}{2}\Gamma_S (\Delta S)^2$	Spot gamma	Spot convexity
$\Delta_{FX} \Delta FX$	FX delta	Currency exposure
$\sum_j a_j \Delta X_j$	Non-spot first-order sensitivities	Rates and volatility-PC exposure
$\frac{1}{2}\sum_j b_j (\Delta X_j)^2$	Non-spot curvature	Volga/rate convexity effects
$\sum_j c_j \Delta S \Delta X_j$	Spot cross sensitivities	Vanna-like interaction effects

Table 6: Interpretation of the Taylor terms used in the risk-factor mapping.

3.5.1 Bump and Reprice

At this point, after obtaining the risk factor shocks, the next step is to find the sensitivities related to the risk factors. These sensitivities are to be calculated locally around their current market state. There exist many viable approaches to this, but choosing one that

balances sensitivity accuracy and computational efficiency is ideal in this case. One of the simpler approaches would be to analytically obtain the sensitivities using a pricing formula like the Black-Scholes. To obtain the simpler delta and gamma this way is simple, but using this approach limits the possibility of doing the same with the more complex sensitivities a_j, b_j, and c_j since the analytical formula for these is not available. Moreover, this approach has been seen to be one of the more inaccurate approaches since it introduces unrealistic assumptions due to using the Black-Scholes framework. The assumption that there is a constant volatility and a constant interest rate during the option life is simply not realistic. Furthermore, it assumes asset prices follow a smooth geometric Brownian motion without sudden, massive gaps.

One better approach is to use what is commonly referred to as the Bump and Reprice approach. The idea is to perturb one or several market inputs by a small amount, reprice the instrument under the bumped market state, and approximate the relevant derivative using finite differences. Let V_0 denote the value of an instrument in the current market. Portfolio-level sensitivities are obtained by computing instrument-level sensitivities and adding them using the corresponding position quantities. If X_j denotes a non-spot risk factor and h_j is the bump size, the first-order sensitivity with respect to X_j is approximated by the central difference

$$a_j \approx \frac{V(X_j + h_j) - V(X_j - h_j)}{2h_j}.$$

This central difference is preferred to a one-sided difference since it is less sensitive to local numerical bias. For European options, the model value V is computed using the option pricing engine with inputs from the current market state. In particular, the pricing input consists of the underlying spot level, the strike, the time to maturity, the relevant discount curve, and an option-specific implied volatility. All of these metrics are at this point known except for the option-specific implied volatility. As previously discussed in 3.2.6, there are difficulties obtaining this implied volatility metric for specific options that do not appear on the grid. The option-specific implied volatility is obtained through the volatility surface interpolation operator,

$$\sigma_{\text{opt}} = \mathcal{I}(\Sigma; \tau, M),$$

where Σ denotes the vector of volatility-surface nodes, τ is the time to maturity and M is the option moneyness. The option value can therefore be written schematically as

$$V = BS(S, K, r, \tau, \mathcal{I}(\Sigma; \tau, M)),$$

where $BS(\cdot)$ denotes the option pricing formula used in the implementation. This should not be interpreted as using a single constant volatility for the whole market. Rather, the pricing model uses a local implied volatility obtained from the current volatility surface. When the volatility surface is bumped, the interpolated implied volatility used in the option price changes as well. Thus, the bump-and-reprice sensitivities are computed consistently with the same volatility-surface representation that is used in the scenario generation and in the full-revaluation benchmark.

Second-order sensitivities are estimated by repricing the instrument at both upward and downward bumps and comparing the result with the unbumped value:

$$b_j \approx \frac{V(X_j + h_j) - 2V_0 + V(X_j - h_j)}{h_j^2}.$$

These terms measure curvature with respect to the chosen risk factor. For implied-volatility principal components, this corresponds to a volatility-convexity effect similar to volga or vomma, but expressed in the principal-component coordinate system rather than with respect to a single implied-volatility quote.

The spot delta is estimated in the same way by bumping the underlying spot level. If h_S denotes the spot bump, then

$$\Delta_S \approx \frac{V(S + h_S) - V(S - h_S)}{2h_S}.$$

The spot gamma is estimated using a higher-order central finite-difference stencil:

$$\Gamma_S \approx \frac{-V(S + 2h_S) + 16V(S + h_S) - 30V_0 + 16V(S - h_S) - V(S - 2h_S)}{12h_S^2}.$$

Using several bumped spot values gives a more stable approximation of the second derivative than the simplest three-point formula, at the cost of additional repricing[Pre+07, Chapter 5]. Cross sensitivities between spot and a non-spot risk factor are the most complicated ones here, but are provided by an estimation by bumping both variables simultaneously. For the cross term c_j, corresponding to the mixed derivative $\partial^2 V / \partial S \partial X_j$, the finite-difference approximation is

$$c_j \approx \frac{V(S + h_S, X_j + h_j) - V(S + h_S, X_j - h_j) - V(S - h_S, X_j + h_j) + V(S - h_S, X_j - h_j)}{4h_S h_j}.$$

This quantity measures how the sensitivity to one risk factor changes when another risk factor is moved. In the context of equity options, the most important example is the spot–volatility interaction. A negative spot shock is often accompanied by an increase in

implied volatility, and the effect on the option value may therefore not be captured by adding a pure spot effect and a pure volatility effect independently. The spot–volatility cross term captures this interaction and is related to a vanna-type sensitivity.

For volatility principal components, the bump is applied in PCA space and then mapped back to the volatility-surface nodes. If $w_{i,j}$ is the loading of node i on principal component j, and s_i is the standard deviation used in the PCA scaling, then a PC bump h_j produces a node-level log-volatility change of

$$\Delta v_i = h_j w_{i,j} s_i.$$

The bumped implied-volatility node is then obtained multiplicatively as

$$\sigma_i^{\text{bumped}} = \sigma_i \exp(\Delta v_i).$$

This ensures that implied volatilities remain positive and that the bump direction is the same as with the PCA representation used in the risk-factor scenario generation. Interest-rate sensitivities are computed by bumping individual curve tenors and rebuilding the curve before repricing. For a tenor k, the first-order rate sensitivity is approximated by

$$\rho_k \approx \frac{V(r^{(k)} + h_r) - V(r^{(k)} - h_r)}{2h_r}.$$

If rate principal components are used instead of individual tenor shocks, the tenor-level sensitivities are projected onto the principal-component directions using the PCA loadings. The same principle is used for spot–rate cross sensitivities.

After all instrument-level sensitivities have been estimated, they are stored together with the position quantity and aggregated at the portfolio level. The key computational advantage is that the repricing step is performed only for the finite set of bumps needed to estimate the sensitivities. Once these sensitivities are available, the Taylor approximation can be evaluated for thousands of simulated risk-factor scenarios without repricing the entire portfolio under each scenario. This is the main efficiency gain of the risk-factor mapping approach compared with full revaluation.

The choice of bump size is important. If the bump is too small, the finite-difference estimate may become dominated by numerical noise from interpolation and floating-point rounding. If the bump is too large, the estimate may no longer represent a local derivative around the current market state. The bump sizes are therefore chosen as a compromise between numerical stability and local accuracy. The bump sizes used in the framework are shown in the results section, and these are obtained using a data-driven approach where

the error of 95-level VaR and 99-level VaR between full repricing and the approximation was minimized by using different bump sizes on 75 randomized SPX-option portfolios, with the amount of options ranging from 6 to 12 in each portfolio. In particular, the objective function:

$$J(h) = \frac{1}{P}\sum_{p=1}^{P}\left(\frac{|\mathrm{VaR}_{95,p}^{\mathrm{Taylor}} - \mathrm{VaR}_{95,p}^{\mathrm{Full}}|}{\mathrm{VaR}_{95,p}^{\mathrm{Full}}} + \frac{|\mathrm{VaR}_{99,p}^{\mathrm{Taylor}} - \mathrm{VaR}_{99,p}^{\mathrm{Full}}|}{\mathrm{VaR}_{99,p}^{\mathrm{Full}}}\right),$$

was minimized to help perceive the data and sizes that would be suitable to use. However, the minimization was not seen as the pure solution to the bump sizes, as these can be chosen differently depending on the prioritization of VaR-level or other robustness metrics of choice. The final chosen bump sizes were then used throughout the validation and backtesting.

3.6 VaR approximation using Monte Carlo scenarios

Finally, after estimating the sensitivities and obtaining the risk-factor scenario shocks, we can provide the final calculations for the VaR. The pipeline makes it possible to estimate both the 99-% confidence level VaR and the 95-% confidence level VaR. The model produces an empirical distribution of one-day portfolio P&L outcomes by bootstrapping the generated scenario shocks in the given Taylor approximation, along with the option-specific sensitivities already calculated (and only needs to be calculated once per option):

$$\Delta\Pi \approx \Delta_S \Delta S + \frac{1}{2}\Gamma_S(\Delta S)^2 + \Delta_{FX}\Delta FX + \sum_j a_j \Delta X_j + \frac{1}{2}\sum_j b_j(\Delta X_j)^2 + \sum_j c_j \Delta S \Delta X_j.$$

Repeating this for all N scenarios gives an empirical distribution of possible one-day portfolio P&L outcomes,

$$\{\Delta\Pi^{(1)}, \Delta\Pi^{(2)}, \ldots, \Delta\Pi^{(N)}\}.$$

The VaR is computed from the loss distribution rather than directly from the P&L distribution. Since losses correspond to negative P&L values, the scenario loss in scenario s is defined as

$$L^{(s)} = -\Delta\Pi^{(s)}.$$

We mainly use a total of 10,000 scenario shocks generated in the empirical implementation. Increasing the amount of bootstrapped scenarios reduces the Monte-Carlo noise but also makes the back-testing and validation slower. It is in our interest to find a common ground between a suitable amount of scenarios generated, so that the VaR-approximation is accurate enough, but at the same time making sure not to make the implementation overly slow, especially with full revaluation computed alongside. For a confidence level α,

the Value-at-Risk is the empirical α-quantile of the simulated loss distribution,

$$\mathrm{VaR}_\alpha = \widehat{Q}_\alpha\left(L^{(1)}, \ldots, L^{(N)}\right).$$

Equivalently, since $L^{(s)} = -\Delta\Pi^{(s)}$, the same quantity can be obtained from the lower tail of the P&L distribution:

$$\mathrm{VaR}_\alpha = -\widehat{Q}_{1-\alpha}\left(\Delta\Pi^{(1)}, \ldots, \Delta\Pi^{(N)}\right).$$

Thus, the 95% VaR is the negative of the 5% quantile of the simulated P&L distribution, while the 99% VaR is the negative of the 1% quantile.

3.6.1 Full Revaluation

For comparison, the same risk-factor scenarios can also be evaluated using full revaluation. In that case, each simulated scenario is first transformed into a shocked market state, including shocked spot levels, yield curves and implied-volatility surfaces. The full portfolio is then repriced under each shocked market state, producing a full-revaluation P&L distribution,

$$\Delta\Pi^{(s)}_{\text{full}} = \Pi(\mathcal{M}^{(s)}) - \Pi(\mathcal{M}_0),$$

where $\mathcal{M}_0$ denotes the current market state and $\mathcal{M}^{(s)}$ denotes the shocked market state in scenario s. The full-revaluation VaR is computed from this distribution using the same empirical quantile procedure as for the Taylor approximation. This makes the Taylor and full-revaluation VaR estimates directly comparable, since they are based on the same underlying scenario shocks.

3.7 Backtesting and model validation

In order to validate the whole model and pipeline a backtesting is constructed. The whole backtest can be divided into two parts. We first validate how well our model approximates the P&L distribution based on the second-order Taylor approximation by comparing the P&L distribution and VaR results to the corresponding full revaluation results. The full revaluation is treated as the baseline model throughout the validation. Here, different methods and steps in the pipeline work are validated. Cubic spline versus bilinear interpolation validation is based on the minimization of the error between the baseline and Taylor models. The same goes for finding the optimal bump sizes in the bump and reprice framework: the decision is data-driven and based on the same minimization of the approximation error between the models based on many different option portfolios. Furthermore, sanity checks are delivered which prove that, in the linear cases with stocks and FX-spots, the VaR and P&L error between the models are approximately 0.

3.7.1 Taylor approximation versus full revaluation

The first validation step evaluates the accuracy of the Taylor-based risk-factor mapping relative to full revaluation. This validation is performed conditional on the same generated risk-factor scenarios. Hence, the comparison isolates the approximation error introduced by the Taylor mapping rather than differences caused by the scenario-generation model.

Let $\Delta\Pi^{(s)}_{\text{Taylor}}$ denote the Taylor-approximated P&L in scenario s, and let $\Delta\Pi^{(s)}_{\text{Full}}$ denote the corresponding full-revaluation P&L obtained by repricing the portfolio under the same shocked market state. The scenario-level approximation error is defined as

$$e^{(s)} = \Delta\Pi^{(s)}_{\text{Taylor}} - \Delta\Pi^{(s)}_{\text{Full}}.$$

The Taylor approximation is considered accurate if these errors are small relative to the scale of the portfolio P&L distribution, especially in the lower tail where VaR is computed. Several diagnostics are used to compare the two P&L distributions. First, the average and root-mean-squared approximation errors are computed,

$$\text{MAE} = \frac{1}{N}\sum_{s=1}^{N}|e^{(s)}|, \qquad \text{RMSE} = \sqrt{\frac{1}{N}\sum_{s=1}^{N}(e^{(s)})^2}.$$

Second, the correlation between $\Delta\Pi_{\text{Taylor}}$ and $\Delta\Pi_{\text{Full}}$ is measured to assess whether the Taylor approximation preserves the ordering and shape of the full-revaluation P&L distribution. Third, the Taylor and full-revaluation VaR estimates are compared directly at the 95% and 99% confidence levels,

$$\Delta\text{VaR}_\alpha = \text{VaR}^{\text{Taylor}}_\alpha - \text{VaR}^{\text{Full}}_\alpha.$$

This quantity measures whether the Taylor approximation underestimates or overestimates the risk estimate relative to full revaluation. The comparison is performed using the same volatility model, the same filtered historical simulation scenarios and the same current market state for both methods. The only difference is the valuation step. In the Taylor approach, the scenario P&L is computed from precomputed sensitivities. In the full-revaluation benchmark, each scenario is mapped to a shocked market state and the portfolio is repriced directly. This makes the comparison a direct test of the risk-factor mapping approximation.

3.7.2 VaR exceedance and clustering backtests

To assess the predictive performance of the VaR models, we conduct daily backtests over the validation period (2022-01-03 to 2026-01-02). For each trading day, the model

produces one-day-ahead VaR forecasts at the 95% and 99% confidence levels. The realized portfolio P&L from actual option prices is compared to the VaR thresholds for each day.

We evaluate the quality of the VaR forecasts using two standard statistical tests described in the background. Firstly, the Kupiec test, which uses observed violation frequency and compares it with the expected amount. Secondly, the Christoffersen–Pelletier duration test for independence of violations over time. This test is based on the Weibull distribution of the durations between consecutive violations; a p-value below 0.05 indicates significant clustering.

In addition to the tests, we analyze the VaR thresholds together with the realized PnL, compute rolling 60-day exceedance rates, examine the state-wise break rates (when state-dependent sampling is used), and inspect the scenario mix for the residuals over time. These visual diagnostics help to interpret the model behavior during calm and turbulent periods.

Six different model specifications are compared. All models share the following common settings: Student-t innovations, volatility-based state metric (when states are used), 10 000 Monte Carlo scenarios, a fixed random seed, and GARCH parameters re-estimated every 100 days along with residual updates to include new market information into the model. Table 7 summarizes the distinctive features of each model. Each model uses both Taylor approximation and full reevaluation to compute VaR estimations; this results in four different estimates in total since both are done on the 95- and 99% level.

Table 7: Overview of the six VaR model configurations. Models 4 and 6 use states for both volatility and covariance. Model 1 uses no states and is a baseline model.

Model	**Volatility model**	**Covariance (residual sampling)**
1	EWMA ($\lambda = 0.94$)	Standard FHS (no states)
2	GARCH(1,1)	State-dependent, soft sampling
3	GJR-GARCH(1,1)	State-dependent, soft sampling
4	Self-switching GJR-GARCH	State-dependent, soft sampling
5	Self-switching GJR-GARCH	State-dependent, hard sampling
6	Self-switching GJR-GARCH	Standard FHS (no states)

For the state-based residual models, the quantile upper-thresholds are set to 40% for stable and 93% for normal. The softmax temperature is 0.1. Model 1 uses no optimization, as well as no states for either volatility or covariance. It can be seen as a simple baseline model to compare the other models with.

4 Results

4.1 Taylor approximation validation

Table 8: Full-period daily Taylor-versus-full-revaluation VaR validation.

Metric	**VaR95**	**VaR99**
Number of valuation dates	994	
Average Taylor VaR	3932.43	5981.92
Average full-revaluation VaR	3900.83	5846.62
Average |Taylor − Full| VaR	37.22	140.21
Median |Taylor − Full| VaR	22.63	103.47
P90 |Taylor − Full| VaR	85.09	309.40
P95 |Taylor − Full| VaR	111.59	399.63
Maximum |Taylor − Full| VaR	984.34	1161.28
Average relative VaR difference	**1.11**%	**2.10**%
P90 relative VaR difference	2.43%	6.07%
Maximum relative VaR difference	7.97%	13.50%
Correlation between Taylor and full VaR	0.9997	0.9992

The full-period daily validation contains 994 valuation dates across the 2022–2026 validation period. In contrast to the sampled validation, this test evaluates the Taylor approximation and the full-revaluation benchmark on every available backtest date. The average absolute Taylor–full VaR difference is 37.22 for VaR95 and 140.21 for VaR99, corresponding to average relative differences of 1.11% and 2.10%, respectively. The correlations between Taylor and full-revaluation VaR estimates are above 0.999 for both confidence levels, indicating that the Taylor approximation closely tracks the full-revaluation benchmark over time.

The largest relative differences occur in isolated tail cases, especially for VaR99. However, these differences do not materially change the backtesting conclusions. Taylor and full revaluation generate very similar exceedance counts over the full validation period, showing that the Taylor approximation preserves the practical risk signal of the full-revaluation model.

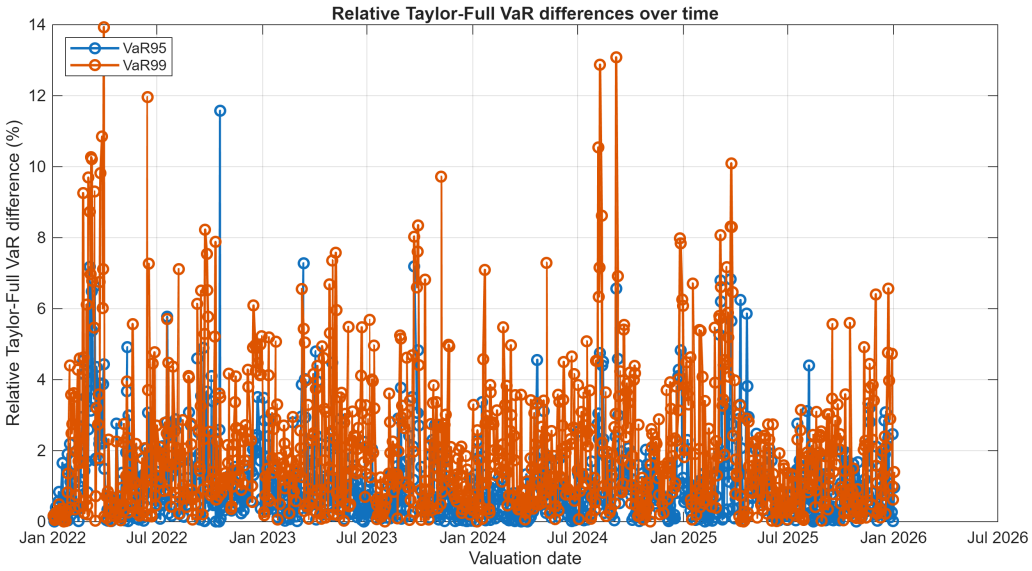


Figure 11: Relative Taylor–full VaR differences over the full 2022–2026 validation period. Relative differences are measured as absolute Taylor–full VaR differences divided by the corresponding full-revaluation VaR.

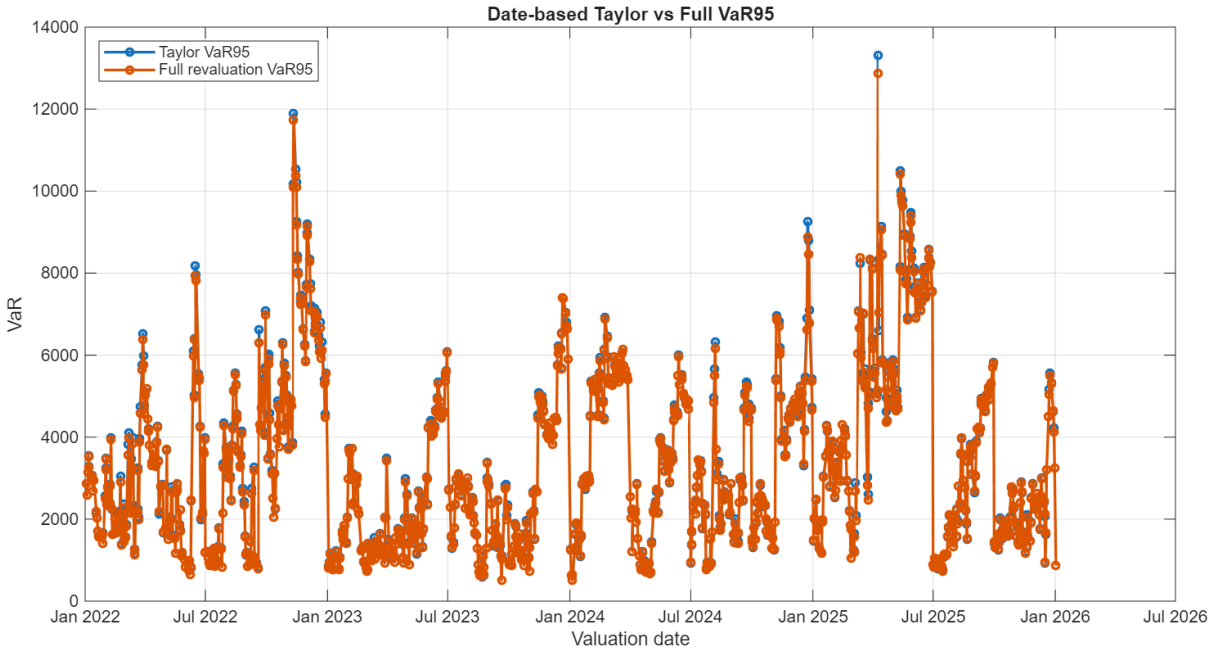


Figure 12: Full-period date-based comparison of Taylor and full-revaluation VaR95 estimates over the 2022–2026 validation period. The two series almost completely overlap, indicating that the Taylor approximation closely tracks the full-revaluation benchmark at the 95% confidence level.

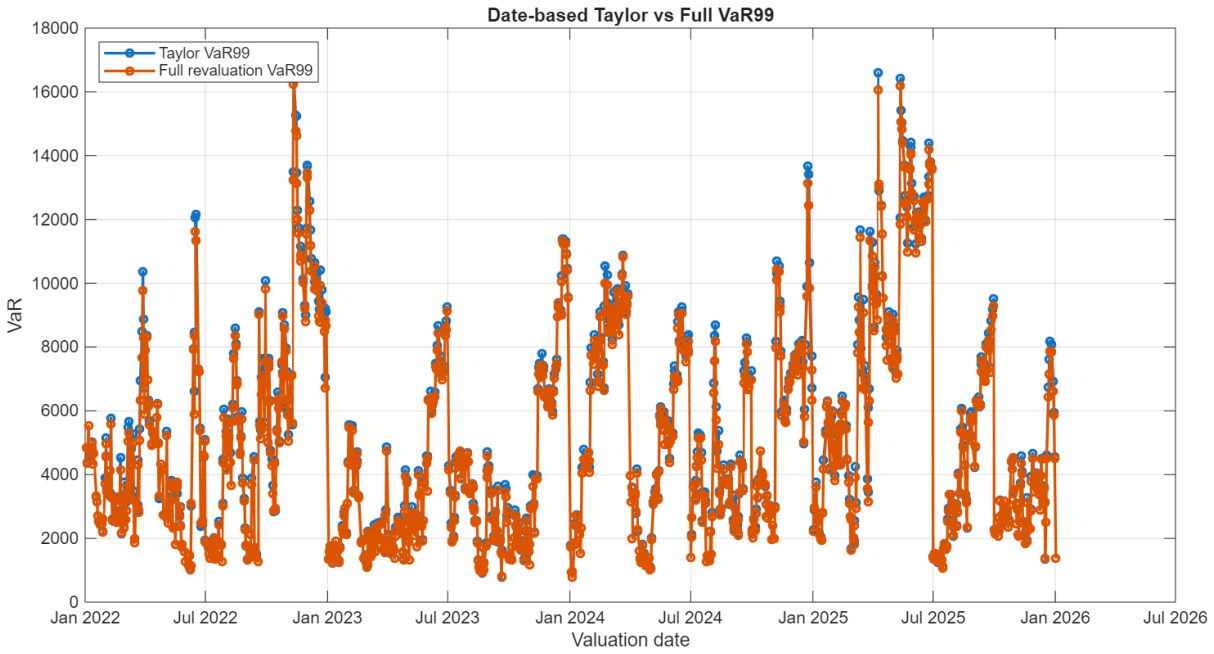


Figure 13: Full-period date-based comparison of Taylor and full-revaluation VaR99 estimates over the 2022–2026 validation period. The agreement remains strong also in the deeper tail, although isolated deviations appear on some valuation dates.

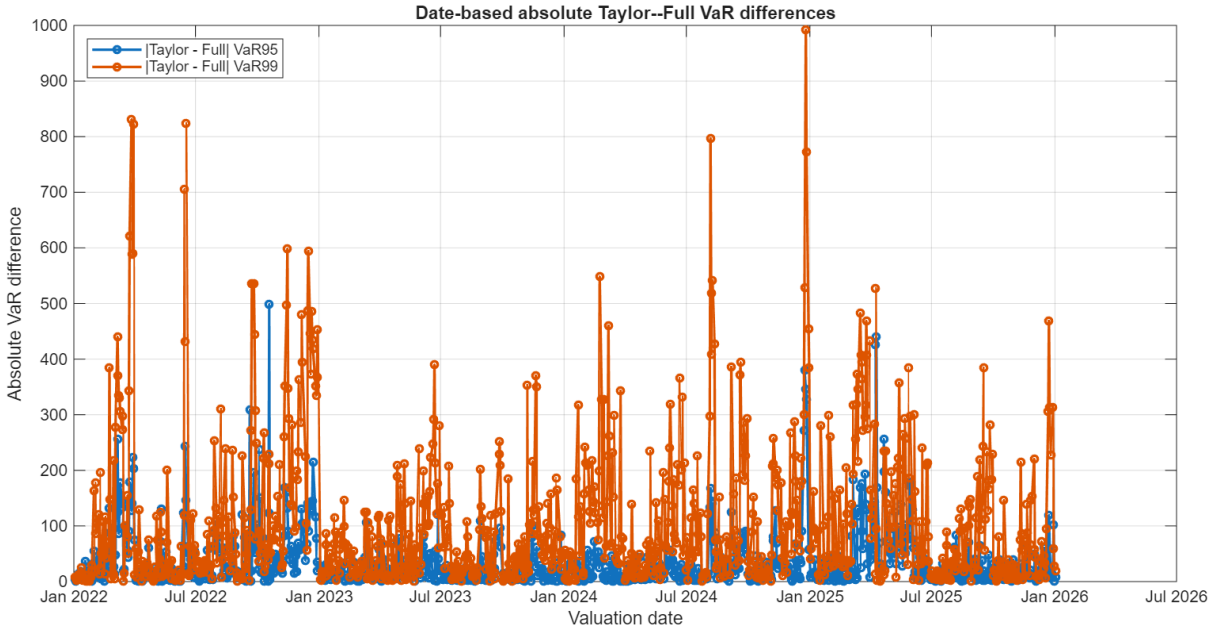


Figure 14: Absolute Taylor–full VaR differences over the full 2022–2026 validation period. Most observations remain at low levels, while the largest absolute deviations occur on isolated dates.

To test whether the low approximation error also holds for more option-heavy portfolios, a separate randomized portfolio validation was performed. In this experiment, 75 synthetic SPX option portfolios were generated across five portfolio families: balanced, short-dated, long-dated, OTM-heavy and ATM-heavy. Each portfolio contains several options with

different strikes, maturities and option types. The same 50,000 risk-factor scenarios were used for Taylor approximation and full revaluation, so the comparison isolates the valuation approximation error.

Table 9: Overall diagnostics for random-portfolio Taylor-versus-full-revaluation validation.

Metric	**Value**
Number of randomized portfolios	75
Number of scenarios	50,000
Average $\lvert\Delta\text{VaR95}\rvert$	7.36
Median $\lvert\Delta\text{VaR95}\rvert$	5.64
P90 $\lvert\Delta\text{VaR95}\rvert$	18.80
Maximum $\lvert\Delta\text{VaR95}\rvert$	50.83
Average $\lvert\Delta\text{VaR99}\rvert$	11.90
Median $\lvert\Delta\text{VaR99}\rvert$	8.54
P90 $\lvert\Delta\text{VaR99}\rvert$	26.71
Maximum $\lvert\Delta\text{VaR99}\rvert$	43.50
Average RMSE	7.62
Average MAE	3.98
Average correlation	0.9989
Average $\lvert\Delta\text{VaR95}\rvert$ / gross value	0.035%
Average $\lvert\Delta\text{VaR99}\rvert$ / gross value	0.062%

Table 10: Random-portfolio Taylor-versus-full-revaluation validation by portfolio family.

Family	N	**Avg.** $\lvert\Delta\text{VaR}_{95}\rvert$	**Avg.** $\lvert\Delta\text{VaR}_{99}\rvert$	**Rel. VaR95**	**Rel. VaR99**	**Avg. corr.**
BALANCED	15	7.43	10.90	1.27%	1.53%	0.9988
SHORT-DATED	15	2.78	10.79	1.05%	1.38%	0.9995
LONG-DATED	15	11.81	12.55	2.93%	2.17%	0.9985
OTM-HEAVY	15	4.24	8.24	2.25%	2.13%	0.9994
ATM-HEAVY	15	10.54	17.03	2.96%	2.18%	0.9981
ALL	75	7.36	11.90	2.09%	1.88%	0.9989

The randomized portfolio validation confirms that the reduced second-order Taylor approximation remains close to full revaluation even for more option-heavy portfolios. Across all 75 portfolios, the average absolute VaR difference is 7.36 for VaR95 and 11.90 for VaR99. Relative to gross market value, the average absolute differences are only

0.035% and 0.062%, respectively. The average Taylor–full P&L correlation is 0.9989, which indicates that the Taylor approximation preserves the shape and ordering of the full-revaluation P&L distribution.

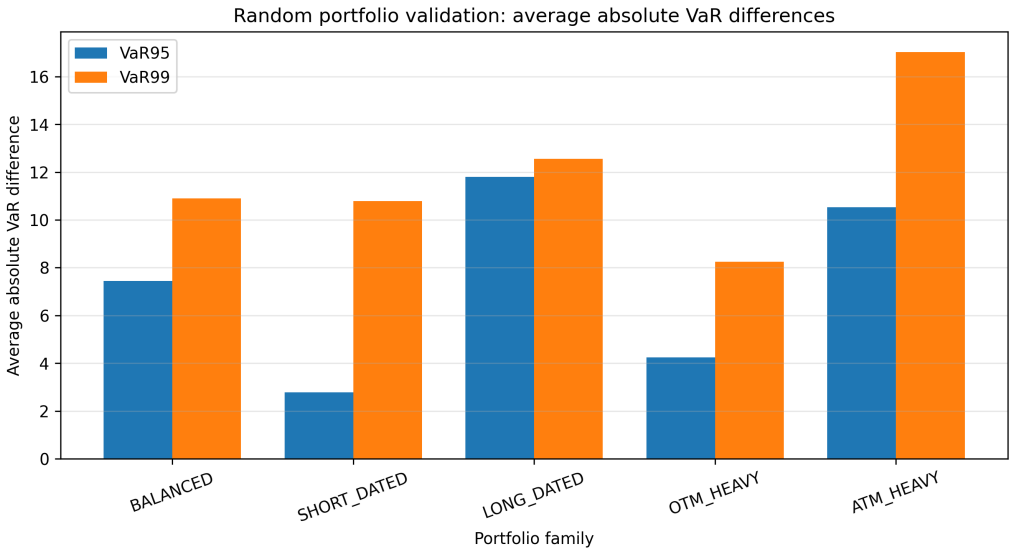


Figure 15: Average absolute Taylor–full VaR difference by randomized portfolio family.

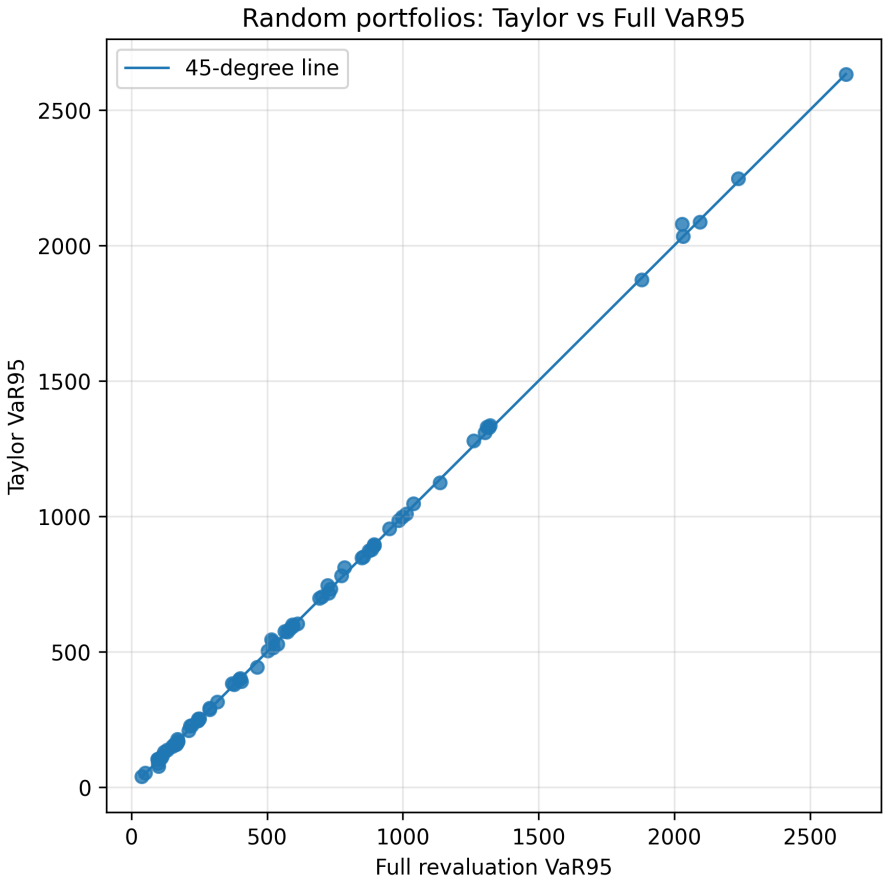


Figure 16: Randomized portfolios: Taylor VaR95 versus full-revaluation VaR95.

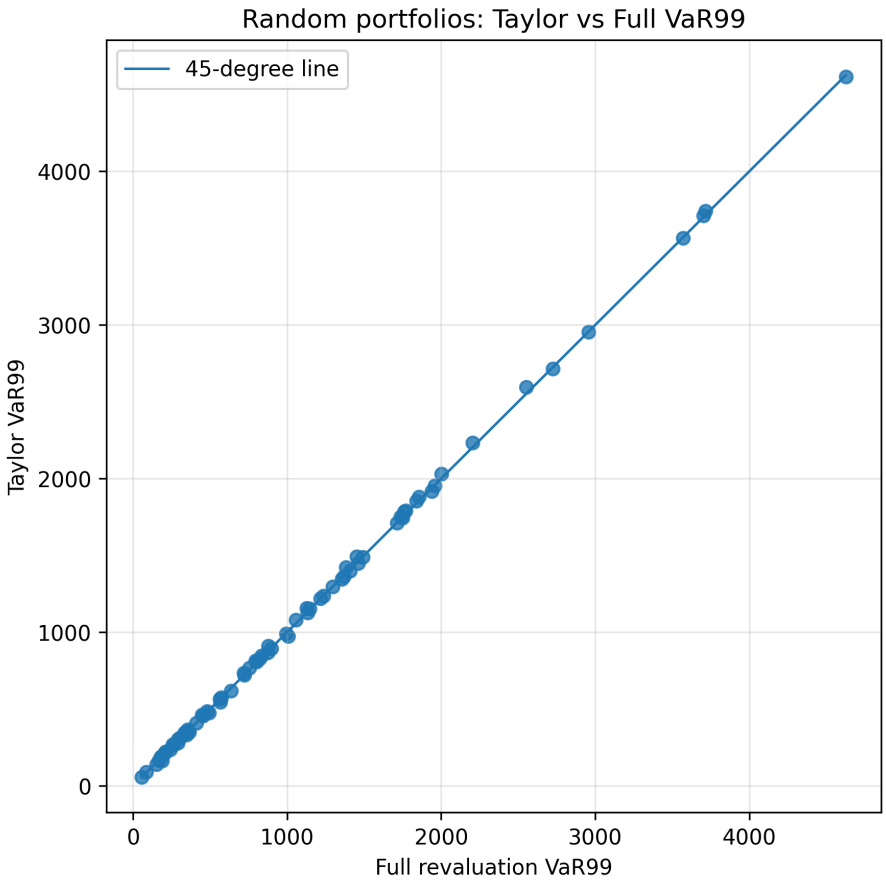


Figure 17: Randomized portfolios: Taylor VaR99 versus full-revaluation VaR99.

4.1.1 Tail validation

Table 11: Tail diagnostics of Taylor-versus-full-revaluation VaR differences. Relative errors are measured as absolute Taylor–full VaR differences divided by full-revaluation VaR.

Validation set	**Avg. VaR95**	**P90 VaR95**	**Max VaR95**	**Avg. VaR99**	**P90 VaR99**	**Max VaR99**
Date-based, full period	1.11%	2.43%	7.97%	2.10%	6.07%	13.50%
Random option portfolios	2.09%	5.15%	20.70%	1.88%	4.12%	13.60%
Bond portfolios	0.042%	0.068%	0.681%	0.062%	0.087%	0.829%

The tail diagnostics show that the average Taylor–full VaR differences understate the largest approximation errors. In the date-based validation, the largest relative differences occur on risky valuation dates, where the scenario shocks move the portfolio further away from the Taylor expansion point. This is expected, since the Taylor approximation is local by construction. For the randomized option portfolios, the largest relative errors occur for portfolios with relatively low full-revaluation VaR levels. In those cases, the

relative error can appear large even though the absolute error and the error relative to gross market value remain small. Bond portfolios exhibit substantially smaller relative errors, confirming that most of the remaining approximation error is driven by option nonlinearity.

Table 12: Worst relative Taylor–full VaR differences.

Validation set	**VaR level**	**Worst case**	**Full VaR**	**Abs. diff.**	**Rel. diff.**	**State / family**
Date-based, full period	VaR95	2025-08-04	2746.56	219.03	7.97%	NORMAL
Date-based, full period	VaR99	2025-12-31	6248.34	843.25	13.50%	STABLE
Random option portfolios	VaR95	ATM_HEAVY_006	99.35	20.56	20.70%	ATM-heavy
Random option portfolios	VaR99	ATM_HEAVY_006	189.17	25.72	13.60%	ATM-heavy
Bond portfolios	VaR95	BARBELL_10	2981.37	20.29	0.681%	Barbell
Bond portfolios	VaR99	BARBELL_10	4953.61	41.09	0.829%	Barbell

The worst-case diagnostics show that the largest relative errors have different interpretations across validation sets. In the date-based full-period validation, the largest relative VaR95 and VaR99 differences are 7.97% and 13.50%, respectively. These are isolated observations in the full 2022–2026 validation period and do not materially affect the overall agreement between Taylor and full revaluation.

For the randomized option portfolios, the maximum relative VaR95 error reaches 20.70%. This occurs for a portfolio with a low full-revaluation VaR level. In absolute terms, the VaR difference is only 20.56, corresponding to approximately 0.081% of gross market value. Hence, the large relative error is mainly caused by a small denominator rather than by a large valuation error. Bond portfolios exhibit substantially smaller worst-case relative errors, confirming that the largest approximation errors arise from option nonlinearity rather than from the general mapping framework.

4.1.2 Non-Option instrument validation

As a sanity check of the risk-factor mapping implementation, the Taylor approximation was also compared with full revaluation for simpler instrument classes. For equity spot and FX spot positions, the Taylor and full-revaluation P&L matched up to numerical precision. For the FX spot test, the mean absolute error was approximately 1.0×10^{-10}, and the VaR95 and VaR99 differences were both below 10^{-9}. For stock positions, the error was exactly zero in the tested scenarios.

Table 13: Instrument-level Taylor-versus-full-revaluation sanity checks. Relative errors are measured against the corresponding full-revaluation VaR estimate.

Instrument class	**Test size**	**Rel. $\lvert\Delta\text{VaR}_{95}\rvert$**	**Rel. $\lvert\Delta\text{VaR}_{99}\rvert$**	**PnL error**	**Comment**
Stock spot	3,000 scenarios	0.0000%	0.0000%	0.0000	Exact linear mapping
FX spot	3,000 scenarios	$< 10^{-10}$%	$< 10^{-10}$%	$\approx 10^{-10}$	Numerical precision
Bonds	75 portfolios	0.042%	0.062%	Low	Very small relative error
Options	75 portfolios	2.09%	1.88%	Moderate	Main nonlinear validation

Table 13 reports relative rather than absolute approximation errors, since absolute VaR differences are not directly comparable across instrument classes with different portfolio sizes and risk levels. The relative error is computed as the absolute Taylor–full VaR difference divided by the corresponding full-revaluation VaR estimate. It should be noted that this validation is strictly a one-day-VaR run on 2026-01-02 with 75 option portfolios, with the range of options varying from 6 to 12.

The results show that the Taylor mapping is exact for stock spot positions and accurate up to numerical precision for FX spot positions. Bond portfolios also exhibit very small relative errors, with average relative VaR differences below 0.1% at both confidence levels. The largest relative approximation errors occur for option portfolios, which is expected since options introduce the strongest nonlinearities through spot curvature, volatility exposure, and spot–volatility interaction terms. Even for the randomized option-heavy portfolios, however, the average relative VaR error remains around 2%, indicating that the Taylor approximation remains close to the full-revaluation benchmark.

4.1.3 Bump sizes

In the case of the sensitivity calculations, the bump sizes are central. It is important to choose bump sizes which do not make the central differences unstable, whilst at the same time making sure the evaluation is properly located not too far away from the local point. To do this, we evaluated a data-driven validation of the bump sizes. These are based on minimizing the absolute VaR error between the baseline full revaluation and the Taylor approximation formula. We use an objective function, based on minimizing the error, to help us perceive the results in an objective manner. The objective function minimization does not need to be followed entirely since choices on bump-sizes can depend on other factors such as prioritizing VaR error at the 99-level or at the 95-level or for any other robustness-based reasons.

As for the final results of the bump sizes, the lowest objective value in the local sweep is

obtained for a slightly smaller gamma configuration. However, the final implementation retains

$$\boxed{h_\Delta = 0.0125S, \qquad h_\Gamma = 0.0275S, \qquad h_{\mathrm{PC}} = 0.20.}$$

Table 14: One-at-a-time bump-size robustness sweep around the baseline configuration.

Config	h_Δ	h_Γ	h_{PC}	**Avg.** $\lvert\Delta\mathrm{VaR}_{95}\rvert$	**Avg.** $\lvert\Delta\mathrm{VaR}_{99}\rvert$	**Objective (%)**
BASE	0.0100	0.0200	0.15	3.01	10.17	1.73
Delta 0.0050	0.0050	0.0200	0.15	3.86	11.13	2.27
Delta 0.0075	0.0075	0.0200	0.15	3.40	10.72	1.99
Delta 0.0125	0.0125	0.0200	0.15	2.87	9.73	1.51
Delta 0.0200	0.0200	0.0200	0.15	6.10	12.62	2.70
Gamma 0.0125	0.0100	0.0125	0.15	2.04	14.71	1.82
Gamma 0.0275	0.0100	0.0275	0.15	6.33	12.90	2.51
Gamma 0.0400	0.0100	0.0400	0.15	11.85	21.41	3.89
Gamma 0.0600	0.0100	0.0600	0.15	17.36	33.20	6.48
PC 0.10	0.0100	0.0200	0.10	3.01	10.17	1.73
PC 0.20	0.0100	0.0200	0.20	3.01	10.17	1.73
PC 0.30	0.0100	0.0200	0.30	3.01	10.17	1.73
PC 0.40	0.0100	0.0200	0.40	3.01	10.17	1.73

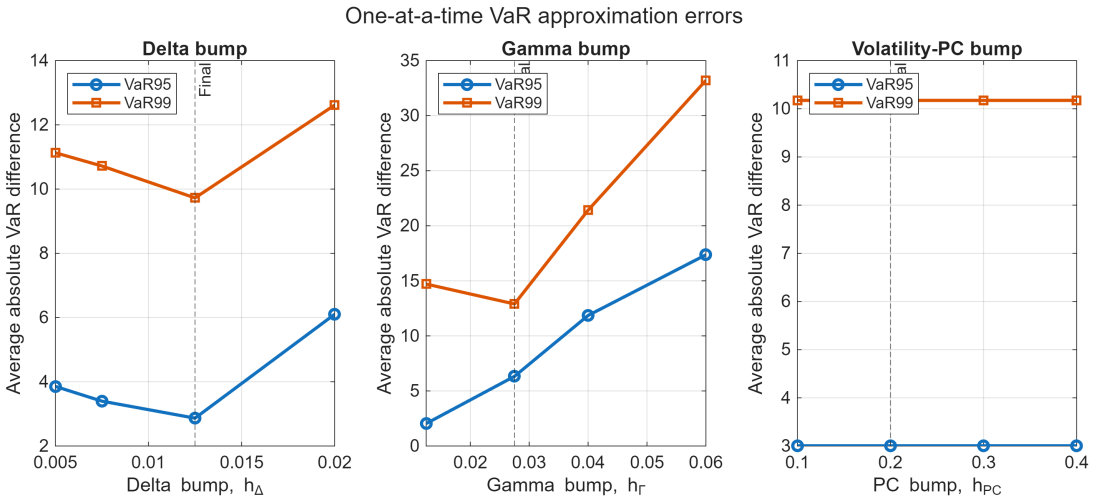


Figure 18: One-at-a-time bump-size robustness sweep. Two bump sizes are held fixed while the third is varied. The final bump configuration is marked by the dashed vertical line in each panel.

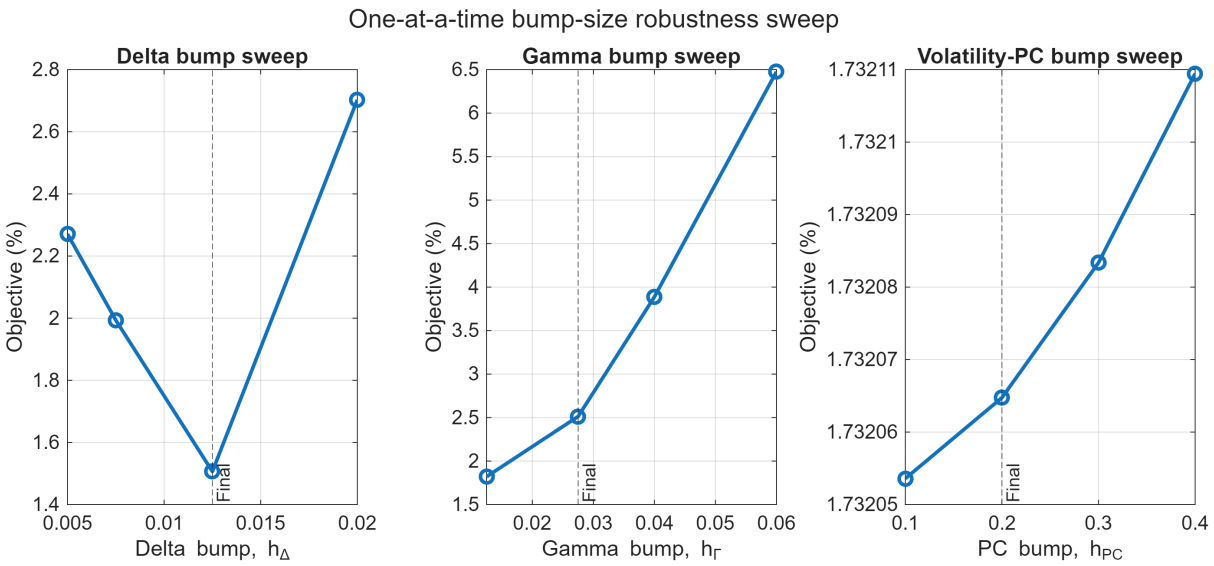


Figure 19: One-at-a-time bump-size robustness sweep. Two bump sizes are held fixed while the third is varied. The final bump configuration is marked by the dashed vertical line in each panel.

4.2 VaR

We evaluate the VaR models using the backtesting framework described in Section 3.7.2. The validation period runs from 2022-01-03 to 2026-01-02, comprising 994 trading days. For each model, we report the unconditional coverage (Kupiec test) and the duration-based independence test (Christoffersen & Pelletier). We also show daily PnL, VaR forecasts, rolling 60-day exceedance rates, scenario mix, and state-wise behavior.

4.2.1 Statistical tests

Table 15 shows the number of VaR violations, the violation rate, and the Kupiec p-value for each model and confidence level. Results are reported for both the Taylor approximation and full revaluation. Table 16 presents the Weibull shape parameter b and the duration test p-value.

Table 15: VaR violation backtesting

Model information	Method	VaR level	Breaks	Break rate	Kupiec p-value
EWMA	Taylor	95%	70 / 994	7.04%	0.0053
($\lambda = 0.94$)	Full revaluation	95%	70 / 994	7.04%	0.0053
No States	Taylor	99%	9 / 994	0.91%	0.7607
	Full revaluation	99%	12 / 994	1.21%	0.5248
GARCH	Taylor	95%	61 / 994	6.14%	0.1117
Standard	Full revaluation	95%	64 / 994	6.44%	0.0459
States	Taylor	99%	10 / 994	1.01%	0.9848
Soft-Sampling	Full revaluation	99%	11 / 994	1.11%	0.7397
GARCH	Taylor	95%	59 / 994	5.94%	0.1881
GJR	Full revaluation	95%	59 / 994	5.94%	0.1881
States	Taylor	99%	9 / 994	0.91%	0.7607
Soft-Sampling	Full revaluation	99%	10 / 994	1.01%	0.9848
Self-Switch GARCH	Taylor	95%	54 / 994	5.43%	0.5369
GJR	Full revaluation	95%	54 / 994	5.43%	0.5369
States	Taylor	99%	11 / 994	1.11%	0.7397
Soft-Sampling	Full revaluation	99%	13 / 994	1.31%	0.3517
Self-Switch GARCH	Taylor	95%	59 / 994	5.94%	0.1881
GJR	Full revaluation	95%	59 / 994	5.94%	0.1881
States	Taylor	99%	11 / 994	1.11%	0.7397
Hard-Sampling	Full revaluation	99%	14 / 994	1.41%	0.2228
Self-Switch GARCH	Taylor	95%	56 / 994	5.63%	0.3684
GJR	Full revaluation	95%	56 / 994	5.63%	0.3684
No States	Taylor	99%	9 / 994	0.91%	0.7607
	Full revaluation	99%	10 / 994	1.01%	0.9848

Table 16: VaR duration backtesting

Model information	Method	VaR level	Breaks	Weibull B value	Duration p-value
EWMA	Taylor	95%	70 / 994	0.8333	0.0347
($\lambda = 0.94$)	Full revaluation	95%	70 / 994	0.8152	0.0158
No States	Taylor	99%	9 / 994	1.2505	0.4855
	Full revaluation	99%	12 / 994	0.9721	0.9083
GARCH	Taylor	95%	61 / 994	0.8223	0.0284
Standard	Full revaluation	95%	64 / 994	0.8186	0.0208
States	Taylor	99%	10 / 994	0.9145	0.7193
Soft-Sampling	Full revaluation	99%	11 / 994	1.0464	0.8549
GARCH	Taylor	95%	59 / 994	0.8623	0.1146
GJR	Full revaluation	95%	59 / 994	0.8623	0.1146
States	Taylor	99%	9 / 994	0.8057	0.4015
Soft-Sampling	Full revaluation	99%	10 / 994	0.9145	0.7193
Self-Switch GARCH	Taylor	95%	54 / 994	0.9707	0.7745
GJR	Full revaluation	95%	54 / 994	0.9480	0.6034
States	Taylor	99%	11 / 994	1.0699	0.8061
Soft-Sampling	Full revaluation	99%	13 / 994	1.0024	0.9918
Self-Switch GARCH	Taylor	95%	59 / 994	0.8838	0.1981
GJR	Full revaluation	95%	59 / 994	0.9037	0.2912
States	Taylor	99%	11 / 994	1.1738	0.5685
Hard-Sampling	Full revaluation	99%	14 / 994	1.1322	0.6004
Self-Switch GARCH	Taylor	95%	56 / 994	0.8739	0.1749
GJR	Full revaluation	95%	56 / 994	0.8739	0.1749
No States	Taylor	99%	9 / 994	0.8268	0.4700
	Full revaluation	99%	10 / 994	0.8313	0.4475

4.2.2 VaR backtest – EWMA

Figures 20 and 21 show the daily PnL, VaR thresholds (95% and 99%), and the 60-day rolling exceedance rate for the EWMA model without states.

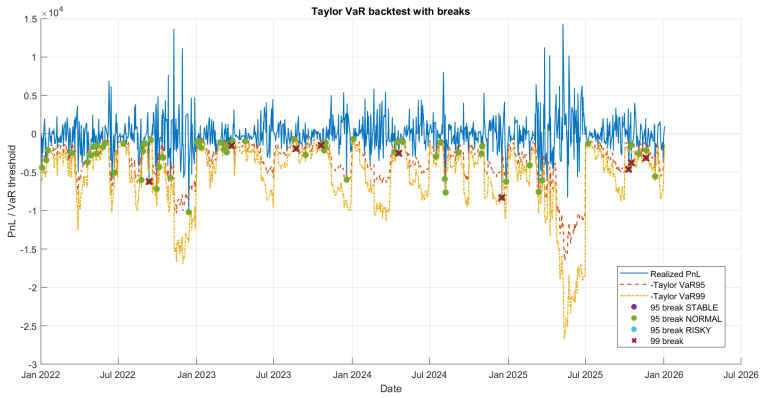


Figure 20: EWMA model – daily PnL, VaR thresholds, and VaR violation.

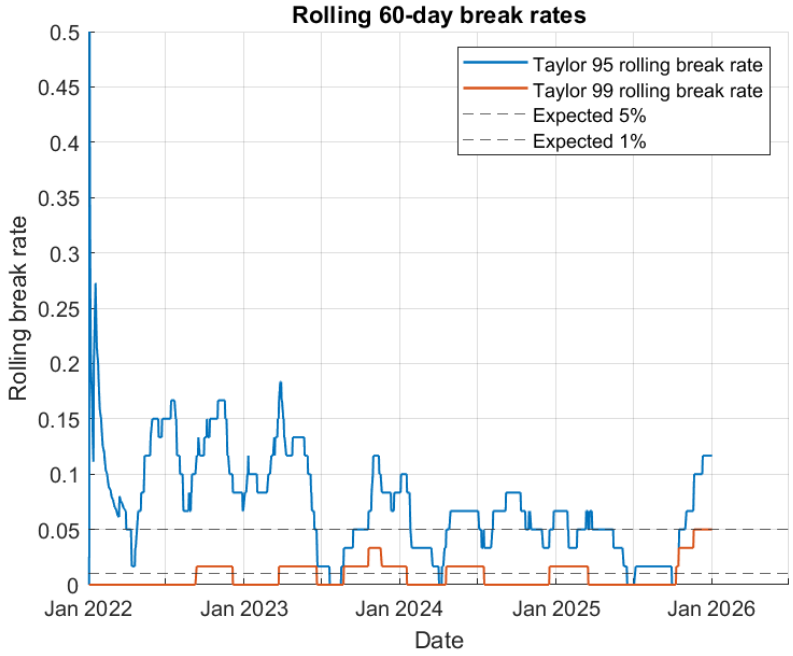


Figure 21: EWMA model – 60-day rolling VaR violation rate.

4.2.3 VaR break backtest - Standard GARCH

Figures 22 and 23 show the daily PnL, VaR thresholds (95% and 99%), and the 60-day rolling exceedance rate for the GARCH model with states and soft sampling. Figures 24 and 25 show the mix of residual state distribution over time as well as state-wise break-rate, respectively.

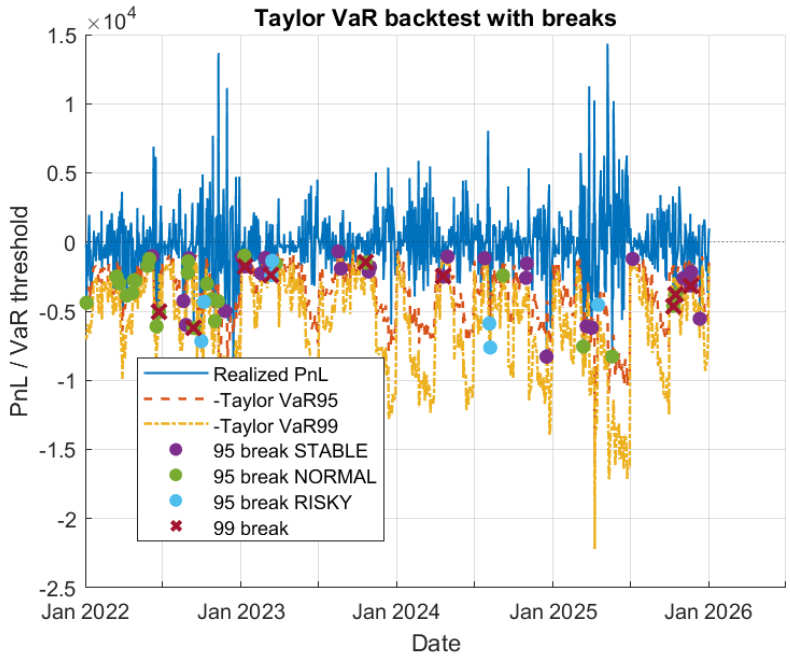


Figure 22: GARCH model – daily PnL, VaR thresholds and VaR violations.

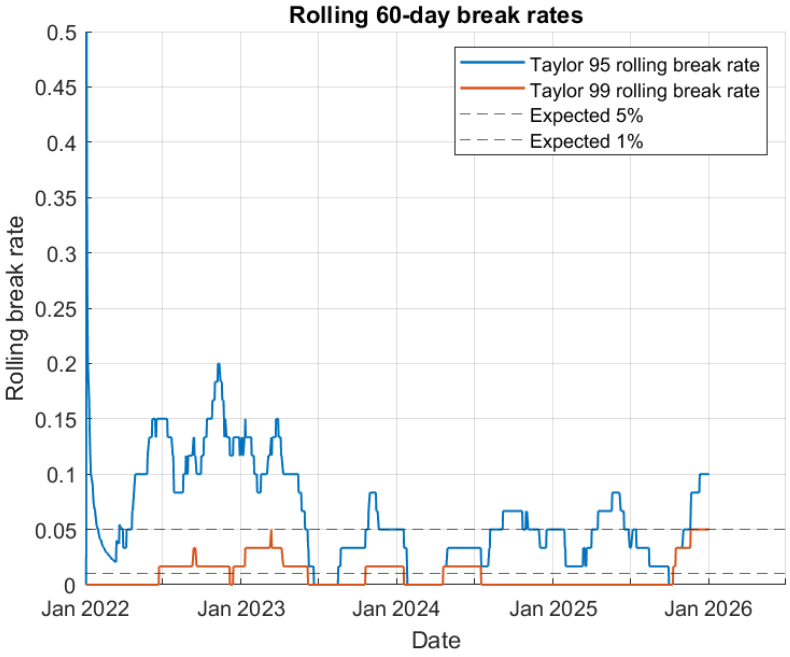


Figure 23: GARCH model – 60-day rolling VaR violation rate for 95- and 99% Taylor VaR estimations.

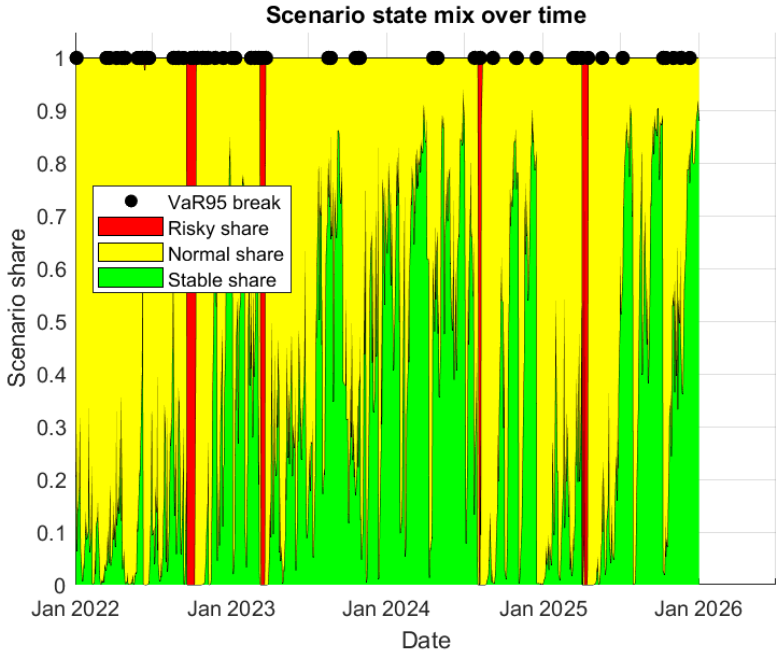


Figure 24: GARCH – Proportion of stable, normal and risky bootstrap draws.

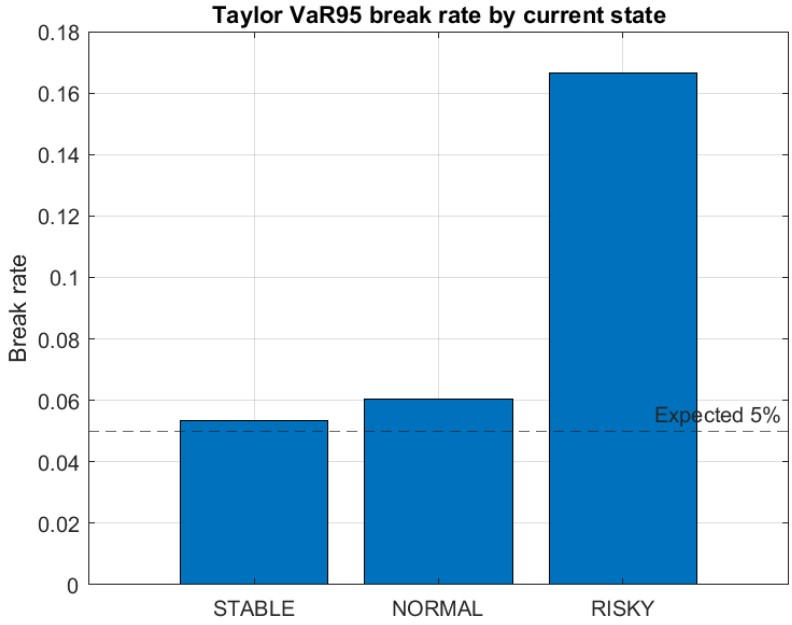


Figure 25: GARCH – VaR break-rate by market state.

4.2.4 VaR break backtest - GARCH GJR

Figures 26 and 27 show the daily PnL, VaR thresholds (95% and 99%), and the 60-day rolling exceedance rate for the GARCH model with states and soft sampling. Figures 28 and 29 show the mix of residual state distribution over time as well as state-wise break-rate, respectively.

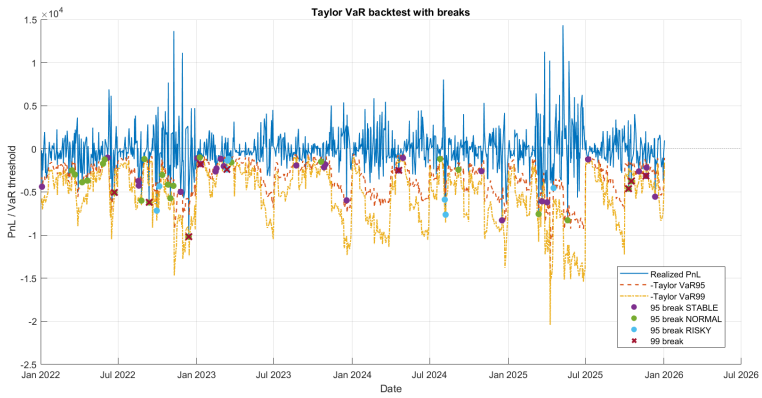


Figure 26: GJR-GARCH model – daily PnL, VaR thresholds and VaR violations.

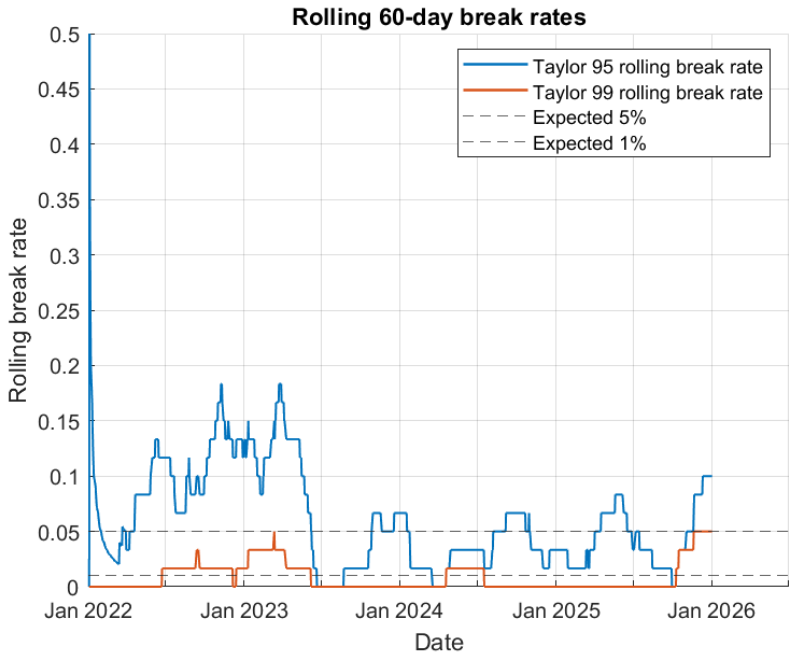


Figure 27: GJR-GARCH model – 60-day rolling VaR violation rate for 95- and 99% Taylor VaR estimations.

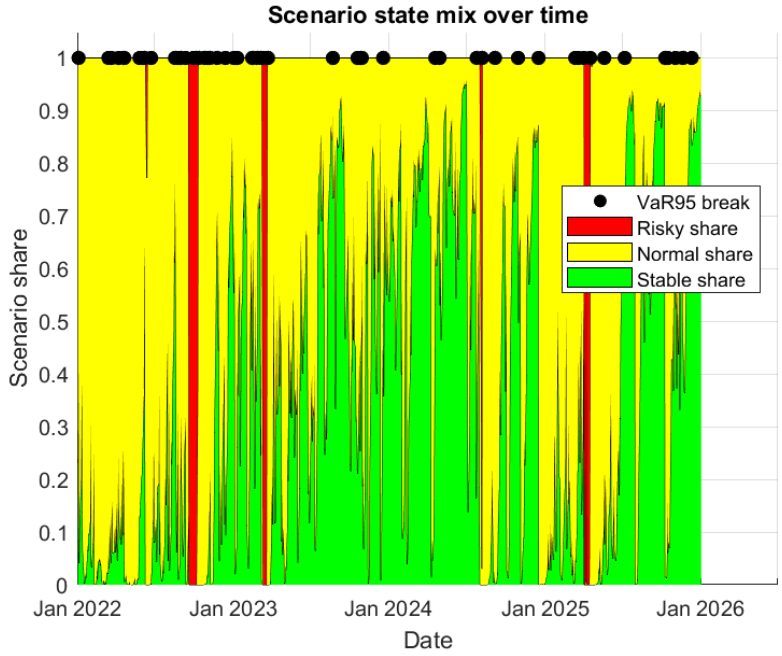


Figure 28: GJR-GARCH – Proportion of stable, normal and risky bootstrap draws.

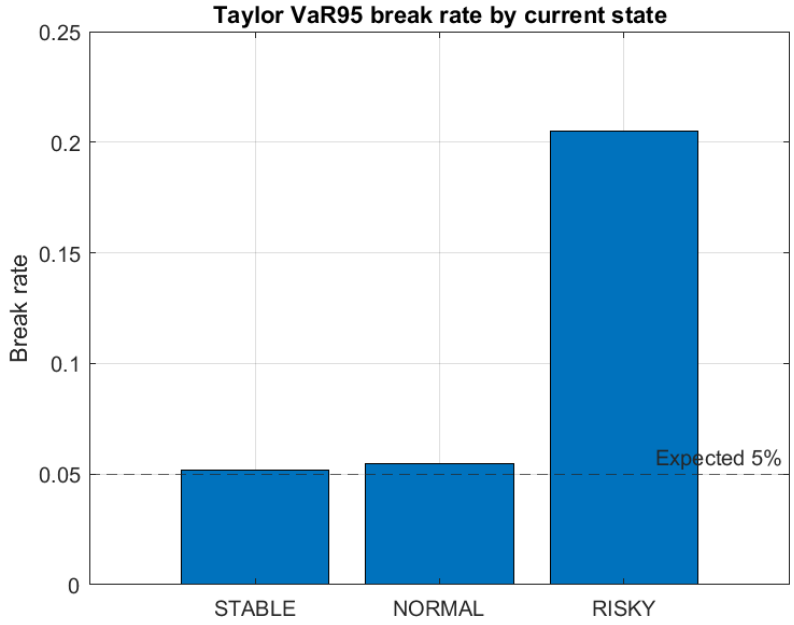


Figure 29: GJR-GARCH – VaR break-rate by market state.

4.2.5 VaR break backtest - Self-Switch Soft-States

Figures 30 and 31 show the daily PnL, VaR thresholds (95% and 99%), and the 60-day rolling exceedance rate for the GARCH model with states and soft sampling. Figures 32 and 33 show the mix of residual state distribution over time as well as state-wise break-rate, respectively.

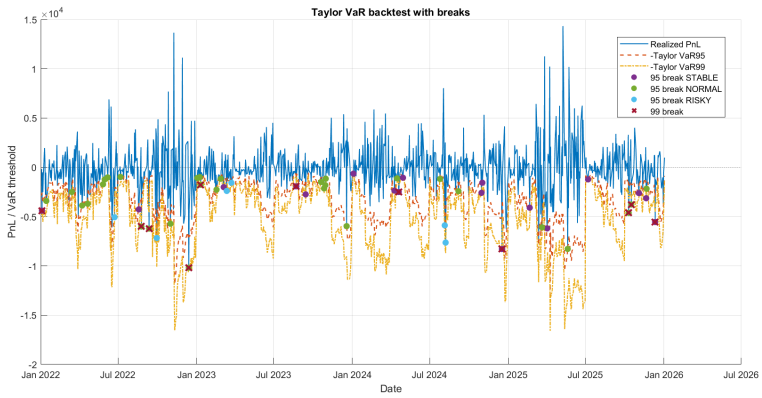


Figure 30: Self-switching GJR (soft sampling) – daily PnL, VaR thresholds and violations.

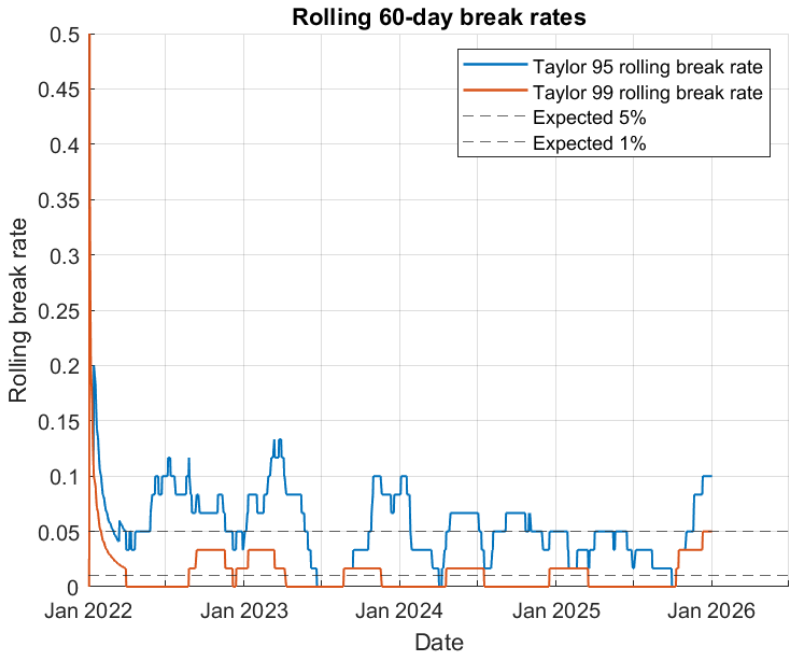


Figure 31: Self-switching GJR (soft sampling) – 60-day rolling VaR violation rate for 95- and 99% Taylor VaR estimations.

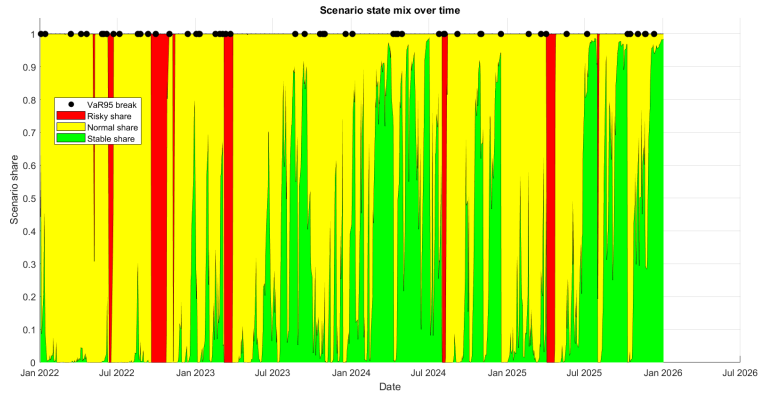


Figure 32: Self-switching GJR (soft sampling) – Proportion of stable, normal and risky bootstrap draws.

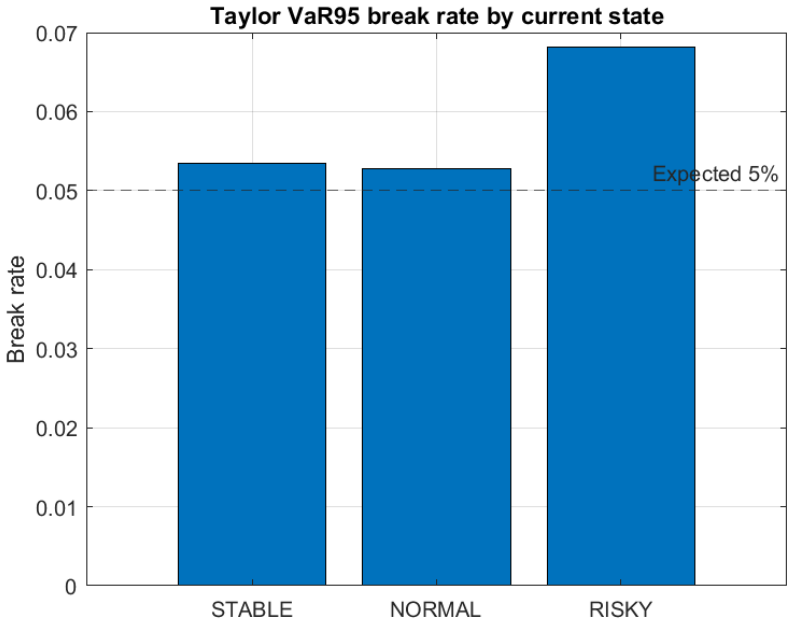


Figure 33: Self-switching GJR (soft sampling) – VaR break-rate by market state.

4.2.6 VaR break backtest - Self-Switch Hard-States

Figures 34 and 35 show the daily PnL, VaR thresholds (95% and 99%), and the 60-day rolling exceedance rate for the GARCH model with states and soft sampling. Figures 36 and 37 show the mix of residual state distribution over time as well as state-wise break-rate, respectively.

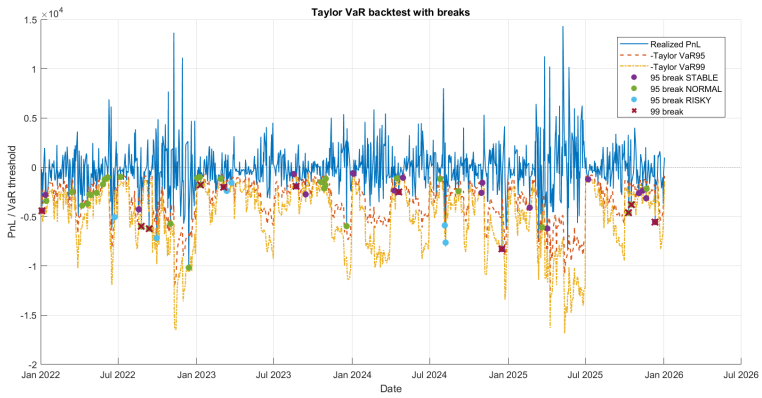


Figure 34: Self-switching GJR (hard sampling) – daily PnL, VaR thresholds and violations.

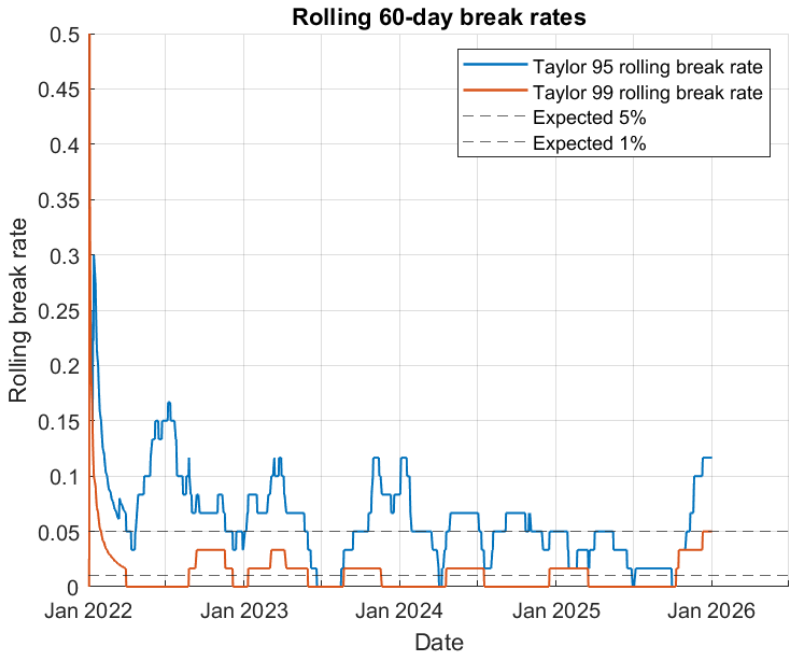


Figure 35: Self-switching GJR (hard sampling) – 60-day rolling VaR violation rate for 95- and 99% Taylor VaR estimations.

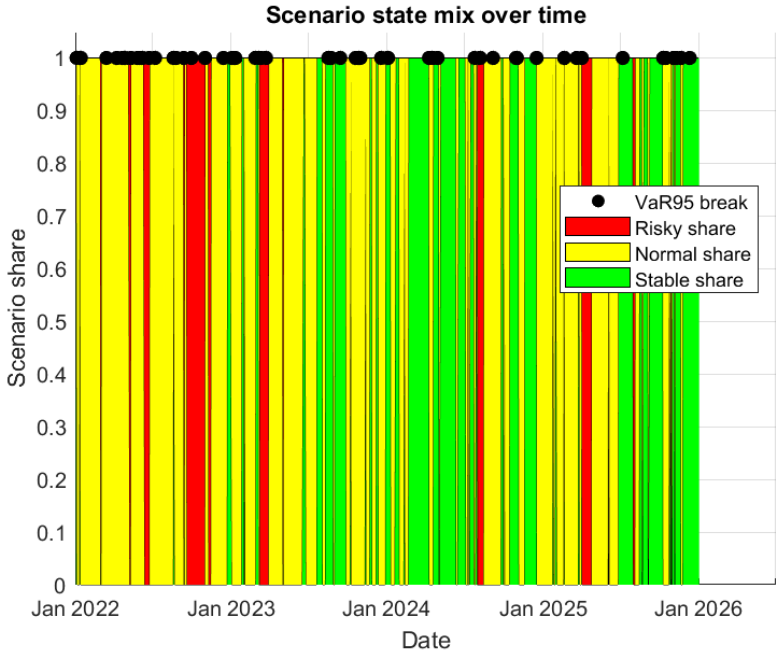


Figure 36: Self-switching GJR (hard sampling) – Proportion of stable, normal and risky bootstrap draws.

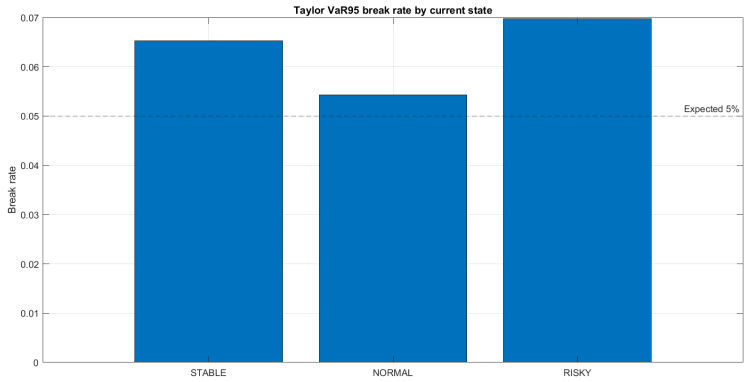


Figure 37: Self-switching GJR (hard sampling) – VaR break-rate by market state.

4.2.7 VaR break backtest - Self-Switch No States

Figures 38 and 39 show the daily PnL, VaR thresholds (95% and 99%), and the 60-day rolling exceedance rate for the GARCH model with states and soft sampling.

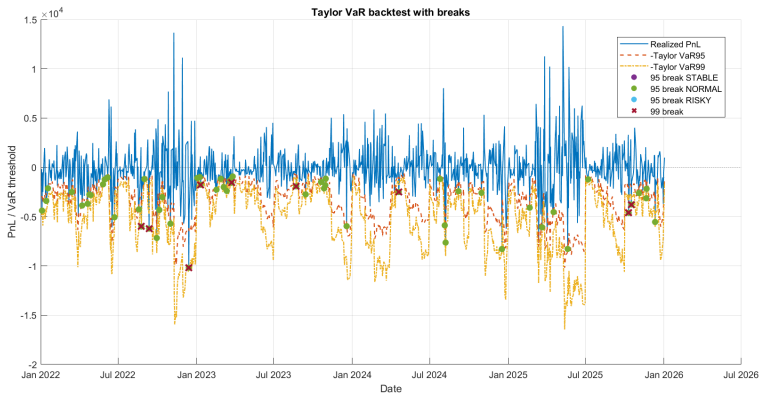


Figure 38: Self-switching GJR (single-state) – daily PnL, VaR thresholds and violations.

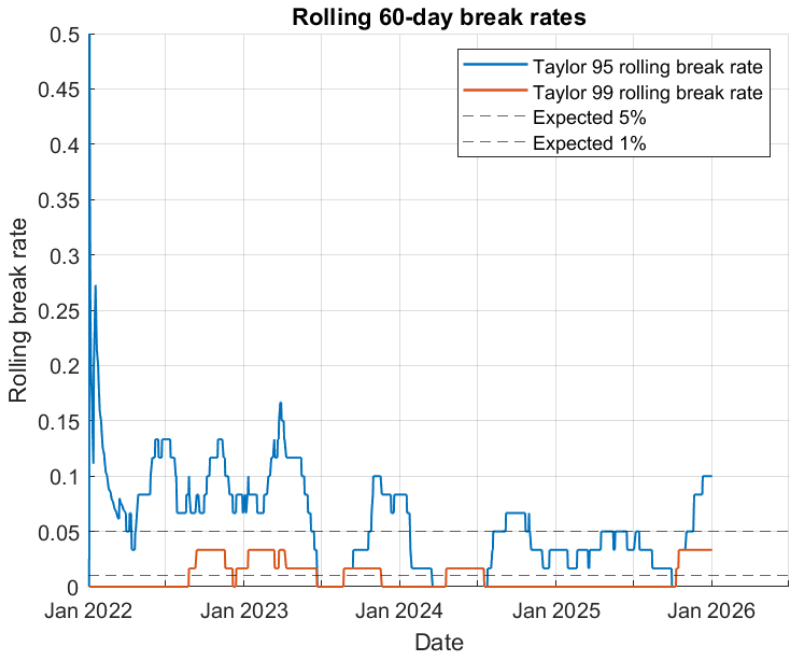


Figure 39: Self-switching GJR (single-state) – 60-day rolling VaR violation rate for 95- and 99% Taylor VaR estimations.

5 Discussion

5.1 Choice of interpolation-technique

As for the interpolation of the implied volatility surface, the first method used in the pipeline was the Bilinear technique. Generally speaking, this method is straightforward to implement and use. The best case to use this is arguably when dealing with fast calculations. Using it on live markets should come with a precaution. The interpolation here is, of course, purely linear, which removes the smoothness on the volatility surface. This can especially become a problem when the surface in question has rather high curvature. At grid points, this often has sharp "kinks", which can cause rather unnatural jumps, especially when it comes to computing the sensitivities - which is the point of using the interpolation technique in the first place.

The go-to method used in the pipeline was chosen to be the Cubic Spline method. In contrast to the bilinear interpolation, the Cubic Spline method often yields stable and continuous sensitivities instead. Another advantage of using Cubic Spline is that this is a global interpolation method, compared to Bilinear, which is only local. This means changing one grid point would only change the neighboring nodes. For Cubic Spline, the whole surface is instead shifted when one grid point is changed. This is especially useful for the sensitivity pipeline used: bumping one PC node shifts the whole surface.

Looking at figure 9 we can immediately spot the smooth switch transition between a downward decrease in implied volatility and an upwards turning point on certain moneyness nodes that Cubic Spline gives us. In contrast, the bilinear interpolation showcases sharp edge points, often at 105% moneyness. This can be a real issue if an underlying asset moves across 105% moneyness; the Gamma and Vega will instantly change values in a discontinuous spike. What can also be seen is that the Cubic Spline has, at some maturity smiles, some minor overshoots. The market data slightly flattens or rises. This might create a distinct convex "hump". This represents a negative probability density (butterfly arbitrage). A trader could buy options on the edges and sell the peak for a mathematically guaranteed risk-free profit.

More specifically, in our case, this would possibly artificially inflate our local derivative ($\frac{\partial \sigma}{\partial K}$). The VaR engine will obtain a slight fake volatility shift for a minor asset price move. In our case, this problem is not significant, since the observed overshoots are very minor, and the non-linear option pipeline case has already been shown to obtain VaR results close to full revaluation. Furthermore, the VaR-exceedance backtest shows that our modeling pipeline, which is strongly affected by the interpolation through bump and

reprice sensitivities, works quite well. It would be quite difficult to imagine a pipeline where the interpolation overshoots significantly and still has results that more than pass the Kupiec and duration tests. However, if one wants to obtain results even closer to full revaluation, one could implement the Monotone Hermite Spline interpolation instead. This has been shown to be less prone to these kinds of overshoots.

When interpolating an implied volatility surface, it is often a good idea to enforce no-arbitrage conditions. This means that the surface does not create theoretical opportunities for risk-free profits. This is particularly important for areas like pricing options or hedging, where an arbitrage-free surface ensures consistent option prices and prevents false indications regarding investing in assets without risk. In the context of VaR estimation, however, the primary goal is accurate risk measurement rather than arbitrage-free pricing. Small arbitrage violations in the interpolated surface are unlikely to significantly distort the portfolio loss distribution or the resulting VaR forecasts. Therefore, for our purpose of fast and efficient risk estimation, we do not impose strict no-arbitrage constraints on the volatility surface. Nevertheless, if the framework were to be used for constructing minimum-risk portfolios or for hedging strategies, an arbitrage-free interpolation method would be recommended.

5.2 Taylor Approximation error

The main purpose of the Taylor-based risk-factor mapping framework is to provide a computationally efficient approximation to full revaluation. According to the results, this goal could be seen as fulfilled as a whole. There were two different tests conducted to validate the Taylor-versus-full-revaluation, and both came to the conclusion that the model VaR at both levels 95 and 99 follows full revaluation closely for the most part.

5.2.1 Overall assessment

Firstly, over the whole validation period with valuation dates between 2022 and 2026, the average relative VaR difference was 1.11% for VaR95 and 2.10% for VaR99. In Market Risk Analysis, Volume IV: Value at Risk Models, Alexander provides extensive empirical tables comparing Delta-Gamma-Vega approximations to full revaluation for exotic and vanilla options portfolios. The book notes that second-order Taylor approximations frequently drift by 3% to 10% in the extreme tails (99% VaR) for complex options due to uncaptured cross-Greeks[Ale09, Chapter 7] (like Vanna and Volga). Our modeling pipeline does include cross-Greeks, such as the spot-vol and spot-yield sensitivities, but does not include, for example, the cross PC volatilities. The reason for this was purely due to the consideration of computational efficiency as something that is important in the work. Of course, having more cross-sensitivity components will reduce the efficiency,

although possibly increasing the accuracy of the engine compared to full revaluation.

The randomized option-portfolio validation leads to the same conclusion. Across the option-heavy validation set, the Taylor approximation reproduces the full-revaluation VaR estimates with small average errors. This is important because the randomized portfolios contain a wider range of strikes, maturities and option types than the rolling backtest portfolio. The results therefore suggest that the approximation is not only accurate for a single rolling portfolio, but also robust across different option exposure profiles.

The good agreement is partly explained by the fact that the implementation includes second-order spot curvature and selected spot–factor cross sensitivities. These terms capture important nonlinear effects such as the interaction between spot moves and volatility-surface shocks. However, the Taylor approximation remains a local approximation around the current market state. It does not include all possible higher-order terms or the full Hessian matrix. Therefore, some approximation error is expected, especially in the 99% tail where scenario shocks are larger.

Moving forward, one key point about our pipeline is that it supports other types of instruments, not only European call and put options. Looking at Table 13, the relative VaR estimate errors were measured on the 95-level and the 99-level also for SPX-stock, bond-portfolios, and the EURUSD-spot. The results here are a good sanity check to see that the Taylor approximation does what is intended from it. For spots, the only risk factor active is of course the spot risk factor. Sensitivities for yields, implied volatilities, and Forex, along with the second-order and cross terms, are all set to zero. This, of course, means that the mapping here is entirely linear, which intuitively yields zero error from full repricing. The same reasoning can be done for the FX spot, which is only a numerical precision from yielding the same zero error. Fixed-income instruments, on the other hand, present a slightly more complex structure. For these asset profiles, the active risk factors are isolated to the interest-rate tenor changes across the yield curve, while all volatility and equity parameters are constrained to zero. The relative error is still extremely small, as is expected. The last instrument supported is the options, which of course yields the most error to full repricing, due to the nature of its non-linearity.

5.2.2 Tail assessment

More information about the approximation error can be extracted from the tail diagnostics than from simple average Taylor–full VaR differences alone. Because VaR estimates are calculated by finding the tail quantiles of the simulated P&L distributions, it matters not just what the average error at the scenario level is, but also how Taylor and full

revaluation orderings of scenarios are different near the 5% and 1% quantile thresholds. As a result, two VaR series may look very similar over time, but isolated larger relative differences could occur in the tail.

It is clear from Figures 12 and 13 that the Taylor and full-revaluation VaR values agree closely over the entire validation period. This illustrates that the Taylor approximation successfully captures the long-term evolution of the full-revaluation benchmark. The graph of relative difference in Figure 11, however, shows that certain dates do display larger relative differences between the two series; this effect is more prominent in the 99% VaR case. This behavior is not unexpected in the tails: if risk factors are significantly shaken in a manner such that they move away from the origin around which the Taylor series is centered, higher-order terms become more important.

The analysis performed in Table 11 quantifies this finding. The average relative difference at the date-based full-period validation is just 1.11% and 2.10% at VaR95 and VaR99, respectively; the P90 errors are also modest, 2.43% and 6.07%, respectively. These average values contrast with larger maximum differences of 7.97% (VaR95) and 13.50% (VaR99). This is consistent: the average and P90 represent the typical error across dates while the maximum value indicates the maximum deviation across the whole set of points.

A somewhat different picture emerges in the randomized option-portfolio validation. Although average errors are not very large (2.09% and 1.88% at VaR95 and VaR99), the maximum relative VaR95 error is extremely large, at 20.70%. Looking at the worst-case date in Table 12, it is apparent that the error in the randomized option validation occurs on date ATM-HEAVY-006, where the full-revaluation VaR95 is only 99.35; the absolute difference in VaR is 20.56, and although small as a fraction of portfolio value, it is extremely large relative to the very small full-revaluation VaR. Because it is a scaling effect due to the very small full-revaluation VaR rather than a systematic error for all scenarios, the maximum error should not be seen as a sign of failure.

The bond-portfolio validation contrasts with the option-portfolio case. With very small maximum relative errors for VaR95 and VaR99 (both below 1%) and extremely small average errors, this validation demonstrates that the general risk-factor mapping framework performs extremely well when the underlying instruments are approximately linear. The relatively larger errors on options stem from the much more strongly nonlinear behavior of options as compared to bonds.

In conclusion, the tail diagnostics demonstrate that while the Taylor approximation may be exact at times in the tails, and while higher-order accuracy is beneficial at VaR99

for option-heavy portfolios, these issues appear to manifest themselves only at isolated points. Over time and overall, the Taylor approximation produces values that are extremely similar to full revaluation at all specified VaR levels, as evidenced by the behavior of VaR exceedances in the validation results.

5.2.3 Optimal bump-sizes

The bump-size validation should be interpreted as a robustness analysis rather than as a strict global optimization. The reason is that the finite-difference bump sizes affect several parts of the Taylor approximation simultaneously. A smaller bump is more local, but may be more sensitive to numerical noise from interpolation and floating-point rounding. A larger bump is less sensitive to such noise, but may no longer represent the local derivative around the current market state. The objective function $J(h)$ is therefore used as a structured way of comparing candidate configurations, but the final choice also considers tail performance, numerical stability and consistency across the full validation framework.

The one-at-a-time sweep in Table 14 shows that the approximation error is most sensitive to the spot-related bump sizes. For the delta bump, the best result in the local sweep is obtained for $h_\Delta = 0.0125S$, which gives the lowest objective value and the lowest average VaR99 error among the tested delta perturbations. Smaller delta bumps lead to higher errors, which suggests that numerical noise starts to dominate when the spot perturbation becomes too small. On the other hand, the largest tested delta bump, $h_\Delta = 0.0200S$, also increases the error, indicating that the finite-difference estimate becomes less local.

The gamma bump displays a clearer trade-off. Very large gamma bumps lead to systematically higher approximation errors, which is consistent with the fact that gamma is a second-order derivative and is therefore more sensitive to the choice of finite-difference step. The smallest tested gamma bump gives the lowest VaR95 error, but it also gives a larger VaR99 error. The final implementation retains $h_\Gamma = 0.0275S$, even though a slightly smaller gamma configuration gives a lower local objective value. This choice is motivated by robustness and consistency with the final validation runs rather than by strict minimization of the local sweep objective. In particular, the selected gamma bump keeps the tail error controlled while avoiding an excessively small second-order stencil.

The size of the volatility-PC bump has almost no effect on the aggregate VaR approximation error over the tested interval. The results for $h_{\mathrm{PC}} = 0.10, 0.20, 0.30$ and 0.40 are essentially identical. This indicates that the finite-difference estimates of the PC sensitivities are numerically stable over the tested range, and that the remaining Taylor–full approximation error is mainly driven by the spot delta and gamma terms rather

than by the PC bump size. Intuitively, this stability arises because the volatility-surface shocks are represented in the reduced PC coordinate system, and the option value is locally smooth in these PC directions. Moreover, the implied volatility at each node is a linear function of the PC scores (the PCA factors are linear combinations of the original log-IV changes), so the finite-difference estimates of the PC sensitivities remain stable over the tested bump range. The final value $h_{\mathrm{PC}} = 0.20$ is therefore retained as a moderate interior choice. Overall, the sweep supports the selected final bump configuration,

$$h_{\Delta} = 0.0125S, \qquad h_{\Gamma} = 0.0275S, \qquad h_{\mathrm{PC}} = 0.20.$$

Although this configuration is not the exact minimizer of the local objective function, it lies in a region where the Taylor approximation error remains small and stable. The final Taylor-versus-full-revaluation validation confirms that these bump sizes produce low approximation errors across both the full date-based validation and the randomized option-portfolio validation. The bump-size study should therefore be viewed as evidence that the final configuration is numerically robust, rather than as a claim of global optimality.

5.3 VaR results

The different models used to calculate daily VaR for the option portfolio produced interesting results. Several general observations can be made, as well as model-specific findings.

5.3.1 Models

The baseline EWMA model without state-dependent residual sampling had low p-values for both the Kupiec test and the duration test. It was the only model that failed the Kupiec test at the 5% significance level for the 95% VaR. The duration test also gave low p-values for both the full revaluation and Taylor approximation, indicating that violations are not independent over time. Figure 20 illustrates this phenomenon too; the cluster of violations between 2022 and 2023 is visible.

The GARCH multi-state model improved upon EWMA. Its Kupiec p-value for the Taylor 95% VaR exceeded 0.1, meaning the null hypothesis of correct unconditional coverage is not rejected at the 90% level. However, the duration test did not improve since the p-values remained below 0.05 for both Taylor and full revaluation at 95%. This suggests that while the GARCH structure helps with unconditional coverage, the residual clustering persists.

The GJR-GARCH multi-state model just barely passed the duration test at the 90% level

while also achieving a solid break rate, with p-values around 0.19 for the unconditional coverage and 0.11 for the duration test. The inclusion of the leverage term γ for the risk factors appears to help with reducing some of the clustering, likely because asymmetric volatility responses to negative shocks are present.

The best model according to the statistical tests was the self-switching GJR-GARCH with soft residual sampling. It produced a break rate of 5.43% at 95%, which is close to the expected 5%, a Weibull shape parameter of 0.97, which is very close to 1, and a duration test p-value of 0.77. This indicates no sign of clustering.

Comparing the three variants of the self-switching model (soft sampling, hard sampling, and no states) reveals the usefulness of state-dependent residual sampling for reducing volatility clustering among the VaR-breaks. Model 6 without any state conditioning had a reasonable break-rate and a lower duration p-value (0.175) than the other two with states. The hard-sampling version achieved a similar p-value for both tests compared to model 6, although still a slightly higher duration test p-value at 0.198. The softmax transition between states does improve on both of these regarding duration on the validation period. Regarding VaR estimation for a project like this, it seems like a good option to include states for both the volatility model parameters as well as the residuals; even more, using a soft sampling method for the residuals seems to be a good option.

5.3.2 Rolling Averages

While statistical tests provide a useful baseline for validation, assessing the model solely by them can be misleading. Even a model that passes the Kupiec and duration tests may exhibit periods where the violation rate is temporarily high or low. Rolling break rates help to detect such local deviations. For comparison, we simulate a perfect VaR model in which violations occur as independent Bernoulli trials with probability 0.05. Figure 40 shows the rolling 60-day break rates for one simulation of 5000 days. Figure 41 shows the first 1000 days of that simulation more zoomed in. As we can see, the rolling rates move randomly around the nominal 5% line. These patterns illustrate that even an ideal model will show periods of above- or below-average violation rates.

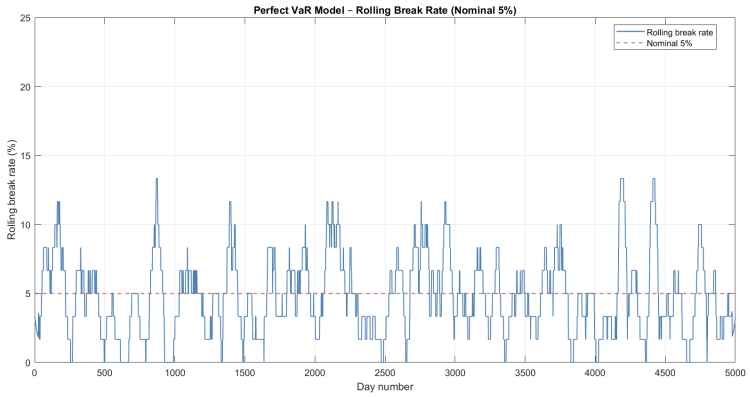


Figure 40: Rolling 60-day break rate for a simulated perfect VaR model (5000 days, nominal 5%).

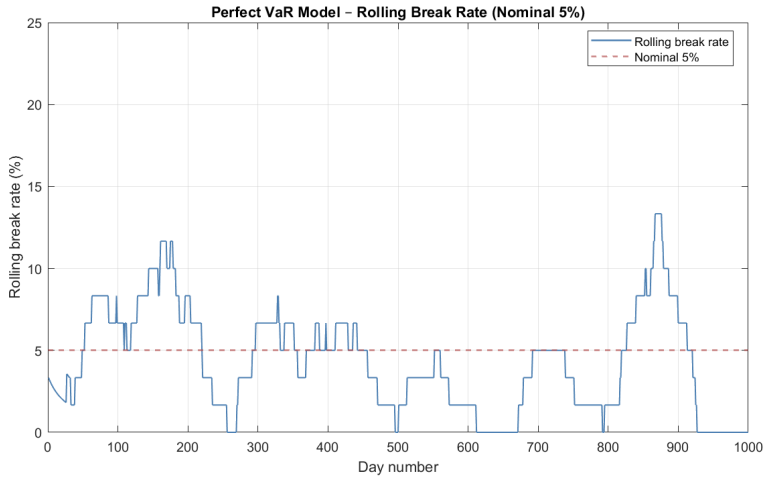


Figure 41: Rolling 60-day break rate for a simulated perfect VaR model (1000 days, nominal 5%).

Looking at the best model, the GJR-GARCH with soft residual sampling, the rolling averages shown in Figure 31 do not appear unrealistic relative to these simulations. There is an all-time high in the first half of 2023, at around a 12-13% break rate. This level is also achieved in the perfect model within 1000 days. If we instead look at the all-time lows, we see that both the used model and the perfect model hit the floor of 0 violations. This is good news, confirming that the results are not unreasonable.

5.3.3 General

A notable result is that every model performed well at the 99% VaR level. Both the break rates and the duration test p-values were acceptable across all models. At only the 95% level do the larger differences appear.

The differences between full revaluation and the Taylor approximation in terms of VaR estimation were minor. In almost all cases, the break counts and duration p-values were very similar, and the Kupiec p-values also matched closely. They did differ to a small extent; however, the difference was not significant enough to make the Taylor models worse over the validation period.

In general, all models exhibited above-average VaR breaches during the second half of 2022, as seen in the rolling exceedance rates and backtest plots. The year 2022 saw several sharp declines in the market, driven by global events. It is possible that the models underestimated risk during this period to some extent. However, even a perfect model would experience periods of a bit more frequent breaks and periods of less over a 60-day window simply due to randomness.

Across models, the risky state consistently showed a higher break rate than the stable and normal states. This suggests that while the models may be able to identify when the market is risky, they may not adjust the risk levels sufficiently to fully compensate. In the self-switching models, this difference was, however, relatively small. This points to the idea that when introducing state-dependent volatility model parameters, it becomes easier to adjust the risk during risky days.

Building a VaR model that perfectly approximates risk every day is challenging. Relying only on past market movements makes it difficult to estimate future risk accurately. Market movements are complex, and by using stochastic models such as GARCH, we inevitably impose certain assumptions. Some assumptions are of course necessary, but this also imposes limitations. With this in mind, the results achieved over the validation period are good. Overall, the self-switching GJR-GARCH model with soft residual sampling during the validation period seems to produce acceptable estimates.

5.4 Practical use

Overall, the results from the VaR backtesting were positive. The fact that the best model had high p-values for both statistical tests is a sign of good potential. The Taylor approximation, in particular, was able to be close to the full reevaluation, meaning the downside of using this framework seems to be small.

The results indicates that the proposed framework is valuable in practice as an efficient approximation engine for VaR calculations for portfolios of nonlinear instruments.

The VaR backtesting results are largely favorable, particularly for the best-performing GJR-GARCH-based specifications. The fact that the top models are not rejected by the unconditional coverage and duration-based independence tests is reassuring in that the generated VaR estimates seem broadly aligned with realized P&L outcomes. These tests should not be interpreted, however, as proof of model correctness. A high p-value merely implies that there is insufficient data to reject the model under the given test.

The Taylor approximation is also demonstrated to be close to full revaluation under the validation tests. This is a critical aspect for risk-management applications where the primary shortcoming of risk-factor mapping is approximation error. The tests suggest this shortcoming is limited in the demonstrated setting. Over the full date-based validation period, the Taylor and full-revaluation VaR estimates are highly correlated, and their exceedance counts are very similar. The randomized option-portfolio validation further reinforces that the approximation remains stable across a broad range of option-heavy portfolios, bolstering the use of Taylor-based risk-factor mapping as a computationally less expensive alternative to full revaluation for situations involving the evaluation of numerous scenarios.

In a practical risk-management context, such a framework will likely be most useful when frequent risk calculations are required, such as for the purpose of daily VaR production, intraday risk monitoring, pre-trade risk checks, and stress testing or scenario-based what-if analysis. While full revaluation may still be used for end-of-day reporting and for the calculation of VaR on portfolios composed of highly exotic options, a Taylor-based mapping method can deliver a fast approximation that captures the primary risk signal without incurring the prohibitive computational costs of full repricing on numerous scenarios, particularly for option portfolios.

In a portfolio consisting of 10,000 options, and let us say a framework using 10,000 historical scenarios, full revaluation would need 100,000,000 pricing calculations. This can be compared to our framework, which, with the use of Bump and Reprice sensitivities and with an approximation formula, would require:

$$10,000 \text{ options} \cdot 91 \text{ calculations} = 910,000 \text{ calculations}.$$

The constant 91 comes from the use of a total of 91 sensitivity-related pricing calculations, including deltas, gammas, volatility PCs, rate tenors and higher order terms. This means that full revaluation is almost 110 times more expensive than the approximation. This can also be seen from an environmental point of view. In large-scale risk systems, VaR calculations may be run daily, intraday or across many desks and portfolios. Reducing the

number of expensive option valuation calls can therefore reduce CPU time and, indirectly, electricity consumption. This should not be interpreted as a direct measurement of energy savings, since actual energy consumption depends on hardware, implementation details, parallelization, data center efficiency and the carbon intensity of the electricity used. However, the comparison shows that the proposed framework has a potential sustainability benefit when applied at scale.

An additional advantage of the method for practical use is the provision of interpretable risk-factor exposures. Portfolio P&L can be decomposed into contributions from spot, FX, interest rate, and volatility-surface risk factors, making it possible to determine which sources of risk are driving the VaR estimate. Use of volatility-surface principal components is particularly helpful, by reducing the high-dimensionality of an implied volatility surface to a manageable set of systemic volatility shocks. This makes the approach tractable without sacrificing exposure to major movements in the volatility surface.

This framework should be applied carefully, considering its potential limitations. The Taylor approximation, by its very nature, is local, and the approximation may become less accurate when scenarios are far from the current state of the market. This can be especially critical in the extreme tails where the impact of significant spot or rate shocks or volatility changes may call for higher-order terms. Although errors are generally managed in the validation, they are not eliminated. As a result, periodic comparison with full revaluation may be necessary, particularly for portfolios with large nonlinear exposures or during turbulent market conditions.

This method is subject to several implementation choices, such as the selection of the specific risk factors to be used, the volatility model, method for state-classification, manner of interpolation of the implied volatility surface, and finite-difference bumping parameters. These will need to be recalibrated regularly if the portfolio or market conditions shift significantly. A production implementation would need to feature regular comparisons between Taylor-based and full revaluation, with close monitoring of exceedance frequency, stress testing based on individually crafted scenarios, etc.

We would like to highlight that, no matter how well a VaR model performs on historical backtests, it is impossible to guarantee that it will perform equally well in the future. Financial markets are hard to predict, and changes in market dynamics may invalidate the assumptions underlying our model. Therefore, the favorable backtest results should not be interpreted as a guarantee of future performance. They provide only evidence that the model could capture historical risk dynamics well; continuous monitoring, periodic recalibration, and regular comparison with other VaR models are a good idea.

In summary, the evidence presented suggests that Taylor-based risk-factor mapping can serve as a practical compromise between full revaluation and less-sophisticated linear VaR approximation. It captures significant nonlinearities through gamma, higher-order volatility terms, and specific cross sensitivities without the significant computational burden of repricing every instrument for every simulated scenario. In this regard, the model can be viewed as an efficient risk calculation engine for repeat scenario valuation applications, with full revaluation serving as a benchmark and validation tool, rather than as the primary calculation method in every instance.

5.5 Future Work

There are a few different future work ideas that we believe would be an interesting extension to this project. Firstly, using PCA to find the risk factors for the implied volatility surface does not have to be the best option. Another option that has shown capability to be more accurate than PCA for the volatility surface is using an autoencoder. This is especially true for implied volatilities in between nodes used for PCA [Nag22].

The most natural extension to this project would be to validate over multiple days, creating a multi-day VaR model. The present implementation focuses on one-day VaR. Extending the model to multi-day horizons would require either simulating risk-factor paths over several days or introducing an appropriate horizon-scaling method. For option portfolios, this is not trivial, since the portfolio must also be aged through time and option maturities, moneyness and sensitivities change over the horizon.

Another important extension would be to include dividends more explicitly in the option pricing framework. In the current implementation, dividend yields are assumed to be zero for simplicity. While this simplification is not expected to dominate the VaR results for the tested portfolios, dividend assumptions may become important for longer-dated equity options or portfolios with large directional exposure. Including continuous dividend yields or dividend curves would make the option-pricing step more realistic.

The self-switch GARCH-GJR was successful enough in capturing the time-varying volatility of the risk factors. There is, of course, no reason to believe that this is the optimal way of doing so. It would be interesting to try and model the volatility with other models and compare them with the self-switch model. One example could be to introduce a Markov-switching GARCH (GS-GARCH) instead. This would fit naturally into the regime structures already implemented by introducing transition probabilities. GS-GARCH has shown the ability to outperform single-regime models too [Ard+18]. During this project,

another two-state GARCH model similar to GS-GARCH was introduced using two sets of parameters for each risk factor. This structure did not perform well on the statistical test and was therefore not included in the final.

The volatility-surface interpolation method could also be developed further. The current implementation uses cubic-spline interpolation, which provides a smooth volatility surface and performed well in the validation. However, alternative interpolation methods such as shape-preserving cubic interpolation, monotone splines, or arbitrage-aware volatility surface construction could be investigated. Such methods may reduce the risk of artificial interpolation curvature while preserving the smoothness needed for stable finite-difference sensitivities.

A further extension would be to replace parts of the bump-and-reprice sensitivity calculation with automatic differentiation. Bump-and-reprice is flexible and easy to validate, but it requires a careful choice of finite-difference bump sizes and can become computationally expensive when many second-order and cross sensitivities are included. Automatic differentiation could provide more accurate local derivatives and reduce the dependence on bump-size calibration. However, implementing automatic differentiation through the full valuation pipeline would require careful treatment of interpolation, curve construction and option-pricing routines. It would also be necessary to evaluate whether the computational gain remains significant when second-order and cross sensitivities are required.

The framework could also be extended to cover a broader set of instruments. Although the empirical validation focuses primarily on SPX option portfolios, the architecture is intended for multi-asset risk-factor mapping. Future work could include a larger set of equity options, FX options, futures, swaps, swaptions, or credit-sensitive instruments. This would make it possible to evaluate how well the Taylor-based mapping performs for portfolios with more diverse nonlinear risk exposures.

Another practical extension would be to make the implementation less hard-coded and more configuration-driven. Some parts of the current framework are tailored to the specific data set, risk-factor names and instrument setup used in the project. A more general production-oriented framework would allow risk factors, volatility surfaces, curves, instruments, bump sizes and validation settings to be specified through external configuration files. This would make the framework easier to extend, test and apply to new portfolios.

The current framework assumes that estimated implied volatilities are available as data to use. This might not always be true in practice. While a lot of data is easy to find publicly for free, implied volatilities might not be. An idea for extending the framework

even more could be to implement an implied volatility estimator given daily option prices. This makes the system usable even if the data available is limited.

Finally, future work could include explicit stress testing in addition to VaR. The current framework generates scenario distributions for VaR estimation, but the same risk-factor mapping structure could also be used for manually specified stress scenarios, such as large equity shocks, volatility-surface shifts, interest-rate curve steepening or combined market stress events. This would be a natural practical extension, since risk managers often use VaR and stress testing together to understand both typical tail losses and extreme but plausible portfolio outcomes.

6 Summary

In this project, we developed a comprehensive Value at Risk framework based on risk factor mapping. The core idea is to express a portfolio's value change as a function of a reduced set of underlying risk factors. This mapping makes the approach scalable to very large and mixed asset portfolios that may contain many options, equities, bonds, and currencies, while keeping the VaR calculation computationally efficient.

The methodology combines dimensionality reduction via principal component analysis of the implied volatility surface, dynamic volatility modeling, and state-dependent residual sampling within a Filtered Historical Simulation framework. The validation uses real prices of European options on the S&P 500 index.

A self-switching GJR-GARCH model uses a global GARCH(1,1) to determine the volatility regime and smoothly blends low-volatility and high-volatility GJR parameters using a linear weight function. The FHS uses bootstrap on a volatility-based market state with stable, normal, and risky choices. It applies softmax sampling to smooth state transitions if chosen. The portfolio P&L is approximated by a second-order Taylor expansion using pre-computed Greeks, which is shown to be almost as accurate as full revaluation while being computationally far more efficient.

Backtesting over a four-year validation period (2022-2026) demonstrates that the self-switching GJR-GARCH model with soft residual sampling achieves break rates very close to the nominal 5% and 1% levels, passes both the Kupiec unconditional coverage test and the Christoffersen–Pelletier duration-based independence test, and exhibits no evidence of violation clustering. In contrast, the EWMA baseline model significantly underestimates risk. The Taylor approximation yields nearly identical backtest results to full revaluation, confirming its suitability for daily VaR calculations. The entire framework is computationally efficient, allowing rapid re-estimation and scenario generation for large portfolios.

Future extensions could include multi-day VaR forecasting and using an autoencoder for implied volatility factor extraction.

References

[Ale08] Carol Alexander. *Pricing, Hedging and Trading Financial Instruments*. 1st ed. John Wiley & Sons, 2008.

[Ale09] Carol Alexander. *Market Risk Analysis, Volume IV: Value at Risk Models*. Chichester, UK: John Wiley & Sons, 2009.

[Ard+18] David Ardia et al. "Forecasting risk with Markov-switching GARCH models: A large-scale performance study". In: *International Journal of Forecasting* 34.4 (2018), pp. 733–747.

[Bas19] Basel Committee on Banking Supervision. *Minimum capital requirements for market risk*. Standard d457. Basel, Switzerland: Bank for International Settlements, 2019. URL: https://www.bis.org/bcbs/publ/d457.pdf.

[Bol86] Tim Bollerslev. "Generalized autoregressive conditional heteroskedasticity". In: *Journal of econometrics* 31.3 (1986), pp. 307–327.

[Bou+01] Jean-Philippe Bouchaud et al. "More stylized facts of financial markets: leverage effect and downside correlations". In: *Quantitative Finance* 1.6 (2001), pp. 575–584.

[BS73] Fischer Black and Myron Scholes. "The Pricing of Options and Corporate Liabilities". In: *The Journal of Political Economy* 81.3 (May 1973), pp. 637–654. URL: http://www.jstor.org/stable/1831029.

[Con01] Rama Cont. "Empirical properties of asset returns: stylized facts and statistical issues". In: *Quantitative Finance* 1.2 (2001), pp. 223–236.

[CP04] Peter Christoffersen and Denis Pelletier. "Backtesting Value-at-Risk: A Duration-Based Approach". In: *Journal of Financial Econometrics* 2.1 (2004), pp. 84–108.

[Die23] Stefan Diehl. *Optimization*. 1st ed. Studentlitteratur, 2023.

[GJR93] Lawrence R Glosten, Ravi Jagannathan, and David E Runkle. "On the relation between the expected value and the volatility of the nominal excess return on stocks". In: *The journal of finance* 48.5 (1993), pp. 1779–1801.

[GM15] Pedro Gurrola-Perez and David Murphy. *Filtered historical simulation Value-at-Risk models and their competitors*. Staff Working Paper 525. London, UK: Bank of England, 2015. URL: https://www.bankofengland.co.uk/working-paper/2015/filtered-historical-simulation-value-at-risk-models-and-their-competitors.

[Hul21] John C. Hull. *Options, Futures, and Other Derivatives*. 11th. Pearson, 2021.

[Hul23] John. C. Hull. *Risk Management and Financial Institutions*. 6th ed. Wiley, 2023.

[Jak21] Andreas Jakobsson. *An Introduction to Time Series Modeling*. 4th ed. Studentlitteratur, 2021.

[Jor06] Philippe Jorion. *Value at Risk*. 3rd ed. McGraw-Hill, 2006.

[JP 96] J.P. Morgan/Reuters. *RiskMetrics—Technical Document*. Tech. rep. New York: J.P. Morgan and Reuters, 1996. URL: https://www.msci.com/documents/10199/5915b101-4206-4ba0-aee2-3449d5c7e95a.

[Kup95] Paul H. Kupiec. "Techniques for Verifying the Accuracy of Risk Measurement Models". In: *The Journal of Derivatives* 3.2 (1995), pp. 73–84.

[LMN15] Erik Lindström, Henrik Madsen, and Jan Nygaard Nielsen. *Statistics for Finance*. 1st ed. Chapman and Hall/CRC, 2015.

[Nag22] Rogier Nagelkerken. "Factor Models for the Implied Volatility Surface". Master's thesis. Erasmus University Rotterdam, Erasmus School of Economics, Nov. 2022.

[PLA26] Pedro Pablo Pérez Velasco, Mengjue Lu, and Daniel Arrieta. "Curved Greeks: A Geometric Layer for Option P&L Adjustments". In: (2026). Working paper.

[Pre+07] William H. Press et al. *Numerical Recipes: The Art of Scientific Computing*. 3rd. Cambridge University Press, 2007.

[Reb04] Riccardo Rebonato. *Volatility and Correlation: The Perfect Hedger and the Fox*. 2nd. Chichester, UK: John Wiley & Sons, 2004.

[Tsa10] Ruey S. Tsay. *Analysis of Financial Time Series*. 3rd. Wiley Series in Probability and Statistics. Hoboken, New Jersey: John Wiley & Sons, 2010.

A GJR-GARCH log-likelihood functions

A.1 Normal assumption

The log-likelihood for a GJR-GARCH(1,1) model with conditionally normal innovations is

$$\ell(\theta) = -\frac{1}{2}\sum_{t=1}^{T}\left(\log(2\pi) + \log(\sigma_t^2) + \frac{\varepsilon_t^2}{\sigma_t^2}\right),$$

where

$$\sigma_t^2 = \omega + (\alpha + \gamma I_{t-1})\varepsilon_{t-1}^2 + \beta\sigma_{t-1}^2,$$

$I_{t-1} = 1$ if $\varepsilon_{t-1} < 0$ and 0 otherwise, and $\theta = (\omega, \alpha, \gamma, \beta)$.

A.2 Student-t assumption

Under the assumption of standardized Student-t innovations with $\nu > 2$ degrees of freedom, the log-likelihood is

$$\ell(\theta) = \sum_{t=1}^{T}\left[\log\Gamma\Big(\frac{\nu+1}{2}\Big) - \log\Gamma\Big(\frac{\nu}{2}\Big) - \frac{1}{2}\log\big(\pi(\nu-2)\sigma_t^2\big) - \frac{\nu+1}{2}\log\left(1 + \frac{\varepsilon_t^2}{(\nu-2)\sigma_t^2}\right)\right],$$

with the same variance recursion as above. The parameter vector includes ν: $\theta = (\omega, \alpha, \gamma, \beta, \nu)$.

B Nelder–Mead method

Algorithm parameters

The following parameters are used in the implementation:

$$\alpha = 2, \qquad \beta = 3, \qquad \gamma = 1.5, \qquad \delta = 0.5.$$

The termination tolerance is set to a simplex size less than 10^{-5}, and the maximum number of iterations is 100 000.

The initial simplexes shown in 17 and 18 are used in the implementation:

Vertex	ω_G	α_G	β_G	ν_G
1	0.10	0.05	0.90	10.0
2	0.12	0.05	0.90	10.0
3	0.10	0.07	0.90	10.0
4	0.10	0.05	0.85	10.0
5	0.10	0.05	0.90	12.0

Table 17: Initial simplex for Student-t GARCH(1,1) parameter optimization.

Vertex	ω_G	α_G	γ_G	β_G	ν_G
1	0.10	0.05	0.05	0.90	10.0
2	0.12	0.05	0.05	0.90	10.0
3	0.10	0.07	0.05	0.90	10.0
4	0.10	0.05	0.07	0.90	10.0
5	0.10	0.05	0.05	0.85	10.0
6	0.10	0.05	0.05	0.90	12.0

Table 18: Initial simplex for Student-t GJR-GARCH(1,1) parameter optimization.

In order to maintain correct parameter restrictions on the GARCH and GJR-GARCH, the value of the negative log-likelihood function was set to infinity for simplexes that break any restriction.

Algorithm pseudocode

The algorithm maintains a simplex $\Delta = \{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_{n+1}\}$ of $n+1$ points. At each iteration, the points are sorted so that

$$f(\mathbf{x}_1) \leq f(\mathbf{x}_2) \leq \cdots \leq f(\mathbf{x}_{n+1}).$$

The centroid $\mathbf{x}_*$ is defined as

$$\mathbf{x}_* = \frac{1}{n} \sum_{i=1}^{n} \mathbf{x}_i.$$

The algorithm proceeds according to the following steps (Algorithm 5 in [Die23]):

Algorithm 1 Nelder–Mead method

Input: initial simplex $\Delta = \{\mathbf{x}_1, \ldots, \mathbf{x}_{n+1}\}$; parameters $\alpha > 1$, $\beta > 2$, $1 < \gamma < 2$, $0 < \delta < 1$
Output: approximate minimizer $\mathbf{x}_1$
while termination criterion not met **do**
    Sort Δ so that $f(\mathbf{x}_1) \leq f(\mathbf{x}_2) \leq \cdots \leq f(\mathbf{x}_{n+1})$
    $\mathbf{x}_* \leftarrow \frac{1}{n} \sum_{i=1}^{n} \mathbf{x}_i$
    $\mathbf{d} \leftarrow \mathbf{x}_* - \mathbf{x}_{n+1}$
    $\mathbf{x}_r \leftarrow \mathbf{x}_{n+1} + \alpha \mathbf{d}$
    if $f(\mathbf{x}_r) < f(\mathbf{x}_n)$ **then**
        Replace $\mathbf{x}_{n+1}$ with $\mathbf{x}_r$ in Δ
        if $f(\mathbf{x}_r) \leq f(\mathbf{x}_1)$ **then**
            $\mathbf{x}_e \leftarrow \mathbf{x}_{n+1} + \beta \mathbf{d}$
            if $f(\mathbf{x}_e) < f(\mathbf{x}_r)$ **then**
                Replace $\mathbf{x}_r$ with $\mathbf{x}_e$ in Δ
            end if
        end if
    else
        if $f(\mathbf{x}_r) < f(\mathbf{x}_{n+1})$ **then**
            $\mathbf{x}_{co} \leftarrow \mathbf{x}_{n+1} + \gamma \mathbf{d}$
            if $f(\mathbf{x}_{co}) \leq f(\mathbf{x}_r)$ **then**
                Replace $\mathbf{x}_{n+1}$ with $\mathbf{x}_{co}$ in Δ
            else
                $\mathbf{x}_{ci} \leftarrow \mathbf{x}_{n+1} + \gamma \mathbf{d}$
                if $f(\mathbf{x}_{ci}) < f(\mathbf{x}_{n+1})$ **then**
                    Replace $\mathbf{x}_{n+1}$ with $\mathbf{x}_{ci}$ in Δ
                else
                    for $i = 2$ to $n + 1$ **do**
                        $\mathbf{x}_i \leftarrow \mathbf{x}_1 + \delta(\mathbf{x}_i - \mathbf{x}_1)$
                    end for
                end if
            end if
        end if
    end if
end while

The termination criterion is based on the size of the simplex, i.e.

$$\max_{i=2,\ldots,n+1} \|\mathbf{x}_i - \mathbf{x}_1\| < \text{tolerance}.$$