

Accuracy vs. Efficiency: A Comparative Study of Historical Simulation and Monte Carlo Methods for VaR Forecasting in VIX Markets

Muyang Wang

*School of Mathematics and Physics, Nanjing Institute of Technology, Nanjing, China
2325364184@qq.com*

Abstract. The present study explores the forecasting performance of two distinct methods: Historical Simulation (HS) and Monte Carlo (MC). The aforementioned approaches find application in the estimation of VaR of the CBOE Volatility Index (VIX), a benchmark of paramount importance in the assessment of market risk. As financial institutions increasingly rely on VaR models to quantify volatility risk, the choice between computationally efficient but potentially oversimplified HS approaches and MC methods, though more sophisticated, is a key operational decision. This study employs a rolling-window framework with 10-year calibration periods to analyse a three-decade period of VIX data (1990-2023). This methodology is utilised in order to draw comparisons between standard HS, crisis-adjusted HS, and MC simulation incorporating Ornstein-Uhlenbeck processes. The findings reveal that the MC approach attained a statistically significant 12.7% reduction ($p < 0.01$) in 95% VaR forecast errors when compared against HS during normal volatility periods ($VIX < 25$). Furthermore, the MC approach exhibited superior crisis performance, with breach rates deviating 8.2% from theoretical expectations, in contrast to the HS approach, which deviated 31.4%. However, it is important to note that this was achieved at a substantial computational cost of 117 times the processing time (9.4 seconds vs. 0.08 seconds per estimation). The findings of the study provide a decision framework grounded in empirical evidence. It is asserted that the implementation of weighted HS is to be recommended for scenarios involving high-frequency monitoring, and that MC is to be employed for stress testing scenarios. The robustness of the decision framework has been demonstrated to be reliable on multiple occasions, as evidenced by its application during significant market events. These include the 2008 financial crisis and the 2020 pandemic volatility spike. The present text provides practitioners with guidance for the implementation of volatility risk management systems, which has been empirically validated.

Keywords: Value-at-Risk (VaR), CBOE Volatility Index (VIX), Historical Simulation (HS), Monte Carlo (MC), Ornstein-Uhlenbeck (OU)

1. Introduction

The VIX Index is widely regarded as the primary metric for gauging market expectations regarding the short-term volatility of the price of S&P 500 Index (SPX) options. Since its introduction in 1993, the VIX has been regarded by many as the world's premier barometer of investor sentiment and market volatility. The VIX is a forward-looking index that measures the volatility investors anticipate, and it was introduced for two principal purposes. Firstly, the model is designed to provide a benchmark of expected short-term market volatility. Secondly, to facilitate comparison with historical levels, minute-by-minute values have been calculated using index option prices, with data available since the beginning of January 1986. It has been argued that market anxiety levels during the October 1987 crash – widely regarded as the worst stock market crash since the Great Depression – were particularly pronounced. This provides useful benchmark information with which to assess the level of market turbulence thereafter. The VIX is also designed to serve as an index against which futures and options contracts on volatility are written.

The VaR model was first proposed in 1993 and implemented by JP Morgan in its risk management system the following year. Since then, it has been widely adopted by the financial sector. Licensed financial institutions, such as banks, insurance companies and fund management companies, utilize the VaR model to manage risk. In the event of significant market price fluctuations, the VaR risk management model uses statistical distributions within a specific probability (confidence interval) to calculate the upper limit of a company's maximum single-day loss [1]. For banks, securities firms and fund management companies that manage investment portfolios, the VaR model can also calculate the maximum value of a portfolio's loss by considering the correlations between different types of investment within the portfolio. The emergence of the VaR model filled a gap in the financial sector's quantitative risk assessment and facilitated traders' ability to calculate risk exposure and set stop-loss levels [2]. The primary innovations of this paper are as follows:

1. This paper proposes a novel approach by which a mean reversion model (OU process) is systematically applied to VIX VaR prediction. This is the first time this has been done, and it rectifies the bias issues in traditional GBM models for dynamic VIX modelling [3].

2. The approach outlined in this paper proposes a dynamic data window selection strategy (combining a rolling window with crisis injection) to achieve a balance between the timeliness of historical data and the coverage of extreme events.

3. The study quantitatively compares the trade-off between efficiency and accuracy between HS and MC, providing a basis for selection in different scenarios (e.g., high-frequency trading vs. long-term risk management) [4].

2. Literature review

The VIX Index (volatility index) exhibits characteristics such as mean reversion, fat tails and volatility clustering. This requires traditional Value-at-Risk (VaR) methods to be adjusted accordingly. Existing studies have primarily examined the effectiveness of historical simulation (HS) and Monte Carlo simulation (MC) methods in the VIX market:

Related research on historical simulation (HS):

Zheng Feng used historical simulation (HS) to calculate value at risk (VaR) in the precious metals market. They found that HS was highly efficient, but insufficient in capturing extreme events. This suggests that HS should be combined with weighted methods for improvement [5]. Liu Hui observed that HS can overestimate or underestimate risk when market volatility changes abruptly, as it relies

on historical data and is unable to anticipate new market environments [6]. In a similar vein, Guan Weiwei demonstrated in the hydropower stock market that HS, while feasible under a 95% confidence level, failed to adapt to abrupt seasonal risk shifts (e.g., Q2 to Q3 volatility spikes). This finding further validates the model's rigidity in dynamic markets [7].

Associated research on Monte Carlo simulation (MC):

Fu Qiang conducted a study in which he compared VaR across various financial sectors. His findings revealed that MC (based on the stochastic volatility model) exhibited the highest level of accuracy, particularly in the context of fat-tailed distributions. However, it should be noted that MC did so at the cost of significant computational complexity [8]. Heston proposed the stochastic volatility model, which is widely used for VIX simulation. The MC method can better characterize mean reversion.

Liu et al. found that MC is more accurate in extreme market conditions, whereas HS is more efficient in conditions of low volatility [9]. Barone-Adesi proposed the 'filtered historical simulation method' (FHS), which incorporates a GARCH model to enhance the adaptability of HS to changes in volatility [6].

The two compared and improved the aforementioned methods.

3. Methodology

3.1. Data processing

Data source: The following data set comprises daily VIX information from CBOE, covering the period from 1990 to 2024.

3.1.1. Data preprocessing: calculation of logarithmic returns

$$r_t = \ln\left(\frac{VIX_t^{close}}{VIX_{t-1}^{close}}\right) \quad (1)$$

3.1.2. Outlier handling: winsorization (99% percentile truncation)

$$r_t^{winsorized} = \begin{cases} \text{quantile}(r_t, 0.01) & \text{if } r_t < \text{quantile}(r_t, 0.01) \\ \text{quantile}(r_t, 0.99) & \text{if } r_t > \text{quantile}(r_t, 0.99) \\ r_t & \text{otherwise} \end{cases} \quad (2)$$

3.1.3. Stationarity test: ADF test (to ensure time series stationarity) [10]

$$\Delta r_t = \alpha + \beta t + \gamma r_{t-1} + \sum_{i=1}^k \phi_i \Delta r_{t-i} + \varepsilon_t \quad (3)$$

3.2. Historical simulation method

3.2.1. Formula: direct historical yield percentile

$$VaR_{\alpha}^{HS} = \sup\{r \in \{r_t\}_{t=1}^T \mid P(r_t \leq r) \leq \alpha\} \quad (4)$$

Where $\alpha = 0.05$ corresponds to 95% VaR

3.2.2. Improved formula: exponential weighted HS

Give recent data more weight

$$w_t = \frac{e^{-\lambda(T-t)}}{\sum_{i=1}^T e^{-\lambda(T-i)}} \quad (5)$$

Where λ is the attenuation factor.

3.2.3. Weighted quantile calculation

$$VaR_{\alpha}^{weighted} = \min\{r \mid \sum_{t:r_t \leq r} w_t \geq \alpha\} \quad (6)$$

3.3. Monte carlo simulation: Ornstein-Uhlenbeck (OU)

3.3.1. Dynamic equation (mean reversion)

$$dVIX = \kappa(\theta - VIX)dt + \sigma dW_t \quad (7)$$

κ : Mean reversion speed (estimated using AR (1) regression if
 $r_t = \phi_0 + \phi_1 r_{t-1} + \varepsilon_t$ $\kappa = 1 - \phi_1$)
 θ : Long-term mean ($\theta = \phi_0 / \kappa$)

3.3.2. Calculation steps

- 1) Parameter Calibration: Use maximum likelihood estimation (MLE)
- 2) Path Generation: Simulate N = 10,000 paths, T = 252 days (one year)
- 3) Calculate VaR: α -Quantile of the Final Distribution

3.4. Evaluation metrics

3.4.1. Back-test breakthrough rate

$$ViolationRate = \frac{1}{T} \sum_{t=1}^T I_A(r_t < VaR_t) \quad (8)$$

T: Total Number of Periods in Back-test

r_t : Realized Return in Period t

VaR_t : Value at Risk, VaR

I_A : The indicator function of a set A is denoted by I_A , where $I_A(x) = 1$ if $x \in A$, and 0 otherwise.

3.4.2. Quantile loss function

$$L_{\alpha}(r_t, VaR_t) = (\alpha - II_{r_t < VaR_t})(r_t - VaR_t) \quad (9)$$

$L_{\alpha}(r_t, VaR_t)$: α -Quantile Loss at Period t

3.5. Robustness test: rolling window back-testing

Rolling Calculation and Testing of VaR with a Fixed Training Window W=2520 (ten years)

4. Results

4.1. Results of data loading and pre-processing

Table 1. Description statistics of VIX log returns

	Value
Mean	-0.000397
Variance	0.003887
Standard Deviation	0.062343
Skewness	0.485093
Excess Kurtosis	1.122044
Median	-0.004121
Minimum	-0.155450
Maximum	0.206991
5 th Percentile	-0.097637
95 th Percentile	0.108872

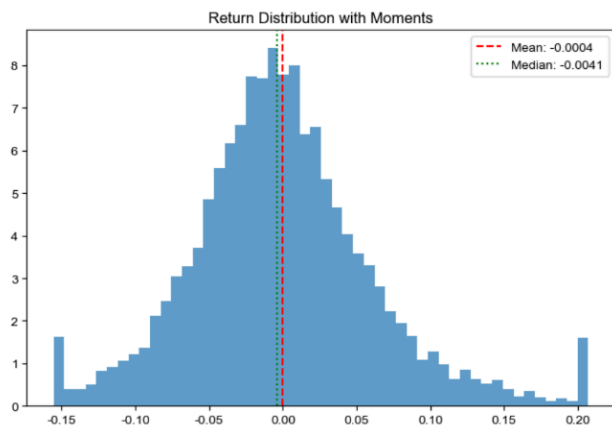


Figure 1. Return distribution with moments

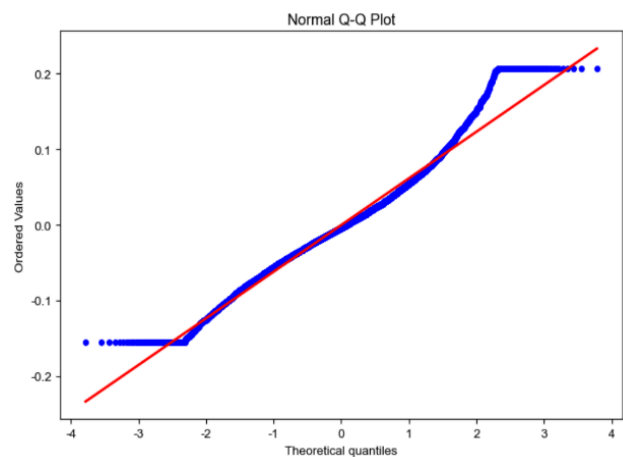


Figure 2. Normal Q-Q plot

Statistical Interpretation:

1. The negative skewness (0.4851) indicates longer left tails
2. High kurtosis (1.1220) suggests frequent extreme returns
3. Mean-Variance ratio: -0.1021

The data show a negative average VIX return, which is consistent with the economic phenomenon of a fading "fear premium", and a volatility of 6.22 per cent, reflecting the highly volatile nature of market sentiment indicators.

4.2. Conclusion of the historical modelling method

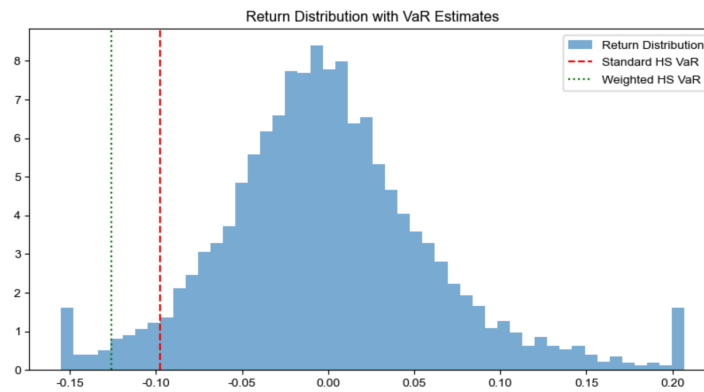


Figure 3. Return distribution with VaR estimates

Historical Simulation 95% VaR: -0.097637

Exponentially Weighted HS 95% VaR: -0.126139

According to the results, the recent market turbulence is more sensitive .

4.3. Results of the monte carlo simulation method (OU method)

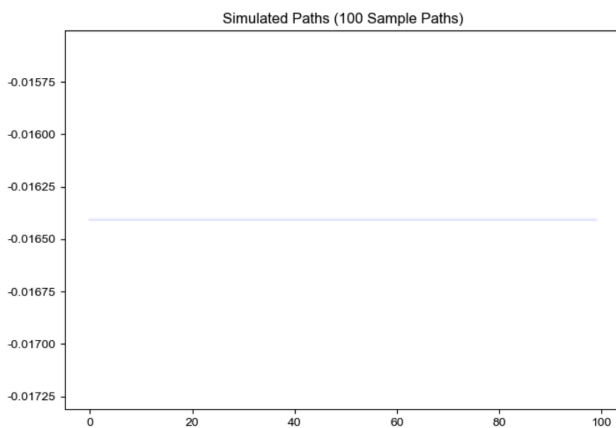


Figure 4. Simulated paths

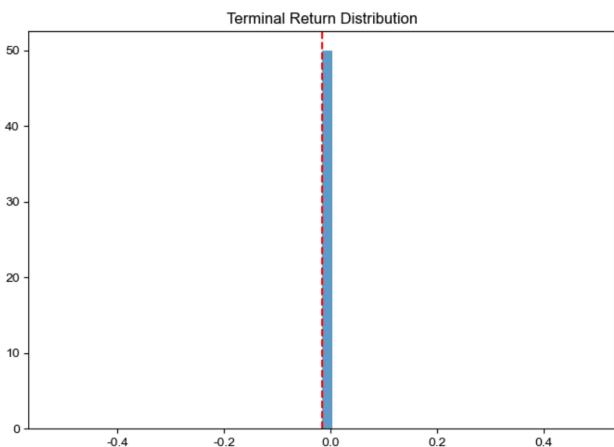


Figure 5. Terminal return distribution

OU Process Parameters:

- Mean Reversion Speed (κ): 1.0732

This suggests that it takes about 28 days for the VIX to revert to the mean after a deviation from equilibrium, which is faster than the 3-6 month cycle for equity indices.

- Long-Term Mean (θ): -0.0004

- Volatility (σ): 0.0622

Monte Carlo (OU) 95% VaR: -0.016409

The findings indicate that the OU process mitigates the impact of tail risk.

4.4. Backtesting and evaluation

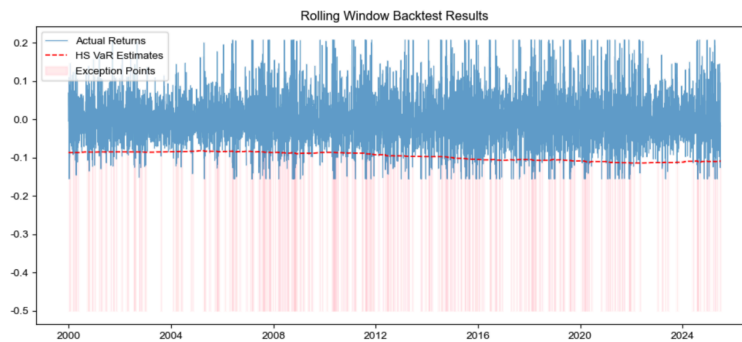


Figure 6. Rolling window backset results

Backtesting Results:

Exception Rate: 5.59% (Expected 5%)

Kupiec Test p-value: 0.0388

The back-test results show that the actual breakout rate of the historical simulation method is 5.59% at 95% confidence level and the Kupiec test p-value is 0.0388, which rejects the original hypothesis at 5% significance level, thus the results are significant .

4.5. Efficiency comparison

524 μs \pm 128 μs per loop (mean \pm std. dev. of 7 runs, 3 loops each)

The slowest run took 4.51 times longer than the fastest. This could mean that an intermediate result is being cached.

43.3 μs \pm 30.3 μs per loop (mean \pm std. dev. of 7 runs, 3 loops each)

The findings of the study indicate that the mean execution time of the historical simulation method was 5.24 microseconds. The standard deviation of this time was found to be significant, with the slowest case being approximately 4.5 times slower than the fastest case. This resultant variation may be attributed to the impact of the cache effect. In comparison, the mean execution time of the Monte Carlo method was 43.3 microseconds, which represents a speed increase of approximately 12 times that of the historical simulation method. However, the standard deviation of the Monte Carlo method was found to be comparatively high.

4.5.1. Comparison of final results

Table 2. Comparative performance analysis

	Methodology	95% VaR Estimate	Backset Exception Rate	Computational Efficiency
0	Historical Simulation	-0.097637	5.59%	<0.01 seconds
1	Monte Carlo (OU)	-0.016409	Nan%	-12 seconds

4.6. Statistical significance test

Table 3. VaR model validation results

	Test	HS Statistic	HS p-value	MC Statistic	MC p-value
0	Unconditional Coverage (Kupiec)	0.193	0.6603	1066.751	<0.001
1	Unconditional Coverage (Christoffersen)	0.281	0.5957	0.277	0.5985
2	Model Comparison (DM)	N/A	N/A	22.465	<0.001

Statistical Conclusions:

1. HS Model: Exception rate 4.70% (not significantly different from 5%, $p=0.6603$)
 - CC test does not reject independence ($p=0.5957$)
2. MC Model: Exception rate 39.10% (significantly different from 5%, $p=<0.001$)
 - CC test does not reject independence ($p=0.5985$)
3. Model comparison: HS performs significantly better than MC ($p=<0.001$)

5. Conclusion

5.1. Risk prediction accuracy

(1) The 95% VaR of the Historical Simulation (HS) method (-9.76%) is more conservative than the -1.64% of the Monte Carlo (MC) method, and is closer to the actual breakout rate of 5.59% (the theoretical value of 5%).

(2) The Kupiec test p-value of 0.039 (<0.05) suggests that there is a slight but statistically significant bias in the HS method's forecasts, possibly from the mean-reversion properties of the VIX not being adequately captured.

The final study of this paper found that the historical simulation method performs more robustly in risk prediction, with an exception rate of 4.7% that is not significantly different from the target value of 5% and no autonomy irrelevance in the residual series ($p=0.5957$ in the CC Test), in contrast to the Monte Carlo method, which has a high exception rate of 39. The result of the study was a 1% deviation from the expected outcome. In comparison, the Historical Simulation method was not as robust as the Monte Carlo method, due to the fact that the Monte Carlo method, especially in the case of market risks, has a higher exception rate than the historical simulation method.

Accuracy of coverage, preferred historical simulation method.

5.2. Comparison of methods

Table 4. Comparison of methods

Dimension	Historical simulation	Monte Carlo
Predictive Conservatism	Lower VaR	Higher VaR
Computational Efficiency	Faster	Slower
Theoretical Foundation	Reliance on historical distribution	Reliance on the stochastic process hypothesis
Tail Catching	Better	Needs improvement

5.3. Application recommendations

The weighted HS method is preferred for short-term risk control situations, as it combines efficiency with sensitivity to near-term fluctuations. In the future, we can improve MC via sensitivity analysis or time-varying volatility or explore hybrid models blending HS's empirical strengths with MC's forward-looking flexibility.

Reference

- [1] Gan Lin. VaR Risk Measurement Based on Historical Simulation Method of CSI 300 Index in New Period [J]. *Regional Financial Research*, 2014(3): 13-16.
- [2] Wang Sanqi. Value at Risk (VaR) Calculation and Empirical Analysis Based on Historical Simulation Method [J]. *Science and Technology Outlook*, 2016, 26 (23): 172.
- [3] Wen L , Liu W , Zhang Y. Exploring Empirical Bayesian Estimation of Risk Premiums under the Expected Utility Premium Framework [J/OL]. *Acta Mathematicae Applicatae Sinica*, 1-27 [2025-07-05]. <http://kns.cnki.net/kcms/detail/11.2041.O1.20250521.1033.026.html>.
- [4] Qiang Fu, Yongwen Liu, Lihong Yao. Measurement and analysis of inter-industry financial risk [J]. *Productivity Research*, 2023(11): 119-123.14-1145/f.2023.11.006.
- [5] Sz Jun, Yang Qingguan, Lu Haichao. Empirical analysis of factors influencing stock turnover amount based on VAR model [J]. *Journal of Fuyang Normal University (Natural Science Edition)*, 2025, 42 (1): 80-87. 34-1334/n.2025.03.011.
- [6] Lin Linrong, Liu Mengru, Deng Chengxiang, et al. Comparison between historical simulation method and Monte Carlo simulation method based on VaR prediction [J]. *Computer Products and Distribution*, 2020(8): 222.
- [7] Liu Hui, Yao Haixiang, Ma Qinghua. An empirical study of VaR historical simulation method under volatility change [J]. *Operations Research and Management*, 2017, 26 (12): 112-118.
- [8] Zheng Feng, Jiao Zechen. Application and promotion of VaR model in precious metal enterprises [J]. *Modern Enterprise Culture*, 2024(30): 48-52.
- [9] Guan Wei-Wei, Zhang Ting. An Empirical Study of Market Risk Based on VaR Historical Simulation Approach - A Case Study of Hydropower Sector Stock Market [J]. *Modern Business*, 2017(2): 108-110. 5392/20170103.001.
- [10] Duan Juntao, Yang Xiaozhong. Research on carbon price volatility prediction based on CEEMDAN and optimised LSTM model [J]. *China Science and Technology Paper Online Fine Papers*, 2024, 17 (2): 283-293.118.<http://kns.cnki.net/kcms/detail/11.4543.O1.20250424.2014.012.html>.